Interest Rate Derivatives

Dr. Graeme West
Financial Modelling Agency
graeme@finmod.co.za
Programme in Advanced Mathematics of Finance
School of Computational and Applied Mathematics
University of the Witwatersrand

November 28, 2004

Contents

Chapter 1

Introduction to Interest Rates

1.1 Introduction

In equity option pricing we make the assumption throughout that interest rates were previsible. This greatly simplifies discussions of hedging and replication, and allows the use of the Black-Scholes analysis and framework. The major implication of this assumption concerns the numeraire in the martingale pricing formula. Recall that the chosen numeraire is the money-market account (continuous investment/borrowing at deterministic interest rates). On application of the assumption, we could bring the numeraire $A(T)$ in the martingale pricing formula: $\frac{f(0)}{A(0)} = \mathbb{E}_0^Q \left[\frac{f(T)}{A(T)} \right]$[1] out of the expectation to evaluate the price of the derivative as: $f(0) = \frac{A(0)}{A(T)} \mathbb{E}_0^Q [f(T)]$.

This equation forms the basis of the Black-Scholes solution. The measure Q is the so-called Risk Neutral Measure, is unique, and implies that the stochastic processes for the underlying

[1]Some notation: $A(T)$ is the amount of money that 1 at time 0 has grown to by time T, by continual reinvestment at the short dated rate: $A(\cdot)$ is called the money market account. Usually this short dated rate is taken to be the overnight rate, so money is placed in the overnight account, and left there, compounding each day. Of course, by this, we mean the usual international meaning of the overnight account, in South Africa, a true overnight rate does not exist, as discussed in (West 2004*b*).

We do not know the value of $A(T)$ at time 0 because interest rates - in reality, and in this course - are stochastic. But, when we price equity derivatives, we assume that it is known.

$Z(t,T)$ is the discount factor for time T as observed at time t. It is known only when we reach t. Often, $t = 0$, in which case we might abbreviate $Z(0,T)$ to $Z(T)$.

$Z(0;t,T)$ is the forward discount factor from time t to time T. As proved in (West 2004*b*), it is equal to $\frac{Z(0,T)}{Z(0,t)}$.

$C(t,T)$ is the capitalisation factor for time T as observed at time t. The same comments apply. $C(0;t,T)$ is the forward capitalisation factor from time t to time T. As proved in (West 2004*b*), it is equal to $\frac{C(0,T)}{C(0,t)}$. Note that $Z(\cdot)$ and $C(\cdot)$ are inverse.

Only in the case that interest rates are previsible (in other words, not) will we have that $A(\cdot) = C(\cdot)$.

(default-free) securities and their derivatives have an expected drift rate of r (the prevailing risk-free rate), under Q. (In fact, under Q, ALL traded securities have an expected return of r!)

Despite our reservations, the model is extensively used, as seen for example in the extensive use of the Black-Scholes formula for short-dated options. This is a consequence of the underlying stochastic process (for the stock) being fairly remote from the stochastic process for the interest rate.

However, this cannot be the case when options and derivatives are written either on interest rates, or on securities whose values are dependent on interest rates (e.g. bond options, swaps, caps, floors etc.) In these cases, it is exactly the fluctuation of interest rates that the option buyer seeks to hedge; so an assumption of constancy of interest rates makes little sense. In particular, the bond option model we saw in (West 2004*b*) is very problematic.

As we saw in (West 2004*b*) vanilla type interest rate derivatives such as deposits (JIBAR deposits), FRAs and swaps do not require interest rate modelling, as they are priced using pure no arbitrage considerations. They can be replicated using instruments that pay certain cash flows in the future and consequently require no statistical modelling at all. They do require an accurate yield curve, though. Producing this is not always as simple a task as it would appear. We saw a (naively) simple method in (West 2004*b*). For more information on the subtle difficulties that can arise, see (Hagan & West 2004).

It will be assumed that you are perfectly familiar with all of the interest rate material dealt with in (West 2004*b*).

In order to price more complicated interest rate products that include optionality, we need to arrive at a statistical model of the evolution of the yield curve. The models we will consider in this course model all interest rates as dependent on a single rate, often termed the short rate. The evolution of the short rate then governs the evolution of all rates along the entire curve. However, this is quite a task, as a change in the yield curve is a complicated phenomenon, since it may undergo combinations of parallel shifts, slope changes and curvature changes.

These models are roughly divided into two categories: equilibrium models and no arbitrage models. In the equilibrium approach, if the model can be trusted to give a fundamentally correct, albeit necessarily simplified, description of economic reality then there will be discrepancies between model and market values. According to the model these discrepancies will reflect trading opportunities! Equilibrium Models attempt to describe the economy of interest rates as a whole. Clearly this approach is quiet abstract, and is not much used - although there are some equilibrium models that can be reformulated as no arbitrage models.

The pure no arbitrage approach seeks to represent the value of a complicated interest rate derivative in terms of vanilla instruments or cash instruments. The prices of these vanilla instruments will be taken as given and any model must recover their actual traded prices. No Arbitrage Models take the market prices of vanilla products as basic building blocks, and infer from them more complicated derivative prices. However, different models could give different prices of the more complicated instruments, even though they recover the same prices for the vanilla instruments!

Another complicating feature is that we cannot buy and sell interest rates. THE UNDERLYING OF THE MODEL IS NOT A TRADEABLE INSTRUMENT. The construction of a "replicating" portfolio requires more thought: it isn't just the delta and cash, as it is for equity: both of these factors don't perform as we would naively like. What we can buy and sell is bonds, whose prices are themselves derivative of interest rates.

Initially, we examine default-free securities and the term structure of interest rates. Once we have done this, we are in a position to model the movements of the yield curve.

1.2 Day count conventions

Denote the generic period between two payment dates as α parts of a year. The rules could be different for bonds and swaps, and even for the floating and fixed legs of the swaps. Moreover, the rules differ by jurisdiction. The relevant markets are

- The bond market: the market for the issuing of treasury bonds. Day count conventions are relevant for the accrual of interest and hence the conversions between clean and dirty price.[2]
- The money market: the market for FRAs and hence the market for the floating leg of swaps.
- The swap market: the market for the fixed leg of swaps.

The day count conventions in the various markets are as follows

	spot/value date	bond	money	fixed swap
RSA	t, $t+3$	Actual/365	Actual/365	Actual/365
USA	$t+2$	Actual/Actual	Actual/360	30/360[3]
UK	t	Actual/Actual	Actual/365	Actual/365
Euro	$t+2$	Actual/Actual	Actual/360	30/360
Japan	$t+2$	Actual/365	Actual/360	Actual/365
Canada	t	Actual/365	Actual/365	Actual/365
Australia	t	Actual/Actual	Actual/365	Actual/365

[2] In fact it is unlikely that there will ever be a need for modelling the US Treasury curve. This is because then the so-called TED spread needs to be measured i.e. the spread between the treasury and AAA curve, which is determined by the swaps. Secondly, the US Treasury is retiring a lot of its debt, so the treasury curve is very illiquid.

where Actual/Actual is rather a tricky day count convention; indeed (International Swaps and Derivatives Association 25 November 1998) clarifies this and points out that at the time there were three conflicting interpretations of this rule. The approach taken thus needs to be recorded when a deal is made. The approaches are a follows:

(i) ISDA Actual/actual (historical). Split the period of interest into the years in which it occurs. For each year, divide the number of actual days in the period by the number of days in that year. Day count is equal to the sum of these fractions.

For example, if the period is from 20 September 2003 to 20 March 2004, and we are now at 13 January 2004, then the period day count is $\frac{102}{365} + \frac{80}{366}$ and the time elapsed day count is $\frac{102}{365} + \frac{13}{366}$.

(ii) AFB Actual/Actual Euro. The numerator is the actual number of days, the denominator is either 365 or 366 depending on whether or not the period includes a 29 February.

In the above example, the period day count is $\frac{182}{366}$ and the elapsed day count is $\frac{115}{366}$.

(iii) ISMA Actual/Actual Bond. This is the actual number of days, divided by the product of the number of days in the period and the number of periods in the year.

In the above example, the period day count is $\frac{1}{2}$ (as there are two periods in a year) and the elapsed day count is $\frac{115}{182 \cdot 2}$.

This may be the most common convention, and is also known as ISMA Rule 251. Thus, for this calculation let the start date of the period of interest be t_1, the date of interest t, and the end date of the period of interest be t_2. Then on an Actual/Actual basis

$$\text{Accrued period} = \frac{t - t_1}{(t_2 - t_1) \cdot \text{round}(\frac{365}{t_2 - t_1}, 0)} \quad (1.1)$$

$$\text{Period remaining} = \frac{t_2 - t}{(t_2 - t_1) \cdot \text{round}(\frac{365}{t_2 - t_1}, 0)} \quad (1.2)$$

The EURIBOR quotation will be a representative rate for euro deposits based on quotations from a pan-European panel of banks. The euro-LIBOR quotation will be a representative rate for euro deposits based on quotations from a panel of 16 banks in the London market. The conventions for both of these are the 'Euro' conventions above.

1.3 Coupon Bonds

Coupon-bearing bond issuers pay regular fixed interest payments to the holder of the bond on specific dates (these are the coupons) as well as the par or face value at maturity. The

[3] Actual/Actual and Actual/360 also occur.

bond value at any time t must be the present value of both its face value and its coupons (the coupon rate is pre-fixed e.g. the r153 has a 13% annual coupon, i.e. 6.5% of par value is paid out to the holder every 6 months).

Suppose a bond pays amounts $c_1, c_2, \ldots, c_n$ at times $t_1, t_2, \ldots, t_n$. (c_i does not necessarily mean coupon here, for example, c_n could be the coupon and bullet.) The current bond price V is the sum of the present values of all payments i.e.

$$V = \sum_{i=1}^{n} c_i Z(0, t_i) \tag{1.3}$$

The act of regarding each coupon as a separate zero ensures arbitrage-free pricing. If this were not true, the coupon bond could be synthetically replicated using zero-coupon bonds. This replication is obviously not practically as straightforward as is made out here.

1.4 Yield-to-Maturity

Coupon-bearing bond prices are often quoted in terms of their yield-to-maturity. This is defined as an interest rate per annum that equates the present value of the bond's associated cash flows to the current market price. There exists a terminology confusion here because the yield-to-maturity of a coupon bond does not correspond directly to any value on the zero-coupon yield curve, in particular, it is not the value at the maturity (or duration, to be defined later) on the yield curve, even taking into account possible conversion between different NAC* quoting methods.

The yield-to-maturity is merely a convenient way of expressing the price of a coupon-bearing bond in terms of a single interest rate. In some rough sense, the YTM represents a weighted average of the interest rates along the current zero-coupon yield curve.

If payments are made annually, then y - the annually compounded yield-to-maturity - is implicitly defined by:

$$V = \sum_{t} \frac{c_t}{(1+y)^t} \tag{1.4}$$

where the payment c_t occurs at time t, and time is measured in years. If payments are made semi-annually (the most common) then y - the semi-annually compounded yield-to-maturity - is implicitly defined by:

$$V = \sum_{t} \frac{c_t}{(1+\frac{y}{2})^{2t}} \tag{1.5}$$

where again time is measured in years.

The market observables are the bond prices V, so the value for y must be calculated numerically. (1.4) and (1.5) cannot be inverted, but Newton's method comes to the rescue. The yield-to-maturity is the holding period return per annum on the coupon bond.

Using the effective yield-to-maturity as the interest rate offered by the bond, implicitly assumes that coupon reinvestment takes place at the semi-annual yield-to-maturity over the bond life. In a world of changing interest rates it is unlikely that this will happen: there is no guarantee that the ytm was an actual market rate at any reinvestment date, even if the yield curve does not change.

1.5 Term Structure of Default-Free Interest Rates

The term structure of interest rates is defined as the relationship between the yield-to-maturity on a zero coupon bond and the bond's maturity. If we are going to price derivatives which have been modelled in continuous-time off of the curve, it makes sense to use continuously-compounded rates from the outset.

Building a yield curve from existing data is a difficult task. The liquidity of the market plays a large part in determining whether the exercise can be done. For US Bond market it is only on-the-run bonds (last auctioned) which will give the most accurate indication of where yields are. In South Africa this task is doubly difficult and requires some econometric artistry. South Africa has a sophisticated and liquid swap market. This makes swap rates a better starting point for a yield curve model.

The results of bootstrapping will be near unique in liquid markets, but there may be significant variation in less liquid markets or markets with fewer inputs.

In so-called normal markets, yield curves are upwardly sloping, with longer term interest rates being higher than short term. A yield curve which is downward sloping is called inverted. A yield curve with one or more turning points is called mixed. Constructing a yield curve consists of solving (1.3) for the discount factors $Z(0,t)$ in a piecewise fashion starting with the shortest maturity instruments and progressing to the longer-dated coupon-bearing instruments. A mixed yield curve is shown in Figure 1.1. The discount factors are also shown.

The South African yield curve typically has two to four turning points. It is often stated that such mixed yield curves are signs of market illiquidity or instability. This is not the case. Supply and demand for the instruments that are used to bootstrap the curve may simply imply such shapes. However, many of the models that we see later in this course are driven by one factor - and intuitively this is clearly best suited to a normal or inverted curve (because, essentially, the model dictates that when the curve moves, it more or less moves in parallel). Thus, these models need to be analysed carefully for their suitability in the South African market: see (Svoboda 2002).

The shape of the graph for $Z(0,t)$ does not reflect the shape of the yield curve in any obvious way. The discount factor curve must - by no arbitrage - be monotonically decreasing whether the yield curve is normal, mixed or inverted. Nevertheless, many bootstrapping and interpolation algorithms for constructing yield curves miss this absolutely fundamental point.

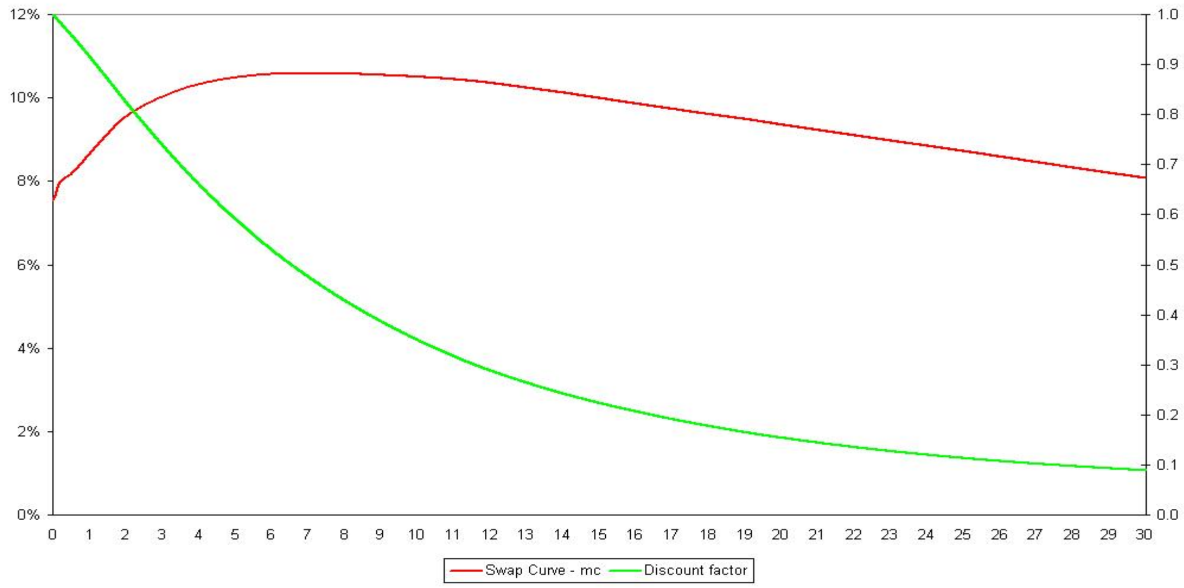

Figure 1.1: The yield curve and the associated discount factor curve

See (Hagan & West 2004).

1.6 The forward curve, the yield curve, and the par bond curve (the swap curve)

Recall that forward rates are very simple to express when using annual, continuously-compounded rates. The forward rate governing the period from T_1 to T_2, denoted $f(0; T_1, T_2)$ satisfies

$$\exp(r_1 T_1) \exp(f(0; T_1, T_2)(T_2 - T_1)) = \exp(r_2 T_2)$$

which implies that

$$f(0; T_1, T_2) = \frac{r_2 T_2 - r_1 T_1}{T_2 - T_1} \tag{1.6}$$

Now we see that the forward rates will lie above the yield curve when the yield curve is normal and below the yield curve when it is inverted.

