

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

**Ned Augenblick**  
*UC Berkeley Haas*

**Eben Lazarus**  
*MIT Sloan*

JANUARY 2022

## Abstract

How restrictive is the assumption of rational expectations in asset markets? To address this, we derive bounds on admissible asset-price variation under general assumptions. The challenge in this task is that valuations reflect both beliefs and preferences. We gain traction by considering market-implied, or risk-neutral, probabilities of future outcomes, and we provide a mapping between the variation in these probabilities and the minimum risk aversion — or slope of the stochastic discount factor — required under rationality. Implementing these bounds empirically using index options, we find that very high risk aversion is needed to rationalize the behavior of risk-neutral beliefs.

*JEL Codes: D84, G14, G41*

---

\*Contact: [ned@haas.berkeley.edu](mailto:ned@haas.berkeley.edu) and [elazarus@mit.edu](mailto:elazarus@mit.edu). We are grateful to Laura Blattner, Jarda Borovička, John Campbell, Gabriel Chodorow-Reich, Emmanuel Farhi, Xavier Gabaix, Niels Gormsen, Sam Hanson, Bryan Kelly, David Laibson, Ian Martin, Matthew Rabin, Andrei Shleifer, Juan Sotes-Paladino, David Sraer, Jim Stock, Adrien Verdelhan, Jessica Wachter, and numerous seminar and conference participants for advice, comments, and helpful discussions.

# 1. Introduction

Asset prices are volatile. Whether they are *too* volatile to reflect rational variation in expected future cash flows, though, is by itself an unresolvable question: the true distribution of future cash-flow streams is unobservable, as are aggregate risk and time preferences. The relevant question is instead what *joint* set of assumptions on (1) beliefs and (2) risk and time preferences — the two elements at the heart of almost all modern theories of asset prices — can be rejected in the data.

Shiller (1981) provides a classic example. He documents excess volatility in equity-index prices relative to a proxy for fundamental value, but this proxy is constructed under the assumption that discount rates over future cash flows are constant over time. Shiller’s test therefore provides a *joint* rejection of (1) rational pricing and (2) constant discount rates. But given the ample evidence that discount rates vary over time, it is unclear what to make of this joint rejection.<sup>1</sup> It would seem that changes in unobserved discount rates are capable of explaining arbitrary movements in asset values, so can any meaningful statements be made about how much rational variation we should observe in the data?

In this paper, we show that, even with a meaningful degree of flexibility in preferences and discount rates, there are still certain bounds on asset-price movements that must hold under the assumption of rational expectations (RE). Our theoretical results apply quite generally to valuation processes for any agent with RE over an asset’s future payoffs; when we then take these bounds to the data, the “agent” in question can be thought of as the market as a whole. Our results accordingly bear on the efficiency of market valuations, as we show that this null can be jointly tested alongside a much less restrictive assumption than those used in past literature. When we implement our bounds empirically, we find that there is so much variation in the asset prices we consider that it is difficult to rationalize the data with plausible assumptions on time and risk preferences.

As in Shiller’s case, we focus our analysis on expectations over the future value of an equity index. But the key feature distinguishing our analysis from previous literature is that we consider the behavior of so-called *risk-neutral (RN) beliefs* over the underlying index’s future price, rather than the behavior of the underlying index itself. The risk-neutral belief distribution can be calculated directly using option prices — options allow for bets over the future asset price, and thus the prices of these bets allow us to back out a probability distribution over this future price — so as is standard, we treat risk-neutral beliefs as observable. These risk-neutral beliefs represent the probability distribution that would be equal to a hypothetical risk-neutral agent’s true (or *physical*) belief distribution about the future asset price, but risk-neutral beliefs are in general distorted relative to the marginal investor’s physical beliefs in the case that the investor is risk-averse. Intuitively, the probability distribution we observe using asset prices will overweight states in which the marginal investor has low wealth (e.g., when the underlying asset has a low return), since the investor will be willing to pay to insure against these high-marginal-utility states.

---

<sup>1</sup>For an overview of time-varying discount rates, see Cochrane (2011). Fama (1991, p. 1586) provides a succinct summary of the issues in interpreting Shiller’s results: “volatility tests are [a] useful way to show that expected returns vary through time,” but they “give no help on the central issue of whether the variation in expected returns is rational.”

We show that statements about the “correct” amount of variation in risk-neutral beliefs under RE require less-restrictive assumptions than statements about variation of the index price itself. Previous analyses focusing on index-price variation require keeping track of some measure of the index’s fundamental value; in contrast, we show that one can place restrictions on the intertemporal behavior of RN beliefs without any knowledge of the asset’s fundamental value, or knowledge of the marginal investor’s underlying physical beliefs. Aside from the maintained assumptions of RE and no arbitrage, our main results require only one general restriction on the stochastic discount factor (SDF), the random variable that determines an asset’s ex-ante price by discounting the asset’s random future cash flows: we assume that the SDF realization does not depend on the path of unobservable state variables realized between a given trading date and the option expiration date. We refer to this assumption as *conditional transition independence (CTI)*, and this assumption is met in many common macro-finance models. Further, we provide sufficient conditions under which our bounds are robust to violations of the CTI assumption.

To demonstrate the logic of our test, it is useful first to consider the simple case in which we can directly observe a person’s subjective belief  $\pi_t(\theta = 1)$  over some binary outcome  $\theta \in \{0, 1\}$ . Suppose we observe that  $\pi_t$  continually moves from 0.10 to 0.90 as  $t$  progresses. While it is of course possible to rationalize this movement ex post by constructing a set of signal realizations from a particular signal data-generating process (DGP), this amount of movement appears intuitively “rare” and might bring into question the hypothesis that the agent has RE. Formalizing this intuition, [Augenblick and Rabin \(2021\)](#) note that, when uncertainty is resolved by some period  $T$ , the expectation of the sum of squared changes in beliefs across all periods (*belief movement*,  $\mathbb{E} \sum_{t=0}^{T-1} (\pi_{t+1} - \pi_t)^2$ ) must equal the agent’s *initial uncertainty* ( $\pi_0(1 - \pi_0)$ ) under RE, *regardless* of the signal DGP. Equivalently, regardless of DGP, the expected *excess movement* — belief movement minus initial uncertainty, which we denote by  $\mathbb{E}[X]$  — must always equal zero. Intuitively, under RE, changes in beliefs must on average correspond to the resolution of uncertainty (from its initial level  $\pi_0(1 - \pi_0)$  to 0); rational belief movements must mean the person is learning something about the outcome in question.<sup>2</sup> If beliefs are instead continually shifting dramatically relative to initial uncertainty, the RE assumption can be statistically rejected at some confidence level.

The main theoretical contribution of this paper is to show how this logic — that movement in beliefs must correspond on average to reduction in uncertainty — can be used to restrict excess movement in *risk-neutral* beliefs,  $\mathbb{E}[X^*]$ , in the general case in which only risk-neutral rather than physical beliefs are observable. Our task becomes considerably more difficult in this case: observable risk-neutral beliefs need not follow a martingale under RE given their distortion relative to physical beliefs, and this means that  $\mathbb{E}[X^*]$  can be non-zero.<sup>3</sup> But given that the distortion

---

<sup>2</sup>As a simple example, take a situation in which two coins are flipped sequentially and a person with RE states her beliefs about the likelihood that she will observe two heads. With probability  $\frac{1}{4}$ , the agent observes two heads, leading to beliefs that move from 25% to 50% to 100%, which has movement of  $\frac{5}{16}$ ; she observes an initial tails with probability  $\frac{1}{4}$ , leading to belief movement of  $\frac{1}{16}$ ; and she observes heads and then tails with probability  $\frac{1}{4}$ , leading to movement of  $\frac{5}{16}$ . As with any signal DGP with a prior of  $\frac{1}{4}$ , the expectation movement is equal to  $\frac{3}{16}$ , the initial uncertainty.

<sup>3</sup>Equivalently, some positive average amount of apparent excess movement in risk-neutral beliefs can be rationalized under RE, as long as the person (or market) exhibits some risk aversion over the time- $T$  outcome.

between risk-neutral and physical beliefs is indexed by risk aversion over the terminal states, we show in this case that the admissible  $\mathbb{E}[X^*]$  under RE can be bounded as a simple function of this risk-aversion value (or, more generally, the slope of the SDF across states). The broad intuition is the same: if a person's valuation for a binary option that pays \$1 in a given state at time  $T$  continually oscillates between \$.10 and \$.90, her unobservable risk aversion renders her precise physical beliefs unidentified, but these extreme valuation changes imply belief movements that must be rare under RE regardless of her exact risk aversion.

Our upper bound for excess RN belief movement  $\mathbb{E}[X^*]$  is tight in the sense that it is possible to construct a (somewhat perverse) signal DGP that produces movement that is arbitrarily close to the bound. But this bound is, by construction, very conservative, and so we also provide a tighter bound that holds under somewhat more restrictive assumptions. In both cases, the bound depends on a risk-aversion parameter  $\phi$  that corresponds to the relative marginal value of a dollar across time- $T$  states. For example, if the person is forming beliefs over two possible terminal consumption states  $C_T \in \{C_{low}, C_{high}\}$  (so that the option payoff is, e.g.,  $Y_T = 1$  if  $C_T = C_{low}$ ), then  $\phi = \frac{U'(C_{low})}{U'(C_{high})}$ . In the empirically relevant case where we observe market-implied RN beliefs over the index return, then  $\phi = \frac{\mathbb{E}[M_{i,T} | R_T = R_{low}]}{\mathbb{E}[M_{i,T} | R_T = R_{high}]}$ , where  $M_{i,T}$  is the stochastic discount factor and  $R_{low}, R_{high}$  are arbitrary index return values. Since our bound for  $\mathbb{E}[X^*]$  depends on this parameter  $\phi$ , any observed average excess movement in the data is informative as to the minimal  $\phi$  required in order for the bound to be satisfied. We are effectively inverting the logic of the joint hypothesis problem: under the assumption of RE, the data is informative as to the minimal  $\phi$  needed to rationalize the observed amount of excess movement in RN beliefs. We are, in other words, asking how restrictive it is — in particular, how much utility curvature one must assume — to maintain the hypothesis of RE.

Our main assumption throughout — the assumption of conditional transition independence (CTI) introduced above — is that  $\phi$  is constant over time for a given belief stream, which generally lasts weeks to months in our empirical exercise. This assumption holds in many leading economic models: given a stable utility function, the marginal value of \$1 in some time- $T$  state generally does not change relative to its marginal value in some other state.<sup>4</sup> Nonetheless, we provide additional results suggesting that reasonable variation in  $\phi$  is unlikely to meaningfully affect our conclusions. This might seem counterintuitive: for example, one might imagine that if  $\phi$  oscillates between 2.9 and 3.1 over and over, the movement of RN beliefs can be unbounded over time even given no movement in physical beliefs, strongly violating our results. This logic, however, ignores the fact that the person must also have rational expectations over  $\phi$ , making such persistent oscillations rare under RE. To demonstrate this principle, we first theoretically prove that if expectations about  $\phi$  evolve as a martingale or supermartingale, our bounds continue to hold. We then numerically simulate hundreds of DGPs in which both the underlying state and  $\phi$  are uncertain, so that  $\phi$  varies within a stream. We find that, across all DGPs, larger initial uncertainty in  $\phi$  (leading to larger average movement in  $\phi$  across the stream) has little to no impact on the average and maximum

---

<sup>4</sup>In asset-pricing terms, while permanent shocks to the SDF may alter  $\mathbb{E}_t[M_{i,T}]$  for all states, they generally do not affect the expected *relative* SDF realizations across possible states.

$\mathbb{E}[X^*]$  statistics. Finally, we simulate a calibrated version of the habit-formation model of [Campbell and Cochrane \(1999\)](#), which implies a changing  $\phi$ . We find that our bounds continue to hold and continue to be conservative.

We then take our bounds to the data using S&P 500 index option prices obtained from OptionMetrics. We use standard methods to infer the distribution of the market's RN beliefs over index returns for each option expiration date in the sample. In order to map to our two-state theoretical setting, we then translate each full distribution into a set of binary RN beliefs  $\pi_t^*(R_T = \theta_j | R_T \in \{\theta_j, \theta_{j+1}\})$ ; these correspond to the RN probability that the index return will be equal to (or in a range close to)  $\theta_j$ , conditional on being either  $\theta_j$  or  $\theta_{j+1}$ . (We set our return states to correspond to five-percentage-point bins for the S&P return.) We then implement our theoretical bounds, which allow us to infer the minimal risk-aversion value  $\phi$  (at each point in the return distribution) needed to rationalize the observed variation in RN beliefs over the index return.

We find that extremely high risk aversion is needed to rationalize the observed excess movement: in many cases, there is in fact no value of  $\phi$  under which the tight version of the bound is met, and the conservative version of the bound generally implies implausibly large values for  $\phi$ . It is thus quite costly (in the sense that a high  $\phi$  is required) to maintain the assumption of RE in light of the degree of excess movement in risk-neutral beliefs over the index return. This suggests that many leading rational frameworks capable of explaining medium-to-low-frequency variation in asset prices have difficulty rationalizing medium-to-high-frequency variation in beliefs.

Given that we conduct our estimation using variation in index options prices, we must also account for the effect of possible market microstructure noise. In order to do so, we derive an additional theoretical result that describes how microstructure noise biases our estimates through its effect on observed excess movement  $X^*$ . This bias depends on the variance of the noise component of observed RN beliefs. We accordingly proceed to estimate this noise variance in the data, by turning to a set of high-frequency index option prices (which we obtain directly from the CBOE) for this purpose. Using these intraday data, we implement the microstructure noise variance estimator proposed by [Li and Linton \(2021\)](#), which is particularly well suited for our purposes. We can then construct an empirical noise correction, removing the effect of noise from  $X^*$  before we conduct our estimation. All of our results are noise-corrected in this way, which help to ensure that our findings do not depend on idiosyncrasies specific to the options market.

We then briefly explore possible explanations for our findings, with the goal of providing some positive directions for alternative models. Conducting regressions of our excess-movement measure on a range of macroeconomic statistics, we find that  $X^*$  has a strong positive relationship with measures of macroeconomic uncertainty and no relationship with measures of liquidity or limits to arbitrage in asset markets. These results provide further evidence against the possibility that factors specific to the option market are the main drivers of our results, but this regression evidence is only suggestive and reduced-form.

**Relation to previous literature and interpretation of results.** In addition to [Shiller \(1981\)](#), we follow, among others, [LeRoy and Porter \(1981\)](#), [De Bondt and Thaler \(1985\)](#), [Campbell and](#)

Shiller (1987), and Stein (1989) in testing for excess volatility in asset prices relative to RE. Kleidon (1986) and Marsh and Merton (1986) emphasize non-stationarity in accounting for apparent excess volatility; much of the literature since then has emphasized time variation in discount rates (Cochrane, 2011). We show that even without imposing any restrictions on the structure of the data-generating process, and imposing only mild restrictions on the variation in discount rates, RE nonetheless restricts the admissible variation in option prices in an empirically testable way.

There are two costs associated with our additional generality. First, we consider derivative prices rather than the behavior of the underlying index directly. In this way, our work is complementary to that of Giglio and Kelly (2018), who document excess volatility in long-maturity claims on equity and currency volatility, inflation swaps, commodity futures, and credit default swaps. Their framework differs from ours in that they achieve identification by parameterizing the DGP for cash flows on the underlying, whereas we restrict the evolution of the SDF. Their parameterization — a low-dimensional affine model under the RN measure — applies well to the term-structure-like claims they consider, but not to claims on the equity index itself, to which our framework does apply. The two frameworks thus provide independent and complementary evidence for excess volatility, and both do so in a manner that allows for discount-rate variation.

Second, rather than allowing for fully binary (rejection vs. non-rejection) empirical tests of RE models, our general framework instead allows for a mapping between the observed asset-price variation and the risk aversion required to rationalize the data. Our results may thus appear similar in spirit to those of Mehra and Prescott (1985), and more generally Hansen and Jagannathan (1991), who find that the SDF must be highly volatile to rationalize the observed excess returns for risky assets. Our results differ from theirs in two respects. First, we obtain our mapping using the second moment of observed returns, while they use the first moment of returns.<sup>5</sup> More importantly, the Hansen–Jagannathan results may in principle be explained by features of the data-generating process for consumption or returns rather than high risk aversion per se; for example, models of rare disasters (e.g., Barro, 2006; Gabaix, 2012) can generate sufficient SDF volatility to rationalize the observed equity premium without requiring high risk aversion. But this is not the case for our results, as we obtain a relationship between local changes in the RN belief distribution and local risk aversion (or the slope of the SDF) at those points of the distribution. If we observe highly variable RN beliefs over the event that the S&P’s 90-day return will be between 5% and 10%, we know that this cannot be attributable to disasters that affect the left tail of the return distribution; instead, we conclude either that risk aversion is very high or that there is a departure from RE.

Finally, our results complement evidence on beliefs obtained from survey data, as, for example, in Greenwood and Shleifer (2014), Bordalo et al. (2020), and De la O and Myers (2021), as well as the results of Augenblick and Rabin (2021) for settings with directly observable beliefs. Another set of related literature endeavors to measure physical beliefs indirectly from options data, for the purpose of examining either expectations or preferences; see Garcia, Ghysels, and Renault (2010)

---

<sup>5</sup>One could instead map between variation in returns and the volatility of SDF volatility (or the heteroskedasticity of the SDF), but we consider the results from our mapping to be somewhat more intuitive than this alternative.

for a survey. As a recent example, [Ross \(2015\)](#) assumes a Markov process for transitions between return states and a transition-independence assumption on the SDF more restrictive than the one we use (see [Borovička, Hansen, and Scheinkman, 2016](#), for a discussion), with which one can back out a distribution of physical beliefs (see also [D’Arienzo, 2020](#)). Our approach differs from this work in that we need not measure physical beliefs at all or know the true data-generating process for returns to conduct our tests, so we accordingly require less structure. We also complement the findings of [van Binsbergen, Brandt, and Koijen \(2012\)](#), who, like us, provide evidence of excess volatility in short-maturity equity derivatives claims.

**Organization.** [Section 2](#) introduces our theoretical framework in a simple two-state setting, which allows for clear derivations and intuition for our main results. [Section 3](#) then extends these results to a more general asset-pricing setting, and [Section 4](#) provides additional results speaking to the robustness and economic interpretation of our bounds. We then implement our bounds empirically in [Section 5](#), which describes our data and presents our results. [Section 6](#) concludes. While we provide derivations for many of our theoretical results in the text, our online appendix contains detailed proofs ([Appendix A](#)) as well as additional technical material ([Appendix B](#)).

## 2. Theoretical Results in a Simple Setting

We first examine risk-neutral (RN) belief movement in a simplified environment with a single individual and two terminal consumption states. This simple setting allows for transparent derivations of our main results and discussion of their economic intuition, before turning to the more general asset-pricing framework considered in [Section 3](#). We begin by describing our setup and reviewing previous results concerning physical belief movement ([Section 2.1](#)), then discuss RN beliefs and the complications introduced by risk aversion ([Section 2.2](#)), and finally provide simple versions of our main theoretical results on RN belief movement ([Section 2.3](#)).

### 2.1 Bayesian Belief Movement: Setup and Results

Before considering asset prices and RN beliefs, we first consider the simplified case in which a person’s beliefs are directly observable. The setup and results in this subsection build on [Augenblick and Rabin \(AR, 2021\)](#).

