University of Illinois, Urbana-Champaign *Professor Tim Johnson*
FIN 512 - Spring 2015 *tcj@illinois.edu*

Lecture Note 7.1: Multiple Sources of Risk

Introduction:

Today we begin a new phase of our course in which we move beyond option pricing problems to investigate more complex applications. In particular, we want to be able to tackle diverse products and markets that we encounter in financial engineering in the real world

This note generalizes the lessons we learned in developing models with stochastic volatilty. The techniques we used there can be used to value derivatives with any number of underlying sources of randomness.

Once we have articulated the problems, we also have to learn how to solve them. We will talk about some issues of implementation. Most crucially, we will discuss how the idea of risk-neutral pricing extends to the multidimensional case. This will give us an extremely powerful tool, that we can begin to use right away.

Outline:

I. Multidimensional Topics
II. Implementability Issues
III. Probabilistic Solutions to General Models
IV. Applications
V. Beyond Monte Carlo
VI. Summary

I. Some new classes of problems.

- Last week we put a lot of effort into figuring out how to cope with a second source of risk in derivatives pricing. Now let's step back and recognize that how useful that was.

  There are many <u>other</u> situations in which we will want to analyze multiple sources of risk.

- Here are a few.

(A) Assessing the other Black-Scholes assumptions.

- The motivation for our investigation last week was to address the sensitivity of the B-S formula to the constant volatility assumption – within a consistent framework.

- We could now proceed the same way to examine the remaining constant parameter assumptions:

  (i) constant interest rates, and

  (ii) constant yields on S.

- Like last week, our method will be to specify new stochastic processes for the things we previously assumed couldn't change.

- If we make the risk-free rate random, the situation is this:

$$\begin{aligned} dS &= \mu(\cdot)S\,dt + \sigma(\cdot)S\,dW^S \\ dr &= m(\cdot)\,dt + b(\cdot)\,dW^r. \end{aligned}$$

- Now we're forgetting about dW^σ. Volatility doesn't have its own source of randomness. (One thing at a time!) In fact, there aren't any more sources of uncertainty besides S and r.

- Ok, we have our option, C, and it varies with both S and r. So just as we did before, we use the two-dimensional Itô rule to write down how it changes. Then we get rid of the two types of risk, just like we did with stochastic volatility. Then ....

- Hey, the whole argument is the exact same!

- So we already know what the PDE will turn out to be:

$$\frac{1}{2}b^2 C_{rr} + \rho b \sigma S C_{rS} + \frac{1}{2}\sigma^2 S^2 C_{SS} + rSC_S - rC + C_t + [m - \lambda b]C_r = 0.$$

- This is the exact same equation we got last week but with r substituted for σ in the partial derivatives.

  - ▶ And now ρ is the correlation between dW^S and dW^r;

  - ▶ and $\lambda = \lambda^r$ is the **market price of interest-rate risk**, which has nothing to do with the price of volatility risk;

  - ▶ and in the old equation r was constant in the 4th and 5th terms.

- The point is, when we derived the stochastic volatility PDE, our argument never really made reference to what σ was.

- So now we know how to handle *any* type of random parameter.

- With *random payout*, y_t, the derivation we used for stochastic σ or r works exactly the same

- As a result, we find basically the same equation:

$$\frac{1}{2}b^2 C_{yy} + \rho b \sigma S C_{yS} + \frac{1}{2}\sigma^2 S^2 C_{SS} + (r-y)SC_S - rC + C_t + [m-\lambda b]C_y = 0$$

  where now

  - ρ is the correlation between dW^S and dW^y;
  - $\lambda = \lambda^y$ is the market price of payout risk;
  - and m and b are the instantaneous drift and diffusion of the payout process.

- Recall how the whole thing worked:

  - The key was there had to be a second security out there that we could use to (continuously) hedge the new risk.
  - And we have to know the new risk's market price, λ.

(B) Derivatives on Non-traded Underlyings

- Our analysis so far has been concerned with the effect of random parameters in valuing derivatives on some asset S.

- But our methods point us to a further generalization:

  ▶ We could price derivatives whose payoff also depends on the random parameter.

  ▶ In fact, they might depend only on it.

  ▶ Our no-arbitrage arguments still works even if there is no traded underlying security – no S – at all.

- For example, interest rate or volatility derivatives.

  ▶ If r is the ONLY random thing affecting the payoffs of some claim, then our PDE is even simpler:

$$\frac{1}{2}b^2 C_{rr} - rC + C_t + [m - \lambda b]C_r = 0.$$

  ▶ All the S terms are gone.

- And we can say more:

  Our approach can value derivatives on any quantity.

- We just happened to discover our PDE by bringing new sources of risk into the problem of valuing an option (on some particular financial asset).
- Now just think of the source of risk itself as anything. It doesn't even have to have any connection to finance. It could be
    - ▶ Weather derivatives.
    - ▶ Telecom/Electricity/Energy derivatives.
    - ▶ Catastrophe derivatives.
    - ▶ A firm's accounting results (like sales).
    - ▶ Whatever you can imagine...

(C) Higher Dimensional Models

- Another direction in which we can exploit the work we have done is to add still more types of risk.
- We have shown how it works for $N = 2$ variables, but nothing stops us from building general $N-$dimensional models with $N = 3, 4, \ldots$
- For example, we could handle:
  - ▶ Currency options with stochastic volatility and domestic and foreign interest rates ($N = 4$).
  - ▶ Multifactor term-structure models, e.g. bond prices that depend on both r and inflation.
  - ▶ Options on portfolios with N correlated assets.
- Solving the PDE takes more computer power. But it can be done.
- One goup of researchers has solved a pricing problem in 360 dimensions!
  - ▶ The product was a collaterallized mortgage obligation (CMO).
  - ▶ This is a security that pools together cash-flows from many underlying loans and re-sells them in packages.
  - ▶ They wanted to let the re-payment rates on each of the underlying mortgages be its own random process.