A bond typically has a number of coupons and a final coupon and notional, or nominal, or bullet. The size of the bullet is a scaling factor, we can always arrange to think of it as 1. If such a bond is issued and has a price in the market of that same nominal value, then it is said to be a par bond.

We can ask ourselves: how big does the coupon of a newly issued bond need to be in order for it to be a par bond? And then, what will be its ytm? So, show that if a newly issued bond prices at par then the ytm is equal to the coupon (with the frequency of the ytm being the frequency of the coupon).

Suppose the bond pays ANNUAL coupons of R_n at times $t_1,\ t_2,\ \ldots,\ t_n$, with the bullet being at time t_n. The inter-coupon times are α_i, so the i^{th} payment is in fact $R_n\alpha_i$. For it to be a par bond, we must have

$$1 = Z(0, t_n) + \sum_{i=1}^{n} R_n \alpha_i Z(0, t_i) \tag{1.7}$$

We now have a function $t_n \to R_n$ which maps maturity dates to the requisite coupon size. This function is called the par bond curve.

But this should look familiar, and we didn't choose R_n as the notation for the annual coupon size by chance. You see that this is EXACTLY the curve of fair swap rates trading in the market.

Now the par coupon rates are a type of average of the zero rates along the life of the bond. Thus, if the yield curve is sloping up, the par curve is below the yield curve, if it is sloping down, the par curve is above.

1.7 The raw interpolation method for yield curve construction

This method corresponds to piecewise constant forward curves. This method is very stable, is trivial to implement, and is usually the starting point. One can often find mistakes in fancier methods by comparing the raw method with the more sophisticated method.

Suppose that the rates $r_1,\ r_2,\ \ldots,\ r_n$ are known at the ordered times $t_1,\ t_2,\ \ldots,\ t_n$. Any interpolation method of the yield curve function $r(t)$ will construct a continuous function $r(t)$ satisfying $r(t_i) = r_i$ for $i = 1,\ 2,\ \ldots,\ n$. By definition, raw interpolation is the method which has constant instantaneous forward rates on every interval $t_{i-1} < t < t_i$.

Note that the instantaneous forward rate $f(t)$ satisfies $f(t) = \frac{d}{dt}(r(t)t)$ (why?). Hence the raw interpolating function satisfies $r(t) = K_i + \frac{C_i}{t}$ for $t_{i-1} < t < t_i$ for some K_i, C_i. Now, since

$$\begin{aligned} r_{i-1}t_{i-1} &= K_i t_{i-1} + C_i \\ r_i t_i &= K_i t_i + C_i \end{aligned}$$

we have immediately (for example, by Cramer's rule)

$$K_i = \frac{r_i t_i - r_{i-1} t_{i-1}}{t_i - t_{i-1}}$$
$$C_i = \frac{(r_{i-1} - r_i) t_{i-1} t_i}{t_i - t_{i-1}}$$

Note now that the instantaneous forward rate on the interval $t_{i-1} < t < t_i$ is equal to the discrete forward rate for that interval. By substituting in, we have that the interpolation formula on that interval is

$$r(t) = \frac{t - t_{i-1}}{t_i - t_{i-1}} \frac{t_i}{t} r_i + \frac{t_i - t}{t_i - t_{i-1}} \frac{t_{i-1}}{t} r_{i-1} \tag{1.8}$$

1.8 Traditional Measures of Interest Rate Risk

Traditionally interest rate (bond) traders used two measures of interest rate risk.

1.8.1 (Macauley) Duration and Modified Duration

Macauley duration is a first-order measure of a bond's riskiness. The Macauley duration is a time (less than the maturity for coupon-bearing bonds) representing the horizon over which the holder of the bond is exposed. The lower the Macauley duration, the less risky the bond. As we will see, it is related to the linear coefficient in a Taylor series expansion for the change in the bond price with respect to a yield shift. Consequently, it is clear why the riskiness of a bond should decrease with decreasing Macauley duration.

Let us start by making the most modern possible definition of Macauley duration. Suppose a bond makes payments c_t at times t. Let the corresponding NACC risk free rates be r_t. Then $V = \sum_t c_t e^{-r_t t}$. Let us decompose the bond into a portfolio with value weights w_t, clearly $w_t = \frac{c_t e^{-r_t t}}{V}$. So $\sum_t w_t = 1$. Then define $D = \sum_t t w_t$. So Macauley duration is the average time at which I receive value in the bond.

Let us now calculate $\frac{dV}{d\overline{r}}$, the derivative w.r.t. parallel shifts in the yield curve. Clearly

$$\frac{dV}{d\overline{r}} = \sum_t -t c_t e^{-r_t t} = \sum_t -t w_t V = -DV \tag{1.9}$$

So we see that Macauley duration is related to the linear coefficient in the series expansion. Macauley duration is a measure of the bonds sensitivity to changes in yield. The minus sign reflects the inverse relationship between yield and bond price.

In fact, bonds are not priced NACC, for example, we are used to pricing bonds with a NACS yield to maturity. What happens then?

As in (1.5), suppose

$$V(y) = \sum_t \frac{c_t}{(1+\frac{y}{2})^{2t}}$$

This time

$$w_t = \frac{\frac{c_t}{1+\frac{y}{2}}^{2t}}{V}$$

and $D = \sum_t tw_t$ is as before. But now something changes. If y shifts by Δy, how does $V(y)$ change? Let $\Delta V(y)$ denote this change:

$$\begin{aligned}
\Delta V(y) &= V(y+\Delta y) - V(y) \\
&= \frac{dV}{dy}\Delta y + \cdots \\
&= -\frac{1}{1+\frac{y}{2}} \sum_t \frac{tc_t}{(1+\frac{y}{2})^{2t}} \Delta y + \cdots
\end{aligned}$$

so it is not true that $\frac{dV}{dy} = -DV$. So what we do is define a new quantity, called modified duration D_m, by:

$$D_m V = \frac{1}{1+\frac{y}{2}} \sum_t \frac{tc_t}{(1+\frac{y}{2})^{2t}} \tag{1.10}$$

so

$$\frac{dV}{dy} = -D_m V \tag{1.11}$$

Also, note that

$$D_m = \frac{D}{1+\frac{y}{2}} \tag{1.12}$$

These formulae have (hopefully) obvious extensions to NACn rates. As $n \to \infty$, the correction between duration and modified duration disappears, and they become the same thing for NACC rates.

Note for calculation purposes the rather convenient

$$D = \sum_t tw_t = \frac{\sum_t tV_t}{\sum_t V_t}$$

where $V_t = w_t V$ is the value of the t^{th} component.

Consider a 4.5 year bond with coupons of 10% and a ytm of 8%. The calculation of duration and modified duration proceeds as in Figure 1.2.

Cash Flows	t	V_t	tV_t
0.05	0.5	0.961538462	0.480769231
0.05	1	0.924556213	0.924556213
0.05	1.5	0.888996359	1.333494538
0.05	2	0.854804191	1.709608382
0.05	2.5	0.821927107	2.054817767
0.05	3	0.790314526	2.370943577
0.05	3.5	0.759917813	2.659712346
0.05	4	0.730690205	2.92276082
1.05	4.5	0.702586736	3.16164031
		7.435331611	17.61830318
	D	2.369538322	
	D_m	2.278402232	

Figure 1.2: duration and modifed duration calculation

1.8.2 Convexity

Of course there is a non-linear relationship between the bond price and the yield-to-maturity. For small shifts in yield, the first order duration calculation is a good measure of the sensitivity of the bond price. If Δy is large, the approximation is no longer accurate.

Using a Taylor expansion to second-order we can define convexity. Let us do it in the NACS setting as before.

$$\begin{aligned}\Delta V &= V\left(y+\Delta y\right) - V\left(y\right) \\ &= \frac{dV}{dy}\Delta y + \tfrac{1}{2}\frac{d^2V}{dy^2}\left(\Delta y\right)^2 + \dots \\ &= -\frac{1}{1+\frac{y}{2}}\sum_t \frac{tc_t}{(1+\frac{y}{2})^{2t}}\Delta y + \frac{1}{2}\frac{1}{(1+\frac{y}{2})^2}\sum_t \frac{t(t+\frac{1}{2})c_t}{(1+\frac{y}{2})^{2t}}(\Delta y)^2 + O\left(\Delta y^3\right)\end{aligned}$$

If convexity C is defined by

$$CV = \frac{1}{(1+\frac{y}{2})^2}\sum_t \frac{t(t+\frac{1}{2})c_t}{(1+\frac{y}{2})^{2t}} \tag{1.13}$$

we have

$$\frac{\Delta V}{V} = -D_m\Delta y + \tfrac{1}{2}C\Delta y^2 \tag{1.14}$$

1.8.3 Problems with these measures

Because the yield-to-maturity is, in some sense, a measure of the average rate of interest represented by the yield curve with respect to the bond in question, information about more subtle yield curve shifts is lost. It is possible for the change in the yield-to-maturity to mask the underlying zero coupon curve fluctuations.

By using duration measurements as interest rate risk measurement across different bonds, it is assumed that each bond experiences the same yield-to-maturity shift as the yield curve moves. This only occurs if the zero coupon curve moves in a parallel fashion i.e. the whole curve moves up or down. We have already discussed how, in general, this is not the case. Furthermore, for large yield shifts, the convexity correction will not capture the price shift, so by using this measure we are trapped in a small-movement, parallel-shift regime. This makes for naive hedging.

Chapter 2

Discrete-time Interest Rate Models

2.1 Introduction

The problem of interest rate derivatives can be approached broadly in one of two ways. We can either model the interest rate process or the bond price process. If we are modelling the interest rate process, then we must decide which interest rate(s) it is that we will model. In either case, there are consistency conditions that must hold because of the structure of the yield curve. In particular, implied forward rates should be positive, and we must be able to recover prices for underlying vanilla instruments (bonds, FRAs and swaps.) If we are to use no arbitrage models, then the underlying instrument prices are exogenous and are used to calibrate the model.

In a discrete-time framework, we can use a similar model to that used in the binomial model of stock price movements to model the bond prices and/or the interest rates. Modelling the price of the bond requires us to ensure that the process is arbitrage-free, consistent with the initial term-structure and obeys the boundary condition - that the price of a maturing bond is par, plus possibly the final coupon.

In this chapter we will model the evolution of the short-rate (defined as the continuously-compounded rate over Δt, the discrete time period for the lattice). **Modelling the short-rate in this fashion assumes that the yield curve evolution is governed by one underlying factor**. Substantial literature is devoted to multifactor models, where either more than one source of uncertainty is modelled (usually principle components or their proxies), or the evolution of the entire forward curve is modelled. One-factor models suffer from restricted degrees of freedom. In particular, they will be hampered by their inability to model yield curve slope and curvature changes and a tendency in the most simple models to violate positivity of interest rates. These shortcomings are sometimes surmountable.

2.2 Introducing Lattice Construction

Definition 1 *The yield curve* ***plus*** *the volatility associated with each* ***forward*** *spot rate along the curve, forms the* ***term structure*** *of interest rates.*

We use binomial trees for the Δt rate, the general notational convention of a tree will be as in Figure 2.1.

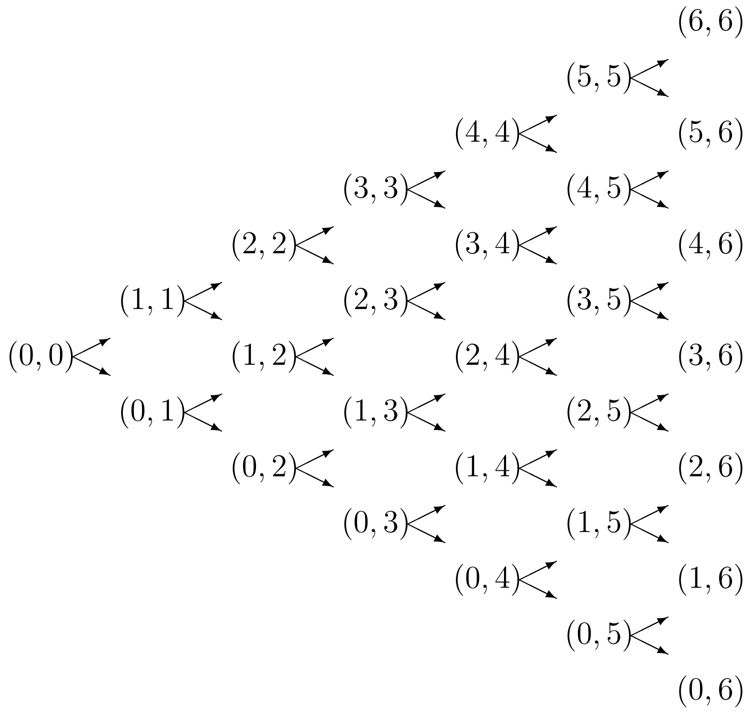

Figure 2.1: General node notation for binomial trees

Consider the following NACC term structure, with $\Delta t = 1$:

i	1	2	3	4	5	6	7	8
r_i	6.1982%	6.4030%	6.8721%	7.0193%	7.2000%	6.9000%	6.9000%	7.0000%
σ_i	1.7000%	1.5000%	1.1000%	1.0000%	1.0000%	1.0000%	1.1000%	

Note: The volatilities are not the bond price volatilities. They are also not the volatilities of the risk free rates that are from my yield curve bootstrap. They are the forward volatilities of the spot rate (over Δt) over the time period $[t_i, t_{i+1}]$, given the spot rates at time $t = 0$. In other words, they are the volatilities for the rate that will be observed at time t_i. **This must be understood carefully for estimation**.

Of course, there is no σ_0: the spot rate for the period from time $0 = t_0$ to time $t_1 = \Delta t$

is known. Similarly, given that the horizon in the above example is t_8, the volatility in the period $[t_8, t_9]$ is irrelevant.

Assume, using the binomial assumption, that the spot rate in one year's time takes one of two values: $r(1,1)$ or $r(0,1)$, corresponding are the discount factors $Z^t(1,1)$ and $Z^t(0,1)$.[1]

2.3 Ho-Lee model

The values we use here will be seen to be consistent with assuming that the spot rate process is **normally distributed**: this is the Ho and Lee model (Ho & Lee 1986).

Using martingale pricing theory on **bond prices** (analogously to stock prices) we know that there exists a unique probability measure $Q = \{\pi\}$ such that the bond price normalised by the money-market account follows a martingale. Furthermore, we know that this strictly implies that there are no arbitrage opportunities in the current bond market. Under the martingale measure, the current forward bond price is the expected price. Then

$$\begin{aligned} Z(0,2) &= \mathbb{E}_0^Q\left[Z(1,2)Z(0,1)\right] \\ &= Z(0,1)\mathbb{E}_0^Q\left[Z(1,2)\right] \\ &= Z(0,1)[\pi Z^t(1,1) + (1-\pi)Z^t(0,1)] \end{aligned} \tag{2.1}$$

Furthermore, the interest rates $r(1,1)$ and $r(0,1)$ must match the volatility term structure. Note that in general, if a variable r can take on two values, a and b, with $a > b$, and with probabilities π and $1-\pi$ respectively, then the variance of r is given by

$$\begin{aligned} \text{variance}\,(r) &= \mathbb{E}^Q\left[r^2\right] - \left(\mathbb{E}^Q\left[r\right]\right)^2 \\ &= \pi a^2 + (1-\pi)b^2 - (\pi a + (1-\pi)b)^2 \\ &= (a-b)^2\pi(1-\pi) \end{aligned} \tag{2.2}$$

So, it is very good idea to assume that $\pi = \frac{1}{2}$. (As usual, the model has three free parameters and is matching two characteristic equations, the mean and variance. So there is one degree of freedom, which we now use.) Then

$$\begin{aligned} Z(0,2) &= Z(0,1)\tfrac{1}{2}[e^{-a} + e^{-b}] \\ \text{stdev}\,(r) &= \tfrac{1}{2}(a-b) \end{aligned}$$

which is two equations in two unknowns, and solves easily:[2] $r(1,1) = 8.3223\%$, $r(1,0) = 4.9223\%$.

[1] The superscript t means 'tree', in order to avoid confusion with the original discount factors from the orginal yield curve. $Z^t(i,j)$ is the discount factor for the following Δt period given that we have evolved to node (i,j). So $Z^t(0,0) = Z(0,t_1)$ (!!).

[2] It doesn't solve easily because it is two in two; that fact means that there should be a solution. The ease of the solution comes from convenient facts of the exponential.

