On the Valuation of Long-Dated Assets

Ian Martin

Stanford University and National Bureau of Economic Research

I show that the pricing of a broad class of long-dated assets is driven by the possibility of extraordinarily bad news. This result does not depend on any assumptions about the existence of disasters, nor does it apply only to assets that hedge bad outcomes; indeed, it applies even to long-dated claims on the market in a lognormal world if the market's Sharpe ratio is higher than its volatility, as appears to be the case in practice.

This paper advances a general principle that the valuation of a broad class of long-dated assets is driven by the possibility of extreme outcomes. Versions of this principle have been developed by Weitzman (1998), in the context of the discount rate appropriate for events that occur in the far-distant future, and by Weitzman (2009), in the form of a "dismal theorem" that highlights the extraordinary influence of extreme events in a certain category of models and whose intended application is to cost-benefit analyses of projects that might mitigate the effects of climate change (see also Gollier 2002, 2008; Weitzman 2007).

But Weitzman's dismal theorem has been criticized by Nordhaus (2011) on the grounds that it relies on particular distributional assumptions and on marginal utility that becomes infinite at zero consumption. If taken seriously, these assumptions imply that society should be prepared to pay almost any price to avoid, say, an astronomically

I am grateful to Dave Backus, Brandon Bates, Jonathan Berk, John Campbell, John Cochrane, Peter DeMarzo, Darrell Duffie, Lars Hansen, Mike Harrison, Ilan Kremer, Steve Ross, Jeremy Stein, Dimitri Vayanos, and Martin Weitzman; to seminar participants at Berkeley, the London School of Economics, and the Stanford Institute for Theoretical Economics conference in 2009; and to the editor, Monika Piazzesi, and anonymous referees for their comments.

[*Journal of Political Economy,* 2012, vol. 120, no. 2]

unlikely but potentially catastrophic meteorite strike, a conclusion that Nordhaus finds unacceptable.

I avoid taking a stand on this issue by starting from an apparently tangential observation that, in its most stripped-down form, relies only on a no-arbitrage assumption. If we write M_t for the stochastic discount factor (SDF) that prices time t payoffs from the perspective of time $t-1$ and R_t for any asset's gross return from time $t-1$ to time t, then it is a fundamental result of asset pricing that the cumulative time- and risk-adjusted return $X_t \equiv M_1 R_1 \cdots M_t R_t$ is a martingale. Although the focus is typically on the expectation, $\mathbb{E}X_t = 1$, one can also ask: What is it about the sample paths of X_t that leads its expectation to equal one? Although this question may seem hopelessly vague, we will see, under minimal assumptions, that X_t has certain distinctive features. In particular, I apply a result of Kakutani (1948)—a theorem on product martingales, not the fixed-point theorem that is used to prove the existence of Nash equilibrium—to show that although X_t has an expected value of one at all horizons, it tends to zero almost surely[1] for generic assets.

Where, then, do such assets get their value—their $\mathbb{E}X_t = 1$—from? I show that in this generic case, X_t occasionally experiences enormous explosions that can be attributed to some combination of explosions in $M_1 \cdots M_t$ and explosions in $R_1 \cdots R_t$. Interestingly, the two alternatives have very different interpretations. The first represents bad news at the aggregate level. I will refer to such outcomes as *aggregate disasters*, although it should be emphasized that I neither assume nor exclude the possibility of rare disasters in the sense of Rietz (1988) and Barro (2006): aggregate disasters in my sense can also unfold slowly in thin-tailed models via a grinding sequence of negative realizations. The second possibility represents good news at the level of the individual asset. In the case of long-dated *hedge assets*—assets such as climate change–mitigation projects whose returns are low when marginal utility is low—valuation is overwhelmingly driven by the possibility of aggregate disaster.

