NBER WORKING PAPER SERIES

THE VALUATION OF LONG-DATED ASSETS

Ian Martin

Working Paper 16219
http://www.nber.org/papers/w16219

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2010

I am grateful to David Backus, Brandon Bates, Jonathan Berk, John Campbell, John Cochrane, Peter DeMarzo, Darrell Duffie, Lars Hansen, Michael Harrison, Ilan Kremer, Steve Ross, Jeremy Stein, Dimitri Vayanos, Martin Weitzman, and seminar participants at Berkeley, LSE, and the Stanford Institute for Theoretical Economics conference on "New Quantitative Models of Asset Markets" for their comments. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The Valuation of Long-Dated Assets
Ian Martin
NBER Working Paper No. 16219
July 2010
JEL No. D53,G1,Q54

ABSTRACT

The expected time- and risk-adjusted cumulative return on any asset equals one at all horizons. Nonetheless, I show that a typical asset's realized time- and risk-adjusted cumulative return tends to zero almost surely. As a corollary, the value of a typical long-dated asset is driven by extreme events: either by good news at the level of the individual asset or by bad news at the aggregate level. In the case of the aggregate market, the fact that its Sharpe ratio is higher than its volatility suggests that bad news is the relevant consideration in practice.

Ian Martin
Graduate School of Business
Stanford University
Stanford, CA 94305
and NBER
ian.martin@gsb.stanford.edu

In the absence of arbitrage, the fundamental equation of asset pricing states that the expected time- and risk-adjusted cumulative return on any asset equals one at all horizons. This paper arrives at, and then interprets, an apparently paradoxical result: for a typical asset, the realized time- and risk-adjusted cumulative return tends to zero with probability one.

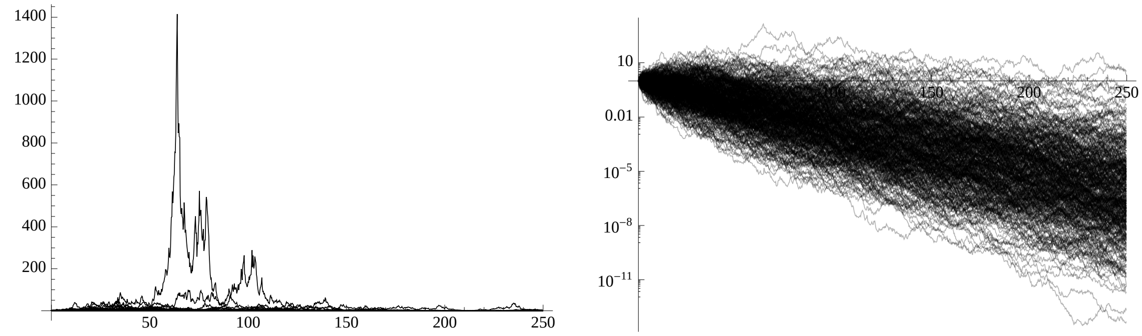
The objects of interest are the martingale $X_t \equiv M_1 R_1 \cdots M_t R_t$, and the random variable $X_\infty \equiv \lim_{t\to\infty} X_t$. ($M_t$ is a stochastic discount factor that prices payoffs at time t from the perspective of time $t-1$; R_t is the gross return on some arbitrary asset from time $t-1$ to time t.) The fundamental asset-pricing equation—$\mathbb{E}_{t-1} M_t R_t = 1$—implies that $\mathbb{E} X_t = 1$ for all finite t, so it is natural to expect that $\mathbb{E} X_\infty = 1$, too. It turns out that this may or may not be true; typically, in fact, it is not, and when it is not, $X_\infty = 0$.[1] I provide a variance criterion that dictates whether an asset is "typical" in this sense.

Where, then, do such assets get their long-run value—their $\mathbb{E} X_t = 1$—from? I show that when $X_\infty = 0$, X_t occasionally experiences enormous explosions that can be attributed to some combination of high $M_1 \cdots M_t$ and high $R_1 \cdots R_t$. The former possibility can be thought of as "bad news" at the aggregate level, and the latter as asset-specific "good news". It is important to emphasize that the existence and importance of such events emerge from the logic of arbitrage-free pricing alone. I neither assume nor exclude the possibility of, say, jumps in asset returns.

The following simple (and well-known) example shows what is going on. Suppose that there is a riskless asset with certain return $R_{f,t} \equiv e^{r_f}$ and a risky asset with return $R_t \equiv e^{\mu-\sigma^2/2+\sigma Z_t}$, where Z_t is standard Normal. $M_t \equiv e^{-r_f-\lambda^2/2-\lambda Z_t}$ is a valid SDF, where λ is the Sharpe ratio $(\mu - r_f)/\sigma$, so $X_t = e^{-(\lambda-\sigma)(Z_1+\cdots+Z_t)-(\lambda-\sigma)^2 t/2}$.

Setting $\sigma = 16\%$ and $\lambda = 50\%$, Figure 1a plots 400 sample paths of X_t over a 250 year horizon. Each sample path starts from $X_0 = 1$. Figure 1b shows the same 400 sample paths plotted on a log scale. Together, the figures illustrate the main results of the paper. First,

[1] This statement holds with probability one, or almost surely. Throughout the paper, I drop such qualifications in the interest of readability.



(a) Linear scale.

(b) Log scale.

Figure 1: 400 sample paths of X_t, plotted against time, over a 250-year horizon.

despite the fact that $\mathbb{E}X_t = 1$ for all t, just two of the 400 sample paths lie above 1 after 250 years. (If the plot were extended, we would see that these paths, too, eventually tend to zero. In the population, the median value of X_t after 250 years is $e^{-(0.50-0.16)^2 \times 250/2} < 10^{-6}$.) Second, this tendency for X_t to approach zero along sample paths is counterbalanced by occasional explosions in X_t: one sample path rises above 1400. The two figures together illustrate the principle that in the long run, extreme events are the dominant influence on asset prices. Third, the *empirical* fact that Sharpe ratios are high—$\lambda > \sigma$—means that in this example explosions in X_t can be attributed to very negative realizations of $Z_1 + \cdots + Z_t$, and hence to explosions in $M_1 \cdots M_t$, that is, to extremely bad news.

In this i.i.d.-lognormal example, the fact that $X_t \to 0$ can be seen as reflecting special properties of Brownian motion. In contrast, I need to impose almost no mathematical structure to derive the main results of this paper, which are presented in Sections 1 and 2. These rest only on a no-arbitrage assumption that leads naturally (in view of Harrison and Kreps (1979)) to the application of martingale methods.

With some extra structure—a conditional lognormality assumption—I am able to show, in the case of the aggregate market, that explosions in X_t can be attributed to bad news, by invoking the empirical fact that the market has a high Sharpe ratio. I also provide a result that characterizes when such explosions can be attributed to bad news in the general case, though the result requires imposition of structure of a different kind, in the shape of a function, κ, that is introduced in Section 3.