- We can summarize what happens to the PDE when we add a new factor, X, as follows:

  1. Add a second derivative term $\frac{1}{2}X^2 C_{XX}$ times the squared volatility of X.

  2. For each other factor, Y, add a cross-derivative term XYC_{XY} times the covariance between X and Y.

  3. Add a first derivative term C_X, whose coefficient is

    (i) X times the risk-free rate minus the payout rate **if** X **is a traded factor**;

    (ii) the drift of X minus the product of the market price of X risk with the diffusion coefficient of X, **if** X **is a non-traded factor**;

- The derivation of this rule just follows the lines of the argument we went through for the stochastic volatility case.

- It is starting to look like we can handle almost unlimited complexity.

II. Implementability Issues

- Before getting carried away, let's just keep track of the practicalities of employing an N-dimensional model on, perhaps, arbitrary quantities.

  1. You have to write down a stochastic model for each random quantity that you believe.

  2. You have to have estimates of the unknown parameters of that model that you trust.

  3. There has to be at least one other security in the market that you can buy and sell to hedge the risk of each new quantity.

  4. The risk of non-hedgeable jumps in the new quantity must be negligible.

  5. You have to have some way of finding the market price of risk for each non-traded quantity.

- Let's think about each of these for a moment.

- Issues 1 and 2 are really about econometrics.

  ▶ For the models to be <u>improvements</u> over simpler models, we have to be confident that:

    * Our specification of the model is a better description of the true underlying process that we are trying to describe – as it will evolve in the future; and

  * We can estimate (and forecast) the new parameters required with a high degree of confidence; and

  * Our uncertainty about them has smaller effects on our ultimate conclusions than a simpler model with fewer parameters.

- This will not always be the case. In fact there are dangers from two directions.

  i. We can't estimate the new model's parameters with any accuracy at all.

  ii. We estimate them "too well" and end up overfitting.

- Overfitting means building a model to capture detailed patterns that occurred in the past, but which were really just chance (or noise).

- When we have N sources of risk, and N models for them, we could have around $N^2/2$ new parameters to estimate.

  * For each new factor we include, we have to specify its correlation with each old one. Then we have to estimate means, variances, and covariances.

  * Depending on the model specification, the degree of estimation error may grow large for higher dimensional models because the data may not be strongly informative about some parameters.

- The task of *model validation* is an important part of risk management in any financial engineering exercise.
  - Firms that use proprietary models typically impose formal procedures for assessing both *estimation risk* and *specification risk*.
  - Even "riskless" arbitrage positions are exposed to these type of risks.

- Issues 3 and 4 on my list above concern the realism of the continuous trading/hedging assumption.
  - Economists call this the *market completeness* condition.
  - It is <u>un</u>realistic if
    - There is not a way to create a pure hedge for each source of risk.
      - For example, an individual homeowner's mortgate repayment decision.
    - The hedge cannot be undertaken from both sides: long and short.
    - Market liquidity is not sufficient to allow frequent re-hedging (without affecting prices).

- ▶ As an aside, jump risk can be incorporated into our PDE framework – provided you are willing to maintain the completeness assumption.
  - ∗ Again, there has to be a tradeable security that allows you to completely hedge against the risk of jumps.
- ▶ And, in that case, you also need to be able to find market prices of jump risk.

- This brings us to issue 5. Recall our earlier discussion of λs.
  - ▶ Once our model involves them, we are no longer in the (nice) position of being able to ignore economics for valuing derivatives.
  - ▶ We need to know something about people's risk/reward preferences for each type of risk.
  - ▶ There is no reason why people's preferences can't change over time!
    - ∗ Our PDE allows us to incorporate <u>changing</u> market prices of risk, by making them functions of other state variables.
    - ∗ Sometimes economic theories can provide guidance.
    - ∗ For example, some macroeconomic models suggest that we could have $\lambda^r = a + b\sqrt{r}$.

- ∗ **Q:** What would the CAPM say about the market price of risk associated with:

  (A) 30-year U.S. Treasury bonds?

  (B) The Korean Won?

  (C) The average Summer temperature in Champaign?

- ▶ As with other unknown parameters, we are actually interested in the values of the λs in the future, not the past.

- To summarize the implementability issues, multidimensional models are exciting, but not always feasible in practice.

- Hence, *more sophisticated models require more sophisticated scrutiny.*

- More complicated isn't necessarily better.

III. Probabalistic Solutions to General Models

- I now want to show you one solution technique for the general multi-dimensional PDE we have developed.

- The idea is to extend to the general case the correspondence between solutions and discounted expectations using suitably adjusted probabilities.
  - ▶ Often, it is possible to find these expectations directly.
  - ▶ Even when it's not, the intuition is important.

- How do we go about "risk-neutralizing" these more complex models?

- To review, when we considered the Black Scholes PDE, we learned that one way to find the solution was *to average all possible terminal payoffs (discounted at the riskless rate), after having adjusted the drift on the underlying process.*

- Specifically, we set the expected percentage change to the risk-free rate (or that rate minus the payout rate).
  - ▶ We interpreted that adjustment as switching to an alternative (false) model in which the stock price was determined by risk-neutrality. Hence we called the pricing approach "taking the discounted risk-neutral expected payoff."

