Disasters implied by equity index options*

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

October 26, 2009

Abstract

We use prices of equity index options to quantify the impact of extreme events on asset returns. We define extreme events as departures from normality of the log of the pricing kernel and summarize their impact with high-order cumulants: skewness, kurtosis, and so on. We show that high-order cumulants are quantitatively important in both representative-agent models with disasters and in a statistical pricing model estimated from equity index options. Option prices thus provide independent confirmation of the impact of extreme events on asset returns, but they imply a more modest distribution of them.

JEL Classification Codes: E44, G12.

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

---

*We thank Michael Brandt, Rodrigo Guimaraes, Sydney Ludvigson, Monika Piazzesi, Romeo Tedongap, Mike Woodford, and Liuren Wu, as well as participants in seminars at the Bank of England, the CEPR symposium on financial markets, the London Business School, the London School of Economics, the NBER summer institute, New York University, SIFR, and the financial econometrics conference in Toulouse. We also thank Mark Broadie for sharing his option pricing code and Vadim Zhitomirsky for research assistance.

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

‡ London Business School and CEPR; mchernov@london.edu.

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

1 Introduction

Barro (2006), Longstaff and Piazzesi (2004), and Rietz (1988) show that disasters — infrequent large declines in aggregate output and consumption — produce dramatic improvement in the ability of representative agent models to reproduce prominent features of US asset returns, including the equity premium. We follow a complementary path, using equity index options to infer the distribution of returns, including extreme events like the disasters apparent in macroeconomic data.

The primary challenge for theories based on disasters lies in estimating their probability and magnitude. Since disasters are, by definition, rare, it is difficult to estimate their distribution reliably from the relatively short history of the US economy. Rietz (1988) simply argues that they are plausible. Longstaff and Piazzesi (2004) argue that disasters based on US experience can explain only about one-half of the equity premium. Barro (2006) and Barro and Ursua (2008) study broader collections of countries, which in principle can tell us about alternative histories the US might have experienced. They show that these histories include occasional drops in output and consumption that are significantly larger than we see in typical business cycles. Equity index options are a useful source of additional information, because their prices tell us how market participants value extreme events, whether they happen in our sample or not. The challenge here lies in distinguishing between true and risk-neutral probabilities. We use a streamlined version of a model estimated by Broadie, Chernov, and Johannes (2007) that identifies both. Roughly speaking, risk-neutral probabilities are identified by option prices (cross-section information) and true probabilities are identified by the distribution of equity returns (time-series information). The resulting estimates provide independent evidence of the quantitative importance of extreme events in US asset returns.

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

The question is how to measure the impact of disasters. We find two statistical concepts helpful here: entropy (a measure of volatility or dispersion) and cumulants (close relatives of moments). Alvarez and Jermann (2005) show that mean excess returns, defined as differences of logs of gross returns, place a lower bound on the entropy of the pricing kernel. If the log of the pricing kernel is normal, then entropy is proportional to its variance. But departures from normality, and disasters in particular, can increase entropy and thereby improve a model's ability to account for observed excess returns. We quantify the impact of departures from normality with high-order cumulants. Disasters and other departures from normality can contribute to entropy by introducing skewness, kurtosis, and so on. These

ideas are laid out in Section 2, where we also show how the pricing kernel is related to the risk-neutral probabilities commonly used in option pricing models.

In Section 3 we illustrate the macro-finance approach to disasters: log consumption growth includes a non-normal component and power utility converts consumption growth into a pricing kernel. We show how infrequent large drops in consumption growth generate positive skewness in the log pricing kernel and increase its entropy. The impact can be large, even with moderate risk aversion. It's important that the departures from normality have this form: adding large positive changes to consumption growth can reduce entropy relative to the normal case.

Do option prices indicate a similar contribution from large adverse events? The answer is, roughly, yes, but the language and modelling approach are quite different. Option pricing models typically express asset prices in terms of risk-neutral probabilities rather than pricing kernels. This is more than language; it governs the choice of model. Where macro-finance models generally start with the true probability distribution of consumption growth and use preferences to deduce the risk-neutral distribution, option pricing models infer both from asset prices. The result is a significantly different functional form for the pricing kernel. In Section 4 we describe the risk-neutral probabilities implied by consumption-based models. In Section 5 we show how data on equity returns and option prices can be used to estimate true and risk-neutral probability distributions. The modelling strategy is to use the same functional form for each, but allow the parameters to differ. We describe how the various parameters are identified and verify the quantitative importance of high-order cumulants. Both consumption- and option-based models imply substantial contributions to entropy from odd high-order cumulants. In this sense, option prices are consistent with the macroeconomic evidence on disasters. Options, however, imply much greater entropy than models designed to reproduce the equity premium alone. Evidently the market places a large premium on whatever risk is involved in selling options.

In Section 6 we explore the differences between the evidence from consumption data and option prices by looking at each from the perspective of the other. If we consider a consumption-based disaster model, how do the option prices compare to those we see in the market? And if we infer consumption growth from option prices, how does it compare to the macroeconomic evidence of disasters? Both of these comparisons suggest that option prices imply more modest disasters than the macroeconomic evidence suggests.

We conclude with a discussion of extensions and related work.

2 Preliminaries

We start with an overview of the tools and evidence used later on. The tools allow us to characterize departures from (log)normality, including disasters, in a convenient way.

Once these tools are developed, we describe some of the evidence they'll be used to explain. Most of this is done for the iid (independent and identically distributed) environments we use in Sections 3 to 6. There are many features of the world that are not iid, but this simplification allows us to focus without distraction on the distribution of returns, particularly the possibility of extreme negative outcomes. We think it's a reasonably good approximation for this purpose, but return to the issue briefly in Section 7.

2.1 Pricing kernels, entropy, and cumulants

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

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

for (gross) returns r^j on all traded assets j. Here E_t denotes the expectation conditional on information available at date t. In the stationary ergodic settings we consider, 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 kernel is tied to macroeconomic quantities such as consumption growth. In this respect, the pricing kernel is a link between macroeconomics and finance.

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

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

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

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

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

for any asset j with positive returns. See Alvarez and Jermann (2005, proof of Proposition 2) and Appendix A.1. Here r^1 is the (gross) return on a one-period risk-free bond, so

the right-hand side is the mean excess return or premium on asset j over the short rate. Inequality (3) therefore transforms estimates of return premiums into estimates of the lower bound of the entropy of the pricing kernel.

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

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

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

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

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

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

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

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

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

This use of the cumulant-generating function is in the spirit of Martin (2008), as are many of its uses in later sections. If $\log m$ is normal, entropy is one-half the variance ($\kappa_2/2$), but in general there will be contributions from skewness ($\kappa_3/3!$), kurtosis ($\kappa_4/4!$), and so on.

As Zin (2002, Section 2) suggests, it's not hard to imagine using high-order cumulants to account for properties of returns that are difficult to explain in lognormal settings. We refer to this as Zin's "never a dull moment" conjecture after a phrase from his paper. We use the following language and metrics to capture this idea. We refer to departures from normality of the log of the pricing kernel as reflecting extreme events and measure their

impact with high-order cumulants. Disasters are a special case in which the extreme events include a significant positive contribution from odd high-order cumulants. This gives us a three-way decomposition of entropy: one-half the variance (the normal term, so to speak) and contributions from odd and even high-order cumulants. We compute odd and even cumulants from the odd and even components of the cumulant-generating function. An arbitrary cumulant-generating function $k(s)$ has odd and even components

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

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

2.2 Risk-neutral probabilities

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

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

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

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

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

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

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

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

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

$$q^p \quad = \quad q^1 E^*[b - q(x)]^+.$$

As we vary b, we trace out the risk-neutral distribution of prices $q(x)$ (Breeden and Litzenberger, 1978).

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

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

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

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

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

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

The expression on the right of (10) is sometimes referred to as the entropy of p^* relative to p, which provides a justification 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

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

The definition of entropy (2) contributes the analog to (6):

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

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

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

2.3 Evidence

Our goal is to put these tools to work in accounting for broad features of macroeconomic and financial data: consumption growth, asset returns, and option prices. Here's a quick overview of US data.

In Table 1 we report familiar evidence on annual consumption growth and equity returns for both a long sample (1889-2006) and a shorter one (1986-2006) that corresponds to our data on options. The numbers are similar to those reported by Alvarez and Jermann (2005, Tables I-III), Barro (2006, Table IV), and Mehra and Prescott (1985, Table 1). There has been some variation over time in the equity premium (larger in the more recent sample) and consumption growth (less volatile in the recent past), but both may be closer to the long sample once we include the most recent observations.

In Table 2 we describe prices of options on S&P 500 contracts. Prices are reported as implied volatilities: values of the volatility parameter of the Black-Scholes-Merton formula that generates the observed option price. This convention allows a simple comparison to the lognormal case, in which volatility is the same at all strike prices. In the table, we report average implied volatilities of options for a range of strike prices. Observations are annual. They cover options of maturities 1, 3, and 12 months. Since the macroeconomic data are annual, the annual maturity is the most natural in this context, but shorter maturities are more informative about extreme events. An option of maturity n months, for example, reflects the n-month distribution of index returns. As we increase n, the standardized distribution becomes more normal and extreme events are relatively less important. Shorter options are also more frequently traded.