My approach is complementary to that of Hansen and Scheinkman (2009), who investigate long-run risk-return relationships in a somewhat more structured (continuous-time, Markov) environment. The two papers focus on quite distinct objects of interest: eigenfunction decompositions as a means of characterizing long-run discount rates in the case of Hansen and Scheinkman (2009), and the importance of rare events and the "edges" of the distribution of sample paths in the case of this paper.

There is also a link to the literature on equivalent martingale measures (Dalang, Morton and Willinger (1990), Schachermayer (1992)). When $X_\infty = 0$, an equivalent martingale measure does not exist, even if there is no arbitrage. The results of this paper attempt to demonstrate what this means in economic terms.

The principle that the value of a long-dated asset may be dictated by extreme outcomes is also explored by Weitzman (1998, 2009) and Gollier (2002) in the context of long-run interest rates and of cost-benefit analyses of environmental projects with payoffs in the distant future. In response, Nordhaus (2009) has suggested that Weitzman's (2009) logic rests on rather special assumptions about functional forms—notably on the properties of utility functions near zero and on the distribution of "consumption" (to be understood broadly) in the left tail. The present paper attempts to place the Weitzman argument on a more general footing, based on very weak assumptions, that is immune to these criticisms.

1 An apparent paradox...

Time is discrete; today is time 0. Consider a sequence of gross returns, R_t, on some limited-liability asset or investment strategy, and suppose that there is no arbitrage. For $t > 0$, we can therefore define M_t to be a stochastic discount factor (SDF) which prices payoffs at time t from the perspective of time $t-1$ (Harrison and Kreps (1979), Hansen and Richard (1987)). Then we have

$$M_t > 0, \quad R_t \geq 0, \quad \text{and} \quad \mathbb{E}_{t-1}(M_t R_t) = 1 \quad \text{for all } t. \tag{1}$$

M_t and R_t are random variables that only become known at time t.

Define the risk-adjusted return X_t, $t = 1, 2, 3, \ldots$, by

$$X_t \equiv M_1 R_1 \cdot M_2 R_2 \cdot \ldots \cdot M_t R_t.$$

It follows from (1) that $\mathbb{E} X_t = 1$ for all t. Moreover, X_t is a non-negative martingale, because

$$\begin{aligned}
\mathbb{E}_{t-1} X_t &= \mathbb{E}_{t-1}\left(M_1 R_1 \cdots M_t R_t\right) \\
&= M_1 R_1 \cdots M_{t-1} R_{t-1}\, \mathbb{E}_{t-1}\left(M_t R_t\right) \\
&= M_1 R_1 \cdots M_{t-1} R_{t-1} \\
&= X_{t-1}.
\end{aligned}$$

As a result, the random variable

$$X_\infty \equiv \lim_{t\to\infty} X_t = \lim_{t\to\infty} M_1 R_1 \cdot M_2 R_2 \cdot \ldots \cdot M_t R_t$$

almost surely exists and is finite, by the martingale convergence theorem of Doob (1953, p. 319). It is tempting to argue that

$$\mathbb{E} X_\infty = \mathbb{E} \lim_{t\to\infty} X_t \stackrel{?}{=} \lim_{t\to\infty} \mathbb{E} X_t = \lim_{t\to\infty} 1 = 1,$$

but, as I now show, the interchange of expectation and limit is not valid in general. The following two Propositions introduce and interpret the *variance criterion*[2]

$$\sum_{t=1}^{\infty} \mathrm{var}_{t-1} \sqrt{M_t R_t}.$$

Proposition 1. *If* $\sum \mathit{var}_{t-1}\sqrt{M_t R_t} = \infty$*, then* $X_\infty = 0$.

If $\sum \mathit{var}_{t-1}\sqrt{M_t R_t} < K$*, for some constant* $K < \infty$*, then* $\mathbb{E} X_\infty = 1$.

[2]Since the variance criterion is a sum of conditional variances, it is a random variable. Therefore the two cases (i) $\sum \mathrm{var}_{t-1}\sqrt{M_t R_t} = \infty$ and (ii) $\sum \mathrm{var}_{t-1}\sqrt{M_t R_t} < K$, for some constant $K < \infty$—should be understood to hold almost surely, as stated in footnote 1.

Proof. Let $a_t \equiv \mathbb{E}_{t-1}\sqrt{M_t R_t}$. By the absence of arbitrage, $\mathbb{E}_{t-1} M_t R_t = 1$, so the conditional form of Jensen's inequality implies that $a_t \leq 1$. Also, we trivially have $a_t > 0$. Define the random variables

$$Y_t = \frac{\sqrt{M_1 R_1}}{a_1} \frac{\sqrt{M_2 R_2}}{a_2} \cdots \frac{\sqrt{M_t R_t}}{a_t};$$

Y_t is then a martingale.

Suppose, first, that $\sum \text{var}_{t-1}\sqrt{M_t R_t} = \infty$ almost surely; equivalently, $\sum \left(1 - a_t^2\right) = \infty$. It follows, by a standard result—see, for example, Theorem 15.5 of Rudin (1987, p. 300)—that $\prod a_t^2 = 0$, and hence $\prod a_t = 0$. (Conversely, if $\sum \left(1 - a_t^2\right) < K$ for some finite constant K, then $\prod a_t^2 > \delta$, for some $\delta > 0$. This fact is used below.) By the martingale convergence theorem, Y_t almost surely has a finite limit Y_∞. But since $Y_\infty = \sqrt{X_\infty} / \prod a_t$, and $\prod a_t = 0$, it must be the case that $X_\infty = 0$.

Alternatively, suppose that (almost surely) $\sum \text{var}_{t-1}\sqrt{M_t R_t} < K$, for some constant $K < \infty$; equivalently, $\sum \left(1 - a_t^2\right) < K$. So $\prod a_t^2 > \delta$, for some $\delta > 0$. We then have $\mathbb{E} Y_t^2 \leq 1/\delta < \infty$, so the martingale Y_t is uniformly bounded in second moment. As a result,

$$\mathbb{E}\left(\max_t X_t\right) \leq \mathbb{E}\left(\max_t Y_t^2\right) \leq 4 \max_t \mathbb{E}\left(Y_t^2\right) < \infty,$$

the second inequality being the $\mathscr{L}^2$ inequality of Doob (1953, p. 317). The random variable $\max_t X_t$ is therefore integrable. Since $\max_t X_t$ dominates X_t, it follows that X_t is uniformly integrable, so $\mathbb{E} X_\infty = 1$ (and we also have $\mathbb{E} \max X_t < \infty$). $\square$

In the above proof, I have adapted the treatment of a result of Kakutani (1948) given by Williams (1995) by generalizing to allow for the empirically relevant case in which asset returns and the stochastic discount factor can be serially dependent.[3]

[3]This modification is not completely costless, since it comes at the expense of a mathematically less elegant result: in the serially independent case, the variance criterion is a real number rather than a random variable (since the conditional variances are variances) so the two alternatives—(i) $\sum \text{var}_{t-1}\sqrt{M_t R_t} = \infty$ or (ii) $\sum \text{var}_{t-1}\sqrt{M_t R_t} < K$, for some constant $K < \infty$—capture all the possibilities, and the above result is a dichotomy. In the serially dependent case, on the other hand, other theoretical possibilities arise: it is possible to construct examples in which, say, $\sum \text{var}_{t-1}\sqrt{M_t R_t} = \infty$ with probability 0.5 and $\sum \text{var}_{t-1}\sqrt{M_t R_t} < K$ with probability 0.5, though such examples do not appear to be relevant in practice.