- But recall the subtle underpinnings of the logic that allowed us to reach this conclusion.

- Then, the justification for this result was that the equation we had to solve made no reference to people's preferences for risk, so that risk neutral people (who value all cashflows as their expected present value) must reach the same solution as anybody else. So we could pretend we were risk-neutral.

- Now, the more general equation we have to solve *does* involve people's preferences. That's what that λ was.
  - ▶ So it's not clear whether we should expect to get away with a similar trick.

- Moreover, intuitively, it would seem very strange to set the drift of some non-traded process (for example, the daily temperature) to the riskless rate, and then use that process to compute expectations.

- Well, luckily for us, there is a theorem in mathematics that tells us precisely what to do.

- I'll describe it in terms of the stochastic volatility model we saw last time. Recall the two processes we specified were

$$\begin{aligned} dS &= \mu S\,dt + \sigma S\,dW^S \\ d\sigma &= \kappa(\sigma_0 - \sigma)\,dt + s_0\sigma\,dW^\sigma. \end{aligned}$$

and the equation we want to solve is:

$$\frac{1}{2}s_0^2\sigma^2 C_{\sigma\sigma} + \rho s_0\sigma^2 S C_{\sigma S} + \frac{1}{2}\sigma^2 S^2 C_{SS} + rSC_S - rC + C_t + [\kappa(\sigma_0 - \sigma) - \lambda s_0\sigma]C_\sigma = 0.$$

- The result we need is a version of what is called the **Feynman-Kač Theorem.** It says this:

  The value of the derivative is the expectation of its discounted payoffs (at rate r), after each of the risk processes has had its drift changed to **whatever multiplies the first derivative term with respect to it in the PDE.**

- In other words, here

(A) Set the drift of the stock to $r \cdot S$, as before (which is the same as setting its expected return to r).

(B) Set the drift of σ to $m() - \lambda b()$.

(C) Don't change anything else. Most notably, this means the volatilities and correlations don't get changed.

- Then all you have to do is figure out the average payoff in present value terms.

- The theorem tells us, once we have shown that our derivative satisfies the PDE, **we can conclude that that adjusted average payoff is the solution.**

- Before looking at examples, I want to call your attention to one subtle point about the recipe. *If the interest rate itself is one of the random variables,* then there is a difference between *the expected payoff, discounted* and *the expected discounted payoff.* The theorem tells us we are supposed to use the latter:
  - ▶ The discount factor goes inside the expectation.

- To put it in terms of equations, let X stand for the vector of all our state variables.

  Then, if the derivative has terminal payoff $C(X_T)$, and $\mathrm{E}^\star$ denotes averaging over future outcomes with the adjusted drifts, then the result says its price at t is

$$\mathrm{E}^\star[DF_{t,T} \cdot C(X_T)]$$

  where

$$DF_{t,T} \equiv e^{-\int_t^T r_u \, du}$$

  is the integrated discount factor.

- Putting it inside the expectation takes into account the possibility that the evolution of X may be correlated with that of r.

- If there are cash-flows $\Gamma_s(X_s)$ at times $t < s < T$, then the formula is

$$\mathrm{E}^\star[\int_t^T DF_{t,s} \cdot \Gamma_s \, ds + DF_{t,T} \cdot C(X_T)]$$

  - ▶ Recall that these cash-flows (such as interest payments) just tack on the additional term Γ to the PDE.

- If we were simulating the process over intervals Δt then the discount factor would be the same as

$$\left(e^{-r_{t_0}\Delta t} \times e^{-r_{t_1}\Delta t} \times e^{-r_{t_2}\Delta t} \times \ldots e^{-r_{t_N}\Delta t}\right)$$

which is approximately equal to

$$\left(\frac{1}{(1+r_{t_0}\Delta t)} \times \frac{1}{(1+r_{t_1}\Delta t)} \times \frac{1}{(1+r_{t_2}\Delta t)} \times \ldots \frac{1}{(1+r_{t_N}\Delta t)}\right)$$

i.e the product of all the little one-period discount factors that will apply between now and T.

- Or, in English, **when interest rates are random, the discount factors between t and T is the product of all the future one-period discount factors multiplied together** – after adjusting the drift of r.

- Of course, if we are using a model with constant r, that is just $e^{-r(T-t)}$.

- Also notice that, if rates are random – but they are independent of the state variables X – then we can observe that

$$\begin{aligned} \mathrm{E}^\star[DF_{t,T} \cdot C(X_T)] &= \mathrm{E}^\star[DF_{t,T}] \cdot \mathrm{E}^\star[C(X_T)] \\ &= B_{t,T} \cdot \mathrm{E}^\star[C(X_T).] \end{aligned}$$

So we don't have to model the bond prices if they don't interact with the variables that determine our cashflows.

• Let's continue with the stochastic volatility example, so that you get the idea of how to use the Feynman-Kač result.

- Our general formula (for a call) requires us to compute

$$B_{t,T} \cdot \mathrm{E}^{\star}[\max(S_T - K, 0)]$$

But it is not clear how to compute that expectation mathematically.

- So just simulate! Here's how it goes.

(A) Adjust the stock and volatility processes according to the rule.

$$\begin{aligned} dS_t &= rS_t\,dt \;+\; \sigma_t S_t\,dW^S \\ d\sigma_t &= [\kappa(\sigma_0 - \sigma_t) - \lambda s_0 \sigma_t]\,dt \;+\; s_0 \sigma_t\,dW^{\sigma}. \end{aligned}$$

(B) Approximate the continuous time model by a discrete time version, i.e. replace dt with Δt and dW^{σ} and dW^S with two normal random variables with mean zero, variance Δt and correlation ρ.

