THE JOURNAL OF FINANCE • VOL. LXVI, NO. 6 • DECEMBER 2011

Disasters Implied by Equity Index Options

DAVID BACKUS, MIKHAIL CHERNOV, and IAN MARTIN*

ABSTRACT

We use equity index options to quantify the distribution of consumption growth disasters. The challenge lies in connecting the risk-neutral distribution of equity returns implied by options to the true distribution of consumption growth. First, we compare pricing kernels constructed from macro-finance and option-pricing models. Second, we compare option prices derived from a macro-finance model to those we observe. Third, we compare the distribution of consumption growth derived from option prices using a macro-finance model to estimates based on macroeconomic data. All three perspectives suggest that options imply smaller probabilities of extreme outcomes than have been estimated from macroeconomic data.

THE FIELD OF MACRO-FINANCE offers an attractive opportunity to explore links between asset returns and the real economy. Many of the most influential papers in the field do just that, most commonly by using macroeconomic data as inputs to models designed to account for asset returns. There is also promise in doing the reverse, that is, in using properties of asset returns to characterize macroeconomic risk. The growing literature on disaster risk is a clear example of where macroeconomic research might benefit from greater input from finance.

Barro (2006), Longstaff and Piazzesi (2004), and Rietz (1988) show that macroeconomic disasters—large declines in aggregate consumption growth—produce dramatic improvement in the ability of representative agent models to reproduce prominent features of U.S. asset returns, including the equity premium. They disagree, however, about their distribution. Rietz (1988) simply chooses an arbitrary distribution and illustrates its impact. Julliard and Ghosh (2008) and Longstaff and Piazzesi (2004) argue that a distribution based on U.S. experience cannot match the equity premium with modest degrees of risk aversion. Barro (2006), Barro and Ursua (2008), and Barro et al. (2009)

*Backus is from New York University and NBER, Chernov is from the London School of Economics and CEPR, and Martin is from Stanford University and NBER. 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, CEPR (Gerzensee meetings), 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 earlier versions.

study broader collections of countries, which in principle can tell us about alternative histories the United States might have experienced. In their panel of international macroeconomic data, the frequency and magnitude of disasters are significantly larger than those we have seen in U.S. history. Their estimated distribution of consumption disasters, which we refer to as the international macroeconomic evidence, has become the industry standard in both macroeconomic and finance research.

The issue, then, is the distribution of extreme negative outcomes, that is, the shape of the left tail of the probability distribution of consumption growth. In virtually all of this research, the distribution is modeled by combining a normal component with a jump component. A jump component, in this context, is simply a mathematical device that produces nonnormal distributions. The debate in the literature concerns the parameter values of this component.

The range of opinion about jump parameters reflects the nature of the problem: they are not easily estimated from the relatively short history of U.S. macroeconomic data. We follow a different but complementary path, estimating them from prices of equity index options. 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. By looking at options with different strike prices, we learn more about the distribution of equity index returns than we do from returns alone. The cyclical behavior of equity returns and the range of strike prices make this class of assets a natural choice. Other cyclical assets, including corporate and government bonds, might be studied in future work.

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. The most salient differences for our purposes are as follows: (i) The macro-finance literature is concerned with properties of consumption growth, while the option-pricing literature is concerned with equity returns. (ii) Macro-finance is concerned with the true or objective probability distribution, while option prices characterize the risk-neutral distribution. (iii) The relation between true and risk-neutral probability distributions is different in the two literatures. In the macro-finance literature, the risk-neutral distribution is derived from the true distribution of consumption growth and the preferences of a representative agent (power utility, for example). In the option-pricing literature, the two distributions are specified separately, which leads to a different structure for the pricing kernel. These differences pose nontrivial challenges to anyone attempting to use evidence from equity options to quantify features of consumption growth.

Our objective is to compare models based on consumption and option data, respectively, but the differences in approach lead us to compare the models from several different perspectives. Each comparison uses a different analytical structure and focuses on a different aspect of the model. In our first comparison, we focus on pricing kernels, which exist in any arbitrage-free model. In macro-finance models, the properties of pricing kernels are closely related to those of consumption growth. A natural first step then is to compare the pricing kernel generated by a representative agent model calibrated to international macroeconomic data to one implied by an estimated option-pricing model. A second comparison focuses on option prices. We follow the macro-finance route, using macroeconomic data and the preferences of a representative agent to value options as well as equity, and compare these option prices to those of an estimated option-pricing model. Here and elsewhere, we report option prices as implied volatility smiles, which represent departures from normality in an intuitive way.

A third comparison delivers on our goal of estimating the distribution of consumption growth disasters from options: we compare the distribution estimated from international macroeconomic data with one derived from option prices. Here we reverse the procedure of the previous comparison, using our estimated implied volatility smile and the preferences of a representative agent to compute the consumption distribution from option prices. By construction, the distribution is consistent with both option prices and the equity premium. This approach differs from the substantial body of work that infers risk aversion from option prices. See, for example, Aït-Sahalia and Lo (2000), Jackwerth (2000), Rosenberg and Engle (2002), and Ziegler (2007). In that literature, preferences are defined directly over wealth, which is proxied by equity, so the distribution of consumption is not studied.

In our comparison of pricing kernels, we find the concept of entropy (a measure of dispersion) useful. Alvarez and Jermann (2005) and Bansal and Lehmann (1997) show that the entropy of the pricing kernel of a given model is the maximum risk premium, defined as the mean excess log return. Entropy can be represented as an infinite sum of the cumulants (close relatives of moments) of the log of the pricing kernel. If the log of the pricing kernel is normal, entropy is proportional to its variance. Departures from lognormality such as those generated by jumps can increase cumulants, and hence entropy, thereby improving a model's ability to account for observed excess returns. Odd cumulants, in particular, reflect the asymmetry inherent in disaster research. Despite the difference in functional forms, entropy can be computed for both the macro-finance and option-pricing models. We find that, while both consumption- and option-based models generate substantial contributions to entropy from odd cumulants, the relative contribution is much smaller in the model based on option prices. We also find that the contribution of the variance is much higher in the option model, which results in greater entropy overall.

In our comparison of option prices, we start with a consumption distribution based on macroeconomic assessments of jumps, link dividends to consumption, price assets with a representative agent with power utility, select risk aversion to match the equity premium, and use the model to value options on equity. We then compare implied volatility smiles derived this way from consumption data to those from an option model estimated by Broadie, Chernov, and Johannes (2007). We find that the volatility smile is much steeper in the consumption-based model.

In our comparison of consumption growth distributions we do the reverse. We take estimates of the risk-neutral jump parameters from Broadie et al.'s (2007) study of equity index options. This ensures that the model matches the implied volatility smile. As before, we use a representative agent with power utility and choose risk aversion to match the equity premium. This allows us to infer the true distribution of consumption growth implied by option prices. The consumption growth distribution implied by option prices agrees with consumption-based estimates of the probability of modest disasters: consumption growth more than three standard deviations to the left of the mean. In this respect, the distribution is similar to U.S. data. (The Great Depression includes one year in which consumption declined by slightly more than three standard deviations). However, the distribution differs significantly for more extreme events: the probability of consumption growth more than five standard deviations to the left of its mean is much smaller than estimates based on international macroeconomic evidence. The parameter values derived from options match not only the equity premium, as in the Barro calibration, but also option prices and moments of consumption growth. We think these parameter values provide a more realistic distribution of disasters in U.S. data than the international macroeconomic evidence used in the literature.

The second and third comparisons are based on a model in which consumption growth and equity returns are perfectly correlated. There is a long tradition of similar models in the macro-finance literature, but the correlation in U.S. data is 0.57. We therefore consider a bivariate model in which the correlation can be set to match the evidence. If we choose parameters based on international macroeconomic evidence, the volatility smile for option prices remains much steeper than we see in the model based on option prices. Conversely, if we choose parameters to match option prices, the probability of large negative realizations of consumption growth (more than five standard deviations to the left of the mean) remains much smaller than suggested by the macroeconomic evidence. It nevertheless matches basic features of equity returns and option prices.

We conclude that option prices imply smaller probabilities of macroeconomic disasters than suggested by Barro and his coauthors. Nevertheless, we would not say that we reject "the Barro model." We would say instead that the model makes a useful point about the role of asymmetries and other departures from normality in asset pricing. Certainly there is evidence of both in prices of equity index options.