Option prices have two features that we examine more closely in Section 5. Similar evidence has been reviewed recently by Bates (2008, Section 1), Drechsler and Yaron (2008, Section 2), and Wu (2006, Section II). The first feature is that implied volatilities are greater than sample standard deviations of returns (compare Table 1). Since prices are increasing in volatility, it implies that option prices are high relative to the lognormal Black-Scholes-Merton benchmark. The second is that implied volatilities, hence option prices, are higher for lower strike prices: the well-known volatility skew. This feature is more evident at shorter maturities, precisely because they are more sensitive to extreme events. It's also intriguing from a disaster perspective, because it suggests market participants value adverse

events more than is implied by a lognormal model. The question for us whether the extra value assigned to bad outcomes corresponds to the disasters documented in macroeconomic research.

Table 3 is a summary of the evidence: a collection of ballpark numbers that we use as targets for theoretical examples. Thus we consider examples in which the log excess return on equity has a mean of 0.0440 (4.40%) and a standard deviation of 0.1500. Similarly, log consumption growth has a mean of 0.0200 and a standard deviation of 0.0350. None of these numbers are definitive, but they give us a starting point for considering the quantitative implications of theoretical models.

3 Disasters in macroeconomic models and data

Barro (2006), Longstaff and Piazzesi (2004), and Rietz (1988) construct representative-agent exchange economies in which infrequent large declines in consumption improve their ability to generate realistic asset returns: illustrations, in other words, of Zin's "never a dull moment" conjecture. We describe their mechanism with two numerical examples that highlight the role of high-order cumulants.

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 $\alpha \geq 0$. If consumption growth is $g_t = c_t/c_{t-1}$, the pricing kernel is

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

With power utility, the properties of the pricing kernel follow from those of consumption growth. Entropy is

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

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

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

See Section 2.1. If log consumption growth is normal, then so is the log of the pricing kernel. Entropy is then one-half the variance of consumption growth times the risk aversion parameter squared. The impact of high-order cumulants depends on $(-\alpha)^j/j!$. The minus sign tells us the 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 helps with the Alvarez-Jermann bound. The magnitudes of high-order cumulants are controlled by $\alpha^j/j!$. Eventually the denominator grows faster than the numerator, but for moderate values of j, risk aversion can magnify the contributions of high-order cumulants relative to the contribution of the variance. Yaron refers to this as a "bazooka": if $\alpha > 1$, even moderate high-order cumulants in log consumption growth can have a large impact on the entropy of the pricing kernel.

We follow Barro (2006) in choosing an iid process for consumption growth so that we can focus on the role played by its distribution. Let consumption growth be

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

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

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

Similarly, the cumulants of $\log m$ are sums of the cumulants of the components:

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

(That's why they call them cumulants: they "[ac]cumulate.") We let $w \sim \mathcal{N}(\mu, \sigma^2)$, so that any contribution to high-order cumulants comes from z.

The question is the behavior of z. We consider two examples. In both cases, parameter values are adapted from Barro (2006), Barro and Ursua (2008), and Barro, Nakamura, Steinsson, and Ursua (2009), who show that sharp downturns are an infrequent but recurring feature of national consumption and output data.

3.1 Example 1: Bernoulli disasters

The simplest example of a disaster process is a Bernoulli random variable. Suppose the second component of consumption growth is

$$z_t \quad = \quad \begin{cases} 0 & \text{with probability } 1-\omega \\ \theta & \text{with probability } \omega. \end{cases} \tag{17}$$

Here ω and $\theta < 0$ represent the probability and magnitude of a sharp drop in consumption growth (a disaster) relative to its mean. If $\theta > 0$ we have the opposite (a boom). Barro (2006) uses a more complex disaster distribution and Rietz (1988) allows time-dependence, but this is enough to make their point: that an infrequent extreme drop in consumption can

have a large impact on asset returns, even when we hold constant the mean and variance of log consumption growth.

With this structure, we can readily compute entropy and cumulants. The entropy of the two components follows from applying its definition (2):

$$L(e^{-\alpha w}) = (-\alpha)^2\sigma^2/2 \tag{18}$$

$$L(e^{-\alpha z}) = \log\left(1-\omega+\omega e^{-\alpha\theta}\right)+\alpha\omega\theta. \tag{19}$$

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

Two numerical examples, reported as the first two columns of Table 4, illustrate the potential quantitative significance of the disaster component. Column (1) has normal log consumption growth ($z = 0$). Column (2) incorporates a Bernoulli disaster. Parameters in both cases are chosen to match the target values in Table 3. The target values for the mean and variance of $\log g$ are $\kappa_1(\log g) = 0.020$ and $\kappa_2(\log g) = 0.035^2$. Finally, we add the disaster component. We set $\omega = 0.01$ and $\theta = -0.3$: a one percent chance of a 30 percent fall in (log) consumption relative to its mean. These choices are somewhat arbitrary, but they're similar to those of Barro and Rietz. Barro (2006) uses a distribution of disasters with an overall probability of 0.017 and a distribution whose mean is similar. Rietz uses smaller probabilities and larger disasters.

With these quasi-realistic numbers, we can explore the ability of the model to satisfy the Alvarez-Jermann bound. The observed equity premium implies that the entropy of the pricing kernel is at least 0.0440. Without disasters (that is, with $\omega = 0$), the logs of consumption growth and the pricing kernel are normal. The mean and variance of log consumption growth imply $\mu = 0.0200$ and $\sigma^2 = 0.035^2$. The Alvarez-Jermann bound implies $\alpha^2\kappa_2(\log g)/2 = \alpha^2 0.035^2/2 \geq 0.0440$ or $\alpha \geq 8.47$. We can satisfy the Alvarez-Jerman bound for the equity premium, but only with a risk aversion parameter greater than 8. There's a range of opinion about this, but some argue that risk aversion this large implies implausible behavior along other dimensions; see, for example, the extensive discussion in Campanale, Castro, Clementi (2007, Section 4.3) and the references cited there.

When we add the disaster component, a smaller risk aversion parameter suffices. We do this holding constant the mean and variance of log consumption growth, so the experiment has a partial derivative flavor: it measures the impact of high-order cumulants, holding constant the mean and variance. We choose μ to equate the mean growth rate to the sample mean: $\mu + \omega\theta = 0.0200$. Similarly, we choose σ to match the sample variance:

$$\sigma^2 + \omega(1-\omega)\theta^2 = \kappa_2(\log g) = 0.035^2.$$

The resulting parameter values are reported in the second column of Table 4. As long as $\omega < 1/2$ and $\theta < 0$, the disaster component z introduces negative skewness and positive

kurtosis into log consumption growth. Both are evident in first panel of Figure 1, where we plot cumulants 2 to 10 for log consumption growth. Each cumulant $\kappa_j(\log g)$ makes a contribution $\kappa_j(\log g)(-\alpha)^j/j!$ to the entropy of the pricing kernel. The next two panels of the figure show how the contributions depend on risk aversion. With $\alpha = 2$, negative skewness in consumption growth is converted 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 to 10, where the contribution of high-order cumulants is now greater than that of the variance. This is Yaron's bazooka in action: even modest high-order cumulants make significant contributions to entropy if α is large enough.

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

It's essential that the extreme events be disasters. If we reverse the sign of θ, the result is the line labelled "booms" in Figure 2. We see that for every value of α, entropy is below even the normal case. The impact of high-order cumulants is apparently negative. Table 5 shows us exactly how this works. With Bernoulli disasters (and $\alpha = 10$), the entropy of the pricing kernel (0.1614) comes from the variance (0.0613), odd high-order cumulants (0.0621), and even high-order cumulants (0.0380). When we switch from disasters to booms, the odd cumulants change sign — see equation (15) — reducing overall entropy. Another example illustrates the role of the probability and magnitude of the disaster. Suppose we halve θ and double ω, with σ adjusting to maintain the variance of consumption growth. Then entropy falls sharply and the contribution of high-order cumulants almost disappears. In this sense, both the low probability and large magnitude in the example are quantitatively important.

We've chosen to focus on the entropy of the pricing kernel, but you get a similar picture in this setting if you look at the equity premium. The short rate $r^1 = 1/q^1 = 1/Em$ is constant in our iid environment. We define "levered equity" as a claim to the dividend $d_t = c_t^\lambda$. This isn't, of course, either equity or levered, but it's a convenient functional form that is widely used in the macro-finance literature to connect consumption growth (the foundation for the pricing kernel) to returns on equity (the asset of interest). In the iid case, the log return is a linear function of log consumption growth:

$$\log r^e_{t+1} \quad = \quad \lambda \log g_{t+1} + \text{constant}. \tag{20}$$

See Appendix A.4. 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 0.15^2, so λ is the ratio of the standard deviation of the excess return (0.15) to the standard deviation of log consumption growth (0.035), approximately 4.3.