(C) Draw one realization of the random sequences.

(D) Compute the terminal payoff, $\max(S_T - K, 0)$, for this one outcome.

(E) Multiply by the discount factor.

(F) Average the result over a whole lot of draws.

- This is the **Monte Carlo** method of evaluating expectations.

IV. Applications.

- Sometimes we don't need to simulate. For some stochastic processes we can compute the adjusted expectations directly.

(A) Forwards.

- Go back to the simple, one-variable model. One fact you need to know is that, if we have any process described by the model

$$\frac{dX}{X} = c_0\,dt + c_1\,dW$$

where c_0 and c_1 are constant, then, if we are at time t, for any horizon T, the expected value of X_T is

$$\mathrm{E}[X_T] = X_t e^{c_0(T-t)} \quad \text{or} \quad \frac{1}{(T-t)} \log \frac{\mathrm{E}[X_T]}{X_t} = c_0.$$

▶ **Example:** A stock pays no dividends and has a constant 9% expected return and constant volatility. Today it is at 100. What is its expected price in 10 years?

∗ The model here is

$$\frac{dS}{S} = 0.09\,dt + \sigma\,dW$$

∗ So the formula says

$$\mathrm{E}[S_T] = 100 \cdot e^{0.09 \cdot 10} = 246$$

- We can use this recipe for expectations to immediately derive some derivatives prices.

- Suppose we want to price an outstanding forward contract with forward price F_0 on a non-dividend paying stock that follows that same type of model.

  ▶ Assume r is constant.

  ▶ Then using our recipe, we first change the expected return to r, and then compute the expected discounted payoff

  $$V_F = \mathrm{E}^{\star}[e^{-r(T-t)}(S_T - F_0)]$$
  $$= e^{-r(T-t)}(\mathrm{E}^{\star}[S_T] - F_0)$$

  where the $\mathrm{E}^{\star}$ notation is to remind us that this isn't our true expectation.

  ▶ Now use the result we just learned to evaluate $\mathrm{E}^{\star}$:

  $$V_F = e^{-r(T-t)}(S_t e^{r(T-t)} - F_0) = B_{t,T}(F_{t,T} - F_0)$$

  which is a formula we saw in week one.

  ▶ Also, if we solve for the forward rate that makes V_F zero, it's given by $F_0 = S_t e^{r(T-t)}$, which we also knew.

- Old stuff there. But what about a forward on, say, the consumer price index? We couldn't do that in week one. But now we at least have a framework for approaching it.

  Let π_t denote the price level index.

  ▶ First, we are going to suppose π follows

  $$d\pi = \iota\pi\, dt + \sigma\pi\, dW.$$

That says, essentially that ι is the expected rate of increase, which is the same as the inflation rate. And now we are supposing ι and σ are constant.

- How do we risk-neutralize this thing? What's our recipe?
- Since the index itself isn't traded, our generic formula say just change the drift to $m - \lambda b$, or, here,

$$\pi\iota - \lambda\pi\sigma$$

- So this makes the adjusted price level process

$$\frac{d\pi}{\pi} = (\iota - \lambda\sigma)\,dt + \sigma\,dW$$

- And this is still just a constant-coefficient lognormal process **assuming λ is a constant.**
- Suppose it is. What is the forward worth?
- Do the same steps as for the stock, replacing S by π:

$$\begin{aligned} V(\pi_t, t) &= \mathrm{E}^{\star}[e^{-r(T-t)}(\pi_T - F_0)] \\ &= e^{-r(T-t)}(\mathrm{E}^{\star}[\pi_T] - F_0) \\ &= e^{-r(T-t)}\left(\pi_t e^{(\iota-\lambda\sigma)(T-t)} - F_0\right) \end{aligned}$$

which is a new result.

- Setting this to zero to solve for the current forward price

$$F = \pi_t e^{(\iota-\lambda\sigma)(T-t)}.$$

- Notice some interesting things about this formula:

  1. **If we know the true process (ι and σ), the forward price would give us λ.** This is an instances of the observation we made last time that we may not need to estimate market prices of risk if we take the prices of other derivatives as given.

  2. **The forward price is the expected spot price if and only if $\lambda = 0$.** This conclusion holds whether or not the process is lognormal. Also, it holds for traded underlyings too. We have already seen that, for currencies, for example:
$$\lambda = (\mu + r_f - r_d)/\sigma$$
So $\lambda = 0$ means $\mu = r_d - r_f$ which means
$$F = S_t \cdot e^{(r_d - r_f)(T-t)} = S_t \cdot e^{\mu \cdot (T-t)} = \mathrm{E}[S_T].$$

(B) Treasury Bonds.

- Consider pricing a zero-coupon riskless bond in a one state-variable model.

- The simplest model of interest rates people use is the **Vasicek model**

$$dr = \kappa(\bar{r} - r)\, dt + b\, dW$$

- This is an Ornstein-Uhlenbeck model. It has several convenient features, including the fact that, given today's rate r_t, the distribution of the rate at every future date is normal:

$$r_T \sim N(r_t\, e^{-\kappa(T-t)} + \bar{r}(1 - e^{-\kappa(T-t)}),\ \frac{b^2}{2\kappa}(1 - e^{-2\kappa(T-t)})\ ).$$

- The risk-neutralized version of the Vasicek process is

$$\begin{aligned} dr &= [\kappa(\bar{r} - r) - \lambda b]\, dt + b\, dW \\ &= \kappa(r^\star - r)\, dt + b\, dW \end{aligned}$$

  where $r^\star \equiv \bar{r} - \lambda b/\kappa$ is the new "long run mean" that results from our risk-neutralization.

