Disasters implied by equity index options*

David Backus,[†] Mikhail Chernov,[‡] and Ian Martin[§]

August 11, 2010

Abstract

We use equity index options to quantify the probability and magnitude of disasters: extreme negative realizations of consumption growth and stock returns. We show that option prices imply smaller probabilities of these extreme outcomes than have been estimated from international macroeconomic data. A useful byproduct is a novel characterization of departures from lognormality in asset pricing models based on high-order cumulants: skewness, excess kurtosis, and so on.

JEL Classification Codes: E44, G12.

Keywords: equity premium, pricing kernel, entropy, cumulants, risk-neutral probabilities, implied volatility.

---

*We thank the many people who have given us advice, including David Bates, Michael Brandt, Rodrigo Guimaraes, Sydney Ludvigson, Monika Piazzesi, Thomas Sargent, Romeo Tedongap, Michael Woodford, and Liuren Wu, as well as participants in seminars at, and conferences sponsored by, the AEA, the Bank of England, Bocconi, the CEPR, the Federal Reserve, Glasgow, Harvard, IDEI/SCOR (conference on extremal events), LBS, Leicester, LSE, Minnesota, MIT, Moscow HSE, NBER (summer institute), NYU, Penn, Princeton, RCEA Rimini (workshop on money and finance), SoFiE (conference on extreme events), SIFR, Toulouse (conference on financial econometrics), and UC Davis. We also thank Campbell Harvey, an associate editor, and a referee for helpful comments on an earlier version.

[†] Stern School of Business, New York University, and NBER; dbackus@stern.nyu.edu.

[‡] London Business School, London School of Economics, and CEPR; M.Chernov@lse.ac.uk.

[§] Graduate School of Business, Stanford University, and NBER; ian.martin@gsb.stanford.edu.

1 Introduction

Barro (2006), Longstaff and Piazzesi (2004), and Rietz (1988) show that macroeconomic disasters — infrequent large declines in aggregate output and consumption — produce dramatic improvement in the ability of representative agent models to reproduce prominent features of US asset returns, including the equity premium. The primary challenge for disaster research lies in estimating their probability and magnitude. Rietz (1988) simply argues that they are plausible. Longstaff and Piazzesi (2004) argue that disasters based on US experience can explain only about one-half of the equity premium. Barro (2006), Barro and Ursua (2008), and Barro, Nakamura, Steinsson, and Ursua (2009) study broader collections of countries, which in principle can tell us about alternative histories the US might have experienced. They show that these histories include occasional drops in output and consumption that are significantly larger than we have seen in US history.

This range of opinion reflects the nature of the problem. Since disasters are rare, it is difficult to estimate their distribution reliably from the relatively short history of the US economy. We follow a complementary path, using equity index options to infer the distribution of consumption growth, including extreme events like the disasters apparent in macroeconomic data. Equity index options are a useful source of additional information here, because their prices tell us how market participants value extreme events, whether they happen in our sample or not. We use the option-pricing model estimated by Broadie, Chernov, and Johannes (2007) to generate an independent estimate of the distribution of disasters in US business cycles. We find that option prices and equity returns imply smaller probabilities of extreme events (more than 5 standard deviations to the left of the mean) than suggested by international macroeconomic evidence.

The idea is straightforward, but the approaches taken in the macro-finance and option-pricing literatures are different enough that it takes some work to put them on a comparable basis. We follow a somewhat unusual path because we think it leads, in the end, to a more direct and transparent assessment of the impact of disasters on asset returns. We start with a pricing kernel, because every asset pricing model has one. We ask, specifically, whether pricing kernels generated by representative agent models with disasters are similar to those implied by option pricing models.

The question is how to measure the impact of disasters on pricing kernels. We find two statistical concepts helpful here: entropy (a measure of dispersion) and cumulants (close relatives of moments). Alvarez and Jermann (2005) and Bansal and Lehmann (1997) show that mean excess returns, defined as differences of logs of gross returns, place a lower bound on the entropy of the pricing kernel. If the log of the pricing kernel is normal, then entropy is proportional to its variance. Departures from lognormality, and disasters in particular, can increase entropy and thereby improve a model's ability to account for observed excess returns. Such departures contribute to entropy by introducing skewness, excess kurtosis, and so on. This line of thought leads to a quantitative assessment of the importance of nonnormal components, including disasters, to asset pricing in general.

In Section 2 we introduce the required tools and take a preliminary look at the evidence. We define entropy and show that it can be divided into components reflecting the variance of the log pricing kernel (the lognormal term) and to odd and even high-order cumulants (skewness and excess kurtosis, for example). We also relate the pricing kernel to risk-neutral probabilities, which are commonly used in option pricing models, and establish facts that are used later to guide quantitative assessments of disasters.

In Section 3 we follow the macro-finance approach to disasters: log consumption growth includes a nonnormal component and power utility converts consumption growth into a pricing kernel. We show how infrequent large drops in consumption growth generate positive skewness in the log pricing kernel and increase its entropy. If we quantify disasters using the international macroeconomic evidence summarized by Barro and his coauthors, the impact is large, even with moderate risk aversion. It's important that the departures from lognormality have this form: adding large positive changes to consumption growth reduces entropy relative to the lognormal case.

Do option prices indicate a similar contribution from large adverse events? The answer is no, but the language and modelling approach are quite different. Option pricing models typically express asset prices in terms of risk-neutral probabilities rather than pricing kernels. This is more than language; it governs the choice of model. Where macro-finance models generally start with the true probability distribution of consumption growth and use preferences to deduce the risk-neutral distribution, option pricing models infer both from asset prices. The result is a significantly different functional form for the pricing kernel. In Section 4 we show how power utility transforms parameters of true distributions into parameters of risk-neutral distributions. In Section 5 we derive the pricing kernel implied by the true and risk-neutral distributions of equity returns estimated by Broadie, Chernov, and Johannes (2007). Roughly speaking, the true distribution is estimated from equity returns and the risk-neutral distribution from option prices. We use their estimated parameters to quantify the contributions of high-order cumulants to the entropy of the pricing kernel. While both consumption- and option-based models generate substantial contributions to entropy from odd high-order cumulants, the relative contribution is much smaller in the model based on option prices. In this sense and others developed later, option prices suggest a smaller role for disasters than the international macroeconomic evidence. Options also imply greater entropy. Evidently the market places a large premium on whatever risk is involved in selling options.

In Section 6 we explore the differences between models based on consumption and option prices by comparing their pricing kernels and by looking at each from the perspective of the other. If we consider a consumption-based disaster model, how do its option prices compare to those generated by an estimated option model? And if we infer consumption growth from option prices, how does it compare to the consumption process estimated from macroeconomic data? Both of these comparisons suggest that option prices imply smaller probabilities of extreme adverse events than we see in international macroeconomic data.

In Section 7 we consider several extensions of our theoretical framework. Our analysis to this point rests on some or all of these three assumptions: consumption growth and equity returns are independent over time, the representative agent has power utility, and equity returns are tightly linked to consumption growth. We explore each in greater depth.

We conclude with a summary and a brief discussion of the value of cumulants and associated tools in finance.

2 Preliminaries

We start with an overview of the tools and evidence used later on. The tools allow us to characterize departures from lognormality, including disasters, in a convenient way. Once these tools are developed, we describe some of the evidence they'll be used to explain.

2.1 Pricing kernels, entropy, and cumulants

One way to express modern asset pricing is with a pricing kernel. In any arbitrage-free environment, there is a positive random variable m that satisfies the pricing relation,

$$E_t \left(m_{t+1} r^j_{t+1} \right) \;\; = \;\; 1, \tag{1}$$

for (gross) returns r^j on all traded assets j. Here E_t denotes the expectation conditional on information available at date t. In stationary ergodic settings, the same relation holds unconditionally as well; that is, with an expectation E based on the ergodic distribution. In finance, the pricing kernel is often a statistical construct designed to account for returns on assets of interest. In macroeconomics, the pricing kernel is tied to macroeconomic quantities such as consumption growth. In this respect, the pricing kernel is a link between macroeconomics and finance.

Asset returns alone tell us some of the properties of the pricing kernel, hence indirectly about macroeconomic fundamentals. A notable example is the Hansen-Jagannathan (1991) bound. We use a similar "entropy bound" derived by Alvarez and Jermann (2005) and Bansal and Lehman (1997). Both bounds relate measures of pricing kernel dispersion to expected differences in returns. With this purpose in mind, we define the entropy of a positive random variable x as

$$L(x) \;\; = \;\; \log Ex - E \log x. \tag{2}$$

We account for this use of the term shortly. Entropy has a number of properties that we use repeatedly. First, entropy is nonnegative and equal to zero only if x is constant (Jensen's inequality). In the familiar lognormal case, where $\log x \sim \mathcal{N}(\kappa_1, \kappa_2)$, entropy is $L(x) = \kappa_2/2$ (one-half the variance of $\log x$). We'll see shortly that $L(x)$ also depends on features of the distribution beyond the first two moments. Second, $L(ax) = L(x)$ for any positive constant a. Third, if x and y are independent, then $L(xy) = L(x) + L(y)$.

The entropy bound relates the entropy of the pricing kernel to expected differences in log returns:

$$L(m) \quad \geq \quad E\left(\log r^j - \log r^1\right) \tag{3}$$

for any asset j with positive returns. See Appendix A.1. Here r^1 is the (gross) return on a one-period risk-free bond, so the right-hand side is the mean excess return or premium on asset j over the short rate. Inequality (3) therefore transforms estimates of return premiums into estimates of the lower bound of the entropy of the pricing kernel.

The beauty of entropy as a dispersion concept for the study of disasters is that it includes a role for the departures from normality they tend to generate. Recall that the moment generating function (if it exists) for a random variable x is defined by

$$h(s; x) \quad = \quad E\left(e^{sx}\right),$$

a function of the real variable s. With enough regularity, the cumulant-generating function, $k(s) = \log h(s)$, has the power series expansion

$$k(s; x) \quad = \quad \log E\left(e^{sx}\right) \quad = \quad \sum_{j=1}^{\infty} \kappa_j(x) s^j / j! \tag{4}$$

for some suitable range of s. This is a Taylor (Maclaurin) series representation of $k(s)$ around $s = 0$ in which the "cumulant" κ_j is the jth derivative of k at $s = 0$. Cumulants are closely related to moments: κ_1 is the mean, κ_2 is the variance, and so on. Skewness γ_1 and excess kurtosis γ_2 are scaled versions of the third and fourth cumulants:

$$\gamma_1 \quad = \quad \kappa_3 / \kappa_2^{3/2}, \quad \gamma_2 \quad = \quad \kappa_4 / \kappa_2^2. \tag{5}$$

The normal distribution has a quadratic cumulant-generating function, which implies zero cumulants after the first two. Nonzero high-order cumulants (κ_j for $j \geq 3$) therefore summarize departures from normality. Note for future reference that $k(s; ax) = k(as; x)$ [replace s with as in (4)]. Therefore if x has cumulants κ_j, ax has cumulants $a^j \kappa_j$.

With this machinery in hand, we can express the entropy of the pricing kernel in terms of the cumulant-generating function and cumulants of $\log m$:

$$\begin{aligned} L(m) \quad &= \quad \log E\left(e^{\log m}\right) - E \log m \\ &= \quad k(1; \log m) - \kappa_1(\log m) \quad = \quad \sum_{j=2}^{\infty} \kappa_j(\log m)/j!. \end{aligned} \tag{6}$$

This use of the cumulant-generating function was suggested by Martin (2009) and recurs throughout the paper. If $\log m$ is normal, entropy is one-half the variance ($\kappa_2/2$), but in general there will be contributions from skewness ($\kappa_3/3!$), excess kurtosis ($\kappa_4/4!$), and so on.

Zin (2002, Section 2) points out that we can use high-order cumulants to account for properties of returns that are difficult to explain in lognormal settings. We implement his insight with a three-way decomposition of entropy: one-half the variance (the lognormal term) and contributions from odd and even high-order cumulants. Although disasters typically show up in both odd and even high-order cumulants, odd cumulants reflect their inherent asymmetry. More generally, the contribution of odd high-order cumulants represents an adaptation and extension of work on skewness preference by Harvey and Siddique (2000) and Kraus and Litzenberger (1976): adaptation because it refers to properties of the log of the pricing kernel rather than its level, and extension because it involves all odd high-order cumulants, not just skewness. We compute odd and even cumulants from the odd and even components of the cumulant-generating function:

$$\begin{aligned} k_{\text{odd}}(s) &= [k(s) - k(-s)]/2 = \sum_{j=1,3,\ldots} \kappa_j(x) s^j / j! \\ k_{\text{even}}(s) &= [k(s) + k(-s)]/2 = \sum_{j=2,4,\ldots} \kappa_j(x) s^j / j!. \end{aligned}$$

Odd and even high-order cumulants follow from subtracting the first and second cumulants, respectively.

2.2 Risk-neutral probabilities

In option pricing models, there is rarely any mention of a pricing kernel, although theory tells us one must exist. Option pricers speak instead of true and risk-neutral probabilities. We use a finite-state iid (independent and identically distributed) setting to show how pricing kernels and risk-neutral probabilities are related.

Consider an iid environment with a finite number of states x that occur with (true) probabilities $p(x)$, positive numbers that represent the frequencies with which different states occur (the data generating process, in other words). With this notation, the pricing relation (1) becomes

$$E\left(mr^j\right) = \sum_x p(x) m(x) r^j(x) = 1$$

for (gross) returns r^j on all assets j. A particularly simple example is a one-period bond, whose price is $q^1 = Em = \sum_x p(x)m(x) = 1/r^1$. Risk-neutral (or better, risk-adjusted) probabilities are

$$p^*(x) = p(x)m(x)/Em = p(x)m(x)/q^1. \tag{7}$$

The p^*s are probabilities in the sense that they are positive and sum to one, but they are not the data generating process. The role of q^1 is to make sure they sum to one. They lead to another version of the pricing relation,

$$q^1 \sum_x p^*(x) r^j(x) \;\;=\;\; q^1 E^* r^j \;\;=\;\; 1, \tag{8}$$

where E^* denotes the expectation computed from risk-neutral probabilities. In (1), the pricing kernel performs two roles: discounting and risk adjustment. In (8) those roles are divided between q^1 and p^*, respectively.

Option pricing is a natural application of this approach. Consider a put option: the option to sell an arbitrary asset with future price $q(x)$ at strike price b. Puts are bets on bad events — the purchaser sells prices below the strike, the seller buys them — so their prices are an indication of how they are valued by the market. If the option's price is q^p (p for put), its return is $r^p(x) = [b - q(x)]^+/q^p$ where $(b-q)^+ \equiv \max\{0, b-q\}$. Equation (8) gives us its price in terms of risk-neutral probabilities:

$$q^p \;\;=\;\; q^1 E^*(b-q)^+.$$

This highlights the role of risk-neutral probabilities in option pricing: As we vary b, we trace out the risk-neutral distribution of prices $q(x)$ (Breeden and Litzenberger, 1978).

But what about the pricing kernel and its entropy? Equation (7) gives us the pricing kernel:

$$m(x) \;\;=\;\; q^1 p^*(x)/p(x). \tag{9}$$

Since q^1 is constant in our iid world, the entropy of the pricing kernel is

$$L(m) \;\;=\;\; L(p^*/p) \;\;=\;\; \log E(p^*/p) - E\log(p^*/p) \;\;=\;\; -E\log(p^*/p). \tag{10}$$

The first equality follows because q^1 is constant [recall $L(ax) = L(x)$]. The second follows from the definition of entropy [equation (2)]. The last one follows from

$$E(p^*/p) \;\;=\;\; \sum_x [p^*(x)/p(x)]p(x) \;\;=\;\; \sum_x p^*(x) \;\;=\;\; 1.$$

The expression on the right of (10) is sometimes referred to as the entropy of p^* relative to p, which accounts for our earlier use of the term.

As before, entropy can be expressed in terms of cumulants. The cumulants in this case are those of $\log(p^*/p)$, whose cumulant-generating function is

$$k[s;\log(p^*/p)] \;\;=\;\; \log E\left(e^{s\log(p^*/p)}\right) \;\;=\;\; \sum_{j=1}^{\infty} \kappa_j[\log(p^*/p)] s^j/j!. \tag{11}$$

The definition of entropy (2) contributes the analog to (6) in which entropy is related to cumulants:

$$\begin{aligned} L(p^*/p) &= k[1; \log(p^*/p)] - \kappa_1[\log(p^*/p)] \\ &= \sum_{j=2}^{\infty} \kappa_j[\log(p^*/p)]/j! \;\; = \;\; -\kappa_1[\log(p^*/p)]. \end{aligned} \tag{12}$$

The second line follows from $k[1; \log(p^*/p)] = \log E(p^*/p) = 0$ (see above). Here we can compute entropy from the first cumulant, but it's matched by an expansion in terms of cumulants 2 and above, just as it was in the analogous expression for $\log m$. All of these cumulants are readily computed from derivatives of the cumulant-generating function (11).

To summarize: we can price assets using either a pricing kernel (m) and true probabilities (p) or the price of a one-period bond (q^1) and risk-neutral probabilities (p^*). The three objects (m, p^*, p) are interconnected: once we know two (and the price of a one-period bond), equation (7) gives us the other. That leaves us with three kinds of cumulants corresponding, respectively, to the true distribution of the random variable x, the risk-neutral distribution, and the true distribution of the log of the pricing kernel. We report all three.