To interpret the result, note that we only have $\sum \text{var}_{t-1}\sqrt{M_t R_t} < \infty$ if the conditional variance of $\sqrt{M_t R_t}$ declines rapidly to zero as $t \to \infty$: in other words, if $M_t R_t$ is roughly constant for large t. The following result makes this idea precise.

Proposition 2. *For $\sum var_{t-1}\sqrt{M_t R_t} = \infty$, it is sufficient (though not necessary) that $M_t R_t \not\to 1$.*

Proof. I will prove that whenever $\sum \text{var}_{t-1}\sqrt{M_t R_t} < \infty$, we have $M_t R_t \to 1$; the result follows. Suppose, then, that $\sum \text{var}_{t-1}\sqrt{M_t R_t} < \infty$. By the conditional form of Chebyshev's inequality,

$$\mathbb{P}_{t-1}\left(\left|\sqrt{M_t R_t} - \mathbb{E}_{t-1}\sqrt{M_t R_t}\right| \geq \varepsilon\right) \leq \frac{\text{var}_{t-1}\sqrt{M_t R_t}}{\varepsilon^2}$$

for arbitrary $\varepsilon > 0$, so

$$\sum_{t=1}^{\infty} \mathbb{P}_{t-1}\left(\left|\sqrt{M_t R_t} - \mathbb{E}_{t-1}\sqrt{M_t R_t}\right| \geq \varepsilon\right) \leq \frac{\sum_{t=1}^{\infty} \text{var}_{t-1}\sqrt{M_t R_t}}{\varepsilon^2} < \infty.$$

By the generalized Borel-Cantelli lemma (see, for example, Neveu (1975, p. 152)), it follows that $\left|\sqrt{M_t R_t} - \mathbb{E}_{t-1}\sqrt{M_t R_t}\right| < \varepsilon$ for all sufficiently large t. Since $\varepsilon > 0$ was arbitrary, we have established that

$$\sqrt{M_t R_t} - \mathbb{E}_{t-1}\sqrt{M_t R_t} \to 0. \tag{2}$$

Furthermore, if $\sum \text{var}_{t-1}\sqrt{M_t R_t} < \infty$, we have $\prod \mathbb{E}_{t-1}\sqrt{M_t R_t} > 0$ so, since $\mathbb{E}_{t-1}\sqrt{M_t R_t} \leq 1$, we must have

$$\mathbb{E}_{t-1}\sqrt{M_t R_t} \to 1. \tag{3}$$

(If not, it would have to be the case that for infinitely many t, $\mathbb{E}_{t-1}\sqrt{M_t R_t} < 1 - \delta$ for some $\delta \in (0,1)$, and hence $\left(\mathbb{E}_{t-1}\sqrt{M_t R_t}\right)^2 < 1 - 2\delta + \delta^2 < 1 - \delta$. But this implies that $\text{var}_{t-1}\sqrt{M_t R_t} > \delta$ for infinitely many t, which contradicts the assumption that $\sum \text{var}_{t-1}\sqrt{M_t R_t} < \infty$.)

It follows from (2) and (3) that $\sqrt{M_t R_t} \to 1$, and hence $M_t R_t \to 1$. $\square$

To understand Proposition 2, suppose that there is an SDF M_t^* and return R_t^* such that $M_t^* R_t^* = 1$. Applying Jensen's inequality to the fundamental asset pricing equation

$\mathbb{E}_{t-1} M_t^* R_t = 1$, for some arbitrary return R_t, we find that $\mathbb{E}_{t-1} \log R_t \leq \mathbb{E}_{t-1} \log (1/M_t^*) = \mathbb{E}_{t-1} \log R_t^*$. That is, R_t^* is the growth-optimal return with maximal expected log return. Moreover, we see that M_t^* is a special SDF, namely the reciprocal of the growth-optimal return (Long (1990)).[4]

Proposition 2 can therefore be interpreted as saying that if either the returns R_t are not asymptotically growth-optimal or the SDF M_t is not asymptotically the reciprocal of the growth-optimal return—or both—then $X_\infty = 0$.[5] This justifies the following terminology:

Definition 1. *We are in the* generic *case if R_t is not asymptotically growth-optimal or M_t is not asymptotically the reciprocal of the growth-optimal return, or both.*

In the generic case, then, $X_\infty = 0$. We are left with an apparent paradox. If such an asset's risk-adjusted return X_t tends to zero almost surely, where does its value—its $\mathbb{E} X_t = 1$—come from? Why isn't it *cheaper*?

2 ...and its resolution

The next result provides a resolution to this apparent paradox by expressing a sense in which such an asset's value can be attributed to outcomes in which X_t explodes.

[4]To see that this *is* an SDF, suppose that there are N assets with returns $R_t^{(i)}$, $i = 1, \ldots, N$. The growth-optimal portfolio is obtained by picking $\alpha_i, i = 1, \ldots, N$ to solve

$$\max_{\{\alpha_i\}} \mathbb{E} \log \sum \alpha_i R_t^{(i)} \quad \text{s.t.} \quad \sum \alpha_i = 1\,.$$

The first-order conditions are that, for each i,

$$\mathbb{E} \, \frac{R_t^{(i)}}{\sum \alpha_j R_t^{(j)}} = \lambda\,.$$

Multiplying both sides of this equation by α_i and summing over i, we find $\lambda = 1$, so

$$\mathbb{E} \, \frac{R_t^{(i)}}{\sum \alpha_j R_t^{(j)}} = 1 \quad \text{for all } i\,,$$

which exhibits $1/\sum \alpha_j R_t^{(j)} = 1/R_t^*$ as a valid SDF.

[5]As a theoretical matter, even if $M_t R_t \to 1$ we may have $\sum \mathrm{var}_{t-1} \sqrt{M_t R_t} = \infty$, and hence $X_\infty = 0$, if the convergence takes place sufficiently slowly. Thus my terminology is conservative.

Proposition 3. *In the generic case, in which $X_\infty = 0$, we have*

$$\mathbb{E} \max X_t = \infty \quad \text{and} \quad \mathbb{E}\left[X_t \log\left(1 + X_t\right)\right] \to \infty \quad \text{as} \quad t \to \infty. \tag{4}$$

In the non-generic case with $\mathbb{E}X_\infty = 1$, we have

$$\mathbb{E} \max X_t < \infty \tag{5}$$

and the following partial converse to the second part of (4): if $M_t R_t$ is bounded, uniformly in t, by some constant (which holds if, for example, the state space is finite) then $\mathbb{E}\left[X_t \log\left(1 + X_t\right)\right]$ remains bounded as $t \to \infty$.