Now extend the lattice from one to two years, and assume that the volatility is time-dependent but not state-dependent (i.e. volatility varies from left to right and not from top to bottom). Furthermore, we will require that the tree recombines. This information is enough to calibrate the tree in closed inductive form.

But first, in order to explore a bit further, some information from heaven: guess that the values for the short-rate at time $t = 2$ are $r(2,2) = 10.8523\%$, $r(1,2) = 7.8583\%$ and $r(0,2) = 4.8583\%$. What do we need to do to check that this heavenly information is correct? First, and easily, the volatility structure:

$$\begin{aligned}\tfrac{1}{2}(r(2,2) - r(1,2)) &= \tfrac{1}{2}(0.108583 - 0.078583) = 0.015 \\ \tfrac{1}{2}(r(1,2) - r(0,2)) &= \tfrac{1}{2}(0.078583 - 0.048583) = 0.015\end{aligned}$$

Now,

$$\begin{aligned}Z(0,3) &= Z(0,1)\mathbb{E}_0^Q\left[Z(1,3)\right] \\ &= Z(0,1)\mathbb{E}_0^Q\left[Z(1,2)Z(2,3)\right] \\ &= Z(0,1)[\tfrac{1}{2}Z^t(1,1)\mathbb{E}_0^Q\left[Z(2,3)|(1,1)\right] + \tfrac{1}{2}Z^t(0,1)\mathbb{E}_0^Q\left[Z(2,3)|(0,1)\right]] \\ &= Z(0,1)[\tfrac{1}{2}Z^t(1,1)\tfrac{1}{2}[Z^t(2,2) + Z^t(1,2)] + \tfrac{1}{2}Z^t(0,1)\tfrac{1}{2}[Z^t(1,2) + Z^t(0,2)]\end{aligned}$$

Exercise 1 *We know from the previous steps the values of $Z(0,1)$, $Z^t(1,1)$ and $Z^t(0,1)$. Check that the heavenly values work in the above equation.*

This makes the model parameters consistent with all the market information provided, and hence the model is arbitrage free.

2.4 Formalising the Lattice Construction

We will have the binomial lattice of discount factors $Z^t(i,j)$, which are the discount factors over the following period if at time j we are in state i. Typically the periods are three months apart. Here i is a state (vertical) index and j is a time (horizontal) index.

Let $\sigma(j)$ be the annualised volatility of the Δt month forward rates for the period $[t_j, t_{j+1}]$. In the absence of implied information, the volatilities can only be calibrated via an historical analysis of the yield curve. This volatility will be calibrated under the Ho-Lee type assumption of normality of interest rates i.e. one takes differences of rates rather than log differences of rates in the calculation estimating standard deviations. Nevertheless, the square root of time rule for volatilities still applies, because we are assuming a Brownian evolution of the rate.

In the Ho-Lee model, there are equal probabilities for evolving to the two subsequent nodes in the tree, and furthermore by a normality assumption the difference in adjacent interest rates at time j for the period from then to time $j+1$ is independent of the state i, so

$$r(i+1,j) = r(i,j) + 2\sigma(j)\sqrt{\Delta t} \tag{2.3}$$

Hence

$$\begin{aligned} Z^t(i+1,j) &= Z^t(i,j)E(j) \; (1 \leq i \leq j-1) && (2.4) \\ E(j) &:= \exp\left[-2\sigma(j)\Delta t^{3/2}\right] && (2.5) \end{aligned}$$

To proceed, define new variables $\lambda(i,j)$. This variable is called the Arrow-Debreu price, and is the price of a security that pays off exactly one if we pass through the node (i,j), the payoff occurring at the moment of passing through. Thus

$$\begin{aligned} \lambda(0,0) &= 1 && (2.6) \\ \lambda(0,j) &= \tfrac{1}{2}\lambda(0,j-1)Z^t(0,j-1) && (2.7) \\ \lambda(i,j) &= \tfrac{1}{2}\lambda(i-1,j-1)Z^t(i-1,j-1) + \tfrac{1}{2}\lambda(i,j-1)Z^t(i,j-1) \; (0 < i < j) && (2.8) \\ \lambda(j,j) &= \tfrac{1}{2}\lambda(j-1,j-1)Z^t(j-1,j-1) && (2.9) \end{aligned}$$

Now, by no arbitrage we have

$$Z(0,(j+1)\Delta t) = \sum_{i=0}^{j} \lambda(i,j)Z^t(i,j) \tag{2.10}$$

and now recalling (2.4) we have

$$Z(0,(j+1)\Delta t) = \sum_{i=0}^{j} \lambda(i,j)Z^t(i,j) = Z^t(0,j)\sum_{i=0}^{j} \lambda(i,j)E(j)^i$$

$$Z^t(0,j) = \frac{Z(0,(j+1)\Delta t)}{\sum_{i=0}^{j} \lambda(i,j)E(j)^i} \tag{2.11}$$

To summarise, our recursion is as follows, at time step j:

(a) Calculate $\lambda(i,j)$ for $0 \leq i \leq j$ from (2.6), (2.7), (2.8) and (2.9).

(b) Calculate $Z^t(0,j)$ from (2.11).

(c) Calculate $Z^t(i,j)$ for $1 \leq i \leq j$ from (2.4).

Exercise 2 *Verify Figure 2.2 for the given data.*

$\lambda(i,j)$	0	1	2	3	4	5	6	7	8
0	1.0000	0.4699	0.2237	0.1065	0.0511	0.0245	0.0122	0.0060	0.0030
1		0.4699	0.4399	0.3099	0.1963	0.1170	0.0692	0.0396	0.0223
2			0.2162	0.3003	0.2830	0.2230	0.1636	0.1113	0.0724
3				0.0970	0.1813	0.2126	0.2061	0.1737	0.1342
4					0.0435	0.1013	0.1460	0.1627	0.1553
5						0.0193	0.0552	0.0914	0.1151
6							0.0087	0.0285	0.0533
7								0.0038	0.0141
8									0.0016

$Z^t(i,j)$	0	1	2	3	4	5	6	7
0	0.9399	0.9520	0.9526	0.9586	0.9606	0.9946	0.9891	0.9971
1		0.9201	0.9244	0.9377	0.9416	0.9749	0.9695	0.9754
2			0.8971	0.9173	0.9229	0.9556	0.9503	0.9541
3				0.8974	0.9046	0.9366	0.9315	0.9334
4					0.8867	0.9181	0.9130	0.9131
5						0.8999	0.8949	0.8932
6							0.8772	0.8738
7								0.8547

$r^t(i,j)$	0	1	2	3	4	5	6	7
0	6.198%	4.922%	4.858%	4.231%	4.023%	0.545%	1.100%	0.295%
1		8.322%	7.858%	6.431%	6.023%	2.545%	3.100%	2.495%
2			10.858%	8.631%	8.023%	4.545%	5.100%	4.695%
3				10.831%	10.023%	6.545%	7.100%	6.895%
4					12.023%	8.545%	9.100%	9.095%
5						10.545%	11.100%	11.295%
6							13.100%	13.495%
7								15.695%

Figure 2.2: The Ho-Lee trees associated with the given data

2.5 Lognormal Distribution

Discrete time models of interest rates where the forward rates are lognormally distributed are usually taken to be some variation of (Black, Derman & Toy 1990). The problem with the normality assumption in the previous section is that it is possible for interest rates to become negative (even if a mean reversion factor is introduced). Clearly, a lognormal assumption precludes this.

Assume now that the logarithm of the spot rate is normally distributed, and the volatility parameter pertains to the logarithms of the interest rates rather than the interest rates themselves. So now

$$\begin{aligned} \ln r(i+1,j) &= \ln r(i,j) + 2\sigma(j)\sqrt{\Delta t} \\ r(i+1,j) &= r(i,j)\exp(2\sigma(j)\sqrt{\Delta t}) \end{aligned} \tag{2.12}$$

Hence

$$Z^t(i+1,j) \quad = \quad Z^t(i,j)^{E(j)} \ (1 \le i \le j-1) \tag{2.13}$$

$$E(j) \quad := \quad \exp\left[2\sigma(j)\sqrt{\Delta t}\right] \tag{2.14}$$

which are calculated in advance. To proceed, again define new variables $\lambda(i,j)$ exactly as before: (2.6), (2.7), (2.8), (2.9) and (2.10) are unchanged. Thus

$$Z(0,(j+1)\Delta t) = \sum_{i=0}^{j} \lambda(i,j) Z^t(i,j) = \sum_{i=0}^{j} \lambda(i,j) Z^t(0,j)^{E(j)^i}$$

which does not have a closed form solution. Hence, we need to solve this for $Z^t(j,0)$ numerically. Let $x = Z^t(0,j)$. Using Newton's method, we solve

$$\begin{aligned} x_1 &= Z^t(0,j-1) \\ x_{n+1} &= x_n - \frac{\sum_{i=0}^{j} \lambda(i,j) x_n^{E(j)^i} - Z(0,(j+1)\Delta t)}{\sum_{i=0}^{j} \lambda(i,j) E(j)^i x_n^{E(j)^i - 1}} \end{aligned}$$

Convergence here is extremely rapid: to double precision in 3 or 4 iterations. The Newton function above is near linear (in a very wide range).

Note that the term structure of volatilities in this lognormal model will reflect higher values than those under the normality assumption. We see this via the following:

$$\begin{aligned} \ln[r+\Delta r] - \ln[r] &= \ln[r(1+\tfrac{\Delta r}{r})] - \ln[r] \\ &= \ln[1+\tfrac{\Delta r}{r}] \\ &\simeq \tfrac{\Delta r}{r} \end{aligned}$$[3]

for small Δ. This implies that,

$$\sigma_{\text{lognormal}} = \frac{\sigma_{\text{normal}}}{r}{}^4 \tag{2.15}$$

Exercise 3 *Assume the following initial data, with $\Delta t = \frac{1}{4}$:*

i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
r_i	*6.1982%*	*6.4030%*	*6.8721%*	*7.0193%*	*7.1000%*	*7.2021%*	*7.3120%*	*7.3000%*
σ_i	*20.0000%*	*18.0000%*	*17.0000%*	*17.0000%*	*17.0000%*	*17.0000%*	*17.0000%*	

Check Figure 2.3.

2.6 Options on Zero-Coupon Bonds

The procedure that we follow is the usual martingale pricing approach. We generate an (arbitrage-free) tree of discount factors and then use risk neutral valuation. Using this basis we can also price complicated interest rate derivatives. In general, the option will have maturity T and will be written on an instrument with maturity T_1 where $(T_1 > T)$.

As a simple example, consider such an option on a zero coupon bill. By arbitrage considerations, the value of the bill at maturity of the option T will be $Z(T, T_1)$. The value of such a bill today is $Z(0; T, T_1)$ ie. the forward discount value. But to value the option, we need to consider the volatility. Then the European call and put boundary conditions are:

$$c[Z(T,T_1), 0; K] = \begin{cases} Z(T,T_1) - K & \text{for} \quad Z(T,T_1) > K \\ 0 & \text{for} \quad Z(T,T_1) \leq K \end{cases} \tag{2.16}$$

$$p[Z(T,T_1), 0; K] = \begin{cases} 0 & \text{for} \quad Z(T,T_1) \geq K \\ K - Z(T,T_1) & \text{for} \quad Z(T,T_1) < K \end{cases} \tag{2.17}$$

respectively.

As an example we will consider a 18 month option on a six month zero-coupon bond i.e. after 18 months we decide whether or not to buy/sell a zero-coupon bond at the strike price; the zero coupon bond pays 1 after 2 years.

We use the Black-Derman-Toy model with the previous data. Given the tree of prices, there is very little to do. All we need to realise is that at time t_6 the bond has a value of

$$V_Z(i,6) = Z^t(i,6)\frac{1}{2}\left[Z^t(i,7) + Z^t(i+1,7)\right] \tag{2.18}$$

[3] Recall that the Taylor series of $\ln(1+x)$ is $x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots$.

[4] So, the classic rule of thumb in South Africa: to get from market volatilities to ballpark Ho-Lee volatilities: knock off a decimal place. Quoted volatilities will be lognormal because use of Black's model is the default: see Chapter 3.

$\lambda(i,j)$	0	1	2	3	4	5	6	7	8
0	1.0000	0.4923	0.2425	0.1193	0.0588	0.0290	0.0143	0.0071	0.0035
1		0.4923	0.4842	0.3568	0.2342	0.1443	0.0854	0.0492	0.0278
2			0.2417	0.3556	0.3496	0.2870	0.2121	0.1465	0.0966
3				0.1181	0.2319	0.2851	0.2806	0.2419	0.1912
4					0.0577	0.1416	0.2087	0.2396	0.2365
5						0.0281	0.0827	0.1422	0.1870
6							0.0136	0.0468	0.0923
7								0.0066	0.0260
8									0.0032

$Z^t(i,j)$	0	1	2	3	4	5	6	7
0	0.9846	0.9852	0.9839	0.9858	0.9871	0.9877	0.9883	0.9903
1		0.9820	0.9808	0.9832	0.9847	0.9854	0.9862	0.9885
2			0.9771	0.9801	0.9819	0.9827	0.9836	0.9864
3				0.9765	0.9785	0.9796	0.9806	0.9839
4					0.9746	0.9758	0.9771	0.9810
5						0.9714	0.9729	0.9775
6							0.9680	0.9734
7								0.9685

$r^t(i,j)$	0	1	2	3	4	5	6	7
0	6.198%	5.950%	6.473%	5.723%	5.213%	4.961%	4.696%	3.894%
1		7.267%	7.750%	6.783%	6.179%	5.880%	5.566%	4.616%
2			9.278%	8.041%	7.325%	6.970%	6.598%	5.471%
3				9.530%	8.682%	8.261%	7.820%	6.485%
4					10.291%	9.792%	9.270%	7.687%
5						11.606%	10.987%	9.111%
6							13.023%	10.799%
7								12.800%

Figure 2.3: The Black-Derman-Toy trees associated with the given data

Thus, the payoff of the option is

$$V^t(i,6) = \max(\eta(V_Z(i,6) - X, 0))$$

The value of the option is

$$V(0) = \sum_{i=0}^{6} \lambda(i,6)V^t(i,6)$$

Alternatively, we can induct backwards through the tree as usual. There are benefits to this approach, as it will enable us to derive hedge ratios, as we shall see shortly.

Verify that, for a call option on a six month zero coupon bond, strike 0.95, we get the tree of option values

0	1	2	3	4	5	6
0.0117	0.0146	0.0175	0.0203	0.0230	0.0255	0.0279
	0.0091	0.0121	0.0152	0.0183	0.0211	0.0238
		0.0065	0.0094	0.0127	0.0160	0.0191
			0.0039	0.0065	0.0099	0.0134
				0.0016	0.0033	0.0068
					0.0000	0.0000
						0.0000

and so a price of 0.0117.

2.6.1 Put-Call Parity

Put-Call parity for bond options is analogous to that of European options written on non-dividend paying stocks. Consider European put and call options written on zero-coupon bonds with $par = 1$. The option maturity is T and the bond maturity is T_1, where $T_1 > T$. Then

$$p[Z(t,T_1), T-t; K] + Z(t,T_1) = c[Z(t,T_1), T-t; K] + KZ(t,T) \tag{2.19}$$

Exercise 4 ***Prove this.***

2.7 Forwards on Zero-Coupon Bonds

What would the forward price be for the above zero coupon bond? Now

$$V^t(i,6) = V_Z(i,6)$$

so the forward value is

$$V(0) = 2^{-6} \sum_{i=0}^{6} \binom{6}{i} V^t(i,6)$$

2.8 Hedging Options

Replicating portfolios for options can be constructed using bonds or forwards. Consider the BDT example as previously. The hedging portfolio must contain two instruments (because of the binomial tree structure). We can choose any two of the zero coupon bonds that can be found in the market. Let us choose bonds maturing at time t_1 and t_2; we seeking hedge quanta n_1 and n_2.

$$\begin{aligned} V(0) &= Z(0,t_1)n_1 + Z(0,t_2)n_2 \\ V^t(0,1) &= n_1 + Z^t(0,1)n_2 \\ V^t(1,1) &= n_1 + Z^t(1,1)n_2 \end{aligned}$$

in other words

$$\begin{aligned} 0.0117 &= 0.9846n_1 + 0.9685n_2 \\ 0.0146 &= n_1 + 0.9852n_2 \\ 0.0091 &= n_1 + 0.9820n_2 \end{aligned}$$

which implies that $n_1 = -1.64$, $n_2 = 1.6793$, by solving the relevant 2×2 system:

$$\begin{bmatrix} 1 & 0.9852 \\ 1 & 0.9820 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \begin{bmatrix} 0.0146 \\ 0.0091 \end{bmatrix}$$

What can you say about the effectiveness of this hedge? How can it be modified to be more robust?