The paper is organized as follows. Section I presents the initial empirical evidence, distinguishes disasters from jump models, and introduces the main analytical tools used in the paper, that is, cumulant-generating functions, entropy, and their connection to the pricing kernel. Section II illustrates the tools using a macro-based model of disasters. Section III shows how to relate true probabilities to risk-neutral probabilities in the context of such a macro-based model. Section IV builds an option-based model of disasters. Section V compares the macro-based and option-based approaches. Section VI concludes. All the technical material is relegated to eight appendices.

Table I

Properties of U.S. Consumption Growth and Asset Returns

Entries are statistics computed from annual observations for the U.S. 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. 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 1-year bond. The 1-year bond is the Treasury security of maturity closest to 1 year. Equity is the S&P 500. Consumption and return data are from Shiller (2009).

Variable	Mean	Std Dev	Skewness	Kurtosis
(a) Consumption growth and returns, annual, 1889–2009				
Consumption growth	0.0198	0.0350	−0.34	1.11
Return on equity	0.0587	0.1795	−0.61	0.43
Excess return on equity	0.0407	0.1812	−0.72	0.91
(b) Consumption growth and returns, annual, 1986–2009				
Consumption growth	0.0178	0.0150	−0.87	0.66
Return on equity	0.0641	0.1845	−1.24	1.56
Excess return on equity	0.0434	0.1808	−1.39	2.01

I. Preliminaries

We start with an overview of the evidence we are trying to understand and the tools we use to shed light on it. The tools allow us to characterize departures from lognormality, including disasters, in a convenient way.

A. Evidence

We provide a quick overview of U.S. evidence on consumption growth, asset returns, and option prices.

In Table I, we report evidence on annual consumption growth and equity returns (the S&P 500 index) for both a long sample (1889 to 2009) and a shorter one (1986 to 2009) that corresponds approximately to the option data used by Broadie et al. (2007). Similar evidence is summarized by Alvarez and Jermann (2005, Tables I to III), Barro (2006, Table IV), and Mehra and Prescott (1985, Table I). 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 (BSM) formula. These volatilities have two well-documented features that we examine more closely in Section IV. Similar evidence has been reviewed recently by Bates (2008, Section I), Drechsler and Yaron (2008, Section II), 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 BSM. 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 that market participants value adverse events more than would be implied by a lognormal model. The question for us is 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 I 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.

B. Jumps and Disasters

We use a jump model as a way to generate departures from normality and, in particular, disasters. We follow Barro (2006) in using a two-component structure for consumption growth $g_t = c_t/c_{t-1}$,

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

with components (w_t, z_t) that are independent of each other and over time. The first component is normal: $w \sim \mathcal{N}(\mu, \sigma^2)$. The second component, the "jump," is a Poisson mixture of normals. Its central ingredient is a 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:

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

If ω is small, the jump model is well approximated by a Bernoulli mixture of normals. In this case, there is at most one jump per unit of time and it occurs with probability ω. But if ω is large, as it is in the option model of Section IV, there can be a significant probability of multiple jumps.

This functional form comes with a number of benefits. One is that it is a flexible functional form that can approximate a wide range of nonnormal behavior. Another is that it is easily scaled to the different time intervals observed in option markets (it is "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).

If the jump model provides a way to represent departures from normality, cumulants provide a way to quantify their magnitude. Recall that the moment-generating function (if it exists) for a random variable x is defined by

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

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

$$k(s; x) = \log E\left(e^{sx}\right) = \sum_{j=1}^{\infty} \kappa_j(x) s^j / j! \tag{3}$$

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 j^{th} 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:

$$\gamma_1 = \kappa_3/\kappa_2^{3/2}, \quad \gamma_2 = \kappa_4/\kappa_2^2. \tag{4}$$

The normal distribution for w has the quadratic cumulant-generating function $\mu s + \sigma^2 s^2/2$, which implies zero cumulants after the first two. Nonzero high-order cumulants (κ_j for $j \geq 3$) therefore summarize departures from normality. We derive the cumulant-generating function for z in Appendix A. Note for future reference that $k(s; ax) = k(as; x)$ (replace s with as in (3)). Therefore, if x has cumulants κ_j, ax has cumulants $a^j \kappa_j$. This use of the cumulant-generating function was suggested by Martin (2009) and recurs throughout the paper.

Since the components are independent, the cumulant-generating function of $\log g$ is the sum of those for w and z. We find cumulants of $\log g$ by taking derivatives of k. The first four are

$$\kappa_1 = \mu + \omega\theta \tag{5}$$

$$\kappa_2 = \sigma^2 + \omega(\theta^2 + \delta^2) \tag{6}$$

$$\kappa_3 = \omega\theta(\theta^2 + 3\delta^2) \tag{7}$$

$$\kappa_4 = \omega(\theta^4 + 6\theta^2\delta^2 + 3\delta^4). \tag{8}$$

Note that cumulants reflect complex combinations of parameters. Negative skewness, for example, requires $\theta < 0$, but its magnitude depends on ω (governing the probability of jumps), θ (the mean jump), and δ (the dispersion of jumps).

It is important to distinguish between jumps (a modeling tool) and disasters (the left tail of a distribution). To be concrete, we define disasters as negative realizations of consumption growth below a threshold $-b$, for large $b > 0$: $\mathcal{D}_b = \{\log g \le -b\}$. The probability of event $\mathcal{D}_b$ implied by the jump model (1) is

$$p(\mathcal{D}_b) = \sum_{j=0}^{\infty} p(\log g \le -b|j) \cdot p(j) = \sum_{j=0}^{\infty} N\left(\frac{-b-\mu-j\theta}{\sqrt{\sigma^2 + j\delta^2}}\right) \cdot \frac{e^{-\omega}\omega^j}{j!}, \tag{9}$$

where $N(\cdot)$ is a cumulative distribution function of a standard normal variable. This expression tells us that a specific disaster probability can be generated by the model in a number of ways. For example, if ω is small, as in Barro (2006), $p(\mathcal{D}_b)$ is approximately equal to a sum of the first two terms ($j = 0$ and 1). Thus, one needs relatively large values of θ and/or δ to obtain a nonnegligible probability value. In contrast, if ω is large, all the terms in the sum contribute to the $p(\mathcal{D}_b)$. Therefore, jump sizes need not be large to generate the same probability value.

This discussion should not be taken to imply that the jump probability and sizes are interchangeable. Different values imply different high-order cumulants of consumption and different properties of assets that are sensitive to high-order cumulants. This is why we need to use information about financial assets that are sensitive to the probabilities of disasters. High-quality corporate bonds could be one example. Presumably, such bonds default only in the case of a disaster. Assuming risk neutrality and zero recovery, the price of a zero-coupon bond is $1 - p(\mathcal{D}_b)$. Equity is a cyclical asset as well, so digital put options on equity will be informative because their price is simply $p(\mathcal{D}_b)$ under risk neutrality. Moving to less exotic derivatives, out-of-the-money puts will be informative about the probability of a disaster as well. In this paper, we focus on regular puts and calls on equity as our source of additional information about disasters. This asset class possesses a nice combination of high-quality data and reliable models. Thus, we are able to extract information from options in a credible manner.

C. Pricing Kernels, Entropy, and Cumulants

So far, we have been discussing the true probabilities of disasters. However, extracting these probabilities from option prices requires that one thinks about the pricing of risks. The pricing kernel is the central tool for this purpose. In this section, we connect the pricing kernel to a concept of entropy (a measure of dispersion) that we think clarifies how the risks of departures from lognormality are valued.