Proof. Inequality (5) was shown in the course of the proof of Proposition 1. Similarly, the first part of (4) must hold because otherwise X_t would be uniformly integrable and we would have $\mathbb{E}X_\infty = 1$.

Next, since $f(x) \equiv (x \log x)_+$ is a convex function,[6] $(X_t \log X_t)_+$ is a submartingale by Jensen's inequality, so $\max \mathbb{E}\left(X_t \log X_t\right)_+ = \lim_{t \to \infty} \mathbb{E}\left(X_t \log X_t\right)_+$. But then, by Propositions IV-2-10 and IV-2-11 of Neveu (1975), the second part of (4) and its partial converse hold with $\mathbb{E}\left[(X_t \log X_t)_+\right]$ replacing $\mathbb{E}\left[X_t \log\left(1 + X_t\right)\right]$.

It remains to be shown that $\lim \mathbb{E}\left[X_t \log\left(1 + X_t\right)\right]$ is infinite iff $\lim \mathbb{E}\left[(X_t \log X_t)_+\right]$ is infinite. But this follows from the observation that when $X_t \geq 1$,

$$X_t \log X_t \leq \left(1 + X_t\right) \log\left(1 + X_t\right) \leq 2X_t \log\left(2X_t\right),$$

together with the fact that $\mathbb{E} \log\left(1 + X_t\right) \leq \mathbb{E}X_t = 1$, since $\log(1 + x) \leq x$. □

The two results in (4) are to be contrasted with the fact that $\mathbb{E}X_t = 1$ for all t. Since $\log\left(1 + X_t\right)$ grows *very* slowly with X_t, the fact that $\mathbb{E}X_t \log(1 + X_t)$ tends to infinity in the generic case indicates that X_t is enormous in some states of the world. (For example, it implies that for any $\varepsilon > 0$, $\mathbb{E}X_t^{1+\varepsilon} \to \infty$.)

The next Proposition considers the probability that $\max X_t$ exceeds some large number N. It places tight bounds on the rate at which this probability declines as N increases. Such events are rare, but not—in the generic case—*very* rare.

[6] I am using the notation $x_+ \equiv \max\{x, 0\}$.

Proposition 4. *In either case, large values of* $\max X_t$ *are rare, in the sense that for any* $N > 0$*,*

$$\mathbb{P}\left(\max X_t \geq N\right) \leq \frac{1}{N}. \tag{6}$$

In the generic case, this result is sharp, in the sense that for any $\varepsilon > 0$ *we can find arbitrarily large* N *such that*

$$\mathbb{P}\left(\max X_t \geq N\right) > \frac{1}{N^{1+\varepsilon}}.$$

Proof. Applying the submartingale inequality of Doob (1953, p. 314) to X_t, we have $N \cdot \mathbb{P}\left(\max_{t \leq T} X_t \geq N\right) \leq \mathbb{E}X_T = 1$, so

$$\mathbb{P}\left(\max_{t \leq T} X_t \geq N\right) \leq \frac{1}{N}.$$

Now, since

$$\mathbf{1}\left[\max_{t \leq T} X_t \geq N\right] \uparrow \mathbf{1}\left[\max_{t} X_t \geq N\right] \quad \text{as} \quad T \uparrow \infty,$$

the first statement follows from the monotone convergence theorem.

Suppose the second statement were false. Then there is an $\varepsilon > 0$ (to be thought of as small) and $C > 1$ (to be thought of as large) such that $\mathbb{P}\left(\max X_t \geq N\right) \leq 1/N^{1+\varepsilon}$ for all $N \geq C$. Since $\max X_t$ is positive, we would then have

$$\begin{aligned}
\mathbb{E}\max X_t &= \int_0^\infty \mathbb{P}\left(\max X_t \geq N\right)\, dN \\
&= \int_0^C \mathbb{P}\left(\max X_t \geq N\right)\, dN + \int_C^\infty \mathbb{P}\left(\max X_t \geq N\right)\, dN \\
&\leq C + \int_C^\infty \frac{1}{N^{1+\varepsilon}}\, dN \\
&< \infty,
\end{aligned}$$

in contradiction with Proposition 3. □

As a corollary of Propositions 3 and 4, Monte Carlo pricing of a long-dated asset may provide an unreliable indication of the asset's value, as this largely depends on states of the world that occur with very low probability. Ignoring, or failing to sample, such states of the world will lead to underpricing of the asset in question: in the case of long-term bonds, the tendency will be to overestimate long-run interest rates.

We have seen that $X_t \to 0$ in the generic case. How fast does convergence take place? To answer this question, it is convenient to introduce stochastic order notation.[7]

Definition 2. *Consider a sequence of random variables Z_t. We write $Z_t = O_p(1)$ if for any $\varepsilon > 0$ there exists a constant N such that*

$$\sup_t \mathbb{P}\left(|Z_t| > N\right) < \varepsilon,$$

and $Z_t = O_p(W_t)$—"Z_t is of the same order of magnitude as W_t"—if $Z_t/W_t = O_p(1)$.

For example, the central limit theorem implies that for i.i.d. random variables K_i with zero mean and finite variance,

$$\frac{1}{t}\sum_{i=1}^{t} K_i = O_p(1/\sqrt{t}),$$

which conveys the idea that the sample mean converges to the population mean at rate $\sqrt{t}$.

Proposition 5. *Recall the definition $a_k \equiv \mathbb{E}_{k-1}\sqrt{M_k R_k}$. We have*

$$X_t = O_p\left(\prod_{k=1}^{t} a_k^2\right).$$

Proof. In the proof of Proposition 1, I defined the non-negative martingale

$$Y_t = \frac{\sqrt{M_1 R_1}}{a_1}\frac{\sqrt{M_2 R_2}}{a_2}\cdots\frac{\sqrt{M_t R_t}}{a_t},$$

which has the almost-sure limit Y_∞ by the martingale convergence theorem. So,

$$\frac{X_t}{\prod_1^t a_k^2} = \frac{M_1 R_1 \cdots M_t R_t}{\prod_1^t a_k^2} \to Y_\infty^2,$$

where convergence is almost-sure; and hence also convergence takes place in distribution. The result follows from Prohorov's theorem. $\square$

[7]See van der Vaart (1998, pp. 12–13) for further details.

To take a simple example, consider an i.i.d. economy, and suppose that the asset of interest is not growth-optimal, so $\mathbb{E}_{t-1}\sqrt{M_t R_t}$ equals some constant $e^{-\delta} < 1$ for all t. Then $X_t = O_p\left(e^{-2\delta t}\right)$: convergence takes place exponentially fast.

3 How do extreme events take place?

In full generality, we have seen that for generic assets, $X_\infty = 0$, an apparently paradoxical result reconciled by the fact that $\mathbb{E}\max X_t = \infty$. That is, there are rare states of the world in which X_t is enormous. In such states, we have

$$M_1 R_1 \cdot M_2 R_2 \cdots M_t R_t \quad \text{very large,}$$

and so we must have some combination of large $M_1 \cdots M_t$ and large $R_1 \cdots R_t$. The former possibility, large $M_1 \cdots M_t$, corresponds roughly to the realization of a disastrously bad state of the world. In a consumption-based model with time-separable utility, for example, $M_1 \cdots M_t$ is large when marginal utility at time t is high. The latter possibility, large $R_1 \cdots R_t$, corresponds to a particularly favorable return realization for the asset in question.