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

3.2 Example 2: Poisson disasters

We turn now to a more realistic model of disasters. The model is a Poisson-normal mixture in which we add a random number of "jumps" to log consumption growth. The distribution over the number of jumps is Poisson and the jumps themselves are normal. The added complexity has a number of benefits. One is that it gives us a better approximation to the empirical distribution of disasters. Another is that this specification is easily scaled to different time intervals. 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. It also allows a more direct comparison to the estimates of option pricing models. In the macro-finance literature, it has been applied by Bates (1988), Martin (2007), and Naik and Lee (1990).

We continue with the two-component structure, equation (16), with one component normal and the other a Poisson-normal mixture. The central ingredient of the second component is a Poisson random variable that takes on nonnegative values j (the number of jumps) with probabilities $e^{-\omega}\omega^j/j!$. Here ω is the average number of jumps per year. Conditional on j, the second component is normal:

$$z_t|j \quad \sim \quad \mathcal{N}(j\theta, j\delta^2) \text{ for } j = 0, 1, 2. \ldots \tag{21}$$

This differs from the Bernoulli model in two respects: there is a positive probability of more than one jump and the jump size has a distribution rather than fixed size. If ω is small, the first is insignificant but the second increases entropy and high-order cumulants. The entropy of this component of the pricing kernel is

$$L(e^{-\alpha z}) \quad = \quad \omega(e^{-\alpha\theta+(\alpha\delta)^2/2} - 1) + \alpha\omega\theta. \tag{22}$$

This and other properties of Poisson-normal mixtures are derived in Appendix A.3. Therefore, the entropy of the pricing kernel is

$$L(m) \quad = \quad (-\alpha\sigma)^2/2 + \omega(e^{-\alpha\theta+(\alpha\delta)^2/2} - 1) + \alpha\omega\theta, \tag{23}$$

the sum of the entropies of the normal and Poisson-normal components.

We illustrate the properties of this example with numbers similar to those used in the Bernoulli example. With $\omega = 0.01$, there is probability 0.9900 of no jumps, 0.0099 of one jump, and 0.0001 of more than one jump. With larger values of ω the probability of multiple jumps can be substantial, but in this example it's miniscule. We set the mean jump size $\theta = -0.3$, the same number we used earlier. The only significant change is the dispersion of jumps: we set $\delta = 0.15$. These parameter values are close to those suggested by Barro, Nakamura, Steinsson, and Ursua (2009). Finally, we choose μ and σ to match the sample mean and variance of log consumption growth. In the model, the mean is $\mu + \omega\theta$ and the variance is $\sigma^2 + \omega(\theta^2 + \delta^2)$. Given the parameters of the second component, we choose μ and σ to match our target values of the mean and variance of log consumption growth. All of these parameters (and more) are listed in column (3) of Table 4.

This example is qualitatively similar to the previous one, but the dispersion in disasters generates greater entropy. The contributions of high-order cumulants are summarized in Figure 4 and Table 5. Figure 5 shows that the model satisfies the Alvarez-Jermann bound at smaller values of α. We match the equity premium with $\alpha = 5.38$, smaller than the value of 6.59 needed for the Bernoulli example; see Tables 4 and 5.

4 Risk-neutral probabilities in representative-agent models

As a warmup for our study of options, we consider the risk-neutral probabilities implied by the examples of the previous section. In general, risk aversion ($\alpha > 0$) generates risk-neutral distributions that are shifted left (more pessimistic) relative to true distributions. The form of this shift depends on the distribution.

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

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

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

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

Thus risk-neutral probabilities have the same form (normal) with mean $\mu^* = \mu - \alpha\sigma^2$ and standard deviation $\sigma^* = \sigma$. The mean shifts the distribution to the left by an amount that depends on risk aversion. The log probability ratio is

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

which implies the cumulant-generating function

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

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

$$L(p^*/p) \;=\; (\alpha\sigma)^2/2,$$

which is what we reported in equation (18).

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

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

Thus p^* is Bernoulli with probability

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

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

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

implies the cumulant-generating function

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

Entropy is therefore

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

which is what we saw in equation (19).

In our final example, consumption growth is a Poisson-normal mixture. The Poisson-normal mixture (21) is based on a state space that includes both the number of jumps and the distribution conditional on the number of jumps: say (j,z). The same logic we used in the other examples then tells us that the risk-neutral distribution has the same form, with parameters

$$\omega^* \;=\; \omega\exp(-\alpha\theta+(\alpha\delta)^2/2),\;\; \theta^* \;=\; \theta-\alpha\delta^2,\;\; \delta^* \;=\; \delta. \tag{24}$$

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 reported in (22).

We won't bother with multi-component models, but they follow similar logic. If log consumption growth is the sum of independent components, then entropy is the sum of the entropies of the components, as in equation (23).

5 Disasters in option models and data

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

5.1 The Merton model

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

$$\log r^e_{t+1} - \log r^1 \quad = \quad w_{t+1} + z_{t+1}. \tag{25}$$

We use the return, rather than the price, because it fits neatly into our iid framework, 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 two 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 $(\omega^*, \theta^*, \delta^*)$.

Related work supports a return process with these features. Ait-Sahalia, Wang, and Yared (2001) report a discrepancy between the risk-neutral density of S&P 500 index returns implied by the cross-section of options and the time series of the underlying asset returns, but conclude that the discrepancy can be resolved by introducing a jump component. One might go on to argue that two jumps are needed: one for macroeconomic disasters and another for more frequent but less extreme financial crashes. However, Bates (2009) studies the US stock market over the period 1926-2006 and shows that a second jump component plays no role in accounting for macroeconomic events like the Depression.

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

$$\begin{aligned} L(m) &= L(p^*/p) \\ &= \frac{1}{2}\left(\frac{\mu-\mu^*}{\sigma}\right)^2 + (\omega^*-\omega) + \omega\left[\log\frac{\omega}{\omega^*} - \log\frac{\delta}{\delta^*} + \frac{(\theta-\theta^*)^2+(\delta^2-\delta^{*2})}{2\delta^{*2}}\right]. \quad (26) \end{aligned}$$

The source of this expression and the corresponding cumulant-generating function are reported in Appendix A.5.

5.2 Parameter values

We use parameter values from Broadie, Chernov, and Johannes (2007), who summarize and extend the existing literature on equity index options. The parameters of the true distribution are estimated from the time series of excess returns on equity. We use the parameters of the Poisson-normal mixture — namely (ω, θ, δ) — reported in Broadie, Chernov, and Johannes (2007, Table I, the line labelled SVJ EJP). These estimates also include stochastic volatility, which we ignore because it conflicts with our iid structure. The estimated jump intensity ω is 1.512, which implies much more frequent jumps than we used in our consumption-based model. With this value, the probability of 0 jumps per year is 0.220, 1 jump per year 0.333, 2 jumps 0.25, 3 jumps 0.13, 4 jumps 0.05, and 7 or more jumps about 0.001. Properties related to extreme events can be difficult to estimate precisely, as we noted in the introduction. The issue is their frequency. If they occur (say) once every hundred years, a long dataset is a necessity. If they're more frequent, as estimates based on US stock returns imply, we can get precise estimates from a finer time interval, even over shorter samples. Given parameters for the Poisson-normal component, μ and σ are chosen to match the mean and variance of excess returns to their target values. 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. The results are reported in Table 4.

The risk-neutral parameters for the Poisson-normal mixture are estimated from the cross section of option prices: specifically, prices of options on the S&P 500 over the period 1987-2003. The depth of the market varies both over time and by the range of strike prices, but there are enough options to allow reasonably precise estimates of the parameters. The numbers we report in Table 4 are from Broadie, Chernov, and Johannes (2007, Table IV, line 5). In practice, option prices identify only the product $\omega^*\theta^*$, so they set $\omega^* = \omega$ and choose θ^* and δ^* to match the level and shape of the implied volatility smile. Finally, μ^* is set to satisfy (8), which implies $\mu^* + \sigma^2/2 + \omega^*[\exp(\theta^* + \delta^{*2}/2) - 1] = 0$.

Figure 6 shows how θ^* and δ^* are identified from the cross section of 3-month option prices. We express prices as implied volatilities and graph them against "moneyness," with higher strike prices to the right. In the data, we measure moneyness as the log of the ratio of the strike price to the spot price. In the figure, the solid line illustrates the slope and shape of 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. They also illustrate how the parameters are identified. Consider three values of θ^*: -0.0259 $(= \theta)$, $+0.0259$, and the estimated value -0.0482. We plot volatility smiles for all three values with $\delta^* = \delta$. Evidently θ^* controls the slope. When it's positive, the smile slopes upward, and when it's a smaller negative value than our estimate the smile is flatter. All of these smiles lie below the estimated one. The value of δ^* affects the level and curvature of the smile. This is evident from the smile for $\delta^* = \delta$, a substantially smaller value. The combination produces an implied volatility smile that closely approximates the slope and curvature we see in the data.

