Backus, Chernov, & Martin

Notes on Disasters and Asset Returns*

(Started: July 20, 2008; Revised: September 23, 2008)

The primary challenge of disaster-based models, and departures from (log)normality in general, is that it's difficult to estimate the parameters of their distributions reliably. How much of this kind of thing is reasonable? Options are a promising source of information, since their prices reflect not only events that occur, but those that market participants think could have occurred. You need a complete model, though, because the risk-neutral probabilities implicit in a cross-section of option prices identify only the risk-neutral distribution — you need time series evidence, too, to nail down the true distribution. The Broadie, Chernov, and Johannes papers on equity index options seem like a good place to start, since they estimate both the true and risk-neutral distributions. The question is what they tell us about the probabilities of extreme events, and how these probabilities compare to those in Barro (QJE, 2006) and Rietz (JME, 1987).

One thought (not clear yet if it's a good one) is to compare the entropy of candidate pricing kernels, since that's a useful summary of their impact on expected excess returns. We sketch some thoughts below to show how this might work.

Preliminaries

1. Alvarez-Jermann (Econometrica, 2005) results. One of the appealing features of their paper is that it uses log returns, which are cleaner theoretically. (Lots of things cancel.) And relative to the Hansen-Jagannathan bound, their "entropy bound" depends on higher-order moments (see below). Like everyone else, they start with the pricing relation,

$$E_t\left(m_{t+1}r_{t+1}\right) \;=\; 1,$$

where m is a pricing kernel and r is the (gross) return on any asset we like. Same thing unconditionally.

Result 1 (high-return asset). Consider an economy with a given pricing kernel. The asset with the highest expected return has return equal to its inverse.

Proof. Since the log is a concave function, the pricing relation and Jensen's inequality imply

$$E\log m_{t+1} + E\log r_{t+1} \;\leq\; \log(1) \;=\; 0,$$

with equality if and only if $m_{t+1}r_{t+1} = 1$ (the return is the inverse of the pricing kernel). Therefore,

$$E\log r_{t+1} \;\leq\; -E\log m_{t+1} \qquad (1)$$

*Working notes, no guarantee of accuracy or sense.

as claimed: no asset has higher expected (log) return than one with return equal to the inverse of the pricing kernel. Thus if we have a kernel, we can put a bound on the expected returns it permits. And if we have a return, that gives us a candidate kernel (its inverse) whose properties we can examine.

Result 2 (entropy bound on pricing kernel). Define the entropy of a positive random variable x as

$$J(x) = \log Ex - E \log x.$$

Entropy is non-negative (Jensen's inequality again) and equal to zero only if x is constant. If $\log x \sim \mathrm{N}(\kappa_1, \kappa_2)$, then $J(x) = \kappa_2/2$. More generally, $J(x)$ picks up higher-order cumulants of $\log x$ (if they exist); see below.

Alvarez and Jermann derive an HJ-like bound based on entropy. Let $b_t^1 = E_t m_{t+1}$ be the price of a one-period bond and $r_{t+1}^1 = 1/b_t^1$ its return. Then the mean log excess return provides a lower bound on the entropy of m:

$$J(m) \geq E\left(\log r_{t+1} - \log r_{t+1}^1\right). \tag{2}$$

This is harder to show, so we won't do it here. [Add: high-return asset hits the bound. Actually, it's a little more complicated if interest rates vary, but that won't come up below.]

2. Cumulants. This is a little sloppy, but outlines a way of thinking about sources of entropy — namely, higher-order cumulants. If we have a random variable x, then the moment generating function (if it exists) is

$$g(s) = E \exp(sx)$$

and the cumulant generating function is $h(s) = \log g(s)$. A useful expansion (conditions?) of the cgf is

$$h(s) = \sum_{j=1}^{\infty} \kappa_j s^j / j!, \tag{3}$$

where the κ_js are cumulants.