To get more intuition for what happens in specific model economies, it is instructive to explore two simple examples that are in a sense polar opposites. For simplicity, I suppose in each case that there is a riskless asset whose return is constant over time.

First, consider a risk-neutral economy. Any asset that is not asymptotically riskless is generic, and the preceding results imply that returns on such assets satisfy

$$\frac{R_1 \cdots R_t}{R_f^t} \to 0 \quad \text{and} \quad \mathbb{E}\max \frac{R_1 \cdots R_t}{R_f^t} = \infty.$$

Since $M_1 \cdots M_t = 1/R_f^t$ is deterministic, the rare explosions that drive the second result can only be attributed to occasional explosions in $R_1 \cdots R_t$. That is, in a risk-neutral economy, the pricing of risky assets is driven by occasional bonanzas: low-probability events in which $R_1 \cdots R_t$ becomes very large.

For the second example, take an economy in which M_t is a nondegenerate random variable for all t, and consider the pricing of an "insurance" asset whose return R_t is a nondecreasing function of M_t. (If the riskless rate is constant then the riskless asset is an insurance asset, for example.) Then, $M_1 R_1 \cdots M_t R_t$ can *only* explode at times when $M_1 \cdots M_t$ explodes, so the pricing of long-dated insurance assets is driven by extreme bad news. This is a more general version of Weitzman's (1998) logic.

What can we say in the case of the aggregate market? From the Hansen-Jagannathan (1991) bound, combined with high available Sharpe ratios and a low riskless rate, it follows that $\sigma(M)$ is large relative to the volatility of the market, $\sigma(R)$. By imposing some more structure on the economy, in the form of a conditional lognormality assumption, we can use this observation to argue that explosions in X_t must be due to explosions in $M_1 \cdots M_t$, and hence to "bad news". It turns out that the critical condition that implies that explosions in X_t correspond to bad news is that the Sharpe ratio of the market is higher than its volatility. In the data, the Sharpe ratio of the market is on the order of 50% while its volatility is on the order of 16%, so this seems an innocuous assumption.

Proposition 6. *Suppose that the market return $R_t \equiv e^{\mu_{t-1}-\sigma_{t-1}^2/2+\sigma_{t-1}Z_t}$ is conditionally lognormal, and that there is a riskless asset with return $R_{f,t} \equiv e^{r_{f,t}}$. Then $M_t \equiv e^{-r_{f,t}-\lambda_{t-1}^2/2-\lambda_{t-1}Z_t}$ is a valid SDF, where $\lambda_t \equiv (\mu_t - r_{f,t+1})/\sigma_t$ is the Sharpe ratio on the market. Finally, suppose that the market Sharpe ratio and volatility satisfy $\lambda_t > \sigma_t + \varepsilon$ almost surely, for some $\varepsilon > 0$.*

Then we are in the generic case, so $X_\infty = 0$ and $\mathbb{E} \max_t X_t = \infty$. Moreover, long-run pricing is driven by the possibility of extremely bad outcomes, in the sense that explosions in X_t are driven by explosions in $M_1 \cdots M_t$.[8]

Proof. We have $M_t R_t = e^{-(\lambda_{t-1}-\sigma_{t-1})Z_t-(\lambda_{t-1}-\sigma_{t-1})^2/2}$, so the variance criterion is

$$\sum \operatorname{var}_{t-1} \sqrt{M_t R_t} = \sum \left(1 - e^{-(\lambda_{t-1}-\sigma_{t-1})^2/4}\right).$$

Since $\lambda_t - \sigma_t > \varepsilon$, the variance criterion is infinite, so without specifying anything further about the properties of λ_{t-1} and σ_{t-1}, we have $X_\infty = 0$. (In practice, we might want λ_{t-1}

[8] The appendix extends this result to allow for multiple risk factors $Z_{j,t}$, $j = 1, \dots, N$.

and σ_{t-1} to be high following realizations of Z_{t-1} or $\sigma_{t-2}Z_{t-1}$ that are negative and large in absolute value.)

By Proposition 3, we also have $\mathbb{E} \max X_t = \infty$. Since $\lambda_{t-1} - \sigma_{t-1} > 0$, $M_t R_t$ is large only if Z_t is negative, so explosions in X_t correspond unambiguously to bad news at the aggregate level (high $M_1 \cdots M_t$) rather than good news at the idiosyncratic level (high $R_1 \cdots R_t$). That is, pricing is driven by the possibility of extremely bad outcomes.[9] $\square$

The simplicity of the above result is largely due to the assumption of conditional log-normality, which amongst other things implies that the higher (conditional) cumulants[10] of $\log M$ and $\log R$ are zero. With non-zero higher cumulants, things become more complicated: it is possible to construct example economies in which (say) M is bounded, while period returns R have a small amount of weight in the extreme right tail, in such a way that $\sigma(M)$ is large (so the maximal Sharpe ratio is high) and $\sigma(R)$ relatively small, and yet explosions in $M_1 R_1 \cdots M_t R_t$ are due to right-tail events in which $R_1 \cdots R_t$ explodes. The goal of the remainder of this section is to refine this intuition, and to develop sufficient conditions that determine whether or not "explosions in X_t are driven by bad news" for a given parametric model, by using the theory of large deviations (and, more specifically, the Gärtner-Ellis theorem).

A natural metric for the extent to which explosions in X_t reflect bad news rather than good news is the conditional probability that $M_1 \cdots M_t > e^{\psi t}$, conditional on the event that $X_t > e^{\phi t}$. (Here ϕ and ψ are fixed growth rates and t is some large time.) Using the notation

$$P_t(\phi, \psi) \equiv \mathbb{P}\left(M_1 \cdots M_t > e^{\psi t} \middle| M_1 R_1 \cdots M_t R_t > e^{\phi t}\right),$$

we can say that bad news dominates consideration in the long run if $P_t(\phi, \psi) \to 1$ as $t \to \infty$. For fixed ϕ, this criterion is more (less) stringent if ψ is high (low).

[9] In the model of Campbell and Cochrane (1999), for example, the conditional standard deviation of the market return is not provided in closed form, but Figures 5 and 6 of the paper suggest that $\lambda_{t-1} - \sigma_{t-1} > 0$.

[10] By higher cumulants, I mean the third, fourth, fifth (etc) cumulants. See Backus, Foresi and Telmer (2001), Martin (2009), and Backus, Chernov and Martin (2009) for more on cumulants.