2.3 Evidence

We will put these tools to work in linking broad features of macroeconomic and financial data: consumption growth, asset returns, and option prices. Here we provide a quick overview of US evidence on each.

In Table 1 we report evidence on annual consumption growth and equity returns (the S&P 500 index) for both a long sample (1889-2009) and a shorter one (1986-2009) that corresponds approximately to the option data used by Broadie, Chernov, and Johannes (2007). Similar evidence is summarized by Alvarez and Jermann (2005, Tables I-III), Barro (2006, Table IV), and Mehra and Prescott (1985, Table 1). In both samples, consumption growth and equity returns exhibit the negative skewness we would expect from occasional disasters. Our estimates of the equity premium (0.0407 in the long sample, 0.0434 in the short sample) are somewhat smaller than those reported elsewhere. One reason is that we measure returns in logs; in levels, the mean excess return on equity is 0.0571 in the long sample and 0.0613 in the short sample. Another reason is that the 2008 return (-0.38 in levels) has a significant impact on the estimated mean, particularly in the short sample.

The next issue is option prices. Options are available on the S&P 500 index and on its futures contracts. Prices are commonly quoted as implied volatilities: the value of the volatility parameter that equates the price with the Black-Scholes-Merton formula. These volatilities have two well-documented features that we examine more closely in Section 5. Similar evidence has been reviewed recently by Bates (2008, Section 1), Drechsler and Yaron

(2008, Section 2), and Wu (2006, Section II). The first feature is that implied volatilities are greater than sample standard deviations of returns. Since prices are increasing in volatility, this implies that options are expensive relative to the lognormal benchmark that underlies Black-Scholes-Merton. As a result, selling options generates high average returns. The second feature is that implied volatilities are higher for lower strike prices: the well-known volatility skew. This feature is intriguing from a disaster perspective, because it suggests market participants value adverse events more than would be implied by a lognormal model. The question for us whether the extra value assigned to bad outcomes corresponds to the disasters documented in macroeconomic research.

In the following sections we use round-number versions of the estimates in Table 1 to illustrate the quantitative importance of disasters. We report the properties of numerical examples in which log consumption growth has a mean of 0.0200 (2%) and a standard deviation of 0.0350 (3.5%). Similarly, the log excess return on equity has a mean of 0.0400 and a standard deviation of 0.1800 and the log return on the one-period bond is 0.0200. Most of these numbers are similar across the long and short samples. The exception is the standard deviation of log consumption growth. We use an estimate based on the long sample because it includes the Great Depression, the one clear disaster in this sample. None of these numbers are definitive, but they are close to the values in the table and give us a starting point for considering the quantitative implications of disasters.

3 Disasters in macroeconomic models and data

Representative-agent exchange economies generate larger risk premiums when we include infrequent large declines in consumption growth. We describe the mechanism with two numerical examples that highlight the role of high-order cumulants. Here and in our study of options we restrict our attention to iid environments. There are many features of the world that are not iid, but this simplification allows us to focus without distraction on the distribution of returns, particularly the possibility of extreme negative outcomes. We think it's a reasonably good approximation for this purpose, but return to the issue briefly in Section 7.

The economic environment consists of preferences for a representative agent and a stochastic process for consumption growth. Preferences are governed by an additive power utility function,

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t),$$

with $u(c) = c^{1-\alpha}/(1-\alpha)$ and $\alpha \geq 0$. We refer to α as risk aversion. If consumption growth is $g_t = c_t/c_{t-1}$, the pricing kernel is

$$\log m_{t+1} \quad = \quad \log \beta - \alpha \log g_{t+1}. \tag{13}$$

With power utility, the second derivative is negative (risk aversion), the third positive (skewness preference), and the fourth negative (kurtosis aversion). The properties of the pricing kernel follow from those of consumption growth. Entropy is

$$L(m) \quad = \quad L(e^{-\alpha \log g}) \tag{14}$$

and the cumulants of $\log m$ are related to those of $\log g$ by

$$\kappa_j(\log m) \quad = \quad \kappa_j(\log g)(-\alpha)^j/j!, \quad j \geq 2. \tag{15}$$

See Section 2.1. If log consumption growth is normal, then so is the log of the pricing kernel. Entropy is then one-half the variance of consumption growth times the risk aversion parameter squared. The impact of high-order cumulants depends on $(-\alpha)^j/j!$. The minus sign tells us that negative odd cumulants of log consumption growth generate positive odd cumulants in the log pricing kernel. Negative skewness in consumption growth, for example, generates positive skewness in the pricing kernel and thus increases the entropy of the pricing kernel. The contributions of high-order cumulants are controlled by the coefficient $\alpha^j/j!$. Eventually the denominator grows faster than the numerator, but for moderate values of j, risk aversion can magnify the contributions of high-order cumulants (those with $j \geq 3$) relative to the variance.

We follow Barro (2006) in using a two-component structure for consumption growth,

$$\log g_{t+1} \quad = \quad w_{t+1} + z_{t+1}, \tag{16}$$

with components (w_t, z_t) that are independent of each other and over time. Since the components are independent, the cumulant-generating function of $\log g$ is the sum of those for w and z. Similarly, the entropy of the pricing kernel is the sum of the entropy of the components,

$$L(m) \quad = \quad L(e^{-\alpha w}) + L(e^{-\alpha z}), \tag{17}$$

and the cumulants of $\log m$ are sums of the cumulants of the components,

$$\kappa_j(\log m) \quad = \quad (-\alpha)^j \kappa_j(w) + (-\alpha)^j \kappa_j(z). \;\; j \geq 2.$$

(That's why they call them cumulants: they "[ac]cumulate.")

In the examples that follow, the first component is normal: $w \sim \mathcal{N}(\mu, \sigma^2)$. The second takes two different forms, but we refer to it generically as the jump component. In discrete time, jumps aren't needed to generate nonnormal random variables, but the terminology is convenient. It's important, however, to distinguish between jumps and disasters. Disasters are large negative realizations of consumption growth. Jumps need not be large, yet in the second example they can still generate extreme realizations if they occur frequently enough.

We choose parameters for the examples in this order. First, we choose parameters for the jump component z to mimic the macroeconomic evidence on disasters documented by Barro (2006), Barro and Ursua (2008), and Barro, Nakamura, Steinsson, and Ursua (2009). Second, we choose parameters for the normal component w to match the mean and variance of log consumption growth reported in Section 2.3 given the parameters of the jump component. Finally, we choose risk aversion to match the equity premium. Given these inputs, we compute the entropy of the pricing kernel and describe the impact of the departures from normality.

3.1 Example 1: Bernoulli jump component

The simplest example of a jump component is a Bernoulli random variable:

$$z_t = \begin{cases} 0 & \text{with probability } 1-\omega \\ \theta & \text{with probability } \omega. \end{cases} \tag{18}$$

Here $\omega > 0$ and $\theta < 0$ represent the probability and magnitude of a drop in consumption growth relative to its mean. The entropy of the two components follows from its definition (2):

$$L(e^{-\alpha w}) = (-\alpha\sigma)^2/2 \tag{19}$$

$$L(e^{-\alpha z}) = \log\left(1-\omega+\omega e^{-\alpha\theta}\right) + \alpha\omega\theta. \tag{20}$$

Both are zero at $\alpha = 0$ and increase with α. The first expression is the usual "one-half the variance" of the lognormal case. The second introduces high-order cumulants; see Appendix A.2.

We can see the quantitative significance of the jump component with numerical examples based on international macroeconomic evidence. Its role is evident in Table 2 in the difference between column (1), the lognormal case, and column (2), which incorporates a Bernoulli jump component. In both cases, the mean and variance of log consumption growth are $\kappa_1(\log g) = 0.020$ and $\kappa_2(\log g) = 0.035^2$. In column (1), we set $\mu = \kappa_1(\log g)$ and $\sigma^2 = \kappa_2(\log g)$. In column (2), we set $\omega = 0.01$ and $\theta = -0.3$: a one percent chance of a 30 percent fall in consumption growth relative to its mean. Given these values, we adjust the parameters of the normal component to maintain the mean and variance: $\mu+\omega\theta = \kappa_1(\log g)$ and $\sigma^2 + \omega(1-\omega)\theta^2 = \kappa_2(\log g)$.

The parameters of the jump component are derived from studies of international macroeconomic data by Barro (2006), Barro and Ursua (2007), and Barro, Nakamura, Steinsson, and Ursua (2009). Each of these studies looks at aggregate output or consumption over the last century or more for 20-plus countries. Martin (2009) uses the empirical distribution reported by Barro (2006) to set $\omega = 0.017$ and $\theta = -0.38$. Barro, Nakamura, Steinsson, and Ursua (2009, Section 6.2) estimate a dynamic model, but argue that its asset pricing implications are the same as an iid model with $\omega = 0.0138$ (corresponding to their p) and $\theta = -0.357$ [corresponding to their $\log(1-b)$]. We use more modest values to avoid overstating the role of jumps and to keep the variance of the normal component positive. These numbers nevertheless suggest what may seem to be an excessively large probability of an extremely bad outcome given US history, but that's what the international evidence implies. We return to this issue when we look at the distribution implied by options.

With these numbers, we can explore the ability of the model to satisfy the entropy bound. The observed equity premium implies that the entropy of the pricing kernel is at least 0.0400. In the lognormal case, the entropy bound implies $\alpha^2 \kappa_2(\log g)/2 = \alpha^2 0.0350^2/2 \geq 0.0400$ or $\alpha \geq 8.08$. We can satisfy the entropy bound for the equity premium, but only with a risk aversion parameter greater than 8. There's a range of opinion about this, but some argue that risk aversion this large implies implausible behavior along other dimensions; see, for example, the discussion in Campanale, Castro, Clementi (2010, Section 4.3) and the references cited there.

When we add the jump component, a smaller risk aversion parameter suffices. Since the mean and variance of log consumption growth are the same, the experiment has a partial derivative flavor: it measures the impact of high-order cumulants, holding constant the mean and variance. As long as $\omega < 1/2$ and $\theta < 0$, the jump component z introduces negative skewness and positive excess kurtosis into log consumption growth. Both are evident in the first panel of Figure 1, where we plot cumulants 2 to 8 for log consumption growth. Each cumulant $\kappa_j(\log g)$ makes a contribution $\kappa_j(\log g)(-\alpha)^j/j!$ to the entropy of the pricing kernel. The next two panels of the figure show how the contributions depend on risk aversion. With $\alpha = 2$, negative skewness in consumption growth translates into a positive contribution to entropy, but the contribution of high-order cumulants overall is small relative to the contribution of the variance. That changes dramatically when we increase α. Even small high-order cumulants make significant contributions to entropy if α is large enough.

Figure 2 gives us another perspective on the same issue: the impact of high-order cumulants on the entropy of the pricing kernel as a function of the risk aversion parameter α. The horizontal line is the lower bound, our estimate of the equity premium in US data. The line labelled "lognormal" is entropy without the jump component. We see, as we noted earlier, that the entropy of the pricing kernel for the lognormal case is below the lower bound until α is above 8. The line labelled "disasters" incorporates the Bernoulli jump component. The difference between the two lines shows that the overall contribution of high-order cumulants is positive and increases sharply with risk aversion. When $\alpha = 2$ the extra terms increase entropy by 16%, but when $\alpha = 8$ the increase is over 100%.

It's essential that the jumps be bad outcomes. If we reverse the sign of θ, the result is the line labelled "booms" in Figure 2. We see that for every value of α, entropy is below even the lognormal case. Table 3 shows us exactly how this works. With Bernoulli jumps (and $\alpha = 10$), the entropy of the pricing kernel (0.1614) comes from the variance (0.0613), odd high-order cumulants (0.0621), and even high-order cumulants (0.0380). When we switch to booms, the odd cumulants change sign — see equation (15) — reducing total entropy. Another example illustrates the role of the probability and magnitude of the disaster. Suppose we halve θ and double ω, with σ adjusting to maintain the variance of consumption growth. Then entropy falls sharply and the contribution of high-order cumulants almost disappears. In this sense, the low probability and the large magnitude in the example are quantitatively important.

We've chosen to focus on the entropy of the pricing kernel, but you get a similar picture if you look at the equity premium. We define levered equity as a claim to the dividend

$$d_t \quad = \quad c_t^{\lambda}. \tag{21}$$

This isn't, of course, either equity or levered, but it's a convenient functional form that is widely used in the macro-finance literature to connect consumption growth (the foundation for the pricing kernel) to returns on equity (the asset of interest). See Abel (1999, Section 2.2). In the iid case, the log excess return is a linear function of log consumption growth:

$$\log r^e_{t+1} - \log r^1_{t+1} \quad = \quad \lambda \log g_{t+1} + \text{constant}. \tag{22}$$

See Appendix A.4. This tight connection between equity returns and consumption growth overstates how closely these two variables are related, but it captures in a simple way the obvious cyclical variation in the stock market. We consider alternatives to equation (22) in Section 7, but for now it's a useful simplification. The leverage parameter λ allows us to control the variance of the equity return separately from the variance of consumption growth and thus to match both. We use an excess return variance of 0.1800^2, so λ is the ratio of the standard deviation of the excess return (0.1800) to the standard deviation of log consumption growth (0.0350), approximately 5.1.

Given a pricing kernel, entropy places an upper bound on the expected excess return of any asset. The asset that hits the bound (the "high-return asset") has return $r_{t+1} = 1/m_{t+1}$. Equity is precisely this asset in this environment when $\alpha = \lambda$, but in other cases the equity premium is strictly less than entropy. We see in Figure 3 that the difference is small in our numerical example for values of α between zero and twelve. The formulas used to generate the figure are reported in Appendix A.4. The parameters, including the value of α that matches the equity premium, are reported in Table 2. As we found with the entropy bound, the lognormal model requires greater risk aversion to account for a given equity premium.

3.2 Example 2: Poisson-normal jump component

Our second model uses a more flexible distribution for log consumption growth: a Poisson mixture of normals. The added complexity has a number of benefits. One is that it gives us a better approximation to the empirical distribution of disasters. Another is that it is easily scaled to the different time intervals observed in option markets (it's "infinitely divisible"). For this reason and others, this specification is commonly used in work on option pricing, where it is referred to as the Merton (1976) model. In the macro-finance literature, it has been applied by Bates (1988), Martin (2007), and Naik and Lee (1990).

As above, log consumption growth has normal and jump components. The central ingredient of the jump component is a Poisson random variable j (the number of jumps) that takes on nonnegative integer values with probabilities $e^{-\omega}\omega^j/j!$. The parameter ω ("jump intensity") is the mean of j. Conditional on j, the jump component is normal:

$$z_t|j \quad \sim \quad \mathcal{N}(j\theta, j\delta^2) \text{ for } j = 0, 1, 2. \ldots. \tag{23}$$

This differs from the Bernoulli model in two respects: there is a positive probability of more than one jump and the jump size has a distribution rather than fixed size. If ω is small, the first is insignificant but the second increases entropy and high-order cumulants. If ω is large, as it is in the option model of Section 5, there can be a significant probability of multiple jumps. Given this structure, the entropy of the jump component of the pricing kernel is

$$L(e^{-\alpha z}) \quad = \quad \omega[e^{-\alpha\theta+(\alpha\delta)^2/2} - 1] + \alpha\omega\theta. \tag{24}$$

Total entropy is the sum of the entropies of the normal and jump components, equations (19) and (24). This and other properties of Poisson-normal mixtures are derived in Appendix A.3.

We illustrate the properties of this example with numbers similar to those used in the Bernoulli example. With $\omega = 0.01$, there is probability 0.9900 of no jumps, 0.0099 of one jump, and 0.0001 of more than one jump. The Poisson process, then, is virtually the same in this respect as the Bernoulli process. The only significant change is the dispersion of jumps. The parameters of the jump component are again based on the studies of Barro (2006), Barro and Ursua (2007), and Barro, Nakamura, Steinsson, and Ursua (2009). If we were to choose the mean and standard deviation to match the empirical distribution estimated by Barro (2006), as Martin (2009) and Wachter (2009) do, we would set $\theta = -0.38$ and $\delta = 0.25$. For the same reasons as before, we choose more modest values: $\theta = -0.30$ and $\delta = 0.15$. Given these values, we choose μ and σ to match our target values for the mean and variance of log consumption growth. In the model, the mean is $\mu + \omega\theta$ and the variance is $\sigma^2 + \omega(\theta^2 + \delta^2)$. The resulting parameter values are listed in column (3) of Table 2.

We see in Table 3 that dispersion in the jump distribution generates greater entropy for any given value of risk aversion than we saw in the Bernoulli example. Of the total entropy of 0.0408, 41% comes from the variance and 38% and 21%, respectively, from odd and even high-order cumulants. The strong contribution from odd cumulants is a clear indication of the important role played by disasters. Figure 4 illustrates the impact on specific cumulants. One consequence is that the model satisfies the entropy bound and matches the equity premium with smaller values of α. See Figure 5.

Both examples increase the probability of extreme negative consumption growth relative to the lognormal benchmark. We see in Table 2 that the probability of log consumption growth more than three standard deviations to the left of its mean is 0.13% in the lognormal case [column (1)] but 1% and 0.9%, respectively, in the Bernoulli and Poisson cases [columns

(2) and (3)]. This corresponds to a drop in consumption of more than 8.5%, something seen only once in US history: in 1931, when consumption fell by 9.9%. Thus a 1% event has occurred once in slightly more than a century of US history. In this respect, the examples correspond roughly to US experience.