- The payoff function for our bond is always one. So, according to our formula, its price must be

$$B_{t,T}(r_t) = \mathrm{E}^\star[DF_{t,T} \cdot 1]$$

  where $DF_{t,T}$ is the riskless discount factor to time T.

- Now that rates are random, this discount factor will be different along each path that r takes.

  ▶ So we need to use

$$DF_{t,T} = e^{-\int_t^T r_u \, du}$$

- So our pricing law just tells us that *zero coupon bond prices are the risk-neutralized expectation of this discount factor.*

- It happens that some mathematician already worked out this expectation for this particular model. The solution is

$$B_{t,T}(r_t) = e^{G_0(t,T) - r_t \, G_1(t,T)}$$

  where

$$\begin{aligned} G_1(t,T) &= \frac{1}{\kappa}(1 - e^{-\kappa(T-t)}) \\ G_0(t,T) &= \frac{1}{\kappa^2}(G_1(t,T) - T + t)(r^\star \kappa^2 - b^2/2) - \frac{(bG_1(t,T))^2}{4\kappa} \end{aligned}$$

  which describes a complete term-structure model in terms of the spot rate and four parameters.

- Do you see how you could now use this to try to extract λ from the government bond curve?

- For more complicated payoffs (like interest-rate options) there aren't closed-form formulas.

- But even without formulas, it is still really simple to just program a routine to simulate future paths of r under the adjusted model, compute the discount factor along each future path, and take the average payoff.
  - ▶ In Vasicek, all we do is change the long-run spot rate before simulating.
  - ▶ **Example:** The volatility of rates is 400 BP/year ($b = 0.04$), the mean reversion, κ, is 30% and the market price of interest rate risk is 15%.
    - ∗ The true long-term rate is 6%.
    - ∗ What is the "risk-adjusted" long term rate?
    - ∗ $r^\star \equiv \bar{r} - \lambda b/\kappa = 0.06 - (0.15\ 0.04/0.30) = 0.04 = 4\%$.
- The point is, the risk adjustment is easy. So pricing bonds is easy!

- If you want to keep the r process from going negative, you could use the **Cox, Ingersoll, Ross** (CIR) specification:

$$dr = \kappa^\star(r^\star - r)\, dt + s\sqrt{r}\; dW$$

which is another version of the CEV model.

- Assuming that this is the risk-neutralized version, the zero-coupon bond prices again have closed form

$$B_{t,T}(r_t) = e^{H_0(t,T)\ - r_t\, H_1(t,T)}$$

where

$$H_1(t,T) = \frac{2\,(\exp(h(T-t)) - 1)}{2h + (\kappa^\star + h)(\exp(h(T-t)) - 1)}$$

$$e^{H_0(t,T)} = \left[\frac{2h\,\exp(\frac{1}{2}(\kappa^\star + h)(T-t))}{2h + (\kappa^\star + h)(\exp(h(T-t)) - 1)}\right]^X$$

with $X = 2\kappa^\star r^\star / s^2$ and $h = \sqrt{(\kappa^\star)^2 + 2s^2}$.

- What does one have to assume about the market price of interest rate risk in order for the CIR model to have the same form (with only differing constants) under the true measure?
  - ▶ What does this mean economically? Is it plausible?

(C) Energy Swaps.

- In many countries there are now deregulated power markets enabling "spot" trading of electricity. However this does not mean electricity is a "traded risk factor" in the sense that we have been using the term.
  - ▶ You cannot buy *and hold* spot electricity in a portfolio.
  - ▶ There is no feasible way to store it – at least with current technology.
  - ▶ And, of course, if it cannot be stored, it can also not be loaned/borrowed. So short selling is impossible.
- However there are still important markets for electricity derivatives, like swaps and options on swaps.
- Even if these cannot be hedged with spot positions, our theory can still price them as long as there is more than one – and that we can find the appropriate market price of risk.
- Suppose we want to compute the no-arbitrage swap price for electricity to be delivered between T_1 and T_2.
- Two important features that we will need to take into account are (i) seasonality in prices, and (ii) long-term price trends.

- So suppose we model the spot price, P, as $f_t + X_t$ where f is the deterministic seasonal mean and X is an OU process with a stochastic long-run mean:

$$dX_t = \kappa(\bar{X}_t - X_t)\ dt\ +\ s\, dW_t$$
$$d\bar{X}_t =\ \bar{s}\, d\bar{W}_t.$$

(Assume these are the specifications under the risk-neutral measure.)

- Then the no-arbitrage fee paid continuously from T_1 to T_2 must equate:

$$\varphi \int_{T_1}^{T_2} e^{-ru}\ du \qquad \text{and} \qquad \int_{T_1}^{T_2} \mathrm{E}_0^\star[P_u] e^{-ru}\ du$$

- The expectation in the right-hand expression is $f_u + \mathrm{E}_0^\star[X_u]$.

- We know how to compute the expectation of an OU process with constant $\bar{X}$. We can find the current expectation if we observe that $X = \bar{X} + Y$ where $Y = X - \bar{X}$ obeys the simple equation $dY = -\kappa Y\ dt + \sigma_Y\ dW^Y$.

  - ▶ Here I use the fact that the diffusion terms of dY are $s\, dW_t - \bar{s}\, d\bar{W}_t$ and the sum of two Brownian motions is still a Brownian motion.