A more prosaic illustration of this result is provided by Warren Buffett in his 2008 letter to the shareholders of Berkshire Hathaway (http://www.berkshirehathaway.com/letters/2008ltr.pdf). Consider selling at-the-money European put options on the S&P 500 index, expiring in 100 years. Plugging reasonable values for volatility and interest rate into the Black-Scholes formula, Buffett calculates that with a notional value of $1 billion, this trade would generate an incoming cash flow of $2.5 million up front. What of the potential outflows? When what he considers an extremely conservative forecast is used (a 1 percent probability

[1] From now on, I drop "almost surely" qualifications in the interest of readability.

of a decline in the index over 100 years and an expected return of −50 percent conditional on a decline) the expected cash outflow in 100 years' time is $5 million. But even in this unfortunate state of the world, the two cash flows would be equivalent to borrowing for 100 years at an interest rate of only 0.7 percent. Buffett concludes that "the Black-Scholes formula, even though it is the standard for establishing the dollar liability for options, produces strange results when the long-term variety are being valued." Here again we are leaning on a particular parametric model. But in Buffett's example, unlike Weitzman's, the model in question—the Black-Scholes model—postulates a particularly nicely behaved world in which the underlying asset price follows a geometric Brownian motion. This makes the phenomenon he describes more striking, not less: long-dated puts are "surprisingly expensive" even in a lognormal model. Again, valuation is driven by the possibility of disaster.

The same may be true even for assets whose returns are low in bad times—and even in a thin-tailed lognormal environment. 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. Then $M_t \equiv e^{-r_f-\lambda^2/2-\lambda Z_t}$ is a valid SDF, where λ is the Sharpe ratio $(\mu - r_f)/\sigma$, so $X_t = e^{-(\lambda-\sigma)(Z_1+\cdots+Z_t)-(\lambda-\sigma)^2 t/2}$. The return volatility of the S&P 500 index is about 16 percent while its Sharpe ratio is about 50 percent, so let $\sigma = 16$ percent and $\lambda = 50$ percent. Figure 1*A* plots 400 sample paths of X_t over a 250-year horizon. Each sample path starts from $X_0 = 1$. Figure 1*B* shows the same 400 sample paths plotted on a log scale. Together the figures illustrate the results of the paper. First, despite the fact that $\mathbb{E}X_t = 1$ for all t, just two of the 400 sample paths lie above one after 250 years. If the plot were extended, we would see that these paths, too, eventually tend to zero. In population, the median value of X_t after 250 years is $e^{-(0.50-0.16)^2\times 250/2} < 10^{-6}$. Second, the tendency for X_t to approach zero along sample paths is counterbalanced by occasional explosions in X_t: one sample path rises above 2,700. The possibility of such sample paths is critically important for the valuation of long-dated assets; these are the paths, lying at the very edge of the distribution, that Buffett neglects to consider. Third, the empirical fact that Sharpe ratios are high—$\lambda > \sigma$—means that in this example explosions in X_t can be attributed to negative realizations of $Z_1 + \cdots + Z_t$ and hence to aggregate disasters.

In the body of the paper, I generalize to a richer conditionally lognormal example and to the nonlognormal case, though the latter requires an economic assumption to link the SDF to the return on the market. In more general parametric models, the theory of large deviations (and specifically the Gärtner-Ellis theorem) can be used in conjunction with the methods of Hansen and Scheinkman (2009) and Han-