5.3 Pricing kernels implied by options

In Table 4 we report the parameters of the option model and some of their implications, and in Table 5 we report the entropy of the pricing kernel and its components. Three features of the option model deserve emphasis.

The first and most important feature of the option model is that odd high-order cumulants make a substantial contribution to entropy. In this respect, the option model agrees with the macroeconomic models of disasters we examined earlier. The contribution of odd high-order cumulants is larger than we saw in macroeconomic models, but smaller as a fraction of entropy.

The second feature is that disasters are more moderate in our option model [column (4)] than in a similar model based on consumption data [column (3)]. The units are different (consumption growth v. returns), so it's not a direct comparision unless the two are linked — as they are in our equation (20). With this caveat in mind, note that standardized measures of skewness and kurtosis are substantially smaller in the option model than in the model based on consumption evidence. This is true whether we look at the true distribution, the risk-neutral distribution, or the distribution of the pricing kernel. The same holds for tail probabilities: probabilities of extreme negative realizations of consumption growth or the return on equity. Outcomes more than 3 and 5 standard deviations to the left of the mean are more likely in the model based on consumption evidence [column (3) of the table] than in the model based on option prices [column (4)].

The third feature is that entropy is significantly higher. While the macro models had entropy less than 0.1, the option model implies entropy of almost 0.8. This reflects, in large part, the high price of options. The prices are high in the sense that selling them generates high average returns; see, for example, Broadie, Chernov, and Johannes (2009). These high average returns imply high entropy via the Alvarez-Jermann bound. Evidently a bound based on the equity premium is too loose: other investment strategies generate higher average excess returns and therefore imply higher entropy.

We look more closely at the differences between consumption- and option-based models in the next section.

6 Comparing consumption- and option-based models

We've seen that disasters implied by options are considerably different from those apparent in consumption data. We explore the difference further in this section, looking at option prices implied by consumption growth, consumption growth implied by option prices, and risk aversion implicit in options data. These comparisons highlight the relations between true and risk-neutral probabilities implied by consumption- and option-based models, respectively.

Consider first the option prices implied by our consumption-based model. In essence, we are taking the true distribution of consumption growth implied by consumption data, computing the risk-neutral distribution by applying power utility, and using the risk-neutral distribution to compute option prices. The only missing link is the connection between consumption growth (the natural random variable for consumption-based models) and equity returns (the natural variable for option-based models). In our environment, the two are connected by (20): equity returns are a log-linear function of consumption growth with slope λ. To convert the consumption growth process to a return process, we multiply θ and δ by λ and keep the jump intensity ω the same. Risk-neutral parameters then follow from (24).

Implied volatility smiles for the consumption-based model are pictured in Figure 7 along with those for the estimated Merton model. Similar consumption-based volatility smiles are reported by Benzoni, Collin-Dufresne, and Goldstein (2005) and Du (2008). What's new is the explicit comparison to an estimated option pricing model. We report two smiles in each case, corresponding to 3-month and 12-month options. The two models are clearly different. The consumption-based model has a steeper smile, greater curvature, and lower at-the-money volatility. The reason, again, is that it has both 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). Stated simply: the difference in extreme outcomes between models based on macroeconomic and options data results in significantly different option prices. This is evident, for example, in the large difference in risk-neutral skewness and kurtosis reported in columns (3) and (4) of Table 4.

Now consider the reverse: the consumption growth process implied by the risk-neutral distribution of equity returns indicated by option prices. Here we are taking the risk-neutral distribution and computing the true distribution using (again) power utility. This procedure places more structure on the problem than we used in the option model — namely, power utility. Given this structure, we can infer the true distribution without relying on the limited time series evidence available for estimating it directly from consumption. Again, we need to rescale the parameters, dividing θ^* and δ^* by λ. Given these values, we compute the parameters of the true distribution using (24). Finally, we set σ to match our target for the standard deviation of log consumption growth and α to match the equity premium. The results are reported in column (5) of Table 4.

The consumption process derived this way from option prices generates disasters in the sense that large declines in consumption growth are substantially more likely than in a lognormal model. They are, however, more moderate than we see in macroeconomic data. This is evident in the standardized measures of skewness and excess kurtosis. It's also evident in the probabilities of tail events. The probability of (at least) 3 standard deviation declines is similar in the two models. Roughly speaking, there's just under a 1 percent chance of a drop in consumption similar to the US in the Depression. However, 5 standard deviation declines are much more likely in the consumption-based model [column (3)] than in the option-based model [column (5)]. The risk aversion parameter must be larger to compensate for the more modest disasters. We see this directly in Table 5, where the contribution of high-order cumulants is much smaller than in the Poisson model based on consumption evidence.

Both of these comparisons use power utility to connect true and risk-neutral parameters, yet the parameter values of the option model are inconsistent with power utility. One example is the difference between δ and δ^*, which is zero with power utility [equation 24)]. Nevertheless, we can derive something like risk aversion from our option model. Note that in (13), risk aversion is implicit in the relation between the pricing kernel and consumption growth:

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

In the option model, the analogous expression is

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

See Leland (1980). In our setting, the second term is λ, so the action is in the first term.

Risk aversion defined this way need not have much to do with the risk aversion of individual agents, but it's a useful way of describing how the market prices risk. Option prices imply that RA depends, in general, on the state; see Appendix A.6. In our case, it's larger for negative returns than for positive ones, with risk aversion of 14 for returns of –10% and 3.5 for returns of +10%. Related work has generated a wide range of patterns, but they all find that risk aversion varies with the state. See, for example, Ait-Sahalia and Lo (2000), Jackwerth (2000), Rosenberg and Engle (2002), and Ziegler (2007). What we find interesting is the possibility that risk premiums on assets might reflect not only disasters in outcomes but pricing of disasters that gives them greater weight than power utility.

7 Summary and extensions

We have described, in a relatively simple theoretical setting, how option prices can be used to infer the probabilities of disasters, including the infrequent sharp declines in consumption

growth documented in macroeconomic data by Barro (2006) and others. Options on the S&P 500 index value bad outcomes more than good ones, and in this sense are similar to the macroeconomic evidence. The disasters implied by option prices are, however, more modest. The analysis that leads to these conclusions leans heavily on three supports: power utility over aggregate consumption, iid consumption growth, and a close connection between aggregate dividends and consumption. Each deserves a closer look.

Perhaps the most interesting extension of the theory is to consider going beyond power utility or even the representative agent framework. Power utility is the workhorse of macroeconomics and finance, but our option model suggests greater aversion to bad outcomes than good ones. If this turns out to be a robust feature of the evidence, it's worth thinking about where it comes from. One possibility is explore alternative preferences, including skewness aversion (Harvey and Siddique, 2000), Chew-Dekel risk preferences (reviewed by Backus, Routledge, and Zin, 2004), state-dependent preferences (Chabi-Yo, Garcia, and Renault, 2008), ambiguity (applied to options by Drechsler, 2008, and Liu, Pan, and Wang, 2005), learning (Shaliastovich, 2008), and habits (Du, 2008). Another promising avenue is heterogeneity across agents. Certainly there is clear evidence of imperfect risk-sharing across individuals and good reason to suspect that this might affect asset prices. Alvarez, Atkeson, and Kehoe (2009), Bates (2008), Chan and Kogan (2002), Guvenen (2009), and Lustig and Van Nieuwerburgh (2005) are notable examples. The question for us is whether these extensions provide a persuasive explanation for prices of equity index options.

Another interesting extension is time-dependence. There's overwhelming evidence that short-term interest rates, implied volatilities, and expected returns on a variety of assets change through time. None of this is consistent with our iid setting. The question is whether time-dependence is quantitatively important in assessing the role of extreme events, particularly their impact on the entropy of the pricing kernel. It's possible the role is small. We know, for example, that the variance of the conditional mean of the pricing kernel is much less than the mean conditional variance; see Cochrane and Hansen (1992, Section 2.7). A similar relation holds for entropy. Nevertheless, recent work by Drechsler and Yaron (2008) and Wachter (2008) suggest that time-variation in the distribution over extreme events can be quantitatively important for asset pricing.

The third extension is to loosen the link between dividends and consumption. We've followed a long tradition in tying dividends to consumption. The tradition is largely a matter of convenience, because it's simpler to have one random variable rather than two. Our focus on the pricing kernel and its entropy reinforces this message, since neither depends on the dividend process. Nevertheless, the work of Bansal and Yaron (2007), Gabaix (2009), and Longstaff and Piazzesi (2004) suggests that this extension has promise in accounting for the behavior of equity prices and returns.

