Derivative Securities – Fall 2012 – Section 10
Notes by Robert V. Kohn, extended and improved by Steve Allen.
Courant Institute of Mathematical Sciences.

This is the third of 3 sections on the pricing of options on interest-based instruments – i.e. pricing of bond options, caps, floors, and swaptions. We start with a brief discussion linking options on a zero-coupon bond to caplets (this is elementary and could logically have been in Section 9). Then we discuss the simplest continuous-time short-rate model, known as Vasicek's model. Finally, we discuss the relationship between forward and futures prices when interest rates are not deterministic. Except for the first segment, this section is more mathematical than most of the course – consider it a hint of what lies ahead, when (for example in the class Continuous Time Finance) the full arsenal of Stochastic Calculus is available for use.

I have not discussed convexity and timing adjustments, but those topics are also important. You can read something about them in Hull's Chapter 29.

****

An option on a zero-coupon bond is essentially a caplet or floorlet. In Section 9 we discussed options on zero-coupon-bonds, caplets, and floorlets separately. However they are closely related. Let me explain why, focusing on the case of a put option on a zero-coupon bond. Recall that such an option gives its holder the right to sell (at time T, the maturity date of the option) a zero-coupon bond (with maturity $T' > T$, worth one dollar at maturity) for price K (the strike price). The option's payoff (at time T) is thus $(K - B(T,T'))_+$.

Let's related this to a caplet on the term lending rate from T to T', with notional principal one dollar and strike R_0. It pays $(R - R_0)_+\Delta t$ at time T', using the notation $R = R(T,T')$ for the term rate and $\Delta t = T' - T$. Therefore the caplet's value at time T is $B(T,T')(R - R_0)_+\Delta t$, or equivalently (remembering the relationship between $B(T,T')$ and $R(T,T')$)

$$\frac{1}{1 + R\Delta t}\max\{(R - R_0)\Delta t, 0\}.$$

Since

$$1 - \frac{1 + R_0\Delta t}{1 + R\Delta t} = \frac{(R - R_0)\Delta t}{1 + R\Delta t}$$

the value of the caplet at time T has the alternative expression

$$(1 + R_0\Delta t)\max\left\{\frac{1}{1 + R_0\Delta t} - \frac{1}{1 + R\Delta t}, 0\right\}$$

which we recognize as

$$(1 + R_0\Delta t)(K - B(T,T'))_+ \quad \text{with } K = \tfrac{1}{1+R_0\Delta t}.$$

Thus, the caplet is equivalent to $(1 + R_0\Delta t)$ puts on a zero-coupon bond, with strike $(1 + R_0\Delta t)^{-1}$.

A similar argument connects a floorlet with calls on a zero-coupon bond. Since a cap is a collection of caplets, it is equivalent to a portfolio of puts on zero-coupon bonds. Similarly, a floor is equivalent to a portfolio of calls on zero-coupon bonds.

****

The simplest continuous-time interest rate model. We discussed in Section 9 how interest-based instruments can be modeled using an interest rate tree. The continuous-time analogue of that discussion is the use of a *short-rate model*. Such a model specifies a stochastic differential equation for the short term interest rate $r(t)$ under the risk-neutral probability, say $dr = \alpha(r,t)\,dt + \beta(r,t)\,dw$. The solution determines the value of a money-market account: it solves $dA/dt = r(t)A$ with $A(0) = 1$, or in other words

$$A(t) = \exp\left(\int_0^t r(s)\,ds\right).$$

The value $V(t)$ of any tradeable is then determined, as discussed in Section 7, by the condition that $V(t)/A(t)$ must be a martingale (under the risk-neutral probability). Applying this to the value $B(t,T)$ of a zero-coupon bond with maturity T and principal one dollar, we get

$$B(t,T) = E_{\rm RN}\left[e^{-\int_t^T r(s)\,ds} \text{ given info at time t}\right]. \tag{1}$$

The simplest example is the Vasicek model, which assumes that under the risk-neutral probability the short rate solves

$$dr = (\theta - ar)\,dt + \sigma\,dw \tag{2}$$

with θ, a, and σ constant and $a > 0$. We have seen something very similar before: if r solves (2) then $y = r - \theta/a$ solves $dy = -ay\,dt + \sigma\,dw$; this is the Ornstein-Uhlenbeck process we discussed in Section 6.

Do we have the right to assume (2)? Yes and no. Yes, in the sense that there the short rate could in principle solve *any* SDE; there is no structural condition, because $r(t)$ is not itself the price of any tradeable. No, in the sense that (2) has only three parameters, so it cannot possibly be calibrated to match the current yield curve. Moreover in this model $r(t)$ can be negative, which shouldn't really happen for an interest rate. So this model is a toy – a cartoon version of a continuous-time short-rate model, which permits us to see the main ideas. (The Hull-White model assumes that $dr = (\theta(t) - ar)\,dt + \sigma\,dw$, where $\theta(t)$ is a deterministic function of t that must be specified. This model can be analyzed using the same ideas we'll apply here to Vasicek. By choosing $\theta(t)$ properly, the Hull-White model can be calibrated to any yield curve at a given time.)