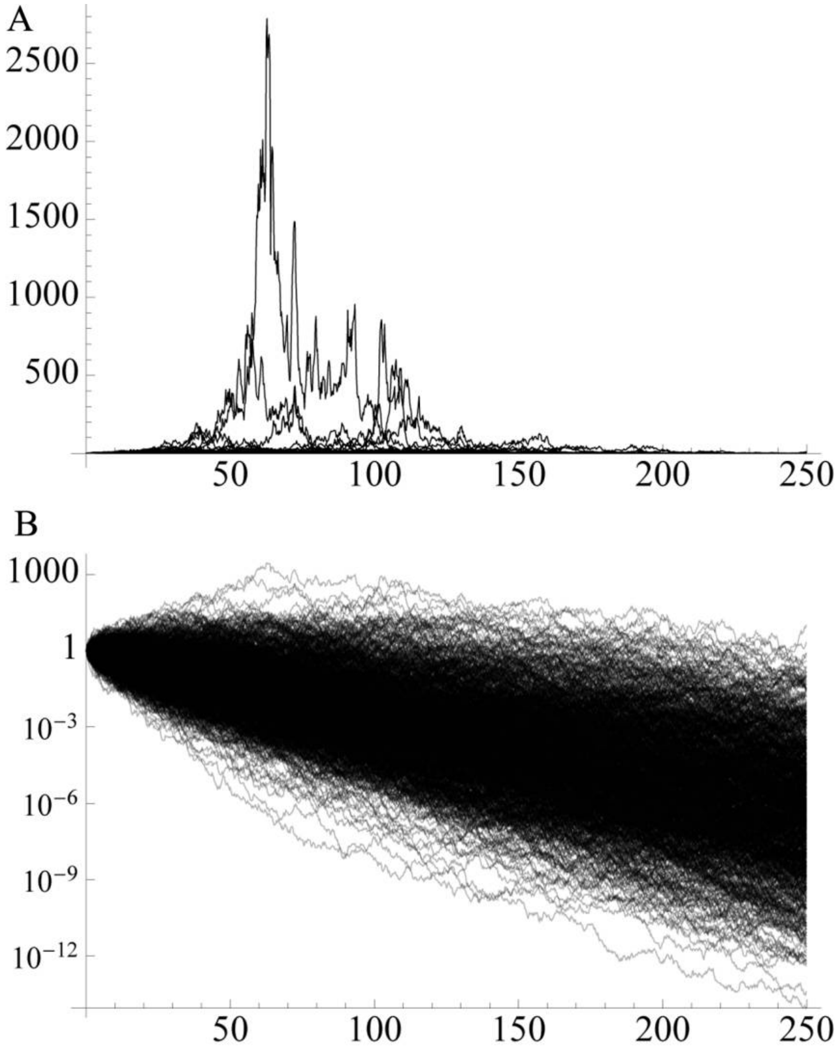

FIG. 1.—400 sample paths of X_t over a 250-year horizon: *A*, linear scale; *B*, log scale

sen (forthcoming) to assess the relative importance of disasters and idiosyncratic good news. I relegate these results to online Appendix B because the earlier results on hedge assets apply very generally and do not require this extra technology. When not included in the text, proofs of the results are in Appendix A.

I. Kakutani and the Long Run

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 an SDF that prices payoffs at time t from the perspective of time $t-1$. Define the *risk-adjusted cumulative return* $X_t \equiv M_1R_1 \cdot M_2R_2 \cdot \cdots \cdot M_tR_t$; then $\mathbb{E}X_t = 1$ for all t and X_t is a nonnegative martingale. As a result, the random variable $X_\infty \equiv \lim_{t\to\infty} X_t$ almost surely exists and is finite, by the martingale convergence theorem of Doob (1953, 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 this interchange of expectation and limit is not valid in general.

Definition 1. We are in the generic case if $M_tR_t \nrightarrow 1$.

To get a feel for this definition, suppose that the SDF M_t and return R_t satisfy $M_tR_t = 1$. Given an arbitrary return $\tilde{R}_t$, we have $\mathbb{E}_{t-1}M_t\tilde{R}_t = 1$, so by Jensen's inequality, $\mathbb{E}_{t-1}\log\tilde{R}_t \le \mathbb{E}_{t-1}\log(1/M_t) = \mathbb{E}_{t-1}\log R_t$. This implies that R_t is the growth-optimal return. Moreover, M_t is a special SDF, namely, the reciprocal of the growth-optimal return (Long 1990). So the generic case applies unless both R_t is asymptotically growth optimal and the SDF M_t is asymptotically the reciprocal of the growth-optimal return.[2]

Result 1. Although $\mathbb{E}X_t = 1$ for all t, in the generic case we have $X_\infty = 0$. On the other hand, (i) $\mathbb{E}\max X_t = \infty$ and (ii) for any $\varepsilon > 0$, $\mathbb{E}X_t^{1+\varepsilon} \to \infty$ as $t \to \infty$.