Hedging can also be done with forwards. The tree of forwards prices is:

0	1	2	3	4	5	6
0.9627	0.9658	0.9686	0.9712	0.9736	0.9758	0.9779
	0.9596	0.9629	0.9660	0.9688	0.9714	0.9738
		0.9562	0.9598	0.9632	0.9662	0.9691
			0.9526	0.9565	0.9601	0.9634
				0.9487	0.9529	0.9568
					0.9444	0.9490
						0.9398

Here, the last column is just the values of (2.18). We then induct to the left: the value in cell (i,j) is just the average of cells $(i,j+1)$ and $(i+1,j+1)$. As rates evolve, so the forward price 'converges' to the actual price after 18 months. Define the delta of the option with respect to the forwards contract in the usual fashion:

$$\Delta = \frac{V^t(0,1) - V^t(1,1)}{V^t_f(0,1) - V^t_f(1,1)} = \frac{0.0146 - 0.0091}{0.9658 - 0.9596} = 0.8763$$

Then replicate the option with a forward position and a cash account, with quanta m_1 and m_2 respectively:

$$\begin{aligned} V(0) &= m_1\, 0 + m_2\, 1 \\ V^t(0,1) &= m_1[V_f^t(0,1) - V_f^t(0,0)] + C(0,1)m_2 \\ V^t(1,1) &= m_1[V_f^t(1,1) - V_f^t(0,0)] + C(0,1)m_2 \end{aligned}$$

in other words

$$\begin{aligned} 0.0117 &= m_2 \\ 0.0146 &= [0.9658 - 0.9627]m_1 + 1.0156m_2 \\ 0.0091 &= [0.9596 - 0.9627]m_1 + 1.0156m_2 \end{aligned}$$

which implies that $m_1 = 0.8763$, $m_2 = 0.0117$. As usual, the hedge ratio is dynamic and we need to re-balance the hedge portfolio at $t = 1$, depending on up/down movement from $t = 0$. And so on.

Chapter 3

Black's Model

We now consider the Black Model for futures/forwards which is the market standard for quoting prices (via implied volatilities). (Black 1976) considered the problem of writing options on commodity futures and this was the first "natural" extension of the Black-Scholes model. Since the Black-Scholes analysis assumes constant (or deterministic) interest rates, and so forward interest rates are realised, it is difficult initially to see how this model applies to interest rate dependent derivatives. However, it can be extended to the case when the underlying variable is no longer a stock but is instead an interest rate. The market has a tendency to re-use the Black-Scholes analysis where and whenever it can, which is the dominant reason for the Black model being regarded as the "market model".

If f is a forward interest rate, it can be shown that it is consistent to assume that

- The discounting process associated with risk-neutral pricing is deterministic. By arbitrage considerations, the discount factors may be determined using the initial yield curve.
- The forward rates are stochastic and log-normally distributed.[1]

For a European option, this means we can used an appropriate version of a Black-Scholes formula. In particular, it can be used to price European bond options, swaptions, caps and floors. The value used for f is the forward value of x, where x can be anything from a bond price to the short rate, or even a short rate spread.

[1]This is subtly different to saying they are subject to Geometric Brownian motion. The assumption is merely that the terminal distribution is lognormal. In order to emphasise this point, σ is often referred to as the volatility measure, rather than just the volatility. σ is that number for which the logarithm of the terminal values is normally distributed with standard deviation $\sigma\sqrt{\tau}$.

3.1 European Bond Options

The clean (quoted) price for a bond is related to the all-in (dirty, cash) price via:

$$\mathbb{A} = \mathbb{C} + \mathbb{I}_A(t) \tag{3.1}$$

where the accrued interest $\mathbb{I}_A(t)$ is the accrued interest as of date t, and is non-zero between coupon dates. The forward price is a all-in price, not a clean price. The forward price is

$$\mathbb{F}_{\mathbb{A}} = C(0,T)\left[\mathbb{A} - \sum_i c_i Z(0,t_i)\right] \tag{3.2}$$

The option strike price $\mathbb{X}$ might be a clean or all-in strike; usually it is clean. If so, we change it to a all-in price by replacing $\mathbb{X}$ with $\mathbb{X}_{\mathbb{A}} = \mathbb{X}_{\mathbb{C}} + \mathbb{I}_{\mathbb{A}}(T)$.

The value of the bond option per unit of nominal is (Hull 2002, §22.2)

$$V_\eta = \eta Z(0,\tau)[\mathbb{F}_{\mathbb{A}} N(\eta d_1) - \mathbb{X}_{\mathbb{A}} N(\eta d_2)] \tag{3.3}$$

$$d_{1,2} = \frac{\ln \frac{\mathbb{F}_{\mathbb{A}}}{\mathbb{X}_{\mathbb{A}}} \pm \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}} \tag{3.4}$$

where $\eta = 1$ stands for a call, $\eta = -1$ for a put. Here σ is the volatility measure of the fair forward all in price.

Example 1 *Consider a* 10*m European call option on a* $1,000,000$ *bond with 9.75 years to maturity. Suppose the coupon is 10% NACS. The clean price is 935,000 and the clean strike price is 1,000,000. We have the following yield curve information:*

Term	*Rate*
3m	*9%*
9m	*9.5%*
10m	*10%*

The 10 month volatility on the bond price is 9%.

Firstly, $\mathbb{X}_{\mathbb{C}} = 1,000,000$ *so* $\mathbb{X}_{\mathbb{A}} = \mathbb{X}_{\mathbb{C}} + \mathbb{I}_{\mathbb{A}}(T) = 1,008,333.33$. *(There will be one month of accrued interest in 10 months time.)*

Secondly, $\mathbb{A} = \mathbb{C} + \mathbb{I}_{\mathbb{A}}(0) = 960,000$. *(There is currently three months of accrued interest.)*

Thirdly, $\mathbb{F}_{\mathbb{A}} = \left[960,000 - 50,000\left[e^{-\frac{3}{12}9\%} + e^{-\frac{9}{12}9.5\%}\right]\right] e^{\frac{10}{12}10\%} = 939,683.97.$

Hence $V_c = 7,968.60$ *and* $V_p = 71,129.06$.

The volatility above is a price volatility measure. However, quoted volatilities are often yield volatility measures. The relationship between the various volatilities of the bond is given via Ito's lemma as

$$\sigma_{\mathbb{A}} = -\frac{\sigma_y y \Delta}{\mathbb{A}} \tag{3.5}$$

$$\sigma_y = -\frac{\sigma_{\mathbb{A}} \mathbb{A}}{y \Delta} \tag{3.6}$$

$$\sigma_{\mathbb{A}} = \frac{\mathbb{C}}{\mathbb{A}} \sigma_{\mathbb{C}} \tag{3.7}$$

How do we see this? Note that

$$dy = \mu y \, dt + \sigma_y y \, dZ$$

is the geometric Brownian motion for the yield y. Now $\mathbb{A} = f(y)$ and so

$$d\mathbb{A} = \cdots \, dt + f'(y) \sigma_y y \, dZ := \cdots \, dt + \frac{\Delta \sigma_y y}{\mathbb{A}} \mathbb{A} \, dZ$$

But, also,

$$d\mathbb{A} = \nu \mathbb{A} \, dt + \sigma_{\mathbb{A}} \mathbb{A} \, dZ$$

and so the result follows - except for a missing minus sign. Why?

3.2 Swaptions

A swaption is an option to enter into a swap. We consider European swaptions. (Bermudan swaptions also exist.) Thus, at a specified time t_0, the holder of the option has the option to enter a swap which commences then (the first payment being one time period later, at t_1, and lasts until time t_n.

Of course, we have two possibilities

(a) a call, which gives the holder the right but not the obligation to receive floating, and pay a fixed rate r_X.

(b) a put, which gives the holder the right but not the obligation to receive a fixed rate r_X, and pay floating.

Let r_f be the fair (par) forward swap rate for the period from t_0 to t_n. The time of payments of the forward starting swap are $t_1, \ t_2, \ \ldots, \ t_n$, where $t_0, \ t_1, \ \ldots, \ t_n$ are successive observation days, for example, quarterly, calculated according to the relevant day count convention and modified following rules. As usual, let $t_i - t_{i-1} = \alpha_i$, measured in years, for $i = 1, \ 2, \ \ldots, \ n$.

Note that in general a swap (forward starting or starting immediately; in the later case $t_0 = 0$) with a fixed rate of R has the fixed leg payments worth

$$V_{\text{fix}} = R \sum_{i=1}^{n} \alpha_i Z(0, t_i) \tag{3.8}$$

while the floating payments are worth

$$V_{\text{float}} = Z(0, t_0) - Z(0, t_n) \tag{3.9}$$

Henece the fair forward swap rate, which equates the fixed and floating leg values, is given by

$$r_f = \frac{Z(0, t_0) - Z(0, t_n)}{\sum_{i=1}^{n} \alpha_i Z(0, t_i)} \tag{3.10}$$

Of course, these values are derived from the existing swap curve.[2] Also, let

$$B = \sum_{i=1}^{n} \alpha_i Z(0, t_i) \tag{3.11}$$

3.2.1 Valuation

The value of the swaption per unit of nominal is (Hull 2002, §22.4)

$$V_\eta = \eta B[r_f N(\eta d_1) - r_X N(\eta d_2)] \tag{3.12}$$

$$d_{1,2} = \frac{\ln \frac{r_f}{r_X} \pm \frac{1}{2}\sigma^2 t_0}{\sigma \sqrt{t_0}} \tag{3.13}$$

where $\eta = 1$ stands for a call, $\eta = -1$ for a put. Here σ is the volatility of the fair forward swap rate, and is typically an implied variable found in the market.

A more sophisticated approach would be to value swaptions by Monte Carlo within the LIBOR Market Model.

3.3 Interest Rate Caps/Floors

See (Hull 2002, §22.3). Assume that an investor is exposed to a floating rate: they lose money if the floating rate rises. The cap rate is a predetermined insurance rate (maximum value) on the floating rate (usually the prevailing 3-month IBOR rate). The relevant 3-month floating

[2] Thus, the fair forward swap rate is dependent upon the bootstrap and interpolation method associated with the construction of the yield curve. Nevertheless, empirically it is found that the choice of interpolation method will only affect the result to less than a basis point, and typically a lot less.

rates are set at the beginning of each period, and interest payments take place at the end of each period, much like a swap, for example. In each three month period the holder of the cap has protection; distinct from the other periods. Hence the cap translates as a series of call options on the 3-month rate (with settlement in arrears of the 3 months).

Let the cap rate (strike) be r_X.

Let the i^{th} reset period from t_{i-1} to t_i be of length α_i as usual, where day count conventions are observed. Suppose the IBOR rate for the period, observed at time t_{i-1}, is r_i. Then the payoff at time t_i is $\alpha_i \max(r_i - r_X, 0)$ per unit notional.

The entire cap is a sequence of caplets for $t_1,\ t_2,\ \ldots,\ t_n$ or, equivalently, the sum of the individual call options. We discount each caplet to the present time using the floating rate curve to determine the relevant discount factors. In an analogous fashion to writing a cap, we can write a floor, where the payout occurs when the floating rate drops below r_X, the floor rate.

3.3.1 Valuation

We value each caplet or floorlet separately off the yield curve using the implied forward rates at $t = 0$, for each time period t_i. Then,

$$V = \sum_{i=1}^{n} V_i \tag{3.14}$$

$$V_i = \eta Z(0, t_i)\alpha_i \left[f(0; t_{i-1}, t_i) N\left(\eta d_1^i\right) - r_X N\left(\eta d_2^i\right)\right] \tag{3.15}$$

$$d_{1,2}^i = \frac{\ln \frac{f(0;t_{i-1},t_i)}{r_X} \pm \frac{1}{2}\sigma_i^2 t_{i-1}}{\sigma_i \sqrt{t_{i-1}}} \tag{3.16}$$

where $\eta = 1$ stands for a cap(let), $\eta = -1$ for a floor(let), and where the floating rate (swap) curve is being used. $f(0; t_{i-1}, t_i)$ is the simple forward rate for the period from t_{i-1} to t_i, so of course

$$1 + \alpha_i f(0; t_{i-1}, t_i) = C(0; t_{i-1}, t_i) = \frac{C(0, t_i)}{C(0, t_{i-1})}$$

where $C(\cdot)$ is the capitalisation function.

Since each caplet is valued separately we expect a different volatility measure for each. Cap volatilities are usually quoted as flat volatilities where the same volatility is used for each caplet, which in some sense will be a weighted average of the individual caplet volatilities. Thus, we use $\sigma = \sigma_i$ for $i = 1,\ 2,\ \ldots,\ n$. Most traders work with independent volatilities for each caplet, though, and these are called forward-forward vols. For no known reason there exists a hump at about 1 year for the forward-forward vols (and, consequently, also for the flat vol, which can be seen as a cumulative average of the forward-forward vols)! This can be observed or backed out of cap prices: see Figure 3.1.

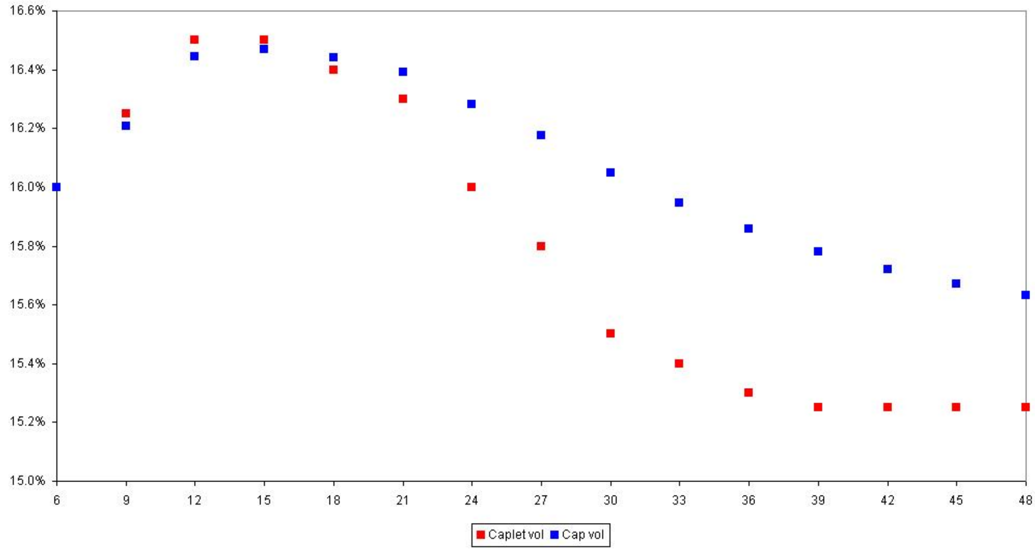

Figure 3.1: Caplet (fwd-fwd) and cap (flat fwd-fwd) volatility term structure

A caplet (a call on an interest rate) is actually a put on a floating-rate bond whose yield is the IBOR floating rate (a put option because of the inverse relationship between yield and bond price).

To see this, each caplet has a payoff in arrears of $\alpha_i \max(r_i - r_X, 0)$. The value of this in advance is

$$\begin{aligned} V(t_{i-1}) &= (1+\alpha_i r_i)^{-1}\alpha_i \max(r_i - r_X, 0) \\ &= \max\left(\frac{\alpha_i r_i - \alpha_i r_X}{1+\alpha_i r_i}, 0\right) \\ &= \max\left(1 - \frac{1+\alpha_i r_X}{1+\alpha_i r_i}, 0\right) \end{aligned}$$

which is a put, with a strike of 1, with the underlying being a zero paying $1+\alpha_i r_X$ at time t_i, with the termination of the option being t_{i-1}.

Likewise, floors are European put options on rates and European call options on the (underlying) floating-rate bond. For example, a floor would be purchased to ensure a minimum yield on a floating-rate bond held long.