A Appendix

A.1 The Alvarez-Jermann bound

The Alvarez-Jermann bound (3) is a byproduct of Proposition 2 in Alvarez and Jermann (2005). The proof goes like this:

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

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

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

$$E \log r \quad \leq \quad -E \log m. \tag{27}$$

  In finance, 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:

$$L(q^1) \quad = \quad \log Eq^1 - E \log q^1 \quad = \quad \log Em + E \log r^1. \tag{28}$$

- Alvarez-Jermann bound. (27) and (28) imply

$$L(m) \quad \geq \quad E\left(\log r^j - \log r^1\right) + L(q^1).$$

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

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

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

where the expression on the right is the Sharpe ratio. The bound depends on

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

where $k(s)$ is the cumulant generating function of $\log m$. Because $k(s)$ depends on the high-order cumulants of $\log m$, the bound does, too. The squared Sharpe ratio is bounded above by

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

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

A.2 Entropy and cumulants of Bernoulli random variables

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

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

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

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

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

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

imply the cumulants

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

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

A.3 Entropy and cumulants of Poisson-normal mixtures

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

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

which ensures 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 directly.

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

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

which implies the cgf

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

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

We find cumulants the usual way, taking derivatives of k. The first five are

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

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

A.4 Equity premium and cumulants with power utility

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 a little obscure, but it's wonderfully compact. The idea and many of the results follow Martin (2008).

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

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

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

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

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

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

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

Thus we have, in compact notation,

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

The equity premium is therefore

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

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

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

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

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

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

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

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

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

A.5 Cumulant-generating functions based on 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 (25), where $z = 0$ and w has true distribution of $\mathcal{N}(\mu, \sigma^2)$ and risk-neutral distribution $\mathcal{N}(\mu^*, \sigma^{*2})$. The density functions are

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

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

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

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

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

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

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

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

Entropy is minus the first derivative evaluated at zero:

$$-\kappa_1(\log p^*/p) = (1/2)[\log\varphi + 1 - \varphi] + (\mu - \mu^*)^2/2\sigma^{*2}. \tag{29}$$

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

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

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

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

Using (29) 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

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

Entropy is minus the first derivative evaluated at zero:

$$\begin{aligned} - & \quad \kappa_1(\log p^*/p) \\ = & \quad (\omega^* - \omega) + \omega[\log(\omega/\omega^*) - 1/2 \cdot \log\varphi + 1/2 \cdot (\varphi - 1)] + \omega(\theta - \theta^*)^2/2\delta^{*2}. \end{aligned} \tag{30}$$

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 (29) with $\varphi = 1$] and the entropy of the Poisson mixture of normals [equation (30)].

A.6 Risk aversion implied by the Merton model

We compute risk aversion as:

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

Given the normal distribution of jumps, the density conditional on the number of jumps j is

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

The (marginal) density for log-returns is

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

Therefore,

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

A similar expression holds for the risk-neutral distribution.

As a result, implied risk aversion is

$$\begin{aligned} \text{RA}/\lambda &= \frac{1}{\sum_{j=0}^{\infty} p^*(\log r^e, j)} \sum_{j=0}^{\infty} p^*(\log r^e, j) \cdot (\log r^e - \mu^* - j\theta^*)/(\sigma^2 + j\delta^{*2}) \\ &- \frac{1}{\sum_{j=0}^{\infty} p(\log r^e, j)} \sum_{j=0}^{\infty} p(\log r^e, j) \cdot (\log r^e - \mu - j\theta)/(\sigma^2 + j\delta^2). \end{aligned} \tag{31}$$

Note that RA is a function of the state through $\log r^e$.

References

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

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

Alvarez, Fernando, Andrew Atkeson, and Patrick J. Kehoe, 2009, "Time-varying risk, interest rates, and exchange rates in general equilibrium," *Review of Economic Studies* 76, 851-878.

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

Backus, David K., Bryan R. Routledge, and Stanley E. Zin, 2004, "Exotic preferences for macroeconomists," *NBER Macroeconomics Annual* 19, 319-390.

Bansal, Ravi, and Amir Yaron, 2007, "The asset pricing-macro nexus and return-cash flow predictability," manuscript, June.

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

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

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

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

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

Bates, David, 2009, "US stock market crash risk, 1926-2006," manuscript, March.

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

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

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

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

Buhlmann, Hans, Freddy Delbaen, Paul Embrechts, and Albert Shiryaev, 1996, "No-

arbitrage, change of measure, and conditional Esscher transforms," *CWI Quarterly* 9, 291-317.

Campanale, Claudio, Rui Castro, Gian Luca Clementi, 2007, "Asset pricing in a production economy with Chew-Dekel preferences," manuscript, April.

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

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

Cochrane, John C., and Lars Peter Hansen, 1992, "Asset pricing explorations for macroeconomics," *NBER Macroeconomics Annual* 7, 115-165.

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

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

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

Gabaix, Xavier, 2009, "Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance," manuscript, March.

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

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

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

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

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

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

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

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

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

Martin, Ian, 2008, "Consumption-based asset pricing with higher cumulants," manuscript, March.

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

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

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

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

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

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

Shiller, Robert, 2007, "Long term stock, bond, interest rate and consumption data," spreadsheet posted on his web site.

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

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

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

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

Table 1
Properties of consumption growth and asset returns

Variable	Mean	Std Dev	Skew	Kurt	Auto
Consumption and returns, 1889-2006					
Consumption growth	0.0200	0.0353	–0.35	1.10	–0.07
Return on one-year bond	0.0182	0.0573	0.03	2.29	0.35
Return on equity	0.0622	0.1737	–0.50	0.18	0.04
Excess return on equity	0.0440	0.1748	–0.60	0.71	0.07
Consumption and returns, 1986-2006					
Consumption growth	0.0186	0.0131	–0.59	–0.20	0.48
Return on one-year bond	0.0221	0.0190	–0.45	–0.68	0.41
Return on equity	0.0845	0.1470	–0.58	–0.52	0.15
Excess return on equity	0.0625	0.1397	–0.67	–0.58	0.17

Notes. Entries are sample moments. Mean is the sample mean, Std Dev is the standard deviation, Skew is skewness, Kurt is excess kurtosis, and Auto is the first autocorrelation. Consumption growth is $\log(c_t/c_{t-1})$ where c is real per capita consumption. Returns are logarithms of gross real returns and the excess return is the difference between the log-returns on equity and the one-year bond. The one-year bond is the treasury security of maturity closest to one year. Equity is the S&P 500. Consumption and return data are from Shiller's web site (Shiller, 2007).

Table 2
Mean implied volatilities on S&P 500 options

	Log of Ratio of Strike Price to Spot Price		
Maturity	–0.04	0.0	0.04
1 month	0.2157	0.1829	0.1653
3 months	0.2052	0.1865	0.1719
12 months	0.1959	0.1858	0.1761

Notes. Implied volatilities are derived from S&P 500 options. The data are annual, 1986-2006, except for 12-month options, which start in 1991.

Table 3
Target values for model economies

Property	Value
Mean of log consumption growth, $E \log g$	0.0200
Standard deviation of log consumption growth, $\mathrm{Var}(\log g)^{1/2}$	0.0350
Equity premium, $E(\log r^e - \log r^1)$	0.0440
Standard deviation of equity excess return, $\mathrm{Var}(\log r^e - \log r^1)^{1/2}$	0.1500
Implied volatility (strike = price)	0.1800

Table 4
Parameter values and properties of model economies

Parameter	Normal Cons Gr (1)	Bernoulli Cons Gr (2)	Poisson Cons Gr (3)	Merton Returns (4)	Implied Cons Gr (5)
Preferences					
α	10.52	6.59	5.38	—	10.07
True distribution					
μ	0.0200	0.0230	0.0230	0.0832	0.0283
σ	0.0350	0.0183	0.0100	0.1377	0.0212
ω	—	0.0100	0.0100	1.5120	1.3864
θ	—	–0.3000	–0.3000	–0.0259	–0.0060
δ	—	—	0.1500	0.0407	0.0229
Risk-neutral distribution					
μ^*	0.0071	0.0208	0.0225	0.0547	0.0238
ω^*	—	0.0680	0.0695	1.5120	1.5120
θ^*	—	–0.4210	–0.4210	–0.0482	–0.0112
δ^*	—	—	0.1500	0.0981	0.0229
Skewness, kurtosis, and tail probabilities					
γ_1 (true)	0	–6.11	–11.02	–0.07	–0.31
γ_2 (true)	0	50.26	145.06	0.05	0.87
γ_1^* (risk-neutral)	0	–3.15	–4.33	–0.32	–0.53
γ_2^* (risk-neutral)	0	8.72	20.20	0.46	0.91
γ_1 ($\log m$)	0	6.11	11.02	–0.08	0.31
γ_2 ($\log m$)	0	50.26	145.06	2.16	0.87
Tail prob (≤ -3 st dev)	0.0013	0.0100	0.0090	0.0040	0.0086
Tail prob (≤ -5 st dev)	0.0000	0.0100	0.0079	0.0000	0.0002
Entropy					
$L(m) = L(p^*/p)$	0.0678	0.0478	0.0449	0.7647	0.0650