Note: (i) $g(0) = 1$ and $h(0) = 0$. (ii) $\kappa_j = h^{(j)}(0)$, the jth derivative of $h(s)$ evaluated at $s = 0$. (iii) The entropy of e^x is $h(1) - \kappa_1$ (ie, all the terms but the first one). We're looking for some notation for this, since it comes up a lot. $h_+(1)$? Other ideas? (iv) The cumulants of ax are $a^j \kappa_j$. (v) If x is normal, then e^x is lognormal and the cumulants of x are zero after the first two. Thus departures from normality show up as nonzero values of high-order cumulants. There are, however, no distributions that have a finite number of cumulants greater than two. Gram-Charlier approximations are a related way to add additional cumulants to the normal.

3. Mehra-Prescott environment. Stationary exchange economy with Markov process for the growth rate of consumption (= output): $x_{t+1} = c_{t+1}/c_t$. If preferences are homogeneous of degree one, everything is a function of the growth rate.

An asset is a claim at a specific date t to a dividend stream d_{t+j} for $j \geq 1$. Some examples:

- One-period riskfree bond. Let $d_{t+1} = 1$ (zero for future periods). Its price is $b^1_t = E_t m_{t+1}$ with return $r^1_{t+1} = 1/b^1_t$.

- Consumption strip. Let $d_{t+1} = c_{t+1}$ (zero for future periods). If the ratio of the price to current consumption is q^s_t, the return is $r^s_{t+1} = x_{t+1}/q^s_t$. (Barro calls this equity, but that's not standard usage. In his iid case it's pretty close though.)

- Consumption stream. Let $d_{t+j} = c_{t+j}$ for all $j \geq 1$. If the price-dividend ratio is q^c, the return is

$$r^c_{t+1} \quad = \quad \frac{x_{t+1}(1+q^c_{t+1})}{q^c_t}.$$

- Levered "equity." Let $d_{t+j} = c^\lambda_{t+j}$ for all $j \geq 1$. If the price-dividend ratio is q^e, the return is

$$r^e_{t+1} \quad = \quad \frac{x^\lambda_{t+1}(1+q^e_{t+1})}{q^e_t}.$$

  The previous asset is a special case with $\lambda = 1$. (We'll call this equity for short; it's what Bansal and Yaron (JF, 2004) and many others call it.)

We'll use power utility, so that $m_{t+1} = \beta x^{-\alpha}_{t+1}$. Then the pricing kernel and returns are either loglinear or close to it, and we can trace the impact of the cumulants of $\log x$ on the pricing kernel and asset returns.

Barro's iid economy

Barro considers the impact of low probability disasters — roughly speaking, extreme negative skewness in consumption growth — on the equity premium when consumption growth is iid. This is a particularly convenient starting point, because prices are constant and the only variation in returns comes from the dividend, itself connected to consumption growth. We work through the math to show how this affects the higher-order cumulants of returns and the entropy of the pricing kernel.

Consider two processes for consumption growth. In the first, we specify arbitrary cumulants for $\log x$: κ_1, κ_2, etc. In the second, $\log x$ has two independent random components:

$$\log x_t \quad = \quad x + \sigma v_t + w_t,$$

where $\{v_t\} \sim \text{NID}(0,1)$ and

$$w_t \quad = \quad \begin{cases} 0 \text{ with probability } 1-p \\ b \text{ with probability } p \end{cases}$$

The idea is that b is negative and represents a disaster. Barro lets the "jump" have a more complicated distribution, but we'll skip that for now. Since v and w are independent, the cumulant generating function is the sum of the cgfs of the components, as are the cumulants. This is a standard result for Levy processes. In both cases, we denote the cumulant generating function of $\log x$ by $h(s)$. This gives us extremely convenient expressions for asset returns and premiums; see the Liuren Wu and Ian Martin references at the end.