In other respects, the examples are more extreme than US history, implying larger departures from lognormality that we've observed. The model implies, for example, skewness of log consumption growth of -11.02 [the entry labelled γ_1(true) in Table 2]. In US data, our estimate is a much more modest -0.34 (the entry labelled skewness in Table 1). Excess kurtosis (γ_2) is similar. This is, of course, Barro's (2006) argument: that what we've seen in US data may not accurately reflect the distribution of what might have happened. That leads us to study options, which in principle reflect the distribution used by market participants.

4 Risk-neutral probabilities in representative-agent models

As a warmup for our study of options, we derive the risk-neutral probabilities implied by the examples of the previous section and use them to compute the risk-neutral parameters reported in Table 2. The state spaces have continuous components, but the logic of Section 2.2 follows with integrals replacing sums where appropriate. In representative-agent models, risk aversion generates risk-neutral distributions that are shifted left (more pessimistic) relative to true distributions. The form of this shift depends on the distribution. More generally, we might think of any such shift as representing something like risk aversion.

Our first example has lognormal consumption growth. Suppose $\log g = w$ with $w \sim \mathcal{N}(\mu, \sigma^2)$. Then

$$p(w) \;=\; (2\pi\sigma^2)^{-1/2} \exp[-(w-\mu)^2/2\sigma^2].$$

The pricing kernel is $m(w) = \beta \exp(-\alpha w)$ and the one-period bond price is $q^1 = Em = \beta \exp[-\alpha\mu + (\alpha\sigma)^2/2]$. Equation (7) gives us the risk-neutral probabilities:

$$p^*(w) \;=\; p(w)m(w)/q^1 \;=\; (2\pi\sigma^2)^{-1/2} \exp[-(w-\mu+\alpha\sigma^2)^2/2\sigma^2].$$

Thus the risk-neutral distribution has the same form (normal) with mean $\mu^* = \mu - \alpha\sigma^2$ and standard deviation $\sigma^* = \sigma$. The former shows us that the distribution shifts to the left by an amount proportional to risk aversion α and risk σ^2. The log probability ratio is

$$\log\left[p^*(w)/p(w)\right] \;=\; [(w-\mu)^2 - (w-\mu^*)^2]/2\sigma^2,$$

which implies the cumulant-generating function

$$k[s; \log(p^*/p)] \;=\; \log E\left(e^{s \log p^*/p}\right) \;=\; \frac{(\mu-\mu^*)^2}{2\sigma^2}(-s+s^2).$$

The cumulants are (evidently) zero after the first two. Entropy follows from equation (12),

$$L(p^*/p) \;=\; \frac{(\mu-\mu^*)^2}{2\sigma^2} \;=\; (\alpha\sigma)^2/2,$$

which is what we reported in equation (19).

Our second example has Bernoulli consumption growth. Let $\log g = z$, with z equal to 0 with probability $1-\omega$ and θ with probability ω. If we ignore the discount factor β (we just saw that it drops out when we compute p^*), the pricing kernel is $m(z) = \exp(-\alpha z)$. The one-period bond price is $q^1 = 1-\omega+\omega\exp(-\alpha\theta)$. Risk-neutral probabilities are

$$p^*(z) \;=\; p(z)m(z)/q^1 \;=\; \begin{cases} (1-\omega)/q^1 & \text{if } z=0 \\ \omega\exp(-\alpha\theta)/q^1 & \text{if } z=\theta. \end{cases}$$

Thus p^* is Bernoulli with probability

$$\omega^* \;=\; \omega e^{-\alpha\theta}/(1-\omega+\omega e^{-\alpha\theta})$$

and magnitude $\theta^* = \theta$. Note that p^* puts more weight on the bad state than p. The probability ratio,

$$p^*(z)/p(z) \;=\; \begin{cases} 1/q^1 & \text{if } z=0 \\ \exp(-\alpha\theta)/q^1 & \text{if } z=\theta, \end{cases}$$

implies the cumulant-generating function

$$k[s;\log(p^*/p)] \;=\; \log\left[(1-\omega)+\omega e^{-s\alpha\theta}\right] - s\log\left[(1-\omega)+\omega e^{-\alpha\theta}\right].$$

Entropy is therefore

$$L(p^*/p) \;=\; (1-\omega)\log q^1 + \omega\log(q^1/e^{-\alpha\theta}) \;=\; \log(1-\omega+\omega e^{-\alpha\theta}) + \alpha\omega\theta,$$

which is what we saw in equation (20).

In our final example, consumption growth follows the Poisson-normal mixture described by equation (23). We derive the risk-neutral distribution from the cumulant-generating function (cgf). This approach works with the other examples, too, but it's particularly convenient here. With power utility, the cgf of the risk-neutral distribution is

$$k^*(s) \;=\; k(s-\alpha) - k(-\alpha).$$

See Appendix A.5. Since $k(s) = \omega[\exp(s\theta + (s\delta)^2/2) - 1]$ (Appendix A.3), we have

$$k^*(s) \;=\; \omega e^{-\alpha\theta+(\alpha\delta)^2/2}\left[e^{s(\theta-\alpha\delta^2)+(s\delta)^2/2} - 1\right]$$

This has the same form as $k(s)$ and describes a Poisson-normal mixture with parameters

$$\omega^* \;=\; \omega e^{-\alpha\theta+(\alpha\delta)^2/2}, \;\; \theta^* \;=\; \theta-\alpha\delta^2, \;\; \delta^* \;=\; \delta. \tag{25}$$

Similar expressions are derived by Bates (1988), Martin (2007), and Naik and Lee (1990). Risk aversion ($\alpha > 0$) places more weight on bad outcomes in two ways: they occur more frequently ($\omega^* > \omega$ if $\theta < 0$) and are on average worse ($\theta^* < \theta$). Entropy is the same as equation (24).

Multi-component models combine these ingredients. If log consumption growth is the sum of independent components, then entropy is the sum of the entropies of the components, as in equation (17).

5 Disasters in option models and data

In the macro-finance literature, pricing kernels are typically constructed as in Section 3: we apply a preference ordering (power utility in our case) to an estimated process for consumption growth (lognormal or otherwise). In the option-pricing literature, pricing kernels are constructed from asset prices alone: we estimate true probabilities from time series data on prices or returns, estimate risk-neutral probabilities from the cross-section of option prices, and compute the pricing kernel from the ratio. The approaches are complementary; they generate pricing kernels from different data. The question is whether they lead to similar conclusions. Do options on US equity indexes imply the same kinds of extreme events that Barro and Rietz suggested? Equity index options are a particularly informative class of assets for this purpose, because they tell us not only the market price of equity returns overall, but the prices of specific outcomes.

5.1 The Merton model

We look at option prices through the lens of the Merton (1976) model, a functional form that has been widely used in the empirical literature on option prices. The starting point is a stochastic process for asset prices or returns. Since we're interested in the return on equity, we let

$$\log r^e_{t+1} - \log r^1 \;\; = \;\; w_{t+1} + z_{t+1}. \qquad (26)$$

We use the return, rather than the price, but the logic is the same either way. As before, the components (w_t, z_t) are independent of each other and over time. Market pricing of risk is built into differences between the true and risk-neutral distributions of the components. We give the distributions the same form, but allow them to have different parameters. The first component, w, has true distribution $\mathcal{N}(\mu, \sigma^2)$ and risk-neutral distribution $\mathcal{N}(\mu^*, \sigma^2)$. By convention, σ is the same in both distributions, a byproduct of its continuous-time origins. The second component, z, is a Poisson-normal mixture. The true distribution has jump intensity ω and the jumps are $\mathcal{N}(\theta, \delta^2)$. The risk-neutral distribution has the same form with parameters $(\omega^*, \theta^*, \delta^*)$. The structure and notation will be familiar from Section 3.2.

The Merton model has been widely used in empirical studies of asset pricing, where the parameters of the jump component provide flexibility over the form of departures from normality. It also scales easily to different time intervals, as we show in Appendix A.7. That's helpful here because it allows us to use the model to price options for a range of maturities. The simplest way to describe this is with the cumulant-generating function, which is proportional to the time interval. Entropy and cumulants scale the same way.

Related work supports a return process with these features. Ait-Sahalia, Wang, and Yared (2001) report a discrepancy between the risk-neutral density of S&P 500 index returns implied by the cross-section of options and the time series of the underlying asset returns, but conclude that the discrepancy can be resolved by introducing a jump component. One might go on to argue that two jumps are needed: one for macroeconomic disasters and another for more frequent but less extreme financial crashes. However, Bates (2010) studies the US stock market over the period 1926-2009 and shows that a second jump component plays no role in accounting for macroeconomic events like the Depression.

Given this structure, the pricing kernel follows from equation (9). Its entropy is

$$\begin{aligned} L(m) &= L(p^*/p) \\ &= \frac{(\mu-\mu^*)^2}{2\sigma^2} + (\omega^*-\omega) + \omega\left[\log\frac{\omega}{\omega^*} - \log\frac{\delta}{\delta^*} + \frac{(\theta-\theta^*)^2 + (\delta^2-\delta^{*2})}{2\delta^{*2}}\right]. \end{aligned} \quad (27)$$

This expression and the corresponding cumulant-generating function are derived in Appendix A.8.

5.2 Parameter values

We use parameter values from Broadie, Chernov, and Johannes (2007), who summarize and extend the existing literature on equity index options. Their estimates also include stochastic volatility. We make volatility constant, but we think the simplification is innocuous for our purposes. For one thing, the volatility smile of our iid model is almost the same as the smile generated by the more general model with the volatility state variable set equal to its mean. For another, the smile in the iid model is very close to the average smile in the stochastic volatility model.

The parameters of the true distribution are estimated from the time series of excess returns on equity. We use the parameters of the Poisson-normal mixture — namely (ω, θ, δ) — reported in Broadie, Chernov, and Johannes (2007, Table I, the line labelled SVJ EJP). The estimated jump intensity ω is 1.512, which implies much more frequent jumps than we used in our consumption-based model. With this value, the probability of 0 jumps per year is 0.220, 1 jump per year 0.333, 2 jumps 0.25, 3 jumps 0.13, 4 jumps 0.05, and 7 or more jumps about 0.001. The jumps have mean $\theta = -0.0259$ and standard deviation $\delta = 0.0407$. Given parameters for the Poisson-normal component, the mean μ and standard deviation

σ of the normal component are chosen to match the mean and variance of excess returns to their target values (0.0400 and 0.1800^2, respectively). In the model, the mean excess return (the equity premium) is $\mu + \omega\theta$, which determines μ. The variance is $\sigma^2 + \omega(\theta^2 + \delta^2)$, which determines σ. All of these numbers are reported in column (4) of Table 2.

The risk-neutral parameters for the Poisson-normal mixture are estimated from the cross section of option prices: specifically, prices of options on the S&P 500 over the period 1987-2003. The depth of the market varies both over time and by the range of strike prices and maturities, but there are enough options to allow reasonably precise estimates of the parameters. The numbers we report in Table 2 are from Broadie, Chernov, and Johannes (2007, Table IV, line 5). In practice, option prices identify only the product $\omega^*\theta^*$, so they set $\omega^* = \omega$ and choose θ^* and δ^* to match the level and shape of the implied volatility smile. Given values for $(\omega^*, \theta^*, \delta^*)$, we set μ^* to satisfy (8), which implies $\mu^* + \sigma^2/2 + \omega^*[\exp(\theta^* + \delta^{*2}/2) - 1] = 0$.

Figure 6 shows how the jump mean θ^* and standard deviation δ^* affect the cross section of 3-month option prices. The relevant formulas are reported in Appendices A.6 and A.7. We express prices as implied volatilities and graph them against "moneyness," with higher strike prices to the right. We measure moneyness as the proportional deviation of the strike from the price: (strike $-$ price)/price. A value of zero is therefore equivalent to an at-the-money option (strike = price) or an option on the return at a strike of zero. We use 3-month rather than 1-year options because departures from lognormality (flat volatility smiles) are more obvious at the shorter maturity. In the figure, the solid line represents the implied volatility smile in the model. Since the model fits extremely well, we can take this as a reasonable representation of the data. The downward slope and convex shape are both evidence of departures from lognormality. The second line illustrates the role of the jump mean θ^*: when we divide it by two, the line is flatter. By making the mean jump size smaller, we reduce the value of out-of-the-money puts. The third line illustrates the role of the jump variance δ^{*2}: when we divide it by two, the smile has less curvature. Both lines lie below the estimated one, so the estimated parameters evidently help to account for the observed premium of implied volatilities over the true standard deviation of equity returns (0.1800 in our model).

5.3 Pricing kernel implied by options

We compute the pricing kernel from the ratio of risk-neutral to true probabilities, as in equation (9). It therefore incorporates evidence on the time series of returns as well as option prices. Its properties are reported in Tables 2 and 3 and Figure 7. We compare it with consumption-based models in the next section, but for now simply note its salient features.

The most striking feature of the pricing kernel is its entropy of 0.7747, more than an order of magnitude larger than the equity premium (0.0400) [column (4) of Table 2]. This reflects, in large part, the high price of options. Prices are high in the sense that selling them generates high average returns; see, for example, the extensive literature review in Broadie, Chernov, and Johannes (2009, Appendix A). These high average returns imply high entropy via the entropy bound, even though the model's parameters are chosen to match the equity premium exactly. Evidently a bound based on the equity premium is too loose: other investment strategies generate significantly higher average excess returns and therefore imply higher entropy.

A second feature is the substantial contribution of high-order cumulants. Entropy of 0.7747 includes contributions of 0.4720 (61%) from the variance and 0.1127 (15%) and 0.1900 (25%) from, respectively, odd and even high-order cumulants (Table 3). Like the smile itself, these numbers verify that departures from the lognormal model are quantitatively important.

Figure 7 illustrates the impact of individual cumulants. The top panel shows that high-order cumulants of equity excess returns are small relative to the variance. We know, however, that the model generates nonzero skewness and excess kurtosis (Table 2). Contributions of high-order cumulants to entropy are reported in the second panel. As we noted above, the contributions are small relative to the variance but quantitatively important. When we divide the jump mean θ^* and variance δ^{*2} by two (the third panel), the contributions decline across the board, much as when we reduce risk aversion in consumption-based models.

6 Comparing consumption- and option-based models

We've seen that macroeconomic data and option prices both suggest significant departures from the lognormal model. In this section, we explore their differences from a number of perspectives, finding in every case that the option-based model implies more modest disasters. They are more modest along almost any dimension you might consider, but specifically these three: contributions to entropy of high-order cumulants are a smaller fraction of the total, skewness and excess kurtosis are smaller (a lot smaller) for both true and risk-neutral distributions, and true probabilities of extreme outcomes (5 standard deviations to the left of the mean) are significantly smaller. Some of these properties are based on the risk-neutral distribution, some on the true distribution, and some on both. The message is the same in every case.

Consider a direct comparison of the pricing kernels of consumption- and option-based models [columns (3) and (4) of Table 2]. One difference is total entropy, which is much larger in the option model. Another is the relative contribution of high-order cumulants, which is significantly smaller in the option model. High-order cumulants contribute 59% of the entropy in the consumption model (the row of Table 3 labelled "Poisson consumption growth, $\alpha = 5.19$") but only 39% in the option model (the row labelled "Merton equity returns"). Most relevant to disaster research, the contribution of odd high-order cumulants is 38% in the consumption model but only 15% in the option model. These numbers indicate significant departures from lognormality in both models, but they are relatively smaller in the option model.

A second comparison between the two models involves option prices implied by the consumption model: how do they compare to those implied by the option model? This comparison focuses on the risk-neutral distribution of returns. We compute that for the consumption model in two steps. First, the parameters of the risk-neutral distribution of log consumption growth follow from power utility and the transformations described in Section 4. The results are reported in column (3) Table 2. Second, equation (22) implies that equity excess returns are a scaled version of consumption growth with scale parameter λ. As before, we set $\lambda = 0.1800/0.0350 \cong 5.1$, the ratio of the standard deviations of equity returns and consumption growth. This scaling leads us to replace the parameters $(\sigma^*, \omega^*, \theta^*, \delta^*)$ with $(\lambda\sigma^*, \omega^*, \lambda\theta^*, \lambda\delta^*)$. See Appendix A.7. The result has the same form as the Merton model but different parameters.

Implied volatility smiles for the consumption- and option-based models are pictured in Figure 8. Similar consumption-based option prices are reported by Benzoni, Collin-Dufresne, and Goldstein (2005) and Du (2008). What's new is the explicit comparison to an estimated option pricing benchmark. As before, we use 3-month options to highlight departures from the lognormal model. The top line (labelled "option-based model") refers to the model based on option prices. It's the same as the top line in Figure 6. The bottom line (labelled "consumption-based model") refers to the model derived from consumption data as described in the previous paragraph. The two models are clearly different. The consumption-based model has a steeper smile, greater curvature, and lower at-the-money volatility. This follows, in part, from its greater risk-neutral skewness and excess kurtosis [columns (3) and (4) of Table 2]. They suggest higher risk-neutral probabilities of large disasters (the left side of the figure) and lower probabilities of less extreme outcomes (the middle and right of the figure). These differences in the underlying distributions result in significantly different option prices.