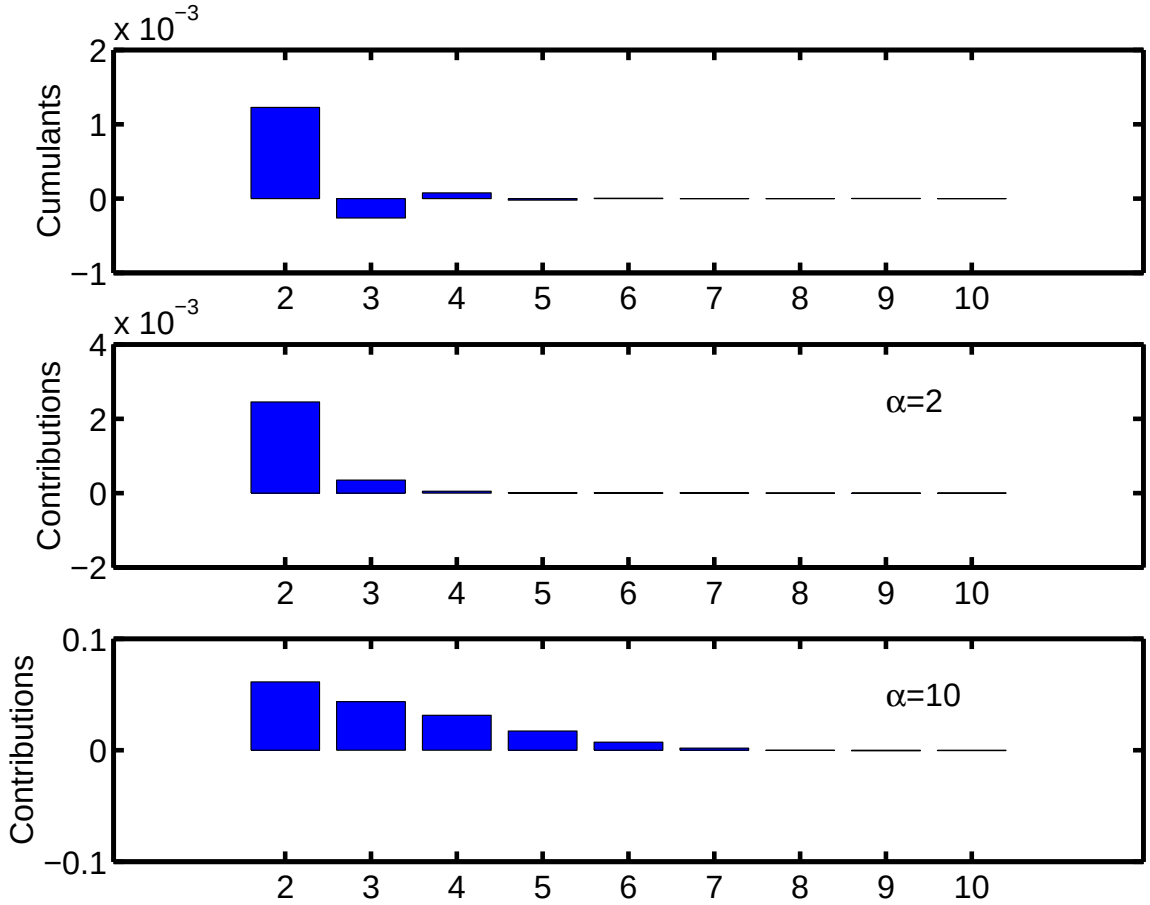
Notes. Entries are parameters and properties of examples with different specifications of disasters. Columns (1)-(3) and (5) are consumption-based models in which log consumption growth has a standard deviation of 0.0350 and risk aversion α is chosen to match the equity premium (0.0440). Column (4) is the Merton model parameterized to option prices and equity returns. Column (5) takes this model, scales the risk-neutral parameters to fit consumption growth, and sets the true parameters by applying the relations implied by power utility [equation (24)]. γ_1 and γ_2 are the traditional measures of standardized skewness and excess kurtosis, defined in equation (5). We report versions for the true distribution of log consumption growth or the log return on equity, the risk-neutral distribution, and the distribution of the pricing kernel. Tail probabilities refer to the probabilities that log consumption growth or the log return on equity are less than -3 and -5 standard deviations, respectively, from their mean.

Table 5
Contributions to entropy of the pricing kernel

Model	Entropy	Variance/2	High-Order Cumulants	
			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
$\alpha = 10.52^*$	0.0678	0.0678	0	0
Bernoulli consumption growth				
$\alpha = 2$	0.0029	0.0025	0.0004	0.0000
$\alpha = 5$	0.0234	0.0153	0.0060	0.0021
$\alpha = 10$	0.1614	0.0613	0.0621	0.0380
$\alpha = 10, \theta = +0.3$ (boom)	0.0372	0.0613	–0.0621	0.0380
$\alpha = 10, \theta = -0.15, \omega = 0.02$	0.0765	0.0613	0.0115	0.0038
$\alpha = 6.59^*$	0.0478	0.0266	0.0147	0.0065
Poisson consumption growth				
$\alpha = 2$	0.0033	0.0025	0.0007	0.0002
$\alpha = 5$	0.0356	0.0153	0.0132	0.0071
$\alpha = 10$	0.5837	0.0613	0.2786	0.2439
$\alpha = 5.38^*$	0.0449	0.0177	0.0173	0.0099
Models fit to option prices				
Merton equity returns	0.7647	0.4699	0.1130	0.1819
Implied consumption growth	0.0650	0.0621	0.0023	0.0006

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

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



Notes. The top panel summarizes the cumulants of log consumption growth, $\kappa_j(\log g)$. The next two panels summarize their contributions to the entropy of the pricing kernel, $\kappa_j(\log m) = (-\alpha)^j \kappa_j(\log g)/j!$, for risk aversion α equal to 2 and 10, respectively.

