Review for Quiz #2

(Started: March 26, 2012; Revised: April 4, 2012)

I'll focus again on the big picture to give you a sense of what we've done and how it fits together.

Asset pricing: summary

This is fundamental, worth repeating.

The no-arbitrage theorem tells us that we can price assets using state prices: break an asset's payoff into state-contingent pieces, multiply each one by its state price, and add them up. That turns states and state prices from hopelessly abstract objects into objects with clear practical value.

The "no-abritrage" theorem says: If asset j has dividend $d^j(z)$ in state z, its price satisfies

$$q^j \;=\; \sum_z q(z) d^j(z). \tag{1}$$

The theorem says we can always find positive state prices $q(z)$ that satisfy this equation for every traded asset j.

There are two versions of this wonderful result that are more commonly used. Both involve modest redefinitions of state prices. The first version is based on a *pricing kernel* m, defined implicitly by $q(z) = p(z)m(z)$. (Solve for m if you like.) The pricing relation (1) becomes

$$q^j \;=\; \sum_z p(z)m(z)d^j(z) \;=\; E(md^j). \tag{2}$$

We've seen that risk premiums stem from covariances of returns with m. It also has a nice link to representative agent models, where m is the agent's marginal rate of substitution.

The second version is based on *risk-neutral probabilities* p^*, defined implicitly by $q(z) = p(z)m(z) = q^1 p^*(z)$. (Solve for p^* if you like.) Here

$$q^1 \;=\; \sum_z p(z)m(z) \;=\; E(m)$$

is the price of a one-period riskfree bond. The pricing relation (1) now turns into

$$q^j \;=\; q^1 \sum_z p^*(z) d^j(z) \;=\; q^1 E^*(d^j), \tag{3}$$

where E^* means the expectation computed from the risk-neutral probabilities. This version is useful in pricing derivatives on an asset, because the risk-neutral distribution is the same for all of them. Once we know p^*, it's a little simpler because we only need the expectation, not the expectation of a product. The reason, of course, is that the product is built into p^*.

Pricing kernels

Here are three approaches to pricing kernels that we've used.

Approach 1: representative agent. Take a representative agent with power utility, see how it values things.

The classic example here is the equity premium. If $g(z)$ is the growth rate of consumption, the pricing kernel is $m(z) = \beta g(z)^{-\alpha}$. Now define equity as a claim to a dividend tied to the same growth rate, such as $d(z) = g(z)^{\lambda}$. If we choose a distribution for g and "reasonable" values for parameters, the question is whether a model of this sort can generate an equity premium similar to what we've observed in the data. The answer is, basically, no. We saw that for returns measured in levels, returns in logs, a two-state distribution, a lognormal distribution, and some others. Unless we have a very large value of α, we don't have a chance. One mechanism that made modest progress was to have strong negative skewness in (log) consumption growth, which tends to increase risk premiums.

Approach 2: HJ bound. The question is where we went wrong. The Hansen-Jagannathan bound suggests that a power utility pricing kernel doesn't have much of a chance to start with. We start with the *Sharpe ratio*, the ratio of the mean of an excess return to its standard deviation. If we look returns on lots of assets, the idea is to look for the asset with the largest Sharpe ratio. Hansen and Jagannathan show that this places a lower bound on the ratio of the standard deviation of the pricing kernel to its mean:

$$\frac{|E(x)|}{Std(x)} \quad \leq \quad \frac{Std(m)}{E(m)}.$$

Roughly speaking, large Sharpe ratios imply large standard deviations of the pricing kernel.

One way to think about the representative agent model's failure with the equity premium is that it doesn't deliver a large enough standard deviation of the pricing kernel.

Approach 3: entropy bound. Similar idea, somewhat different implementation. We define the entropy of the pricing kernel by

$$H(m) \quad = \quad \log E(m) - E \log m.$$

This is a measure of dispersion: it's nonnegative, and strictly positive unless m is constant. The bound here tells us that

$$E \log r^j - \log r^1 \quad \leq \quad H(m).$$

The difficulty here with the representative agent model is that it doesn't generate enough entropy in the pricing kernel.

