Interest Rate Models

Craig Pirrong
Bauer College of Business
University of Houston

April 14, 2021

The basic building block for interest rate modeling is a zero coupon bond, i.e., a security that pays \$1 at maturity, with no intervening cash flows.

Denote the time-t price of a zero that matures at T as $P_{t,T}$. Further, denote $p_{t,T} = \ln P_{t,T}$. The continuously compounded "yield" on the zero is:

$$Y_{t,T} = \frac{-p_{t,T}}{T-t}$$

That is:

$$P_{t,T} = e^{-Y_{t,T}(T-t)}$$

The "yield curve" is a locus of points plotting the relationship between yield and time to maturity. Yield curves can assume a variety of shapes– upward sloping, downward sloping, and humped. This is also called the "term structure" of interest rates.

Interest rate–"fixed income"–products are the most important securities, and the most heavily traded derivatives (e.g., interest rate swaps). Therefore, there has been extensive work on modeling these instruments, which involves modeling the yield curve.

The rates $Y_{t,T}$ are "spot" rates, which are used to discount cash flows from T back to t. An important concept is that of the forward rate, which is the rate at which one discounts back between two dates in the future, say T_1 and $T_2 > T_1$:

$$F_{t,T_1,T_2} = -\frac{p_{t,T_2} - p_{t,T_1}}{T_2 - T_1}$$

The instantaneous forward rate is:

$$f_{t,T} = -\frac{dp_{t,T}}{dT}$$

Note:

$$P_{t,T} = e^{-\int_t^T f_{t,s} ds}$$

and thus:

$$Y_{t,T} = \frac{\int_t^T f_{t,s} ds}{T - t}$$

There are two basic strands of modeling. The older strand models the dynamics of a theoretical abstraction–the instantaneous risk free (i.e., default free) interest rate r_t–and then explores the implications of these dynamics for the pricing of zeros. The new strand–"market models"–model the dynamics of observable rates, including Libor rates and swap yields.

1 Spot Rate Models

Spot rate models characterize the dynamics of the instantaneous, condtinuously compounded rate r_t. The general form of the simplest kind of model is:

$$dr_t = \alpha(r_t, t)dt + \sigma(r_t, t)dW_t$$

In this framework, the price of a zero is:

$$P_{t,T} = \tilde{E}(e^{-\int_t^T r_s ds})$$

Different models have different choices of $\alpha(r_t, t)$ and $\sigma(r_t, t)$.

The proto-model is that of Vasicek, which posits:

$$dr_t = \phi(\bar{r} - r_t)dt + \sigma dW_t$$

Here $\bar{r}$ is a long-run mean interest rate to which the spot rate reverts at rate $\phi > 0$. That is, the rate trends down when it is above the long term mean, and trends up when it is below the long run mean.

Note that rates can become negative in this model. Once upon a time that was considered a bug. Now it is a feature!

This model can generate yield curves of various shapes. Traders consider the model deficient because it cannot fit the prices of all zero coupon bonds (i.e., it cannot match the term structure exactly). This is because there are only four things to adjust–r_t, ϕ, $\bar{r}$ and σ–but a continuum of bonds. We will discuss ways to address this issue later.

Another model is Cox-Ingersoll-Ross:

$$dr_t = \phi(\bar{r} - r_t)dt + \sigma_r r_t^{.5} dW_t$$

Here, interest rates cannot become negative: there is a reflecting barrier at zero, because if the rate ever hits zero the mean reversion term causes it to rise deterministically.

2 Valuing Zeros in Spot Rate Models: The PDE Approach

A zero is just a derivative of the spot rate. Therefore, our standard PDE approach is applicable. Thus, in the Vasicek model, a zero solves the PDE:

$$rP = \frac{\partial P}{\partial t} + [\phi(\bar{r} - r_t) - \lambda\sigma]\frac{\partial P}{\partial r} + .5\sigma^2\frac{\partial^2 P}{\partial r^2}$$

Through a change in the time variables:

$$rP = -\frac{\partial P}{\partial T} + [\phi(\bar{r} - r_t) - \lambda\sigma]\frac{\partial P}{\partial r} + .5\sigma^2\frac{\partial^2 P}{\partial r^2}$$

where λ is a market price of risk, which is necessary because r_t is not traded.