Time is discrete and indexed by  $t \in \{0, 1, 2, \dots, T\}$ . At the beginning of each period, a person observes a signal  $s_t \in S$  regarding two mutually exclusive and exhaustive events, which we call *states*  $\theta \in \{0, 1\}$ . The data-generating process (DGP) is general: signals are drawn from the discrete signal distribution  $DGP(s_t | \theta, H_{t-1})$ , where  $H_t$  represents the history of signal realizations through  $t$ . Define  $\mathbb{P}(H_T)$  to be the probability of observing history  $H_T$  induced by the DGP, and write  $\mathbb{E}[\cdot] \equiv \mathbb{E}^{\mathbb{P}}[\cdot]$  for the expectation under  $\mathbb{P}$ . The person’s (*physical or subjective*) belief in state 1 (vs. state 0) at time  $t$  given the DGP and history  $H_t$  is denoted by  $\pi_t(H_t)$ . The belief stream  $\pi(H_t) = [\pi_0, \pi_1(s_1), \pi_2(\{s_1, s_2\}), \dots]$  is the collection of beliefs given history  $H_t$ . We will commonly suppress the dependence of these objects on  $H_t$  to simplify notation. Given our empirical

setting, we focus entirely on *resolving* streams in which the person achieves certainty about the true state by period  $T$  with probability 1:  $\pi_T(H_T) = \theta \in \{0, 1\}$  for all  $H_T$ .<sup>6</sup>

Total *belief movement* of  $\pi$  is defined as the sum of squared changes in beliefs across all periods:

$$m(\pi) \equiv \sum_{t=0}^{T-1} (\pi_{t+1} - \pi_t)^2. \quad (1)$$

*Initial uncertainty* of  $\pi$  is defined as the time-0 variance of the Bernoulli random variable  $\mathbb{1}\{\theta = 1\}$ :

$$u_0(\pi) \equiv (1 - \pi_0)\pi_0. \quad (2)$$

Given that we focus on resolving belief streams for which  $\pi_T \in \{0, 1\}$ , final uncertainty is always zero ( $u_T = (1 - \pi_T)\pi_T = 0$  a.s.). Initial uncertainty  $u_0$  is therefore equal to the total amount of uncertainty reduction for  $\pi$ ,  $u_0 - u_T$ , which is helpful in interpreting some of the following results.

Our main variable of interest is the difference between movement and initial uncertainty, which — for reasons that will become clear — we call *excess movement*:

$$X(\pi) \equiv m(\pi) - u_0(\pi). \quad (3)$$

Throughout the paper, we maintain the assumption that the person's beliefs are a martingale with respect to the true DGP:

**ASSUMPTION 1** (*Martingale Beliefs*). Beliefs satisfy  $\pi_t(H_t) = \mathbb{E}[\pi_{t+1} | H_t]$  for any  $H_t$ .

A person with rational expectations (RE) over  $\theta$  — equivalently, someone with a correct prior who updates using Bayes' rule according to the true DGP — will satisfy this assumption. Given our focus on resolving streams for which  $\pi_T = \theta$ , this assumption is in fact equivalent to RE, as  $\pi_t = \mathbb{E}[\pi_T | H_t] = \mathbb{E}[\theta | H_t] = \mathbb{P}(\theta = 1 | H_t)$ .<sup>7</sup> We thus often refer to [Assumption 1](#) as RE.

To clarify our setting, [Table 1](#) lists some examples of possible signals, belief streams, likelihoods of those belief streams, movement, and uncertainty resolution given three simple binary-signal DGPs that we will return to in the subsequent sections. (Only the first five columns of the table are relevant for now, as the remaining three refer to risk-neutral beliefs as defined and discussed below.) For example, in the first DGP, the person has a rational prior of  $\pi_0 = .25$ ; at  $t = 1$ , she observes  $s_1 \in \{l, h\}$  with symmetric relative likelihood ratios (or *precision*) for each of the two signals,  $L_h \equiv \frac{\mathbb{P}[s_1=h|\theta=1]}{\mathbb{P}[s_1=h|\theta=0]} = 3$  and  $L_l \equiv \frac{\mathbb{P}[s_1=l|\theta=0]}{\mathbb{P}[s_1=l|\theta=1]} = 3$ ; and then at  $t = 2$ ,  $s_2 \in \{L, H\}$  fully reveals the state (with signal “H” revealing state 1 and vice versa). In the first row, if the person first observes  $s_1 = l$  and then  $s_2 = L$ , she will first update to  $\pi_1 = .1$  and then  $\pi_2 = 0$ . The ex ante likelihood of this signal realization for this DGP is  $\frac{9}{16} = .5625$  and the resultant belief stream  $[\pi_0, \pi_1, \pi_2] = [.25, .1, 0]$  has movement of  $(.25 - .1)^2 + (.1 - 0)^2 = .0325$  and initial uncertainty of  $\pi_0(1 - \pi_0) = .1875$ . The

<sup>6</sup>As discussed in [AR](#), a number of statements can also be made about non-resolving (sub)streams. In our setting, these statements would often be excessively conservative and add little benefit to our empirical analysis.

<sup>7</sup>While one can construct cases for which beliefs violate RE but are a martingale under  $\mathbb{P}$ , such cases are degenerate (e.g.,  $\pi_t = 1$  for all  $H_t$ ) and ruled out by  $\pi_T = \theta$ . But for now we assume RE only over  $\theta$ , which is weaker than a full rationality assumption, as beliefs over individual histories  $H_T$  could still differ from  $\mathbb{P}(H_T)$ .



second example uses an asymmetric DGP in which the person either sees a high signal and learns with certainty that state 1 is true or sees a low signal and learns that state 0 is more likely. As a result, movements towards a belief of 1 are much larger than movements toward 0. The third example is the opposite with a resolving low signal and noisy high signal.

Under the RE assumption, belief movement and initial uncertainty are related according to the following proposition, which restates a main result in [AR](#).

**PROPOSITION 1** ([Augenblick and Rabin, 2021](#)). *Under Assumption 1, for any DGP, expected total belief movement must equal initial uncertainty. Expected excess movement in beliefs must therefore be zero:*

$$\mathbb{E}[X] = 0. \tag{4}$$

As discussed in [AR](#), [Proposition 1](#) is a straightforward implication of the assumption of martingale beliefs.<sup>8</sup> It motivates referring to  $X$  as *excess movement*. As an example, consider the first DGP in [Table 1](#): the four signal realizations  $[l, L]$ ,  $[l, H]$ ,  $[h, L]$ ,  $[h, H]$  have likelihood .5625, .0625, .1875, .1875, respectively, with respective movement statistics .0325, .8325, .3125, and .3125. Therefore, expected movement is .1875, which equals initial uncertainty, so expected excess movement is zero. Note that the other two DGPs in the table also start with a prior of .25 and initial uncertainty of .1875. [Proposition 1](#) therefore implies that, although these DGPs produce different streams with different movement statistics and likelihoods, the expected movement will be .1875 and expected excess movement will be zero.

[AR](#) then use this relationship to create simple tests of RE if one can directly observe beliefs. Given a set of observed belief streams, one can straightforwardly calculate the sample average of the empirical excess movement statistic and statistically test if it differs from zero. The restriction reflects the intuition that if the person’s beliefs are moving, this movement must on average correspond to learning about the true terminal state (in the sense that uncertainty is resolved from its initial value to 0). Rewriting  $\mathbb{E}[X] = 0$  as  $\mathbb{E}[\sum_{t=0}^{T-1} (1 - 2\pi_t)(\pi_{t+1} - \pi_t)] = 0$  (see [footnote 8](#)), it is apparent that expected belief movements toward 0.5 (the point of highest uncertainty) lead to a positive  $\mathbb{E}[X]$  statistic, and vice versa. So it could be the case that  $\mathbb{E}[\pi_{t+1} - \pi_t] = 0$  unconditionally, but a test based on [Proposition 1](#) would still reject the null of RE if, for instance, low values of  $\pi_t$  (i.e.,  $1 - 2\pi_t > 0$ , or  $\pi_t < 0.5$ ) tend to be revised upward ( $\pi_{t+1} - \pi_t > 0$ ), and high values tend to be revised downward.

## 2.2 Risk-Neutral Beliefs: Setup

[Proposition 1](#) shows that one can make statements about excess belief movement under RE when beliefs are observable. We now consider the situation in which beliefs are not directly observable, but it is possible to observe the stream of the person’s willingness to pay for an Arrow-Debreu

---

<sup>8</sup>To see this, rewrite  $X$  as  $\sum_{t=0}^{T-1} (2\pi_t - 1)(\pi_t - \pi_{t+1})$ . Using the law of iterated expectations on each term in the sum,  $\mathbb{E}[(2\pi_t - 1)(\pi_t - \pi_{t+1})] = \mathbb{E}[(2\pi_t - 1)(\pi_t - \mathbb{E}[\pi_{t+1}|\pi_t])]$ , which must be zero under the martingale assumption. This result has appeared in other forms in past literature; for one example, see [Barndorff-Nielsen and Shephard \(2001\)](#).

security that pays \$1 (1 unit of the numeraire consumption good) in period  $T$  if state  $\theta$  is realized. We denote this valuation by  $q_t(\theta|H_t)$  for each  $\theta \in \{0, 1\}$ . An object analogous to  $q_t(\theta|H_t)$  will be observable using options data for suitably defined states, but we postpone this additional formalism to [Section 3](#) so as to maintain focus on the main economic intuition for now. Our main contribution will be to develop a semi-parametric test of RE using only data on valuation streams.

To start, consider the simple case in which the person values consumption at all periods and in all states equally (i.e., she is risk-neutral and does not discount future consumption). In this case,  $q_t(1|H_t) = \pi(H_t)$  and  $q_t(0|H_t) = 1 - \pi(H_t)$ . Beliefs are thus directly observable from asset values, and we can apply [Proposition 1](#) as a test of the rationality of these beliefs.

The problem of identifying excess movement in asset valuations becomes significantly more complicated when the person does not weight consumption equally across time and states, as valuations no longer correspond directly to beliefs. To see this, assume now that the agent has time-separable utility, with concave period utility function  $U(C_t)$ , and exponentially discounts future consumption with discount factor  $\beta$ . Assume that  $\theta$  determines period- $T$  consumption  $C_{T,\theta} \equiv C_T(\theta)$ . At an interior optimum, valuations are

$$q_t(1|H_t) = \frac{\beta^{T-t}U'(C_{T,1})}{U'(C_t)}\pi_t, \quad q_t(0|H_t) = \frac{\beta^{T-t}U'(C_{T,0})}{U'(C_t)}(1 - \pi_t). \quad (5)$$

Valuations thus no longer reveal beliefs.

Following asset-pricing convention, define the *stochastic discount factor* (SDF)  $M_{t,T}(\theta)$  as the ratio of valuation to probability for state  $\theta$ :

$$M_{t,T}(1) \equiv \frac{q_t(1|H_t)}{\pi_t(H_t)}, \quad M_{t,T}(0) \equiv \frac{q_t(0|H_t)}{1 - \pi_t(H_t)}.$$

Continuing from (5),  $M_{t,T}(\theta) = \frac{\beta^{T-t}U'(C_{T,\theta})}{U'(C_t)}$  encodes the relative valuation of a marginal dollar in state  $\theta$  in period  $T$  versus one in period  $t$ . In our environment, the *SDF ratio*  $\phi_t$  will be particularly important:

$$\phi_t \equiv \frac{M_{t,T}(1)}{M_{t,T}(0)}, \quad (6)$$

which encodes the slope of the stochastic discount factor across the two states. In consumption terms, continuing the above example, we have

$$\phi_t = \frac{\frac{\beta^{T-t}U'(C_{T,1})}{U'(C_t)}}{\frac{\beta^{T-t}U'(C_{T,0})}{U'(C_t)}} = \frac{U'(C_{T,1})}{U'(C_{T,0})}, \quad (7)$$

so  $\phi_t$  can be thought of as the marginal rate of substitution across the states. Assume that  $\theta = 1$  corresponds to the low-consumption state relative to  $\theta = 0$ , or  $C_{T,1} < C_{T,0}$ ; this is without loss of generality, as the states can be relabeled arbitrarily. Given this labeling, with concave utility, we have  $U'(C_{T,1}) \geq U'(C_{T,0})$  and thus  $\phi_t \geq 1$ . We will maintain the convention of labeling state 1

(corresponding to the belief  $\pi_t$ ) as the “bad” state, so we state it as an assumption:

**ASSUMPTION 2** (*Labeling of States*).  $\phi_t \geq 1$  with probability 1 for all  $t$ .

In (7), given that  $\theta$  determines period- $T$  consumption,  $\phi_t$  is in fact constant across  $t$ . The person’s beliefs may be changing about the relative likelihood of the two states, but her beliefs about the relative marginal utility outcomes conditional on each of the two states being realized are not. This assumption is central to our results:

**ASSUMPTION 3** (*Constant SDF Ratio*).  $\phi_t = \phi$  is constant with probability 1 for all  $t$ .

We discuss this assumption in more detail, for general settings, in [Section 3](#). Intuitively, what is required is that the “severity” of the bad state relative to the good state is constant, so that the relative valuation of a marginal dollar in the two states is constant. This follows naturally in the current example from the assumption of stable time-separable utility and fixed state-contingent consumption values. This illustrates the manner in which restrictions on risk-neutral belief variation require weaker assumptions than restrictions on the underlying asset price: an Arrow-Debreu state price (and associated RN belief) depends on marginal utility in a single state, whereas the price of a consumption claim depends on the probability-weighted sum of marginal utilities over *all* states. Assuming constant discount rates allows for identification in the latter context (e.g., [Shiller, 1981](#)), but we need not make this assumption when working with RN beliefs. RN beliefs also eliminate dependence on time- $t$  marginal utility by means of a convenient normalization. [Appendix B.1](#) (in our online appendix) discusses the relationship between RN beliefs and discount rates in greater detail. We also later consider how our bounds are affected when  $\phi_t$  is time-varying; this robustness discussion is postponed to [Section 4.3](#).

For further intuition on  $\phi$ , we can conduct a Taylor expansion of  $U'(C_{T,0})$  around  $C_{T,1}$  as  $U'(C_{T,0}) = U'(C_{T,1}) + U''(C_{T,1})(C_{T,0} - C_{T,1}) + \mathcal{O}((C_{T,0} - C_{T,1})^2)$ . Then rearranging and using (7), to first order,

$$\gamma(C_{T,1}) \equiv -\frac{C_{T,1}U''(C_{T,1})}{U'(C_{T,1})} = \frac{\phi - 1}{(C_{T,0} - C_{T,1})/C_{T,1}}. \quad (8)$$

Relative risk aversion  $\gamma$  depends on the ratio of marginal utilities across states  $\phi$  relative to the percent consumption gap across states. Thus up to a scaling constant  $(C_{T,0} - C_{T,1})/C_{T,1}$ , the value  $\phi$  can be thought of as an index of risk aversion.

Given that beliefs are not directly inferable from asset valuations without knowledge of  $\phi$ , one cannot in general apply [Proposition 1](#) in the current context. We instead aim to derive bounds on excess movement in *risk-neutral* (RN) beliefs, which are defined (following convention) as

$$\begin{aligned} \pi_t^*(H_t) &\equiv \frac{q_t(1|H_t)}{q_t(0|H_t) + q_t(1|H_t)} \\ &= \frac{U'(C_{T,1})}{\mathbb{E}_t[U'(C_T)]} \pi_t(H_t) = \frac{\phi \pi_t(H_t)}{1 + (\phi - 1)\pi_t(H_t)}, \end{aligned} \quad (9)$$

with the second line following from (5) and (7). RN beliefs for  $\theta = 0$  can be similarly defined as  $\frac{q_t(0|H_t)}{q_t(0|H_t)+q_t(1|H_t)} = 1 - \pi_t^*$ , so the two states' RN beliefs are positive and sum to 1 by construction. Risk-neutral beliefs are so named because they can be interpreted as the subjective beliefs for a fictitious risk-neutral agent, as they coincide with actual subjective beliefs in the case that the person is risk-neutral. In general, they represent a pseudo-belief distribution, reflecting a combination of beliefs and risk preferences. Given  $\phi \geq 1$ , the bad-state RN belief  $\pi_t^*$  in general exceeds the person's subjective belief  $\pi_t$ : the person is willing to pay more than the actuarially fair value for a bad-state consumption claim given her high marginal utility in that state, and vice versa.

Returning to Table 1, the sixth column provides RN beliefs for the three example DGPs under the calibration  $\phi = 3$ . The person has a prior of .25 in each DGP, so using (9), the RN prior is  $\pi_0^* = .5$ . Intuitively, the person perceives state 0 as three times as likely as state 1, but values a marginal dollar in state 1 three times as much as in state 0, so  $q_t(0|H_t) = q_t(1|H_t)$  and  $\pi_0^* = .5$ .

One can also invert (9) to solve for  $\pi_t$  as a function of  $\pi_t^*$  and  $\phi$ , the solution to which we denote by  $\pi_t(\pi_t^*, \phi)$ :

$$\pi_t(\pi_t^*, \phi) = \frac{\pi_t^*}{\phi + (1 - \phi)\pi_t^*}. \quad (10)$$

While  $\pi_t$  is a time-varying unobservable, (10) clarifies that it can be represented as a function of the observable  $\pi_t^*$  and a single unknown parameter  $\phi$ .

We define the RN belief stream  $\pi^*(H_t)$ , RN belief movement  $m^*(\pi^*)$ , RN initial uncertainty  $u_0^*(\pi^*)$ , and RN excess movement  $X^*(\pi^*)$  as in Section 2.1, but with RN beliefs  $\pi_t^*$  in the place of physical beliefs  $\pi_t$ . It will also be useful to define the *RN measure* as

$$\mathbb{P}^*(H_T) \equiv \begin{cases} \mathbb{P}(H_T) \frac{\pi_0^*}{\pi_0} & \text{if } \pi_T(H_T) = 1 \\ \mathbb{P}(H_T) \frac{1 - \pi_0^*}{1 - \pi_0} & \text{if } \pi_T(H_T) = 0, \end{cases} \quad (11)$$

where  $\mathbb{P}(H_T)$  is the probability of observing history  $H_T$  under *DGP*. We show in Section 3 that (11) follows from the usual definition of the RN measure in a general asset-pricing setting (see Lemma A.1 in Appendix A). For now, it suffices to think of  $\mathbb{P}^*$  as representing the change of measure that adjusts the frequency of each path of signal realizations such that a person with RN beliefs has rational expectations: defining  $\mathbb{E}^*[\cdot]$  to be the expectation under  $\mathbb{P}^*$ , we have from (11) that

$$\pi_t^*(H_t) = \mathbb{E}^*[\pi_{t+1}^*(H_{t+1})|H_t]. \quad (12)$$

RN beliefs are not well-calibrated under the physical measure; for instance, in the first example in Table 1, the person has a RN prior of .5 but ends up with a belief of 1 only 25% of the time. Instead, the RN measure adjusts the weights on signal realizations to .375, .125, .125, and .375, respectively, so that the person ends up with a belief of 1 with 50% RN probability and is well-calibrated.

Given (12), the expectation of movement of RN beliefs under the RN measure must equal RN

initial uncertainty and therefore expected excess RN movement must be zero:

$$\mathbb{E}^*[X^*] = 0. \quad (13)$$

We wish to make statements similar to [Proposition 1](#), but applicable to the observable asset values. The fundamental challenge in doing so is twofold. First, asset values allow for the calculation of RN beliefs rather than physical beliefs; that is, we observe  $\pi_t^*$  rather than  $\pi_t$ , so we cannot construct  $\mathbb{E}[X]$  in [equation \(4\)](#). But second, the frequency of the observed RN belief streams used to construct expected excess movement is determined by the physical measure rather than the RN measure; that is, we cannot directly observe an empirical counterpart to  $\mathbb{E}^*[X^*]$  in [equation \(13\)](#).

Instead, we can observe a counterpart to  $\mathbb{E}[X^*]$ . But the distortion in RN relative to subjective beliefs can cause RN movement to differ from RN initial uncertainty on average, so  $\mathbb{E}^*[X^*] \neq 0$ . Thus even with rational physical beliefs, one can observe, for example, what appears to be excess movement in RN beliefs implied by valuations. (The third example in [Table 1](#), discussed further below, is one such case.) So if we naïvely test for RE using [Proposition 1](#) on observed RN (rather than actual) beliefs, we may spuriously conclude that beliefs are excessively volatile.

### 2.3 Risk-Neutral Beliefs: Results

We now proceed to derive theoretical bounds for  $\mathbb{E}[X^*]$  given a particular value of the unobservable parameter  $\phi$ . For any observed value for  $\mathbb{E}[X^*]$ , we can then place restrictions on the admissible values of  $\phi$  under the null of RE by inverting the bounds. We sketch derivations for our main results in the text, and full formal proofs for all results are provided in [Appendix A](#).

#### Main Results

To understand the object  $\mathbb{E}[X^*]$ , we start with a key observation using [\(11\)](#):

$$\mathbb{P}^*(H_T|\theta) = \mathbb{P}(H_T|\theta). \quad (14)$$

Compared to the physical measure, the RN measure places higher likelihood of all signal histories resolving in state 1, but does so proportionally, so that likelihoods of signal histories *conditional on state 1* do not change.<sup>9</sup> While this is implied directly by the definition [\(11\)](#), it again applies under the usual definition of the RN measure in a general setting (see [Lemma A.2](#) in [Appendix A](#)).