Here are some ballpark estimates of the first two cumulants for annual data: mean ($\kappa_1 = 0.018$) and variance ($\kappa_2 = 0.035^2$). For the disaster model, we'll use $p = 0.017$ and $b = -0.25$. That implies $\kappa_1 = x + pb$ or $x = \kappa_1 - pb = 0.0222$. Similarly, $\kappa_2 = \sigma^2 + p(1-p)b^2$ or $\sigma^2 = [\kappa_2 - p(1-p)b^2] = 0.0134^2$. There's no particular logic for these parameters, other than to allow us to see concrete examples of how disasters affect asset prices. [Comment: This is a modest disaster/jump by Barro's standards, but if you make it much larger, you drive the variance of the first component below zero.] [Warning: numbers quick and dirty, no guarantees.]

Now some asset pricing:

- Pricing kernel. The pricing kernel is $m = \beta x^{-\alpha}$, which has entropy $J(m) = h(-\alpha) + \alpha\kappa_1$. Its terms are $\alpha^2\kappa_2/2$, $-\alpha^3\kappa_3/3!$, $\alpha^4\kappa_4/4!$, and so on. You can see positive contributions from the variance and kurtosis of consumption growth and a negative contribution of skewness (or, rather, negative skewness of consumption growth leads to higher entropy of the kernel). All of these effects are magnified by α: as α rises, the impact of higher-order cumulants increases much more. In the lognormal case with $\alpha = 4$, $J(m) = \alpha^2 0.035^2/2 = 0.0098$. In the disaster version,

$$\begin{aligned} E\log m &= \log\beta\alpha(x+pb) \;=\; \log\beta - 0.0720 \\ Em &= \beta\exp(-\alpha x + \alpha^2\sigma^2/2)\,[(1-p) + p\exp(-\alpha b)] \;=\; \beta\, 0.9724. \end{aligned}$$

  That gives us $J(m) = 0.0132$, which doesn't leave much room for an equity premium.

- Riskfree rate. The price of a one-period riskfree bond in this economy is $b^1 = Em = E(\beta x^{-\alpha})$. With arbitrary cumulants, the log of the riskfree rate is

$$\log r^1 \;=\; -\log b^1 \;=\; -\log\beta - \sum_{j=1}^{\infty}(-\alpha)^j\kappa_j/j! \;=\; -\log\beta - h(-\alpha).$$

  You can see, for example, that negative skewness and positive kurtosis in consumption growth both lower the riskfree rate. For the lognormal model, we get

$$\log r^1 \;=\; -\log\beta + \alpha\kappa_1 - \alpha^2\kappa_2/2 \;=\; -\log\beta + 0.0622.$$

  (You can see that this is too high; we'd need a discount factor above one to get it down.) For the disaster model, $\log r^1 = -\log\beta + 0.0588$, so it brings the riskfree rate down a little. Neither leaves much room for a risk premium, since the difference with $-E\log m$ is 0.0098 in the lognormal case and 0.0134 in the disaster case. Both premiums hit the entropy bound exactly, which reflects (in part) the constant interest rate.

- Consumption strip. The price is $q^s = E(\beta x^{1-\alpha})$, so $\log q^s = \log \beta + h(1-\alpha)$. The expected return is $E \log r^s = E \log x - \log q^s$, so the expected excess return is

$$E \log r^s - \log r^1 \quad = \quad \kappa_1 - h(1-\alpha) + h(-\alpha) \;=\; \sum_{j=2}^{\infty} [(-\alpha)^j - (1-\alpha)^j] \kappa_j / j!.$$

  In the lognormal case, this is $(2\alpha - 1)\kappa_2/2 = 0.0043$ (43 bps!). You can see that this is not only small, it's less than half the largest premium possible.

  The formula illustrates the impact of high-order cumulants. The coefficient of the skewness term is $3\alpha(1-\alpha) - 1 < 0$, so negative skewness makes the expected excess return higher. The coefficient of the kurtosis term is positive as long as $\alpha > 1$, so that makes a positive contribution as well. We see their impact in the disaster case. The price is

$$\begin{aligned} q^s &= E(\beta x^{1-\alpha}) \\ &= \beta \exp[(1-\alpha)x + (1-\alpha)^2\sigma^2/2)]\,(1 - p + p \exp[(1-\alpha)b]) \;=\; 0.9540. \end{aligned}$$

  and the (log) return is

$$E \log r^s \quad = \quad E \log x - \log q^s \;=\; \kappa_1 - \log q^s \;=\; -\log \beta + 0.0651.$$

  That gives us a premium of 0.0064, larger than the lognormal model, but still small — and, again, well below the bound.