This PDE can be solved by making an inspired guess for the zero price:

$$P_{t,T} = e^{A(t,T)-B(t,T)r_t}$$

Plugging this into the PDE, and solving subject to the boundary conditions:

$$P_{T,T} = 1$$

$$P_{t,T}(0) = 1;$$

$$\lim_{r_t \to \infty} P_{t,T}(r_t) = 0;$$

will allow us to determine $A(t,T)$ and $B(t,T)$.

To simplify the notation, assume $t = 0$ and write $A(T)$ and $B(T)$. That is, rather than using a maturity date, use T to denote time to maturity.

Note that the boundary condition at maturity will be satisfied if:

$$A(0) - B(0)r = 0$$

Since this must hold for all r, $A(0) = 0$ and $B(0) = 0$.

Given the guess,

$$\frac{1}{P}\frac{\partial P}{\partial r} = -B(T)$$

$$\frac{1}{P}\frac{\partial^2 P}{\partial r^2} = -B^2(T)$$

$$\frac{1}{P}\frac{\partial P}{\partial T} = A'(T) - B'(T)r$$

Substituting into the PDE:

$$-B(T)\phi(\bar{r} - r) + .5B^2(T)\sigma^2 - A'(T) + B'(T)r - r = -B(T)\lambda\sigma$$

Again, this must hold for all r, which requires:

$$A'(T) = .5B^2(T)\sigma^2 - (\phi\bar{r} - \lambda\sigma)B(T)$$

and

$$B'(T) = 1 - B(T)\phi$$

This is a system of ordinary differential equations, which can be solved separately, solving the second one first:

$$\frac{dB}{dT} = 1 - \phi B$$

$$\int_0^T \frac{dB}{1-\phi B} = \int_0^T dT$$

$$-\frac{1}{\phi}\ln(1 - \phi B(T)) = T$$

which implies:

$$B(T) = \frac{1}{\phi}(1 - e^{-\phi T})$$

Now integrate the first expression:

$$A(T) = .5\sigma^2 \int_0^T B^2(T)dT - (\phi\bar{r} - \lambda\sigma)\int_0^T B(T)dT + C$$

where C is an integration constant that we will choose to satsify the boundary condition $A(0) = 0$. We can substitute for $B(T)$ based on our previous solution:

$$A(T) = \frac{.5\sigma^2}{\phi^2}\int_0^T (1 - 2e^{-\phi T} + e^{-2\phi T})dT - (\phi\bar{r} - \frac{\lambda\sigma}{\phi})\int_0^T (1 - e^{-\phi T})dT + C$$

Simplifying the integrals:

$$A(T) = \frac{.5\sigma^2}{\phi^2}(T + \frac{2e^{-\phi T}}{\phi} - \frac{e^{-2\phi T}}{2\phi}) - (\bar{r} - \frac{\lambda\sigma}{\phi})(T + \frac{e^{-\phi T}}{\phi}) + C$$

To solve for C:

$$A(0) = 0 = \frac{.5\sigma^2}{\phi^2}(\frac{2}{\phi} - \frac{1}{2\phi}) - (\bar{r} - \frac{\lambda\sigma}{\phi})\frac{1}{\phi}$$

Solving for C and substituting:

$$A(T) = \frac{.5\sigma^2}{\phi^2}(T + \frac{2e^{-\phi T} - 1}{\phi} - \frac{e^{-2\phi T} - 1}{2\phi}) - (\bar{r} - \frac{\lambda\sigma}{\phi})(T + \frac{e^{-\phi T} - 1}{\phi})$$

Now, we could stop there, but let's continue in a sadomasochistic fashion, by noting:

$$B^2(T) = \frac{1}{\phi^2}(1 - 2e^{-\phi T} + e^{-2\phi T})$$

$$\phi B^2(T) = 2\frac{1 - e^{-\phi T}}{\phi} + \frac{e^{-2\phi T} - 1}{\phi}$$

$$\phi B^2(T) - 2B(T) = \frac{e^{-2\phi T} - 1}{\phi}$$

Which, based on substituting back into our expression for $A(T)$:

$$A(T) = \frac{\sigma^2}{2\phi^2}(T - 2B(T) - .5\phi B^2(T) + B(T)) - (\bar{r} - \frac{\lambda\sigma}{\phi})(T - B(T)$$

Simplifying further gives:

$$A(T) = -\frac{\sigma^2}{4\phi^2}B^2(T) - (\bar{r} - \frac{\lambda\sigma}{\phi} - \frac{\sigma^2}{2\phi^2})(T - B(T))$$

Of course, remembering the connection between PDE and martingale methods, we should be able to derive the same expression by evaluating:

$$P_{t,T} = \tilde{E}e^{-\int_t^T(\bar{r} - r_s)ds - \int_t^T \sigma dW_s}$$

This will be left as a future exercise, which should be straightforward after we implement this method for a more complicated model: the Hull-White model.