Now consider the reverse: the consumption growth process implied by option prices. Here the focus is on the true distribution of consumption growth. For the option model, this involves taking the risk-neutral distribution of returns implied by option prices and computing the true distribution of consumption growth. We no longer need an estimate of the true distribution from returns data. Instead, we use power utility to link true to risk-neutral parameters and equation (22) to link consumption growth to returns. This imposes different structure than the option model on the connection between true and risk-neutral parameters. For example, the restriction $\omega = \omega^*$ used in option model no longer holds. The calculations are analogous to the previous comparison. First, we use the scale parameter λ to transform risk-neutral parameters for equity returns into risk-neutral parameters for consumption growth. This involves replacing $(\omega^*, \theta^*, \delta^*)$ in column (4) of Table 2 with $(\omega^*, \theta^*/\lambda, \delta^*/\lambda)$ for reasons outlined in Appendix A.7. Second, we use (25), a consequence of power utility, to compute the parameters (ω, θ, δ) of the true distribution. Finally, we set σ to match our target value for the standard deviation of log consumption growth and α to match the equity premium.

The consumption process derived this way from option prices [column (5) of Table 2] has the same Poisson-normal distribution as the consumption process estimated from international macroeconomic evidence [column (3)]. The parameters, however, are much different. In the consumption-based model, there is a small chance (governed by the jump intensity $\omega = 0.01$) of a large jump (the mean jump $\theta = -0.3$ is 8.6 standard deviations). In the option-based model, there is a larger chance (jump intensity is $\omega = 1.3987$) of a much smaller jump (the mean jump $\theta = -0.0074$ is 0.21 standard deviations). Both models generate disasters in the sense that the probabilities of tail events are much larger than in the lognormal case [column (1)]. The probability of a 3 standard deviation drop in consumption, similar to the US in the Great Depression, is about 1% in each case. However, declines in consumption of more than 5 standard deviations are much more likely in the consumption-based model (probability 0.0079) than in the option-based model (0.0001). Events of this magnitude have not been observed in US history, so the models disagree on events that have never occurred. The difference in tail probabilities is reflected in their cumulants. Skewness and excess kurtosis are -11.01 and 145.06, respectively, in the consumption-based model, but only -0.28 and 0.48 in the option-based model.

7 Extensions

We have described, in a relatively simple theoretical setting, how option prices can be used to infer probabilities of extreme outcomes, including the infrequent sharp declines in consumption growth documented in international macroeconomic data by Barro and others. We find that the distribution of outcomes implied by option prices is less extreme than the macroeconomic evidence suggests. The analysis that leads to this conclusion leans heavily on three supports: iid consumption growth and returns, power utility over aggregate consumption, and a close connection between consumption growth and equity returns. Each deserves a closer look.

Consider the iid assumption. Our objective is to characterize the unconditional distribution of consumption growth, particularly the distribution of large adverse outcomes. The question is whether the kinds of time-dependence we see in asset prices are quantitatively important in assessing the role of extreme events. It's hard to make a definitive statement without knowing the precise form of time-dependence, but there's good reason to think its impact could be small. The leading example in this context is stochastic volatility, a central feature of the option pricing model estimated by Broadie, Chernov, and Johannes (2007). However, average implied volatility smiles from this model are very close to those from an iid model in which the variance is set equal to its mean. Furthermore, stochastic volatility has little impact on the probabilities of tail events, which is our interest here.

Perhaps the most interesting extension is to go beyond power utility or even the representative agent framework. The option model in Section 5 gives us a hint of the features such an extension should have. Taken literally, the parameter values of that model are inconsistent with power utility. The difference between δ and δ^*, for example, is zero with power utility [see equation (25)]. Nevertheless, we can derive something like risk aversion for the model. In (13), risk aversion is implicit in the relation between the pricing kernel and consumption growth:

$$\alpha \quad = \quad -\frac{\partial \log m}{\partial \log g}.$$

In the option model, the analogous expression is

$$\text{RA} \quad = \quad -\frac{\partial \log(p^*/p)}{\partial \log r^e} \cdot \frac{\partial \log r^e}{\partial \log g}, \tag{28}$$

which in principle can vary with $\log r^e$. See Leland (1980). In our setting, the second term is the constant λ, so any variation in implied risk aversion RA comes from the first term.

Risk aversion defined this way equals the coefficient of relative risk aversion when assets are priced by a representative agent with power utility. More generally, we find it a useful way of describing how risk is priced in any arbitrage-free model. In our option-based model RA depends, in general, on the state; see Appendix A.9. Figure 9 shows that with our parameter values, RA is larger for negative returns than for positive ones, with risk aversion of 9.2 for returns of -10% and 2.9 for returns of $+10\%$. Related work has generated a wide range of patterns with different methodologies, but they all find that risk aversion varies with the state. See, for example, Ait-Sahalia and Lo (2000), Bakshi, Madan, and Panayotov (2010), Jackwerth (2000), Rosenberg and Engle (2002), and Ziegler (2007). What we find interesting is the possibility that risk premiums on assets might reflect not only extreme outcomes but pricing that gives such outcomes greater weight than power utility.

Power utility is the workhorse of macroeconomics and finance, but our option model suggests greater aversion to bad outcomes than good ones. If this turns out to be a robust feature of the evidence, it's worth thinking about where it comes from. One possibility is explore alternative preferences, including skewness preference (Harvey and Siddique, 2000), recursive preferences (Garcia, Luger, and Renault, 2002, and Wachter, 2009), state-dependent preferences (Chabi-Yo, Garcia, and Renault, 2008), ambiguity (Drechsler, 2008, and Liu, Pan, and Wang, 2005), learning (Shaliastovich, 2008), and habits (Bekaert and Engstrom, 2010, and Du, 2008). Another promising avenue is heterogeneity across agents. Certainly there is strong evidence of imperfect risk-sharing across individuals and good reason to suspect that this affects asset prices. Bates (2008), Chan and Kogan (2002), Guvenen (2009), Longstaff and Wang (2008), and Lustig and Van Nieuwerburgh (2005) are notable examples. The question for future work is whether these extensions provide a persuasive explanation for prices of equity index options.

Finally, we loosen the tight connection between consumption growth and equity returns. We've followed a long tradition in tying dividends to consumption. The tradition is largely a matter of convenience, because it's simpler to have one random variable rather than two, but it's motivated by the evidence connecting equity returns to business cycles. Still, it's a relatively simple matter to explore the possibility of imperfect correlation between consumption growth and equity returns. Figure 10 shows how the two are related in US data. There's a clear association between business cycles (represented by consumption growth) and equity returns, but the association isn't perfect. The sample correlation is 0.566.

Consider, then, a bivariate model of consumption growth and equity returns with arbitrary correlation between the two. Does allowing imperfect correlation affect our conclusion that consumption evidence implies more extreme outcomes than option prices? The answer is no, but it's worth working through the details. Let $x' = (x_1, x_2) = (\log g, \log r^e - \log r^1)$ follow

$$x_{t+1} \;\; = \;\; w_{t+1} + z_{t+1},$$

with w and z each bivariate and iid over time. The first component is bivariate normal: $w_t \sim \mathcal{N}(\mu, \Sigma)$, where μ and Σ have elements μ_i and σ_{ij}, respectively. The second component is a Poisson mixture of bivariate normals similar to the model of Section 3.2. Jumps occur with Poisson intensity ω. Each jump generates a draw from the bivariate normal distribution $\mathcal{N}(\theta, \Delta)$, where θ and Δ have elements θ_i and δ_{ij}. This process is a special case of one used by Ait-Sahalia, Cacho-Diaz, and Hurd (2009) and similar to consumption-dividend processes used by Gabaix (2010) and Longstaff and Piazzesi (2004). Option pricing follows directly from the risk-neutral distribution over equity returns. With appropriate redefinition of parameters, option pricing has the same structure as Section 5 and Appendix A.6.

We choose parameter values that reproduce the mean and variance of log consumption growth, the mean and variance of the log excess return on equity, and the correlation between them. Consider the components in turn. We use the parameters reported in column (3) of Table 2 for the (true) consumption process. Here we label them $(\mu_1, \sigma_{11}, \omega, \theta_1, \delta_{11})$. With the exception of μ_2, the parameters of the return process are scaled versions of the same numbers: $\sigma_{22} = \lambda^2 \sigma_{11}$, $\theta_2 = \lambda \theta_1$, and $\delta_{22} = \lambda^2 \delta_{11}$. The jump intensity ω is, of course, the same. The mean μ_2 is then chosen to match the equity premium: $\mu_2 + \omega\theta_2 = 0.0400$. The correlation depends on the mean jumps and the correlations $\rho_w = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$ and $\rho_z = \delta_{12}/(\delta_{11}\delta_{22})^{1/2}$. We set $\rho_w = \rho_z = \rho$ to match the correlation in the data. Option prices follow from the risk-neutral parameters. With power utility, $\mu_2^* = \mu_2 - \alpha\sigma_{12}$, $\omega^* = \omega \exp(-\alpha\theta_1 + \alpha^2\delta_{11}/2)$, and $\theta_2^* = \theta_2 - \alpha\delta_{12}$. We choose risk aversion α to satisfy the arbitrage restriction, $\mu_2^* + \sigma_{22}/2 + \omega^*[\exp(\theta_2^* + \delta_{22}/2) - 1]$. See Appendix A.10.

If consumption growth and equity returns are perfectly correlated, this procedure reproduces the calculation of consumption-based option prices in Figure 8. For the correlation observed in the data, the result is the middle line in the same figure. We see that the two consumption-based lines are similar to each other and notably different from the smile estimated from option prices. It appears, then, that perfect correlation between consumption growth and the return on equity is not the source of the sharp difference between volatility smiles based on option prices and consumption data.

One might ask, in the end, whether options reliably identify extreme events (disasters) in consumption. We think the answer is yes, but let's run through the argument. There are two degrees of separation between equity options and consumption: options identify the risk-neutral distribution of returns and we're interested in the true distribution of consumption growth. The first link is between true and risk-neutral returns. We've seen that a direct estimate of the true distribution of returns shares with the estimated risk-neutral distribution modest values of skewness and excess kurtosis [column (4) of Table 2]. The second link is between returns and consumption growth. When the two are perfectly correlated, as they are in equation (22), their skewness, excess kurtosis, and tail probabilities (measured in standard deviation units) are the same. The contrast along these dimensions between distributions based on international consumption data [column (3)] and option prices [columns (4) and (5)] is striking. So, too, is the difference between implied volatilities based on option prices and consumption data (Figure 8). Finally, the evidence on the relation between returns and consumption growth (Figure 10) and the limited impact of imperfect correlation on option prices (Figure 8) suggest that option prices are a reasonably good indicator of the likelihood of disasters in consumption growth. It's possible that future work using other methods will detect a significant difference between the tail behavior of consumption growth and equity returns. In the meantime, we think the evidence indicates that equity index options imply smaller probabilities of consumption disasters than the international macroeconomic evidence.

8 Final remarks

We have used prices of equity index options to infer probabilities of large negative realizations of consumption growth that we can compare to the macroeconomic evidence summarized by Barro and coauthors. The exercise faces two issues from the start: equity index options refer to equity, not consumption, and they reflect risk-neutral rather than true probabilities. We address these issues in a number of ways. All of them indicate that the probabilities of extreme adverse events implied by option prices are smaller than we see in international macroeconomic data. On the broader question of the role of disasters in asset pricing, we find that a pricing kernel constructed from option prices includes substantial contributions from high-order cumulants. In this sense, the departures from lognormality suggested by the disaster literature remain quantitatively important in the option pricing model. Furthermore, the contribution of odd high-order cumulants is suggestive of the market pricing of return asymmetries noted in research on skewness preference.

This exercise had two useful byproducts. One is the reminder that matching the equity premium may not be enough. There's growing evidence that other trading strategies can generate average returns that are substantially higher, so that models designed to account for the equity premium may not be able to account for higher returns on other assets. The other is the value of transform methods, particularly cumulant-generating functions. These are not new to finance, but they are nevertheless extremely helpful. We find them not only a source of intuition and compact notation, but a convenient approach to a number of practical problems.

References

Abel, Andrew B., 1999, "Risk premia and term premia in general equilibrium," *Journal of Monetary Economics* 43, 3-33.

Ait-Sahalia, Yacine, Julio Cacho-Diaz and T.R. Hurd, 2009, "Portfolio choice with jumps: a closed-form solution," *Annals of Applied Probability* 19, 556-584.

Ait-Sahalia, Yacine, and Andrew W. Lo, 2000, "Nonparametric risk management and implied risk aversion," *Journal of Econometrics* 94, 9-51.

Ait-Sahalia, Yacine, Yubo Wang, and Francis Yared, 2001, "Do option markets correctly price the probabilities of movement of the underlying price?," *Journal of Econometrics* 102, 67-110.

Alvarez, Fernando, and Urban Jermann, 2005, "Using asset prices to measure the persistence of the marginal utility of wealth," *Econometrica* 73, 1977-2016.

Bakshi, Gurdip, Dilip Madan, and George Panayotov, 2010, "Returns of claims on the upside and the viability of U-shaped pricing kernels," *Journal of Financial Economics* 97, 130-154.

Bansal, Ravi, and Bruce N. Lehmann, 1997, "Growth-optimal portfolio restrictions on asset pricing models," *Macroeconomic Dynamics* 1, 333-354.

Barro, Robert J., 2006, "Rare disasters and asset markets in the twentieth century," *Quarterly Journal of Economics* 121, 823-867.

Barro, Robert J., Emi Nakamura, Jon Steinsson, and Jose F. Ursua, 2009, "Crises and recoveries in an empirical model of consumption disasters," manuscript, June.

Barro, Robert J., and Jose F. Ursua, 2008, "Macroeconomic crises since 1870," *Brookings Papers on Economic Activity*, 255-335.

Bates, David, 1988, "Pricing options under jump-diffusion processes," manuscript.

Bates, David, 2008, "The market for crash risk," *Journal of Economic Dynamics and Control* 32, 2291-2231.

Bates, David, 2010, "US stock market crash risk, 1926-2009," manuscript, January.

Bekaert, Geert, and Eric Engstrom, 2010, "Asset return dynamics under bad environment-good environment fundamentals," manuscript, June.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2005, "Can standard preferences explain the prices of out-of-the-money S&P 500 put options?," manuscript, September.

Breeden, Douglas T. and Robert H. Litzenberger, 1978, "Prices of state-contingent claims implicit in option prices," *Journal of Business* 51, 621-651.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, "Model specification and risk premia: evidence from futures options," *Journal of Finance* 62, 1453-1490.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2009, "Understanding index option returns," *Review of Financial Studies*, in press.

Buhlmann, Hans, Freddy Delbaen, Paul Embrechts, and Albert Shiryaev, 1996, "No-arbitrage, change of measure, and conditional Esscher transforms," *CWI Quarterly* 9, 291-317.

Campanale, Claudio, Rui Castro, Gian Luca Clementi, 2010, "Asset pricing in a production economy with Chew-Dekel preferences," *Review of Economic Dynamics* 13, 379-402.

Chan, Yeung Lewis, and Leonid Kogan, 2002, "Catching up with the Joneses: heterogeneous preferences and the dynamics of asset prices," *Journal of Political Economy* 110, 1255-1285.

Chabi-Yo, Fousseni, Rene Garcia, and Eric Renault, 2008, "State-dependence can explain the risk aversion puzzle," *Review of Financial Studies* 21, 973-1011.

Drechsler, Itamar, 2008, "Uncertainty, time-varying fear, and asset prices," manuscript, December.

Drechsler, Itamar, and Amir Yaron, 2008, "What's vol got to do with it?" manuscript, December.

Du, Du, 2008, "How bad will the potential economic disasters be? Evidences from S&P 500 index options data," manuscript, June.

Gabaix, Xavier, 2010, "Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance," manuscript, March.

Garcia, Rene, Richard Luger, and Eric Renault, 2003, "Empirical assessment of an intertemporal option pricing model with latent variables," *Journal of Econometrics* 116, 49-83.

Guvenen, Fatih, 2009, "A parsimonious macroeconomic model for asset pricing," *Econometrica*, forthcoming.

Hansen, Lars Peter, and Ravi Jagannathn, 1991, "Implications of security market data for models of dynamic economies," *Journal of Political Economy* 99, 225-262.

Harvey, Campbell R., and Akhtar Siddique, 2000, "Conditional skewness in asset pricing tests," *Journal of Finance* 55, 1263-1295.

Jackwerth, Jens Carsten, 2000, "Recovering risk aversion from option prices and realized returns," *Review of Financial Studies* 13, 433-451.

Kraus, Alan, and Robert H. Litzenberger, 1976, "Skewness preference and the valuation of risk assets," *Journal of Finance* 31, 1085-1100.

Leland, Hayne E., 1980, "Who should buy portfolio insurance?," *Journal of Finance* 35, 581-594.

Liu, Jun, Jun Pan, and Tan Wang, 2005, "An equilibrium model of rare-event premia and its implication for option smirks," *Review of Financial Studies* 18, 131-164.

Longstaff, Francis A., and Monika Piazzesi, 2004, "Corporate earnings and the equity premium," *Journal of Financial Economics* 74, 401-421.

Longstaff, Francis A., and Jiang Wang, 2008, "Asset pricing and the credit market," manuscript, April.