- Levered equity. In the iid case, the price is constant and satisfies

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

  You can see immediately that high-return asset has $\lambda = \alpha$ (Alvarez-Jermann Result 1). The "price ratio" in general is

$$\log\left[q^e/(1+q^e)\right] \quad = \quad \log \beta + h(\lambda - \alpha),$$

  so the expected return is

$$\begin{aligned} E \log r^e &= \lambda E \log x - \log\left[q^e/(1+q^e)\right] \\ &= \lambda \kappa_1 - \log \beta - h(\lambda - \alpha). \end{aligned}$$

  The expected excess return is

$$E \log r^e - \log r^1 \quad = \quad (\lambda - \alpha)\kappa_1 - h(\lambda - \alpha) + h(-\alpha).$$

  Holding the pricing kernel fixed (hence α), the first two terms are minus the entropy of $x^{\lambda-\alpha}$, which hits a max of zero at $\lambda = \alpha$. The excess return in this case is $h(-\alpha)$ (need to kill off the initial term??).

- Risk-neutral probabilities ... Work these out, show how options work.

Related papers

Stan Zin's paper: "Are behavioral asset-pricing models structural?" JME, 2002. The highlight is the section heading, "Never a dull moment."

Liuren Wu's papers on cumulants in jump models. One example is "Dampened power law," J Bus, 2006. He also recommends Polimenis, "Skewness Correction for Asset Pricing," SSRN, 2006.

Ian Martin, a Harvard PhD student who is going to Stanford, refers to two related papers on his website, one on cumulants, the other on using option prices to infer disaster probabilities. We might want to track them down. See

`http://www.people.fas.harvard.edu/~iwmartin/papers.html`

Random thoughts on what to do next

1. One idea is to come up with simple summaries of models that we can compare. One possibility is the entropy of the pricing kernel. Even better, graph the components of entropy: $\kappa_2, 2$, $\kappa_3/3!$, etc. Then we could compare (say) the kernel in Barro's paper with that in the estimated model of Broadie, Chernov, and Johannes. We'll have to think about whether that makes sense. Our guess is that if a model has too much entropy, some assets will be priced incorrectly. We could develop this further, perhaps looking at options. Or we could think about representative agent models that approximate the estimated model. In short, we'll have to think about the point of this project.

2. We could also look at equity returns. Compute mean gross return, mean log return, entropy of return, standard deviation of log return, "excess entropy" due to higher cumulants. Set $\log m_t = -\log r^e_{t+1}$ (it prices equity right) and compute entropy and excess entropy. This should give us some idea what equity returns imply, although it has the same weakness as macro data: extreme events don't happen very often. Still, it's interesting to contrast equity returns with consumption growth.

3. Look at options. In Barro's example, we'd guess that the highest-return asset has a significantly higher return that his equity strips. In that sense, his model is too much of a good thing. Where does that show up? One guess is out of the money calls, which are more sensitive to the underlying than the underlying itself. Does his model overvalue them? Should we look at equity options more generally?

4. Does Barro report the distribution of contractions he uses in his paper? Might be there, but we haven't found it.

Possible outline

1. Intro. The challenge of the disaster line of work is that they're infrequent, so it's hard to estimate their distribution reliably. We use options, which price events that

need not occur. The challenge here is to distinguish between true and risk-neutral probabilities...

Punchline?

2. Barro's disasters. Show how this works.

3.

Appendix: Old notes on true and risk-neutral distributions

Notes from June 2007. Show how we connect pricing kernel representation to so-called P and Q measures. This started with a comment by Liuren Wu.