One of the nice things about entropy as a measure is that it incorporates things like skewness and excess kurtosis naturally. We can think of the first term in entropy as the cgf of $\log m$ evaluated at $s = 1$:

$$\log E(m) \quad = \quad \log E\left(e^{\log m}\right) \quad = \quad \kappa_1 + \kappa_2/2 + \kappa_3/3! + \kappa_4/4! + \cdots.$$

When we subtract the mean, we get entropy:

$$H(m) \;=\; \log E(m) - E \log m \;=\; \kappa_2/2 + \kappa_3/3! + \kappa_4/4! + \cdots .$$

In the lognormal case, only κ_2 is nonzero. Otherwise, the other cumulants play a role.

Option pricing

As before, we have several components that combine to give us a complete picture. The notation below isn't self-contained, you may need to go back to the notes.

Options. Options are the right to buy or sell an asset (the "underlying") at a fixed price k (the "strike") at (or by) some future date. We'll label the dates t ("now") and $t+1$ (or $t+\tau$) ("later"). The question is what that right is worth now. We're going to value it using risk-neutral probabilities, since we know they can price anything.

Put-call parity. For European options, there's a connection between prices of put options (the right to sell) and call options (the right to buy):

$$\underbrace{q_t^c}_{\text{buy call}} \;-\; \underbrace{q_t^p}_{\text{sell put}} \;+\; \underbrace{q_t^\tau k}_{\text{present value of strike}} \;=\; \underbrace{s_t}_{\text{buy stock}} .$$

This holds for every strike k. It works pretty well in practice, as we saw in Lab Report #5.

BSM formula and implied volatility. The standard textbook formula for a call option is

$$\begin{aligned} q_t^c &= s_t N(d) - q_t^\tau b N(d - \tau^{1/2}\sigma) \\ d &= \frac{\log(s_t/q_t^\tau k) + \tau\sigma^2/2}{\tau^{1/2}\sigma}. \end{aligned}$$

Everything is observable here except "volatility" σ, which we can back out from the price. That is: given σ, we use the formula to compute the price. But if we know the price, we can reverse the process and compute implied volatility: the value of σ for which the formula delivers the observed price.

One of the reliable facts about option prices is that volatility varies with the strike. The shape of the line in a graph of volatility against the strike price is referred to as the *volatility smile*. We'll think of it as a convenient way to summarize option prices.

Risk-neutral pricing. We've seen that any cash flow can be valued using (3). Options are no different. The question is generally what the risk-neutral distribution is.

Let's be specific. The "dividend" for an option is the future price of the underlying s_{t+1}. The risk-neutral pricing relation (3) then gives us

$$s_t \;=\; q_t^1 E^*(s_{t+1}).$$

Given a risk-neutral distribution, this gives us s_t. Usually we observe s_t, and this gives rise to a restriction on the risk-neutral distribution that we call the *no-arbitrage condition*.

Here's an example. Suppose $y_{t+1} = \log s_{t+1} \sim \mathcal{N}(\kappa_1, \kappa_2)$. Then the pricing relation implies

$$s_t = q_t^1 e^{\kappa_1 + \kappa_2/2}.$$

Usually we choose κ_1 to satisfy the equation given information about the other components.

BSM from lognormal risk-neutral underlying. With these ingredients, we can derive the BSM formula using a lognormal risk-neutral distribution of the underlying. A put price is at strike k is

$$q_t^p = q_t^1 E^*(k - e^{y_{t+1}})^+ = q_t^1 \int_{-\infty}^{\log k} (k - e^{y_{t+1}})(2\pi\kappa_2)^{-1/2} \exp[-(y_{t+1} - \kappa_1)^2/2\kappa_2] dy_{t+1}.$$

Integrating gives us

$$\begin{aligned} q_t^p &= q_t^1 k N(d) - q_t^1 e^{\kappa_1 + \kappa_2/2} N(d - \kappa_2^{1/2}) \\ d &= (\log k - \kappa_1)/\kappa_2^{1/2}. \end{aligned}$$

We get the BSM formula for $\tau = 1$ by using the no-arbitrage condition.

Beyond BSM. It's not hard to get option formulas that differ from BSM, we simply start with a risk-neutral distribution of the underlying that's not lognormal. There are lots of examples. We used "normal mixtures," which have some of the analytical convenience of normality but provide more flexibility.