In any arbitrage-free environment, there is a positive random variable m (the pricing kernel) that satisfies the pricing relation,

$$E_t\left(m_{t+1}r^j_{t+1}\right) = 1, \tag{10}$$

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, and 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

$$L(x) = \log Ex - E\log x. \tag{11}$$

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 will 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,

$$L(m) \geq E(\log r^j - \log r^1), \tag{12}$$

for any asset j with positive returns. See Appendix B. 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 (12) 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 that they tend to generate. We can express the entropy of the pricing kernel in terms of the cumulant-generating function and cumulants of $\log m$:

$$\begin{aligned} L(m) &= \log E(e^{\log m}) - E\log m \\ &= k(1;\log m) - \kappa_1(\log m) = \sum_{j=2}^{\infty}\kappa_j(\log m)/j!. \end{aligned} \tag{13}$$

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 II) 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:

$$k_{\text{odd}}(s) = [k(s) - k(-s)]/2 = \sum_{j=1,3,\ldots} \kappa_j(x) s^j/j!$$

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

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

D. 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 (10) becomes

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

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

$$p^*(x) = p(x)m(x)/Em = p(x)m(x)/q^1. \tag{14}$$

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,

$$q^1 \sum_x p^*(x) r^j(x) = q^1 E^* r^j = 1, \tag{15}$$

where E^* denotes the expectation computed from risk-neutral probabilities. In (10), the pricing kernel performs two roles: discounting and risk adjustment. In (15), 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 (15) gives us its price in terms of risk-neutral probabilities:

$$q^p = q^1 E^*(b - q)^+. \tag{16}$$

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 (14) gives us the pricing kernel:

$$m(x) = q^1 p^*(x)/p(x). \tag{17}$$

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

$$L(m) = L(p^*/p) = \log E(p^*/p) - E \log(p^*/p) = -E \log(p^*/p). \tag{18}$$

The first equality follows because q^1 is constant (recall that $L(ax) = L(x)$). The second follows from the definition of entropy (equation (11)). 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 (18) 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

$$k[s; \log(p^*/p)] = \log E(e^{s \log(p^*/p)}) = \sum_{j=1}^{\infty} \kappa_j [\log(p^*/p)] s^j/j!. \tag{19}$$

The definition of entropy (11) contributes the analog to (13) in which entropy is related to cumulants:

$$\begin{aligned} L(p^*/p) &= k[1; \log(p^*/p)] - \kappa_1[\log(p^*/p)] \\ &= \sum_{j=2}^{\infty} \kappa_j[\log(p^*/p)]/j! = -\kappa_1[\log(p^*/p)]. \end{aligned} \tag{20}$$

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 is matched by an expansion in terms of cumulants two 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 (19).

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 one-period bond price), equation (14) 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.

II. 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 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 (reported in Table I). We think it is a reasonably good approximation for this purpose, but return to the issue briefly in Section V.

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. The pricing kernel is

$$\log m_{t+1} = \log \beta - \alpha \log g_{t+1}. \tag{21}$$

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 and cumulants follow from the two-component process (1) for consumption growth and the pricing kernel (21). Entropy is

$$L(m) = L(e^{-\alpha \log g}) = L(e^{-\alpha w}) + L(e^{-\alpha z}). \tag{22}$$

The entropy of the components follows from its definition (11):

$$L(e^{-\alpha w}) = (-\alpha\sigma)^2/2 \tag{23}$$

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

See Appendix A. The cumulants of $\log m$ are related to those of $\log g$ by

$$\kappa_j(\log m) = \kappa_j(\log g)(-\alpha)^j/j! = (-\alpha)^j \kappa_j(w)/j! + (-\alpha)^j \kappa_j(z)/j! \tag{25}$$

for $j \geq 1$. See Section I.C.

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 can see the quantitative significance of the jump component with numerical examples based on international macroeconomic evidence. Its role is evident in Table II in the difference between column (1), the lognormal case, and column (2), which incorporates a Poisson 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 $\omega = 0.01$, $\theta = -0.3$, and $\delta = 0.15$, which implies a 1% chance of a 30% fall (on average) in consumption growth relative to its mean. Given these values, we adjust the parameters of the normal component to maintain the mean and variance, whose theoretical values are given in (5) and (6).

The parameters of the jump component are derived from studies of international macroeconomic data by Barro (2006), Barro and Ursua (2008), and Barro et al. (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 $\omega = 0.017$, $\theta = -0.38$, and $\delta = 0.25$. Wachter (2009) uses exactly the same specification of jump sizes as Barro (2006). Barro et al. (2009, Section VI.2) estimate a dynamic model, but argue that its asset pricing implications are the same as an iid model with $\omega = 0.0138$

Table II

Parameter Values and Properties of Model Economies

Entries are parameters and properties of the models considered. The labels at the top of the columns describe the model used and the variable on which it is based. Columns (1) to (2) and (4) to (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. We assume that equity is a levered claim on consumption with leverage $\lambda = 5.1$ that is selected to match volatility of returns (0.1800). Risk aversion α is chosen to match the mean equity premium (0.0400). Column (3) is the Merton model parameterized to option prices and equity returns. Columns (4) and (5) take this model and infer true parameters of consumption growth by applying the relations implied by power utility. γ_1 and γ_2 are the measures of skewness and excess kurtosis, defined in equation (4). We report these 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, $p(\mathcal{D}_b)$, are defined in equation (9) with $-b = \kappa_1 - n\kappa_2^{1/2}$, $n = 3$ or 5.

Parameter	Lognormal Cons Gr (1)	Poisson Cons Gr (2)	Merton Returns (3)	Implied Cons Gr (4)	Bivariate Implied Cons Gr (5)
Preferences					
α (risk aversion)	8.92	5.19	—	8.70	14.94
True distribution					
μ (mean)	0.0200	0.0230	0.0792	0.0303	0.0269
σ (volatility)	0.0350	0.0100	0.1699	0.0253	0.0265
ω (jump intensity)	—	0.0100	1.5120	1.3987	1.3452
θ (mean jump size)	—	−0.3000	−0.0259	−0.0074	−0.0051
δ (volatility of jumps)	—	0.1500	0.0407	0.0191	0.0191
Risk-neutral distribution					
μ^* (mean)	0.0091	0.0225	0.0584	0.0247	0.0233
ω^* (jump intensity)	—	0.0642	1.5120	1.5120	1.5120
θ^* (mean jump size)	—	−0.4168	−0.0542	−0.0105	−0.0105
δ^* (volatility of jumps)	—	0.1500	0.0981	0.0191	0.0191
Skewness, excess kurtosis, and tail probabilities					
γ_1 (true skewness)	0	−11.02	−0.04	−0.28	−0.19
γ_2 (true kurtosis)	0	145.06	0.02	0.48	0.41
γ_1^* (risk-neutral skewness)	0	−4.51	−0.25	−0.38	−0.36
γ_2^* (risk-neutral kurtosis)	0	21.96	0.30	0.53	0.49
γ_1 ($\log m$ skewness)	0	11.02	−0.12	0.28	0.18
γ_2 ($\log m$ kurtosis)	0	145.06	2.21	0.48	0.41
$p(\mathcal{D}_b)$ (tail prob ≤ -3 st dev)	0.0013	0.0090	0.0032	0.0081	0.0056
$p(\mathcal{D}_b)$ (tail prob ≤ -5 st dev)	0.0000	0.0079	0.0000	0.0001	0.00005
Entropy					
$L(m) = L(p^*/p)$	0.0487	0.0400	0.7747	0.0478	0.1424

(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 U.S. history, but that is what the international evidence implies. We return to this issue when we look at the distribution implied by options.

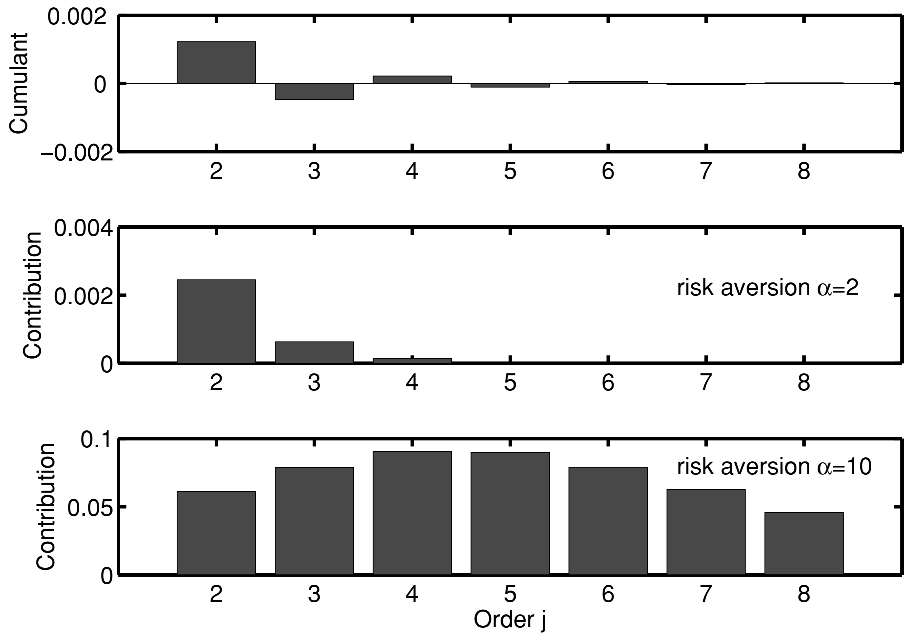

Figure 1. Poisson jumps: cumulants of log consumption growth and contributions to entropy. The panels graph terms in the power series expansion of entropy, equation (13), for the consumption-based asset pricing model with a Poisson-normal jump component. The top panel plots the j^{th} cumulant of log consumption growth, $\kappa_j(\log g)$, against its order j. The next two panels plot the contribution to entropy of the j^{th} 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 III and in column (2) of Table II.