If a generic 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? The answer is that such an asset's value can be attributed to the presence of a small number of extreme sample paths along which X_t explodes, as demonstrated by the fact that $\mathbb{E}\max X_t = \infty$ and $\mathbb{E}X_t^{1+\varepsilon} \to \infty$.[3]

The next result 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. In the generic case, such events are rare, but not *very* rare.

Result 2. In either case, explosions in X_t are rare in the sense that, for any $N > 0$, $\mathbb{P}(\max X_t \ge N) \le 1/N$. In the generic case this result is sharp in the sense that for any $\varepsilon > 0$ we can find arbitrarily large N such that $\mathbb{P}(\max X_t \ge N) > 1/N^{1+\varepsilon}$.

[2] In a complete market, the SDF is unique and so must equal the reciprocal of the growth-optimal return (which is always an SDF). The generic case applies as long as R_t is not asymptotically growth optimal.

[3] The proof of result 1 actually establishes the stronger result that $\mathbb{E}X_t\log(1 + X_t) \to \infty$.

For generic assets, there are rare states of the world in which X_t is enormous. In such states, we have $M_1 R_1 \cdot M_2 R_2 \cdots M_t R_t$ 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 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 series of returns for the asset in question. To get more intuition for what happens in specific model economies, I now consider two examples that are in a sense polar opposites.

First, consider a risk-neutral economy with a constant riskless rate R_f. Any asset that is not asymptotically riskless is generic, and from the above results we have

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

Since $M_1 \cdots M_t = 1/R_f^t$ is deterministic, the rare explosions that drive the second result can be attributed only 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 a hedge asset whose return R_t is a (weakly) increasing function of M_t. Then $M_1 R_1 \cdots M_t R_t$ can explode only at times when $M_1 \cdots M_t$ explodes, so we have the following result.

RESULT 3. Pricing of long-dated hedge assets is driven by the possibility of aggregate disaster.

This is a more general version of Weitzman's (1998) logic. What can we say in the case of a risky asset such as 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. 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 percent whereas its volatility is on the order of 16 percent, so this is a mild assumption.

RESULT 4. 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 the riskless return is $R_{f,t} \equiv e^{r_{f,t}}$. 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, the pricing of long-dated claims on the market is driven by the possibility of extremely bad outcomes, in the sense that explosions in X_t are driven by explosions in $M_1 \cdots M_t$.

This result applies to a wide range of models, and it can be extended to allow for multiple risk factors $Z_{j,t}$, $j = 1, \ldots, N$, and for imperfect correlation between $\log R_t$ and $\log M_t$ (see online App. B). It does, however, depend in a more important way on the conditional lognormality assumption, which implies that the higher conditional cumulants of log M and log R are zero.[4] With nonzero higher cumulants, it becomes possible to construct examples in which (say) M is bounded whereas R has a small amount of weight in the extreme right tail, in such a way that SDF volatility is large (so the maximal Sharpe ratio is high) and return volatility 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. In Appendix B, I show how to use the theory of large deviations to determine whether or not explosions are driven by bad news in any given parametric model.

A simpler approach is to note that in the lognormal setting of result 4, the log SDF and log market return are tightly linked by a relation $\log M_t = \alpha_{t-1} - \gamma_{t-1} \log R_t$; the assumption that the market's Sharpe ratio is greater than its volatility is equivalent to the restriction $\gamma_{t-1} > 1$. To generalize result 4 to the nonlognormal setting in a tractable and somewhat flexible way, we can assume that this latter relationship holds without imposing the requirement that R_t and M_t are lognormal. A sufficient condition for this to be the case is that there is a representative agent with power utility $u(\cdot)$ and risk aversion greater than one, who chooses portfolio weights $\{w_i\}$ to solve the maximization problem

$$\max_{\{w_i\}} \mathbb{E}_t u\left(\sum_i w_i R_{t+1}^{(i)}\right) \quad \text{subject to} \quad \sum_i w_i = 1 \tag{1}$$

and who ends up holding the market, so that $\sum w_i R_{t+1}^{(i)} = R_{t+1}$. The first-order conditions for this problem imply that $u'(R_{t+1})/\mathbb{E}_t[R_{t+1} u'(R_{t+1})]$ is an SDF and hence that $\log M_{t+1} = \alpha_t - \gamma_t \log R_{t+1}$, where α_t is an unimportant quantity known at time t and $\gamma_t > 1$ is the agent's relative risk aversion. Imposing this assumption on preferences gives us some traction by pinning down the behavior of the higher cumulants, and we have the following result.