A key feature of (2) is *mean reversion*: if $r > \theta/a$ then the drift term is negative, while if $r < \theta/a$ then the drift term is positive. So $r(t)$ has a tendency to return to θ/a, though noise (the dw term) keeps pushing it away. This feature is reasonable for a short-rate model, since interest rates rarely stay very high or very low for a long time.

Hull-White shares with Vasicek the flaw that $r(t)$ can be negative. There are variants (for example, the Cox-Ingersoll-Ross short-rate model) that avoid this problem. However, Hull-White and its variants have a more serious shortcoming: while they can be calibrated to today's yield curve, there is little freedom to insert information about how the yield curve is expected to *evolve*. To include such information, one must take an entirely different approach, such as that of Heath-Jarrow-Morton.

Enough perspective. Let's return to Vasicek model (2). Our goals are to

(i) find explicit formulas for the discount factors $B(t,T)$ under this model; and

(ii) show that the hypothesis underneath Black's formula is valid for this model.

Getting started: we gave an explicit solution for the Ornstein-Uhlenbeck process in Section 6. A similar procedure leads to an explicit formula for the solution of (2):

$$d(e^{at}r) = e^{at}\,dr + ae^{at}r\,dt = \theta e^{at}\,dt + e^{at}\sigma\,dw,$$

so

$$e^{at}r(t) = r(0) + \theta\int_0^t e^{as}\,ds + \sigma\int_0^t e^{as}dw(s).$$

which simplifies to

$$r(t) = r(0)e^{-at} + \frac{\theta}{a}(1-e^{-at}) + \sigma\int_0^t e^{-a(t-s)}\,dw(s). \tag{3}$$

That calculation could have started at any time; thus

$$r(t) = r(s)e^{-a(t-s)} + \frac{\theta}{a}(1-e^{-a(t-s)}) + \sigma\int_s^t e^{-a(t-\tau)}\,dw(\tau). \tag{4}$$

We observe from (3) that $r(t)$ is Gaussian (the stochastic integral is Gaussian, because each term of the approximating Riemann sum is Gaussian, and sums of Gaussians are Gaussian.) Its mean is clearly

$$E[r(t)] = r(0)e^{-at} + \frac{\theta}{a}(1-e^{-at}),$$

and the variance is

$$Var\,[r(t)] = \sigma^2 E\left[\left(\int_0^t e^{-a(t-s)}\,dw(s)\right)^2\right] = \sigma^2\int_0^t e^{-2a(t-s)}\,ds = \frac{\sigma^2}{2a}(1-e^{-2at}).$$

We can use the preceding calculation to see that $B(t,T)$ has lognormal statistics (under the risk-neutral probability). Indeed, substitution of (4) (with t and s interchanged) into (1) gives

$$B(t,T) = C(t,T)e^{-D(t,T)r(t)} \tag{5}$$

with

$$D(t,T) = \int_t^T e^{-a(s-t)}\,ds, \quad \text{and} \quad C(t,T) = E\left[e^{-\int_t^T \left\{\frac{\theta}{a}(1-e^{-a(s-t)})+\sigma\int_t^s e^{-a(s-\tau)}\,dw(\tau)\right\}ds}\right].$$

Since $C(t,T)$ and $D(t,T)$ are deterministic and $r(t)$ is Gaussian, (5) shows that $B(t,T)$ is lognormal.

We promised an explicit formula for $B(t,T)$; in the one obtained just above, $C(t,T)$ is pretty explicit since integral is easily evaluated, but $D(t,T)$ is not so explicit. To do better, we use the following **Claim**: *Suppose a function $V(t,r)$ solves the PDE*

$$V_t + (\theta - ar)V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0 \tag{6}$$

with final-time condition $V(T,r) = 1$ for all r at $t = T$. Then $V(t,r(t)) = B(t,T)$ for all $t < T$. To justify the claim, observe that it is certainly true when $t = T$, since $V(T,r(T)) = 1 = B(T,T)$. So it is sufficient to show that $V(t,r(t))/A(t)$ is a martingale, where $A(t)$ is the balance of the money market account. Making repeated use of the Ito calculus, we have that if $z(t) = V(t,r(t))/A(t)$ then

$$dz = A^{-1}dV(t,r(t)) - A^{-2}V(t,r(t))dA = A^{-1}(dV(t,r(t)) - rV\,dt)$$

and

$$\begin{aligned} dV(t,r(t)) &= V_t\,dt + V_r dr + \tfrac{1}{2}V_{rr}dr dr \\ &= [V_t + (\theta - ar)V_r + \tfrac{1}{2}\sigma^2 V_{rr}]\,dt + \sigma V_r\,dw. \end{aligned}$$

Thus, if V solves (6) then

$$dz = A^{-1}V_r\sigma\,dw$$

i.e. $z(t)$ is a martingale. This proves the claim.