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 eight. There is 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, and Clementi (2010, Section IV.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. The jump component 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 two to eight 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

Table III

Components of Entropy for Model Economies

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 risk aversion α that matches the observed equity premium.

			High-Order Cumulants	
Model	Entropy	Variance/2	Odd	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
Poisson-normal 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
$\alpha = 10$, $\theta = +0.3$ (boom)	0.0265	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 (univariate)	0.0478	0.0464	0.0013	0.0002
Implied consumption growth (bivariate)	0.1424	0.1368	0.0044	0.0013

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 (see also Table III).

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 U.S. data. The line labeled "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 eight. The line labeled "disasters" incorporates the 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. Table III implies that when $\alpha = 2$ the extra terms increase entropy by 32%, but when $\alpha = 10$ the increase is 850%.

It is essential that the jumps be bad outcomes. If we reverse the sign of θ, so that the mean jump is positive, the result is the line labeled "booms" in Figure 2. We see that, for every value of α, entropy is below even the lognormal

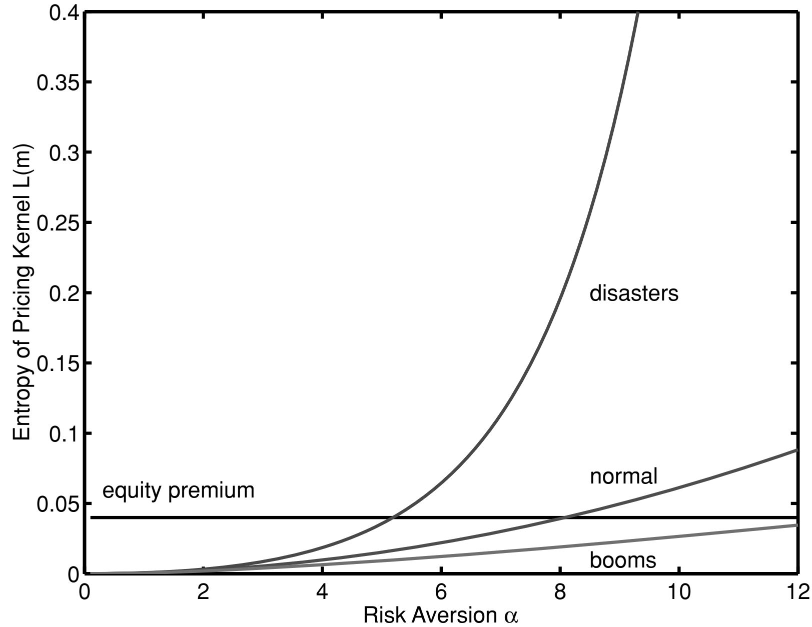

Figure 2. Poisson jumps: entropy of the pricing kernel. The lines represent versions of the consumption-based asset pricing model with Poisson jumps outlined in Section III. The middle line is the lognormal model: log consumption growth is normal and there is no jump component. The top line shows how a jump component (infrequent large negative realizations) increases entropy. The bottom line shows how this changes when the jump is positive (infrequent large positive realizations).

case. Table III shows us exactly how this works. With jumps (and $\alpha = 10$), the entropy of the pricing kernel (0.5837) comes from the variance (0.0613), odd high-order cumulants (0.2786), and even high-order cumulants (0.2439). When we switch to booms, the odd cumulants change sign—see equation (25)—reducing total entropy.

The jump model increases the probability of extreme negative values of consumption growth $\mathcal{D}_b$ relative to the lognormal benchmark. We see in Table II that the probability of log consumption growth more than three standard deviations to the left of its mean ($-b{=}\kappa_1{-}3\kappa_2^{1/2}$ in (9)) is 0.13% in the lognormal case (column (1)) but 0.9% in the Poisson case (column (2)). This corresponds to a drop in consumption of more than 8.5%, something seen only once in U.S. history—in 1931, when consumption fell by 9.9%. Thus, a 1% event has occurred once in slightly more than a century of U.S. history. In this respect, the examples correspond roughly to U.S. experience.

In other respects, the examples are more extreme than U.S. history, implying larger departures from lognormality that we have observed. The model implies, for example, skewness of log consumption growth of -11.02 (the entry labeled

γ_1(true) in Table II). In U.S. data, our estimate is a much more modest -0.34 (the entry labeled skewness in Table I). Excess kurtosis (γ_2) is similar. This is, of course, Barro's (2006) argument: that what we have seen in U.S. data may not accurately reflect the distribution of what might have happened. That discussion leads us to study options, which in principle reflect the distribution used by market participants.

III. 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 II. The state spaces have continuous components, but the logic of Section I.D 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 we have

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

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 (14) 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

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

which implies the cumulant-generating function

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

The cumulants are (evidently) zero after the first two. Entropy follows from equation (20),

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

which is what we reported in equation (23).

In our second example, consumption growth follows the Poisson-normal mixture described by equation (2). We derive the risk-neutral distribution from the cumulant-generating function. This approach works with the previous example, too, but it is particularly convenient here. With power utility, the cumulant-generating function of the risk-neutral distribution is

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

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

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

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

$$\omega^* = \omega e^{-\alpha\theta+(\alpha\delta)^2/2}, \quad \theta^* = \theta - \alpha\delta^2, \quad \delta^* = \delta. \tag{26}$$

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 (22).

IV. Disasters in Option Models and Data

In the macro-finance literature, pricing kernels are typically constructed as in Section III: 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 U.S. 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.

A. 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 are interested in the return on equity, we let

$$\log r^e_{t+1} - \log r^1 = w_{t+1} + z_{t+1}. \tag{27}$$

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 I.B. Appendix D summarizes the option valuation in the framework of this model.

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 E. That is 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. Aït-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 U.S. stock market over the period 1926 to 2009 and shows that a second jump component plays no role in accounting for macroeconomic events like the Great Depression.

Given this structure, the pricing kernel follows from equation (17). Its entropy is

$$\begin{aligned} L(m) &= L(p^*/p) \\ &= \frac{(\mu - \mu^*)^2}{2\sigma^2} + (\omega^* - \omega) + \omega \left[\log \frac{\omega}{\omega^*} - \log \frac{\delta}{\delta^*} + \frac{(\theta - \theta^*)^2 + (\delta^2 - \delta^{*2})}{2\delta^{*2}} \right]. \end{aligned} \tag{28}$$

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

B. Parameter Values

We use parameter values from Broadie et al. (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 et al. (2007, Table I, the line labeled 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 zero jumps per year is 0.220, one jump per year 0.333, two jumps 0.25, three jumps 0.13, four jumps 0.05, and seven 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 II.

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 to 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 II are from Broadie et al. (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 (15), which implies $\mu^* + \sigma^2/2 + \omega^*[\exp(\theta^* + \delta^{*2}/2) - 1] = 0$.