[Equation \(14\)](#) implies that conditional expectations under the two respective measures are equal,  $\mathbb{E}^*[X^*|\theta] = \mathbb{E}[X^*|\theta]$ , which implies

$$\begin{aligned} \mathbb{E}^*[X^*] &= \pi_0^* \cdot \mathbb{E}^*[X^*|\theta = 1] + (1 - \pi_0^*) \cdot \mathbb{E}^*[X^*|\theta = 0] \\ &= \pi_0^* \cdot \mathbb{E}[X^*|\theta = 1] + (1 - \pi_0^*) \cdot \mathbb{E}[X^*|\theta = 0] = 0, \end{aligned} \quad (15)$$

---

<sup>9</sup>For example, in the first DGP of [Table 1](#), the signal histories  $[l, H]$  and  $[h, H]$  that imply state 1 occur in the physical measure with probability .0625 and .1875 respectively, so that  $\mathbb{P}([l, H]|\theta = 1) = .25$ . In the RN measure, these are increased by a factor of  $\pi_0^*/\pi_0 = 2$  to .125 and .375, respectively. Consequently,  $\mathbb{P}^*([l, H]|\theta = 1)$  remains .25.

where the last equality applies (13). For  $\mathbb{E}[X^*]$ , it is useful to separate this expectation as in (15):

$$\mathbb{E}[X^*] = \pi_0 \cdot \mathbb{E}[X^*|\theta = 1] + (1 - \pi_0) \cdot \mathbb{E}[X^*|\theta = 0]. \quad (16)$$

Finally, it will be helpful to define the difference in conditional expected  $X^*$  as

$$\Delta \equiv \mathbb{E}[X^*|\theta = 0] - \mathbb{E}[X^*|\theta = 1]. \quad (17)$$

As movement must always be positive, using equation (15),<sup>10</sup> we have that

$$\Delta \leq \pi_0^*. \quad (18)$$

Combining (15) and (16), one can express  $\mathbb{E}[X^*]$  as follows. This expression in turn generates a set of bounds relating objects from observable asset valuations to the unobservable structural parameter  $\phi$ . We assume throughout that [Assumptions 1–3](#) hold.

**PROPOSITION 2.** *For any DGP,*

$$\begin{aligned} \mathbb{E}[X^*] &= (\pi_0^* - \pi_0)(\Delta) \\ &= \left( \pi_0^* - \frac{\pi_0^*}{\phi + (1 - \phi)\pi_0^*} \right) (\mathbb{E}[X^*|\theta = 0] - \mathbb{E}[X^*|\theta = 1]). \end{aligned}$$

**PROPOSITION 3.** *For any DGP and any value for  $\Delta$ ,*

$$\mathbb{E}[X^*] \leq (\pi_0^* - \pi_0)\pi_0^*.$$

*Equivalently,*

$$\mathbb{E}[X^*] \leq \left( 1 - \frac{1}{\phi + (1 - \phi)\pi_0^*} \right) \pi_0^{*2}. \quad (19)$$

**COROLLARY 1.** *For any DGP and any values  $\Delta$  and  $\phi$ ,*

$$\mathbb{E}[X^*] \leq \pi_0^{*2}.$$

Additionally, if one is willing to make an assumption on the sign of  $\Delta$  (discussed shortly), the following stronger bound applies.

**COROLLARY 2.** *If  $\mathbb{E}[X^*|\theta = 0] \leq \mathbb{E}[X^*|\theta = 1]$ , for any DGP and any value for  $\phi$ ,*

$$\mathbb{E}[X^*] \leq 0.$$

[Proposition 2](#) shows that the expectation of RN excess movement  $\mathbb{E}[X^*]$  is equal to the product

---

<sup>10</sup>Equation (15) can be rewritten as  $\pi_0^* \cdot \mathbb{E}[m^*|\theta = 1] + (1 - \pi_0^*) \cdot \mathbb{E}[m^*|\theta = 0] = \pi_0^*(1 - \pi_0^*)$ . Thus, fixing  $\pi_0^*$ ,  $\mathbb{E}[m^*|\theta = 0]$  decreases monotonically in  $\mathbb{E}[m^*|\theta = 1]$ . Given that  $\mathbb{E}[m^*|\theta = 1] \geq 0$ ,  $\mathbb{E}[m^*|\theta = 0]$  must be bounded above by  $\pi_0^*$ . Therefore,  $\Delta$ , which can be rewritten as  $\mathbb{E}[m^*|\theta = 0] - \mathbb{E}[m^*|\theta = 1]$ , must be bounded above by  $\pi_0^*$ .

of  $\pi_0^* - \pi_0$  and  $\Delta \equiv \mathbb{E}[X^*|\theta = 0] - \mathbb{E}[X^*|\theta = 1]$ . Note that if  $\phi = 1$ , then  $\pi_0^* = \pi_0$  and therefore  $\mathbb{E}[X^*] = 0$ . As  $\phi$  rises,  $\pi_0$  drops further below  $\pi_0^*$  and  $\mathbb{E}[X^*]$  deviates from 0: the greater is risk aversion, the more one can observe excess RN movement differ from zero on average under RE. The size of this deviation, and whether it is positive or negative, depends on  $\Delta$ . In other words, all DGPs with the same  $\Delta$  will produce the same level of  $\mathbb{E}[X^*]$  for a given  $\phi$ .

Regardless of the signal distribution,  $\Delta$  is bounded above as in (18). Plugging this into [Proposition 2](#) yields [Proposition 3](#). The version of the bound expressed in (19) is one of our main results. It gives a bound for  $\mathbb{E}[X^*]$  as a function of  $\pi_0^*$  and  $\phi$ , regardless of  $\Delta$ . Equivalently, it relates the unobserved structural parameter  $\phi$ , which corresponds to the slope of the SDF across the two states as in (6), to a set of observable values. Under risk neutrality ( $\phi = 1$ ), this upper bound becomes zero, following [Proposition 1](#). But this bound is otherwise positive, and the *admissible excess movement* in RN beliefs given by the right side of the inequality increases monotonically in  $\phi$ . Movement in RN beliefs must still correspond on average to the agent learning something about the true terminal state, but the bias in RN beliefs relative to subjective beliefs induced by risk aversion allows for positive excess movement in those observed beliefs under RE. This result thus formalizes a more general notion of the admissible amount of belief volatility under rationality, this time as an increasing function of risk aversion across the two states.

The bound in [Proposition 3](#) is very conservative, as it implicitly holds under a “worst-case” DGP with an extreme value for  $\Delta$ . Higher RN priors yield more “room” for downward belief movement, so the worst case  $\Delta = \pi_0^*$  rises monotonically. For  $\phi > 1$ , this generates asymmetry in the bound around  $\pi_0^* = .5$ , as lower values of  $\pi_0^*$  generally yield a lower upper bound for  $\mathbb{E}[X^*]$ .<sup>11</sup> But the bound in (19) is also non-monotonic in  $\pi_0^*$  for  $1 < \phi < \infty$  (in particular for  $\pi_0^* > 0.5$ ): for a given  $\phi$ , the first term  $\pi_0^* - \pi_0$  does not rise monotonically, exerting a countervailing force relative to the monotonic increase in the second term  $\pi_0^*$ .<sup>12</sup>

While the bound in [Proposition 3](#) maps between observed values and the slope of the SDF required to rationalize those values, the fact that RN beliefs are bounded between 0 and 1 by construction implies that the bound is well-defined even for infinitely large risk aversion: there is only so far that RN beliefs can be distorted relative to subjective beliefs. Taking  $\phi \rightarrow \infty$  in (19) yields [Corollary 1](#), which provides our most conservative bound for  $\mathbb{E}[X^*]$ . In the limit as  $\phi \rightarrow \infty$ ,  $\pi_0 \rightarrow 0$  for any  $\pi_0^*$ , so  $\pi_0^* - \pi_0 \rightarrow \pi_0^*$  and the bound in [Proposition 3](#) approaches  $\pi_0^{*2}$ . That is,  $\mathbb{E}[X^*] > \pi_0^{*2}$  simply cannot be rationalized under RE given constant  $\phi$ . For low  $\pi_0^*$ , this bound again does not allow large positive  $\mathbb{E}[X^*]$ . For example, if  $\pi_0 = .2$ , RN excess movement can only rise to .04, regardless of  $\phi$  or the DGP. That is, even arbitrarily high risk aversion does not allow for large amounts of excess movement in RN beliefs implied by asset valuations.

Taken together, [Proposition 3](#) and [Corollary 1](#) characterize the maximal admissible excess movement in RN beliefs as a function of  $\phi$  for any RN prior. [Figure 1](#) provides a graphical illustration of these bounds. Starting from the bottom of the chart, the thick purple line corresponds

<sup>11</sup>For example, with  $\pi_0^* = .2$  and  $\phi = 2$ , (19) gives  $\mathbb{E}[X^*] \leq .018$ . This implies the maximal expected RN movement,  $\max \mathbb{E}[m^*] = 0.178$ , is barely above  $\mathbb{E}[m] = 0.16$  for physical beliefs in the analogous case with  $\pi_0 = .2$ .

<sup>12</sup>For example, take  $\phi = 3$ . If  $\pi_0^* = .5$ , then  $\pi_0 = .25$  and  $\pi_0^* - \pi_0 = .25$ . But when  $\pi_0^* = .99$ ,  $\pi_0^* - \pi_0 = .02$ .

to the bound for  $\phi = 1$ : in this case,  $\mathbb{E}[X^*] = \mathbb{E}[X] = 0$  regardless of the prior or DGP, from [Proposition 1](#). The thin dashed gray lines correspond to arbitrarily selected DGPs in the case of  $\phi = 3$ . While there can be positive RN excess movement, this is not necessarily the case for all possible DGPs. Taking the envelope over all of these processes for  $\phi = 3$  yields the bound from [Proposition 3](#), which is shown in the thick blue line. It is asymmetric around 0.5 as well as non-monotonic, as described above. Finally, the thick red line shows the bound for the limiting case  $\phi \rightarrow \infty$ , which is equal to the squared RN prior from [Corollary 1](#).

Finally, when  $\Delta < 0$ ,  $\mathbb{E}[X^*]$  is decreasing in  $\phi$  and therefore  $\mathbb{E}[X^*] < 0$  for any  $\phi > 1$ . Consequently, as formalized in [Corollary 2](#), the highest excess movement is  $\mathbb{E}[X^*] = 0$ .

The above conclusions can also be seen in the examples in [Table 1](#). For the first example, the RN prior is .5 and the signals are symmetric, so that  $\Delta = 0$  and  $\mathbb{E}[X^*] = 0$ . For the second example, the DGP is asymmetric with large movements upwards, so that expected movement conditional on  $\theta = 1$  is large (.405) compared to  $\theta = 0$  (.095), so that  $\Delta$  is negative (-.31) and therefore  $\mathbb{E}[X^*]$  is negative. The converse occurs in the third example, which is asymmetric in the opposite way and therefore leads to  $\Delta = -.31$ . These examples point to the relationship of  $\pi_0^*$  and the DGP with the sign of  $\Delta$ . We now study this relationship in more detail.

### What Determines $\Delta$ ?

[Proposition 2](#) tells us that the deviation of  $\mathbb{E}[X^*]$  from 0 depends on the product of  $\pi_0^* - \pi_0$  and  $\Delta \equiv \mathbb{E}[X^*|\theta = 0] - \mathbb{E}[X^*|\theta = 1]$ . The difference  $\pi_0^* - \pi_0$  is always positive (strictly positive, for  $\phi > 1$ ) and increases in  $\phi$ . But how is the sign and magnitude of  $\Delta$  related to the RN prior and DGP? To answer this question, we provide two theoretical results and briefly summarize a set of numerical simulations discussed in detail in [Appendix B.2](#).

First, given the arbitrary labeling of the two states, there is no reason to expect under RE that  $\Delta$  should take a particular sign:

**PROPOSITION 4.** *Fixing  $\phi$ , for every RN prior and DGP that leads to a given  $\Delta$ , there exists a different RN prior and DGP that leads to  $-\Delta$ .*

Intuitively, for any  $\pi_0^*$  and DGP with some  $\Delta$ , the RN prior  $1 - \pi_0^*$  with the “reversed” DGP will necessarily lead to  $-\Delta$ . Consequently, there is no reason to assume that  $\mathbb{E}[X^*]$  is more likely than not to be positive given  $\phi > 1$  under RE.

Next, we summarize a set of numerical simulations for a large family of DGPs as described in [Appendix B.2](#). The results there suggest some intuitive comparative statics. First, when the DGP changes such that downward movements become larger, expected RN movement conditional on state 0 rises, so that  $\Delta$  rises and  $\mathbb{E}[X^*]$  rises. Second, when upward movements become larger, the opposite occurs and  $\mathbb{E}[X^*]$  falls. Third, as  $\pi_0^*$  rises,  $\Delta$  rises. Intuitively, a higher  $\pi_0^*$  allows more movement potential toward state 0 (and less toward state 1), so that expected movement given state 0 rises (and movement given state 1 falls), leading to a higher  $\Delta$ . For example, when the DGP is symmetric,  $\Delta < 0$  if  $\pi_0^* < .5$ ,  $\Delta = 0$  if  $\pi_0^* = .5$ , and  $\Delta > 0$  if  $\pi_0^* > .5$ . Combining



these statements suggests that, for example, low  $\pi_0^*$  should lead to a negative  $\Delta$  unless the DGP is extremely asymmetric.<sup>13</sup> And as we show later, the empirical DGPs in our setting appear largely symmetric, so that we estimate  $\Delta < 0$  for  $\pi_0^* < .5$ .

These results also suggest that extreme values of  $\Delta$  only occur in highly asymmetric DGPs where movements in one direction are large and movements in the other direction are tiny. In fact, our upper bound is attainable (in the limit) given the most asymmetric DGP possible:

**PROPOSITION 5.** *There exists a sequence of DGPs, indexed by  $T$ , for which  $\mathbb{E}[X^*]$  approaches the bound in Proposition 3 as  $T \rightarrow \infty$ . For each DGP in this sequence, downward movements ( $\pi_{t+1}^* < \pi_t^*$ ) are resolving ( $\pi_{t+1}^* = 0$ ) and thus as large as possible, while upward movements are small ( $\pi_{t+1}^* - \pi_t^* \rightarrow 0$  as  $T \rightarrow \infty$ ). Meanwhile, the bound holds with strict inequality for any  $T < \infty$  as long as  $\phi > 1$  and  $\pi_0^* \in (0, 1)$ .*

One implication of this result is that the bound in Proposition 3 is approximately tight, as one can construct a DGP for which  $\mathbb{E}[X^*]$  is close to the bound for large  $T$ . Perhaps more important, though, is that it points to the bound’s conservatism: it holds under a somewhat perverse DGP that can be thought of as a “rare bonanzas” process, where with small probability the person receives news that the bad state ( $\theta = 1$ ) will not be realized (so  $\pi_{t+1}^* = 0$ ), and otherwise there is mostly uninformative bad news that increases  $\pi_{t+1}^*$  slightly. More reasonable DGPs, or  $T \ll \infty$ , will give lower  $\mathbb{E}[X^*]$ . That said, the conservative bound has the advantage of being very simple and not requiring any estimation of  $\Delta$ . And as we show below, empirical excess movement is in fact so high that even these conservative bounds are often violated for reasonable values of  $\phi$ .

### What Generates Bound Violations?

We have assumed throughout that Assumption 1 holds, which in general requires both a correct subjective prior  $\pi_0 = \mathbb{P}_0(\theta = 1)$  and rational updating using the true signal distribution  $DGP(s_t | \theta, H_{t-1})$  as the likelihood. One natural question is whether an incorrect prior by itself can generate violations of the upper bound for excess RN movement, or whether excessive movement in general requires incorrect updating.<sup>14</sup> The following proposition, which is straightforward to show, makes clear that the latter is likely to be necessary for a violation of the bound.

**PROPOSITION 6.** *Assume that Assumptions 2–3 continue to hold. In place of Assumption 1, however, assume that the person has an incorrect prior,  $\pi_0 \neq \mathbb{P}_0(\theta = 1)$ , but updates correctly, in the sense that  $\pi_t(H_t) \propto \pi_{t-1}(H_{t-1})DGP(s_t | \theta = 1, H_{t-1})$ . Define  $\check{\phi} \equiv \phi L$ , where the person’s  $\phi$  is as in (6) and where  $L \equiv \frac{\pi_0/(1-\pi_0)}{\mathbb{P}_0(\theta=1)/(1-\mathbb{P}_0(\theta=1))}$  indexes the prior belief distortion, with  $0 < L < \infty$ . Then:*

- (i) *For all  $H_t$ , the person’s RN beliefs  $\pi_t^*$  are equivalent to the RN beliefs of a fictitious agent whose physical beliefs  $\{\check{\pi}_t\}$  satisfy Assumption 1 but who has  $\check{\phi}$  in place of  $\phi$ .*

<sup>13</sup>For example, in the large class of DGPs we simulate, only 2% of DGPs have a positive  $\Delta$  when  $\pi_0^* < .25$ . For  $\pi_0^* < .5$ , it is 11%. When  $\pi_0^*$  is low, the only DGPs in which  $\Delta > 0$  are asymmetric and extreme. For example, when  $\pi_0^* = .25$ ,  $\Delta > 0$  only occurs if the likelihood ratio of the downward signals is at least two times the likelihood ratio of the upward signals. Thus Corollary 2 (which requires  $\Delta < 0$ ) appears likely to be applicable for low values of  $\pi_0^*$ .

<sup>14</sup>One might think that the distinction between the person’s prior and likelihood is artificial both formally and empirically, but we think the result here nonetheless helps clarify what kinds of belief streams violate the bounds.

- (ii) If  $\check{\phi} \geq 1$ , then the restrictions on  $\mathbb{E}[X^*]$  above ([Propositions 2–3](#), [Corollaries 1–2](#)) continue to hold, with  $\check{\phi}$  in place of  $\phi$  and  $\check{\pi}_0$  in place of  $\pi_0$ . In particular, one cannot in this case have  $\mathbb{E}[X^*] > \pi_0^{*2}$ .
- (iii) If  $\check{\phi} < 1$  so that  $\pi_0^* < \mathbb{P}_0(\theta = 1) = \check{\pi}_0$ , then [Proposition 2](#) continues to hold, but the bound in [Proposition 3](#) becomes  $\mathbb{E}[X^*] \leq (\check{\pi}_0 - \pi_0^*)(1 - \pi_0^*)$ , and [Corollary 1](#) becomes  $\mathbb{E}[X^*] \leq (1 - \pi_0^*)^2$ . Thus regardless of  $\check{\phi}$ , it must be the case that  $\mathbb{E}[X^*] \leq \max(\pi_0^{*2}, (1 - \pi_0^*)^2)$ .

Part (i) formalizes that risk aversion is isomorphic to an incorrect prior, in that both have the same effect on  $\pi_t^*$  relative to the true  $\mathbb{P}_t(\theta = 1)$ . Thus with a suitably altered value of  $\phi$ , the bounds generally cover the case of an incorrect prior, as in part (ii). The only case in which this argument requires slight amendment is when the prior is so downwardly distorted that  $\pi_0^* < \mathbb{P}_0(\theta = 1)$ . Even in this case, though, a slightly altered version of [Corollary 1](#) still applies, as in part (iii). An incorrect prior acts as a one-time belief distortion; while reverting to the correct belief in this case does require some excess movement, this is generally not sufficient for a full violation of the bound in [Proposition 3](#). In general, then, incorrect updating behavior must be present in such a violation, and the restriction imposed by our bound implies that this incorrect updating behavior necessarily entails excessive volatility in beliefs relative to the degree of uncertainty resolution over time.

### 3. Generalized Theoretical Results for Equilibrium Asset Prices

While the setting considered in the previous section is useful for providing intuition, it is also artificial: one cannot obtain a single person’s valuation of Arrow-Debreu claims in observational data alone; there are more than two possible states; and the realized state determines more than just consumption. We thus now consider a general many-state framework for equilibrium asset prices and show how our results extend to this case. We first set up the framework and define notation ([Section 3.1](#)), then discuss the main assumption used to derive our bounds ([Section 3.2](#)), and finally describe how our results apply in this case ([Section 3.3](#)). This generalized setting will also allow us to more thoroughly consider the robustness and interpretation of our results, as we do in [Section 4](#).