3 The Hull-White Model

As noted earlier, traders don't like the fact that the Vasicek model (or the CIR model, for that matter) can't fit the prices of every zero exactly. So what to do? What to do?

No problem!: Add an infinite number of degrees of freedom to fit an infinite number of bond prices!

That's what Hull-White do. They posit:

$$dr_t = (\Theta_t - ar_t)dt + \sigma dW_t$$

where Θ_t is an arbitrary function. As we will see, we can choose Θ_t to fit all bond prices exactly.

As it turns out, working directly with this equation is somewhat cumbersome, so we will transform variables to make the problem isomorphic to one that is easy to solve–an Orenstein-Uhlenbeck ("OU") process:

$$dX_t = -\mu dt + \sigma dW_t$$

To solve this SDE, posit:

$$X_t = a(t)[X_0 + \int_0^t b(s)dW_s]$$

Then, applying Ito's Lemma:

$$dX_t = a'(t)[X_0 + \int_0^t b(s)dW_s]dt + a(t)b(t)dW_t$$

$$dX_t = \frac{a'(t)X_t}{a(t)}dt + a(t)b(t)dW_t$$

Now we match coefficients. First, we get an ordinary differential equation:

$$-\mu = \frac{a'(t)}{a(t)}$$

$$a(t)b(t) = \sigma$$

The boundary conditions for the ODE is $a(0) = 1$. Solving the ODE subject to this condition gives: $a(t) = e^{-\mu t}$. Thus:

$$e^{-\mu t}b(t) = \sigma$$

so

$$b(t) = e^{\mu t}\sigma$$

Thus, the solution to the SDE is:

$$X_t = e^{-\mu t}X_0 + \int_0^t e^{-\mu(t-s)}dW_s$$

Now that we know how to solve for OU processes, we can implement the following transformation:

$$y_s = r_s - \alpha(s)$$

where:

$$\alpha(s) = e^{-as}(r_0 + \int_0^s e^{au}\Theta_u du)$$

which implies:

$$dy_s = dr_s - d\alpha(s) = dr_s + ae^{-as}(r_0 + \int_0^s e^{au}\Theta_u du)ds - e^{-as}e^{as}\Theta_s ds$$

Thus:

$$dy_s = dr_s + a\alpha(s)ds - \Theta_s ds$$

Substituting from our original equation for dr_s, we get:

$$dy_t = (\Theta_t - ar_t)dt + \sigma dW_t - \Theta_t dt + a\alpha(t)$$

$$dy_t = -a(r_t - \alpha(t))dt + \sigma dW_t$$

$$dy_t = -ay_t + \sigma dW_t$$

which is an OU process. Solving for this process, given that $y_0 = 0$:

$$y_T = e^{-aT} y_0 + \sigma e^{-aT} \int_0^T e^{as} dW_s$$

$$y_T = \sigma e^{-aT} \int_0^T e^{as} dW_s$$

Recall:

$$P(0,T) = \tilde{E}_0 e^{-\int_0^T r_t dt} = \tilde{E}_0 e^{-\int_0^T y_t dt - \int_0^T \alpha(t) dt}$$

$$P(0,T) = \tilde{E}_0 e^{-\int_0^T r_t dt} = e^{-\int_0^T \alpha(t) dt} \tilde{E}_0 e^{-\int_0^T y_t dt}$$

Since

$$y_t = \sigma e^{-at} \int_0^t e^{au} dW_u$$

$$\int_0^T y_t dt = \sigma \int_0^T \int_0^t e^{-at} e^{au} dW_u dt$$

We now change the order of integration. Note $0 \leq u \leq t$, and $0 \leq t \leq T$. This implies $0 \leq u \leq T$, and $u \leq t \leq T$. Thus:

$$\int_0^T y_t dt = \sigma \int_0^T \int_u^T e^{a(u-t)} dt dW_u$$

First evaluate the inner integral:

$$\int_0^T \int_u^T e^{a(u-t)} dt = e^{au} \frac{e^{-au} - e^{-at}}{a} = \frac{1 - e^{-a(T-u)}}{a}$$

$$\int_0^T y_t dt = \sigma \int_0^T \frac{1 - e^{-a(T-u)}}{a} dW_u$$

Note this is an Ito Integral, and we can use the Ito Isometry to evaluate it given that we know the expectation of the exponential of a normal variate is the exponential of one-half its variance.