Some notation: let

$$\kappa(\theta_M, \theta_R) \equiv \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[(M_1 \cdots M_t)^{\theta_M} (R_1 \cdots R_t)^{\theta_R}\right] . \tag{7}$$

I assume that $\kappa(\theta_M, \theta_R)$ is finite and continuously differentiable for all $\theta_M, \theta_R \in \mathbb{R}$, and write $\kappa_M(\cdot,\cdot)$ and $\kappa_R(\cdot,\cdot)$ for the partial derivatives of κ with respect to its first and second argument, respectively. If the vectors $(\log M_t, \log R_t)$ are i.i.d. for all t, then the definition (7) reduces to $\kappa(\theta_M, \theta_R) = \log \mathbb{E}\left(M_1^{\theta_M} R_1^{\theta_R}\right)$, so $\kappa(\cdot,\cdot)$ is the cumulant-generating function of the random vector $(\log M_t, \log R_t)$.

Proposition 7. *Let θ_M^* and θ_R^* solve the equations*

$$\begin{aligned} \kappa_M(\theta_M^*, \theta_R^*) &= \psi \\ \kappa_R(\theta_M^*, \theta_R^*) &= \phi - \psi \,. \end{aligned}$$

Then $P_t(\phi, \psi) \to 1$ as $t \to \infty$ if $\theta_M^* < \theta_R^*$ and $P_t(\phi, \psi) \to 0$ as $t \to \infty$ if $\theta_M^* > \theta_R^*$.

Proof. See appendix. □

To link this result to the earlier results of this section, consider the simple special case in which $\kappa(\theta_M, \theta_R) = \mu_M\theta_M + \mu_R\theta_R + \sigma_{MM}\theta_M^2/2 + \sigma_{MR}\theta_M\theta_R + \sigma_{RR}\theta_R^2/2$. This case arises if—but not only if[11]—the vector $(\log M_t, \log R_t)$ is i.i.d. bivariate Normal with mean (μ_M, μ_R) and covariance matrix $\left(\begin{smallmatrix}\sigma_{MM} & \sigma_{MR} \\ \sigma_{MR} & \sigma_{RR}\end{smallmatrix}\right)$. By Proposition 7, $P_t(\phi, \psi) \to 1$ if

$$\left(\frac{\sigma_{MM} + \sigma_{MR}}{\sigma_{MM} + 2\sigma_{MR} + \sigma_{RR}}\right)\phi - \mu_R - \sigma_{RR}/2 - \sigma_{MR}/2 > \psi \,.$$

Fixing $\psi > 0$, this inequality is satisfied for sufficiently large ϕ, so long as

$$\sigma_{MM} + \sigma_{MR} > 0. \tag{8}$$

In the "insurance asset" case, (8) holds because $\sigma_{MR} \geq 0$. For risky assets with $\sigma_{MR} < 0$, (8) may still hold if σ_{MM} is sufficiently large relative to σ_{RR}: in the case considered in the introduction, for example, (8) is equivalent to $\lambda > \sigma$.

[11] Very roughly, the assumption is that the economy looks lognormal over long time periods.

4 Applications

I now present two examples to illustrate the applicability of these results.

4.1 A generalization of a traditional result

Suppose that the SDF is the reciprocal of the growth-optimal return, $M_t = 1/R_t^*$, but that R_t is not asymptotically growth-optimal. Then $M_t R_t \not\to 1$; this is an example of the generic case.

In this context, Proposition 1 amounts to the statement that $R_1 \cdots R_t/(R_1^* \cdots R_t^*) \to 0$ as $t \to \infty$: with probability one, the growth-optimal portfolio outperforms any non-growth-optimal portfolio by an arbitrary amount in the long run. It can therefore be thought of as extending the traditional results of Latané (1959), Samuelson (1971) and Markowitz (1976) to the non-i.i.d. case. Of greater interest, it demonstrates that these traditional results can be extended to SDFs $M_t \neq 1/R_t^*$. This is important because it is often desirable to work with SDFs that are more easily interpretable than $1/R_t^*$—for example, with SDFs proportional to the marginal value of wealth.

We also have a new result: $\mathbb{E} \max \left[R_1 \cdots R_t/(R_1^* \cdots R_t^*)\right] = \infty$. In the short run, the growth-optimal portfolio can hugely underperform. The probability of N-fold underperformance is at most $1/N$; on the other hand, for any $\varepsilon > 0$ we can find large N such that the probability of N-fold underperformance is at least $1/N^{1+\varepsilon}$.

4.2 The consumption path of a utility-maximizing investor

Suppose that there is an unconstrained investor in the economy who maximizes $\mathbb{E} \sum \beta^t u(C_t)$ for some concave, differentiable utility function $u(\cdot)$ and subjective discount factor β. The investor's marginal rate of substitution is then a valid SDF, and the above results imply that in the generic case,

$$\beta^t \frac{u'(C_t)}{u'(C_0)} R_1 \cdots R_t \to 0 \tag{9}$$

and yet

$$\mathbb{E}\max\left[\beta^t \frac{u'(C_t)}{u'(C_0)} R_1 \cdots R_t\right] = \infty. \tag{10}$$

For these equations to hold when applied to a riskless asset with time-t return $R_{f,t}$, for example, it is enough that pricing is not asymptotically risk-neutral, so $M_t R_{f,t} \not\to 1$. Suppose that this is so, and that the riskless rate is constant, $R_{f,t} = R_f$. Furthermore, suppose the investor is sufficiently patient that $\beta R_f \geq 1$. Then (9) implies that

$$u'(C_t) \to 0.$$

In particular, if $u(\cdot)$ satisfies the Inada conditions, then consumption tends to infinity in the long run. This is a result of Chamberlain and Wilson (2000): here, though, the result emerges as a special case of the more general results presented previously. Moreover, the observation that almost sure convergence to zero is inextricably linked with occasional explosions in X_t appears to be new.[12]

Conversely, if the investor is impatient, with $\beta R_f \leq 1$, then (10) implies that $\mathbb{E}\left[\max u'(C_t)\right] = \infty$, or equivalently—assuming $u'' < 0$—that $\mathbb{E}\left[u'(\min C_t)\right] = \infty$.

5 Conclusion

The absence of arbitrage implies that expected risk-adjusted returns on all assets equal one at all horizons. Proposition 1 provides a variance criterion that determines whether the realized risk-adjusted return on an asset tends to zero. Proposition 2 demonstrates that this is the relevant case unless (i) the asset is asymptotically growth-optimal *and* (ii) the SDF is asymptotically the reciprocal of the growth-optimal return. These apparently paradoxical findings are resolved by the fact that realized risk-adjusted returns explode (Proposition 3) occasionally (Proposition 4). Proposition 5 characterizes the speed of convergence of risk-adjusted returns.

[12]We can also strengthen the finding that $u'(C_t) \to 0$ by applying (9) to the growth-optimal asset, to conclude that $\beta^t R_1^* \cdots R_t^* u'(C_t) \to 0$. This is stronger because $R_1^* \cdots R_t^* / R_f^t \to \infty$, so $\beta^t R_1^* \cdots R_t^* \to \infty$.

In general, then, as a theoretical matter, explosions in risk-adjusted returns can be attributed either to spectacular outperformance of the asset in question, or to disastrously bad news at the aggregate level. I couple this observation with the empirical fact that the market has a high Sharpe ratio to argue that disasters are the relevant consideration in practice. As a corollary, cost-benefit analyses of long-dated assets, such as the payoffs to environmental projects, should pay special attention to worst-case scenarios; calculations based on back-of-the-envelope logic, or on small Monte-Carlo exercises, are likely to underestimate the value of such projects.