Figure 3 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 D and E. 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

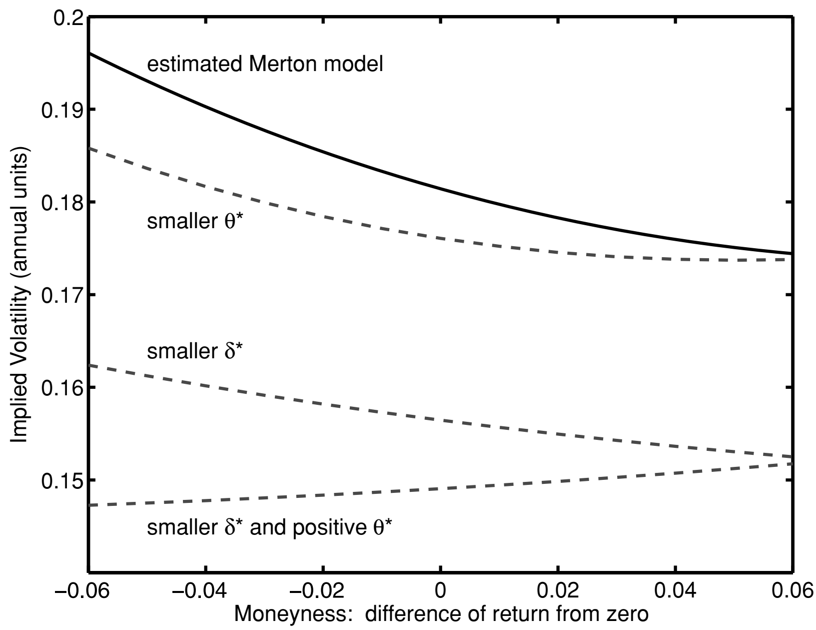

Figure 3. Option model: implied volatility smiles for 3-month options. The lines represent implied "volatility smiles" for the Merton model with estimated parameters and some alternatives. Moneyness is measured as the proportional difference of the strike from the price, (strike − price)/price. For the solid line the parameters are those reported in column (4) of Table II. For the second line, we divide the jump mean θ^* by two. For the bottom line, we divide the jump variance δ^{*2} by two.

help to account for the observed premium of implied volatilities over the true standard deviation of equity returns (0.1800 in our model).

C. Pricing Kernel Implied by Options

We compute the pricing kernel from the ratio of risk neutral to true probabilities, as in equation (17). It therefore incorporates evidence on the time series of returns (the source of information about p) as well as option prices (the source of information about p^*). The properties of the pricing kernel are reported in Tables II and III and Figure 4. We compare the pricing kernel 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, which is more than an order of magnitude larger than the equity premium (0.0400) (column (3) of Table II). 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 et al. (2009, Appendix A). These high average returns imply high entropy via the entropy

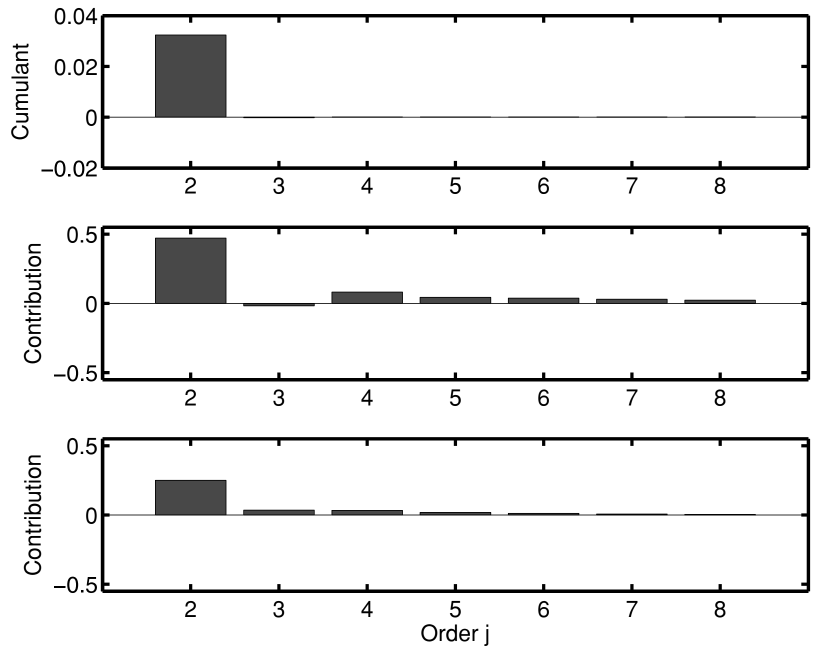

Figure 4. Option model: cumulants of equity returns and contributions to entropy. The figure summarizes properties of the estimated Merton model using parameters reported in column (4) of Table II. The top panel shows cumulants of the log excess return on equity based on its (estimated) true distribution. 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 j^{th} cumulant of $\log m$. The third panel is the same with θ^* and δ^{*2} each divided by two.

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.

The primary source of entropy in this case is the variance of the implied log-pricing kernel: the contribution of the variance is 0.4720, or 61% of the total. High-order cumulants also make significant contributions: 0.1127 (15%) and 0.1900 (25%) from, respectively, odd and even high-order cumulants (Table III). Like the smile itself, these numbers verify that departures from the lognormal model are quantitatively important.

Figure 4 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 II). Contributions of high-order cumulants to entropy are reported in the second panel. As we noted, 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 of Figure 4), the contributions decline across the board, much as when we reduce risk aversion in consumption-based models.

V. Comparing Macroeconomic and Option Models

So far, we show that both option prices and international macroeconomic data suggest significant departures from the lognormal model. Here we explore their differences. As we have seen, the macro- and option-based modeling approaches have two degrees of separation: the latter characterizes the risk-neutral distribution of equity returns, while the former is concerned with the true distribution of consumption growth. The challenge is to link these two objects. Since there is no standard resolution of this problem, we compare the two approaches along the three dimensions suggested by equation (17): the true distribution, the risk-neutral distribution, and their ratio, the pricing kernel. Each gives us a different perspective on the two sources of data and is based on different theoretical structure. First, we compare the macro-based pricing kernel to one derived from option prices and returns (Section I.D). Second, we compare the macro-based risk-neutral distribution of equity returns (Sections II and III) to the same distribution based on option prices (Section IV). Third, we compare the true distribution of consumption growth evident in international macroeconomic data to the same distribution derived from option prices.

We need one final piece of theory to connect equity returns to consumption growth. Equity returns then allow us to calibrate the risk aversion parameter α. We define levered equity as a claim to the dividend

$$d_t = c_t^{\lambda}. \tag{29}$$

This is not, of course, either equity or levered, but it is 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 II.2). In the iid case, the log excess return is a linear function of log consumption growth:

$$\log r_{t+1}^{e} - \log r_{t+1}^{1} = \lambda \log g_{t+1} + \text{constant}. \tag{30}$$

See Appendix G. 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 later in this section, but for now it serves as a useful simplification. Given these assumptions, the equity premium has a compact representation in terms of the cumulant-generating function of consumption growth:

$$E\left(\log r_{t+1}^{e} - \log r_{t+1}^{1}\right) = \lambda \kappa_1(\log g) + k(-\alpha; \log g) - k(\lambda - \alpha; \log g). \tag{31}$$

See Appendix G.

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.

A. Comparing Pricing Kernels

We start with a direct comparison of the pricing kernels derived in Sections II and IV. Their entropies and the relevant parameter values are reported in columns (2) and (3) of Table II. In the macro or consumption-based model, we set $\alpha = 5.19$ to match the target equity premium of 0.0400, computed from equation (31).

One clear difference between the two models is their 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 III labeled "Poisson consumption growth, $\alpha = 5.19$") but only 39% in the option model (the row labeled "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.

This is, to be sure, a somewhat odd result. Options imply lots of entropy, which is, after all, one of the benefits of introducing disasters into asset pricing models. But it does so primarily by increasing the variance of the log pricing kernel. The relative contribution of odd high-order cumulants remains important but smaller than we see in the consumption-based model.