Specifically:

$$var(\int_0^T y_t dt) = \sigma^2 \int_0^T \frac{(1 - e^{-a(T-u)})^2}{a^2} du$$

Note:

$$\frac{\sigma^2(1-e^{-a(T-u)})^2}{a^2} = \frac{\sigma^2}{a^2}(1-2e^{a(u-T)}+e^{2a(u-T)})$$

The integral of the first term is easy:

$$\frac{\sigma^2 T}{a^2}$$

The integral of the second term is:

$$-\frac{2}{a^2}(1-e^{-aT})$$

The integral of the third term is:

$$\frac{\sigma^2}{2a^3}(1-e^{-2aT})$$

Putting it all together, we get that the variance of $\int_0^T y_t dt$ is:

$$\hat{\sigma}_T^2 = \sigma^2 \frac{2aT - 4(1-e^{-aT}) + 1 - 2^{-2aT}}{2a^3}$$

Thus:

$$P_{0,T} = e^{-\int_0^T \alpha(t)dt} E e^{\hat{\sigma}_T \tilde{Z}}$$

where Z is a standard normal variate. This implies:

$$P_{0,T} = e^{-\int_0^T \alpha(t)dt + .5\hat{\sigma}_T^2}$$

Take logs of both sides:

$$p_{0,T} = -\int_0^T \alpha(t)dt + .5\hat{\sigma}_T^2$$

To recover $\alpha(T)$ take the derivative with respect to T (using Leibniz's Rule):

$$\frac{\partial p_{0,T}}{\partial T} = -\alpha(T) + \frac{\sigma^2}{2a^2}(1-e^{-aT})^2$$

which implies:

$$\alpha(T) = -\frac{\partial p_{0,T}}{\partial T} + \frac{\sigma^2}{2a^2}(1-e^{-aT})^2$$

Recalling the definition of the instantaneous forward yield:

$$\alpha(T) = f_T + \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2$$

Thus, if we know forward rates for all maturities, we know $\alpha(T)$ for all maturities.

Recalling the definition of $\alpha(T)$:

$$\alpha(T) = e^{-aT}(r_0 + \int_0^T e^{au}\Theta_u du) = f_T + \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2$$

We can no recover Θ_t by taking the derivative with respect to T and applying Leibniz's Rule:

$$-ae^{-aT}(r_0 + \int_0^T e^{au}\Theta_u du) + e^{-aT}e^{at}\Theta_t = \frac{\partial f_T}{\partial T} - \frac{a\sigma^2}{2a^2}(1 - e^{-aT})e^{-aT}$$

Recognizing that

$$e^{-aT}(r_0 + \int_0^T e^{au}\Theta_u du) = -\frac{\partial p_{0,T}}{\partial T} = f_T$$

and substituting, we get:

$$-af_T + \Theta_T = \frac{\partial f_T}{\partial T} - \frac{a\sigma^2}{2a^2}(1 - e^{-aT})e^{-aT}$$

Which (finally!) implies:

$$\Theta_T = af_T + \frac{\partial f_T}{\partial T} - \frac{a\sigma^2}{2a^2}(1 - e^{-aT})e^{-aT}$$

That is, Θ_T is the sum of the speed of mean reversion multiplied by the instantaneous forward rate at T, f_T, plus the slope of the forward curve at T, and an adjustment term that depends on T and the speed of mean reversion.

The good news here is that if we have the yield curve, we can choose Θ_t to fit it exactly. The bad news is that if we have the yield curve, we can choose Θ_t to fit it exactly. Exact fitting means misfitting. But traders like neatness, so here we are.