Result 5. Suppose that $\log M_{t+1} = \alpha_t - \gamma_t \log R_{t+1}$, where $\gamma_t > 1$. Then

[4] See Backus, Chernov, and Martin (2011) and Martin (forthcoming) for an extended discussion of cumulants.

$$\left(\frac{M_{t+1}}{\mathbb{E}_t M_{t+1}}\right)^{1-(1/\gamma_t)} \leq M_{t+1} R_{t+1} \leq \left(\frac{R_{t+1}}{\mathbb{E}_t R_{t+1}}\right)^{1-\gamma_t},$$

so $M_{t+1} R_{t+1}$ can explode only if $R_{t+1}/\mathbb{E}_t R_{t+1}$ is small. That is, explosions in X_t correspond to bad news. In contrast, if $M_{t+1} R_{t+1}$ is small, then $M_{t+1}/\mathbb{E}_t M_{t+1}$ must also be small.

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 so that we are in the generic case. Result 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. This extends the traditional results of Latané (1959), Breiman (1960), and Markowitz (1976) to the non–independent and identically distributed (i.i.d.) case. Such an extension has already been provided by Algoet and Cover (1988); however, none of these authors emphasize the corollary—in the spirit of Samuelson (1971)—that since

$$\mathbb{E} \max \left[R_1 \cdots R_t/(R_1^* \cdots R_t^*)\right] = \infty,$$

the growth-optimal portfolio can hugely underperform in the short run.

Of greater interest, though, we have seen that if markets are incomplete (so that the SDF is not unique), this traditional result can be extended to SDFs $M_t \neq 1/R_t^*$. This is interesting because it is often desirable to work with SDFs that are more easily interpretable than $1/R_t^*$, such as SDFs proportional to the marginal value of wealth.

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 strictly 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,

$$\beta^t \frac{u'(C_t)}{u'(C_0)} R_1 \cdots R_t \to 0 \tag{2}$$

and yet

$$\mathbb{E} \max \left[\beta^t \frac{u'(C_t)}{u'(C_0)} R_1 \cdots R_t\right] = \infty. \tag{3}$$

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 $M_t R_{f,t} \nrightarrow 1$. Suppose that this is so and that the riskless rate is constant, $R_{f,t} = R_f$. Furthermore, suppose that the investor is sufficiently patient that $\beta R_f \geq 1$. Then (2) 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.[5] This is a result of Chamberlain and Wilson (2000)—see also the textbook treatment of Ljungqvist and Sargent (2004)—but here the result emerges as a special case of the more general result 1. Moreover, the link between almost-sure convergence to zero and explosions in $\beta^t[u'(C_t)/u'(C_0)]R_1\cdots R_t$ appears to be new. Conversely, if the investor is impatient, $\beta R_f \leq 1$, then we can conclude from (3) that $\mathbb{E}[\max u'(C_t)] = \infty$; equivalently, $\mathbb{E}[u'(\min C_t)] = \infty$.

II. Conclusion

This paper bears on a collection of superficially unrelated observations that can be summarized loosely as follows.

i. Latané, Breiman, Markowitz, Algoet and Cover, and others: The growth-optimal portfolio outperforms any non-growth-optimal portfolio by an arbitrary amount in the long run.
ii. Chamberlain and Wilson: The consumption of patient utility-maximizing investors tends to infinity in the long run.
iii. Weitzman: Tail events exert an extraordinary influence on long-run interest rates (under particular assumptions about the driving stochastic process and agents' utility functions).
iv. Buffett: The Black-Scholes model makes apparently strange predictions for long-dated put option prices.