To get the desired explicit formula, we should obviously seek a solution of (6) of the form

$$V(t,r) = C(t,T)e^{-D(t,T)r}.$$

Substituting this into the PDE, we see that C and D must satisfy

$$C_t - \theta CD + \tfrac{1}{2}\sigma^2 CD^2 = 0 \quad \text{and} \quad D_t - aD + 1 = 0$$

for $t < T$, with final-time conditions

$$C(T,T) = 1 \quad \text{and} \quad D(T,T) = 0.$$

Solving for D first, then C, we get

$$D(t,T) = \frac{1}{a}(1 - e^{-a(T-t)})$$

and

$$C(t,T) = \exp\left[\left(\frac{\theta}{a} - \frac{\sigma^2}{2a^2}\right)(D(t,T) - T + t) - \frac{\sigma^2}{4a}D^2(t,T)\right].$$

The desired explicit formula for $B(t,T)$ is now

$$B(t,T) = C(t,T)e^{-D(t,T)r(t)}.$$

We turn now to the validity of Black's formula, for options on zero-coupon bonds. Based on the discussion in Section 9, our task is to show that[1] $B(T,T')$ is lognormal under the forward-risk-neutral measure. (Remember: this is the measure under which tradeables normalized by $B(t,T)$ are martingales.) We already know it is lognormal under the risk-neutral measure, but here we're interested in a different measure, associated with a different numeraire.

We discussed change of numeraire in Section 9 on a tree; we now discuss how it works in the continuous time setting. The risk-neutral measure is associated with the money-market account $A(t)$ as numeraire. Let N be another numeraire; our goal is to understand the revised probability under which $V(t)/N(t)$ is a martingale whenever V is the value of a tradeable. (When $N(t) = B(t,T)$ this is the forward-risk-neutral measure.) We can only use tradeables as numeraires, so (under the original, risk-neutral probability) the SDE for N is

$$dN = rN\,dt + \sigma_N N\,dw.$$

By Ito we have

$$d(A/N) = A\,d(N^{-1}) + N^{-1}\,dA$$

(there is no $dAd(N^{-1})$ term since $dA = rA\,dt$). A bit of algebra gives

$$d(A/N) = (A/N)\sigma_N^2\,dt - (A/N)\sigma_N\,dw.$$

When we use N as numeraire, the associated probability is characterized by the fact that A/N is a martingale, i.e.

$$d(A/N) = -(A/N)\sigma_N\,d\overline{w}$$

where $\overline{w}$ is a Brownian motion under the new probability. Evidently,

$$d\overline{w} = -\sigma_N\,dt + dw.$$

What's the point of all this? We need to write the SDE for the Vasicek process under the forward-risk-neutral probability, i.e. when the numeraire is $B(t,T)$. Recall that $B(t,T) = C(t,T)e^{-D(t,T)r(t)}$, so from Ito the volatility of $B(t,T)$ is $-D(t,T)\sigma$. Therefore the preceding calculation gives

$$d\overline{w} = \sigma D(t,T)\,dt + dw.$$

We conclude that the SDE for the interest rate is

$$dr = (\theta - ar)\,dt + \sigma\,dw = [\theta - ar - \sigma^2 D(t,T)]\,dt + \sigma\,d\overline{w}$$

where $\overline{w}$ is a Brownian motion under the forward-risk-neutral probability. To show that $B(T,T')$ is lognormal under the forward-risk-neutral probability, we have simply to repeat the argument we have earlier (when we showed that $B(t,T)$ is lognormal under the risk-neutral probability), but using the SDE

$$dr = [\theta - ar - \sigma^2 D(t,T)]\,dt + \sigma\,d\overline{w}$$

[1] This used to say $B(t,T)$; corrected 11/28/2012. Similar corrections done below without further footnotes.

in place of (2). The calculation is very similar to the one we did before (the fact that the drift depends on time doesn't disturb the calculation at all). Briefly:[2] one substitutes the ansatz $B(t,T') = E(t,T')e^{-F(t,T')r(t)}$ into the analogue of (6) to get ODE's for $E(t,T')$ and $F(t,T')$. Explicit formulas are a bit tedious to obtain – but they're not needed. Since $E(T,T')$ and $F(T,T')$ are deterministic and $r(T)$ is Gaussian, we see (just as in the risk-neutral setting) that $B(T,T')$ is lognormal under the forward-risk-neutral probability.