Lustig, Hanno N., and Stijn G. Van Nieuwerburgh, 2005, "Housing collateral, consumption insurance, and risk premia: an empirical perspective, " *Journal of Finance* 60, 1167-1219.

Martin, Ian, 2007, "Disasters and asset pricing: evidence from option markets," manuscript, February.

Martin, Ian, 2009, "Consumption-based asset pricing with higher cumulants," manuscript, January.

Mehra, Rajnish, and Edward C. Prescott, 1985, "The equity premium: a puzzle," *Journal of Monetary Economics* 15, 145-161.

Merton, Robert C., 1976, "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics* 3, 125-144.

Naik, Vasanttilak, and Moon Lee, 1990, "General equilibrium pricing of options on the market portfolio with discontinuous returns," *Review of Financial Studies* 3, 493-521.

Rietz, Thomas A., 1988, "The equity risk premium: a solution," *Journal of Monetary Economics* 22, 117-131.

Rosenberg, Joshua V., and Robert F. Engle, 2002, "Empirical pricing kernels," *Journal of Financial Economics* 64, 341-372.

Shaliastovich, Ivan, 2008, "Learning, confidence and option prices," manuscript.

Shiller, Robert, 2009, "Long term stock, bond, interest rate and consumption data," spreadsheet posted on his web site.

Wachter, Jessica, 2009, "Can time-varying risk of rare disasters explain aggregate stock market volatility?," manuscript, March.

Wu, Liuren, 2006, "Dampened power law: reconciling the tail behavior of financial security returns," *Journal of Business* 79, 1445-1474.

Ziegler, Alexandre, 2007, "Why does implied risk aversion smile?," *Review of Financial Studies* 20, 859-904.

Zin, Stanley, 2002, "Are behavioral models of asset-pricing structural?," *Journal of Monetary Economics* 49, 215-228.

Table 1
Properties of consumption growth and asset returns

Variable	Mean	Std Dev	Skewness	Kurtosis	Autocorr
(a) Consumption growth and returns, annual, 1889-2009					
Consumption growth	0.0198	0.0350	−0.34	1.11	−0.06
Return on one-year bond	0.0180	0.0566	0.04	2.40	0.34
Return on equity	0.0587	0.1795	−0.61	0.43	0.02
Excess return on equity	0.0407	0.1812	−0.72	0.91	0.04
(b) Consumption growth and returns, annual, 1986-2009					
Consumption growth	0.0178	0.0150	−0.87	0.66	0.59
Return on one-year bond	0.0207	0.0185	−0.37	−0.85	0.27
Return on equity	0.0641	0.1845	−1.24	1.56	0.00
Excess return on equity	0.0434	0.1808	−1.39	2.01	−0.03

Notes. Entries are statistics computed from annual observations for the US economy. Mean is the sample mean, Std Dev is the standard deviation, Skewness is the standard measure of skewness, Kurtosis is the standard measure of excess kurtosis, and Autocorr is the first autocorrelation. Consumption growth is $\log(c_t/c_{t-1})$ where c is real per capita consumption. Returns are logarithms of gross real returns and the excess return is the difference between the log-returns on equity and the one-year bond. The one-year bond is the Treasury security of maturity closest to one year. Equity is the S&P 500. Consumption and return data are from Shiller (2009).

Table 2
Parameter values and properties of model economies

Parameter	Lognormal Cons Gr (1)	Bernoulli Cons Gr (2)	Poisson Cons Gr (3)	Merton Returns (4)	Implied Cons Gr (5)
Preferences					
α	8.92	6.21	5.19	—	8.70
True distribution					
μ	0.0200	0.0230	0.0230	0.0792	0.0303
σ	0.0350	0.0183	0.0100	0.1699	0.0253
ω	—	0.0100	0.0100	1.5120	1.3987
θ	—	−0.3000	−0.3000	−0.0259	−0.0074
δ	—	—	0.1500	0.0407	0.0191
Risk-neutral distribution					
μ^*	0.0091	0.0209	0.0225	0.0584	0.0247
ω^*	—	0.0610	0.0642	1.5120	1.5120
θ^*	—	−0.3000	−0.4168	−0.0542	−0.0105
δ^*	—	—	0.1500	0.0981	0.0191
Skewness, excess kurtosis, and tail probabilities					
γ_1 (true skewness)	0	−6.11	−11.02	−0.04	−0.28
γ_2 (true kurtosis)	0	50.26	145.06	0.02	0.48
γ_1^* (risk-neutral skewness)	0	−3.34	−4.51	−0.25	−0.38
γ_2^* (risk-neutral kurtosis)	0	10.10	21.96	0.30	0.53
γ_1 ($\log m$ skewness)	0	6.11	11.02	−0.12	0.28
γ_2 ($\log m$ kurtosis)	0	50.26	145.06	2.21	0.48
Tail prob (≤ -3 st dev)	0.0013	0.0100	0.0090	0.0032	0.0081
Tail prob (≤ -5 st dev)	0.0000	0.0100	0.0079	0.0000	0.0001
Entropy					
$L(m) = L(p^*/p)$	0.0487	0.0408	0.0400	0.7747	0.0478

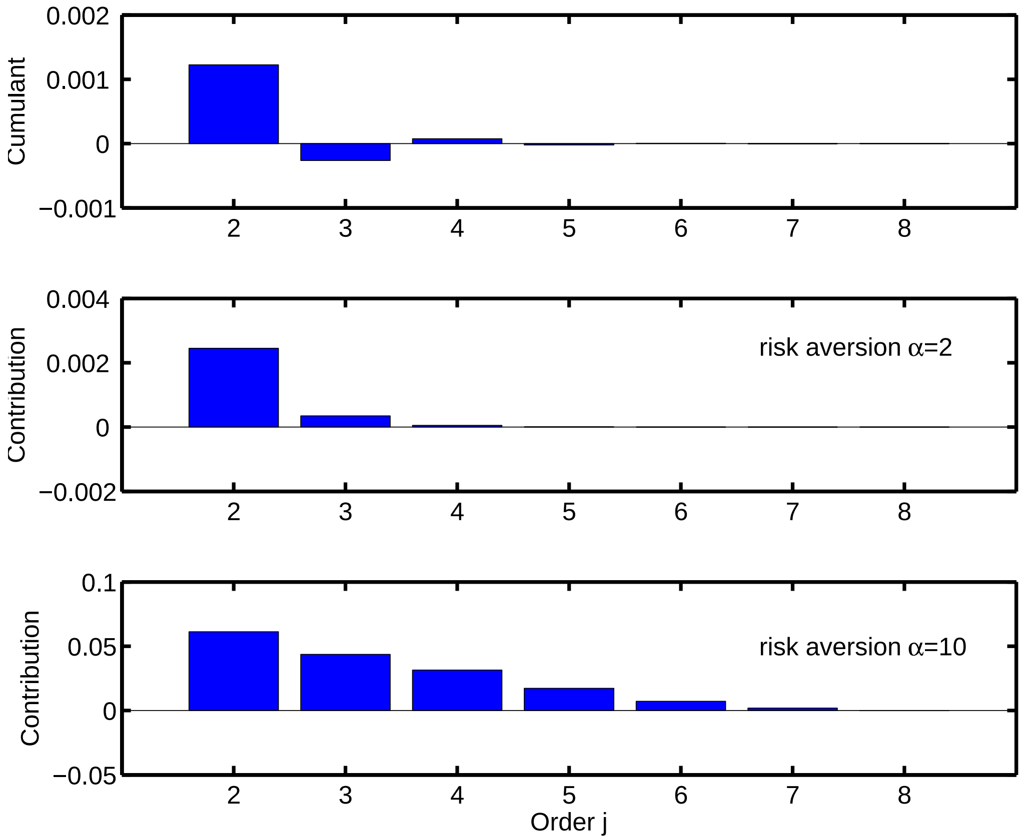
Notes. Entries are parameters and properties of examples with different specifications of disasters. The labels at the top of the columns describe the model used and the variable on which it is based. Columns (1)-(3) and (5) are based on consumption growth. In each one, log consumption growth has a mean of 0.0200 and a standard deviation of 0.0350. Risk aversion α is chosen to match the mean equity premium (0.0400). Column (4) is the Merton model parameterized to option prices and equity returns. Column (5) takes this model, scales the risk-neutral parameters to fit consumption growth, and sets the true parameters by applying the relations implied by power utility [equation (25)]. γ_1 and γ_2 are the traditional measures of skewness and excess kurtosis, defined in equation (5). We report versions for the true distribution of log consumption growth or the log return on equity, the risk-neutral distribution, and the distribution of the pricing kernel. Tail probabilities refer to the probabilities that log consumption growth or the log return on equity are less than −3 and −5 standard deviations, respectively, from their mean.

Table 3
Components of entropy for model economies

Model	Entropy	Variance/2	High-Order Cumulants Odd	High-Order Cumulants Even
Normal consumption growth				
$\alpha = 2$	0.0025	0.0025	0	0
$\alpha = 5$	0.0153	0.0153	0	0
$\alpha = 10$	0.0613	0.0613	0	0
$\alpha = 8.92^*$	0.0487	0.0487	0	0
Bernoulli consumption growth				
$\alpha = 2$	0.0029	0.0025	0.0004	0.0000
$\alpha = 5$	0.0234	0.0153	0.0060	0.0021
$\alpha = 10$	0.1614	0.0613	0.0621	0.0380
$\alpha = 10, \theta = +0.3$ (boom)	0.0372	0.0613	–0.0621	0.0380
$\alpha = 10, \theta = -0.15, \omega = 0.02$	0.0765	0.0613	0.0115	0.0038
$\alpha = 6.21^*$	0.0408	0.0236	0.0121	0.0051
Poisson consumption growth				
$\alpha = 2$	0.0033	0.0025	0.0007	0.0002
$\alpha = 5$	0.0356	0.0153	0.0132	0.0071
$\alpha = 10$	0.5837	0.0613	0.2786	0.2439
$\alpha = 5.19^*$	0.0400	0.0165	0.0151	0.0084
Models fit to option prices				
Merton equity returns	0.7747	0.4720	0.1127	0.1900
Implied consumption growth	0.0478	0.0464	0.0013	0.0002

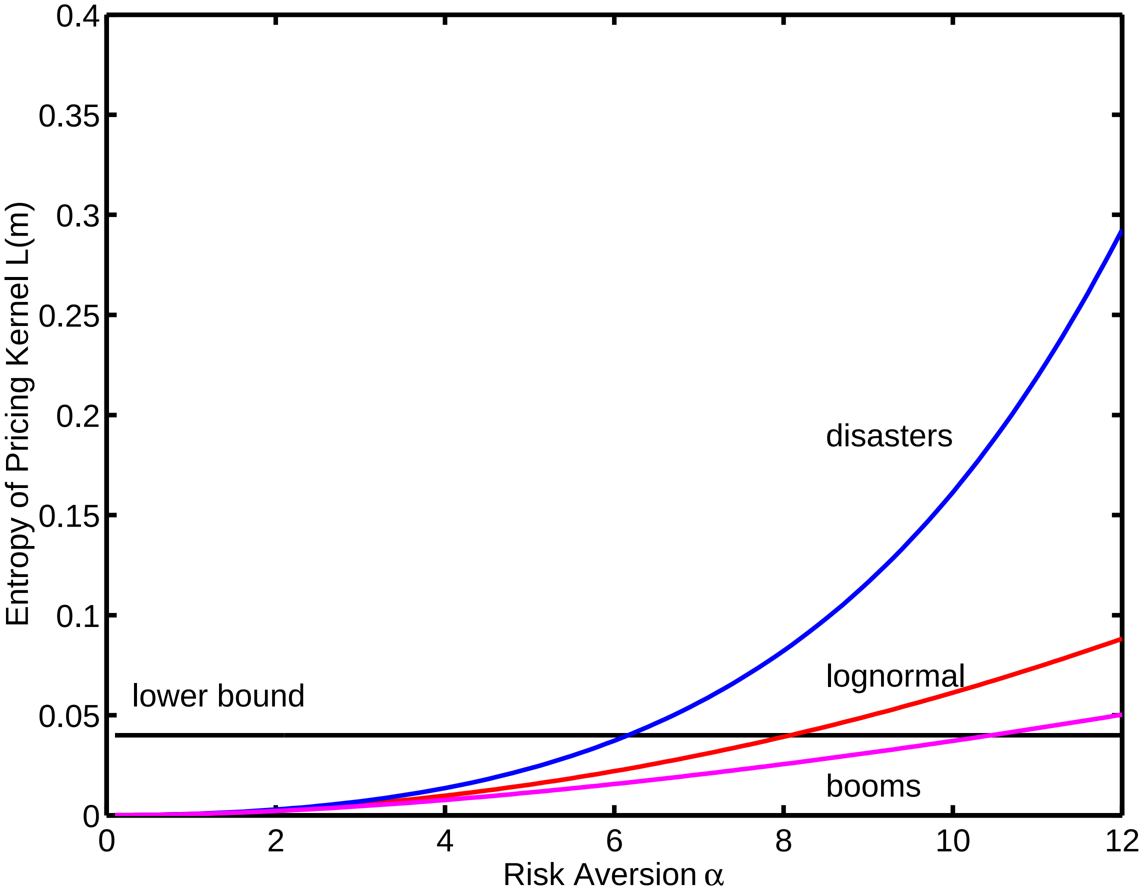
Notes. Entries include entropy of the pricing kernel and its components for a variety of models. Entropy is the sum of contributions from the variance and from odd and even high-order cumulants (those of order $j \geq 3$). An asterisk denotes a value of α that matches the observed equity premium.

Figure 1
Bernoulli jumps: cumulants of log consumption growth and contributions to entropy



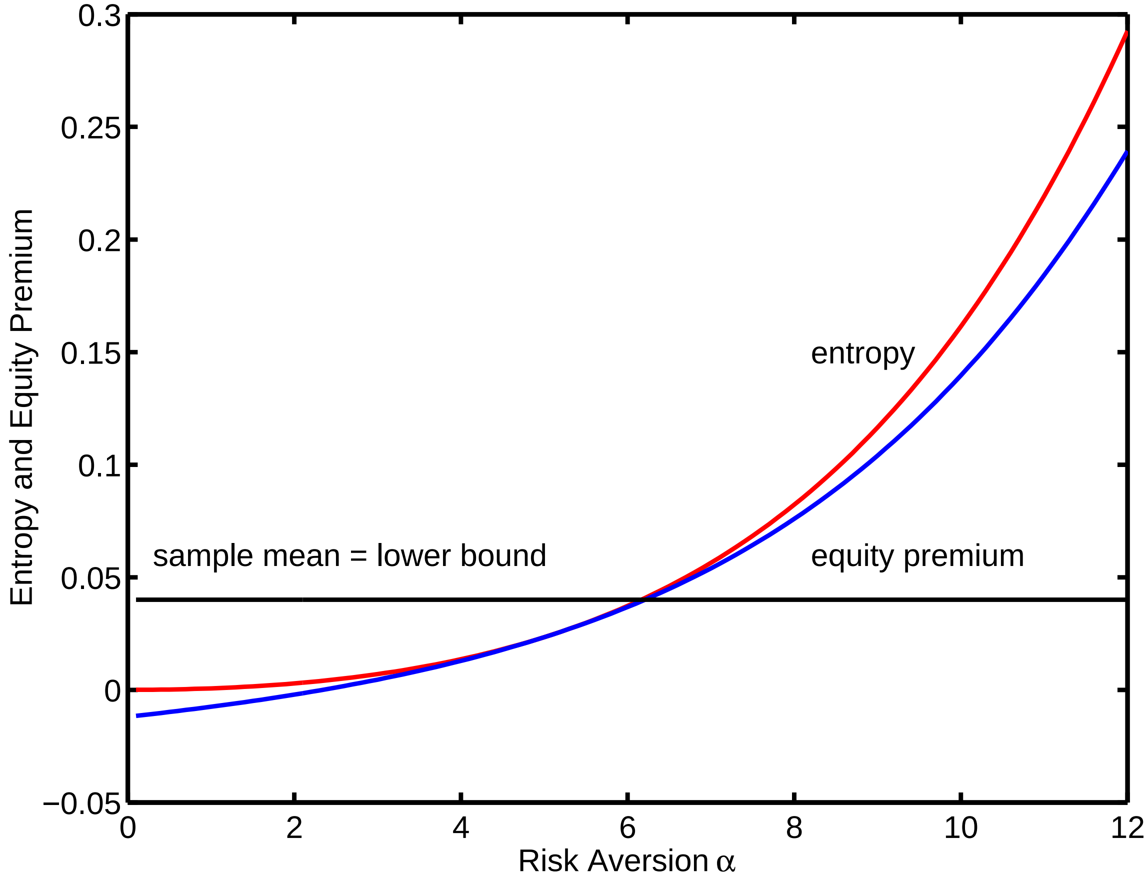
Notes. The panels graph terms in the power series expansion of entropy, equation (6), for the consumption-based asset pricing model with a Bernoulli jump component. The top panel plots the jth cumulant of log consumption growth, $\kappa_j(\log g)$, against its order j. The next two panels plot the contribution to entropy of the jth term, $\kappa_j(\log m)/j! = (-\alpha)^j \kappa_j(\log g)/j!$, against j for risk aversion α equal to 2 and 10, respectively. The model and parameter values are reported in Section 3.1 and column (2) of Table 2.