B. Comparing Option Prices

Our second comparison involves option prices, which reflect the risk-neutral distribution of equity returns. We compute option prices 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 equation (26). The results are reported in column (2) of Table II. Second, equation (30) implies that equity excess returns are a scaled version of consumption growth with scale parameter λ. This scaling leads us to replace the parameters (σ^*, ω^*, θ^*, δ^*) with ($\lambda\sigma^*$, ω^*, $\lambda\theta^*$, $\lambda\delta^*$). See Appendix E. 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 5 . Similar consumption-based option prices are reported by Benzoni, Collin-Dufresne, and Goldstein (2011) and Du (2008). What is 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 (labeled "option-based model") refers to the model based on option prices. It is the same as the top line in Figure 3. The bottom line (labeled "consumption-based model") refers to the model derived from consumption

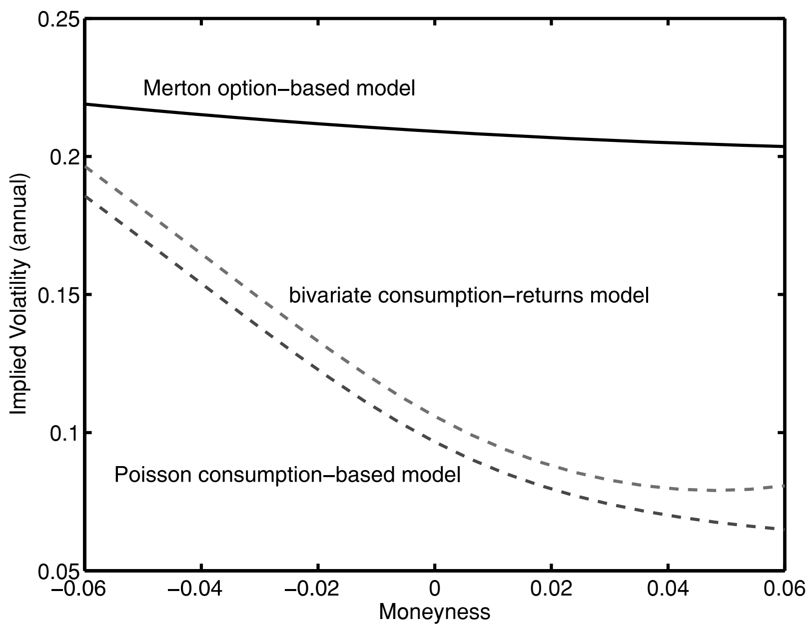

Figure 5. Implied volatility smiles based on option and consumption data. The lines represent implied volatility "smiles" for the Merton model with three different sets of parameters. The top line comes from Figure 3 and uses parameters estimated from option prices. The shape looks different from that in Figure 3 because of the difference in ranges on the y-axes. The bottom line uses parameters estimated from consumption data, extrapolated to equity returns using equation (30). The middle line is based on a bivariate model of consumption growth and equity returns with correlation chosen to match U.S. data.

data as described in the previous paragraph. The implications of the two sets of parameter values are clearly different. The consumption-based calibration 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 II). 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.

C. Comparing Consumption Disasters

Now consider the reverse, that is, the true distribution of consumption growth implied by option prices. How does it compare to the distribution estimated from international macroeconomic data? 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 (30) to link consumption growth to returns. This imposes different structure from the option model on the connection between true and risk-neutral parameters. For example, the restriction $\omega = \omega^*$ used in the 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 in the Merton model into risk-neutral parameters for consumption growth. This involves replacing (ω^*, θ^*, δ^*) in column (3) of Table II with (ω^*, θ^*/λ, δ^*/λ), as outlined in Appendix E. Second, expressions (26) link the risk-neutral parameters of the consumption distribution to their counterparts (ω, θ, δ) for the true distribution. Finally, μ and σ are selected to match our target values for the mean and standard deviation of log consumption growth; see equations (5) and (6). Risk aversion α is then chosen to match the equity premium. The jump parameters ω and θ depend on α, the two true normal parameters μ and σ depend on the true jump parameters, and α depends on all the true parameters. Thus, we have a system of five equations in five unknowns.

The consumption process derived this way from option prices (column (4) of Table II) has the same Poisson-normal distribution as the consumption process estimated from international macroeconomic data (column (2)). 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 of consumption growth). 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 three standard deviation drop in consumption ($-b = \kappa_1 - 3\kappa_2^{1/2}$ in (9)), similar to the United States in the Great Depression, is about 1% in each case. However, declines in consumption of more than five 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 U.S. 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. The latter are much closer to the sample statistics reported in Table I than the former.

This leaves us with an alternative to the popular "Barro" calibration of disasters. The parameter values in column (4) have three useful properties: (i) they match the equity premium, (ii) they match the implied volatility smile of option prices, and (iii) they are consistent with U.S. consumption data. Properties (i) and (ii) are true by construction; (iii) is an implication.

The mechanism for matching these properties is different from the Barro calibration: while leverage λ is the same as in the Poisson consumption-based model (column (2) of Table II), risk aversion is not much smaller than in the lognormal case (column (1) of Table II) and much larger than in the Poisson model. So, the difference in jump parameters leads to the difference in risk aversion

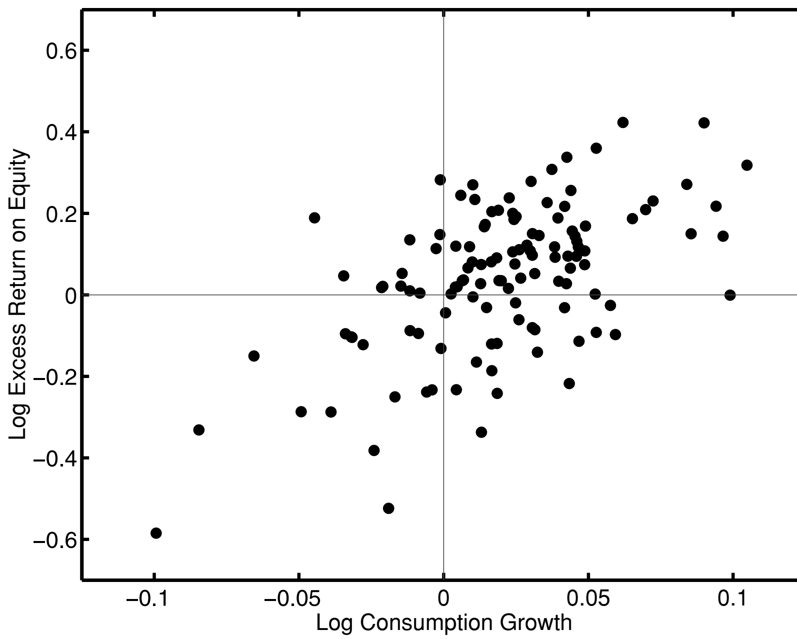

Figure 6. Equity returns and consumption growth in U.S. data. The dots represent annual observations of the log excess return on equity and the log of per capita consumption growth over the period 1889 to 2009. The data come from Shiller (2009), updated as needed.

that is required to match equity premium. Relatively high risk aversion does not imply that nonnormalities do not matter: the log-normal model would not be able to match items (ii) and (iii).

Finally, our alternative implied calibration does not solve all the problems. There is a tension in the model highlighted by the fact that the entropy of the implied calibration is not much higher than that in the Poisson consumption-based model and much lower than that in the Merton option-based model. This suggests that the implied calibration will have difficulty matching option risk premia (option excess returns).

D. A Bivariate Model of Consumption and Equity Returns

Our comparisons of option prices and consumption growth rely on the tight connection between consumption growth and equity returns imposed by (30). We follow a long tradition in doing so, but consider an alternative here that is more closely based on the evidence. Figure 6 shows that consumption growth and equity returns have been strongly correlated, but the sample correlation is 0.566, not one.

Consider, then, a bivariate model of consumption growth and dividends (we will come to returns shortly) with arbitrary correlation between the two. Does allowing imperfect correlation affect our conclusion that option prices imply smaller probabilities of disasters than international macroeconomic data? The answer is no, but we think it is worth working through the details.

Let $\log g^\top = (\log g_1, \log g_2)$ be a vector whose elements are log consumption and dividend growth: $g_{1t+1} = c_{t+1}/c_t$ and $g_{2t+1} = d_{t+1}/d_t$. A bivariate jump process analogous to (1) is

$$\log g_{t+1} = \mu^g + w_{t+1} + z_{t+1},$$

where w and z are bivariate and independent of each other and over time. The intercept μ^g has elements μ_i^g. The first component is bivariate normal: $w_t \sim \mathcal{N}(0, \Sigma)$, where Σ has elements σ_{ij}. The second component is a Poisson mixture of bivariate normals. 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 that used by Aït-Sahalia, Cacho-Diaz, and Hurd (2009) and similar to consumption-dividend processes used by Gabaix (2010) and Longstaff and Piazzesi (2004). This bivariate process is a modest generalization of (29). It allows us to maintain the scaling performed by (29) without imposing a correlation of one. Scaling includes $\mu_2^g = \lambda \mu^g{}_1$, $\sigma_{22} = \lambda^2 \sigma_{11}$, $\theta_2 = \lambda \theta_1$, and $\delta_{22} = \lambda^2 \delta_{11}$.