****

Forwards versus futures. We showed in Section 1 that when interest rates are deterministic, the forward price of a tradeable is equal to its futures price (for the same delivery date). In a stochastic interest rate environment this is false. We can see why, and get a handle on the relationship between the two, by considering the following argument. (This material is drawn from Section 12.3 of *Quantitative Modeling of Derivative Securities* by M. Avellaneda and P. Laurence, Chapman & Hall, 1999.)

A typical interest rate future involves 3-month Eurodollar contracts: at the contract's maturity the holder must make a 3-month loan to the counterparty, at interest rate equal to the 3-month-term LIBOR rate. We have called this rate $R(t,T)$; here T=t + 3 months, and t is the maturity date of the futures contract. We know that the associated futures price $\tilde{f}_0(t,T)$ – which determines the daily settlements during the course of the contract – is a martingale under the risk-neutral probability, in other words

$$\tilde{f}_0(t,T) = E_{\mathrm{RN}}\left[R(t,T) \text{ given info at time t}\right].$$

Let us seek a similar representation for the forward term rate $f_0(t,T)$, defined as usual by

$$\frac{1}{1+f_0(t,T)\Delta T} = F_0(t,T) = \frac{B(0,T)}{B(0,t)}$$

with $\Delta T = T - t$. Solving for $f_0(t,T)$ gives

$$f_0(t,T) = \frac{1}{\Delta T \cdot B(0,T)}(B(0,t) - B(0,T)).$$

Using

$$B(0,t) = E_{\mathrm{RN}}\left[e^{-\int_0^t r(s)\,ds}\right]$$

and the analogous expression for $B(0,T)$, we get

$$\begin{aligned} f_0(t,T) &= \frac{1}{\Delta T \cdot B(0,T)} E_{\mathrm{RN}}\left[e^{-\int_0^t r(s)\,ds} - e^{-\int_0^T r(s)\,ds}\right] \\ &= \frac{1}{B(0,T)} E_{\mathrm{RN}}\left[e^{-\int_0^t r(s)\,ds} \cdot \frac{1 - e^{-\int_t^T r(s)\,ds}}{\Delta T}\right] \\ &= \frac{1}{B(0,T)} E_{\mathrm{RN}}\left[e^{-\int_0^t r(s)\,ds} \cdot \frac{1 - B(t,T)}{\Delta T}\right], \end{aligned}$$

[2] Better explanation inserted here, 11/28/2012.

making use in the last step of the fact that risk-neutral expectations are determined working backward in time. Now, the relation $B(t,T) = 1/[1 + R(t,T)\Delta T]$ can be rewritten as

$$\frac{1 - B(t,T)}{\Delta T} = R(t,T)B(t,T),$$

so we have shown that

$$\begin{aligned} f_0(t,T) &= \frac{1}{B(0,T)} E_{\mathrm{RN}} \left[e^{-\int_0^t r(s)\,ds} R(t,T)B(t,T) \right] \\ &= \frac{1}{B(0,T)} E_{\mathrm{RN}} \left[e^{-\int_0^t r(s)\,ds} R(t,T) e^{-\int_t^T r(s)\,ds} \right], \end{aligned}$$

using once more the fact that risk-neutral expectations are determined working backward in time. Combining the two exponential terms, we conclude finally that

$$f_0(t,T) = \frac{E_{\mathrm{RN}} \left[R(t,T) e^{-\int_0^T r(s)\,ds} \right]}{E_{\mathrm{RN}} \left[e^{-\int_0^T r(s)\,ds} \right]}.$$

Thus the forward rate $f_0(t,T)$ is *not* the risk-neutral expectation of the term rate $R(t,T)$. Rather it is the expectation of $R(t,T)$ with respect to a different probability measure, the one obtained by weighting each path by $\exp\left(-\int_0^T r(s)\,ds\right)$.

It is clear from this calculation that forward rates and futures prices are different. We can also see something about the relation between the two. In fact, writing $R = R(t,T)$ and $D = \exp\left(-\int_0^T r(s)\,ds\right)$ we have

$$\text{forward rate} - \text{futures price} = \frac{E[RD] - E[R]E[D]}{E[D]}.$$

where E represents risk-neutral expectation. If R and D were independent the right hand side would be zero and forward rates would equal futures prices. In general however we should expect R and D to be negatively correlated, since R is a term interest rate and D is a discount factor. Recognizing that $E[RD] - E[R]E[D]$ is the covariance of R and D, we conclude that this expression should normally be negative, implying that

$$\text{forward rate} < \text{futures price}.$$

This is in fact what is observed (the difference is relatively small). A scheme for adjusting the futures price to obtain the forward rate is sometimes called a “convexity adjustment”. It should be clear from our analysis that different models of stochastic interest rate dynamics should lead to different convexity adjustment rules in this context.