3.3.2 Greeks

PV01

The PV01 is the p&l that arises from a one basis point parallel move in the yield curve. Thus

$$\text{PV01} = V(\underline{r}) - V(\underline{r} + 0.0001) \tag{3.17}$$

Thus, we have

$$\text{PV01} \approx \tfrac{1}{10000}\rho \tag{3.18}$$

where

$$\rho = \frac{\partial V}{\partial \underline{r}} \tag{3.19}$$

$$\begin{aligned}
\frac{\partial V_i}{\partial \underline{r}} &= -t_i V_i + Z(0,t_i)\alpha_i \eta \left[f(0;t_{i-1},t_i) N'(\eta d_1^i) \frac{\partial}{\partial \underline{r}} d_1^i + N(\eta d_1^i) - r_X N'(\eta d_2^i) \frac{\partial}{\partial \underline{r}} d_2^i \right] \\
&= -t_i V_i + Z(0,t_i)\alpha_i \eta N(\eta d_1^i) && (3.20) \\
\frac{\partial V}{\partial \underline{r}} &= \sum_{i=1}^{n} \frac{\partial V_i}{\partial \underline{r}} && (3.21)
\end{aligned}$$

Here $\sigma = \sigma_i$ for $i = 1,\ 2,\ \ldots,\ n$.

3.3.3 Delta hedging caps/floors with swaps

We require the 'with delta' value of a cap or floor. Often the client wants to do the deal 'with delta', which means that the 'linear' hedge comes with the option trade. Thus we quote delta based on hedging the cap/floor with a (forward starting) swap with the same dates, basis and frequency. Then

$$\Delta = \frac{\frac{\partial V(\text{cap})}{\partial \underline{r}}}{\frac{\partial V(\text{swap})}{\partial \underline{r}}} \tag{3.22}$$

The numerator is found in (3.21). For the denominator: the value of the swap fixed receiver, fixed rate R having been set, is

$$V(\text{swap}) = R \sum_{i=1}^{n} \alpha_i Z(0,t_i) - Z(0,t_0) + Z(0,t_n) \tag{3.23}$$

as seen in (3.8), (3.9). So the derivative is

$$\frac{\partial V(\text{swap})}{\partial \underline{r}} = -R \sum_{i=1}^{n} \alpha_i t_i Z(0,t_i) + t_0 Z(0,t_0) - t_n Z(0,t_n) \tag{3.24}$$

Vega

Vega is $\frac{\partial V}{\partial \sigma}$. It is the Greek w.r.t. the cap volatility; it is not a Greek with respect to the caplet volatilities. Here $\sigma = \sigma_i$ for $i = 1,\ 2,\ \ldots,\ n$.

$$\begin{aligned} \frac{\partial V_i}{\partial \sigma} &= Z(0,t_i)\alpha_i \left[f(0;t_{i-1},t_i)N'(d_1^i)\frac{\partial}{\partial \sigma}d_1^i - r_X N'(d_2^i)\frac{\partial}{\partial \sigma}d_2^i \right] \\ &= Z(0,t_i)\alpha_i f(0;t_{i-1},t_i)N'(d_1^i)\sqrt{t_{i-1}} \qquad (3.25) \\ \frac{\partial V}{\partial \sigma} &= \sum_{i=1}^{n} \frac{\partial V_i}{\partial \sigma} \qquad (3.26) \end{aligned}$$

Bucket (Caplet/Floorlet) Vega

For sensitivities to caplet volatilities, we would be looking at a bucket risk type of scenario. This would be $\frac{\partial V}{\partial \sigma_i}$, and this only makes sense of course in the case where we have individual forward-forward volatilities rather than just a flat cap volatility.

$$\frac{\partial V_i}{\partial \sigma_i} = Z(0,t_i)\alpha_i f(0;t_{i-1},t_i)N'(d_1^i)\sqrt{t_{i-1}}$$

where now it is the caplet volatility σ_i being used, not the cap volatility σ.

Chapter 4

One factor continuous-time interest rate models

This chapter is derived from components of (Wilmott 2000, Chapters 40, 41), (Svoboda 2002), (Svoboda 2003); amongst others.

The fundamental complicating factor in interest rate models is the non-traded nature of rates. Coupled to this is the non-linear, inverse relationship between bond prices and yields. Building and modelling a yield curve is difficult enough - see (Hagan & West 2004) - creating hedging/replicating positions off of it compounds the problem.

If a derivative is dependent on one or more interest rates, the rather neat consequences of the Black-Scholes model for equity derivatives, where the expected rate of return μ drops out of the pricing formula and is 'replaced' by risk-neutral valuation and a return of r, will almost certainly not be valid. What is more, in that Black-Scholes differential equation it was exactly that r which was a constant, it is now a variable.

We develop models of the short rate. The short rate itself is quite a theoretical concept: it is the yield on a bond which trades now and matures at time dt. In practice, a rate that truly exists, such as the overnight, one month or even three month rate, will be used as a surrogate for this rate. Although recall from (West 2004*b*) that we would probably not want to use the overnight rate in the South African market.

In actual fact the model of choice nowadays is arguably the LIBOR Market model (LMM, BGM) (Brace, Gątarek & Musiela 1997) where the underlying variable is the discrete three month rate. But in this chapter we only consider one factor models of the short rate.

We will now have a new variable, known as the market price of risk; knowing the market price of risk is equivalent to knowing the expected rate of return.

4.1 Derivatives Modelled on a Single Stochastic Variable

Assume that the short rate can be modelled as an Itô process of the following kind:

$$dr = \mu(r,t)dt + \sigma(r,t)dz \tag{4.1}$$

where dz has the usual properties. (4.1) is sufficiently general for a broad spectrum of possible models.

Recall that if $f = f(r,t)$ is a sufficiently well behaved function then Itô 's lemma tells us

$$\begin{aligned} df &= \frac{\partial f}{\partial r}\,dr + \left(\frac{\partial f}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 f}{\partial r^2}\right)dt & (4.2)\\ &= \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial r}\mu + \tfrac{1}{2}\sigma^2\frac{\partial^2 f}{\partial r^2}\right)dt + \frac{\partial f}{\partial r}\sigma\, dz & (4.3)\end{aligned}$$

Assume that two traded derivatives $V_1(r,t)$ and $V_2(r,t)$ are dependent only on $r(t)$ and t, and that their evolution can be described through Ito's Lemma. (In other words, V_1 and V_2 are sufficiently well behaved.) As a typical example, V_i could be a zero coupon bond with a specified maturity. Although then V_i is given by the discount factor which is obtained from the bootstrapped yield curve, we have that within our model V_i is (hopefully fairly accurately) given as a function of the short rate. Alternatively, V_i could be a 'true' derivative.

The underlying stochastic process r is non-traded, so the only possible way of creating a riskless portfolio is by using both V_1 and V_2. Let $\Pi(t) = V_1 - \Delta V_2$. Then,

$$\begin{aligned} d\Pi &= dV_1 - \Delta dV_2 \\ &= \left[\frac{\partial V_1}{\partial t}dt + \frac{\partial V_1}{\partial r}dr + \tfrac{1}{2}\sigma^2\frac{\partial^2 V_1}{\partial r^2}dt\right] - \Delta\left[\frac{\partial V_2}{\partial t}dt + \frac{\partial V_2}{\partial r}dr + \tfrac{1}{2}\sigma^2\frac{\partial^2 V_2}{\partial r^2}dt\right] & (4.4)\end{aligned}$$

by Itô 's Lemma. If

$$\Delta = \frac{\frac{\partial V_1}{\partial r}}{\frac{\partial V_2}{\partial r}}$$

then there is no risk, as all terms in dr cancel. Thus $d\Pi dt = r\Pi dt$ and so

$$r\left[V_1 - \frac{\frac{\partial V_1}{\partial r}}{\frac{\partial V_2}{\partial r}}V_2\right]dt = r\Pi dt = d\Pi dt = \left[\frac{\partial V_1}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 V_1}{\partial r^2} - \frac{\frac{\partial V_1}{\partial r}}{\frac{\partial V_2}{\partial r}}\left[\frac{\partial V_2}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 V_2}{\partial r^2}\right]\right]dt \tag{4.5}$$

which implies that

$$r\left[V_1 - \frac{\frac{\partial V_1}{\partial r}}{\frac{\partial V_2}{\partial r}}V_2\right] = \frac{\partial V_1}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 V_1}{\partial r^2} - \frac{\frac{\partial V_1}{\partial r}}{\frac{\partial V_2}{\partial r}}\left[\frac{\partial V_2}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 V_2}{\partial r^2}\right] \tag{4.6}$$

We can separate V_1 terms from V_2 terms, to give:

$$\frac{\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V_1}{\partial r^2} - rV_1}{\frac{\partial V_1}{\partial r}} = \frac{\frac{\partial V_2}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V_2}{\partial r^2} - rV_2}{\frac{\partial V_2}{\partial r}} \tag{4.7}$$

Note that, as in the usual separation of variables technique, the left-hand side is a function of V_1 and the right-hand side a function of V_2. Consequently, either side is a function only of $r(t)$ and t. Let,

$$\frac{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial r^2} - rV}{\frac{\partial V}{\partial r}} = a(r(t), t) = \lambda(r,t)\sigma(r,t) - \mu(r,t) \tag{4.8}$$

for some $\lambda(r,t)$, which for the moment is unspecified. It is always possible to represent $a(r,t)$ in the above form, using $\sigma(r,t)$ and $\mu(r,t)$ as basis functions, provided $\sigma(r,t) \neq 0$.

Then, (4.8) for all derivatives dependent only on $r(t)$ (including bonds) is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial r^2} + (\mu - \lambda\sigma)\frac{\partial V}{\partial r} - rV = 0 \tag{4.9}$$

This is the "Black-Scholes" equation for interest rate derivatives when the yield curve is modelled by a one-factor process of the type (4.1). Note that a vanilla bond is a derivative; options on bonds are derivatives of derivatives. The underlying here is the short rate. It is possible to apply (4.9) to bonds that have coupons by adding $K(r,t)$ to the left-hand side, where $K(r,t)\,dt$ represents the amount of coupon received in the period dt. (This may be continuous or discrete; in the latter case we will be using a finite difference scheme with specified jumps in value to the bond as it goes ex coupon.)

We require two "spatial" boundary conditions and one final boundary condition to fully specify the model. For example, we know that if $V(r,t,T)$ is the price of a bond with par value 1, $V(r,T,T) = 1$. Other examples would be the terminal intrinsic value of a vanilla option, etc.

Note that the Ito process (4.1) has been 'transformed' into the risk neutral form

$$dr = (\mu - \lambda\sigma)dt + \sigma dz \tag{4.10}$$

The risk neutral drift here is $\mu - \lambda\sigma$; the real world drift is of course μ.

This gives us access to non-traded underlying variables if we can estimate the market price of risk.

4.2 What is the market price of risk?

The function $\lambda(r,t)$ in (4.9) is called the market price of risk associated with the short rate $r(t)$. Note that

$$\begin{aligned} dV &= \left[\frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2\frac{\partial^2 V}{\partial r^2} + \mu\frac{\partial V}{\partial r}\right] dt + \sigma\frac{\partial V}{\partial r}dz \\ &= \left[\lambda\sigma\frac{\partial V}{\partial r} + rV\right] dt + \sigma\frac{\partial V}{\partial r}dz \\ \Rightarrow dV - rV dt &= \sigma\frac{\partial V}{\partial r}\left[\lambda dt + dz\right] \end{aligned}$$

Hence λ can be interpreted as the deterministic market price of risk.

4.3 One factor equilibrium models

These early equilibrium models are based on a mathematical model of the economy. They focus on describing and explaining the interest rate term structure. However, a fundamental problem with the equilibrium approach is that the models may not be arbitrage free, in other words, they fail to price even the vanilla inputs trading in the market. This is to be expected: in some sense these instruments are not necessarily inputs to the models. As a consequence it is unlikely that these models will be used.

4.3.1 GBM model

(Dothan 1978), (Rendelman & Bartter 1980)

This is simply

$$dr = \mu r dt + \sigma r dz \tag{4.11}$$

which is the same as the usual log-normal stock price model. However, an obvious problem that arises is that of mean-reversion: interest rates appear to be pulled back to some long-run average level over time. Clearly there is no mean-reversion in this model. In fact, the rate may rise or fall exponentially.

4.3.2 Vasicek model

(Vasicek 1977), (Jamshidian 1989)

$$dr = \mu(b - r)dt + \sigma dz \tag{4.12}$$

The Vasicek model attempts to model the mean-reversion property. The rate $r(t)$ "reverts" to a long-run mean b at a rate μ. The long-run rate b may even be stochastic, which will give us a two factor model. Unfortunately, the Vasicek model permits negative interest rates.

4.3.3 Cox-Ingersoll-Ross model

(Cox, Ingersoll & Ross 1985)

$$dr = \mu(b - r)dt + \sigma\sqrt{r}dz \tag{4.13}$$

The CIR model is an adaptation of the Vasicek model aimed at ensuring that rates are always positive. It can be shown that if $2\mu b \geq \sigma^2$ then the origin $r = 0$ cannot be reached i.e. the upward drift is sufficiently strong near the origin to exclude negative interest rates.

Although the intuition is equilibrium based, the CIR model can be calibrated from the initial term structure, which would then make it a no-arbitrage model.

4.4 Two factor equilibrium models

In (Brennan & Schwartz 1979), (Brennan & Schwartz 1982) the process for the short rate reverts to a long rate, which is turn follows a stochastic process. The long rate is chosen as the yield on a perpetual bond.

(Longstaff & Schwartz 1992) starts with a general equilibrium model of the economy and derives a term structure model where is there is stochastic volatility. This model is analytically quite tractible.

4.5 No-arbitrage Short Rate Models

In order to address the shortcomings of equilibrium models, a variety of no-arbitrage models have been postulated.

(4.9) is the bond pricing equation for an arbitrary model. In addition there will be the boundary conditions as already mentioned. In order to progress, we want to specify the risk neutral drift $\mu - \lambda\sigma$ and the volatility σ to arrive at tractable models i.e. models where (at the very least) the price of zero coupon bonds can be found analytically.[1] In other words, we will attempt to solve prices under the process (4.10).

[1]This will give us a check on the model, by comparing these results to our bootstrap, and in fact, might turn the whole problem inside-out, and become the calibration procedure.

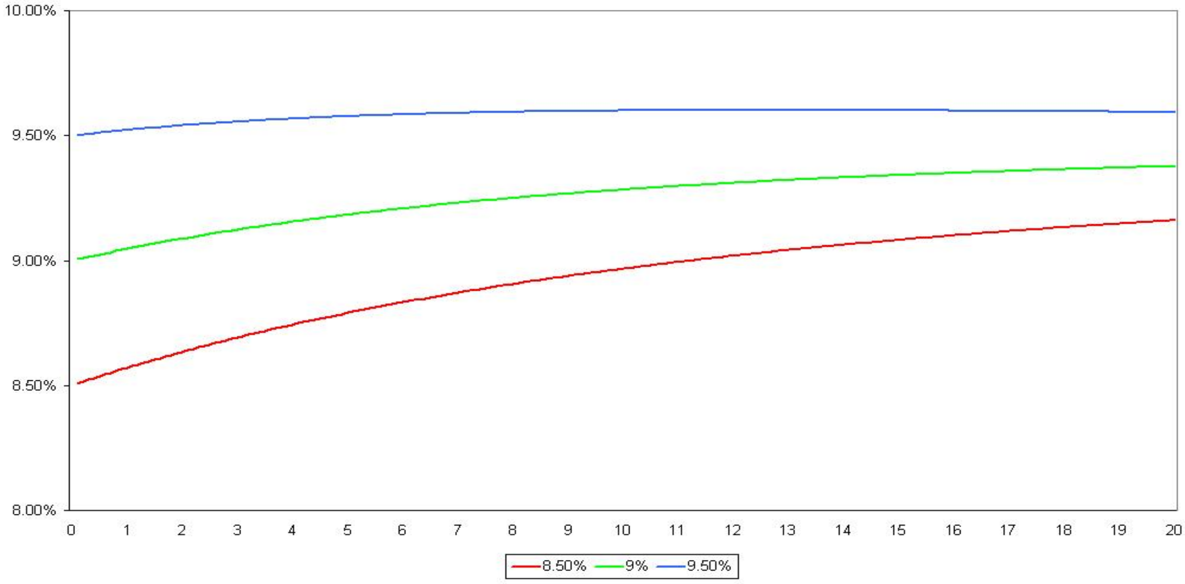

Figure 4.1: Term structures with three possible values of the short rate r.

Let the value of a zero-coupon bond be $Z(r,t,T)$ and let its dependency on the one-factor short rate be:

$$Z(r,t,T) = \exp(A(t,T) - rB(t,T)) \tag{4.14}$$

The class of models that result from this assumption are referred to as affine.