Figure 2
Bernoulli jumps: entropy of the pricing kernel



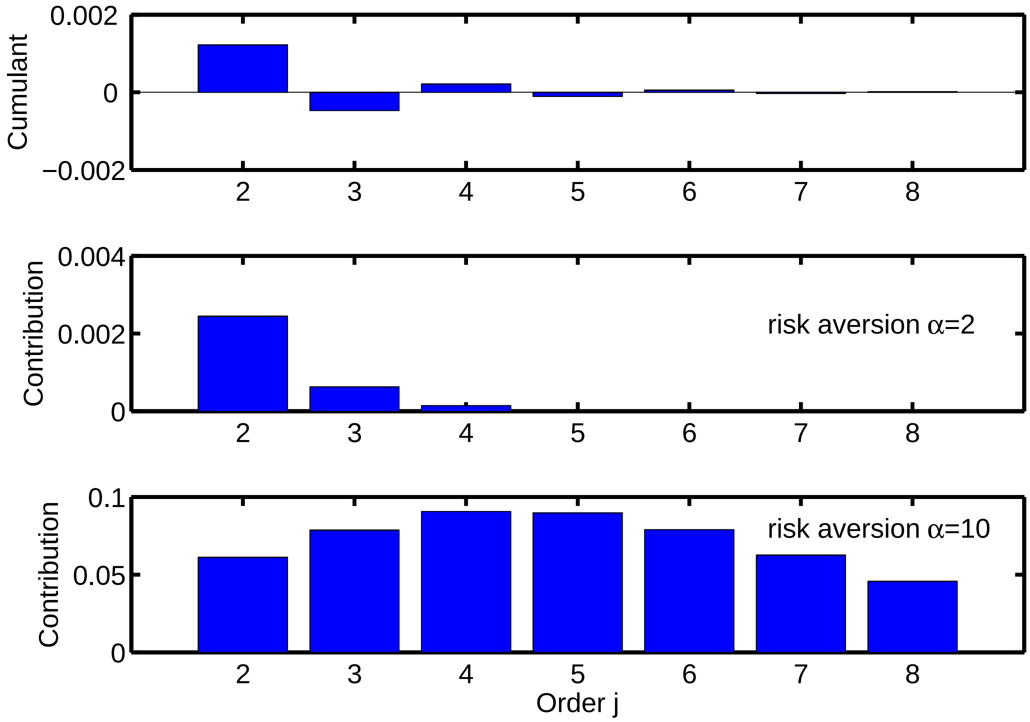
Notes. The lines represent versions of the consumption-based asset pricing model with Bernoulli jumps outlined in Section 3.1. The middle line is the lognormal model: log consumption growth is normal and there is no jump component. The top line shows how a Bernoulli component (an infrequent large negative realization) increases entropy. The bottom line shows how this changes when the jump is positive (an infrequent large positive realization).

Figure 3
Bernoulli jumps: entropy and the equity premium



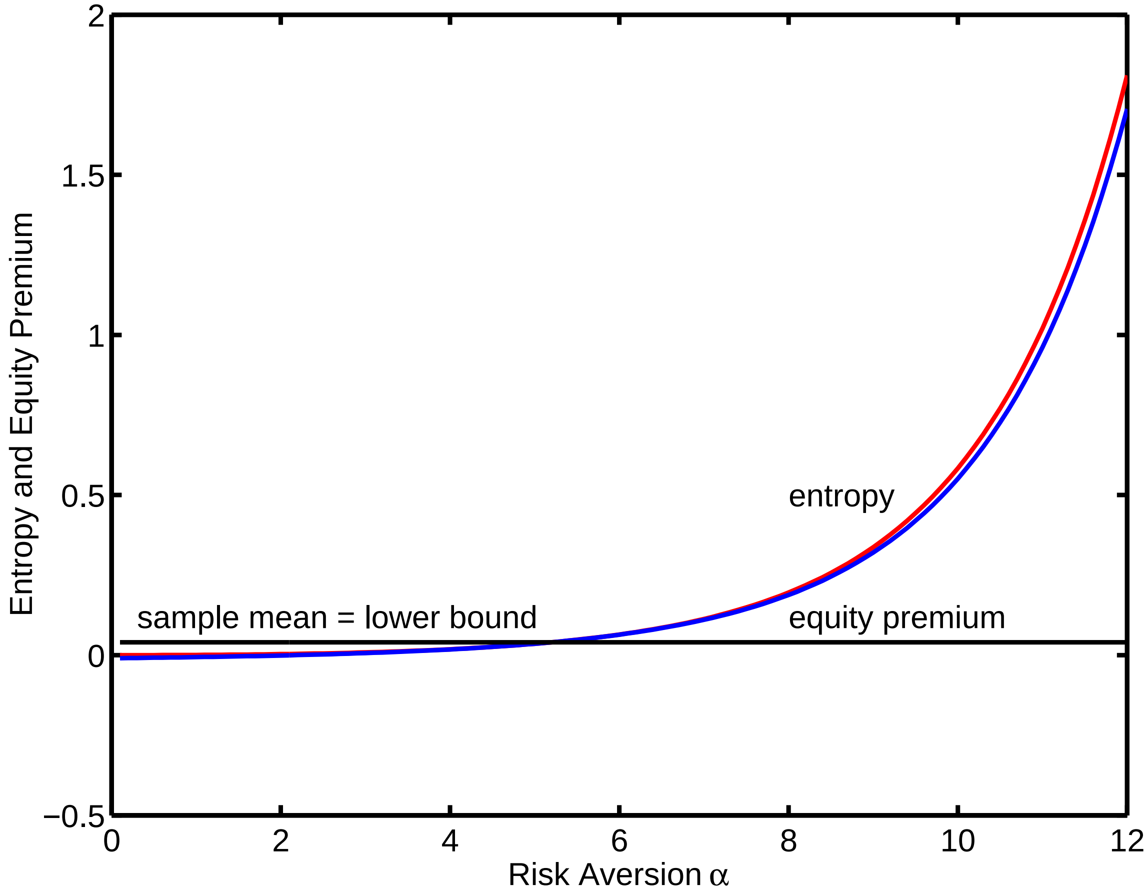
Notes. The lines represent entropy of the pricing kernel and the equity premium for the consumption-based asset pricing model with a Bernoulli jump component (infrequent large negative realization of consumption growth). Equity is defined by equation (21) and its excess return by (22). The leverage parameter is $\lambda = 5.1$.

Figure 4
Poisson jumps: cumulants of log consumption growth and contributions to entropy



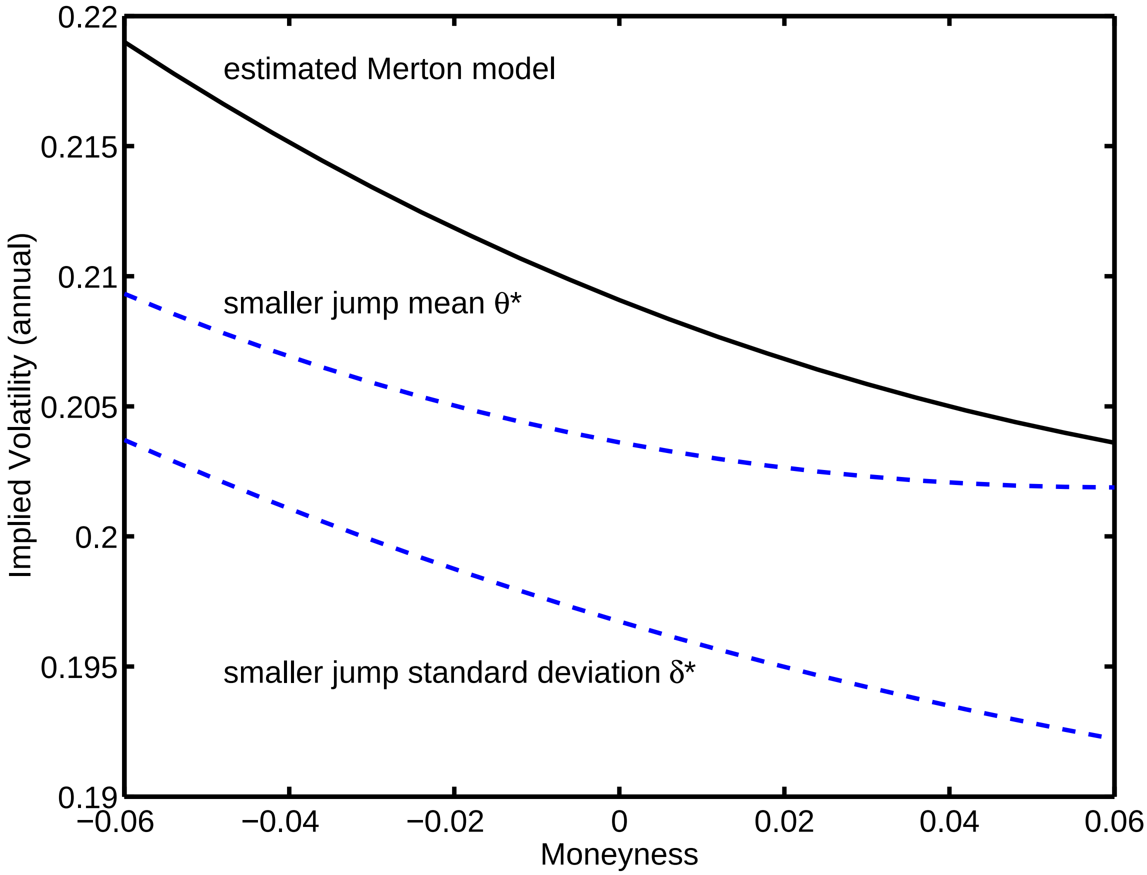
Notes. The panels graph terms in the power series expansion of entropy, equation (6), for the consumption-based asset pricing model with a Poisson-normal jump component. The top panel plots the jth cumulant of log consumption growth, $\kappa_j(\log g)$, against its order j. The next two panels plot the contribution to entropy of the jth term, $\kappa_j(\log m)/j! = (-\alpha)^j \kappa_j(\log g)/j!$, against j for risk aversion α equal to 2 and 10, respectively. The model and parameter values are reported in Section 3.2 and column (3) of Table 2.

Figure 5
Poisson jumps: entropy and equity premium



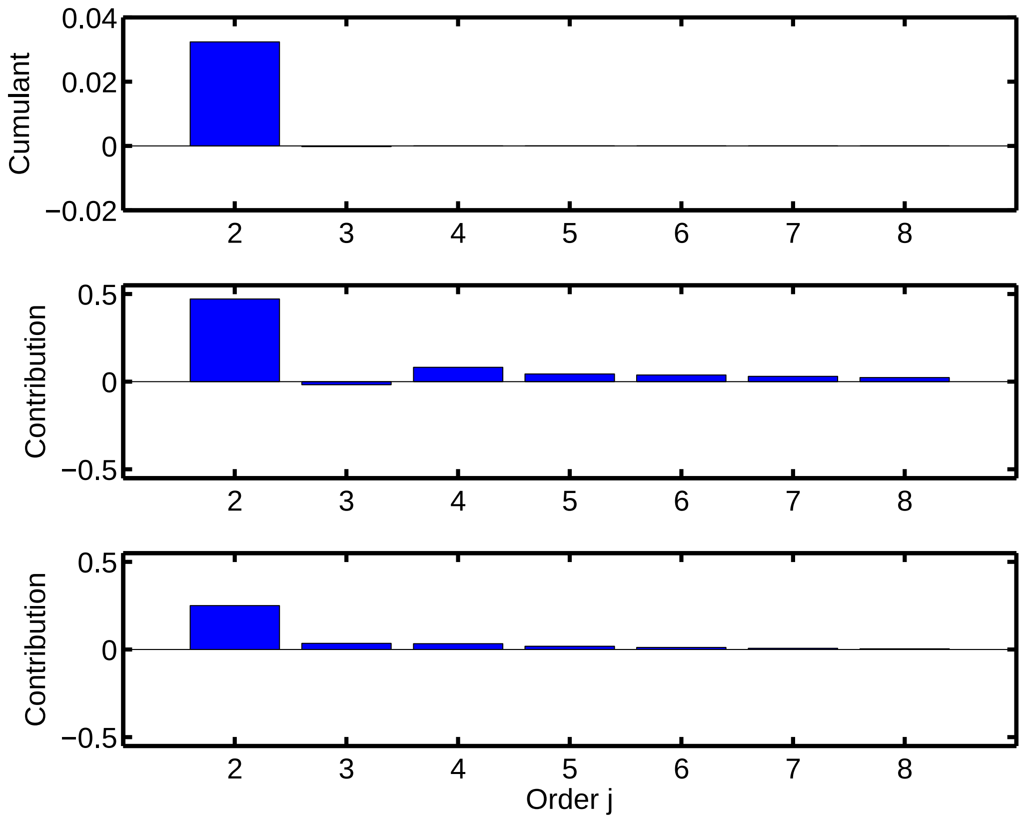
Notes. The lines represent entropy of the pricing kernel and the equity premium for the consumption-based asset pricing model with a Poisson-normal jump component. Equity is defined by equation (21) and its excess return by (22). The leverage parameter is $\lambda = 5.1$.

Figure 6
Option model: implied volatility smiles for 3-month options



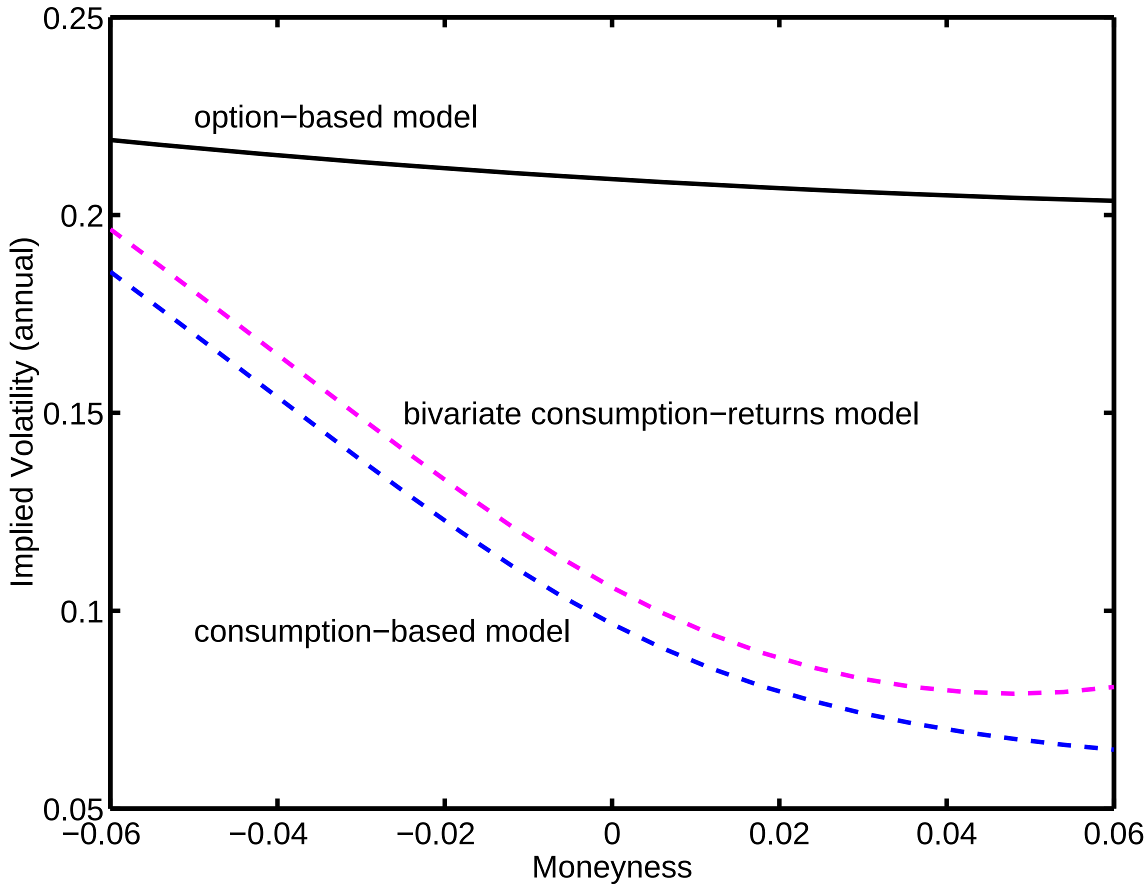
Notes. The lines represent implied volatility "smiles" for the Merton model with estimated parameters and some alternatives. Moneyness is measured as the proportional difference of the strike from the price, (strike − price)/price. For the solid line the parameters are those reported in column (4) of Table 2. For the second line, we have divided the jump mean θ^* by two. For the bottom line, we have divided the jump variance δ^{*2} by two.