With this process for consumption and dividend growth, the joint process for consumption growth and equity returns has a similar structure. If

$$x^\top = (x_1, x_2) = (\log g_1, \log r^e - \log r^1),$$

then

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

Here w and z are the same as above and μ has elements

$$(\mu_1, \mu_2) = \left(\mu_1^g, E\left(\log r_{t+1}^e - \log r_{t+1}^1\right) - \omega\lambda\theta_1\right).$$

The equity premium is

$$E\left(\log r_{t+1}^e - \log r_{t+1}^1\right) = \lambda\left(\mu_1^g + \omega\theta_1\right) + k((-\alpha, 0)'; \log g) - k((-\alpha, 1)'; \log g). \quad (32)$$

See Appendix G.

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 (2) of Table II for the (true) consumption process. Here we label them $(\mu_1, \sigma_{11}^{1/2}, \omega, \theta_1, \delta_{11}^{1/2})$. The parameters of the return process are scaled versions of the same numbers as described above. The jump intensity ω is, of course, the same. We choose risk aversion α to match the equity premium (32). The correlation depends on the mean jumps and the correlations $\rho_w = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$ and $\rho_z = \delta_{12}/(\delta_{11}\delta_{22})^{1/2}$. If $\rho_w = \rho_z = \rho$, then the correlation between consumption and returns is

$$\mathrm{Corr}(x_1, x_2) = \frac{\rho\sigma_{11} + \omega\big(\theta_1^2 + \rho\delta_{11}\big)}{\sigma_{11} + \omega\big(\theta_1^2 + \delta_{11}\big)}. \tag{33}$$

See Appendix H. We set ρ to match the correlation in the data (0.566).

Option pricing follows directly from the risk-neutral distribution over equity returns. With power utility, we have

$$\omega^* = \omega e^{-\alpha\theta_1 + \alpha^2\delta_{11}/2}, \quad \theta_2^* = \theta_2 - \alpha\delta_{12}, \quad \delta_{22}^* = \delta_{22}.$$

See Appendix H. With appropriate redefinition of parameters, option pricing has the same structure as in Section IV and Appendix D.

If consumption growth and equity returns are perfectly correlated, this procedure reproduces the calculation of consumption-based option prices in Figure 5. 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 the perfect correlation between consumption growth and equity returns is not the source of the sharp difference between volatility smiles based on option prices and consumption data.

Now consider the consumption growth process implied by option prices in this bivariate setting. First, we use the scale parameter λ to transform risk-neutral parameters for equity returns into risk-neutral parameters for consumption growth. This involves setting $(\omega^*, \theta_1^*, \delta_{11}^*) = (\omega^*, \theta^*/\lambda, \delta^{*2}/\lambda^2)$ for reasons outlined in Appendix E, where $(\omega^*, \theta^*, \delta^*)$ are parameters estimated from option prices and reported in column (3) of Table II. Second, the expressions

$$\omega^* = \omega e^{-\alpha\theta_1 + \alpha^2\delta_{11}/2}, \quad \theta_1^* = \theta_1 - \alpha\delta_{11}, \quad \delta_{11}^* = \delta_{11}$$

from Appendix H link the risk-neutral parameters of consumption to their counterparts $(\omega, \theta_1, \delta_{11})$ of the true distribution. The parameters of the normal component of consumption growth, μ_1 and σ_{11}, are chosen to match our target values for the mean and standard deviation of log consumption growth:

$$\begin{aligned} 0.0200 &= \kappa_1(\log g_1) = \mu_1 + \omega\theta_1 \\ 0.0350^2 &= \kappa_2(\log g_1) = \sigma_{11} + \omega\big(\theta_1^2 + \delta_{11}\big). \end{aligned}$$

The correlation ρ is chosen to match the sample correlation between consumption growth and equity returns, as in (33). Risk aversion α is chosen to match the equity premium (32). The two true jump parameters ω and θ_1 depend on α, the two true normal parameters μ_1 and σ_{11} depend on the true jump parameters, ρ depends on true jump parameters and σ_{11}, and α depends on all the true parameters. Thus, we have to solve a system of six equations in six unknowns.

As in the univariate case, the consumption process derived this way from option prices (column (5) of Table II) has the same Poisson-normal distribution as the consumption process estimated from international macroeconomic evidence (column (2)). The parameters are, however, different. With the exception of risk aversion, the difference between columns (2) and (5) is similar to the difference between columns (2) and (4), the latter being the univariate case. The need for higher risk aversion reflects the lower correlation between dividends and consumption. Again, it seems that allowing a smaller correlation does not have a large impact on the consumption process derived from option prices.

E. Discussion

One might ask, in the end, whether options reliably identify extreme events (disasters) in consumption. We think the answer is yes, but let us 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 are interested in the true distribution of consumption growth. The first link is between true and risk-neutral returns. We have 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 (3) of Table II). The second link is between returns and consumption growth. When the two are perfectly correlated, as they are in equation (30), 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 (2)) and option prices (columns (3)–(5)) is striking. So, too, is the difference between implied volatilities based on option prices and consumption data (Figure 5). Finally, the evidence on the relation between equity returns and consumption growth (Figure 6) and the limited impact of imperfect correlation on option prices (Figure 5) and on option-implied consumption growth (columns (4) and (5) of Table II) suggest that option prices are a reasonably good indicator of the likelihood of disasters in consumption growth. It is 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.

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. Nonetheless, this option-implied distribution is capable of matching the equity premium, option prices, and consumption moments in a simple setting of iid consumption growth and power utility over aggregate consumption.

Can we make further progress by relaxing these assumptions? In general, the answer must be yes. The most recent generation of dynamic asset pricing models is capable of capturing important features of the data by combining exotic preferences with non-iid consumption dynamics. However, this work focuses on intertemporal properties of the asset prices, such as return predictability, volatility persistence, and yield curve modeling. Here we focus on the drivers of the unconditional moments of consumption and asset prices listed in Table I. Although this is not the last word on the subject, we think the models we consider are well suited to the task. They are also relatively transparent, which we think adds to our understanding of their properties.

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 is hard to make a definitive statement without knowing the precise form of time dependence, but there is 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 et al. (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.

Power utility is the workhorse of macroeconomics and finance, but our option model suggests much higher entropy than implied by macro models with these preferences, even the ones that were calibrated to match option prices. One possible remedy is to explore alternative preferences, including skewness preferences (Harvey and Siddique (2000)), recursive preferences (Garcia, Luger, and Renault (2003) 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.

VI. Final Remarks

In this paper, we use 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. This 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. Each approach indicates that the probabilities of extreme adverse events implied by option prices are smaller than we see in international macroeconomic data. Turning to 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 leads us to two useful byproducts. One is a reminder that matching the equity premium may not be enough. Growing evidence suggests 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 byproduct is a reminder of the value of transform methods, particularly cumulant-generating functions. These are not new to finance, but they are nevertheless extremely helpful. We find that they are not only a source of intuition and compact notation, but also a convenient approach to a number of practical problems.

Appendix A: The Cumulant-Generating Function of Poisson-Normal Mixtures

We look at a Poisson-normal mixture shortly, but it is useful to start with a Poisson random variable 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 moment-generating function for the mixture is the probability-weighted average,

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

which implies the cumulant-generating function

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

The same approach can be used for jumps with other distributions. If we set $\theta = 1$ and $\delta = 0$, we get the cumulant-generating function 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.

Appendix B: The Entropy Bound

The entropy bound (12) 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 (B3) below). We derive the bound as follows:

- *Bound on mean log return.* Since log is a concave function, Jensen's inequality and the unconditional version of the pricing relation (10) imply that for any positive return r,

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

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

  $$E \log r \;\le\; -E \log m. \tag{B1}$$

  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:

  $$L(q^1) = \log Eq^1 - E \log q^1 = \log Em + E \log r^1. \tag{B2}$$

- *Entropy bound.* Expressions (B1) and (B2) imply

  $$L(m) \ge E\left(\log r^j - \log r^1\right) + L(q^1). \tag{B3}$$

Inequality (12) follows from $L(q^1) \geq 0$ (entropy is nonnegative). In practice, $L(q^1)$ is small; in the iid case, it is 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

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

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

$$\begin{aligned} Em &= E\big(e^{\log m}\big) = e^{k(1)} \\ \text{Var}(m) &= E(m^2) - (Em)^2 = e^{k(2)} - e^{2k(1)}. \end{aligned}$$

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

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

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 $k(1) - k'(0)$.

Appendix C: 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 is 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

$$k^*(s) = k(s-\alpha) - k(-\alpha). \tag{C1}$$

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.