4 Multifactor Affine Models

The Vasicek, CIR, and Hull-White are single factor "affine" models in which the instantaneous spot rate is the only state variable, and (log) bond prices are an affine function of that rate. Multifactor models specify multiple (usually latent) factors, $\mathbf{y_t}$:

$$\mathbf{y_t} = \phi(\bar{\mathbf{y}} - \mathbf{y_t})dt + \Sigma \mathbf{dW_t}$$

where $\mathbf{y_t}$ is an $n \times 1$ vector. The spot rate is a linear function of these factors:

$$r = \delta_0 + \delta' \mathbf{y_t}$$

Furthermore:

$$\mathbf{dW_{it}} = \sqrt{\alpha_i + \beta_i \mathbf{y_t}} dZ_{it}$$

where dZ_{it}, $i = 1, \ldots, n$ are uncorrelated Brownian motions.

Though obviously more complicated, these models can be solved in the same ways as the single factor models, by solvng a PDE or by expectation. Note that it is first necessary to change measures:

$$\mathbf{d\tilde{W}_{it}} = \sqrt{\alpha_i + \beta_i \mathbf{y_t}}(dZ_{it} - \lambda_i dt)$$

Note that the value of a zero is $P_{t,T}(\mathbf{y_t})$, and thus, there are first partial derivatives and second partial derivatives for each factor, and cross partials. This makes things messier, but the basic mechanics can be used. First, posit:

$$P_T = e^{\mathbf{A}(T) - \mathbf{B}'(T)\mathbf{y}}$$

Note that $\mathbf{A}(T)$ and $\mathbf{B}(T)$ are $n \times 1$ vectors. Then, plug this expression into the PDE. Simplifying implies systems of ordinary differential equations:

$$\frac{\partial \mathbf{B}(T)}{\partial T} = -\phi' \mathbf{B}(T) - \sum_{i=1}^{n} [\mathbf{B}(T)\lambda_i + .5(\Sigma' \mathbf{B}(T))_i^2]\beta_i + \delta$$

$$\frac{\partial \mathbf{A}(T)}{\partial T} = \sum_{i=1}^{n} [\mathbf{B}(T)\lambda_i + .5(\Sigma' \mathbf{B}(T))_i^2]\alpha_i - \mathbf{B}'(T)\phi\bar{\mathbf{y}} - \delta_0$$

Such systems of ODEs are readily solved numerically.

5 Market Models

Practitioners came to dislike affine models because the state variables do not correspond to anything actually quoted or traded in the market: Go to your bank and ask them what the current continuously compounded spot rate is!

Historically, in the market the two most important traded rates were LIBOR–the London Interbank Offered Rate–and swap rates. LIBOR is on its way to being phased out, but there no "term" rate replacement on offer.

To address practitioners' complaints, quants developed "market models" which specify observable, traded rates as the state variables. Thus, we have LIBOR Market Models and Swap Market Models. Let's consider each in turn.

5.1 LIBOR Market Models

LIBOR is intended to represent the cost of unsecured interbank borrowing, and hence reflect the marginal cost of funds to large financial institutions. LIBOR is (for now!) quoted for 1-3 weeks, and 1-12 months. Moreover, there are LIBOR forward rates: for example, there is the rate for borrowing lending between 3 and 6 months from now. These forward rates extend for periods years into the future. (The CME Eurodollar Futures contract prices basically give three month forward rates for periods extending out ten years.)

Now for some notation. Consider dates $T_1, T_2, \ldots, T_i, \ldots T_N$. Call $L_i(t)$ the forward rate quoted at t ("today") that involves borrowing at T_i and repaying at T_{i+1}. $L_i(t)$ is an annualized rate, and $[T_i, T_{i+1}]$ is not necessarily a year. The convention in LIBOR is that interest is calculated on an actual/360 basis, meaning that fractions of a year are calculated by counting the actual number of days between T_i and T_{i+1} and dividing by 360. Call this year fraction α_i. Then, borrowing \$1 at T_i requires repayment of $1 + \alpha_i L_i(t)$ at T_{i+1}.

This is a forward agreement: the parties agree to $L_i(t)$ at t, but no money changes hands then. This is an agreement to borrow or lend in the future.

Call $D_i(t)$ the time t price of a zero maturing at T_i and $D_{i+1}(t)$ as the time t price of the zero maturing at T_{i+1}. Thus:

$$0 = D_i(t) - (1 + \alpha_i L_i(t))D_{i+1}(t)$$

That is, the present value of the cash flows on the forward borrowing must equal zero, because no money changes hands at t.