Figure 7
Option model: cumulants of equity returns and contributions to entropy



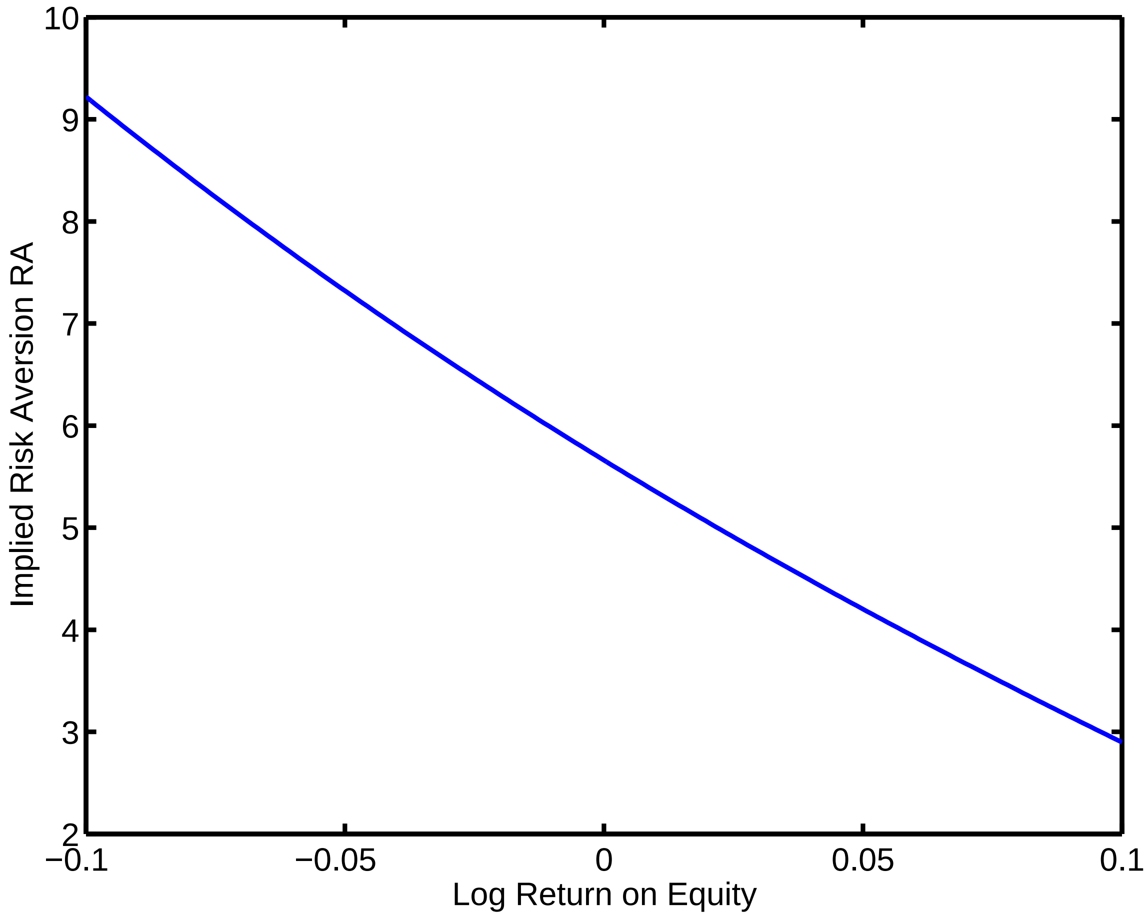
Notes. The figure summarizes properties of the estimated Merton model using parameters reported in column (4) of Table 2. The top panel shows cumulants of the log excess return on equity based on its (estimated) true distribution. The second panel shows contributions to entropy of the cumulants of $\log m$. The contribution of order j is $\kappa_j(\log m)/j!$ where $\kappa_j(\log m)$ is the jth cumulant of $\log m$. The third panel is the same with θ^* and δ^{*2} each divided by two.

Figure 8
Option models: implied volatility smiles based on option and consumption data



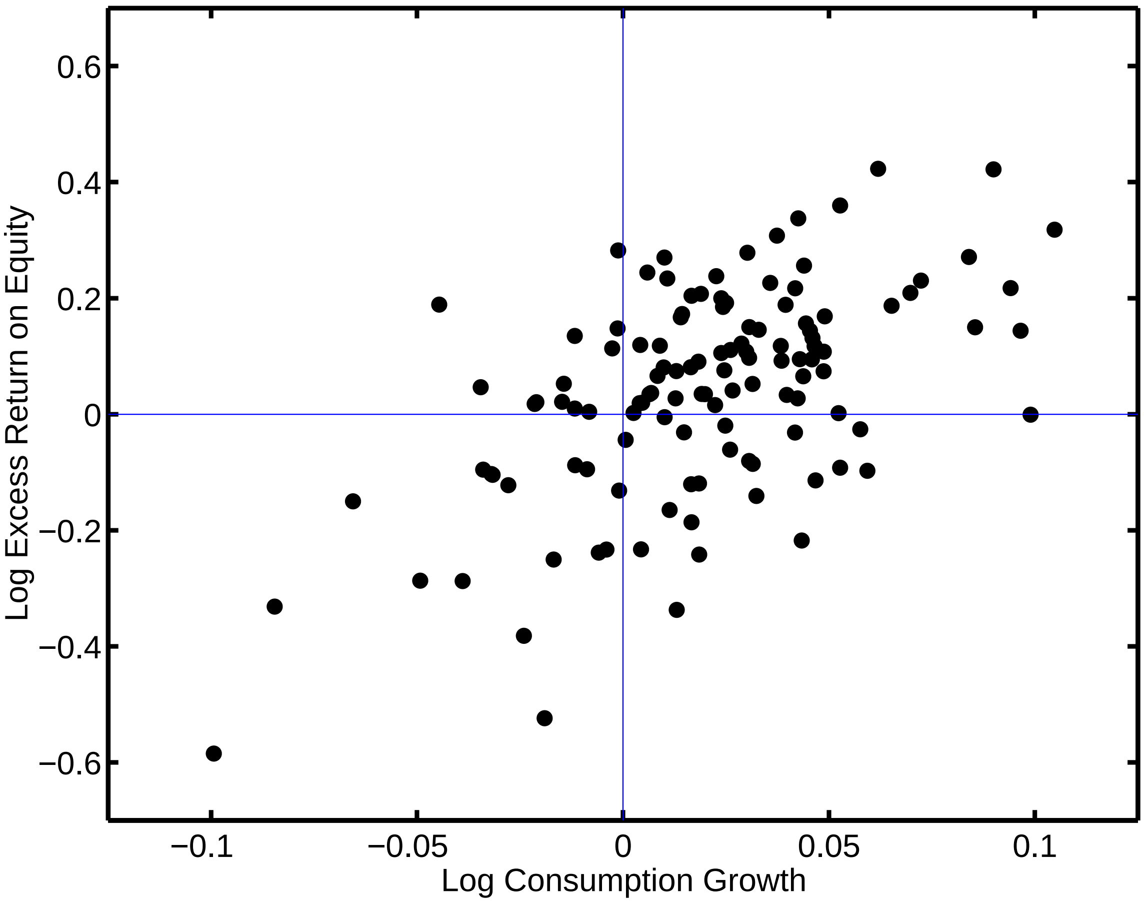
Notes. The lines represent implied volatility "smiles" for the Merton model with three different sets of parameters. The top line is the one we saw in Figure 6 and uses parameters estimated from option prices. The bottom line uses parameters estimated from consumption data, extrapolated to equity returns using equation (22). The middle line is based on a bivariate model of consumption growth and equity returns with correlation chosen to match US data.

Figure 9
Option model: implied risk aversion v. equity return



Notes. The line represents implied risk aversion RA for the option model, defined in equation (28).

Figure 10
Equity returns and consumption growth in US data



Notes. The dots represent annual observations of the log excess return on equity and the log of per capita consumption growth over the period 1889-2009. The data come from Shiller (2009), updated as needed.

A Appendix

A.1 The entropy bound

The entropy bound (3) is derived by Alvarez and Jermann (2005) as a byproduct of their Proposition 2. Bansal and Lehmann (1997, Section 2.3) have a similar result that treats variation in the short rate differently [the term $L(q^1)$ in (31) below]. We derive the bound like this:

- Bound on mean log return. Since log is a concave function, Jensen's inequality and the unconditional version of the pricing relation (1) imply that for any positive return r,

  $$E \log m + E \log r \;\; \leq \;\; \log(1) \;=\; 0,$$

  with equality if and only if $mr = 1$. Therefore no asset has higher expected (log) return than the inverse of the pricing kernel:

  $$E \log r \;\; \leq \;\; -E \log m. \tag{29}$$

  The asset with this return is sometimes called the "growth optimal portfolio." We call it the "high-return asset."

- Short rate. A one-period (risk-free) bond has price $q_t^1 = E_t m_{t+1}$, so its return is $r_{t+1}^1 = 1/E_t m_{t+1}$.

- Entropy of the one-period bond price. With the bound in mind, our next step is to express $E \log r^1$ in terms of unconditional moments. The entropy of the one-period bond price does the trick:

  $$L(q^1) \;\; = \;\; \log E q^1 - E \log q^1 \;\; = \;\; \log E m + E \log r^1. \tag{30}$$

- Entropy bound. (29) and (30) imply

  $$L(m) \;\; \geq \;\; E\left(\log r^j - \log r^1\right) + L(q^1). \tag{31}$$

  Inequality (3) follows from $L(q^1) \geq 0$ (entropy is nonnegative). In practice, $L(q^1)$ is small; in the iid case, it's zero.

We find the loglinear perspective of the entropy bound convenient, but the familiar Hansen-Jagannathan bound also depends (implicitly) on high-order cumulants of $\log m$. The bound is

$$\mathrm{Var}(m)^{1/2}/Em \;\; \geq \;\; E\left(r^j - r^1\right)/\mathrm{Var}(r^j - r^1)^{1/2} \;\; = \;\; \mathrm{SR},$$

where SR is the Sharpe ratio. If $k(s)$ is the cumulant generating function for $\log m$, the bound depends on

$$\begin{aligned} Em &= E\left(e^{\log m}\right) &= e^{k(1)} \\ \mathrm{Var}(m) &= E(m^2) - (Em)^2 &= e^{k(2)} - e^{2k(1)}, \end{aligned}$$

Since $k(1)$ and $k(2)$ involve high-order cumulants of $\log m$, the bound does, too. The squared Sharpe ratio is bounded below by

$$\mathrm{Var}(m)/E(m)^2 \quad = \quad e^{k(2)-2k(1)} - 1.$$

If the cumulants are small (true for a small enough time interval), this is approximately $k(2) - 2k(1)$. Expressed in similar form, entropy is $k(1) - k'(0)$.

A.2 Entropy and cumulants of Bernoulli random variables

We derive the entropy and cumulants of a Bernoulli random variable, as in Section 3. Let z take on the values 0 and 1 with probabilities $1-\omega$ and ω. Entropy follows from its definition (2):

$$L(e^z) \quad = \quad \log\left(1 - \omega + \omega e^1\right) - \omega.$$

Cumulants can be used to quantify the contribution of specific terms. The cumulant-generating function for w is

$$k(s) \quad = \quad \log E e^{sz} \quad = \quad \log(1 - \omega + \omega e^s).$$

Cumulants are derivatives evaluated at $s = 0$: $\kappa_j = k^{(j)}(0)$. The derivatives

$$\begin{aligned} k^{(1)}(s) &= e^{-k(s)}\omega e^s \\ k^{(2)}(s) &= k^{(1)}(s)[1 - k^{(1)}(s)] \\ k^{(3)}(s) &= k^{(2)}(s)[1 - 2k^{(1)}(s)] \\ k^{(4)}(s) &= k^{(3)}(s)[1 - 2k^{(1)}(s)] - 2[k^{(2)}(s)]^2 \\ k^{(5)}(s) &= k^{(4)}(s)[1 - 2k^{(1)}(s)] - 6k^{(2)}(s)k^{(3)}(s) \end{aligned}$$

imply the cumulants

$$\begin{aligned} \kappa_1 &= \omega \\ \kappa_2 &= \kappa_1(1-\kappa_1) \quad = \quad \omega(1-\omega) \\ \kappa_3 &= \kappa_1(1-\kappa_1)(1-2\kappa_1) \quad = \quad \omega(1-\omega)(1-2\omega) \\ \kappa_4 &= \kappa_3(1-2\kappa_1) - 2(\kappa_2)^2 \quad = \quad \omega(1-\omega)(6\omega^2 - 6\omega + 1) \\ \kappa_5 &= \kappa_4(1-2\kappa_1) - 6\kappa_2\kappa_3 \quad = \quad \omega(1-\omega)(1-2\omega)(12\omega^2 - 12\omega + 1). \end{aligned}$$

It's evident that odd moments come from $\omega \neq 1/2$. The example in Section 3 is the same random variable multiplied by θ.

A.3 Entropy and cumulants of Poisson-normal mixtures

We'll look at a Poisson-normal mixture shortly, but it's useful to start with a Poisson random variable z that equals j with probability $e^{-\omega}\omega^j/j!$ for $j = 0, 1, 2, \ldots$. Recall that the power series representation of the exponential function is

$$e^{\omega} \;=\; \sum_{j=0}^{\infty} \omega^j/j!.$$

From this we see that the probabilities sum to one. The moment-generating function is

$$h(s) \;=\; \sum_{j=0}^{\infty} e^{-\omega}\omega^j/j! e^{sj} \;=\; \sum_{j=0}^{\infty} e^{-\omega}(\omega e^s)^j/j! \;=\; \exp[\omega(e^s - 1)].$$

The cumulant-generating function is therefore

$$k(s) \;=\; \log h(s) \;=\; \omega(e^s - 1).$$

Cumulants follow by differentiating.

The Poisson-normal mixture has a similar structure. Conditional on j, z is normal with mean $j\theta$ and variance $j\delta^2$. The conditional moment-generating function is $\exp[(s\theta + s^2\delta^2/2)j]$. The mgf for the mixture is the probability-weighted average,

$$h(s) \;=\; \sum_{j=0}^{\infty} e^{-\omega}\omega^j/j! \exp[(s\theta + s^2\delta^2/2)j] \;=\; \exp\left(\omega[e^{s\theta+(s\delta)^2/2} - 1]\right),$$

which implies the cgf

$$k(s) \;=\; \omega\left[e^{s\theta+(s\delta)^2/2} - 1\right].$$

The same approach can be used for jumps with other distributions. If we set $\theta = 1$ and $\delta = 0$, we get the cgf of the original Poisson.

We find cumulants by taking derivatives of k. The first five are

$$\begin{aligned}
\kappa_1 &= \omega\theta \\
\kappa_2 &= \omega(\theta^2 + \delta^2) \\
\kappa_3 &= \omega\theta(\theta^2 + 3\delta^2) \\
\kappa_4 &= \omega(\theta^4 + 6\theta^2\delta^2 + 3\delta^4) \\
\kappa_5 &= \omega\theta(\theta^4 + 10\theta^2\delta^2 + 15\delta^4).
\end{aligned}$$

Here you can see that the sign of the odd moments is governed by the sign of θ. Negative odd cumulants evidently require $\theta < 0$.

A.4 Equity premium

Most of our analysis is loglinear, which allows us to express asset prices and returns as functions of cumulant-generating functions of (say) the log of consumption growth. The notation is wonderfully compact. The idea and many of the results follow Martin (2009).

Let's start with the short rate. A one-period risk-free bond sells at price $q_t^1 = E_t m_{t+1}$ and has return $r_{t+1}^1 = 1/q_t^1 = 1/E_t m_{t+1}$. In the iid case, the short rate is constant and equals

$$\begin{aligned} \log r^1 &= -\log E(m) \\ &= -\log \beta - \log E\left(e^{-\alpha \log g}\right) \;=\; -\log\beta - k(-\alpha; \log g). \end{aligned}$$

The second equality is based on the definition of the pricing kernel, equation (13). The last one follows from the definition of the cumulant-generating function k, equation (4).

We now turn to equity, defined as a claim to a dividend process $d_t = c_t^\lambda$. If the price-dividend ratio on this claim is q^e, the return is

$$r_{t+1}^e = g_{t+1}^\lambda (1 + q_{t+1}^e)/q_t^e.$$

In the iid case, q^e is constant. The pricing relation (1) and our power utility pricing kernel (13) then imply

$$q^e/(1+q^e) = E\left(\beta g^{\lambda-\alpha}\right) = \beta E\left(e^{(\lambda-\alpha)\log g}\right).$$

Thus we have, in compact notation,

$$\begin{aligned} \log\left[q^e/(1+q^e)\right] &= \log\beta + k(\lambda-\alpha; \log g) \\ \log r_{t+1}^e &= \lambda \log g_{t+1} - \log\beta - k(\lambda-\alpha;\log g) \\ \log r_{t+1}^1 &= -\log\beta - k(-\alpha;\log g) \\ \log r_{t+1}^e - \log r_{t+1}^1 &= \lambda\log g_{t+1} + k(-\alpha;\log g) - k(\lambda-\alpha;\log g). \end{aligned}$$

The equity premium is therefore

$$\begin{aligned} E\left(\log r_{t+1}^e - \log r_{t+1}^1\right) &= \lambda\kappa_1(\log g) + k(-\alpha;\log g) - k(\lambda-\alpha;\log g) \\ &= L(e^{-\alpha\log g}) - L(e^{(\lambda-\alpha)\log g}) \\ &= \sum_{j=2}^{\infty} \kappa_j(\log g)[(-\alpha)^j - (\lambda-\alpha)^j]/j!. \end{aligned}$$

The second line follows because the first-order cumulants cancel. The third is the usual cumulant expansion of entropy. They tell us that the equity premium is the entropy of the pricing kernel minus a penalty (entropy must be positive). It hits its maximum when $\lambda = \alpha$, in which case equity is the high return asset.