  - ▶ In this case, $\sigma_Y^2 = s^2 + \bar{s}^2 + 2\rho s\ \bar{s}$, but we will not need this.

- So now we know $\mathrm{E}_0^\star[X_u] = \mathrm{E}_0^\star[\bar{X}_u + Y_u] = \mathrm{E}_0^\star[\bar{X}_u] + = \mathrm{E}_0^\star[Y_u] = \bar{X}_0 + Y_0 e^{-\kappa u} = \bar{X}_0(1 - e^{-\kappa u}) + X_0 e^{-\kappa u}$ – which is the same as for an ordinary OU!

- Completing our problem, we find

$$\varphi = \frac{\int_{T_1}^{T_2} [f_u + \bar{X}_0 + (X_0 - \bar{X}_0)e^{-\kappa u}]e^{-ru}\ du}{\int_{T_1}^{T_2}\ e^{-ru}\ du}$$

- This is the sum of (i) the long run level today, (ii) the weighted average seasonal mean over the delivery period, and (iii) a weighted mean-reversion term equal to

$$(X_0 - \bar{X}_0)\frac{\int_{T_1}^{T_2} e^{-(r+\kappa)u}\ du}{\int_{T_1}^{T_2}\ e^{-ru}\ du}$$

- Another realistic feature of electricity markets is *seasonal volatility*.
  - ▶ How would that have affected the calculation in this example?

(D) Path-Dependent Options.

- Monte Carlo techniques can easily be used to value types of exotic derivatives we haven't even tried to price before.

- A path-dependent option is a claim whose payoff is a function of what happened to the underlying over the time period t to T. The most popular examples are

  Asian Options. These are options which pay the difference between the *average* price of the underlying over the time period and the strike. Alternatively, *average strike options* gives you the right to buy/sell the underlying at the average.

  Lookback Options. Are similar except, instead of the right to buy/sell at the average, they confer the right to sell/buy the the high or low. So a lookback put essentially guarantees you the ability to "sell at the top".

- Now we have to be careful before trying to value these things: *Are we sure their values obey our PDE?* And, if so, what are the state variables and what are their prices of risk?

  ▶ Remember, we are only allowed to apply our expectations recipe if we know the value of our claim is a function of some continuous, Brownian motion type processes, and hence satisifies the PDE.

  ▶ For example, the right to buy at the low has payoff of $S_T - L_T$ where L_T is the low between t and T. It is NOT the right type of process. (Draw it).

- Luckily, $Z_t \equiv S_t - L_t$ IS a type of diffusion, called a *reflected Brownian motion*, which does fall into our category of acceptable processes.
- And, although Z_t is not the price of a traded asset, it is instantaneously perfectly correlated with S_t. This means that there is really only one source of randomness. (So we do not need to find a λ for Z, for example).

- The tricky thing about these claims is that you have to keep track of the whole history of the underlying's price to know what they are worth at the end.
- But for Monte Carlo methods this is no problem at all! You are already building entire future paths. So it's no extra work (or very little) to figure out what the high was, say, over each path.
- So the payoffs are just as simple to compute as for plain calls and puts. And again all you do is compute a whole bunch of them and average the (discounted) result.

(E) Options on Strategies.

- Exotic options like those above still just have one underlying. But we can handle much more than that.

- Consider an option on a portfolio of stocks *whose weights change over time.*

- Suppose you are a fund manager and you want to insure a portfolio whose holdings (shares of each stock) is held fixed. Then your percentage allocations vary with their performances.

- For example, suppose there are just two stocks A and B with prices S_A and S_B, and you hold N_A shares of A and N_B shares of B.

- Then, even if they are independent, the volatility of your portfolio is
$$\frac{\sqrt{N_A^2 S_A^2 \sigma_A^2 + N_B^2 S_B^2 \sigma_B^2}}{N_A S_A + N_B S_B}$$
which is a complicated, time varying function of S_A, and S_B.

- But we know how to price options on it! Just set the expected returns of both stocks to r (no prices of risk here) and start simulating.

- You can extend this idea to price derivatives on *arbitrary* time-varying strategies, as long as you can specify the portfolio weights as functions of the state variables.

V. Beyond Monte Carlo

- Before you jump to the conclusion that you now know how to handle any derivatives problems, I have to give you the bad news: Monte Carlo techniques have a few practical drawbacks.

- Roughly, these fall into two categories: computational difficulties, and inherent limitations.

- The computational difficulties arise because the averages you calculate are, by definition, *approximations*. That means there is always approximation error. That means you have to (a) keep track of the error; and (b) make it small.

- In fact there are two sources of approximation error:
  - ▶ The error that comes from only being able to approximate an expectation with an average of a finite number of draws; and

  - ▶ The error that comes from approximating a continuous-time process by a discrete-time one.

- In principle, we can deal with both issues with enough computer power. Just scale up the number of paths, and scale down the time interval Δt.

- And, in fact, the second type of error declines quickly as you shrink the time interval. It is essentially zero when you use $\Delta t =$ one day.

- But, an unfortunate mathematical law tells us that the first type of error goes down at a slow rate. Specifically, the error in the average for a payoff function $g(S_T)$ with M simulations of S is approximately

$$\sqrt{\frac{\mathrm{Var}[g(S_T)]}{M}}.$$

The numerator is the variance of the payoffs experienced over the M draws. It doesn't change (much) with M. The problem is the denominator only gets large very slowly.