Note that this means that the entire yield curve is determined by the current value of the short rate r, as follows:

$$\exp(-\mathbb{r}(r,t,T)(T-t)) = Z(r,t,T) = \exp(A(t,T) - rB(t,T))$$

and so

$$\mathbb{r}(r,t,T) = -\frac{A(t,T) - rB(t,T)}{T-t} \tag{4.15}$$

So, as the short rate r varies, the entire yield curve $\mathbb{r}(r,t,T)$ varies.

Note that

$$\frac{\partial Z}{\partial t} = Z\left(\frac{\partial A}{\partial t} - r\frac{\partial B}{\partial t}\right)$$
$$\frac{\partial Z}{\partial r} = -B(t,T)Z$$
$$\frac{\partial^2 Z}{\partial r^2} = B(t,T)^2 Z$$

and so, substituting (4.14) into (4.9), and dividing by Z, we get

$$\frac{\partial A}{\partial t} - r\frac{\partial B}{\partial t} + \tfrac{1}{2}\sigma^2 B^2 - (\mu - \lambda\sigma)B - r = 0. \tag{4.16}$$

If we differentiate this equation twice with respect to r, and divide through by B, we arrive at

$$\tfrac{1}{2}\frac{\partial^2}{\partial r^2}\sigma^2 B - \frac{\partial^2}{\partial r^2}(\mu - \lambda\sigma) = 0.$$

In this the expiry date T is arbitrary, so the coefficients must be separately zero:

$$\frac{\partial^2}{\partial r^2}\sigma^2 = 0$$
$$\frac{\partial^2}{\partial r^2}(\mu - \lambda\sigma) = 0$$

and so we can find functions $\alpha(t)$, $\beta(t)$, $\eta(t)$ and $\gamma(t)$ which are function of t alone (and not of r) such that

$$\sigma^2(r,t) = \alpha(t)r + \beta(t) \tag{4.17}$$
$$\mu(r,t) - \lambda(r,t)\sigma(r,t) = \eta(t) - \gamma(t)r \tag{4.18}$$

Substituting (4.17) and (4.18) into (4.16) we have

$$\frac{\partial A}{\partial t} - r\frac{\partial B}{\partial t} + \tfrac{1}{2}(\alpha(t)r + \beta(t))B^2 - (\eta(t) - \gamma(t)r)B - r = 0$$

and so

$$\frac{\partial A}{\partial t} = \eta(t)B - \tfrac{1}{2}\beta(t)B^2 \tag{4.19}$$
$$\frac{\partial B}{\partial t} = -1 + \gamma(t)B + \tfrac{1}{2}\alpha(t)B^2 \tag{4.20}$$

(4.20) is what is known as a Ricatti equation, it is an equation in B which does not involve A. Having solved (4.20) we then insert this solution into (4.19) and do straightforward integration to obtain A. In general the system will be solved numerically; however, sometimes it is possible to find closed form solutions, and we will focus on these cases.

The boundary conditions must be:

$$A(T,T) = 0$$
$$B(T,T) = 0$$

Our risk neutral single factor stochastic process has become

$$dr = (\eta(t) - \gamma(t)r)dt + \sqrt{\alpha(t)r + \beta(t)}dz \tag{4.21}$$

Let us start with the simplest case: all of α, β, η, γ are constants. Then

$$\frac{dB}{dt} = -1 + \gamma B + \tfrac{1}{2}\alpha B^2$$

and so (after some unpleasant calculus)

$$\begin{aligned} B(t,T) &= \frac{2(\exp(\psi_1(T-t)) - 1)}{(\gamma + \psi_1)(\exp(\psi_1(T-t)) - 1) + 2\psi_1} \\ \psi_1 &= \sqrt{\gamma^2 + 2\alpha} \end{aligned}$$

Now, we can write

$$\frac{dA}{dB} = \frac{\eta B - \frac{1}{2}\beta B^2}{\frac{1}{2}\alpha B^2 + \gamma B - 1}$$

and so

$$\begin{aligned} \tfrac{1}{2}\alpha A &= a\psi_2 \ln(a - B) + (\psi_2 + \tfrac{1}{2}\beta) b \ln \frac{B + b}{b} - \tfrac{1}{2}\beta B - a\psi_2 \ln a \\ \psi_2 &= \frac{\eta - \frac{1}{2}a\beta}{a + b} \\ b, a &= \frac{\pm\gamma + \sqrt{\gamma^2 + 2\alpha}}{\alpha} \end{aligned}$$

Note that here we have incorporated the final conditions.

Obviously A and B are functions of $\tau = T - t$, and not functions of t and T separately.

We now consider varying models of (4.17) and (4.18) which give rise to tractable solutions to (4.21) (with a relevant boundary value).

4.5.1 The Vasicek Model

The Vasicek model has all four parameters constant, with $\alpha = 0$, and (of necessity) $\beta > 0$. Thus the volatility is time and state independent and equal to $\sqrt{\beta}$. So

$$dr = (\eta - \gamma r)dt + \sqrt{\beta} dz \tag{4.22}$$

The short rate mean reverts to $\frac{\eta}{\gamma}$ at a 'speed' of γ. This model is very tractable, and there are explicit solutions for a number of derivatives based on it.

Now (4.19) and (4.20) become

$$\begin{aligned} \frac{dA}{dt} &= \eta B - \tfrac{1}{2}\beta B^2 \\ \frac{dB}{dt} &= -1 + \gamma B \end{aligned}$$

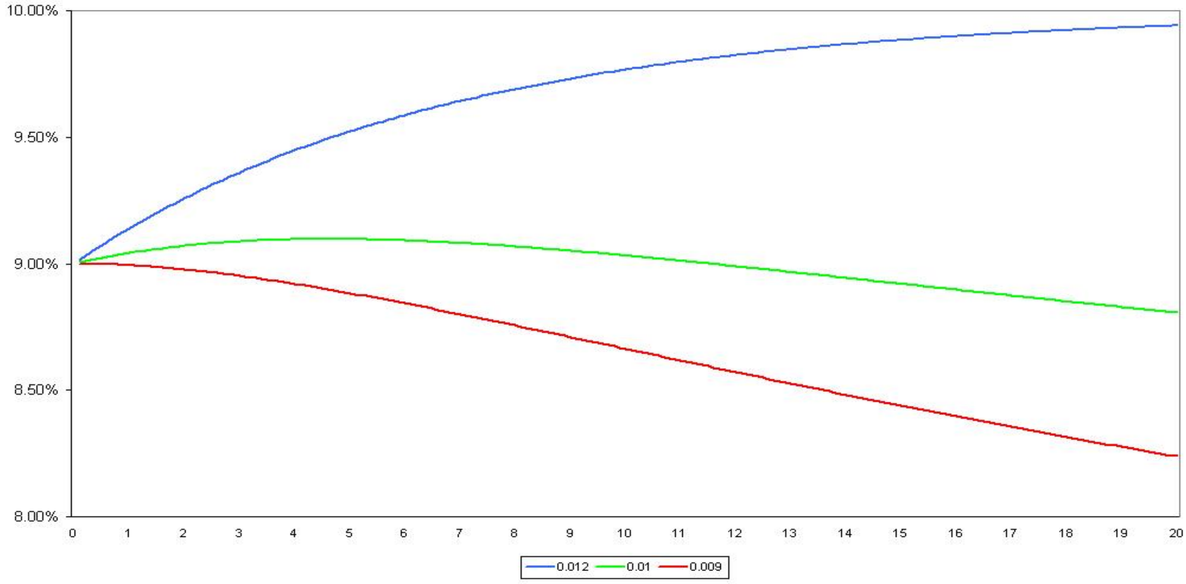

Figure 4.2: Term structures under the Vasicek model: $\beta = 0.04\%$, $\gamma = 10\%$, short rate $= 9\%$ and three possible values of η.

(there is no dependency on r in the parameters - t is the only variable in play, and partial differentiation is the same as total differentiation) and so (again, after some work)

$$B(t,T) = \frac{1}{\gamma}(1 - \exp(-\gamma(T-t)))$$
$$A(t,T) = \frac{(\frac{1}{2}\beta - \eta\gamma)((T-t) - B)}{\gamma^2} - \frac{\beta B^2}{4\gamma}$$

It is possible to generate increasing, decreasing or humped curves under this model. See Figure 4.2.

Note that it is possible to obtain negative rates under this model.

Under the Vasicek model we can value European options on default-free bonds exactly (Jamshidian 1989). Consider a call option on an T_2-maturity discount bond with exercise price K and expiration $T_1 < T_2$. Using risk-neutral valuation, the vanilla option price is:

$$V_{\pm}[r,t,T_1,T_2;K] = Z(r,t,T_1)\mathbb{E}_t^Q\left[\max(0, \pm(Z-K))\right]$$

where Z is the zero coupon price for expiry T_2 as observed at time T_1. It has current forward value

$$f(Z) = \frac{Z(r,t,T_2)}{Z(r,t,T_1)}.$$

It turns out that we obtain a 'Black-like' option pricing formula:

$$c[r,t,T_1,T_2;K] = Z(r,t,T_2)N(d_+) - KZ(r,t,T_1)N(d_-) \tag{4.23}$$

$$p[r,t,T_1,T_2;K] = KZ(r,t,T_1)N(-d_-) - Z(r,t,T_2)N(-d_+) \tag{4.24}$$

$$\sigma_p^2 = \frac{\beta(1-\exp(-2\gamma(T_1-t)))(1-\exp(-\gamma(T_2-T_1)))^2}{2\gamma^3} \tag{4.25}$$

$$d_\pm = \frac{\ln\frac{f(Z)}{K} \pm \frac{1}{2}\sigma_p^2}{\sigma_p} \tag{4.26}$$

Remarkably, the valuation of options on coupon-bearing bonds is also possible. Although the model was originally posited with the underlying stochastic process being Vasicek's model, it can be applied to any derivative where the value of zero-coupon bonds is dependent on the short rate only. The intuition is that a set of mini-options on each of the coupons and the bond can be constructed which will have the same value as the total option. Because the movements of the entire yield curve are perfectly correlated there is currently no additional value to having this decomposition. Thus, the value of the option on the coupon bearing bond can be decomposed into a portfolio of options on zero coupon bonds.

This procedure only works if the option is European.

Consider a European call with exercise price K and maturity T, on a coupon-bearing bond. Suppose that the bond provides cashflows $c_1,\ c_2,\ \ldots,\ c_n$ at times $t_1,\ t_2,\ \ldots,\ t_n$ after the maturity T of the option. The payoff of the option at time T is

$$\max\left[0, \sum_{i=1}^{n} c_i Z(r,T,t_i) - K\right] \tag{4.27}$$

where r is the short rate as observed at time T; of course unknown at time t. Let r^* be the value of the short rate at time T which causes the coupon-bearing bond price to equal the strike price. r^* is the solution of

$$K = \sum_{i=1}^{n} c_i Z(r^*,T,t_i) \tag{4.28}$$

We calculate r^* now using Newton's method. (Just using goalseek in excel will suffice for simple problems.) Since zero prices are decreasing functions of r, it follows that the option expires in the money for all $r < r^*$, and out of the money for all $r > r^*$. Then

$$\begin{aligned}
\text{payoff} &= \max\left[0, \sum_{i=1}^{n} c_i Z(r, T, t_i) - K\right] \\
&= \max\left[0, \sum_{i=1}^{n} c_i Z(r, T, t_i) - \sum_{i=1}^{n} c_i Z(r^*, T, t_i)\right] \\
&= \sum_{i=1}^{n} c_i \max[0, Z(r, T, t_i) - Z(r^*, T, t_i)]
\end{aligned}$$

Thus the call on the coupon-bearing bond is reduced to the sum of separate calls on the underlying zeros.

Example 2 *Suppose that* $\sigma = \sqrt{\beta} = 0.02$, $\gamma = 0.1$, $\eta = 0.01$ *and the short rate be 10%. Consider a 3-year European put option on a 5-year bond with a coupon of 10% NACS, strike price* $K = 98$*, par 100.*

At the end of 3 years there will be four cashflows remaining, namely

time	*amount*	*discount factor*
3.5	*5*	$\exp(A(3;3.50) - r(3)B(3;3.5))$
4	*5*	$\exp(A(3;4) - r(3)B(3;4))$
4.6	*5*	$\exp(A(3;4.5) - r(3)B(3;4.5))$
5	*105*	$\exp(A(3;5) - r(3)B(3;5))$

where the functional form of $A(t,T)$ *and* $B(t,T)$ *are known - the only unknown here is* $r(3)$*, the value of the short rate that will actually be observed at time 3.*

Solving (4.28) for $r^* = r(3)$*, where* $K = 98$ *gives* $r^* = 0.1095222$*.*

time	*A*	*B*	*Z*	*Cash*	*Value*
0	*0*	*0*	*1.000000*		
0.5	*-0.001*	*0.4877*	*0.946830*	*5*	*4.734149*
1	*-0.005*	*0.9516*	*0.896731*	*5*	*4.483653*
1.5	*-0.011*	*1.3929*	*0.849538*	*5*	*4.247691*
2	*-0.018*	*1.8127*	*0.805091*	*105*	*84.534506*
					98

The option is the sum of four options:

time	*Cash*	*Strike*
3.5	*5*	*4.734149*
4	*5*	*4.483653*
4.5	*5*	*4.247691*
5	*105*	*84.53451*

We value each as follows:

time	*A*	*B*	*Z*	*Cash*	*Strike*	σ_p	d_1	d_2	*V*
3.5	*-0.052*	*2.9531*	*0.706252*	*5*	*4.734149*	*1.465%*	*0.376364*	*0.361713*	*0.012449*
4	*-0.067*	*3.2968*	*0.672465*	*5*	*4.483653*	*2.859%*	*0.3903*	*0.361713*	*0.022830*
4.5	*-0.083*	*3.6237*	*0.640436*	*5*	*4.247691*	*4.184%*	*0.403556*	*0.361713*	*0.031429*
5	*-0.101*	*3.9347*	*0.610074*	*105*	*84.53451*	*5.445%*	*0.416166*	*0.361713*	*0.808417*
									0.875125

4.5.2 The Cox-Ingersoll-Ross Model

The (Cox et al. 1985) model has all four parameters constant, but this time $\beta = 0$ and $\alpha > 0$. In this case (4.9) becomes:

$$dr = (\eta - \gamma r)dt + \sqrt{\alpha r}dz \tag{4.29}$$

The spot rate again has a mean reversion level of $\frac{\eta}{\gamma}$, with reversion occurring at a rate of γ but, unlike the Vasicek model, the steady-state distribution is not a special case of the normal distribution (since the volatility is now rate dependent). The CIR model does not permit negative interest rates, provided the technical condition $\eta > \frac{1}{2}\alpha$ holds. CIR give explicit solutions for some interest rate derivatives but the solutions are complicated and sometimes involve integrals that do not have exact solutions (and have to be estimated numerically).

For the pricing of discount bonds under the affine model (4.14) the bond pricing equations (4.20) and (4.19) reduce to:

$$\frac{dB}{dt} = \tfrac{1}{2}\alpha B^2 + \gamma B - 1$$
$$\frac{dA}{dt} = \eta B$$

which has the solution:

$$Z(t,T) = \frac{2(\exp(\psi_1(T-t)) - 1)}{(\gamma + \psi_1)(\exp(\psi_1(T-t)) - 1) + 2\psi_1}$$
$$\psi_1 = \sqrt{\gamma^2 + 2\alpha}$$
$$A(t,T) = \frac{2\eta}{\alpha}\ln\frac{2\psi_1\exp((\gamma+\psi_1)(T-t)/2)}{(\psi_1+\gamma)(\exp(\psi_1(T-t)) - 1) + 2\psi_1}$$

For specific values of α, γ and η we can generate yield curves of various shapes. Again, A and B are functions only of $T - t$. We can use a similar approach to Jamshidian's to value European style options on coupon-bearing bonds.

Consider a call option on an T_2-maturity discount bond with exercise price K and expiration $T_1 < T_2$. Using risk-neutral valuation, the vanilla option price is:

$$V_{\pm}[r,t,T_1,T_2;K] = Z(r,t;T_1)\mathbb{E}_t^Q\left[\max(0, \pm(Z-K))\right]$$

where

$$f(Z) = \frac{Z(r,t,T_2)}{Z(r,t,T_1)}$$

Then,

$$V_c[r,t,T_1,T_2;K] = Z(r,t,T_1)\left[f(Z)\chi^2(h_1) - K\chi^2(h_2)\right]$$

where χ^2 is the non-central chi-squared distribution. The parameters h_1 and h_2 have a complicated dependence on α, γ and η.