Moments of risk-neutral distributions

Risk-neutral probabilities. Let everything be a function of the state z: the pricing kernel is $m(z)$, the true probability $p(z)$, and the return is $r(z)$. The fundamental pricing relation is

$$\sum_z m(z)p(z)r(z) \;=\; E\,(mr) \;=\; 1.$$

If we define the risk-neutral probabilities

$$p^*(z) \;=\; m(z)p(z)/E(m),$$

the pricing relation becomes

$$E(m)\sum_z p^*(z)r(z) \;=\; E(m)E^*(r).$$

In general, p^* can be anything: we can use m to transform p into any distribution we want.

Cumulant generating functions. Consider a random variable $q(z)$. Its cgf is

$$k(s) \;=\; \log\left(\sum_z p(z)\exp[sq(z)]\right) \;=\; \log E\exp(sq).$$

Cumulants follow from derivatives of k evaluated at $s = 0$. Similarly, the cgf using the risk-neutral probabilities is

$$k^*(s) \;=\; \log\left(\sum_z p^*(z)\exp[sq(z)]\right) \;=\; \log E^*\exp(sq).$$

Again, there's no necessary connection between k and k^*.

Representative agent with power utility. Let $x(z)$ be the log of consumption growth. Then the pricing kernel (=mrs) is

$$m(z) \;=\; \beta\exp[-\alpha x(z)].$$

Now let's compare true and risk-neutral probabilities. The true probabilities are whatever we want; let's say they imply the cgf for x: $k(s) = \log E[\exp(sx)]$. The risk-neutral probabilities are

$$p^*(z) \;=\; m(z)p(z)/E(m) \;=\; \frac{\beta\exp[-\alpha x(z)]p(z)}{\beta\exp[-k(-\alpha)]} \;=\; \exp[-\alpha x(z) - k(-\alpha)]p(z).$$

The cgf for the risk-neutral probabilities is therefore

$$k^*(s) \;=\; \log E^*\exp(sx) \;=\; \log E\exp[(s-\alpha)x - k(s)] \;=\; k(s-\alpha) - k(-\alpha).$$

Only the first term matters for the cumulants; ie, $k(s-\alpha)$.

This translates into properties of returns if the latter are related to x. Typically, the return on the "market" is $\exp(x)$, or something like that, so properties of x translate into properties of the (log of the) market return.

Liuren's notes

Your example,

$$\begin{aligned}\ln m_{t+1} &= \beta - \alpha\varepsilon_{t+1} \\ \ln R_{t+1} &= r_{t+1} = \mu + \varepsilon_{t+1}\end{aligned}$$

I agree with you that we can use $\exp(-r_f) = E_t[m_{t+1}]$ to determine the short rate and use $1 = E_t[m_{t+1}R_{t+1}]$ to determine the risk premium (mean). So let's say we did these exercise and know what β and μ is already (so we can ignore them now). For simplicity, I assume zero rates, then $1 = E_t(m_{t+1})$ and $\beta = -k(-\alpha)$ where $k(c)$ is the cumulant exponent of ε_{t+1} so that

$$E_t[m_{t+1}] = \exp(\beta)\exp(k(-\alpha)) = 1.$$

The cumulant exponent is defined by

$$k(s) = \ln E_t[\exp(s\varepsilon)].$$

In math jargon, I call $\exp(-\alpha\varepsilon - k(-\alpha))$ an exponential martingale. It defines the measure change as follows.