The results of the paper unify and extend these observations. The growth-optimal portfolio result i—reformulated as the equivalent statement that $X_t \to 0$ for a particular SDF, namely, the reciprocal of the growth-optimal return—is extended to generic SDFs; result ii is an immediate corollary. Observations iii and iv are the flip side of i and ii: tail events exert an extraordinary influence on long-run discount rates for hedge assets.

My results support Weitzman's basic position without requiring the strong assumptions to which Nordhaus (2011) objects. They also show that the phenomenon Buffett discusses is less puzzling than it might seem and, in particular, is not a quirk of the Black-Scholes model.

The starting point of the paper is that despite the fact that the absence of arbitrage implies that expected risk-adjusted returns on all assets equal one at all horizons, realized risk-adjusted returns tend to zero unless (i) the asset in question is asymptotically growth optimal and (ii) the SDF is asymptotically the reciprocal of the growth-optimal return. The key ingredient of this result is a theorem of Kakutani (1948). For economic applications, the flip side of this result is more interesting:

[5] In fact, applying (2) to the growth-optimal asset, we have the stronger result that $\beta^t R_1^* \cdots R_t^* u'(C_t) \to 0$.

realized risk-adjusted returns occasionally explode. Such explosions in risk-adjusted returns must be due to one of two sources that have sharply differing implications: either asset-specific good news (large R) or aggregate disaster (large M).

Aggregate disasters are the relevant consideration for hedge assets such as put options, riskless indexed bonds, or climate change–mitigation projects. (I also argue that bad news is the relevant consideration when valuing a long-dated claim on the market, as a result of the empirical fact that the market's Sharpe ratio is higher than its volatility.) It follows that cost-benefit analyses of long-dated assets, such as the payoffs to environmental projects, should pay special attention to worst-case scenarios to avoid underestimating the value of such projects. There is a stark contrast between the typical sample paths visible in figure 1*B* and the rare, extreme sample paths, visible in figure 1*A*, that are the dominant influence on long-run pricing.

Appendix A

The proof of result 1 divides into two steps. The first is a straightforward adaptation of a result of Kakutani (1948).[6] The goal is to establish that if $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} = \infty$, then $X_\infty = 0$, whereas if $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} < K$, for some constant K, then $\mathbb{E}X_\infty = 1$. Since $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t}$ is a key diagnostic, I refer to it as the *variance criterion*. The second step shows that if $M_t R_t \nrightarrow 1$, then $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} = \infty$.

Proof of Result 1

Step 1: 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 $Y_t = \sqrt{X_t}/(a_1 \cdots a_t)$, and note that Y_t is a martingale.

First, suppose that $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} = \infty$; equivalently, $\sum (1 - a_t^2) = \infty$. It follows, by a standard result—see, for example, theorem 15.5 of Rudin (1987, 300)—that $\prod a_t^2 = 0$, and hence $\prod a_t = 0$. (Conversely, if $\sum [1 - a_t^2] < K$ 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 $X_\infty = 0$. We must also have $\mathbb{E}\max X_t = \infty$: if not, X_t would be uniformly integrable and we would have $\mathbb{E}X_\infty = 1$.

Alternatively, suppose that $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} < K$ for some constant $K < \infty$; equivalently, $\sum (1 - a_t^2) < K$. 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, $\mathbb{E}(\max_t X_t) \leq \mathbb{E}(\max_t Y_t^2) \leq 4\max_t \mathbb{E}(Y_t^2) < \infty$, the second inequality being the $\mathcal{L}^2$

[6] The textbook of Williams (1995) provides a particularly clear treatment of this result.

inequality of Doob (1953, 317). Since $\max_t X_t$ dominates X_t, it follows that X_t is uniformly integrable, so $\mathbb{E}X_\infty = 1$.