This implies:

$$L_i(t) = \frac{D_i(t) - D_{i+1}(t)}{\alpha_i D_{i+1}(t)}$$

Note that this is a ratio of traded asset prices. Now we can utilize our understanding of numeraires to make $L_i(t)$ into a martingale. Specifically, we know that $L_i(t)$ is a martingale if we choose the zero maturing at the *maturity* of the forward loan the numeraire. Moreover, if we assume that $L_i(t)$ follows a geometric Brownian motion, we can therefore write:

$$\frac{dL_i(t)}{L_i(t)} = \sigma_i dW_{it}^{i+1}$$

Here, W_{it}^{i+1} is a Brownian motion. The superscript denotes the measure under which it is a Brownian motion, which is the measure corresponding to D_{i+1} as the numeraire.

We can now price some contingent claims. One common LIBOR-related contingent claim is a caplet, i.e., an option that provides a cap on the interest rate. Consider a caplet maturing at T_i, that has a payoff based on $L_i(t)$. To match the payoffs of the option to the payoff to the underlying forward loan, the payoff to this claim occurs at T_{i+1}, and is:

$$\max \alpha_i [L_i(T_i) - K, 0]$$

where K is the strike price, and $L_i(T_i)$ is the spot $[T_i, T_{i+1}]$ LIBOR rate at T_i.

We can price this using our martingale technique:

$$c(L_i(t), t) = D_{i+1}(t) E^{i+1} \frac{\max \alpha_i [L_i(T_i) - K, 0]}{1}$$

There is a 1 in the denominator, because this is the value of the numeraire at T_{i+1}. The superscript on the expectation operator again denotes that this is an expectation taken with respect to the "terminal measure" corresponding to the D_{i+1} numeraire.

This problem is isomorphic to the Black-Scholes model, and we can therefore show:

$$c(L_i(t), t) = \alpha_i D_{i+1}(t) [L_i(t) N(d_1) - K N(d_2)]$$

where

$$d_1 = \frac{\frac{L_i(t)}{K} + \int_t^T \sigma_i^2(s) ds}{\sqrt{\int_t^T \sigma_i^2(s) ds}}$$

and

$$d_2 = d_2 - \sqrt{\int_t^T \sigma_i^2(s) ds}$$

If σ_i is a constant,

$$\int_t^T \sigma_i^2(s) ds = \sigma_i^2 (T_i - t)$$

Note that we can only choose to make a single forward rate a martingale, because we only can choose one maturity zero as numeraire. However, many continent claims are dependent on multiple LIBOR rates. For example, variable rate mortgages may have payoffs that depend on 15 one-year LIBOR rates over the life of the loan. We can't make all these rates martingales: we can make one such. Once we choose which one, we have to adjust the drifts of the other rates.

One way to do this is to make the last-maturing rate of interest a martingale by choosing as numeraire the zero maturing on the maturity date of the loan corresponding to that rate. We then work backwards to progressively earlier rates.

For example, if we choose $L_i(t)$ to be a martingale, what about $L_{i-1}(t)$?

$$L_{i-1} = \frac{D_{i-1} - D_i}{\alpha_{i-1} D_i}$$

Now multiply both sides by D_i/D_{i+1}:

$$\frac{D_i L_{i-1}}{D_{i+1}} = \frac{D_{i-1} - D_i}{\alpha_{i-1} D_{i+1}}$$

Note that if we choose D_i as the numeraire, we can make L_{i-1} a martingale:

$$dL_{i-1} = \sigma_{i-1} L_{i-1} dW_{i-1}^i$$

where the subscript on the Brownian motion refers to the starting date corresponding to the rate, and the superscript refers to the numeraire. To change measure to the $i+1$ numeraire:

$$dW_{i-1}^{i+1} = dW_{i-1}^i + \lambda dt$$

Further:

$$\frac{D_i L_{i-1}}{D_{i+1}} = L_{i-1}(1 + \alpha_i L_i)$$

Moreover,

$$dL_{i-1}(1 + \alpha_i L_i) = (1 + \alpha_i L_i) dL_{i-1} + L_{i-1} \alpha_i dL_i + \alpha_i dL_i dL_{i-1}$$