6 Bibliography

Backus, D. K., Chernov, M. and I. W. R. Martin (2009), "Disasters Implied by Equity Index Options," NBER Working Paper No. w15240.

Backus, D. K., Foresi, S. and C. I. Telmer (2001), "Affine Term Structure Models and the Forward Premium Anomaly," *Journal of Finance*, 56:1:279–304.

Bansal, R. and A. Yaron (2004), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59:4:1481–1509.

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*, 107:2:205–251.

Campbell, J. Y. and T. Vuolteenaho (2004), "Bad Beta, Good Beta," *American Economic Review*, 94:5:1249–1275.

Chamberlain, G. and C. A. Wilson (2000), "Optimal Intertemporal Consumption under Uncertainty," *Review of Economic Dynamics*, 3:365–395.

Dalang, R. C., A. Morton and W. Willinger (1990), "Equivalent Martingale Measures and No-Arbitrage in Stochastic Securities Market Models," *Stochastics and Stochastic Reports*, 29:185–201.

Dembo, A., and O. Zeitouni (1998), *Large Deviations Techniques and Applications*, Springer.

Doob, J. L. (1953), *Stochastic Processes*, John Wiley & Sons, New York.

Gollier, C. (2002), "Discounting an Uncertain Future," *Journal of Public Economics*, 85:149–166.

Hansen, L. P. and R. Jagannathan (1991), "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99:2:225–262.

Hansen, L. P. and S. F. Richard (1987), "The Role of Conditioning Information in Deducing

Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55:587–614.

Hansen, L. P. and J. A. Scheinkman (2009), "Long-Term Risk: An Operator Approach," *Econometrica*, 77:1:177–234.

Harrison, M. and D. Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20:381–408.

Kakutani, S. (1948), "On Equivalence of Infinite Product Measures," *Annals of Mathematics*, 49:214–224.

Latané, H. A. (1959), "Criteria for Choice Among Risky Ventures," *Journal of Political Economy*, 67:2:144–155.

Long, J. B. (1990), "The Numeraire Portfolio," *Journal of Financial Economics*, 26:29–69.

Markowitz, H. M. (1976), "Investment for the Long Run: New Evidence for an Old Rule," *Journal of Finance*, 31:5:1273–1286.

Martin, I. W. R. (2009), "Consumption-Based Asset Pricing with Higher Cumulants," Stanford GSB working paper.

Neveu, J. (1975), *Discrete-Parameter Martingales*, trans. T. P. Speed, North-Holland, Amsterdam.

Nordhaus, W. D. (2009), "An Analysis of the Dismal Theorem," working paper, Yale University.

Samuelson, P. A. (1971), "The 'Fallacy' of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling," *Proceedings of the National Academy of Sciences of the United States of America*, 68:10:2493–2496.

Schachermayer, W. (1994), "Martingale Measures for Discrete-Time Processes with Infinite Horizon," *Mathematical Finance*, 4:1:25–55.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press, UK.

Weitzman, M. L. (1998), "Why the Far-Distant Future Should Be Discounted At Its Lowest Possible Rate," *Journal of Environmental Economics and Management*, 36:201–208.

Weitzman, M. L. (2009), "On Modeling and Interpreting the Economics of Catastrophic Climate Change," *Review of Economics and Statistics*, 91:1:1–19.

Williams, D. (1995), *Probability with Martingales*, Cambridge University Press, Cambridge, UK.

A Appendix

A.1 Extension of Proposition 6 to the N-factor case

Suppose that the asset of interest loads on multiple conditionally Normal risk factors $Z_{j,t}$, indexed by $j = 1, \ldots, N$. Suppose, for example, that

$$R_t = \exp\left\{\mu_{t-1} + \boldsymbol{\beta}'_{t-1}\boldsymbol{Z}_t - (1/2)\boldsymbol{\beta}'_{t-1}\boldsymbol{V}_{t-1}\boldsymbol{\beta}_{t-1}\right\} ,$$

where $\boldsymbol{Z_t} = (Z_{1,t}, \ldots, Z_{N,t})$ is a vector of risk factors with conditional covariance matrix $\boldsymbol{V}_t$, and $\boldsymbol{\beta}_{t-1} = (\beta_{1,t-1}, \ldots, \beta_{N,t-1})$ is a vector of loadings on the N risk factors at time $t-1$. I assume that the signs on factors are chosen so that $\beta_{j,t} > 0$ for all j and t, so a large positive value of $Z_{j,t}$ is always good news for the asset.[13] I subtract off the variance term in the exponential so that $\mathbb{E}_{t-1}R_t = e^{\mu_{t-1}}$. For simplicity, suppose also that there is a riskless asset with return $R_{f,t} = e^{r_{f,t}}$.

Writing $\boldsymbol{\lambda}_{t-1} = (\lambda_{1,t-1}, \ldots, \lambda_{N,t-1})$ for the vector of risk prices, the SDF

$$M_t = \exp\left\{-r_{f,t} - \boldsymbol{\lambda}'_{t-1}\boldsymbol{Z}_t - (1/2)\boldsymbol{\lambda}'_{t-1}\boldsymbol{V}_{t-1}\boldsymbol{\lambda}_{t-1}\right\},$$

is valid so long as the risk premium, the price of risk, $\boldsymbol{\lambda}_{t-1}$, and the quantity of risk, $\boldsymbol{V}_{t-1}\boldsymbol{\beta}_{t-1}$, are linked by the relationship $\mu_{t-1} - r_{f,t} = \boldsymbol{\beta}'_{t-1}\boldsymbol{V}_{t-1}\boldsymbol{\lambda}_{t-1}$. It follows that

$$M_tR_t = \exp\left\{-\left(\boldsymbol{\lambda}_{t-1} - \boldsymbol{\beta}_{t-1}\right)'\boldsymbol{Z}_t - (1/2)\left(\boldsymbol{\lambda}_{t-1} - \boldsymbol{\beta}_{t-1}\right)'\boldsymbol{V}_{t-1}\left(\boldsymbol{\lambda}_{t-1} - \boldsymbol{\beta}_{t-1}\right)\right\}.$$

So, if $\lambda_{j,t-1} - \beta_{j,t-1}$ is almost surely positive (respectively, negative) then factor j is important in the long run due to the possibility of long sequences of negative $Z_{j,t}$, representing disasters (respectively, positive $Z_{j,t}$, representing bonanzas).