Step 2: I will prove the contrapositive: if $\sum \mathrm{Var}_{t-1} \sqrt{M_t R_t} < \infty$, then $M_t R_t \to 1$. Suppose, then, that $\sum \mathrm{Var}_{t-1} \sqrt{M_t R_t} < \infty$. The conditional version of Chebyshev's inequality implies that $\mathbb{P}_{t-1}(|\sqrt{M_t R_t} - \mathbb{E}_{t-1}\sqrt{M_t R_t}| \geq \varepsilon) \leq (1/\varepsilon^2)\,\mathrm{Var}_{t-1}\sqrt{M_t R_t}$ for arbitrary $\varepsilon > 0$, so

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

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

$$\sqrt{M_t R_t} - \mathbb{E}_{t-1}\sqrt{M_t R_t} \to 0. \tag{A1}$$

Furthermore, if $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} < \infty$, 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

$$\mathbb{E}_{t-1}\sqrt{M_t R_t} \to 1. \tag{A2}$$

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 $(\mathbb{E}_{t-1}\sqrt{M_t R_t})^2 < 1 - 2\delta + \delta^2 < 1 - \delta$. But then $\mathrm{Var}_{t-1}\sqrt{M_t R_t} > \delta$ for infinitely many t, which contradicts the assumption that $\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} < \infty$.

It follows from (A1) and (A2) that $\sqrt{M_t R_t} \to 1$, and hence $M_t R_t \to 1$.

For the second part of the result, observe that since $f(x) \equiv (x \log x)_+$ is convex,[7] $(X_t \log X_t)_+$ is a submartingale by Jensen's inequality, and so

$$\max \mathbb{E}(X_t \log X_t)_+ = \lim_{t\to\infty} \mathbb{E}(X_t \log X_t)_+.$$

But then, by results IV-2-10 and IV-2-11 of Neveu (1975), the second part of the result holds with $\mathbb{E}[(X_t \log X_t)_+]$ replacing $\mathbb{E}[X_t \log(1 + X_t)]$. It remains to check that $\lim \mathbb{E}[X_t \log(1 + X_t)]$ is infinite iff $\lim \mathbb{E}[(X_t \log X_t)_+]$ is infinite. This follows from two facts: (i) if $X_t \geq 1$, then

$$X_t \log X_t \leq (1 + X_t)\log(1 + X_t) \leq 2X_t \log(2X_t);$$

and (ii) $\mathbb{E}\log(1 + X_t) \leq \mathbb{E}X_t = 1$, since $\log(1 + x) \leq x$. QED

Proof of Result 2

By Doob's (1953, 314) submartingale inequality, $N\mathbb{P}(\max_{t\leq T} X_t \geq N) \leq \mathbb{E}X_T = 1$, so $\mathbb{P}(\max_{t\leq T} X_t \geq N) \leq 1/N$. The first statement follows by the monotone convergence theorem, because $\mathbf{1}[\max_{t\leq T} X_t \geq N] \uparrow \mathbf{1}[\max_t X_t \geq N]$ as $T \uparrow \infty$.

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

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

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

But this would contradict result 1. QED

Proof of Result 4

We have $M_t R_t = e^{-(\lambda_{t-1}-\sigma_{t-1})Z_t - (\lambda_{t-1}-\sigma_{t-1})^2/2}$, so

$$\sum \mathrm{Var}_{t-1}\sqrt{M_t R_t} = \sum\,[1 - e^{-(\lambda_{t-1}-\sigma_{t-1})^2/4}].$$

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} 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 result 1, we 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. QED

Proof of Result 5

The assumption about the form taken by the SDF implies that

$$M_{t+1}R_{t+1} = e^{\alpha_t}R_{t+1}^{1-\gamma_t} = e^{\alpha_t/\gamma_t}M_{t+1}^{1-(1/\gamma_t)}.$$

Dividing through by conditional means, we have

$$M_{t+1}R_{t+1} = \frac{R_{t+1}^{1-\gamma_t}}{\mathbb{E}_t R_{t+1}^{1-\gamma_t}} = \frac{M_{t+1}^{1-(1/\gamma_t)}}{\mathbb{E}_t M_{t+1}^{1-(1/\gamma_t)}}.$$

Now $1-\gamma_t < 0$ and $1 - 1/\gamma_t \in (0,\ 1)$, so by Jensen's inequality $\mathbb{E}_t R_{t+1}^{1-\gamma_t} \geq (\mathbb{E}_t R_{t+1})^{1-\gamma_t}$ and $\mathbb{E}_t M_{t+1}^{1-(1/\gamma_t)} \leq (\mathbb{E}_t M_{t+1})^{1-(1/\gamma_t)}$. The result follows. QED

References

Algoet, Paul H., and Thomas M. Cover. 1988. "Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment." *Ann. Probability* 16 (April): 876–98.