#### 3.1 Setup and Notation

##### Preliminaries: Probability Space, Prices, and Risk-Neutral Probabilities

Time is again indexed by  $t \in \{0, 1, 2, \dots\}$ , and we consider a discrete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with filtration  $\{H_t\}$ .<sup>15</sup> A realization of the elementary state is denoted by  $\omega \in \Omega$ . To make our results empirically implementable, we will be concerned with the ex-dividend value of the market index,  $V_t^m: \Omega \rightarrow \mathbb{R}_+$ , on some option expiration date  $T$ . (We will later extend the notation to allow for multiple option expiration dates.) A European call option on the index with strike price  $K$  has date- $T$  payoff  $(V_T^m - K)^+ = \max(V_T^m - K, 0)$ , and its time- $t$  price is  $q_{t,K}^m$ . Assume without loss of

<sup>15</sup>We could extend to continuous state spaces with additional technicalities, but do not do so given that empirical implementation requires discretization and that our results are better understood for probabilities than densities. We note as well that objects analogous to those in [Section 2](#) are given the same denotation in this section.

generality that these option prices are observable for some set of strike prices  $\mathcal{K} \subseteq \mathbb{R}_+$  beginning at date 0. Assuming the absence of arbitrage, there exists a strictly positive *stochastic discount factor* (SDF)  $M_{t,T}$  such that option prices satisfy  $q_{t,K}^m = \mathbb{E}_t[M_{t,T}(V_T^m - K)^+]$ , where  $\mathbb{E}_t[\cdot] \equiv \mathbb{E}^{\mathbb{P}}[\cdot | H_t]$ .

Option prices will be of interest for inferring a distribution over the change in value of the market index from 0 to  $T$  (rather than consumption, for which options are not directly traded). We say that *return state*  $\theta \in \Theta \subset \mathbb{R}_+$  is realized for the market as of date  $T$  if  $R_T^m \equiv V_T^m / V_0^m = \theta$ . The measure  $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$  governs the *objective* or *physical* probabilities of these return states. In this general case, the *risk-neutral (RN) measure* is defined by the change of measure

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} \Big|_{H_t} = \frac{M_{t,T}}{\mathbb{E}_t[M_{t,T}]} \quad (20)$$

and expectations under  $\mathbb{P}^*$  are denoted by  $\mathbb{E}^*[\cdot]$ .<sup>16</sup> Using this definition of  $\mathbb{P}^*$ , the RN probability of return state  $\theta$  is

$$\mathbb{P}_t^*(R_T^m = \theta) = \frac{\mathbb{E}_t[M_{t,T} | R_T^m = \theta]}{\mathbb{E}_t[M_{t,T}]} \mathbb{P}_t(R_T^m = \theta). \quad (21)$$

The RN pricing equation  $q_{t,K}^m = \mathbb{E}_t^*[(V_T^m - K)^+] / R_{t,T}^f$  (see [footnote 16](#)) can be used to show that the date- $t$  option prices  $\{q_{t,K}^m\}_{K \in \mathcal{K}}$  reveal the set of RN probabilities  $\{\mathbb{P}_t^*(R_T^m = \theta)\}_{\theta \in \Theta}$ . Assume that the set of return states  $\Theta$  is ordered such that  $\theta_1 < \theta_2 < \dots < \theta_J$ , and assume for notational simplicity that the set of traded option strikes  $\mathcal{K}$  coincides with the set of possible date- $T$  index values (i.e.,  $\mathcal{K} = \{K_j\}_{j=1, \dots, J}$ , with  $K_j = V_0^m \theta_j$ ). We can then back out RN probabilities from option prices as follows:

$$\mathbb{P}_t^*(R_T^m = \theta_j) = R_{t,T}^f \left[ \frac{q_{t,K_{j+1}}^m - q_{t,K_j}^m}{K_{j+1} - K_j} - \frac{q_{t,K_j}^m - q_{t,K_{j-1}}^m}{K_j - K_{j-1}} \right]. \quad (22)$$

See Appendix A.2 for a derivation of this result, which follows directly from a discrete-state application of the classic result of [Breedon and Litzenberger \(1978\)](#).<sup>17</sup>

## Beliefs

Aside from assuming no arbitrage, we have not yet taken a stance on the market structure or subjective beliefs underlying prices and RN probabilities. We could in principle pursue a strict mapping from [Section 2](#) to the current case, by assuming a setting in which all individual traders in the economy have common beliefs satisfying RE with respect to the return state. The assumptions required to generate such an equilibrium are well studied in the literature on information and asset

<sup>16</sup>Since  $\text{Price}_t = \mathbb{E}_t[M_{t,T} \text{Payoff}_T]$  for any asset,  $R_{t,T}^f = \mathbb{E}_t[M_{t,T}]^{-1}$ , where  $R_{t,T}^f$  is the gross risk-free rate corresponding to an asset with payoff 1 at  $T$ . One can thus rewrite  $\text{Price}_t = \mathbb{E}_t^*[\text{Payoff}_T] / R_{t,T}^f$ , so  $\mathbb{P}^*$  incorporates the risk adjustment needed to discount  $T$ -payoffs at the risk-free rate from  $t$  to  $T$ . This definition of  $\mathbb{P}^*$  is sometimes referred to as the  $T$ -forward measure (e.g., [Jamshidian, 1989](#), [Qin and Linetsky, 2017](#)).

<sup>17</sup>As can be seen from the result,  $\mathcal{K}$  need not coincide with  $V_0^m \times \Theta$  to back out  $\mathbb{P}_t^*(R_T^m = \theta_j)$ , as strikes at (near)  $V_0^m \theta_{j-1}$ ,  $V_0^m \theta_j$ , and  $V_0^m \theta_{j+1}$  are sufficient for exact (approximate) measurement.

prices following [Radner \(1979\)](#) and [Milgrom and Stokey \(1982\)](#).<sup>18</sup> But rather than maintaining this microeconomic focus on testing the rationality of individual beliefs, we prefer an interpretation in which the “agent” in question is the market as a whole (or alternatively, the marginal trader); such an interpretation requires no auxiliary assumptions, and our resultant tests are informative about the efficiency of market valuations. We thus cannot and do not seek to rule out explanations for our empirical results based on belief heterogeneity, where all individual traders begin with different priors (about which they agree to disagree) and update rationally, but where the identity of the marginal trader changes with the wealth distribution (as, e.g., in [Martin and Papadimitriou, 2021](#)).

We therefore assume prices correspond to valuations for an agent (“the market”) who, at the beginning of each period, observes a signal vector  $s_t \in \mathcal{S}$  drawn from the distribution  $DGP(s_t | \theta, H_{t-1}) = \mathbb{P}_{t-1}(s_t | \theta)$ , where  $\theta$  is the return state realized at  $T$  and  $H_t = \sigma(s_\tau, 0 \leq \tau \leq t)$  is the Borel  $\sigma$ -algebra representing the history of signal realizations. The agent’s subjective belief distribution over return states is denoted by  $\Pi_{t,T} = \{\pi_t(R_T^m = \theta)\}_{\theta \in \Theta}$ , where  $\pi_t(R_T^m = \theta) \geq 0 \forall \theta \in \Theta$  and  $\sum_{\theta \in \Theta} \pi_t(R_T^m = \theta) = 1$ . More generally, for any random variable  $Y(\omega)$ , the agent attaches subjective probability  $\pi_t(Y = y)$  to the outcome  $Y = y$ . We generalize [Assumption 1](#) as:

**ASSUMPTION 1’ (RE).** For any random variable  $Y$ , beliefs satisfy  $\pi_t(Y = y) = \mathbb{P}_t(Y = y)$  with probability 1 for all  $t$ .

An immediate implication of this assumption, from the law of iterated expectations, is that beliefs satisfy a martingale condition analogous to [Assumption 1](#),  $\pi_t(R_T^m = \theta) = \mathbb{E}_t[\pi_{t+1}(R_T^m = \theta)]$  for all  $\theta \in \Theta$ . As in [Section 2](#), this martingale condition for beliefs over returns is all that is required for our main results to carry through. The full-RE generalization stated in [Assumption 1’](#), though, will be useful for streamlining some of the remaining discussion, as it further implies that *all* conditional expectations — including over the SDF — are a martingale under the objective measure.

Given [Assumption 1’](#), we can define the RN belief distribution without explicitly restricting the agent’s utility or constraint set by applying the same change of measure as in (20), using the general SDF  $M_{t,T}$ . This yields the RN belief-distribution  $\Pi_{t,T}^* = \{\pi_t^*(R_T^m = \theta)\}_{\theta \in \Theta}$  such that  $\pi_t^*(R_T^m = \theta) = \frac{\mathbb{E}_t[M_{t,T} | R_T^m = \theta]}{\mathbb{E}_t[M_{t,T}]} \pi_t(R_T^m = \theta)$  as in (21), and thus (22) tells us that option prices reveal the agent’s RN beliefs as given here.

### Localization: Conditional Beliefs, SDF Ratio, and Excess Movement

To align with the analysis in [Section 2](#), we consider the behavior of *conditional* RN beliefs over adjacent pairs of return states. That is, rather than directly considering the full distribution  $\Pi_{t,T}^*$ , we instead consider restrictions on the behavior of the individual entries in  $\{\tilde{\pi}_{t,j}^*\}_{j=1,\dots,J-1}$ , defined by

$$\tilde{\pi}_{t,j}^* \equiv \pi_t^*(R_T^m = \theta_j | R_T^m \in \{\theta_j, \theta_{j+1}\}) = \frac{\pi_t^*(R_T^m = \theta_j)}{\pi_t^*(R_T^m = \theta_j) + \pi_t^*(R_T^m = \theta_{j+1})}, \quad (23)$$

<sup>18</sup>Complete markets and a common-prior assumption, for example, are sufficient: prices in general reveal information (including private signals) in a rational expectations equilibrium, giving common posteriors. Results under alternative conditions have also been studied extensively (to take one example, see [Blume, Coury, and Easley, 2006](#)).

for  $\pi_t^*(R_T^m = \theta_j) + \pi_t^*(R_T^m = \theta_{j+1}) > 0$ . In words,  $\tilde{\pi}_{t,j}^*$  is the RN belief that return state  $\theta_j$  will be realized, conditional on either  $\theta_j$  or  $\theta_{j+1}$ . This binary localization will be useful for two reasons: (i) it will allow us to apply results from the two-state setting of [Section 2](#), and (ii) the main identifying assumption used to derive our tests is less restrictive than it would be without such a transformation (as discussed in [Section 3.2](#) below). Conditional physical beliefs  $\tilde{\pi}_{t,j}$  are defined analogously, and the expectation under the conditional physical measure is  $\tilde{\mathbb{E}}_t[\cdot] \equiv \mathbb{E}_t[\cdot | R_T^m \in \{\theta_j, \theta_{j+1}\}]$ .

In this context, the analogue to  $\phi_t$  as defined in (6) is

$$\phi_{t,j} \equiv \frac{\mathbb{E}_t[M_{t,T} | R_T^m = \theta_j]}{\mathbb{E}_t[M_{t,T} | R_T^m = \theta_{j+1}]}, \quad (24)$$

which encodes the slope of the SDF across the adjacent return states  $\theta_j, \theta_{j+1}$ . In a representative-agent economy with consumption  $C_t$ , time-separable utility, and time discount factor  $\beta$ , the SDF is  $M_{t,T} = \beta^{T-t} U'(C_T) / U'(C_t)$ , and thus  $\phi_{t,j} = \mathbb{E}_t[U'(C_T) | R_T^m = \theta_j] / \mathbb{E}_t[U'(C_T) | R_T^m = \theta_{j+1}]$ , similar to the marginal rate of substitution in (7). But the representation in (24) is general and does not require a representative-agent structure (though we make periodic reference to such an economy for interpretation). Return state  $\theta_j$  corresponds to state 1 (vs. state 0 for  $\theta_{j+1}$ ) in [Section 2](#), which motivates our use of the index  $j$  to refer to variables conditioned on the pair  $(\theta_j, \theta_{j+1})$ . We maintain the convention of labeling  $\theta_j$  as the “bad” (i.e., low) return state, so that [Assumption 2](#) becomes:

**ASSUMPTION 2'** (*Labeling of Return States*).  $\phi_{t,j} \geq 1$  with probability 1 for all  $t, j$ , where the set of return states  $\Theta$  is ordered such that  $\theta_1 < \theta_2 < \dots < \theta_J$ .

Using (23)–(24), we can write the RN odds ratio as  $\frac{\tilde{\pi}_{t,j}^*}{1 - \tilde{\pi}_{t,j}^*} = \phi_{t,j} \frac{\tilde{\pi}_{t,j}}{1 - \tilde{\pi}_{t,j}}$ , from which it is clear that (9) and (10) hold in this context for  $\tilde{\pi}_{t,j}^*, \tilde{\pi}_{t,j}$ , and  $\phi_{t,j}$ . Given the resolving RN belief stream  $\pi_j^* = [\tilde{\pi}_{0,j}^*, \dots, \tilde{\pi}_{T,j}^*]$ , RN belief movement  $m_j^*$ , RN initial uncertainty  $u_{0,j}^*$ , and RN excess movement  $X_j^*$  are as defined in (1)–(3), with  $\tilde{\pi}_{t,j}^*$  in place of  $\pi_t$ . We commonly suppress the dependence on  $j$  (writing, e.g.,  $X^*$ ) when considering an arbitrary state pair.

### 3.2 Restriction on the SDF

We now confront the joint hypothesis problem, restating the main assumption ([Assumption 3](#)) we make alongside the maintained hypothesis of RE ([Assumption 1](#)). While the usefulness of this assumption was described briefly in [Section 2](#), discussing it in this context will clarify why it is appealing economically: while it does impose a restriction on the SDF, the restriction is sufficiently general to allow for identification in a broad class of models. First, restating the assumption:

**ASSUMPTION 3'** (*Constant SDF Ratio, or Conditional Transition Independence*). We say the SDF satisfies *conditional transition independence (CTI)* for the return-state pair  $(\theta_j, \theta_{j+1})$  if  $\phi_{t,j}$  is constant with probability 1 for all  $t$ . We assume CTI is satisfied for all *interior* state pairs,  $j = 2, 3, \dots, J - 2$ .

#### REMARKS:

1. As in [Section 2](#), this assumption requires that changes in the RN odds ratio arise from changes

in the subjective belief  $\tilde{\pi}_{t,j}$  rather than changes in the expected relative severity of the adjacent return states  $\theta_j$  and  $\theta_{j+1}$ .

2. We exclude the extreme state pairs  $(\theta_1, \theta_2)$  and  $(\theta_{J-1}, \theta_J)$  from the constant- $\phi$  requirement. This will be useful theoretically: thinking of  $\theta_1$  and  $\theta_J$  as tail return states, we are allowing for time-varying disaster (or positive-jump) severity.<sup>19</sup> This accordingly does *not* require that all changes in the underlying RN belief distribution arise from changes in subjective beliefs  $\pi_t(R_T^m = \theta_j)$ : there may be simultaneous time variation in the values in the numerator and denominator in (24), and the value  $\mathbb{E}_t[M_{t,T}]$  need not be constant.
3. The assumption corresponds to a notion of transition or path independence because it implies that the realization of  $M_{t,T}$  in return state  $\theta_j$ , relative to  $\theta_{j+1}$ , depend in expectation only on the return state pair and not on the path of any variables realized between  $t$  and  $T$ . This intuition is formalized in Lemma A.1, and it motivates the CTI label we use to reference it.<sup>20</sup>
4. If, analogous to the case in Section 2, a representative agent's utility depends only on the maturity value of the market index, then  $\phi_{t,j}$  is constant.<sup>21</sup> But the assumption holds in more general cases as well. For example, permanent shocks to the SDF that raise  $M_{t,T}$  in all states are admissible under the assumption. Given that permanent shocks to the SDF are empirically much more important for SDF variation than transitory shocks (Alvarez and Jermann, 2005), the assumption appears reasonable to a first order. (In addition, some transitory shocks are also admissible; see the first two examples below.)

We now turn to a set of examples to illustrate the restriction concretely.

**EXAMPLE 1.** Assume an economy with a one-dimensional state variable  $A_t: \Omega \rightarrow \mathbb{R}$  (e.g., productivity, consumption, volatility), with  $dV_t^m/dA_t > 0$ . This process may be non-stationary but is assumed to satisfy the Markov property. Assume further that there exists a representative agent with time-separable utility over the consumption process  $\{C_t(A_t)\}$  and that the market index pays dividends  $\{D_t(A_t)\}$ , where these are arbitrary but yield a stationary price-dividend ratio. Then CTI holds for any two adjacent return states. If, in addition, consumption or consumption growth is i.i.d. over  $t$ , then CTI holds as well if the agent instead has Epstein–Zin recursive utility.

While the assumption of a scalar Markov forcing process in this example is restrictive, it nonetheless encompasses some simple leading cases that are useful for intuition. For example, with log consumption as the state variable ( $A_t = c_t \equiv \log(C_t)$ ), we could have  $c_t = g + \rho c_{t-1} + h(c_{t-1})\varepsilon_t$  where  $\varepsilon_t$  is i.i.d. with arbitrary distribution. This example allows for time variation in discount rates and risk premia (see Appendix B.1), and it also illustrates that both temporary and permanent

<sup>19</sup>Empirically, return states will be defined as narrow intervals for the index return with the exception of  $\theta_1$  and  $\theta_J$ , which contain the left and right tails of the distribution, respectively; it is likely that, e.g.,  $\mathbb{E}[M_{t,T} | R_T^m \leq -20\%]$  changes over time as the distribution of returns conditional on  $R_T^m \leq -20\%$  changes.

<sup>20</sup>“Conditional” is used to underscore that it requires only constancy of the ratio of conditional expectations of the SDF, rather than requiring deterministic state-contingent SDF realizations (as in the transition independence assumptions used by, among others, Ross, 2015, which are thus more restrictive). Borovička, Hansen, and Scheinkman (2016) show that Ross's assumptions rule out permanent shocks to the SDF, while ours does not.

<sup>21</sup>This holds more generally as long as there is some agent whose indirect utility can be written as a function only of the terminal index value (e.g., an investor retiring at date  $T$  with savings fully invested in the market).

shocks to the SDF are admissible in principle.<sup>22</sup> The same is true of the following example, which considers a well-known (and more fully specified) model featuring time-varying risk premia and permanent SDF shocks (see [Bakshi and Chabi-Yo, 2012](#)), and which has been advanced as a rationalization of the excess-volatility puzzle.

**EXAMPLE 2.** Consider the variable rare-disasters model of [Gabaix \(2012\)](#), described fully in Appendix B.3. Given any market-index option horizon  $T$  and any (small) positive value  $\delta$ , there exists a return state  $\underline{\theta}$  such that for all  $\theta_j \geq \underline{\theta}$ , the conditional probability of having realized at least one disaster over the life of the option is negligible:  $\mathbb{P}_0(\sum_{t=1}^T \mathbb{1}\{\text{disaster}_t\} > 0 \mid R_T^m \geq \underline{\theta}) < \delta$ . For all  $\theta_j \geq \underline{\theta}$ , CTI holds for the state pair  $(\theta_j, \theta_{j+1})$  up to a negligible error, as  $\phi_{t,j} = \phi_j + \eta_t$  for  $\phi_j$  constant and  $\eta_t = o_p(1)$  as  $\delta \rightarrow 0$ .

This result implies that for an economy described by this model, we need only focus attention on RN beliefs across return states  $(\theta_j, \theta_{j+1})$  for which there is negligible probability of having realized a disaster conditional on reaching  $\theta_j$ . This helps illustrate the use of the localization provided by considering conditional beliefs, and this insight guides our empirical analysis in [Section 5](#).

Finally, our last example helps illustrate the types of settings in which CTI does *not* hold.

**EXAMPLE 3.** Consider the external habit-formation model of [Campbell and Cochrane \(1999\)](#), described fully in Appendix B.4. Under the assumptions stated there, CTI fails to hold for all  $j$ .

With habit formation, the path of consumption always matters in a manner not fully accounted for by conditioning on the return state. We note, though, that nothing about our theoretical framework requires considering beliefs over *return* states: all our results would apply for beliefs over elementary states in  $\Omega$ , or, in the [Campbell–Cochrane](#) case, over the joint realization of the terminal consumption and *surplus* consumption values. Empirical implementation in this case, though, is infeasible. We consider return states specifically in our theory because this allows us to map directly between the theory and data. We later consider the effects of the violation of CTI implied by the [Campbell–Cochrane](#) model in a calibrated simulation study.