Figure 2
Bernoulli disasters: entropy of the pricing kernel

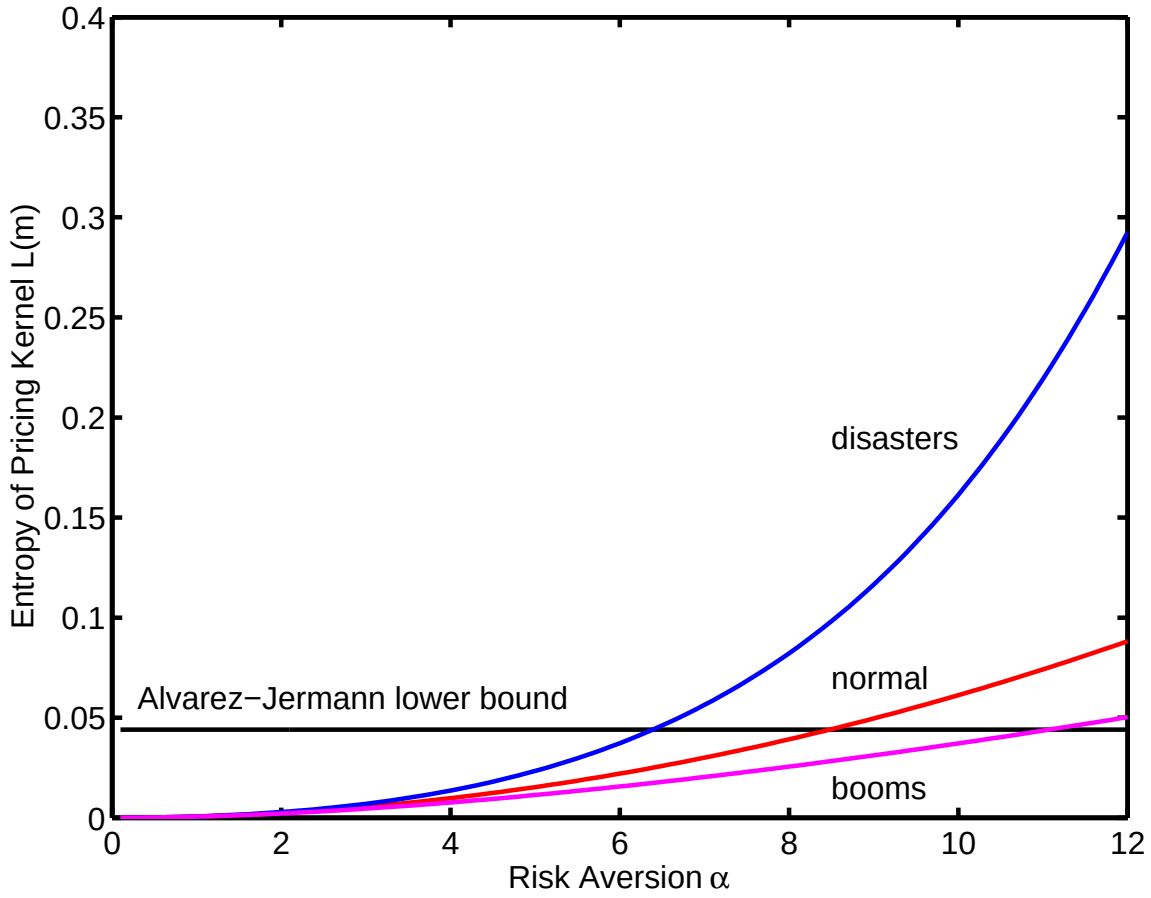


Figure 3
Bernoulli disasters: entropy and the equity premium

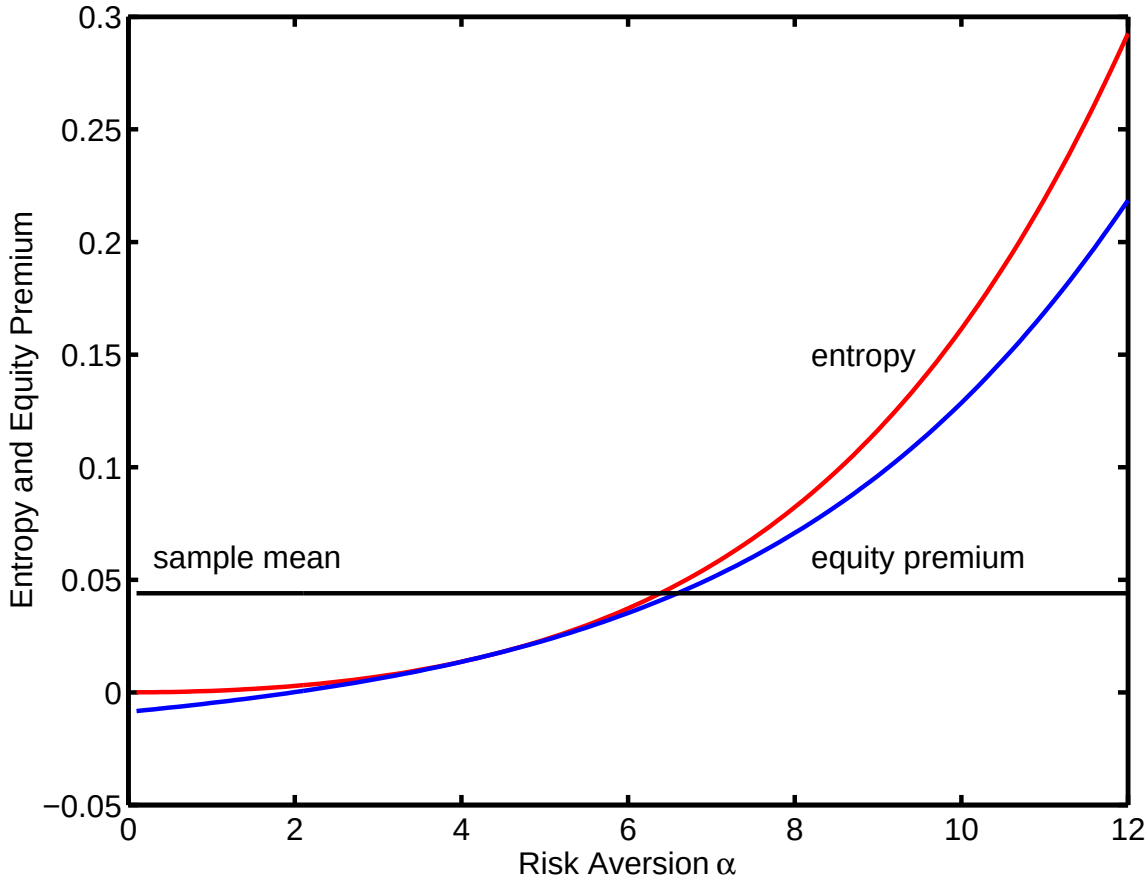
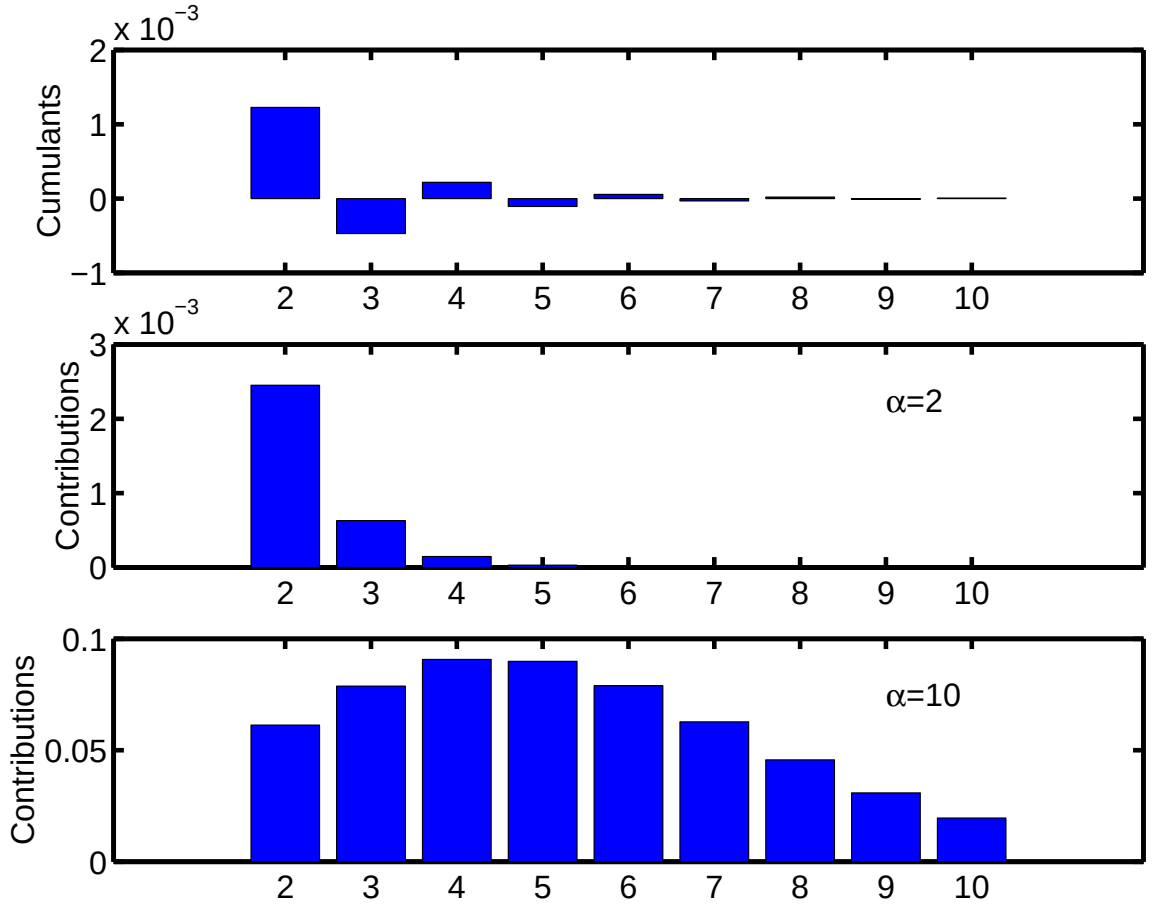


Figure 4
Poisson disasters: cumulants of log consumption growth and contributions to entropy



Notes. See Figure 1.