Making our Girsanov substitution:

$$(1+\alpha_i L_i)\sigma_{i-1}L_{i-1}(dW_{i-1}^{i+1}-\lambda dt)+L_{i-1}\alpha_i\sigma_i L_i dW_i^{i+1}+\alpha_i\sigma_{i-1}\sigma_i L_{i-1}L_i\rho dt$$

where $\rho dt = dW_{i-1}^{i+1}dW_i^{i+1}$

Note that $L_{i-1}(1+\alpha_i L_i)$ is a ratio of asset prices, with the numeraire in the denominator. Thus, it is a martingale under the $i+1$ measure. To be a martingale, the drift must equal zero:

$$-(1+\alpha_i L_i)\sigma_{i-1}L_{i-1}\lambda+\alpha_i\sigma_{i-1}\sigma_i L_{i-1}L_i\rho = 0$$

which implies:

$$\lambda = \frac{\alpha_i\sigma_i\rho L_i}{1+\alpha_i L_i}$$

Which produces:

$$dL_{i-1} = -\frac{\alpha_i\sigma_i\sigma_{i-1}\rho L_i L_{i-1}dt}{1+\alpha_i L_i}+\sigma_{i-1}L_{i-1}dW_{i-1}^{i+1}$$

Now consider L_{i-2}. Repeating the foregoing implies:

$$dL_{i-2} = -\frac{\alpha_{i-1}\sigma_{i-1}\sigma_{i-2}\rho L_{i-1}L_{i-2}dt}{1+\alpha_{i-1}L_{i-1}}+\sigma_{i-2}L_{i-2}dW_{i-2}^{i}$$

where for simplicity I have assumed that the correlation all rates is the same for each pair. Substituting

$$dW_{i-2}^{i} = dW_{i-2}-\frac{\alpha_i\sigma_i L_i}{1+\alpha_i L_i}$$

produces:

$$dL_{i-2} = -\frac{\alpha_{i-1}\sigma_i\sigma_{i-2}\rho L_{i-1}L_{i-2}dt}{1+\alpha_{i-1}L_{i-1}}-\frac{\alpha_i\sigma_i L_i\sigma_{i-2}L_{i-2}\rho dt}{1+\alpha_i L_i}+\sigma_{i-2}L_{i-2}dW_{i-2}^{i+1}$$

We can continue recursively in this fashion to derive dL_{i-2}, dL_{i-3}, and so on, adding a new term to the drift each time.

The only practical way to implement this is via simulation. That is, to value say the 15 variable rate mortgage, we have to simulate paths for each of the rates, using these drift-adjusted equations.

5.2 Swap Market Models

As the name suggests, interest rate swaps are contracts for exchanging–“swapping”–cash flows based on interest rates. A “vanilla” swap has two “legs”–a floating leg and a fixed leg–and two payers–a floating payer and a fixed payer. The fixed payer agrees to pay a set (fixed) amount on a sequence of payment dates. The floating payer pays on these same payment dates a variable amount which depends on some short-term interest rate. Historically, this rate has been LIBOR. Note that the fixed price payer receives floating, and the floating price payer receives fixed. It is common to refer to positions as pay fixed/receive floating or pay floating/receive fixed.

A vanilla swap specifies a notional principal amount, e.g., $100 million. This amount never changes hands, but is merely used to determine the size of the fixed and floating payments. The swap also specifies the payment frequency, e.g., semi-annual or quarterly. This payment frequency determines the maturity of the rate used to determine the floating payment. For example, a semi-annual swap floating payment is based on a six-month interest rate.

Consider a swap entered at T_n, with payment dates $\{T_{n+1}, \dots T_N\}$. If $n > 0$, this is a “forward starting” swap. Further, α_i is the fraction of a year between T_i and T_{i+1}. (Most swaps use the actual/360 day count convention.) Assume the fixed rate is K (annualized)

The mechanics of the swap work as follows. On payment date $i + 1$, the fixed payer pays $\alpha_i K$. If the floating payment is based on LIBOR, the floating payer pays $\alpha_i L_i(T_i)$. Note that the floating payment is based on the reference floating rate quoted *on the previous payment date.* This matches up swap cash flows with cash flows on LIBOR-based loans. (There is a type of swap called the LIBOR in Arrears swap where the payment on a date is based on the rate quoted on the payment date.)