### 3.3 Main Results in the General Setting

Having completed the formal setup and discussion of our main assumption, we turn now to our main results in this more general asset-pricing setting. The bulk of the work in this case is, it turns out, in the setup and notation, as all our main results apply with appropriate relabeling:

**PROPOSITION 7.** *Under no arbitrage and [Assumptions 1'–3'](#), for  $j = 2, 3, \dots, J - 2$ , [Propositions 1–6](#) and [Corollaries 1–2](#) continue to hold, with  $\tilde{\pi}_{t,j}^*$  replacing  $\pi_t^*$ ,  $\tilde{\pi}_{t,j}$  replacing  $\pi_t$ ,  $X_j^*$  replacing  $X^*$ ,  $\phi_j$  replacing  $\phi$ ,  $\tilde{\mathbb{E}}_0[\cdot]$  replacing  $\mathbb{E}[\cdot]$ , and with  $\Delta_j \equiv \tilde{\mathbb{E}}_0[X_j^* \mid R_T^m = \theta_{j+1}] - \tilde{\mathbb{E}}_0[X_j^* \mid R_T^m = \theta_j]$  replacing  $\Delta$ .*

The main theoretical complication in applying the results in [Section 2](#) to this setting is in proving that the RN measure, defined in a standard manner in (20), satisfies the property that pathwise

<sup>22</sup>For example, with CRRA utility and i.i.d. consumption-growth shocks, there are only permanent shocks to marginal utility; meanwhile, an economy with i.i.d. consumption has purely transitory shocks and a fixed  $\mathbb{E}_t[U_T^l \mid R_T^m = \theta_j] \forall j$ .

probabilities share the same value  $\tilde{\mathbb{P}}^*(H_T)/\tilde{\mathbb{P}}(H_T)$  conditional on  $R_T^m$ , as in (11). This result — along with its main implication, that  $\tilde{\mathbb{E}}^*[X_j^* | R_T^m = \theta_j] = \tilde{\mathbb{E}}[X_j^* | R_T^m = \theta_j]$ , as after (14) — is proven to hold in Appendix A.2 for a measure that is observationally equivalent to the RN measure, which is sufficient for our purposes.

The economic intuition for these results is largely identical to the intuition discussed in Section 2. Two additional points, though, are worth discussing in this setting. First, in the case of a bound violation  $\tilde{\mathbb{E}}[X_j^*] > \tilde{\pi}_{0,j}^2$ , the conclusion would not necessarily be that there are violations of the no-arbitrage condition; instead, there would be no rational-expectations SDF process satisfying CTI capable of generating the observed excess movement in RN beliefs. The *actual* SDF process translating between objective probabilities and RN beliefs would in this case include a market-level belief distortion that induces excess volatility.<sup>23</sup>

Second, while the results above are convenient to express in terms of the SDF slope  $\phi_j$  given that this allows for closed-form solutions that can be applied across a wide range of structural models regardless of the origin of the SDF, the results also admit an interpretation in terms of the approximate required risk-aversion value for a fictitious representative agent with utility over the terminal value of the market index. Analogous to equation (8) in Section 2, we have the following result. To avoid repetition, for the remaining results we will continue to assume no arbitrage and that Assumptions 1'–3' hold for  $j = 2, 3, \dots, J - 2$ , unless stated otherwise.

**PROPOSITION 8.** *Assume additionally that there is a representative agent with (indirect) utility over time- $T$  wealth, with wealth equal to the market index value, and denote  $V_j^m \equiv V_0^m \theta_j$ . Then local relative risk aversion  $\gamma_j \equiv -V_j^m U''(V_j^m)/U'(V_j^m)$  is given to a first order around return state  $\theta_j$  by*

$$\gamma_j = \frac{\phi_j - 1}{(V_{j+1}^m - V_j^m)/V_j^m}. \quad (25)$$

As in Section 2,  $\gamma_j$  is proportional to  $\phi_j - 1$ , as this gives the percent decrease in marginal utility in moving from low-return state  $\theta_j$  to high-return state  $\theta_{j+1}$ . To calculate relative risk aversion, this change in marginal utility must be normalized by the percent wealth increase  $(V_{j+1}^m - V_j^m)/V_j^m$  in moving from  $\theta_j$  to  $\theta_{j+1}$ , which is also equal to the percent return deviation  $(\theta_{j+1} - \theta_j)/\theta_j$  between the two states. If, for example,  $\theta_j = 1$ ,  $\theta_{j+1} = 1.05$ , then a value  $\phi_j = 1.5$  implies  $\gamma_j = 10$ .

## 4. Theoretical Results on Robustness and Empirical Implementation

Building on Section 3, we now turn to a set of additional robustness results. First, to obtain a direct mapping from theory to data, we consider the observability of the statistics in our bounds (Section 4.1) and how to account for possible mismeasurement or market microstructure noise (Section 4.2). We then extend our analysis to consider cases in which CTI is violated (Section 4.3).

<sup>23</sup>Such a finding would be close in spirit to a violation of the “good-deal bounds” of Cochrane and Saá-Requejo (2000): even if the no-arbitrage condition holds, there would be an investment strategy with a large Sharpe ratio under the objective measure — in particular, betting on mean reversion in RN beliefs — that is nonetheless not traded away.



## 4.1 Aggregating Over Belief Streams

The discussion to this point has taken as given that all the values in our bounds for expected excess RN movement are observable aside from  $\phi_j$ . But this has elided one issue: the bounds are stated with respect to date-0 expectations conditional on the RN prior, but we observe only one draw  $X_j^*$  per expiration date rather than the ex-ante expectation of this statistic for a given  $\tilde{\pi}_{0,j}^*$ . We now show how one can aggregate over multiple RN belief streams — i.e., over multiple option expiration dates — so as to implement the bounds empirically. In doing so, we allow  $\phi_j$  to change as well across belief streams, though we continue to assume that it is fixed within a given stream.<sup>24</sup>

To do so, we must further generalize the environment slightly. We now assume that we can observe index option prices for some set of  $N$  option expiration dates  $\mathcal{T} \equiv \{T_i\}_{i=1,\dots,N}$ , so that belief streams (and their DGPs) are indexed by  $i$ . We accordingly add the subscript  $i$  to refer to objects corresponding to expiration date  $T_i$ :  $\phi_{i,j}$  is the SDF ratio for date  $T_i$  and state pair  $(\theta_j, \theta_{j+1})$ , RN beliefs are  $\tilde{\pi}_{i,j}^*$ , and RN excess movement is  $X_{i,j}^*$ . To simplify notation for now, in order to focus specifically on the issue of aggregation for an arbitrary state pair, we maintain the  $i$  subscripts but otherwise revert to the notation from [Section 2](#) (e.g., using  $X_i^*$  in place of  $X_{i,j}^*$ ).

Due to Jensen's inequality, we cannot simply insert  $\mathbb{E}[\phi_i]$  in place of  $\phi_i$  or  $\mathbb{E}[\pi_{0,i}^*]$  in place of  $\pi_{0,i}^*$  (where these expectations are implicitly taken over DGPs  $i$ ) when taking the expectation of both sides of the results in [Propositions 2](#) and [3](#) (or their counterparts in [Proposition 7](#)). However, we can show that the following modified generalizations do hold:

**PROPOSITION 9.** Define  $\bar{\phi} \equiv \max_{\pi_{0,i}^*} \mathbb{E}[\phi_i | \pi_{0,i}^*]$ . We have:

- (i) GENERALIZATION OF [PROPOSITION 2](#): If  $\text{Cov}(\pi_{0,i}, \Delta_i) = 0$ , and  $\pi_{0,i}^*$  is constant across  $i$  (i.e., fixing a given  $\pi_{0,i}^*$ ), then over all streams,

$$\mathbb{E}[X_i^*] \leq \max \left\{ 0, \left( \pi_{0,i}^* - \frac{\pi_{0,i}^*}{\mathbb{E}[\phi_i] + (1 - \mathbb{E}[\phi_i])\pi_{0,i}^*} \right) \mathbb{E}[\Delta_i] \right\}. \quad (26)$$

- (ii) GENERALIZATION OF [PROPOSITION 3](#): Over all streams, we have without any additional assumptions that

$$\mathbb{E}[X_i^*] \leq \mathbb{E} \left[ \left( \pi_{0,i}^* - \frac{\pi_{0,i}^*}{\bar{\phi} + (1 - \bar{\phi})\pi_{0,i}^*} \right) \pi_{0,i}^* \right], \quad (27)$$

or, fixing a given  $\pi_{0,i}^*$ ,  $\mathbb{E}[X_i^*] \leq \left( \pi_{0,i}^* - \frac{\pi_{0,i}^*}{\mathbb{E}[\phi_i] + (1 - \mathbb{E}[\phi_i])\pi_{0,i}^*} \right) \pi_{0,i}^*$ .

- (iii) GENERALIZATION OF [COROLLARY 1](#): Over all streams, without any additional assumptions,

$$\mathbb{E}[X_i^*] \leq \mathbb{E} \left[ \pi_{0,i}^{*2} \right].$$

<sup>24</sup>While it might be palatable to assume that  $\phi_j$  is constant *within* a belief stream (over the course of months), it becomes more challenging to assume that  $\phi_j$  is constant *across* belief streams (over the course of years).

(iv) GENERALIZATION OF COROLLARY 2: If  $\Delta_i \leq 0$  for all  $i$ , then over all streams,

$$\mathbb{E}[X_i^*] \leq 0.$$

First consider part (ii), which generalizes Proposition 3. This result effectively applies Jensen’s inequality for one of several variables, using the fact that the second partial derivative of the bound with respect to  $\phi_{i,j}$  is negative. The bound is accordingly even more conservative than the original bound as stated for a single stream, as it effectively accounts for the worst-case possible relationship between RN priors and  $\phi_i$ . We also state the bound fixing a given value of  $\pi_{0,i}^*$ , because in some of our empirical results we report  $\mathbb{E}[X^*]$  across streams conditional on a particular value of this RN prior.<sup>25</sup> This bound is quite conservative as well; for example, if  $\pi_{0,i}^* = 0.5$  for  $i = 1, 2$ , with  $\phi_1 = 1$  and  $\phi_2 = 99$ , then under the maximizing DGP for  $\mathbb{E}[X^*]$  (from Proposition 5), the bound would lead us to infer  $\mathbb{E}[\phi_i] \geq 2.9$ , when in fact  $\mathbb{E}[\phi_i] = 50$ .

Taking  $\bar{\phi} \rightarrow \infty$  in (27) generates the result in part (iii), which indicates that Corollary 1 does not require any modification in taking unconditional expectations over streams. Similarly, Corollary 2 does not depend on  $\phi$  and therefore is robust to  $\phi$  changing across streams as long as  $\Delta_i \leq 0$ .

Finally, returning to part (i), the analogue to Proposition 2 requires an additional assumption. The original formula includes the product of  $\pi_0^* - \pi_0$  and  $\Delta$ . Therefore, when averaging over DGPs, the covariance between  $\pi_{0,i}$  and  $\Delta_i$  across DGPs affects the bounds (which otherwise follow the logic for part (ii)). For simplicity, the proposition assumes that this covariance is zero (though one could adjust this bound under different assumed values for the covariance). To understand this assumption, recall from Section 2.3 that, fixing  $\pi_0^*$ ,  $\Delta$  is determined by the asymmetry of the likelihood ratios of the upward and downward signals in the DGP. Assuming zero covariance is thus equivalent to assuming no relationship between the asymmetry of the DGP and  $\phi$ . Although we have no reason to presume a particular relationship between the DGP asymmetry and  $\phi$ , our main empirical results apply the more conservative bound in part (ii), which has the benefit of requiring no further assumptions. Part (i) also holds only for a fixed  $\pi_{0,i}^*$ ; this is sufficient for our purposes, as our empirical results for this less-conservative bound will generally be conditional on a given  $\pi_{0,i}^*$ , in particular in Figure 2.

The bounds stated in the proposition are now empirically implementable: in particular, we can measure a sample counterpart of the average excess-movement statistic on the left side and the RN priors on the right side, and the minimum  $\bar{\phi}$  that solves the bound in (27) is then a conservative estimate of the maximal conditional-mean SDF slope for the return-state pair in question.<sup>26</sup> If no such  $\bar{\phi}$  solves the bound, then the bound in part (iii) is violated. Even simpler, the bound in part (iv) only requires comparing  $\mathbb{E}[X_i^*]$  to zero.

The results and discussion above simplified notation by dropping the state-pair subscript  $j$ , but the values  $\bar{\phi}_j$  are likely to vary over  $j$ . The same argument used in Proposition 9 to take expectations

<sup>25</sup>These constant- $\pi_{0,i}^*$  bounds can equivalently be read as bounds for the conditional expectation  $\mathbb{E}[X_i^* | \pi_{0,i}^*]$  given  $\mathbb{E}[\phi_i | \pi_{0,i}^*]$ . We write them in the simpler format presented above (“fixing a given  $\pi_{0,i}^*$ ”) only to limit notational complexity.

<sup>26</sup>Further, if  $\mathbb{E}[\phi_i | \pi_{0,i}^*] = \mathbb{E}[\phi_i]$ , as might be expected to hold approximately, then  $\bar{\phi} = \mathbb{E}[\phi_i]$ .

over  $i$ , though, can also be applied to take expectations over  $j$ , thereby obtaining a single estimate  $\overline{\phi}$  aggregated over both streams *and* return states (for all states meeting CTI) if desired.

## 4.2 Robustness to Measurement Error

The final issue we confront in making our bounds empirically implementable is the possibility of mismeasurement or microstructure noise in RN beliefs. Our bounds provide a minimum value of  $\phi$  required to rationalize the observed variation in RN beliefs; if some of this variation is in fact arising due to noise, then we may overestimate this required  $\phi$ . A simple correction can be applied to our bounds to account for this issue, as shown in the following result. Given that noise arises period-by-period, we first define one-period analogues for our statistics: denote RN movement between  $t$  and  $t + 1$  by  $m_{t,t+1}^* \equiv (\pi_t^* - \pi_{t+1}^*)^2$ , RN uncertainty at  $t$  by  $u_t^* \equiv (1 - \pi_{t+1}^*)\pi_t^*$ , and RN excess movement between  $t$  and  $t + 1$  as  $X_{t,t+1}^* \equiv m_{t,t+1}^* - (u_t^* - u_{t+1}^*)$  (where we now omit  $i$  and  $j$  subscripts for clarity). Similar to AR (Section II.E), we then have the following:

**PROPOSITION 10.** *Assume that the observed  $\widehat{\pi}_t^*$  is measured with error with respect to the true  $\pi_t^*$ :*

$$\widehat{\pi}_t^* = \pi_t^* + \epsilon_t, \quad (28)$$

where  $\mathbb{E}[\epsilon_t] = 0$ ,  $\mathbb{E}[\epsilon_t \epsilon_{t+1}] = 0$ , and  $\mathbb{E}[\epsilon_{t+k} \pi_{t+k'}^*] = 0$  for  $k, k' \in \{0, 1\}$ . Denoting the observed one-period RN excess movement statistic by  $\widehat{X}_{t,t+1}^*$ , its relation to the true value  $X_{t,t+1}^*$  in expectation is

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

We can thus subtract  $2\text{Var}(\epsilon_{t,j})$  from each period's observed excess-movement statistic to obtain an unbiased true excess movement value, which can then be used in our bounds after summing over the full stream. If measurement error is positively correlated over time rather than uncorrelated, this will reduce the upward bias in measured  $X^*$ .<sup>27</sup> We discuss estimation of  $\text{Var}(\epsilon_t)$  in Section 5.

## 4.3 Robustness to Violations of CTI

Proposition 10 allows for a correction with respect to empirical misspecification; we turn now to a result that speaks to the possibility of theoretical misspecification in the form of violations of CTI. While this assumption of a constant  $\phi$  within belief streams is met in some commonly used theoretical frameworks, it is a knife-edge restriction that is unlikely to be met exactly. In this section, we show theoretically that within a large class of environments, our main results continue to hold even with a time-varying  $\phi_t$ . We also consider a set of simulations of other environments and find that realistic violations of CTI appear unlikely to affect our results more than minimally.

The fact that time-varying  $\phi_t$  does not produce significant excess movement might seem

---

<sup>27</sup>One might worry about negatively correlated measurement error in the case of bid-ask bounce, but as we will see in our empirical estimation, the autocorrelation values for estimated microstructure noise have long died out at a one-day lag, and we use end-of-day data to construct our statistics.

unintuitive: one might worry that even small oscillations in  $\phi_t$  could cause unbounded movement in RN beliefs and dramatically invalidate our conclusions. For example, suppose that  $\pi_t$  is constant at .5 and  $\phi_t$  changes back and forth from 1 to 1.5 repeatedly. Without any movement in physical beliefs,  $\pi_0^*$  will vary repeatedly between .5 and .6, leading to unbounded movement as  $T \rightarrow \infty$ . Crucially, however, this argument overlooks a core insight of the paper: given rational expectations,  $\phi_t$  cannot undergo such repeated oscillation if its variance is bounded, because it is itself a function of martingale conditional expectations.

More formally, taking  $\phi_t$  as defined in (24) (again suppressing  $i, j$  subscripts for clarity), write  $\phi_t = \phi_t^{(1)}/\phi_t^{(0)}$ , where  $\phi_t^{(1)} \equiv \mathbb{E}_t[M_{t,T} | R_T^m = \theta_j]$  is the numerator and  $\phi_t^{(0)} \equiv \mathbb{E}_t[M_{t,T} | R_T^m = \theta_{j+1}]$  is the denominator of (24). We now drop [Assumption 3'](#) (CTI) and let  $\phi_t$  vary, but [Assumption 1'](#) implies that  $\phi_t^{(\ell)} = \mathbb{E}_t[\phi_{t+1}^{(\ell)}]$  for  $\ell = 0, 1$ . In spite of this restriction, though, allowing  $\phi_t$  to vary meaningfully complicates the situation, as RN beliefs become a non-linear function of time-varying subjective beliefs *and* conditional SDF expectations (see, e.g., (9)), all of which have unrestricted DGPs. Surprisingly, though, our bounds still hold as stated over a large class of such DGPs:

**PROPOSITION 11.** *Assume no arbitrage and [Assumptions 1'–2'](#) continue to hold, but in place of [Assumption 3'](#), assume only that  $\text{Var}(\phi_t) < \infty$ . If  $\phi_t$  evolves as a martingale or supermartingale ( $\mathbb{E}_t[\phi_{t+1}] \leq \phi_t$ ), then the bounds in [Proposition 3](#) and [Corollary 1](#) (and their counterparts in [Propositions 7](#) and [9](#)) continue to apply, with  $\phi_0$  replacing  $\phi$ .*

The proof of [Proposition 11](#) demonstrates that the variation in  $\pi_0^*$  arising when  $\phi_t$  is a non-degenerate supermartingale always *strictly lowers* the expected excess belief movement period by period, rendering the main bound in (19) even more conservative.<sup>28</sup> Whether  $\phi_t$  is a supermartingale depends on the sign of  $\text{Cov}_t(\phi_{t+1}, \phi_{t+1}^{(0)})$ : from the definition of this covariance, write

$$\mathbb{E}_t[\phi_{t+1}] = \phi_t - \frac{\text{Cov}_t(\phi_{t+1}, \phi_{t+1}^{(0)})}{\phi_{t+1}^{(0)}},$$

so  $\mathbb{E}_t[\phi_{t+1}] \leq \phi_t$  if and only if  $\text{Cov}_t(\phi_{t+1}, \phi_{t+1}^{(0)}) = \text{Cov}_t(\phi_{t+1}, \mathbb{E}_t[M_{t,T} | R_T^m = \theta_{j+1}]) \geq 0$ . Thinking of  $M_{t,T}$  as proportional to  $U'(C_T)$  as in [Section 2](#), this requires that expected terminal risk aversion (encoded in  $\phi_{t+1}$ , as in [Proposition 8](#)) be positively correlated with expected marginal utility (MU) in the high-consumption state. This is an intuitively reasonable restriction, as it implies bad news (in the form of higher expected MU) generally arrives at the same time for both states, with the expected low-consumption MU increasing more than its high-consumption counterpart. The converse ( $\text{Cov}_t(\phi_{t+1}, \phi_{t+1}^{(0)}) < 0$ ), by contrast, requires risk aversion to increase in general in response to *good* news about MU in the good state.<sup>29</sup>

If one wishes to consider such a setting in which  $\phi_t$  does not evolve as a supermartingale, though, it is difficult to make analytical statements given the complexity of such a setting. Instead,

<sup>28</sup>Further, given the continuity of our bounds, if  $\phi_t$  is a submartingale but close to a martingale, then our results also hold up to a small additive error.