Appendix D: Risk-Neutral Option Pricing

We review option pricing in the Merton model, starting with its primary ingredient, the BSM 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 log-normal: $\log r \sim \mathcal{N}(\log r^1 + \kappa_1, \kappa_2)$. The pricing relation (15) implies the restriction $\kappa_1 + \kappa_2/2 = 0$, which we will 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

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

where

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

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

$$q^p(b) = f(b; \kappa_1, \kappa_2) = q^1 b N(d) - \exp(\kappa_1 + \kappa_2/2) N\big(d - \kappa_2^{1/2}\big). \tag{D1}$$

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 IV.A 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 $\kappa_{2j} = \sigma^2 + j\delta^2$.

Appendix E: 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 1 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 (30)). The general result follows from this property of cumulant-generating functions: $k(s;\, \lambda x) = k(\lambda s;\, x)$. If x (think log consumption growth) has the Poisson-normal structure of Section I.B, its cumulant-generating function is

$$k(s; x) = \mu s + (\sigma s)^2/2 + \omega[e^{\theta s + (\delta s)^2/2} - 1]. \tag{E1}$$

The cumulant-generating function 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 as the cumulant-generating function of consumption growth 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 (C1), their cumulant-generating functions are connected by

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

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

The second issue concerns the time interval. In an iid setting, suppose the cumulant-generating function (E1) applies to the distribution over a unit time interval. The cumulant-generating function for an arbitrary time interval $\tau > 0$, if it exists, is the cumulant-generating function 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 cumulant-generating function has the same form as (E1) with (μ, σ^2, ω, θ, δ^2) replaced by ($\tau\mu$, $\tau\sigma^2$, $\tau\omega$, θ, δ^2).

Appendix F: 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 (27), 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

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

This differs from the examples in Section IV 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 et al. (1996). The risk-neutral pricing relation (15) 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

$$\begin{aligned}
h(s;\log p^*/p) &= E(e^{s\log p^*/p}) \\
&= \int_{-\infty}^{\infty} p^*(w)^s p(w)^{1-s} dw \\
&= (2\pi\sigma^2)^{-1/2}\varphi^{s/2}\int_{-\infty}^{\infty} \exp\{-[(1-s)(w-\mu)^2 + s\varphi^2(w-\mu^*)^2]/2\sigma^2\} dw \\
&= \varphi^{s/2}[1-s(1-\varphi)]^{-1/2}\exp\left(\frac{s(s-1)(\mu^*-\mu)^2}{2\sigma^{*2}[1-s(1-\varphi)]}\right)
\end{aligned} \tag{F1}$$

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:

$$-\kappa_1(\log p^*/p) = (1/2)[\log\varphi + 1 - \varphi] + (\mu-\mu^*)^2/2\sigma^{*2}. \tag{F2}$$

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

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

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

Now let us 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

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

Using (F1) we have

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

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

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

Entropy is minus the first derivative evaluated at zero:

$$\begin{aligned} &- \kappa_1(\log p^*/p) \\ &= (\omega^* - \omega) + \omega[\log(\omega/\omega^*) - 1/2 \cdot \log\varphi + 1/2 \cdot (\varphi - 1)] + \omega(\theta - \theta^*)^2/2\delta^{*2}. \end{aligned} \tag{F3}$$

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 (F2) with $\varphi = 1$) and the entropy of the Poisson mixture of normals (equation (F3)).

Appendix G: 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).

Perfectly Correlated Consumption and Dividend Growth

Let us 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

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

The second equality is based on the definition of the pricing kernel, equation (21). The last one follows from the definition of the cumulant-generating function k, equation (3).

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 again constant. The pricing relation (10) and our power utility pricing kernel (21) then imply

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

Thus, we have, in compact notation,

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

The equity premium is therefore

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

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 Bivariate Model of Consumption and Dividend Growth

As before, a one-period risk-free bond sells at price $q^1_t = E_t m_{t+1}$ and has return $r^1_{t+1} = 1/q^1_t = 1/E_t m_{t+1}$. In the iid case, the short rate is constant and equals

$$\begin{aligned}
\log r^1 &= -\log E(m)\\
&= -\log\beta - \log E(e^{-\alpha\log g_1}) = -\log\beta - k((-\alpha,0)';\log g).
\end{aligned}$$

The last equality follows from using a cumulant-generating function of a bivariate variable $\log g = (\log g_1, \log g_2)'$ given by

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

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

$$r^e_{t+1} = e^{\log g_{2t+1}}\left(1+q^e_{t+1}\right)/q^e_t.$$

In the iid case, q^e is constant. The pricing relation (10) and our power utility pricing kernel (21) then imply

$$q^e/(1+q^e) = E(\beta e^{-\alpha\log g_1 + \log g_2}).$$

Thus, we have

$$\begin{aligned}\log[q^e/(1+q^e)] &= \log\beta + k((-\alpha,1)';\log g)\\ \log r^e_{t+1} &= \log g_{2t+1} - \log\beta - k((-\alpha,1)';\log g)\\ \log r^1_{t+1} &= -\log\beta - k((-\alpha,0)';\log g)\\ \log r^e_{t+1} - \log r^1_{t+1} &= \log g_{2t+1} + k((-\alpha,0)';\log g) - k((-\alpha,1)';\log g)\end{aligned}$$

and the equity premium is

$$E\left(\log r^e_{t+1} - \log r^1_{t+1}\right) = \kappa_1(\log g_2) + k((-\alpha,0)';\log g) - k((-\alpha,1)';\log g).$$

Appendix H: 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:

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

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 I.B, 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) cumulant-generating function is

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

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

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

The correlation is

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

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 cumulant-generating function 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 C. The cumulant-generating functions 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 cumulant-generating function.) Thus, the cumulant-generating function for x_2 is $k(0, s_2)$. The risk-neutral cumulant-generating function for the log equity excess return is therefore

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

with the implicit definitions

$$\begin{aligned} &\mu_2^* = \mu_2 - \alpha\sigma_{12}, \quad \theta_2^* = \theta_2 - \alpha\delta_{12}, \quad \delta_{22}^* = \delta_{22} \\ &\omega^* = \omega e^{-\alpha\theta_1+\alpha^2\delta_{11}/2}. \end{aligned}$$

This has the same form as the Merton model with suitably defined parameters. Options can therefore be priced using the methods of Appendix D. Similarly, $k^*(s_1, 0)$ implies the risk-neutral distribution of consumption growth:

$$\mu_1^* = \mu_1 - \alpha\sigma_{11}, \quad \theta_1^* = \theta_1 - \alpha\delta_{11}, \quad \delta_{11}^* = \delta_{11}.$$

REFERENCES

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

Aït-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.

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

Aït-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.

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, Working paper, Harvard University.

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, Working paper, Wharton School.

Bates, David, 2008, The market for crash risk, *Journal of Economic Dynamics and Control* 32, 2291–2321.

Bates, David, 2010, U.S. stock market crash risk, 1926–2009, Working paper, University of Iowa.

Bekaert, Geert, and Eric Engstrom, 2010, Asset return dynamics under bad environment-good environment fundamentals, Working paper, Columbia University.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2011, Explaining asset pricing puzzles associated with the 1987 market crash, *Journal of Financial Economics* 101, 552–573.

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* 22, 4493–4529.

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.

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

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.

Drechsler, Itamar, 2008, Uncertainty, time-varying fear, and asset prices, Working paper, New York University.

Drechsler, Itamar, and Amir Yaron, 2008, What's vol got to do with it?, Working paper, New York University.

Du, Du, 2008, How bad will the potential economic disasters be? Evidences from S&P 500 index options data, Working paper, HKUST.

Gabaix, Xavier, 2010, Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance, Working paper, New York University.

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* 77, 1711–1740.

Hansen, Lars Peter, and Ravi Jagannathan, 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.

Julliard, Christian, and Anisha Ghosh, 2008, Can rare events explain the equity premium puzzle? Working paper, London School of Economics.

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

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, Working paper, UCLA.

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, Working paper, Harvard University.

Martin, Ian, 2009, Consumption-based asset pricing with higher cumulants, Working paper, Stanford University.

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, Working paper, Wharton School.

Shiller, Robert, 2009, Long term stock, bond, interest rate and consumption data, spreadsheet posted on his website http://www.econ.yale.edu/s̃hiller/data.htm (accessed June 9, 2010).

Wachter, Jessica, 2009, Can time-varying risk of rare disasters explain aggregate stock market volatility?, Working paper, Wharton School.

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.