A.5 Risk neutral distributions with power utility

A similar approach reveals the connection between true and risk-neutral cumulants of log consumption growth $\log g = w$ (w because it's easier to type). The cumulant generating function for the true distribution is

$$k(s) \;=\; \log E\left(e^{sw}\right).$$

The pricing kernel is $m(w) = \beta e^{-\alpha w}$, which implies $q^1 = \beta k(-\alpha)$. Risk-neutral probabilities are $p^*(w) = p(w)m(w)/q^1 = p(w)e^{-\alpha w}/k(-\alpha)$. The cumulant generating function is therefore

$$k^*(s) \;=\; k(s-\alpha) - k(-\alpha). \tag{32}$$

This is a standard math result. We find its cumulants by differentiating:

$$\kappa_n^* \;=\; \sum_{j=0}^{\infty} \kappa_{n+j}(-\alpha)^j/j!.$$

Thus risk-neutral cumulants depend on higher-order true cumulants. Positive excess kurtosis, for example, reduces risk-neutral skewness.

A.6 Risk-neutral option pricing

We review option pricing in the Merton model, starting with its primary ingredient, the Black-Scholes-Merton formula. For convenience, we define options on returns rather than prices and drop the time subscripts. All of the parameters in what follows refer to the risk-neutral distribution.

Let the risk-neutral distribution of the return on an arbitrary asset be lognormal: $\log r \sim \mathcal{N}(\log r^1 + \kappa_1, \kappa_2)$. The pricing relation (8) implies the restriction $\kappa_1 + \kappa_2/2 = 0$, which we'll hold in reserve. The BSM formula is the solution to

$$q^p \;=\; q^1 E^*(b-r)^+.$$

The implicit integral on the right includes the terms

$$\begin{aligned} q^1 b \,\text{Prob}(r \le b) \;&=\; q^1 b N(d) \\ -q^1 E^*(r | r \le b) \;&=\; -q^1 \int_{-\infty}^{\log b} e^{\log r}(2\pi\kappa_2)^{-1/2} \exp[-(\log r - \log r^1 - \kappa_1)^2/2\kappa_2]\, d\log r \\ &=\; -\exp(\kappa_1 + \kappa_2/2) N(d - \kappa_2^{1/2}), \end{aligned}$$

with

$$d \;=\; (\log b - \log r^1 - \kappa_1)/\kappa_2^{1/2}$$

and N is the standard normal cdf. We use this to define the function

$$q^p(b) \;=\; f(b;\kappa_1,\kappa_2) \;=\; q^1 bN(d) - \exp(\kappa_1+\kappa_2/2)N(d-\kappa_2^{1/2}). \qquad (33)$$

In the conventional BSM formula, we set $\kappa_1 + \kappa_2/2 = 0$ and simplify, but this version is more useful in what follows.

The Merton model with normal jumps is a Poisson-weighted average of BSM option prices. The model is described in Section 3.2 and has (risk-neutral) parameters $(\mu, \sigma, \omega, \theta, \delta)$. The first two pertain to the normal component, the remainder to the Poisson-normal mixture. Option prices in this setting are

$$q^p(b) \;=\; \sum_{j=0}^{\infty}(e^{-\omega}\omega^j/j!)f(b;\kappa_{1j},\kappa_{2j})$$

with $\kappa_{1j} = \mu + j\theta$ and $\kappa_{2j} = \sigma^2 + j\delta^2$.

A.7 Two scaling issues

Two scaling issues come up in the paper. The first is the relation between equity returns and consumption growth: for most of the paper, log equity returns are a linear function of log consumption growth. The second is time: option prices for intervals other than one year depend on the distribution of returns over other time intervals.

Consider the relation between the distributions of x and λx for some scale factor λ. Consumption growth and equity returns have this structure if we ignore intercepts [equation (22)]. The general result follows from this property of cgfs: $k(s;\lambda x) = k(\lambda s; x)$. If x (think log consumption growth) has the Poisson-normal structure of Section 3.2, its cgf is

$$k(s;x) \;=\; \mu s + (\sigma s)^2/2 + \omega[e^{\theta s + (\delta s)^2/2} - 1]. \qquad (34)$$

The cgf for λx (think excess returns) is therefore

$$k(s;\lambda x) \;=\; \mu\lambda s + (\sigma\lambda s)^2/2 + \omega[e^{\theta\lambda s + (\delta\lambda s)^2/2} - 1].$$

This has the same form with $(\mu, \sigma, \omega, \theta, \delta)$ replaced by $(\lambda\mu, \lambda\sigma, \omega, \lambda\theta, \lambda\delta)$. A similar result applies to the relation between the true distribution of $x = \log g$ and the risk-neutral distribution of λx with power utility. Given (32), their cgfs are connected by

$$k^*(s;\lambda x) \;=\; k^*(\lambda s;x) \;=\; k(\lambda s - \alpha; x) - k(-\alpha; x).$$

In words, we compute the cgf of the risk-neutral distribution of λx by, first, computing the cgf of the risk-neutral distribution of x and, second, scaling by λ. It's important the steps be done in that order.

The second issue concerns the time interval. In an iid setting, suppose the cgf (34) applies to the distribution over a unit time interval. The cgf for an arbitrary time interval $\tau > 0$, if it exists, is the cgf for a time interval of one multiplied by τ. In the Poisson-normal case, we have

$$k(s;\tau) \;=\; \tau\mu s + \tau(\sigma s)^2/2 + \tau\omega[e^{\theta s + (\delta s)^2/2} - 1].$$

The cgf has the same form as (34) with $(\mu, \sigma^2, \omega, \theta, \delta^2)$ replaced by $(\tau\mu, \tau\sigma^2, \tau\omega, \theta, \delta^2)$.

A.8 Cumulant-generating functions based on true and risk-neutral probabilities

We derive the salient features of models in which the true and risk-neutral distributions are Poisson mixtures of normals with different parameters.

We start with a normal example that serves as a component of the Poisson mixture. Let the log return follow (26), where $z = 0$ and w has true distribution of $\mathcal{N}(\mu, \sigma^2)$ and risk-neutral distribution $\mathcal{N}(\mu^*, \sigma^{*2})$. The density functions are

$$\begin{aligned} p(w) &= (2\pi\sigma^2)^{-1/2}\exp[-(w-\mu)^2/2\sigma^2] \\ p^*(w) &= (2\pi\sigma^{*2})^{-1/2}\exp[-(w-\mu^*)^2/2\sigma^{*2}]. \end{aligned}$$

This differs from the examples in Section 4 in allowing the variance to differ between the two distributions. In continuous time, $\sigma^* = \sigma$ is needed to assure absolute continuity of the true and risk-neutral probability measures with respect to each other. In discrete time, there is no such requirement; see, for example, Buhlmann, Delbaen, Elbrechts, and Shiryaev (1996). The risk-neutral pricing relation (8) implies $\mu^* + \sigma^{*2}/2 = 0$.

We can derive all of the relevant properties from these inputs. The log probability ratio is

$$\log[p^*(w)/p(w)] \;=\; (1/2)\log\varphi + [(w-\mu)^2 - \varphi(w-\mu^*)^2]/2\sigma^2,$$

where $\varphi = \sigma^2/\sigma^{*2} > 0$. The moment-generating function of the log probability ratio is

$$\begin{aligned} h(s;\log p^*/p) &= E\left(e^{s\log p^*/p}\right) \\ &= \int_{-\infty}^{\infty} p^*(w)^s p(w)^{1-s} dw \\ &= (2\pi\sigma^2)^{-1/2}\varphi^{s/2}\int_{-\infty}^{\infty}\exp\{-[(1-s)(w-\mu)^2 + s\varphi^2(w-\mu^*)^2]/2\sigma^2\}dw \\ &= \varphi^{s/2}[1-s(1-\varphi)]^{-1/2}\exp\left(\frac{s(s-1)(\mu^*-\mu)^2}{2\sigma^{*2}[1-s(1-\varphi)]}\right) \end{aligned}$$

for $1 - s(1-\varphi) > 0$ (automatically satisfied if $s = 0$ or $s = 1$). The last line follows from completing the square. Thus the cumulant-generating function is

$$k(s;\log p^*/p) \;=\; (s/2)\log\varphi - (1/2)\log[1-s(1-\varphi)] + \left(\frac{s(s-1)(\mu^*-\mu)^2}{2\sigma^{*2}[1-s(1-\varphi)]}\right).$$

Entropy is minus the first derivative evaluated at zero:

$$-\kappa_1(\log p^*/p) \quad = \quad (1/2)[\log \varphi + 1 - \varphi] + (\mu - \mu^*)^2/2\sigma^{*2}. \qquad (35)$$

If $\varphi = 1$ ($\sigma^* = \sigma$), we have

$$k(s; \log p^*/p) \quad = \quad s(s-1)(\mu^* - \mu)^2/2\sigma^2,$$

and the only nonzero cumulants are the first two. Otherwise, high-order cumulants are generally nonzero.

Now let's ignore the normal component and focus on z. Both the true and risk-neutral distributions have Poisson arrivals and normal jumps, but the parameters differ. Conditional on a number of jumps j, the density functions are

$$\begin{aligned} p(z|j) &= e^{-\omega}\omega^j/j! \cdot (2\pi j\delta^2)^{-1/2} \exp[-(z_j - j\theta)^2/(2j\delta^2)] \\ p^*(z|j) &= e^{-\omega^*}\omega^{*j}/j! \cdot (2\pi j\delta^{*2})^{-1/2} \exp[-(z_j - j\theta^*)^2/(2j\delta^{*2})]. \end{aligned}$$

The moment generating function for $\log p^*/p$ is

$$h(s; \log p^*/p) \quad = \quad \sum_{j=0}^{\infty} e^{-\omega}\omega^j/j! \left[e^{s(\omega-\omega^*)+js\log(\omega^*/\omega)} h(s;z)^j \right].$$

Using (35) we have

$$h(s;z) \quad = \quad \varphi^{s/2}[1 - s(1-\varphi)]^{-1/2} \exp\left(\frac{s(s-1)(\theta^* - \theta)^2}{2\delta^{*2}[1 - s(1-\varphi)]} \right),$$

where $\varphi = \delta^2/\delta^{*2}$. Therefore the cumulant-generating function is

$$\begin{aligned} k(s; \log p^*/p) &= s(\omega - \omega^*) \\ &+ \omega \left[(\omega^*/\omega)^s \varphi^{s/2} [1 - s(1-\varphi)]^{-1/2} \exp\left(\frac{s(s-1)(\theta^* - \theta)^2}{2\delta^{*2}[1 - s(1-\varphi)]} \right) - 1 \right]. \end{aligned}$$

Entropy is minus the first derivative evaluated at zero:

$$\begin{aligned} - & \kappa_1(\log p^*/p) \qquad (36) \\ = & (\omega^* - \omega) + \omega[\log(\omega/\omega^*) - 1/2 \cdot \log \varphi + 1/2 \cdot (\varphi - 1)] + \omega(\theta - \theta^*)^2/2\delta^{*2}. \end{aligned}$$

Because the normal and Poisson mixture components are independent, their cumulant-generating functions are additive. Therefore, the entropy for the full model is the sum of the entropy of the normal case [equation (35) with $\varphi = 1$] and the entropy of the Poisson mixture of normals [equation (36)].

A.9 Risk aversion implied by the Merton model

We define the implied risk aversion of an asset pricing model by

$$\text{RA} \quad = \quad -\frac{\partial \log(p^*/p)}{\partial \log g} \quad = \quad -\frac{\partial \log(p^*/p)}{\partial \log r^e} \cdot \frac{\partial \log r^e}{\partial \log g} \quad = \quad \left(\frac{1}{p} \cdot \frac{\partial p}{\partial \log r^e} - \frac{1}{p^*} \cdot \frac{\partial p^*}{\partial \log r^e} \right) \lambda.$$

We compute this for the Poisson-normal mixture used in Section 5. The probability of j jumps is

$$p(j) \quad = \quad e^{-\omega} \omega^j / j!.$$

Conditional on j, the density for the log return is

$$p(\log r^e | j) \quad = \quad [2\pi(\sigma^2 + j\delta^2)]^{-1/2} \exp\left\{ -(\log r^e - \mu - j\theta)^2 / [2(\sigma^2 + j\delta^2)] \right\}.$$

The marginal density for log returns is therefore

$$p(\log r^e) \quad = \quad \sum_{j=0}^{\infty} p(j) p(\log r^e | j).$$

Its derivative is

$$\frac{\partial p(\log r^e)}{\partial \log r^e} \quad = \quad \sum_{j=0}^{\infty} p(j) \frac{\partial p(\log r^e | j)}{\partial \log r^e} \quad = \quad -\sum_{j=0}^{\infty} p(j) p(\log r^e | j)(\log r^e - \mu - j\theta)/(\sigma^2 + j\delta^2).$$

A similar expression holds for the risk-neutral distribution. Risk aversion is therefore

$$\begin{aligned} \text{RA}/\lambda \quad = \quad & p^*(\log r^e)^{-1} \sum_{j=0}^{\infty} p^*(j) p^*(\log r^e | j)(\log r^e - \mu^* - j\theta^*)/(\sigma^2 + j\delta^{*2}) \\ & - p(\log r^e)^{-1} \sum_{j=0}^{\infty} p(j) p(\log r^e | j)(\log r^e - \mu - j\theta)/(\sigma^2 + j\delta^2). \qquad (37) \end{aligned}$$

This is a function, in general, of $\log r^e$.

A.10 Risk-neutral distribution of returns in a bivariate model

Consider a bivariate model in which equity returns are (potentially) less closely tied to consumption growth. Let $x' = (x_1, x_2) = (\log g, \log r^e - \log r^1)$ have the two-component structure used throughout the paper:

$$x_{t+1} \quad = \quad w_{t+1} + z_{t+1}.$$

The first component is bivariate normal: $w_t \sim \mathcal{N}(\mu, \Sigma)$ where μ and Σ have elements μ_i and σ_{ij}, respectively. The second component is a Poisson mixture of bivariate normals. As in Section 3.2, jumps occur with Poisson intensity ω so that the probability of j jumps is $e^{-\omega}\omega^j/j!$. Each jump adds a draw from the bivariate normal distribution $\mathcal{N}(\theta, \Delta)$.

Option pricing (indeed equity pricing) requires us to deal with the bivariate distribution of equity returns and the pricing kernel. We derive the risk-neutral distribution of returns from its cumulant-generating function. The (joint) cgf is

$$k(s) \;=\; E(e^{s'x}) \;=\; s'\mu + s'\Sigma s/2 + \omega\left[e^{s'\theta + s'\Delta s/2} - 1\right].$$

The logic here is virtually identical to the univariate case outlined in Appendix A.3. Derivatives of this expression lead to these formulas:

$$\begin{aligned} \text{Var}(x_i) &= \sigma_{ii} + \omega(\theta_i^2 + \delta_{ii}) \\ \text{Cov}(x_1, x_2) &= \sigma_{12} + \omega(\theta_1\theta_2 + \delta_{12}). \end{aligned}$$

The correlation is

$$\begin{aligned} \text{Corr}(x_1, x_2) &= \frac{\sigma_{12} + \omega(\theta_1\theta_2 + \delta_{12})}{[\sigma_{11} + \omega(\theta_1^2 + \delta_{11})]^{1/2}[\sigma_{22} + \omega(\theta_2^2 + \delta_{22})]^{1/2}} \\ &= \frac{\rho_w(\sigma_{11}\sigma_{22})^{1/2} + \omega[\theta_1\theta_2 + \rho_z(\delta_{11}\delta_{22})^{1/2}]}{[\sigma_{11} + \omega(\theta_1^2 + \delta_{11})]^{1/2}[\sigma_{22} + \omega(\theta_2^2 + \delta_{22})]^{1/2}} \end{aligned}$$

with $\rho_w = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$ and $\rho_z = \delta_{12}/(\delta_{11}\delta_{22})^{1/2}$.

The remaining step is to find the risk-neutral distribution for returns. With power utility, the cgf corresponding to the risk-neutral distribution is

$$k^*(s_1, s_2) \;=\; k(s_1 - \alpha, s_2) - k(-\alpha, 0).$$

The logic is analogous to Appendix A.5. The cgfs corresponding to marginal distributions follow from setting the other elements of s equal to zero. (This follows from the definitions of the marginal distribution and cgf.) Thus the cgf for x_2 is $k(0, s_2)$. The risk-neutral cgf for the log equity excess return is therefore

$$\begin{aligned} k^*(0, s_2) &= s_2(\mu_2 - \alpha\sigma_{12}) + s_2^2\sigma_{22}/2 + \omega e^{-\alpha\theta_1 + \alpha^2\delta_{11}/2}\left[e^{s_2(\theta_2 - \alpha\delta_{12}) + s_2^2\delta_{22}/2} - 1\right] \\ &= s_2\mu_2^* + s_2^2\sigma_{22}/2 + \omega^*\left[e^{s_2\theta_2^* + s_2^2\delta_{22}/2} - 1\right] \end{aligned}$$

with the implicit definitions

$$\mu_2^* \;=\; \mu_2 - \alpha\sigma_{12}, \;\; \omega^* \;=\; \omega e^{-\alpha\theta_1 + \alpha^2\delta_{11}/2}, \;\; \theta_2^* \;=\; \theta_2 - \alpha\delta_{12}.$$

This has the same form as the Merton model with suitably defined parameters. Options can therefore be priced using the methods of Appendix A.6.