<sup>29</sup>For further discussion of such a restriction in a different context, see [Lazarus \(2019\)](#).

we numerically simulate a set of binary DGPs featuring strict submartingale  $\phi_t$  processes alongside learning about the return state. Results for this exercise are reported in Appendix B.5. Even in a parameterization with extremely high  $\phi$  uncertainty — in which  $\phi$  can vary between 1 and 9, corresponding to  $\gamma$  varying between 0 and 160 —  $\mathbb{E}[X^*]$  rises by at most 0.015, and it breaches the bound for  $\phi = 3$  (equal to  $\phi_0$ ) less than one percent of the time in such DGPs.

Finally, to consider how time-varying  $\phi_t$  affects our results in a less abstract environment, we simulate the [Campbell and Cochrane \(1999\)](#) habit-formation model, which violates CTI, using the baseline calibration proposed by [Campbell and Cochrane](#). We find in this case that  $\phi_t$  is closely approximated by a martingale and therefore does not generate additional excess movement. Our bounds thus continue to hold and continue to be conservative. See Appendix B.6 for details.

Given our theoretical and numerical results, we conclude that allowing  $\phi_t$  to change within a stream is unlikely to impact our results and bounds meaningfully. That said, we can imagine that there exists a set of perverse distributions with negatively correlated values of  $\phi_{t+1}^{(1)}$  and  $\phi_{t+1}^{(0)}$  which, combined with near-worst-case DGPs for information about the return state, might further increase excess movement. Our strong suspicion based on the above results, though, is that any realistic DGP will have a small impact.

## 5. Empirical Estimation and Results

Our theory leads to bounds on the variation in risk-neutral beliefs over the value of the market index, which we proceed now to measure in the data. We begin by describing how we map from theory to data ([Section 5.1](#)) and how we estimate microstructure noise ([Section 5.2](#)), and then we summarize the data ([Section 5.3](#)) before turning to our empirical results ([Section 5.4](#)). We conclude with a brief discussion of the reduced-form correlates of RN excess movement ([Section 5.5](#)).

### 5.1 Data and Risk-Neutral Distribution

**Data.** Our main source for S&P 500 index options data is the OptionMetrics database, which provides end-of-day prices for European call and put options for all strike prices and option expiration dates traded on the Chicago Board Options Exchange (CBOE). The sample runs from January 1996 through December 2018. We augment this data with intraday price quotes obtained directly from the CBOE for a subset of trading days in our sample, in order to account for market microstructure noise; this additional data is described further in [Section 5.2](#).

We apply standard filters to remove outliers and options with poor trading liquidity from the OptionMetrics data, with details provided in Appendix B.7.<sup>30</sup> Two aspects of this data cleaning bear mention here. First, while our bounds apply for belief streams of arbitrary length, we follow past literature (e.g., [Christoffersen, Heston, and Jacobs, 2013](#); [Martin, 2017](#)) and consider options with

---

<sup>30</sup>The raw dataset contains 12.4 million option prices; the filtered data, which are then used to measure the RN distribution, contain 4.3 million option prices corresponding to 5,537 trading dates and 991 option expiration dates. The majority of the difference is attributable to our use of only out-of-the-money call and put strikes.

maturity of at most one year. Second, after transforming the filtered prices to RN beliefs (described below), we keep only conditional RN belief observations  $\tilde{\pi}_{t,i,j}^*$  for which the non-conditional beliefs satisfy  $\pi_t^*(R_{T_i}^m = \theta_j) + \pi_t^*(R_{T_i}^m = \theta_{j+1}) \geq 5\%$ , as conditional beliefs  $\tilde{\pi}_{t,i,j}^*$  are likely to be particularly susceptible to mismeasurement when the underlying beliefs are close to zero.

**Empirical return space.** For our baseline estimation, we define the return state space  $\Theta$  in terms of log excess return intervals:

$$\Theta = R_{0_i, T_i}^f \exp\{(-\infty, -0.2], (-0.2, -0.15], (-0.15, -0.1], \dots, (0.1, 0.15], (0.15, 0.2], (0.2, \infty)\},$$

where  $R_{0_i, T_i}^f$  is the gross risk-free rate from  $0_i$  to  $T_i$ .<sup>31</sup> In words, return state 1 is realized if the log excess S&P return from  $0_i$  to  $T_i$  is less than  $-0.2 \approx -20\%$ ; state 2 is realized if the excess return is in the five-percentage-point bin between  $-0.2$  and  $-0.15$ ; and so on. Abusing notation slightly, we typically refer to states by the right end of their associated log excess-return bin:  $\theta_1 = -0.2$ ,  $\theta_2 = -0.15$ ,  $\dots$ ,  $\theta_9 = 0.2$ ,  $\theta_{10} = \infty$ . Following the labeling convention in [Assumption 2'](#), we define the binary conditional beliefs to be used in our tests as  $\tilde{\pi}_{t,i,j}^* \equiv \pi_t^*(R_{T_i}^m = \theta_j | R_{T_i}^m \in \{\theta_j, \theta_{j+1}\})$ , so  $\tilde{\pi}_{t,i,j}^*$  corresponds to the probability that the low state  $j$  (e.g.,  $\theta_2 = -15\%$ ) will be realized, conditional on  $j$  or  $j+1$  being realized (in this case, conditional on an excess return between  $-20\%$  and  $-10\%$ ). We again use of the right end of the return bin for  $j$  as shorthand in referencing the conditional belief and excess movement for state pair  $(\theta_j, \theta_{j+1})$ , corresponding to the midpoint of the two return bins.

Our use of five-percentage-point ranges for return states reflects a desire to balance (i) measurement accuracy for the RN beliefs and (ii) plausibility of our assumption of constant  $\phi_j$  (CTI). Wider bins lead to greater measurement accuracy, but make it less likely that there are no changes in the expected SDF realization conditional on a given return state  $\theta_j$  relative to  $\theta_{j+1}$ . We report empirical estimates below for all adjacent state pairs for completeness, but as discussed in [Remark 2](#) after [Assumption 3'](#), it is unlikely that CTI is met for the extreme state pairs ( $\theta_1$  relative to  $\theta_2$ , and  $\theta_9 = \theta_{J-1}$  vs.  $\theta_{10} = \theta_J$ ). Our focus is thus on the *interior* state pairs with low-return states  $\theta_2, \dots, \theta_8$ ; in particular, when we aggregate our state-by-state estimates of  $\phi_j$  required to rationalize the data into a single average value  $\bar{\phi}$  across states, we use only these interior states.<sup>32</sup>

**Risk-neutral beliefs.** To extract a risk-neutral distribution over the return states in  $\Theta$  from the observed option cross-sections, we use standard tools from the option-pricing literature. Our starting point is equation (22), which tells us how to map from option prices to RN beliefs. We use this to construct a smooth RN distribution for returns, largely following the technique proposed by [Malz \(2014\)](#); Appendix B.7 provides a detailed description. With the RN beliefs  $\pi_{t,i,j}^*$  in hand, we can then calculate conditional beliefs straightforwardly as  $\tilde{\pi}_{t,i,j}^* = \pi_{t,i,j}^* / (\pi_{t,i,j}^* + \pi_{t,i,j+1}^*)$ . We then use the resulting conditional RN belief streams to calculate the excess movement statistics  $X_{i,j}^*$

<sup>31</sup>We use excess returns for convenience of interpretation. Following [van Binsbergen, Diamond, and Grotteria \(2022\)](#), we measure  $R_{0_i, T_i}^f$  directly from the options prices by applying the put-call parity relationship; again see Appendix B.7.

<sup>32</sup>This yields an additional de facto data filter, as we are effectively considering only option strikes with moneyness between 0.8 and 1.2 (following, e.g., [Constantinides, Jackwerth, and Savov, 2013](#)).

needed to implement our bounds. Our general results in [Section 3](#) restrict the expectation of  $X_{i,j}^*$  conditional on state  $\theta_j$  or  $\theta_{j+1}$  being realized — in particular, [Propositions 7](#) and [9](#) give bounds for  $\tilde{\mathbb{E}}[X_{i,j}^*]$  — and we accordingly keep only observations for which  $\tilde{\pi}_{T,i,j}^* = 0$  or 1 ex post; for example, if the total excess return on the market over the life of option contract  $i$  is -14%, then we keep only  $X_{i,2}^*$  ( $\theta_2$  ranges from -20% to -15% return, so  $\tilde{\pi}_{T,i,2}^* = 0$ ) and  $X_{i,3}^*$  ( $\tilde{\pi}_{T,i,3}^* = 1$ ).

**Simplifying notation.** To this point, we have taken care in this section to use detailed notation to clarify that our conditional statistics  $(\tilde{\pi}_{t,i,j}^*, X_{i,j}^*)$  depend on the contract  $i$  and state pair  $j$ . Having clarified this, though, in what follows we generally suppress the cumbersome use of  $i, j$ , and  $\tilde{\cdot}$ , and simply write  $\pi_t^*$  in place of  $\tilde{\pi}_{t,i,j}^*$ ,  $X^*$  for  $X_{i,j}^*$ , and so on (as, e.g., in [Section 4.2](#)). Similarly, we often drop the “conditional” qualifier when referring to conditional RN belief statistics.

## 5.2 Noise Estimation

As in [Proposition 10](#), we also wish to account for measurement error stemming from possible microstructure noise in our estimation.<sup>33</sup> That result shows that unlike in the classical errors-in-variables regression model (which leads to attenuation), measurement error in our case can increase the observed variation  $X^*$  and thereby lead to an upward bias in the estimated SDF slope needed to rationalize the data. With noise described by  $\hat{\pi}_t^* = \pi_t^* + \epsilon_t$  as in [\(28\)](#), [Proposition 10](#) tells us that we must estimate  $\text{Var}(\epsilon_t)$  in order to eliminate this bias. We turn to a sample of high-frequency option prices to estimate this noise variance in our RN beliefs data.

Specifically, we obtain minute-by-minute price quotes on S&P index options for a subset of trading days directly from the CBOE. For each available option expiration date on each such trading day, we recalculate the RN belief distribution at the end of each minute using exactly the same procedure as described in [Section 5.1](#). As this requires calculating 390 sets of RN beliefs for each trading day (9:30 AM–4:00 PM), this procedure would be computationally infeasible if applied to our entire sample of 5,537 trading days (each of which has an average of 11 available option expiration dates, generating 60,543  $(t, T)$  combinations). We accordingly select 30 trading days at random from within our available sample period, and use the minute-by-minute quotes to calculate intraday RN distributions for these days.<sup>34</sup>

We then use tools from the literature on microstructure noise to estimate  $\text{Var}(\epsilon_t)$  using these intraday data. The basic intuition for this estimation strategy — as described, for example, by [Zhang, Mykland, and Ait-Sahalia \(2005\)](#) — is best understood by assuming temporarily that the noise  $\epsilon_t$  in [\(28\)](#) is i.i.d., while the true  $\pi_t^*$  changes smoothly over time (i.e.,  $\eta_{t+h} \equiv \pi_{t+h}^* - \pi_t^* \rightarrow 0$  and  $\text{Var}(\eta_{t+h}) \rightarrow 0$  as  $h \rightarrow 0$ , as would be the case for an Itô process). Imagine calculating movement using the observed beliefs,  $(\hat{\pi}_{t+h}^* - \hat{\pi}_t^*)^2$ , with less and less time  $h$  between consecutive observations. As one decreases  $h$  to 0, the noise swamps the true variation: since  $(\pi_{t+h}^* - \pi_t^*)^2 \rightarrow 0$ ,

<sup>33</sup>For example, transient demand pressure in the option market ([Bollen and Whaley, 2004](#)) may cause variation in observed RN beliefs unrelated to the underlying index dynamics.

<sup>34</sup>This yields an intraday data set roughly twice as large as the original one, as  $30 \times 390 \times 11 \approx 130,000$ .

we have  $\mathbb{E}[(\hat{\pi}_{t+h}^* - \hat{\pi}_t^*)^2] \rightarrow 2\text{Var}(\epsilon_t)$ . Thus in this simple example,  $\text{Var}(\epsilon_t)$  can be estimated by calculating the quadratic variation in RN beliefs sampled at a high frequency.

In practice, one would expect the data to contain both non-i.i.d. noise  $\epsilon_t$  and jumps in the true process  $\pi_t^*$ , and it is desirable to use a noise estimation method that is robust to these features. One such estimator for  $\text{Var}(\epsilon_t)$  is the ReMeDI (“Realized moMents of Disjoint Increments”) estimator proposed by [Li and Linton \(2021\)](#). This estimator effectively takes the average product of *disjoint increments* of the observed process,  $(\hat{\pi}_t^* - \hat{\pi}_{t-h}^*)(\hat{\pi}_t^* - \hat{\pi}_{t+h}^*)$ .<sup>35</sup> The idea is that even if the true process features jumps so that  $\mathbb{E}[(\pi_{t+h}^* - \pi_t^*)^2] > 0$ , its increments over non-overlapping windows are still approximately uncorrelated. [Li and Linton \(2021\)](#) show that this estimator is consistent for  $\text{Var}(\epsilon_t)$  under general semimartingale processes for  $\pi_t^*$  (i.e.,  $\pi_t^*$  can feature jumps but must be of bounded variation), and for quite general dependent noise processes as long as the autocovariances for  $\epsilon_t$  decay to 0 sufficiently quickly relative to the increment width (see their Theorem 4.1 for precise conditions).<sup>36</sup> It also performs well in simulations and empirical applications.

Using this ReMeDI estimator on our minute-by-minute data, we estimate  $\text{Var}(\epsilon_t) = \text{Var}(\epsilon_{t,i,j})$  separately for each combination of trading day  $t$ , expiration date  $T_i$ , and return state pair  $j$  in our intraday sample.<sup>37</sup> We then match the noise estimates (which are obtained for a subsample of days) to the observed excess movement observations in our original data; see Appendix B.8 for details on this procedure. Finally, we subtract  $2\widehat{\text{Var}}(\epsilon_{t,i,j})$  from  $\widehat{X}_{t,t+1,i,j}^*$  to obtain a *noise-adjusted* estimate of one-day excess movement following [Proposition 10](#), and we sum these noise-adjusted one-day values over the full stream to obtain noise-adjusted estimates of  $X_{i,j}^*$ .

We discuss the magnitude of the noise estimates in the next subsection alongside descriptive statistics for the excess movement values. The ReMeDI procedure also allows for estimation of the intraday autocovariances of the noise  $\epsilon_t$ . These autocovariances are estimated to be positive for small lag values, but they die out quickly and are precisely estimated near zero for noise observations more than an hour apart. This justifies the assumption in [Proposition 10](#) that end-of-day noise observations are uncorrelated,  $\mathbb{E}[\epsilon_t \epsilon_{t+1}] = 0$ , as ultimately we care about noise only to the extent that it affects our excess movement statistics at a daily frequency.

Our main results in [Section 5.4](#) use the noise-adjusted excess movement data. All standard errors and confidence intervals are based on a bootstrap procedure (detailed in [Section 5.4](#)) that accounts for the sampling uncertainty in the above noise estimation and averaging procedure.

<sup>35</sup>More formally and specifically, our ReMeDI estimator (following the replication code provided kindly by [Li and Linton](#)) is  $\widehat{\text{Var}}(\epsilon_t) = \frac{1}{N_{\epsilon,n}} \sum_{i=2k_n}^{N_{\epsilon,n}-k_n} (\hat{\pi}_{t_i}^* - \hat{\pi}_{t_i-2k_n}^*)(\hat{\pi}_{t_i}^* - \hat{\pi}_{t_i+k_n}^*)$ , where  $N_{\epsilon,n}$  is the number of observations over a fixed time span (in our case, one trading day) and  $k_n$  is a tuning parameter, with  $k_n, N_{\epsilon,n} \rightarrow \infty$  and  $k_n/N_{\epsilon,n} \rightarrow 0$  as  $n \rightarrow \infty$ . We select  $k_n$  for each return state using the algorithm in Section F.1 of the Online Appendix of [Li and Linton \(2021\)](#).

<sup>36</sup>While jumps induce upward bias in a naïve estimator based on  $(\hat{\pi}_{t+h}^* - \hat{\pi}_t^*)^2$ , positively autocorrelated noise goes in the other direction. To address this, the ReMeDI estimator uses a window width ( $k_n$  in [footnote 35](#)) that is sufficiently wide even as one decreases the time  $h$  between observations. The second (downward) bias in the naïve estimator seems predominant in our data: the ReMeDI noise estimates are about 50% higher than the naïve estimates.

<sup>37</sup>For this exercise, to increase our available observations, we do not condition on the ex post state being  $\theta_j$  or  $\theta_{j+1}$ .



### 5.3 Excess Movement: Descriptive Statistics and Figures

We first provide summary statistics and plots describing the RN beliefs and excess movement values observed in our data. [Table 2](#) summarizes the average excess movement values  $\bar{X}^*$  overall (across all interior state pairs and expiration dates) and by subsample. Excess movement is difficult to interpret without some normalization; we should, for example, expect movement to increase alongside initial uncertainty under Bayesian updating. The first two columns thus divide  $\bar{X}^*$  by average initial uncertainty  $\bar{u}_0$ . As in [Augenblick and Rabin \(2021, Section II.D\)](#), the resulting normalized statistic can be interpreted as the percent by which average movement exceeds initial uncertainty and thus uncertainty resolution. (For example, a value of 1 corresponds to 100% more movement than uncertainty resolution.) These values are quite high in our data: for the noise-adjusted statistics, there is on average 123% more movement than initial uncertainty. The splits by bin show that these values decrease for return states in the middle of the distribution. The beginning of the sample features high but noisy  $X^*$  statistics, but these averages remain high until the most recent subsample. Finally, higher priors  $\pi_0^*$  correspond with greater excess movement, which accords with our theoretical bounds.

The next two columns of [Table 2](#) instead normalize  $\bar{X}^*$  by the average contract length  $\bar{T}$ , so the resulting statistics can be interpreted as excess movement per day. For a rough understanding of the actual variation in RN beliefs corresponding to these values, consider a pair  $(t, t + 1)$  for which  $\pi_{t+1}^* = 1 - \pi_t^*$ , so that there is no uncertainty resolution from  $t$  to  $t + 1$ . One-day excess movement and movement therefore coincide, so that excess movement is equal to the squared change in beliefs; for example, the noise-adjusted average of 0.0038 corresponds to a raw change of  $\sqrt{0.0038} \approx 0.06$  (continuing the example above,  $\pi_t^* = 0.47$ , and then  $\pi_{t+1}^* = 0.53$ ). Under this normalization, there is now no clear pattern for average excess movement across bins, as the more-extreme return states also tend to have longer contract lengths.<sup>38</sup> This pattern is in fact quite consistent in our data: longer contract lengths tend to coincide with more excess movement, as RN beliefs bounce up and down over the length of a contract. This general pattern is inconsistent with the null of Bayesian updating, which tells us that excess movement in subjective beliefs should not depend at all on the horizon at which uncertainty is resolved. Meanwhile, for the splits by date and by prior, the basic patterns discussed above are still present under this normalization by  $\bar{T}$ .

Comparing the raw and noise-adjusted values makes clear that despite the substantial excess movement in the noise-adjusted statistics, noise does represent a meaningful portion (about 1/3) of the raw  $X^*$  data. The raw and noise-adjusted mean for  $\bar{X}^*/\bar{T}$  differ by  $0.0059 - 0.0038 \approx 0.002$ , so  $\widehat{\text{Var}}(\epsilon_t) \approx 0.002/2 = 0.001$ . The standard deviation of  $\epsilon_t$  is thus roughly 0.03 per day. For the splits by state, noise variances tend to be lowest for return states near the center of the distribution, as is intuitive.