- For instance, if you start with 1000 paths and discover that your error is 10%, and your boss tells you to get it down to 1%, you have to draw

$$\frac{\text{old error}}{\text{new error}} = \frac{.10}{0.01} = 10 = \sqrt{\frac{\text{new } M}{\text{old } M}}$$

which means new $M = 100\times$ old M or 100,000 draws!

- As you know from your other courses, there are various numerical tricks for making things better.

- Even with state-of-the-art technology, though, Monte Carlo techniques may still be too slow for:

(A) Valuing lots of derivatives in real-time; and

(B) Valuing securities whose payoff variance is high, such as AAA coporate bonds or deep out-of-the-money options.

- In addition, the nature of derivatives problems also imposes some *inherent* limitations on the use of Monte Carlo.

- There are two things you need to be aware of:

1. **It's very hard to value American options with Monte Carlo.** This seems strange, given all the other exotics we can do. But Americans are different, because the early exercise floor imposes a boundary condition that depends on the value of the derivative at the time. If you think about a grid, there we valued by going backwards in time. So at each date we knew what the option's value was. But now we are trying to value by simulating forward in time. And, at intermediate time-steps, we don't have any way of knowing what the derivative will turn out to be worth.

2. **Monte Carlo only delivers the price of the derivative.** For nearly every derivative application, we need to know more than the price. We need all the hedge ratios to tell us how to replicate it, or take away our risk. Monte Carlo tells you what the thing is worth starting from the world right now. It can't tell you what it would be worth if things changed a little bit (like a small move in the underlying). To get those, we just have to change the starting conditions a little bit *and run the whole thing again.* If we want to know hedge ratios with respect to 5 quantities (perhaps volatility, interest rates, time, and two underlyings), we have to repeat the process 5 times.

- So sometimes we have to revert to other PDE solution techniques.

- Let me briefly sketch one alternative approach: **finite difference methods**.

- Suppose we have an asset, C, whose payoff we know at time-T, and which obeys:

$$-C_t = \frac{1}{2}s^2\sigma^2 C_{\sigma\sigma} + \frac{1}{2}\sigma^2 S^2 C_{SS} + rSC_S - rC + \widetilde{\kappa}(\widetilde{\sigma} - \sigma)C_\sigma.$$

- If we approximate dt as Δt, then we can view $C_t = \frac{\partial C}{\partial t}$ as $[C(t) - C(t - \Delta t)]/\Delta t$ and view the PDE as specifying a recipe for moving backwards in time.
  - ▶ Conceptually, if we know the solution at T we could evaluate all the partial derivatives using that solution.
  - ▶ Then we could just iterate back:

    $C(t - \Delta t) = C(t) + \Delta t\,(\text{all the other terms evaluated at T})$.
  - ▶ Then, just like in the binomial model, we'd repeat the process again for time $t - 2\Delta t$, and continue back to today.

- Finite difference methods are ways of making this practical by choosing a discrete grid of $[S, \sigma]$ points and then, at any point in the grid, say (S_i, σ_j), approximating, e.g.,

$$\frac{\partial^2 C}{\partial S^2} \approx \frac{C(S_{i+1}, \sigma_j) - 2C(S_i, \sigma_j) + C(S_{i-1}, \sigma_j)}{(\delta S)^2}$$

and

$$\frac{\partial C}{\partial \sigma} \approx \frac{C(S_i, \sigma_{j+1}) - C(S_i, \sigma_{j-1})}{2\Delta\sigma}.$$

And so on.

- While this is all fine conceptually (as long as the grid spacing is small enough), it turns out the simple procedure I outlined above is not very good numerically.

- It doesn't keep the solutions smooth enough.
- And little errors can cause it to blow up.

- To ensure that the solutions are stable, one usually has to evaluate some of the approximate partial derivatives using the grid values at the earlier time step.
  - These are called *implicit* methods, because you are solving for those earlier time step values at the same time that you are using them implicitly to approximate the derivatives.

- To illustrate, suppose we are at time t_{n-1} and we are solving for the values along one column of our grid with $\sigma = \sigma_j$ fixed.
  - Then we can write a *linear system* of equations that solves for C for all the S_i values at once, as follows.
  - First evaluate $\frac{\partial C}{\partial \sigma}$ and $\frac{\partial^2 C}{\partial \sigma^2}$ using the simple explicit replacements using the C function at t_n. Call the sum of these terms W_j^0 (a column vector).
  - Then bring all the other terms over to the other side:

$$C_{i,j} - \Delta t \left(\frac{1}{2} \sigma_j^2 S_i^2 \frac{C_{i+1,j} - 2C_{i,j} + C_{i-1,j}}{(\Delta S)^2} + rS_i \frac{C_{i+1,j} - C_{i-1,j}}{2\Delta\sigma} - rC_{i,j} \right)$$
$$= C_{i,j}(t) + \Delta t \; W_{i,j}^0.$$

where all the Cs on the left are the values at t_{n-1} that we have to solve for.