The value of the fixed payments in the swap is:

$$V_{fix}(t) = K \sum_{i=n}^{N-1} \alpha_i D_{i+1}(t)$$

What about the value of the floating payments? How do we estimate the uncertain cash flows? Well, we can use a hedging strategy. If I lend \$1 at LIBOR at T_n (the starting date of the swap), keep the interest payments, and reinvest the principal at LIBOR when the loan matures, and do that for each payment date, the interest payments match the floating payments on the swap. The value of the floating payments is therefore equal to the cost of implementing this strategy, which involves a cash flow of -1 at T_n and +1 at T_N. The value of this is:

$$V_{float} = D_n(t) - D_N(t)$$

The swap involves no cash flows at t (when it is negotiated), so the present value of the floating and fixed payments must be equal:

$$V_{fix}(t) = K \sum_{i=n}^{N-1} \alpha_i D_{i+1}(t) = V_{float} = D_n(t) - D_N(t)$$

This implies:

$$K = \frac{D_n(t) - D_N(t)}{\sum_{i=n}^{N-1} \alpha_i D_{i+1}(t)}$$

Note that is a ratio of traded asset prices. Therefore, if we choose $\sum_{i=n}^{N-1} \alpha_i D_{i+1}(t)$ to be the numeraire, K is a martingale under the equivalent measure.

$P_{n+1,N} := \sum_{i=n}^{N-1} \alpha_i D_{i+1}(t)$ is referred to as the dollar value of a basis point (DVBP) or price value of a basis point (PVBP). It is the amount by which the value of the swap changes for a one basis point change in interest rates.

Let's modify the notation a little. Denote $y_{n,N}(t)$ as the fixed rate in a swap quoted at time t, where the swap starts at $T_n \geq t$, and ends at T_N. Further, let's add some structure, and assume that $y_{n,N}(t)$ follows a geometric Brownian motion. Choosing $P_{n+1,N}$ as the numeraire:

$$dy_{n,N}(t) = \sigma_{n,N}(t) y_{n,N}(t) dW_t^{n+1,N}$$

where the superscript on the Brownian motion denotes the measure.

We now can price "swaptions," which are options to enter into a swap at a future date. A "payer" swaption with strike price K allows the option owner to enter into a swap where she pays a fixed rate K. A "receiver" swaption allows the option owner to enter into a swap where she receives the fixed rate K.

Note that two swaps with identical payment dates (and same floating reference rate) have identical floating cash flows. Assume the option expires at T. The holder of the payer swaption can therefore obtain the following payoff:

$$P_{n+1,N}(T) \max[0, y_{n,N}(T) - K]$$

This is because he can exercise the option, agree to pay K every payment date, and simultaneously enter into a receiver swaption at the rate $y_{n,N}(T)$, thereby locking in a cash flow of $y_{n,N}(T) - K$ every payment period.

Similarly, the owner of the receiver swaption obtains:

$$P_{n+1,N}(T) \max[0, K - y_{n,N}(T)]$$

Note that given the assumed dynamics of $y_{n,N}$, the problem is isomorphic to the Black-Scholes-Merton problem. For a payer swaption,

$$PS_{n,N,T}(t) = P_{n+1,N}(t) E^{n+1,N} \max\{\frac{P_{n+1,N}(T)[y_{n,N}(T) - K]}{P_{n+1,N}(T)}, 0\}$$

$$PS_{n,N,T} = P_{n+1,N}(t)E^{n+1,N} \max\{y_{n,N}(T) - K, 0\}$$

Using the same type of analysis used in the derivation of the BSM model, we get:

$$PS_{n,N,T}(t) = P_{n+1,N}(t)[y_{n,N}(t)N(d_1) - KN(d_2)$$

with:

$$d_1 = \frac{\ln(\frac{y_{n,N}(t)}{K}) + .5\Sigma^2_{n,N}}{\Sigma^2_{n,N}}$$

and

$$d_2 = d_1 - \Sigma^2_{n,N}$$

where

$$\Sigma^2_{n,N} = \int_t^T \sigma^2 n, N(s)ds$$

This is called the Black swaption formula.

Just like with the LMM, there are are multiple swap rates, corresponding to different maturities, starting dates, etc., but we can only make a single rate a martingale. However, there are products that have payoffs that may depend on multiple swap rates. Just as with the LMM, once we choose a measure that turns one particular rate into a martingale, we have to formulate a drift for other swap rates under this measure. This is an even more tedious exercise than for the LMM, so I will leave it to your imagination.