Next, [Figure 2](#) provides a visual summary of the  $X^*$  statistics relative to the bounds. The blue curves describe the raw and noise-adjusted local-average  $X^*$  statistics as one varies the RN prior  $\pi_0^*$ ,

<sup>38</sup>Recall that an observation for a given state pair is only included conditional on the realized return being in one of the two states. Longer contracts are likelier to generate greater absolute returns, explaining this positive covariance.

calculated using all interior state pairs and streams; these curves are the same in both panels. As one would expect, there is very little excess movement for RN priors near 0 or 1, but excess movement is positive for intermediate  $\pi_0^*$  values for which there is greater initial uncertainty. We compare these values to the theoretical bounds under different levels of  $\phi$  (or more precisely,  $\mathbb{E}[\phi_i]$ ) for each value of  $\pi_0^*$ , which are shown in gray. Panel (a) uses the tighter bound from [Proposition 9\(i\)](#), which we implement by estimating  $\Delta$  using local averages for the conditional expectations  $\mathbb{E}[X^* | \theta, \pi_0^*]$  in (17) over  $\pi_0^*$ .<sup>39</sup> Panel (b) uses the conservative bound from the second inequality in [Proposition 9\(ii\)](#), so the gray lines in this panel align with the solid bound lines in [Figure 1](#).

Across the two panels, it is evident that the values of  $X^*$  observed in the data exceed both sets of bounds, other than for high RN priors  $\pi_0^*$  and for large values of  $\phi$ . For panel (a), we estimate  $\hat{\Delta} = 0$  for  $\pi_0^* \approx 0.5$ ,  $\hat{\Delta} < 0$  for  $\pi_0^*$  below this cutoff, and  $\hat{\Delta} > 0$  above it, as is evident from the bounds crossing zero at  $\pi_0^* \approx 0.5$ . As discussed in [Section 2.3](#) (see [pages 15–16](#)), this indicates that the DGP is close to symmetric, with equally informative signals for the two states  $(\theta_j, \theta_{j+1})$  on average (and thus roughly equal-sized upward and downward movements of  $\pi_i^*$ ). As the bounds in [Proposition 9](#) apply for each possible  $\pi_0^*$ , the positive point estimates for  $\mathbb{E}[X^* | \pi_0^*]$  clearly violate the bounds for  $\pi_0^* < 0.5$ , which are (at most) 0 for all  $\phi$ . And while the empirical curves are closer to the more-conservative bounds in panel (b), the noise-adjusted estimates still exceed  $\pi_0^{*2}$  (the bound for  $\phi = \infty$ ) for  $\pi_0^*$  less than about 2/3.

These figures do not, however, integrate over  $\pi_0^*$  to provide a single statistic summarizing excess movement relative to the bounds over all streams. More importantly, they do not include any measures of statistical uncertainty necessary to make inferential statements or conduct hypothesis tests. To address these two issues, we move on to our main estimation and results.

## 5.4 Main Results

We turn now to the empirical implementation of our theoretical bounds. Given our sample of noise-adjusted excess movement statistics and corresponding RN priors, each possible value of  $\bar{\phi}$  leads to a *residual* excess movement value  $e_i(\bar{\phi}) = X_i^* - \text{bound}(\pi_{0,i}^*, \bar{\phi})$  for contract  $i$  (where we continue to suppress dependence on state pair  $j$ ). We calculate two versions of this residual corresponding to the bound in part (i) and the unconditional bound in part (ii) of [Proposition 9](#) ([equations \(26\)](#) and [\(27\)](#), respectively):

$$\begin{aligned} e_i^\Delta(\bar{\phi}) &= X_i^* - \max\left\{0, \left(\pi_{0,i}^* - \frac{\pi_{0,i}^*}{\bar{\phi} + (1-\bar{\phi})\pi_{0,i}^*}\right) \hat{\Delta}_i\right\}, \\ e_i^{\text{main}}(\bar{\phi}) &= X_i^* - \left(\pi_{0,i}^* - \frac{\pi_{0,i}^*}{\bar{\phi} + (1-\bar{\phi})\pi_{0,i}^*}\right) \pi_{0,i}^*. \end{aligned} \tag{29}$$

<sup>39</sup>We slightly modify the formula in [Proposition 9\(i\)](#) for the figure: we use  $\mathbb{E}[X_i^*] \leq (\pi_{0,i}^* - \frac{\pi_{0,i}^*}{\mathbb{E}[\phi_i] + (1-\mathbb{E}[\phi_i])\pi_{0,i}^*})\mathbb{E}[\Delta_i]$ , rather than cutting the bounds off at 0, to clarify that  $\hat{\Delta} < 0$  for  $\pi_0^* < 0.5$ .

The first version corresponds to the tighter bound from [Proposition 2](#), which requires a smoothed estimate of  $\Delta_i$  as calculated for [Figure 2](#). This version accounts for the estimated DGP (through  $\widehat{\Delta}_i$ ) and thus conveys some useful preliminary information, but strictly speaking, part (i) of the proposition applies only conditional on a given  $\pi_{0,i}^*$  and only under the unverifiable condition that  $\text{Cov}(\pi_{0,i}, \Delta_i) = 0$ .<sup>40</sup> Thus the second version, which implements the more conservative bound from [Proposition 3](#) and which applies unconditionally, is the basis for our main set of results. Both residuals can be directly calculated using our noise-adjusted data for each possible value of  $\bar{\phi}$ .

We first present sample averages of these residual statistics for a range of values of  $\bar{\phi}$ , both for each individual return state pair  $j$  and aggregated across all interior state pairs. As these values have no natural scaling, we present them in the form of  $t$ -statistics,  $t_{\bar{e}} = \overline{e_i(\bar{\phi})} / \widehat{SE}_{\bar{e}}$ , where  $\overline{e_i(\bar{\phi})}$  is the sample average of residual excess movement  $e_i(\bar{\phi})$  and  $\widehat{SE}_{\bar{e}}$  is its standard error. We calculate these standard errors using a block bootstrap with a block length of one month, with each block containing (i) raw excess movement statistics and priors for all contracts expiring in a given month, and (ii) noise variance estimates for any trading days in our intraday sample that fall in the same month. For each resampled data set, we use the set of  $(X_i^*, \tilde{\pi}_{0,i}^*, \{\widehat{\text{Var}}(\epsilon_{t,i})\})$  values to recalculate noise-adjusted excess movement and residual values  $e_i(\bar{\phi})$ .<sup>41</sup> The bootstrap accordingly accounts for sampling uncertainty in all statistics used to calculate noise-adjusted excess movement and  $e_i(\bar{\phi})$ . We conduct 10,000 such draws, from which we calculate standard errors as the standard deviation of  $\overline{e_i(\bar{\phi})}$  across bootstrap draws.

These residual  $t$ -statistics are presented in [Table 3](#) for the two versions of the residual in (29), both overall (for all interior state pairs) and by individual return state pair. Since  $\mathbb{E}[e_i(\bar{\phi})] = 0$  under RE given a correctly specified value  $\bar{\phi}$ , these  $t$ -statistics tell us how far the residuals are from being consistent with any hypothesized null value for the SDF slope. Positive numbers correspond to the data exhibiting too much excess movement to be consistent with a given value of  $\bar{\phi}$ , and vice versa for negative numbers. In panel (a), the only negative  $t$ -statistics for the tighter bound are for RN beliefs over  $(\theta_j, \theta_{j+1}) = ([0\%, 5\%], [5\%, 10\%])$  (i.e.,  $j = 6$ ) for  $\bar{\phi} > 10$ ; all other  $t$ -statistics, including the overall values, are positive (and generally large in magnitude), indicating no value of  $\bar{\phi}$  is consistent with the degree of observed excess RN movement. In panel (b), the average conservative bound residuals are smaller in magnitude but also generally positive, other than for large  $\bar{\phi}$  (with magnitudes discussed below) and for return-state pairs in the middle of the distribution.

As admissible excess movement  $\mathbb{E}[X_i^*]$  is monotonically increasing in the unobserved parameter  $\bar{\phi}$ , our main empirical exercise is to estimate the lower bound for this SDF slope such that the bound for  $\mathbb{E}[X_i^*]$  is satisfied. This lower bound for  $\bar{\phi}$  is estimated as the minimal value for which the average residual value  $\overline{e_i(\bar{\phi})}$  is zero, so that we are effectively finding the root of the function traced out in [Table 3](#). Given that the tighter bound is generally not satisfied even for  $\bar{\phi} = \infty$  (and since it holds only under restrictive assumptions), we now confine attention to the conservative

<sup>40</sup>It is also stated for  $\mathbb{E}[\phi_i]$  rather than  $\bar{\phi}$ , but the bound still applies using the more conservative value  $\bar{\phi}$  as well.

<sup>41</sup>For the residual  $e_i^{\Delta}(\bar{\phi})$  corresponding to the tighter bound, we also re-estimate  $\widehat{\Delta}_i$  in each bootstrap draw by calculating the same local average (with respect to  $\pi_{0,i}^*$ ) as in [Figure 2\(a\)](#), and then evaluating it at the observed  $\pi_{0,i}^*$ .

bound from [Proposition 9\(ii\)](#). The estimated  $\bar{\phi}$  can be interpreted as the minimal SDF slope for which the amount of observed excess movement in RN beliefs can be rationalized; it is accordingly an index of the restrictiveness of the joint assumptions of RE and CTI (or, given [Proposition 11](#), RE and supermartingale  $\phi_t$ ). As above, we estimate this SDF slope both for each individual state pair ( $\bar{\phi} = \bar{\phi}_j$ ) and over all interior state pairs ( $\bar{\phi} = \bar{\phi}$ ). To make the estimates for  $\bar{\phi}$  more readily interpretable, we also use [Proposition 8](#) to translate them into local relative risk aversion values  $\bar{\gamma}$  for a fictitious representative agent who consumes the market.

[Table 4](#) presents our main results. Whenever there is *no* value of  $\bar{\phi}$  for which the bound for  $\mathbb{E}[X_i^*]$  is satisfied — i.e., from [Proposition 9\(iii\)](#), when we estimate  $\mathbb{E}[X_i^*] > \mathbb{E}[\pi_{0,i}^*{}^2]$  — we write  $\bar{\phi} = \infty$ . In brackets below each point estimate, we provide the lower bound of a one-sided 95% confidence interval (CI) for the parameter in question.<sup>42</sup> Starting from the same bootstrap resampling procedure as described just above, we obtain these CIs by inverting a one-sided test: the CI lower bound  $\hat{\phi}_{LB}$  is the minimal  $\bar{\phi}$  such that  $e_i^{\text{main}}(\bar{\phi}) = 0$  is not rejected at the 5% level using the bootstrapped data (see [Appendix B.9](#) for further details). The first column of the table provides estimates for the minimal SDF slopes  $\bar{\phi}$ . The second column translates these figures into relative risk aversion values using  $\bar{\gamma} = \frac{(\bar{\phi}-1)}{0.05} = (\bar{\phi} - 1) \times 20$  from [\(25\)](#), since adjacent return states have midpoints 5 percent (more precisely, 5 log points) apart by construction.

Starting with the overall estimates in the first row: the point estimate for the conservative lower bound for  $\bar{\phi}$  is slightly greater than 50, in line with the *t*-statistic in [Table 3\(b\)](#) for  $\bar{\phi} = 50$  (which is slightly positive but close to zero). This translates to an extraordinarily high estimated lower bound for  $\bar{\gamma}$  of 1,075. Values of  $\bar{\phi}$  below 9.8 (for  $\bar{\gamma}$ , below 175) are rejected at the 5% level. We conclude that extremely high risk aversion is needed to rationalize the large degree of excess movement in RN beliefs observed in the data.

Moving to the local estimates for the individual return-state pairs, all but two of the point estimates of the lower bounds for  $\bar{\phi}_j$  and  $\bar{\gamma}_j$  are infinite, indicating that no amount of utility curvature (or SDF slope) can rationalize the observed excess movement such that the bounds are satisfied. For many of the states ( $j = 2, 3, 4$ , and  $8$ ), their associated confidence intervals also have lower bounds of  $\infty$  (or more precisely, are empty), indicating outright model rejection. Only for RN beliefs over the state pairs in the middle of the distribution — in particular,  $j = 5$  and  $6$ , corresponding to excess-return midpoints of 0% and 5%, respectively — have finite point estimates and confidence intervals that contain reasonable risk-aversion values of about 20. RN beliefs over these intermediate return states are thus comparatively well-behaved; for all other states, there is so much mean-reverting variation in RN beliefs that the bounds are only met for implausibly large values of  $\bar{\phi}_j$ , if at all.

In light of [Proposition 6](#), we conclude that belief revisions are excessively volatile in all cases for which the data cannot be rationalized with finite risk aversion, as these findings cannot in general be produced solely by miscalibrated priors. Further, the large local risk-aversion bounds at every point of the return distribution and the extremely large overall estimates (which use only beliefs

---

<sup>42</sup>The set identification implied by our theoretical bound motivates our use of these one-sided intervals.

over the excess-return states between -20% and +20%) imply that no feature of the true underlying data-generating process (e.g., volatility in the left tail of the return distribution) can by itself be responsible for these findings of excess belief volatility.

## 5.5 Suggestive Evidence on the Correlates of RN Excess Movement

We conclude our empirical analysis by briefly considering reduced-form evidence on the macroeconomic and financial correlates of RN excess movement. Online Appendix Table A.1 shows the results of a set of time-series regressions to this end. The dependent variable is the quarterly average of noise-adjusted RN excess movement  $X_{t,t+1,i,j}^*$ . We consider as regressors: (1) proxies for option liquidity and trading activity; (2) measures of volatility and macroeconomic uncertainty; and (3) statistics related to index returns and valuations. Across all specifications, the liquidity- and noise-related variables (bid-ask spreads and volume) have coefficients that are both economically and statistically small, which gives 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, as is intuitive (and further evidence that our measure is in fact reflective of excess volatility). Lagged S&P 500 returns and valuation ratios are also positively related to excess movement. 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 much of the quarterly variation in excess movement. That said, this regression evidence is only suggestive and reduced-form.

## 6. Conclusion

We consider a general theoretical framework in which we show that the assumption of rational expectations imposes testable restrictions on the time variation in risk-neutral beliefs as expressed in asset values. Unlike in much of the previous literature, these results do not require any restrictive assumptions on the data-generating process, and they allow for time variation in discount rates. Further, by using asset values, we do not require direct measures of physical beliefs over future outcomes, and our bounds exploit intertemporal consistency requirements of rational beliefs without the need for the econometrician to know what agents' beliefs "should" be under RE.

When taken to the data using observed asset prices, our bounds provide evidence on the rationality of the market as a whole. Using risk-neutral beliefs over the future return on the S&P 500 index measured from options data, we find that under our assumption of conditional transition independence, extremely high risk aversion is needed to rationalize the variation in these beliefs. We conclude that the RE assumption appears to be quite restrictive. In many cases, no amount of risk aversion (or SDF slope) is capable of rationalizing the behavior of RN beliefs.

While our bounds are informative in ruling out classes of models (or modeling assumptions), this paper only provides basic suggestive evidence as to the potential causes of our bound violations. We believe, though, that there are numerous feasible ways to make progress on this

question. Conducting additional tests on the empirical correlates of excess movement, as well as generalizing our analysis to alternative asset classes, may provide useful additional information. Further, detailed data on changes in individual portfolios could allow for tests on the rationality of individual beliefs, which would help distinguish between micro and macro explanations for the observed excess movement in RN beliefs. We leave these possibilities to future work.

## References

- ALVAREZ, F. AND U. J. JERMANN (2005): "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73, 1977–2016.
- AUGENBLICK, N. AND M. RABIN (2021): "Belief Movement, Uncertainty Reduction, and Rational Updating," *Quarterly Journal of Economics*, 136, 933–985.
- BAKSHI, G. AND F. CHABI-YO (2012): "Variance Bounds on the Permanent and Transitory Components of Stochastic Discount Factors," *Journal of Financial Economics*, 105, 191–208.
- BARNDORFF-NIELSEN, O. E. AND N. SHEPHARD (2001): "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society, Series B*, 63, 167–241.
- BARRO, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121, 823–866.
- VAN BINSBERGEN, J. H., M. BRANDT, AND R. KOIJEN (2012): "On the Timing and Pricing of Dividends," *American Economic Review*, 102, 1596–1618.
- VAN BINSBERGEN, J. H., W. F. DIAMOND, AND M. GROTTERRA (2022): "Risk-Free Interest Rates," *Journal of Financial Economics*, 143, 1–29.
- BLUME, L., T. COURY, AND D. EASLEY (2006): "Information, Trade and Incomplete Markets," *Economic Theory*, 29, 379–394.
- BOLLEN, N. P. AND R. E. WHALEY (2004): "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?" *Journal of Finance*, 59, 711–753.
- BORDALO, P., N. GENNAIOLI, Y. MA, AND A. SHLEIFER (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–2782.
- BOROVIČKA, J., L. P. HANSEN, AND J. A. SCHEINKMAN (2016): "Misspecified Recovery," *Journal of Finance*, 71, 2493–2544.
- BREEDEN, D. T. AND R. H. LITZENBERGER (1978): "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, 621–651.
- 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.
- CAMPBELL, J. Y. AND R. J. SHILLER (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062–1088.
- CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2013): "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel," *Review of Financial Studies*, 26, 1962–2006.
- COCHRANE, J. H. (2011): "Presidential Address: Discount Rates," *Journal of Finance*, 66, 1047–1108.
- COCHRANE, J. H. AND J. SAÁ-REQUEJO (2000): "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets," *Journal of Political Economy*, 108, 79–119.
- CONSTANTINIDES, G. M., J. C. JACKWERTH, AND A. SAVOV (2013): "The Puzzle of Index Option Returns," *Review of Asset Pricing Studies*, 3, 229–257.
- D'ARIENZO, D. (2020): "Maturity Increasing Overreaction and Bond Market Puzzles," *Working Paper*.

- DE BONDT, W. F. M. AND R. THALER (1985): "Does the Stock Market Overreact?" *Journal of Finance*, 40, 793–805.
- DE LA O, R. AND S. MYERS (2021): "Subjective Cash Flow and Discount Rate Expectations," *Journal of Finance*, 76, 1339–1387.
- FAMA, E. F. (1991): "Efficient Capital Markets: II," *Journal of Finance*, 46, 1575–1617.
- GABAIX, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *Quarterly Journal of Economics*, 127, 645–700.
- GARCIA, R., E. GHYSELS, AND É. RENAULT (2010): "The Econometrics of Option Pricing," in *Handbook of Financial Econometrics*, ed. by Y. Aït-Sahalia and L. P. Hansen, Amsterdam: Elsevier, vol. 1, chap. 9, 479–552.
- GIGLIO, S. AND B. KELLY (2018): "Excess Volatility: Beyond Discount Rates," *Quarterly Journal of Economics*, 133, 71–127.
- GREENWOOD, R. AND A. SHLEIFER (2014): "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 27, 714–746.
- HANSEN, L. P. AND R. JAGANNATHAN (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.
- JAMSHIDIAN, F. (1989): "An Exact Bond Option Formula," *Journal of Finance*, 44, 205–209.
- KLEIDON, A. W. (1986): "Variance Bounds Tests and Stock Price Valuation Models," *Journal of Political Economy*, 94, 953–1001.
- LAZARUS, E. (2019): "Horizon-Dependent Risk Pricing: Evidence from Short-Dated Options," *Working Paper*.
- LEROY, S. F. AND R. D. PORTER (1981): "The Present-Value Relation: Tests Based on Implied Variance Bounds," *Econometrica*, 49, 555–574.
- LI, Z. M. AND O. B. LINTON (2021): "A ReMeDI for Microstructure Noise," *Forthcoming, Econometrica*.
- MALZ, A. M. (2014): "A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions," *Federal Reserve Bank of New York Staff Report No. 677*.
- MARSH, T. A. AND R. C. MERTON (1986): "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," *American Economic Review*, 76, 483–498.
- MARTIN, I. (2017): "What Is the Expected Return on the Market?" *Quarterly Journal of Economics*, 132, 367–433.
- MARTIN, I. AND D. PAPADIMITRIOU (2021): "Sentiment and Speculation in a Market with Heterogeneous Beliefs," *Working Paper*.
- MEHRA, R. AND E. C. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.
- MILGROM, P. AND N. STOKEY (1982): "Information, Trade, and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.
- QIN, L. AND V. LINETSKY (2017): "Long-Term Risk: A Martingale Approach," *Econometrica*, 85, 299–312.
- RADNER, R. (1979): "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices," *Econometrica*, 47, 655.
- ROSS, S. (2015): "The Recovery Theorem," *Journal of Finance*, 70, 615–648.
- SHILLER, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, 71, 421–436.
- STEIN, J. (1989): "Overreactions in the Options Market," *Journal of Finance*, 44, 1011–1023.
- ZHANG, L., P. A. MYKLAND, AND Y. AÏT-SAHALIA (2005): "A Tale of Two Time Scales," *Journal of the American Statistical Association*, 100, 1394–1411.