In the two-beta model of Campbell and Vuolteenaho (2004), two factors drive market returns: $Z_{1,t} = N_{CF,t}$ "cashflow news" and $Z_{2,t} = -N_{DR,t}$ "discount-rate news". In my notation, the market return has unit loading on each factor, so $\beta_{CF,t} = \beta_{DR,t} = 1$. Equation (8) of Campbell and Vuolteenaho's paper expresses the fact that the price of cashflow

[13] The loss of generality here—the asset's factor loading cannot change sign over time—simplifies subsequent interpretation.

news risk, $\lambda_{CF,t}$, equals the coefficient of risk aversion, γ, while the price of discount-rate news risk, $\lambda_{DR,t}$, is equal to one. Thus, whenever risk aversion is greater than one, so $\lambda_{CF,t} - \beta_{CF,t} = \gamma - 1 > 0$, the dominant concern in the long run is the possibility of cashflow disaster. On the other hand, discount-rate news has no long-run impact in this model, since $\lambda_{DR,t} - \beta_{DR,t} = 0$. In fact, in any model in which price-dividend ratios are stationary, so discount-rate news has no long-run impact on asset prices, this logic implies that the price of discount-rate risk cannot systematically be either greater or less than one.

In the long-run risks model of Bansal and Yaron (2004), there are again two priced risk factors: an expected consumption growth factor (e) and a consumption volatility factor (w). Using the notation of Bansal and Yaron, it can be seen that $\lambda_{m,e} > \beta_{m,e}$ if and only if risk aversion γ is greater than the "leverage ratio" ϕ, which holds in their calibration. Similarly, $\lambda_{m,w} < \beta_{m,w} < 0$.[14] Thus long-run pricing is driven by the possibility of disastrously low shocks to the expected consumption growth factor and disastrously high shocks to the consumption volatility factor.

A.2 Proof of Proposition 7

Proof. By Bayes' rule,

$$\begin{aligned} P_t(\phi,\psi) &= \frac{\mathbb{P}\left(G_{M,t} > \psi \text{ and } G_{M,t} + G_{R,t} > \phi\right)}{\mathbb{P}\left(G_{M,t} + G_{R,t} > \phi\right)} \\ &= \frac{\mathbb{P}(A_t)}{\mathbb{P}(A_t) + \mathbb{P}(B_t)}, \end{aligned}$$

where $G_{M,t} \equiv \frac{1}{t}\sum_1^t \log M_i$, $G_{R,t} \equiv \frac{1}{t}\sum_1^t \log R_i$, and A_t and B_t are the (disjoint) events "$G_{M,t} > \psi$ and $G_{M,t} + G_{R,t} > \phi$" and "$G_{M,t} < \psi$ and $G_{M,t} + G_{R,t} > \phi$".

When $\phi > 0$, $\mathbb{P}(A_t) + \mathbb{P}(B_t)$ tends to zero as $t \to \infty$. (To see this, note that $\mathbb{P}(A_t) + \mathbb{P}(B_t) = \mathbb{P}(M_1 \cdots R_t > e^{\phi t})$. Now pick arbitrary $\varepsilon > 0$. As a corollary of the first part of Proposition 6, if we take T large enough that $e^{\phi T} > 1/\varepsilon$, then $\mathbb{P}(M_1 \cdots R_t > e^{\phi t}) < \varepsilon$ for all $t > T$. That is, $\mathbb{P}(A_t) + \mathbb{P}(B_t) \to 0$.) Since $\mathbb{P}(A_t) + \mathbb{P}(B_t)$ tends to zero, $\mathbb{P}(A_t)$ and $\mathbb{P}(B_t)$ must each tend to zero.

[14] Since $\beta_{m,w} < 0$, in conflict with my earlier notational assumption, it is indeed the case that when $\lambda_{m,w} < \beta_{m,w}$, explosions in X_t occur at times of disaster.

The goal is now to analyze the rates at which $\mathbb{P}(A_t)$ and $\mathbb{P}(B_t)$ tend to zero. We will have $P_t(\phi,\psi) \to 1$ if $\mathbb{P}(B_t)$ tends to zero at a faster rate than $\mathbb{P}(A_t)$, and conversely $P_t(\phi,\psi) \to 0$ if $\mathbb{P}(A_t)$ tends to zero faster than $\mathbb{P}(B_t)$. So we must find a condition that ensures that $\mathbb{P}(B_t)$ tends to zero faster than $\mathbb{P}(A_t)$:

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}(\overline{B}_t) \leq \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}(A_t)\,, \tag{11}$$

where $\overline{B}_t$ is the event "$G_{M,t} \leq \psi$ and $G_{M,t} + G_{R,t} \geq \phi$". (The argument for the converse condition, which ensures that $\mathbb{P}(A_t) \to 0$ faster than $\mathbb{P}(B_t) \to 0$, is very similar, so is omitted.)

Let $\kappa^*(x_M, x_R) \equiv \sup_{\theta_M,\theta_R \in \mathbb{R}} x_M\theta_M + x_R\theta_R - \kappa(\theta_M, \theta_R)$, the Fenchel-Legendre transform of $\kappa(\cdot,\cdot)$. The Gärtner-Ellis theorem[15] implies that (11) holds if

$$\inf_{\substack{x_M > \psi \\ x_M + x_R \geq \phi}} \kappa^*(x_M, x_R) \leq \inf_{\substack{x_M \leq \psi \\ x_M + x_R \geq \phi}} \kappa^*(x_M, x_R)\,. \tag{12}$$

The function κ^* has the following properties: (i) it is convex (by Lemma 2.3.9 of Dembo and Zeitouni (1998, p. 46)); (ii) $\kappa^*(x_M, x_R) \geq 0$ (since it is at least as large as $x_M \cdot 0 + x_R \cdot 0 - \kappa(0,0) = 0$); (iii) $\kappa^*(x_M, x_R) \geq x_M + x_R$ (since it is at least as large as $x_M \cdot 1 + x_R \cdot 1 - \kappa(1,1) = x_M + x_R$); (iv) $\kappa^*(\mu_M, \mu_R) = 0$ where $\mu_M \equiv \kappa_M(0,0)$ and $\mu_R \equiv \kappa_R(0,0)$, so κ^* attains its global minimum at (μ_M, μ_R).

From (iii) and (iv), $\mu_M + \mu_R \leq 0$, so $(\mu_M, \mu_R) \notin \{(x_M, x_R) : x_M + x_R \geq \phi\}$. It follows by convexity that κ^* attains its minimum over $\{(x_M, x_R) : x_M + x_R \geq \phi\}$ on the boundary of the set, i.e. on the line $\{(x_M, x_R) : x_M + x_R = \phi\}$. The question is then whether the minimum is attained for x_M greater than ψ or less than ψ. Setting $f(x) \equiv \kappa^*(x, \phi - x)$, (12) is satisfied if $f'(\psi) < 0$, or equivalently $\kappa^*_M(\psi, \phi - \psi) < \kappa^*_R(\psi, \phi - \psi)$, where κ^*_M denotes the derivative of κ^* with respect to its first argument, and similarly for κ^*_R. The result follows by the envelope theorem. □

[15] For a proof of the theorem, see Theorem 2.3.6 in Dembo and Zeitouni (1998, p. 44). The simplified version of the theorem outlined in Remark (c) (p. 45) suffices, due to the assumption that $\kappa(\theta_M, \theta_R) < \infty$ for all $\theta_M, \theta_R \in \mathbb{R}$.