Backus, David K., Mikhail Chernov, and Ian W. R. Martin. 2011. "Disasters Implied by Equity Index Options." *J. Finance* 66 (December): 1969–2012.

Barro, Robert J. 2006. "Rare Disasters and Asset Markets in the Twentieth Century." *Q.J.E.* 121 (August): 823–66.

Breiman, Leo. 1960. "Investment Policies for Expanding Businesses Optimal in a Long-Run Sense." *Naval Res. Logistics Q.* 7:647–51.

Chamberlain, Gary, and Charles A. Wilson. 2000. "Optimal Intertemporal Consumption under Uncertainty." *Rev. Econ. Dynamics* 3:365–95.

Doob, Joseph L. 1953. *Stochastic Processes.* New York: Wiley.

Gollier, Christian. 2002. "Discounting an Uncertain Future." *J. Public Econ.* 85: 149–66.

———. 2008. "Discounting with Fat-Tailed Economic Growth." *J. Risk and Uncertainty* 37:171–86.

Hansen, Lars Peter. Forthcoming. "Dynamic Valuation Decomposition within Stochastic Economies." *Econometrica.*

Hansen, Lars Peter, and Ravi Jagannathan. 1991. "Implications of Security Market Data for Models of Dynamic Economies." *J.P.E.* 99 (April): 225–62.

Hansen, Lars Peter, and José A. Scheinkman. 2009. "Long-Term Risk: An Operator Approach." *Econometrica* 77 (January): 177–234.

Kakutani, Shizuo. 1948. "On Equivalence of Infinite Product Measures." *Ann. Math.* 49 (January): 214–24.

Latané, Henry A. 1959. "Criteria for Choice among Risky Ventures." *J.P.E.* 67 (April): 144–55.

Ljungqvist, Lars, and Thomas J. Sargent. 2004. *Recursive Macroeconomic Theory.* 2nd ed. Cambridge, MA: MIT Press.

Long, John B. 1990. "The Numeraire Portfolio." *J. Financial Econ.* 26:29–69.

Markowitz, Harry M. 1976. "Investment for the Long Run: New Evidence for an Old Rule." *J. Finance* 31 (December): 1273–86.

Martin, Ian W. R. Forthcoming. "Consumption-Based Asset Pricing with Higher Cumulants." *Rev. Econ. Studies.*

Neveu, Jacques. 1975. *Discrete-Parameter Martingales.* Amsterdam: North-Holland.

Nordhaus, William D. 2011. "The Economics of Tail Events with an Application to Climate Change." *Rev. Environmental Econ. and Policy* 5 (Summer): 240–57.

Rietz, Thomas A. 1988. "The Equity Premium: A Solution." *J. Monetary Econ.* 22: 117–31.

Rudin, Walter. 1987. *Real and Complex Analysis.* 3rd ed. New York: McGraw-Hill.

Samuelson, Paul A. 1971. "The 'Fallacy' of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling." *Proc. Nat. Acad. Sci. USA* 68 (October): 2493–96.

Weitzman, Martin L. 1998. "Why the Far-Distant Future Should Be Discounted at Its Lowest Possible Rate." *J. Environmental Econ. and Management* 36:201–8.

———. 2007. "Subjective Expectations and Asset-Return Puzzles." *A.E.R.* 97 (September): 1102–30.

———. 2009. "On Modeling and Interpreting the Economics of Catastrophic Climate Change." *Rev. Econ. and Statis.* 91 (February): 1–19.

Williams, David. 1995. *Probability with Martingales.* Cambridge: Cambridge Univ. Press.