Consider a generic time-$(t+1)$ state-coningent payoff function $\pi(\varepsilon_{t+1})$. I write it as a function of ε_{t+1} because ε_{t+1} is the only risk we have here. You can think of an option or something for the payoff. Then, we can write its time-t value as

$$V_t = E_t^P[m_{t+1}\pi(\varepsilon_{t+1})].$$

where the expectation is under P (statistical measure). Alternatively, we assume there is a Q probabiliyty measure such that

$$V_t = E_t^Q[\pi(\varepsilon_{t+1})].$$

Compare the two equations, we have

$$\int \pi(\varepsilon)\exp(-\alpha\varepsilon - k(-\alpha)) f^P(\varepsilon)\,d\varepsilon = \int \pi(\varepsilon) f^Q(\varepsilon)\,d\varepsilon,$$

where f^P and f^Q are the probability density functions. You can see that the two density functions are linked by,

$$f^Q = \exp(-\alpha\varepsilon - k(-\alpha)) f^P(\varepsilon).$$

First, f^Q is obviously still a density function: It is positive and integrates to 1 because $E_t[\exp(-\alpha\varepsilon - k(-\alpha))] = 1$.

Second, with positive α, the exponential function $\exp(-\alpha\varepsilon)$ intuitively make the left tail thicker and the right tail thinner and hence makes the distribution more negatively skewed, except under the very special normal distribution case. In that case, it only changes the mean, not the tail. For example, if $f^P(\varepsilon)$ is symmetric $\alpha-$stable distribution, then the exponential function would make the distribution "dampened" alpha stable under Q and negatively skewed.

More generally, let's assume the following cumulant exponent under P:

$$k(s) = \frac{1}{2}s^2 + \frac{1}{6}\gamma_1 s^3 + \frac{1}{24}\gamma_2 s^4$$

so that ε has zero mean, unit variance, skewness γ_1 and kurtosis γ_2 under P–You are very familar with this expansion stuff. I assume no higher cumulants for simplicity. Now let's try to solve its cumulant function under Q

$$\begin{aligned} k_Q(s) &= \ln E_t^Q[\exp(s\varepsilon)] \\ &= \ln E_t^P[\exp(-\alpha\varepsilon - k(-\alpha)) + s\varepsilon] \\ &= k(s-\alpha) - k(-\alpha) \\ &= \frac{1}{2}(s-\alpha)^2 + \frac{1}{6}\gamma_1(s-\alpha)^3 + \frac{1}{24}\gamma_2(s-\alpha)^4 - \frac{1}{2}(\alpha)^2 + \frac{1}{6}\gamma_1(-\alpha)^3 + \frac{1}{24}\gamma_2(-\alpha)^4 \end{aligned}$$

Taking derivatives, we get all the cumulants under Q. The mean is

$$\begin{aligned} c_1^Q &= k_Q'(s)\Big|_{s=0} \\ &= (s-\alpha) + \frac{1}{2}\gamma_1(s-\alpha)^2 + \frac{1}{6}\gamma_2(s-\alpha)^3 \\ &= -\alpha + \frac{1}{2}\gamma_1\alpha^2 - \frac{1}{6}\gamma_2\alpha^3. \end{aligned}$$

The variance is

$$\begin{aligned} c_2^Q &= k_Q''(s)\Big|_{s=0} \\ &= 1 + \gamma_1(s-\alpha) + \frac{1}{2}\gamma_2(s-\alpha)^2 \\ &= 1 - \gamma_1\alpha + \frac{1}{2}\gamma_2\alpha^2 \end{aligned}$$

Note that the varance increases under the risk-neutral measure if the statistical distribution is negatively skewed. This partly explains the "variance risk premia": Risk-neutral variance is higher than historical variance.

Now let's go to the skewness part:

$$\begin{aligned} c_3^Q &= k_Q'''(s)\Big|_{s=0} \\ &= \gamma_1 + \gamma_2(s-\alpha) \\ &= \gamma_1 - \gamma_2\alpha, \end{aligned}$$

which says that **even if the statistical distribution is symmetric** $(\gamma_1 = 0)$**, the risk-neutral distribution is negatively skewed as long as (1) risk aversion is positive** $(\alpha > 0)$**and (2) the statistical distribution has fat tail** $(\gamma_2 > 0)$**.**

In my example, the fourth cumulant does not change because $c_5 = 0$, but ...

See various papers on Liuren Wu's home page:

http://faculty.baruch.cuny.edu/lwu/