## Tables and Figures

**Table 1: Example DGPs, Beliefs, Movement and Uncertainty Resolution Statistics**

Signals [ $s_1, s_2, \dots$ ]	Stream Likelihood $\mathbb{P}([s_1, \dots])$	Physical Beliefs [ $\pi_0, \pi_1, \dots$ ]	Movement & Init. Uncertainty m $u_0$		RN Beliefs ( $\phi = 3$ ) [ $\pi_0^*, \pi_1^*, \dots$ ]	RN Movement & Init. Uncertainty m* $u_0^*$	
<i>(i) Symmetric signals with <math>L_l \equiv \frac{\mathbb{P}[s_t=l \theta=0]}{\mathbb{P}[s_t=l \theta=1]} = 3</math> and <math>L_h \equiv \frac{\mathbb{P}[s_t=h \theta=1]}{\mathbb{P}[s_t=h \theta=0]} = 3</math>, and then resolution</i>							
[ $l, l$ ]	.5625	[.25, .1, 0]	.0325	.1875	[.5, .25, 0]	.125	.25
[ $l, H$ ]	.0625	[.25, .1, 1]	.8325	.1875	[.5, .25, 1]	.625	.25
[ $h, L$ ]	.1875	[.25, .5, 0]	.3125	.1875	[.5, .75, 0]	.625	.25
[ $h, H$ ]	.1875	[.25, .5, 1]	.3125	.1875	[.5, .75, 1]	.125	.25
$\mathbb{E}[m] = .1875 = u_0$ $\implies \mathbb{E}[X] = 0$				$\mathbb{E}[m^*] = .25 = u_0^*$ $\implies \mathbb{E}[X^*] = 0,$ $\Delta = 0$			
<i>(ii) Asymmetric signals with <math>L_l = 3</math> and <math>L_h \rightarrow \infty</math>, and then resolution</i>							
[ $l, l, L$ ]	.75	[.25, .1, .04, 0]	.0279	.1875	[.5, .25, .1, 0]	.095	.25
[ $H$ ]	.1667	[.25, 1, 1, 1]	.5625	.1875	[.5, 1, 1, 1]	.250	.25
[ $l, H$ ]	.0567	[.25, .1, 1, 1]	.8325	.1875	[.5, .25, 1, 1]	.625	.25
[ $l, l, H$ ]	.0278	[.25, .1, .04, 1]	.9565	.1875	[.5, .25, .1, 1]	.895	.25
$\mathbb{E}[m] = .1875 = u_0$ $\implies \mathbb{E}[X] = 0$				$\mathbb{E}[m^*] = .1725, u_0^* = .25$ $\implies \mathbb{E}[X^*] = -.0775,$ $\Delta = -.31$			
<i>(iii) Asymmetric signals with <math>L_l \rightarrow \infty</math> and <math>L_h = 3</math>, and then resolution</i>							
[ $h, h, H$ ]	.25	[.25, .5, .75, 1]	.1875	.1875	[.5, .75, .9, 1]	.095	.25
[ $L$ ]	.5	[.25, 0, 0, 0]	.0625	.1875	[.5, .0, 0, 0]	.250	.25
[ $h, L$ ]	.1667	[.25, .5, 0, 0]	.3125	.1875	[.5, .75, 0, 0]	.625	.25
[ $h, h, L$ ]	.0833	[.25, .5, .75, 1]	.6875	.1875	[.5, .75, .9, 0]	.895	.25
$\mathbb{E}[m] = .1875 = u_0$ $\implies \mathbb{E}[X] = 0$				$\mathbb{E}[m^*] = .3275, u_0^* = .25$ $\implies \mathbb{E}[X^*] = .0775,$ $\Delta = .31$			

*Note:* This table provides theoretical examples for physical and risk-neutral belief statistics under rational expectations for three possible DGPs, as described in [Section 2.1](#).



**Table 2: Descriptive Statistics for Excess Movement**

	$\bar{X}^*/\bar{u}_0$		$\bar{X}^*/\bar{T}$		$\bar{u}_0$	$\bar{T}$	N (Obs.)
	Raw	Noise-Adj.	Raw	Noise-Adj.			
<b>Overall mean:</b> <i>[Bootstrapped SE]</i>	1.89 [0.25]	1.23 [0.22]	0.0059 [0.0015]	0.0038 [0.0013]	0.18 [0.00]	56 [2]	1,809
<i>By return state:</i>							
1 (-20%)	5.83 [1.18]	4.83 [1.05]	0.0049 [0.0027]	0.0041 [0.0024]	0.17 [0.01]	200 [20]	26
2 (-15%)	11.61 [3.32]	5.70 [3.06]	0.0180 [0.0096]	0.0088 [0.0083]	0.22 [0.01]	141 [25]	19
3 (-10%)	5.76 [0.99]	2.37 [1.07]	0.0151 [0.0059]	0.0062 [0.0051]	0.21 [0.01]	81 [12]	49
4 (-5%)	2.67 [0.59]	1.39 [0.50]	0.0088 [0.0038]	0.0046 [0.0029]	0.14 [0.01]	42 [5]	272
5 (0%)	0.70 [0.16]	0.47 [0.14]	0.0045 [0.0019]	0.0030 [0.0017]	0.23 [0.00]	37 [2]	700
6 (+5%)	1.71 [0.35]	1.14 [0.34]	0.0039 [0.0015]	0.0026 [0.0014]	0.11 [0.01]	49 [3]	567
7 (+10%)	3.87 [1.00]	2.92 [1.03]	0.0053 [0.0023]	0.0040 [0.0023]	0.18 [0.01]	129 [9]	144
8 (+15%)	5.65 [1.48]	5.26 [1.48]	0.0060 [0.0027]	0.0056 [0.0027]	0.21 [0.01]	200 [11]	58
9 (+20%)	3.44 [0.89]	2.09 [1.27]	0.0032 [0.0015]	0.0020 [0.0020]	0.22 [0.01]	232 [9]	36
<i>By date:</i>							
1996–2000	10.89 [2.24]	9.67 [2.17]	0.0211 [0.0074]	0.0187 [0.0072]	0.21 [0.01]	107 [11]	109
2001–2005	1.75 [0.51]	0.55 [0.40]	0.0042 [0.0021]	0.0013 [0.0015]	0.22 [0.01]	90 [11]	112
2006–2010	1.25 [0.22]	0.68 [0.19]	0.0065 [0.0026]	0.0035 [0.0021]	0.17 [0.00]	32 [4]	502
2011–2015	1.75 [0.36]	1.09 [0.28]	0.0050 [0.0024]	0.0031 [0.0019]	0.19 [0.00]	67 [5]	530
2016–2018	0.36 [0.21]	-0.11 [0.14]	0.0011 [0.0017]	-0.0003 [0.0009]	0.16 [0.00]	50 [3]	556
<i>By <math>\pi_0^*</math>:</i>							
0–0.25	1.01 [0.60]	0.30 [0.55]	0.0055 [0.0048]	0.0017 [0.0047]	0.09 [0.01]	16 [2]	185
0.25–0.5	1.58 [0.21]	0.91 [0.17]	0.0067 [0.0017]	0.0039 [0.0014]	0.23 [0.00]	55 [3]	883
0.5–0.75	2.84 [0.61]	2.19 [0.58]	0.0053 [0.0020]	0.0041 [0.0019]	0.23 [0.00]	123 [7]	284
0.75–1	2.54 [0.95]	1.88 [0.90]	0.0048 [0.0031]	0.0036 [0.0029]	0.06 [0.00]	31 [2]	457

*Notes:* Empirical conditional means of risk-neutral excess movement  $\bar{X}^* \equiv \widehat{\mathbb{E}}[X_{i,j}^*]$  are calculated over all interior state pairs  $j = 2, \dots, 8$ , aside from averages by bin, which are calculated for each state pair separately. Standard errors are estimated using a block bootstrap for the normalized statistic  $\bar{X}^*/\bar{u}_0$  or  $\bar{X}^*/\bar{T}$ , with a block size of one month (where contracts are classified by the month in which they expire) and 10,000 draws.

**Table 3: Residual Excess Movement  $t$ -Statistics for Different  $\bar{\phi}$** 

<b>(a) Tighter Bound from Proposition 9(i)</b>						
	$\bar{\phi} = 1$	2	5	10	50	$\infty$
Overall $t$ -stat.:	5.19	5.15	4.98	4.64	3.03	0.78
<i>By return state:</i>						
1 (-20%)	4.06	3.69	3.04	2.57	1.94	1.72
2 (-15%)	2.13	2.13	2.13	2.13	2.14	2.14
3 (-10%)	2.23	2.23	2.23	2.23	2.23	2.23
4 (-5%)	2.78	2.78	2.77	2.77	2.76	2.76
5 (0%)	3.09	3.09	3.08	3.08	3.06	3.06
6 (+5%)	3.26	2.66	1.27	0.04	-1.71	-2.42
7 (+10%)	2.82	2.73	2.51	2.29	1.70	1.23
8 (+15%)	3.53	3.49	3.42	3.37	3.29	3.25
9 (+20%)	1.63	1.58	1.52	1.48	1.44	1.43

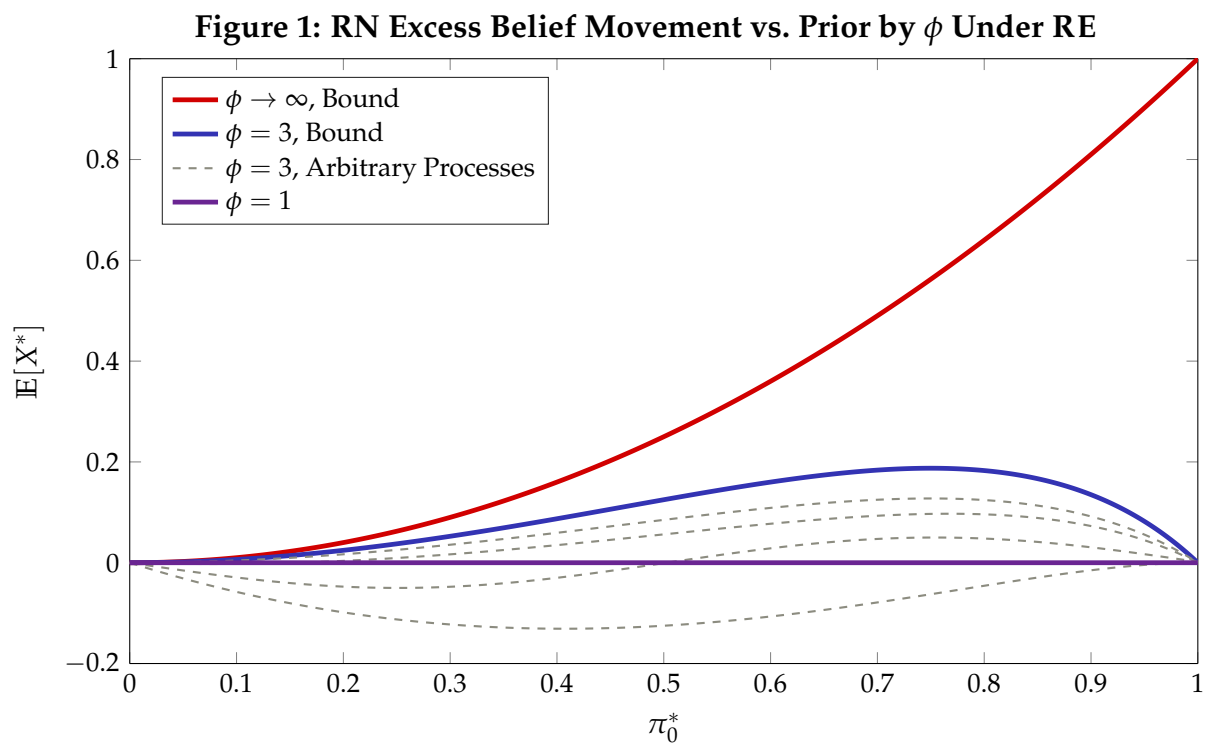
<b>(b) Conservative Bound from Proposition 9(ii)</b>						
	$\bar{\phi} = 1$	2	5	10	50	$\infty$
Overall $t$ -stat.:	5.19	4.08	2.68	1.75	0.07	-1.48
<i>By return state:</i>						
1 (-20%)	4.06	3.58	2.79	2.22	1.47	1.21
2 (-15%)	2.13	2.03	1.94	1.90	1.87	1.86
3 (-10%)	2.23	2.03	1.86	1.79	1.72	1.70
4 (-5%)	2.78	2.51	2.28	2.18	2.09	2.06
5 (0%)	3.09	1.86	0.74	0.26	-0.19	-0.31
6 (+5%)	3.26	1.96	-0.12	-1.77	-5.22	-8.39
7 (+10%)	2.82	2.38	1.73	1.28	0.50	0.06
8 (+15%)	3.53	3.20	2.77	2.51	2.21	2.10
9 (+20%)	1.63	1.36	1.05	0.89	0.72	0.68

*Notes:* Both panels report  $t$ -statistics for the average residuals  $\overline{e_i(\bar{\phi})}$  in (29) for different values of  $\bar{\phi}$ . Panel (a) uses  $e_i^{\Delta}(\bar{\phi})$ , with  $\Delta_i$  estimated as in panel (a) of Figure 2. Panel (b) uses  $e_i^{\text{main}}(\bar{\phi})$ . Standard errors are estimated using a block bootstrap with block size of one month and 10,000 draws. All statistics are calculated using conditional means of noise-adjusted excess movement.

**Table 4: Main Estimation Results**

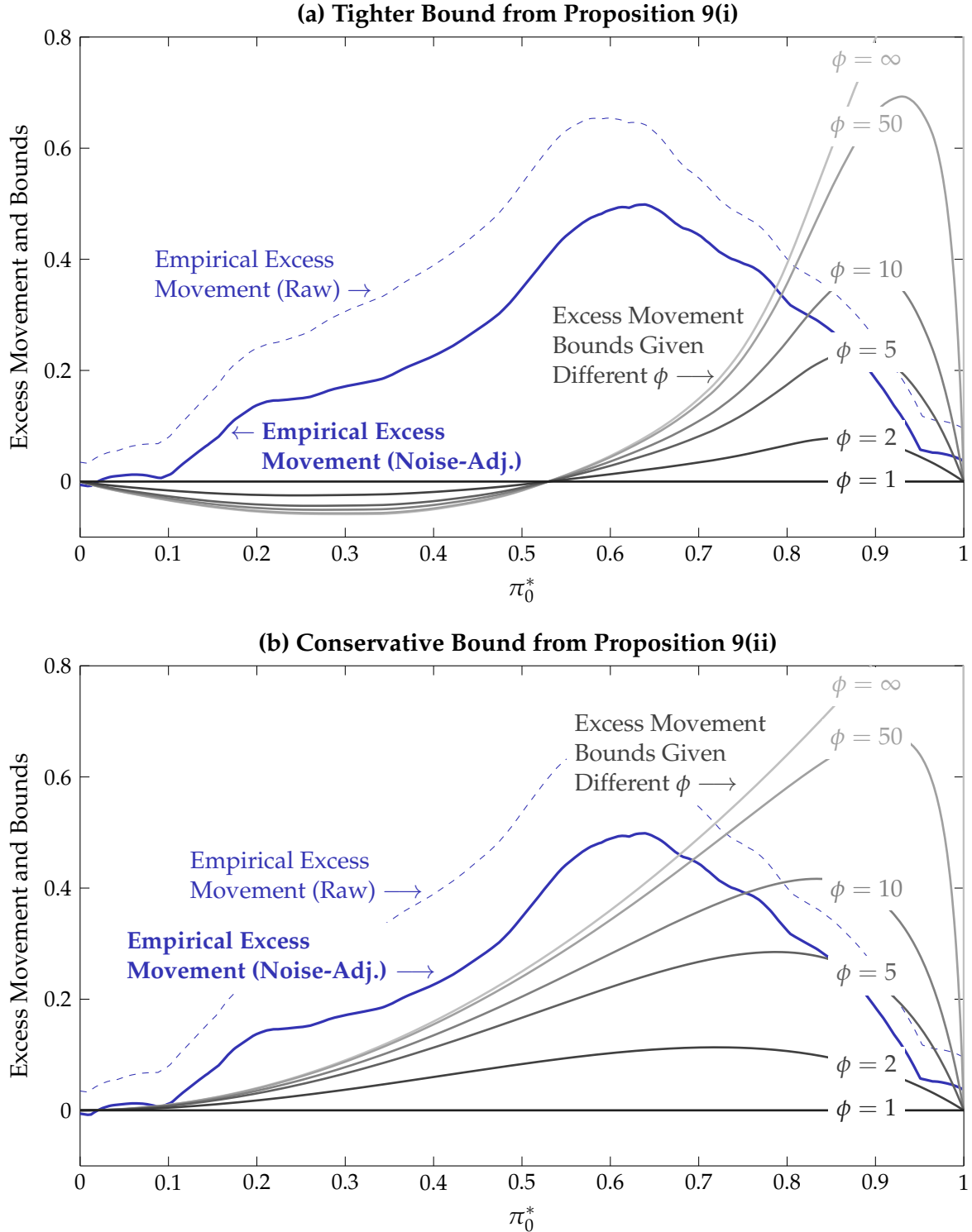
	<i>Conservative Lower Bound for:</i>	
	SDF Slope $\bar{\phi}$	RRA $\bar{\gamma}$
Overall bound: <i>[95% CI Lower Bound]</i>	54.7 [9.8]	1,075 [175]
<i>By return state:</i>		
1 (-20%)	$\infty$ [24.2]	$\infty$ [464]
2 (-15%)	$\infty$ [ $\infty$ ]	$\infty$ [ $\infty$ ]
3 (-10%)	$\infty$ [ $\infty$ ]	$\infty$ [ $\infty$ ]
4 (-5%)	$\infty$ [ $\infty$ ]	$\infty$ [ $\infty$ ]
5 (0%)	19.4 [2.1]	368 [22]
6 (+5%)	4.8 [2.2]	75 [24]
7 (+10%)	$\infty$ [4.6]	$\infty$ [73]
8 (+15%)	$\infty$ [ $\infty$ ]	$\infty$ [ $\infty$ ]
9 (+20%)	$\infty$ [1.0]	$\infty$ [1]

*Notes:* The first column reports estimates for the minimal value of  $\bar{\phi}$  satisfying the conservative bound for excess movement in [Proposition 9\(ii\)](#). These estimates are translated to relative risk aversion values  $\bar{\gamma}$  using [Proposition 8](#), as shown in the second column. Point estimates are obtained by finding the value  $\bar{\phi}$  such that  $e_i^{\text{main}}(\bar{\phi}) = 0$  in (29). Confidence interval lower bounds are obtained by inverting a test for  $\bar{\phi}$  using bootstrapped data; see Appendix B.9 for details. All estimates use conditional means of noise-adjusted excess movement.



*Note:* Theoretical bounds are obtained from the formulas in [Proposition 3](#) and [Corollary 1](#).

**Figure 2: Excess Movement vs. RN Prior: Data and Theoretical Bounds**



*Notes:* Empirical excess movement curves are kernel-weighted local averages (Epanechnikov kernel, bandwidth for  $\pi_0^*$  of 0.07) over all interior state pairs  $j = 2, \dots, 8$ . All statistics are estimates of conditional means  $\tilde{\mathbb{E}}[\cdot]$  for RN beliefs  $\tilde{\pi}_{i,i,j}^*$ , and theoretical curves correspond to  $\phi \equiv \tilde{\mathbb{E}}[\phi_{i,j}]$ , with notation simplified for clarity. Bounds for (a) obtain  $\hat{\Delta}$  using a kernel-weighted local average over  $\pi_0^*$  for each of the two terms in (17), with  $\hat{\Delta}$  then plugged into the inequality in Proposition 9(i). Bounds for (b) use only the second inequality in Proposition 9(ii).