Figure 5
Poisson disasters: entropy and equity premium

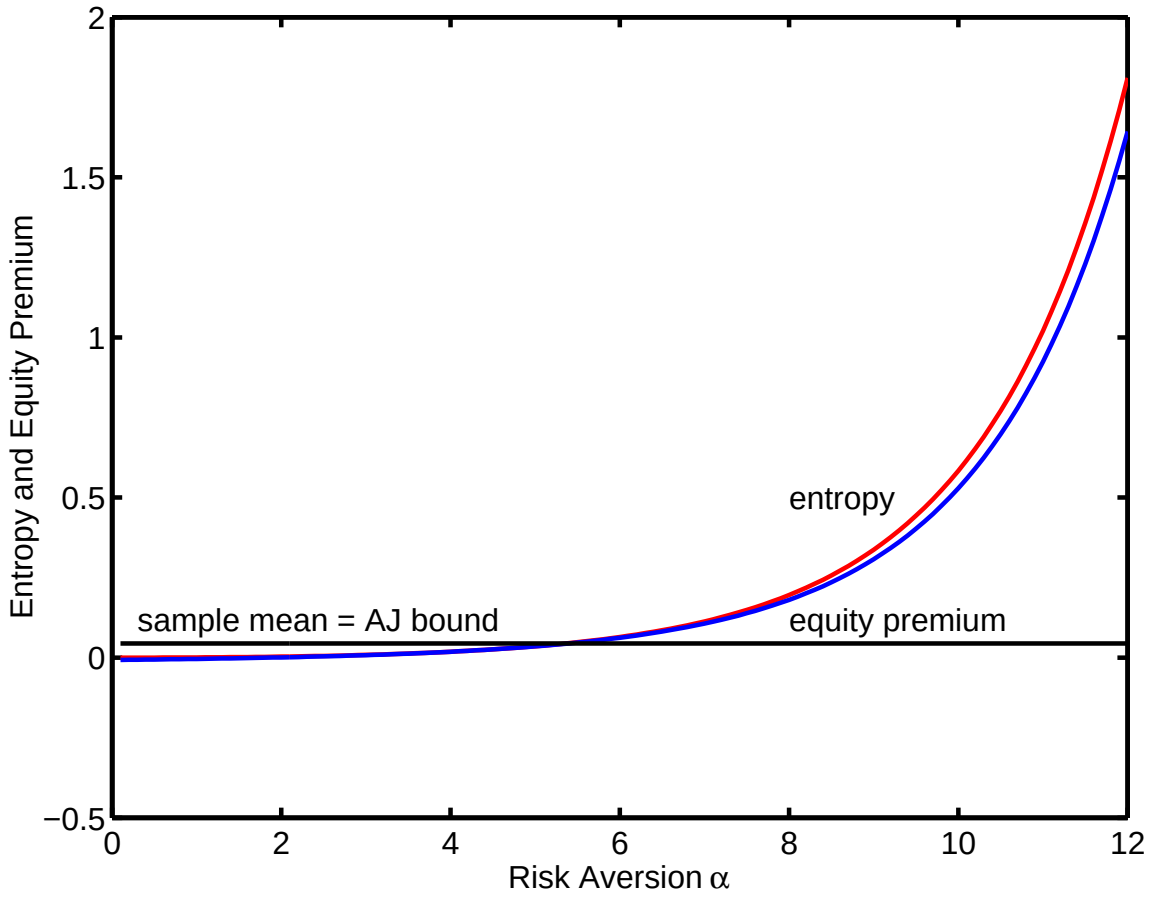
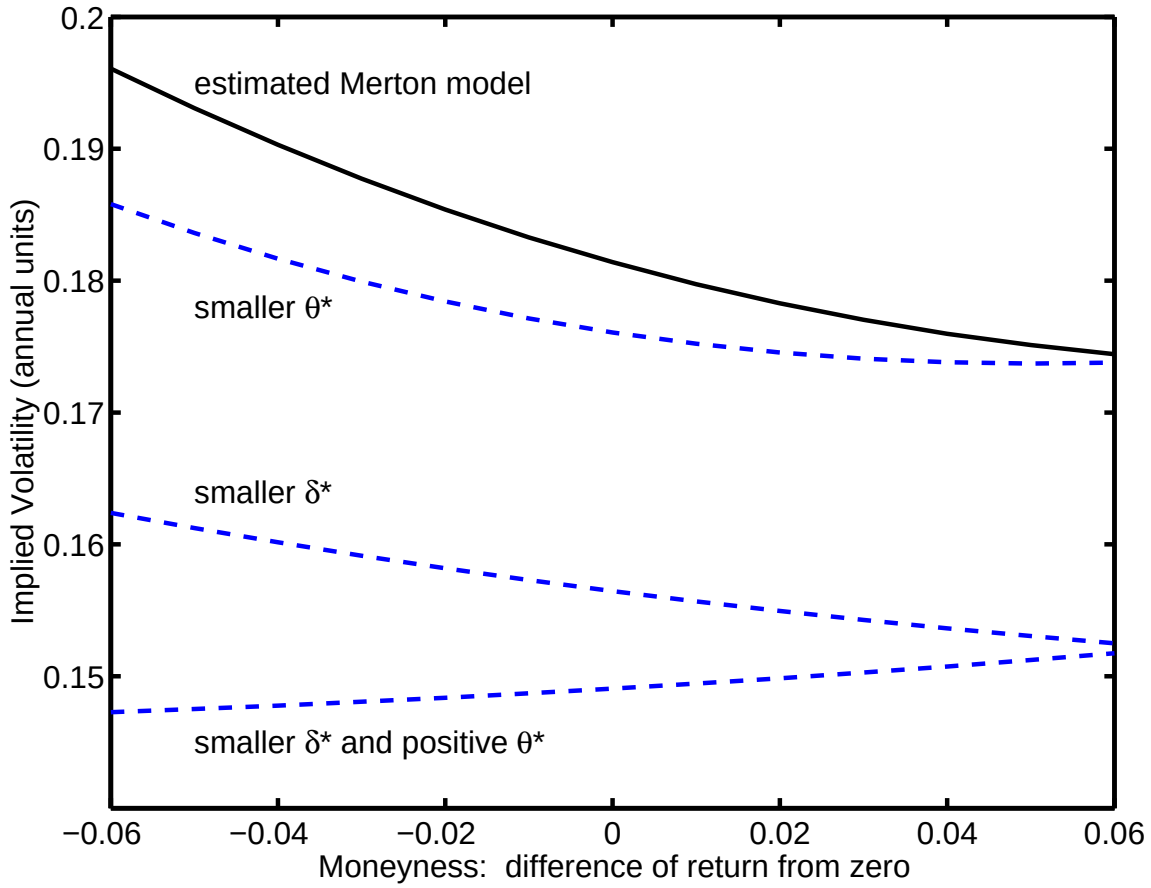


Figure 6
Implied volatility smiles for 3-month options



Notes. The lines represent implied volatility "smiles" for the Merton model with estimated parameters and some alternatives. Moneyness is measured as the difference of the return from zero or, equivalently, the proportional difference of the strike from the price.

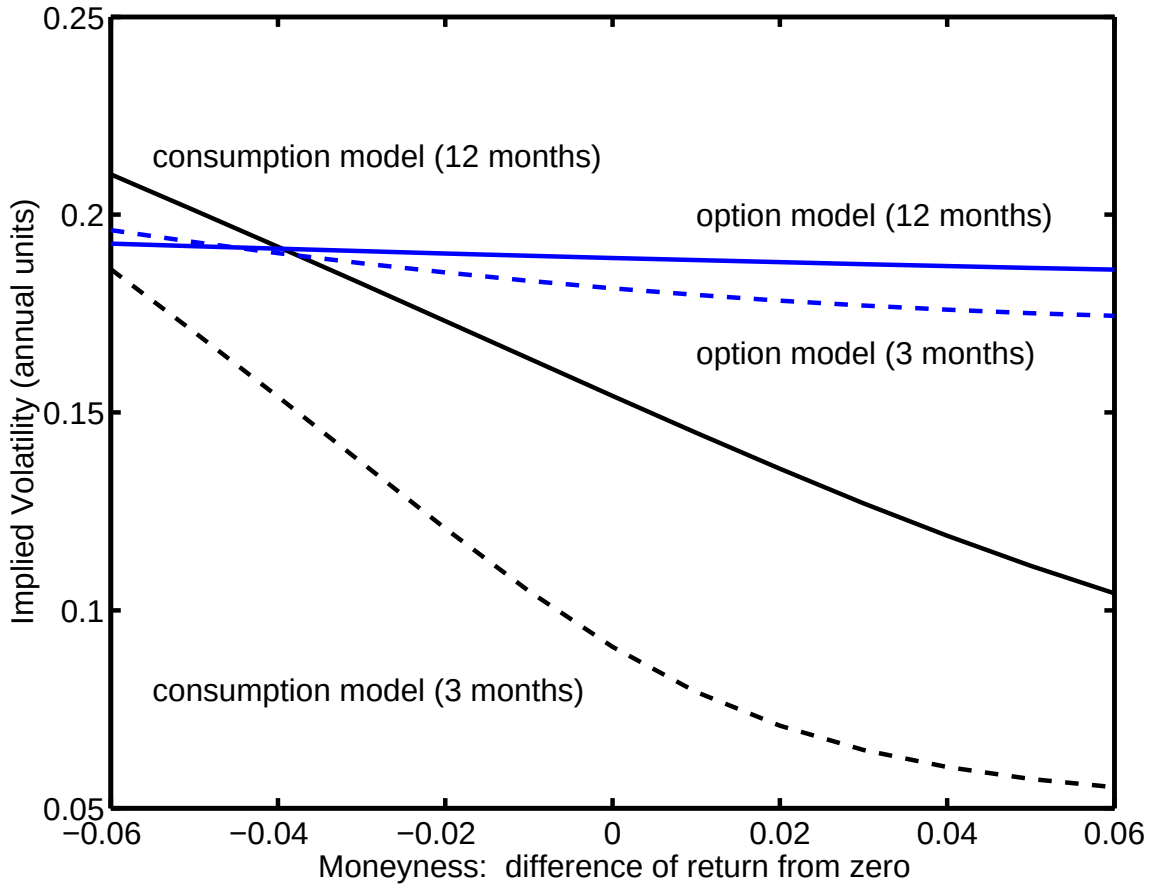



Figure 7
Implied volatility smiles based on option and consumption data