4.5.3 The Continuous time version of the Ho-Lee model

(Rebonato 1998) presents a simple analysis by which the continuous time equivalent of any discrete time model, modelled within a binomial lattice, may be found.

Given that we are in state $(i, j-1)$, we can only move in the next time step to state (i,j) or state $(i+1, j)$. Given the assumption that the short term interest rate follows a Gaussian process we have

$$r(i+1,j) = r(i,j) + 2\sigma(j)\sqrt{\Delta t} \tag{4.30}$$

where $\sigma(j)$ is the volatility of the one period rate that is earned in the period $[j\,\Delta t, (j+1)\,\Delta t]$.

Let $r_m(i,j)$ be the expected interest rate at time j given where we are aat time $j-1$, hence:

$$\begin{aligned} r_m(i,j) &= \tfrac{1}{2}\left[r(i+1,j) + r(i,j)\right] \\ \Rightarrow \quad r(i+1,j) &= r_m(i,j) + \sigma(j)\sqrt{\Delta t} \\ r(i,j) &= r_m(i,j) - \sigma(j)\sqrt{\Delta t} \end{aligned}$$

and so, in continuous time, we may write:

$$r(t) = \mu(r,t) + \sigma(t)z(t)$$

If we assume that the volatility is also not time dependent that we may put this in our favoured affine form: we have $\beta > 0$ constant, $\alpha = 0 = \gamma$, but η a function of time:

$$dr = \eta(t)dt + \sqrt{\beta}dz \tag{4.31}$$

The above equation is the continuous-time version of the discrete time model of (Ho & Lee 1986). There is no mean reversion in the model which means that the process for r has unbounded variance. The function $\eta(t)$ can be determined analytically and calibrated from the initial term structure of interest rates, as follows:

(4.19) and (4.20) have become

$$\frac{\partial A}{\partial t} = \eta(t)B - \tfrac{1}{2}\beta B^2 \tag{4.32}$$

$$\frac{\partial B}{\partial t} = -1 \tag{4.33}$$

and so

$$Z(t,T) = T - t \tag{4.34}$$

$$A(t,T) = -\int_t^T \eta(s)(T-s)\,ds + \tfrac{1}{6}\beta(T-t)^3 \tag{4.35}$$

We can calibrate $\eta(\cdot)$ so that the observed yield curve (found from our yield curve bootstrap) fits exactly. The best way to work with this is to think of t as today, so $t = 0$ and that the variable in play is T. Thus

$$-T\mathbb{r}(r,0,T) = A(0,T) - rZ(0,T)$$

We use (4.34) and (4.35) in this, and then differentiate twice w.r.t. T. Here we remember the standard formula for differentiating under an integral sign (Abramowitz & Stegun 1974):

$$\frac{d}{dT}\int_{a(T)}^{b(T)} f(s,T)\,ds = f(b(T),T)\frac{db(T)}{dT} - f(a(T),T)\frac{da(T)}{dT} + \int_{a(T)}^{b(T)} \frac{d}{dT}f(s,T)\,ds \tag{4.36}$$

The result is

$$\eta(T) = \frac{\partial^2}{\partial T^2}[\mathbb{r}(r,0,T)T] + \beta T \tag{4.37}$$

We will see more of this type of program in §4.6. Of course, here the interpolated curve needs to be twice differentiable. Most cubic spline interpolation schemes satisfy this property.

Under the Ho-Lee model we can value European options on default-free bonds in exactly the same way as before. Consider a call option on an T_2-maturity discount bond with exercise price K and expiration $T_1 < T_2$. Using risk-neutral valuation, the vanilla option price is:

$$V_\pm[r,t,T_1,T_2;K] = Z(r,t,T_1)\mathbb{E}_t^Q\left[\max(0,\pm(Z-K))\right]$$

where Z is the zero coupon price for expiry T_2 as observed at time T_1. It has current forward value

$$f(Z) = \frac{Z(r,t,T_2)}{Z(r,t,T_1)}.$$

It turns out that we obtain a 'Black-like' option pricing formula:

$$c[r,t,T_1,T_2;K] = Z(r,t,T_2)N(d_+) - KZ(r,t,T_1)N(d_-) \tag{4.38}$$

$$p[r,t,T_1,T_2;K] = KZ(r,t,T_1)N(-d_-) - Z(r,t,T_2)N(-d_+) \tag{4.39}$$

$$\sigma_p^2 = \sigma(T_2 - T_1)^2 T_1 \tag{4.40}$$

$$d_\pm = \frac{\ln\frac{f(Z)}{K} \pm \frac{1}{2}\sigma_p^2}{\sigma_p} \tag{4.41}$$

4.5.4 The Continuous time version of the Black-Derman-Toy model

Again, we can use (Rebonato 1998) for the derivation of the continuous equivalent of the (Black et al. 1990) model.

In the (Black et al. 1990) model, a lognormal distribution of the short term interest rate is assumed i.e. $\ln r(t)$ is normally distributed. At each node $(i, j-1)$ we have 2 possible states of the world and interest rates denoted $r(i,j)$ and $r(i+1,j)$ that we can evolve to. The mean short term interest rate at this time may then be calculated as:

$$\ln r_m(i,j) = \tfrac{1}{2}\left[\ln r(i,j) + \ln r(i+1,j)\right]$$

Also

$$r(i+1,j) = r(i,j)\exp\left[2\sigma(j)\sqrt{\Delta t}\right]$$

and hence, we find each of the two possible rates as an offset from the median rate of interest, $r_m(i,j)$

$$\begin{aligned} r(i+1,j) &= r_m(i,j)\exp\left[\sigma(j)\sqrt{\Delta t}\right] \\ r(i,j) &= r_m(i,j)\exp\left[-\sigma(j)\sqrt{\Delta t}\right] \end{aligned}$$

So, the correct continuous analogue of this is as follows

$$r(t) = u(t)\exp\left(\sigma(t)z(t)\right) \tag{4.42}$$

where

$$\begin{aligned} u(t) &\sim \text{time } t \text{ median of the short term interest rate distribution,} \\ \sigma(t) &\sim \text{short term interest rate volatility at time } t, \\ z(t) &\sim \text{standard Brownian motion.} \end{aligned}$$

To examine the nature of the stochastic process driving the short term interest rate we must examine the evolution of $\ln r(t)$ where

$$\begin{aligned} d\ln r(t) &= \frac{\partial \ln r(t)}{\partial r}\,dr + \frac{1}{2}\frac{\partial^2 \ln r(t)}{\partial r^2}\,(dr)^2 \\ &= \tfrac{1}{r}\,dr - \tfrac{1}{2r^2}\,(dr)^2 \end{aligned} \tag{4.43}$$

Since $r(t) = r(t, z(t))$ Ito's Lemma gives:

$$dr = \frac{\partial r}{\partial t}\,dt + \frac{\partial r}{\partial z}\,dz + \frac{1}{2}\frac{\partial^2 r}{\partial z^2}\,dzdz \tag{4.44}$$

where

$$
\begin{aligned}
\frac{\partial r}{\partial t}\,dt &= \frac{\partial u(t)}{\partial t}\,\exp\left(\sigma(t)z(t)\right)dt + u(t)\frac{\partial \sigma(t)z(t)}{\partial t}\,\exp\left(\sigma(t)z(t)\right)dt \\
&= \frac{\partial u(t)}{\partial t}\,\exp\left(\sigma(t)z(t)\right)dt + u(t)\left[z(t)\frac{\partial \sigma(t)}{\partial t}\right]\exp\left(\sigma(t)z(t)\right)dt \\
\frac{\partial r}{\partial z}\,dz &= u(t)\sigma(t)\exp\left(\sigma(t)z(t)\right)dz \\
\frac{\partial^2 r}{\partial z^2}\,dzdz &= u(t)\sigma^2(t)\exp\left(\sigma(t)z(t)\right)dt
\end{aligned}
$$

Substituting into (4.43) we have

$$
\begin{aligned}
d\ln r(t) &= \frac{1}{u(t)\exp\left(\sigma(t)z(t)\right)}\left(\frac{\partial u(t)}{\partial t}\,\exp\left(\sigma(t)z(t)\right)dt + u(t)\left[z(t)\frac{\partial \sigma(t)}{\partial t}\right]\exp\left(\sigma(t)z(t)\right)dt\right. \\
&\quad \left. + u(t)\sigma(t)\exp\left(\sigma(t)z(t)\right)dz + \tfrac{1}{2}\,u(t)\sigma^2(t)\exp\left(\sigma(t)z(t)\right)dt\right) \\
&\quad - \frac{1}{2u^2(t)\exp\left(2\sigma(t)z(t)\right)}\,u^2(t)\sigma^2(t)\exp\left(2\sigma(t)z(t)\right)dt \\
&= \frac{\partial \ln u(t)}{\partial t}\,dt + z(t)\frac{\partial \sigma(t)}{\partial t}\,dt + \sigma(t)dz + \tfrac{1}{2}\,\sigma^2(t)dt - \tfrac{1}{2}\,\sigma^2(t)dt \\
&=^{2} \frac{\partial \ln u(t)}{\partial t}\,dt + \frac{\ln r(t) - \ln u(t)}{\sigma(t)}\frac{\partial \sigma(t)}{\partial t}\,dt + \sigma(t)dz \\
&:= \left[\theta(t) + \frac{\sigma'(t)}{\sigma(t)}\ln r(t)\right]dt + \sigma(t)dz
\end{aligned}
$$

(Black & Karasinski 1991) is the general version of the above:

$$
d(\ln r(t)) = [\theta(t) + a(t)\ln r(t)]dt + \sigma(t)dz
$$

Note that these are NOT affine models.

It is tempting to work one-factor models in order to match not only the term structure but also to match the price of caps (liquid instruments) and swaptions. However, this amounts to over-parameterisation and leads to a non-stationary volatility structure. The corresponding volatility term structure implied by the model is unlikely to be anything like the existing volatility structure and can create derivative mispricing.

---

[2] From (4.42) we have:

$$
\begin{aligned}
r(t) &= u(t)\exp\left(\sigma(t)z(t)\right) \\
\Rightarrow\quad z(t) &= \frac{\ln r(t) - \ln u(t)}{\sigma(t)}
\end{aligned}
$$

4.5.5 The Extensions of Hull & White

(Hull & White 1990) extended the Vasicek and CIR models to include time-dependency in the drift and volatility parameters. The extended Vasicek model is

$$dr = (\eta(t) - \gamma r(t))dt + \sqrt{\beta}dz \tag{4.45}$$

and the extended CIR model is

$$dr = (\eta(t) - \gamma r(t))dt + \sqrt{\alpha r(t)}dz \tag{4.46}$$

The extended Vasicek model with constant volatility is analytically tractable and results in a richer volatility structure for forward rates, spot rates and discount bonds.

(4.19) and (4.20) have become

$$\begin{aligned}\frac{\partial A}{\partial t} &= \eta(t)B - \tfrac{1}{2}\alpha B^2 \\ \frac{\partial B}{\partial t} &= -1 + \beta B\end{aligned}$$

and so

$$\begin{aligned}Z(t,T) &= \tfrac{1}{\gamma}(1 - e^{-\gamma(T-t)}) \\ A(t,T) &= -\int_t^T \eta(s)Z(s,T)\,ds + \tfrac{1}{2}\beta\int_t^T Z(s,T)^2\,ds\end{aligned}$$

and manipulating as we did in §4.5.3 we get

$$\int_0^T \eta(s)Z(s,T)\,ds = \tfrac{1}{2}\beta\int_0^T Z(s,T)^2\,ds - rZ(0,T) + \mathbb{r}(r,0,T)T$$

Assuming that γ and β have been estimated statistically, we need to derive the function $\eta(T)$ at time $t = 0$, today.

As before we differentiate twice w.r.t. T. In §4.5.3, when we did so, $\eta(T)$ emerged on the left hand side. This time, it is only slightly more complicated. The result is

$$\eta(T) = \frac{\partial^2}{\partial T^2}[T\mathbb{r}(r,0,T)] + \gamma\frac{\partial}{\partial T}[T\mathbb{r}(r,0,T)] + \tfrac{1}{2}\frac{\beta}{\gamma}(1 - e^{-2\gamma T}) \tag{4.47}$$

4.6 Modelling the Yield Curve

This is based closely on (Wilmott 2000, §41.4).

Why do we want to fit a yield curve using the method of §4.5.3 or §4.5.5? We need to price our derivatives, but as importantly, we need to hedge them. A model is useless if the hedge instruments do not price nearly perfectly within the model. We thus require a model which prices all observable instruments exactly. Assuming that the observable instruments are exactly the zero coupon bonds (more precisely, the instruments that are used in the bootstrap) these approaches will then satisfy this requirement.

The problem is that if we later re-bootstrap the yield curve, and then refit the $\eta(T)$ function, we will find that it has changed materially. In theory, we should only move into the function i.e. the vertical axis should simply move to the right. Thus, the theory is wrong (although not necessarily not useful).

What about choices of α, β and γ? So, the theme now will be: make reasonable choices of these parameters, and then the function $\eta(T)$ follows algorithmically. Then we have the model completely specified, and we can price derivatives accordingly.

- The volatility is $\sqrt{\alpha r + \beta}$. Hence there is a lower bound on the short rate of $-\frac{\beta}{\alpha}$, at least for those models where $\alpha \neq 0$. So, for example, we could consider the smallest value $r_{\min}$ the short rate has been in a suitable period of history, and hope to arrange that $r_{\min} = -\frac{\beta}{\alpha}$. A proxy for the short rate might be the overnight rate.
- Likewise, let us measure the historical volatility σ_h of the short rate (using an EWMA scheme or some such). Then $\sigma_h = \sqrt{\alpha r + \beta}$. Now we have two equations in the two unknowns α and β, which we can solve.

Now let us examine the behaviour of the yield curve at the short end, with T not much bigger than t. Let us suppose we have a Taylor series expansion

$$Z(r,t,T) = 1 + a(r)(T-t) + b(r)(T-t)^2 + c(r)(T-t)^3 + \cdots$$

We substitute this into the derivative pricing equation (4.9), in a second order expansion around T, as follows:

$$\begin{aligned} 0 \approx\ & -a - 2b(T-t) - 3c(T-t)^2 \\ & + \tfrac{1}{2}\left[\sigma^2 - 2(T-t)\sigma \left.\frac{\partial \sigma}{\partial t}\right| T\right]\left[\frac{\partial^2 a}{\partial r^2}(T-t) + \frac{\partial^2 b}{\partial r^2}(T-t)^2\right] \\ & + \left[\mu - \lambda\sigma - (T-t)\left.\frac{\partial}{\partial t}(\mu - \lambda\sigma)\right| T\right]\left[\frac{\partial a}{\partial r}(T-t) + \frac{\partial b}{\partial r}(T-t)^2\right] \\ & - r(1 + a(T-t) + b(T-t)^2) \end{aligned}$$

and so, equating powers of $T-t$, we have

$$\begin{aligned}
a(r) &= -r \\
b(r) &= \tfrac{1}{2}r^2 - \tfrac{1}{2}(\mu - \lambda\sigma) \\
c(r) &= \tfrac{1}{12}\sigma^2 \frac{\partial^2}{\partial r^2}(r^2 - r(\mu - \lambda\sigma)) - \tfrac{1}{6}(\mu - \lambda\sigma)\frac{\partial}{\partial r}(r^2 - r(\mu - \lambda\sigma)) \\
&\quad - \tfrac{1}{3}\frac{\partial}{\partial t}(\mu - \lambda\sigma) + \tfrac{1}{6}r^2(r^2 - r(\mu - \lambda\sigma))
\end{aligned}$$

where $\mu = \mu(r,T)$, $\lambda = \lambda(r,T)$, $\sigma = \sigma(r,T)$.

Again, using the second order Taylor series expansion we have

$$\mathbb{r}(r,t,T) = -a + (\tfrac{1}{2}a^2 - b)(T-t) + (ab - c - \tfrac{1}{3}a^3)(T-t)^2 + \cdots$$

in other words, the yield curve has

- NACC rate at the short end equal to the short rate of the short rate model
- slope at the short end equal to $\frac{1}{2}a^2 - b = \frac{1}{2}(\mu - \lambda\sigma)$
- curvature at the short end 'equal' to $ab - c - \frac{1}{3}a^3$.

Thus, the slope of the yield curve depends on the risk neutral drift, and vise versa. This slope is often very large positive or negative, but typically the curve flattens out for longer maturities. Thus, there is a large curvature.