- This is a linear equation involving $C_{i_1,j}$, $C_{i,j}$, and $C_{i-1,j}$.
  - $*$ And we have one such equation for each i.
  - $*$ So if we stack all N_i equations, we have a complete system $W_j = K_j\ C_j$ that looks like:

$$\begin{bmatrix} \vdots \\ \vdots \\ W_{i,j} \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} & & \ddots & & & & \\ & 0 & k^+_{i+1,j} & k^0_{i+1,j} & k^-_{i+1,j} & 0 & \cdots \\ \cdots & 0 & k^+_{i,j} & k^0_{i,j} & k^-_{i,j} & 0 & \cdots \\ & \cdots & 0 & k^+_{i-1,j} & k^0_{i-1,j} & k^-_{i-1,j} & 0 \\ & & & & & \ddots & \end{bmatrix} \cdot \begin{bmatrix} \vdots \\ C_{i+1,j} \\ C_{i,j} \\ C_{i-1,j} \\ \vdots \end{bmatrix}$$

where $W_{i,j} = C_{i,j}(t) + \Delta t\ W^0_{i,j}$ and
$k^+_{i,j} = -\sigma_j^2 S_i^2 \frac{\Delta t}{2\Delta S^2} - rS_i \frac{\Delta t}{2\Delta S}$ and
$k^-_{i,j} = -\sigma_j^2 S_i^2 \frac{\Delta t}{2\Delta S^2} + rS_i \frac{\Delta t}{2\Delta S}$ and
$k^0_{i,j} = 1 + r\Delta t - \sigma_j^2 S_i^2 \frac{\Delta t}{2\Delta S^2}$.

  - $*$ In fact we don't even have to invert the matrix K_j here because it is a tri-diagonal form that allows the system to be solved quickly by back-substitution. (MATLAB does this automatically via their backslash matrix division operator.)
- Using this approach we can solve the linear system for each value j of our grid and thus fill in the time t_{n-1} grid.

- I skipped a step here however: we actually can't build the system as I described at the very top and very bottom points of our i-vector.
  - The $i = 1$ and $i = N_i$ lines will involve C values that are off the grid.
  - So the PDE only actually yields $N_i - 2$ equations.
  - We have to supply two more equations to close the system.
  - If we know the value exactly on one of the boundaries, that provides an additional equation.
  - Otherwise, another possibility is to make the grid large enough that we feel confident that the *curvature is going to zero* near the boundary. (Of course we need to test this assumption.)
    - Then we would have an additional equation that said, for example,

$$[1, \quad -2, \quad 1] \quad \begin{bmatrix} C_{N_i,j} \\ C_{N_i-1,j} \\ C_{N_i-2,j} \end{bmatrix} = 0.$$

- Finally, the routine I just described probably struck you as odd becuase it treated the σ derivatives differently from the S derivatives.

- In fact, the *alternating direction implicit* algorithm (ADI) is a 2-step procedure:

  1. To go from t_n to t_{n-1}, do the steps above.

  2. Then, to go from t_{n-1} to t_{n-2}, switch i and j. That is, evaluate the S partials explicitly, and solve the implicit system along line of fixed S, evaluating the σ partials implicitly.

- Finite difference methods restore many of the benefits of binomial trees, because, by working backwards in time, it is easy to evaluate early-exercise type decisions.

VI. Summary

- We learned that, in principle, we can handle derivative problems involving any number of sources of risk. We discussed several types of applications:

- In practice, these models are tricky to fit and require strong assumptions. **Use the simplest model you can get away with.**

- The method of pricing derivatives by taking "risk-neutral" discounted expectations applies to all problems which can be described in terms of our fundamental PDE

- The general transformation is to change the expected returns of all traded processes to r (or $r - y$), and to change the drifts of all non-traded processes to $m - \lambda b$. Also:
  - ▶ Volatilities and correlations don't get changed.
  - ▶ The discount factor goes inside the expectation.

- When we can't figure out the exact expectation, we can often approximate it accurately by Monte Carlo simulation.

- When Monte Carlo isn't practical, we may need finite differences or other approaches.

Lecture Note 7.1: Summary of Notation

Symbol	Page	Meaning
dr	*p2*	*change in instantaneous risk-free rate over* dt
$m(), b()$	*p2*	*drift and diffusion functions of* dr
ρ, λ	*p3*	*correlation between* r *and* S*, and market price of* r *risk*
m, b, ρ, λ	*p4*	*analogous parameters when payout* dy *is random*
X	*p8*	*an arbitrary risk factor obeying a diffusion process*
C_X, *etc*	*p8*	*sensitivity* $(\partial C/\partial X)$ *of security* C *to* X
$DF_{t,T}$	*p15*	*discount factor to apply to random payouts at time* T *when future interest rates are unknown*
$\mathrm{E}^{\star}[\ ]$	*p15*	*expectation using risk-neutralized processes*
$C(), \Gamma()$	*p15*	*payoff and cash-flow as a function of state-variables* X
c_0, c_1	*p19*	*expected growth and volatility of generic process* dX
$\mathrm{E}[\ ]$	*p19*	*expectation using true model for* X
V_F	*p20*	*value of an old forward contract at forward price* F_0 *when the current spot price is* S_t *at* t
π, ι	*p20*	*consumer price index and inflation rate*
σ, λ	*p21*	*volatility of price index and market price of* π *risk*
κ, r_0	*p23*	*mean-reversion rate and long-run level of riskless rate* r
b, λ	*p23*	*diffusion coefficient of riskless rate and market price of* r *ris*
$r^{\star}$	*p23*	*risk-neutralized long-run level of riskless rate*
$\int_t^T r_u\, du$	*p23*	*sum of future instantaneous rates between* t *and* T
G_0, G_1	*p24*	*coefficient functions in Vasicek bond solution*
Z_t	*p30*	*difference between* S_t *and the low,* L_t*, between 0 and* t
S_A, S_B, N_A, N_B	*p32*	*prices and number of shares of two stocks* A *and* B
Γ_{t_i}	*p33*	i*th coupon payment of structured note.*
$F(T)$	*p34*	*Principal repayment on rainbow note*
$\mathrm{Var}[g(S_T)]$	*p36*	*variance of realized payoffs* $g(S_T)$ *in a sample of* M *random paths of* S