When fitting Ho-Lee or Hull-White-Vasicek, one typically finds the following: the value of $\eta(0)$ is very large. This is because the yield curve slope is large. The slope of $\eta(T)$ at $T=0$ is large and negative. This is because the curvature of the yield curve is large.

If the theory held true, then as time passed, the vertical axis should simply move to the right, with no other change to the function. In particular, the value and slope of $\eta(\cdot)$ at $T=0$ should reduce in size. This does not in fact happen: rather, both the axis and the function move in time. The basic shape of the yield curve is maintained; forward rates are not realised.

No one factor model will capture the high slope and curvature that is usual for yield curves. They may give reasonable results when the yield curve is quite flat. In general one needs to go to two factor or multi-factor models.

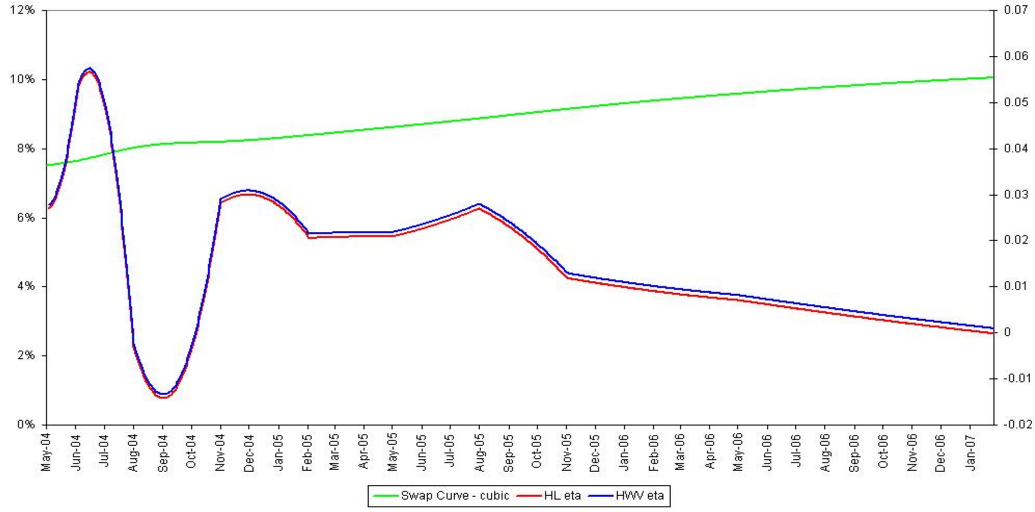

Figure 4.3: The Ho-Lee and Hull-White-Vasicek $\eta(\cdot)$ functions on 19 May 2004

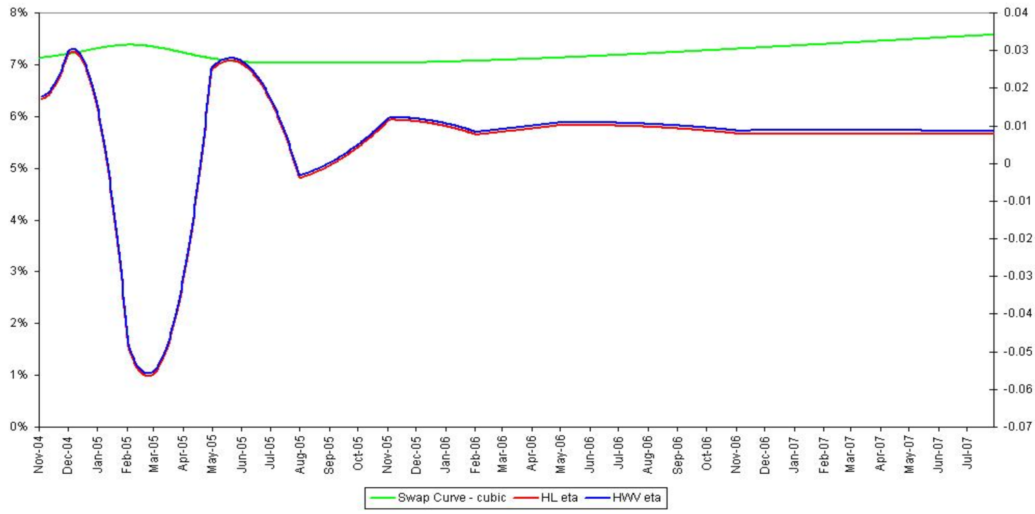

Figure 4.4: The Ho-Lee and Hull-White-Vasicek $\eta(\cdot)$ functions on 19 Nov 2004

Chapter 5

Credit Risk in interest rate derivatives

5.1 Credit risky bonds

A capital market portfolio receives mark-to-market treatment, probably on a daily basis, using readily available price and ratings change data. Portfolio theory is well-defined at this point. On the other hand, loan portfolios are usually comprised of illiquid assets that tend to be held to maturity; the relevant events are credit changes and defaults.

Default risk is the risk that a counterparty or country will become unable to meet its financial obligations. In the event of a default, the firm with the credit exposure will incur a loss equal to the owed amount minus whatever the firm can recover from a foreclosure, liquidation, or restructuring. Default risk is generally associated with instruments that tend to be held to maturity, such as loans.

The credit spread is the excess return that the market demands in exchange for assuming credit exposure associated with a given counterparty or country. Spread risk is the risk of loss resulting from a change in the level of the credit spread used in the mark-to-market of a financial instrument such as a bond or a credit derivative. One might think of spread risk in terms of measuring partial default risk, since an obligor could be considered more likely to default if its credit rating is lowered. Spread risk is often seen as a type of market risk, however, it makes sense to consider it as a credit risk.

In addition, most loans contain embedded optionality, in the form of unique terms or the ability to prepay. This contrast tends to complicate the management of overall credit risk.

Let us now see how the price of credit risky bonds allows us to calculate the risk neutral probability of default. Consider

$$\begin{aligned} r(T) &\sim \text{ zero coupon bond yield for term } T \\ y(T) &\sim \text{ zero coupon bond yield for term } T \\ s(T) &\sim \text{ zero coupon credit spread for term } T \\ P(T) &\sim \text{ probability of default by time } T \\ \Omega(T) &\sim \text{ recovery percentage at time } T \text{ on a default} \end{aligned}$$

Note that $y(T) = r(T) + s(T)$. We have the equation of value

$$\begin{aligned} e^{-y(T)T} &= (1 - P(T) + P(T)\zeta(T))e^{-r(T)T} \\ \Rightarrow e^{-s(T)T} &= 1 - P(T)(1 - \Omega(T)) \end{aligned} \tag{5.1}$$

Now, in principle we know $s(T)$ (why?) and so we have a single equation in two unknowns. $\Omega(T)$ is typically estimated from market data (the price at which defaulted bonds are trading in the market), and so $P(T)$ is inferred. Alternatively, there are bootstrap mechanisms where the pair $P(T)$ and $\Omega(T)$ of term structures in inferred simultaneously inductively; see (Hull 2002, §26.1) for example.

This is a risk-neutral probability of default. In reality, the historically measured (real world) default probabilities are typically lower. The premium is usually explained as being due to undiversifiable, systemic, risk.

When pricing derivatives we use the risk neutral probability. When performing scenario analysis, or calculating risk measures, we use the real world probability.

A concept related to probability of default is the hazard rate. The hazard rate $h(t)$ is the intensity of default at time t i.e. $h(t)\delta t$ is the probability of default in the time interval $[t, t+\delta t]$, conditional on there being no default until time t. $P(T)$ defined above is tantamount to the existence of a default pdf $p(t)$ such that $P(T) = \int_0^T p(t)\, dt$. In other words, $p(t)\, dt$ is the unconditional probability of default in the time interval $[t, t + \delta t]$ as seen at time 0.

Note that either $h(\cdot)$ or $p(\cdot)$ or $P(\cdot)$ can be used to describe the world, as

$$p(T) = h(T) \exp\left(-\int_0^T h(t)\, dt\right) \tag{5.2}$$

In practice the default probability approach dominates, but the hazard rate approach is preferable when, for example, building default into a finite difference scheme - as the differential equation being solved holds instantaneously. See §5.4.5 for example.

5.2 CreditMetrics ©

This model was developed in 1997 by JP Morgan (J.P.Morgan & Reuters 1997); it was the first readily available portfolio model for evaluating credit risk. It is a 'portfolio approach' model which was developed to address portfolio diversification and to rectify 'concentration risk' - to have large exposures to a single counterparty or group of counterparties. CreditMetrics is a methodology for assessing portfolio Value At Risk due to changes in obligor credit quality.

There is not much of a connection to RiskMetrics, in the sense that the risks of simultaneous market and credit losses are not dealt with. This can be seen as a potential drawback since losses are generated by credit-rating changes only and not by market movements. This results in no uncertainty over market exposures. However, CreditMetrics does look to a horizon and construct a distribution of historically estimated credit outcomes (rating migrations including potential default). Each credit quality migration is weighted by its likelihood (transition matrix analysis). Each outcome has an estimate of change in value (given by either credit spreads or studies of recovery rates in default). We then aggregate volatilities across the portfolio, applying estimates of correlation. Thus, although the relevant time horizon is usually longer for credit risk, with CreditMetrics we compute credit risk on a comparable basis with market risk.

One of the fundamental techniques is migration analysis, that is, the study of changes in the credit quality of names through time. J.P. Morgan developed transition matrices for this purpose as early as 1987. CreditMetrics has since built upon a broad literature of work, which applies migration analysis to credit risk evaluation. The basic approach is via modelling the credit rating. By modelling the credit rating one can answer the basic question of what is the probability of a particular name defaulting by a certain time, or in a certain time interval. More generally we can ask what is the probability of them being downgraded - this is of interest as this will negatively impact the value of any bonds or the like that they have outstanding. More subtly, we can hope to have aggregation mechanisms where we can answer the same questions for a pair or basket of names - this will lead into the world of copulas.

Deriving the credit rating is a process - fairly non-transparent - that is performed by the major credit rating agencies: Standard & Poor's, Moody's, Fitch. Alternatively (and this is especially the case in South Africa, where international credit rating agency expertise can be patchy) the credit rating may be a bank's own private rating for the credit risk.

As an example, S&P use the symbols AAA, AA, A, BBB, BB, B, CCC, and D, where AAA denotes the very best credits, decreasing in order of quality, until CCC which denotes the absolute worst credit quality short of actual default, and D denotes D that has occurred. D is an absorbing state i.e. a state which cannot be escaped from. The rating agencies update their ratings at least yearly and more frequently if there has been some news that could affect the rating.

The transition matrix (also published by the above) is the matrix of probabilities $M = [m_{ij}]$ where m_{ij} denotes the probability of moving from rating category i to rating category j in the next discrete time period, which is taken to be one year. See Figure 5.1.

There will be a Markov assumption, so the transition matrix for time α is M^α. This is

	probability of rating at year end (percent)							
	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	93.66	5.83	0.4	0.08	0.03	0	0	0
AA	0.66	91.72	6.94	0.49	0.06	0.09	0.02	0.01
A	0.07	2.25	91.76	5.19	0.49	0.2	0.01	0.04
BBB	0.03	0.25	4.83	89.26	4.44	0.81	0.16	0.22
BB	0.03	0.07	0.44	6.67	83.31	7.47	1.05	0.98
B	0	0.1	0.33	0.46	5.77	84.19	3.87	5.3
CCC	0.16	0	0.31	0.93	2	10.74	63.96	21.94
Default	0	0	0	0	0	0	0	100

Figure 5.1: One-year ratings migration probabilities based upon bond rating data from 1981-2000. Data is adjusted for rating withdrawals. Source: Standard & Poor's.

clear for $\alpha \in \mathbb{N}$; it can be made to make sense for other α too, using matrix mathematics. Note that the Markov assumption can be weak because

- The rating process might be name specific as well and rating specific. Rating's volatility will be a function of the name and the sector of the ecomomy they are working in.
- Credit ratings have momentum: if a credit rating has changed recently, there is a greater probability than otherwise that another change in the same direction will follow shortly.
- Cyclical economic factors will be a significant contributing factor to default probabilities. If the economy is healthy/unhealthy, ratings might drift up/down - this clearly is a poor model of reality. One alternative is to have a model that allows for a transition matrix in a healthy economy and another in an unhealthy economy, with a separate (Markov?) model determining the state of the economy.

See (Skora 2000).

Note that the above transition matrices are real-world transition matrices - they are derived from empirical historical analysis. What about risk-neutral transition matrices? We will model these from bond spreads.

Suppose that using the bond spreads we find that the annual risk neutral default probabilities, for a four state world, are as in Figure 5.2. Let $\underline{d}_1,\ \underline{d}_2,\ \ldots,\ \underline{d}_5$ be the columns of this table.

Let the unknown 4×4 risk neutral transition matrix be M. Note that $[M^n]_{i4} = d_{i,n}$. We have an overspecified system: M has 9 degrees of freedom, and there are 15 conditions to be satisfied. Furthermore, there are non-linear constraints: $1 \geq M_{ij} \geq 0$ for $1 \leq i, j \leq 4$.

	1	2	3	4	5
A	0.45%	1.00%	2.00%	2.50%	3.00%
B	1.66%	4.00%	6.00%	8.00%	10.00%
C	3.00%	6.00%	9.00%	11.00%	14.00%
D	100.00%	100.00%	100.00%	100.00%	100.00%

Figure 5.2: Spread implied cumulative default probabilities for year n, $1 \leq n \leq 5$

	A	B	C	D
A	96.92%	2.52%	0.00%	0.56%
B	0.14%	89.19%	8.74%	1.92%
C	2.17%	1.34%	93.42%	3.08%
D	0.00%	0.00%	0.00%	100.00%

Figure 5.3: Resulting risk neutral transition matrix

Thus, we cannot solve the system exactly, but can only hope to find a 'best fit' solution. Thus, we need to minimise the distance between $[M^n]_{i4}$ and $d_{i,n}$. A suitable error function is

$$\sum_{n=1}^{5} \sum_{i=1}^{3} ([M^n]_{i4} - d_{i,n})^2$$

Using solver to minimise the error function, including the constraints, we have the annual transition matrix seen in Figure 5.3. The values are intuitively sensible.

Note that solver is not the most robust piece of software ever built. The solution does not appear to be unique, nor especially stable, and is quite dependent on the seed guess at the transition matrix.

5.3 Credit from equity pricing: the Merton Model and KMV

The Merton approach assumes that the value of a company's equity is determined solely by the value of the firm's assets and its debt outstanding.

The method is based on stock prices rather than bond yields. This is advantageous, especially in South Africa, which essentially does not have a corporate bond market.

The mathematics of the Merton model is as follows:

- Let the firm's value be V.

- Let the firm have one bond B in issue that matures at T and has face value K ie. the debt of the company is a bullet.
- Let the equity of the firm at T be $S(T)$.

At the end of the period, if the firm is worth more than K then the bond is paid in full. If the firm is worth less than the bond, default occurs and the bond-holders receive V only. So,

$$S(T) = \max(V(T) - K, 0).$$

Hence, the equity of a company is represented by a call option on its assets, the exercise price is the face value of the debt of the company.

The Black-Scholes analysis gives the value of the equity as

$$S(0) = V(0)N(d_1) - Ke^{-rT}N(d_2)$$
$$d_{1,2} = \frac{\ln\frac{V(0)}{K} + \left(r \pm \frac{1}{2}\sigma_V^2\right)(T)}{\sigma_V\sqrt{T}}$$

The risk neutral probability of default is $N(-d_2)$. Also, from Ito's lemma,

$$\sigma_S S(0) = \frac{\partial S}{\partial V}\sigma_V V(0) = N(d_1)\sigma_V V(0)$$

Now, given $S(0)$ and σ_S, we have two (highly non-linear) equations in the two unknowns $V(0)$ and σ_V. These can be solved using the two-dimensional Newton method, and we deduce $V(0)$, hence $B(0) = V(0) - S(0)$, and $N(-d_2)$.

One of the biggest drawbacks of this initial theory was the assumption of constant interest rates - this resulted in poor test results. Another disadvantage is that sovereign credit risk cannot be measured with this model as a country cannot have a "stock price'.

KMV is a private company, based in San Francisco, that provides credit risk and portfolio management products to more than 25,000 companies around the world. The credit risk theory that KMV use is based on a version of the Merton model, but is well developed beyond the original Merton model. One feature that is important is that the distributions in the Merton model, typically normal, are callibrated to actual historical market data in the KMV model. This data and the distributions are 'encoded" within the software; updates are provided to clients on a monthly basis.

The driving force in KMV's credit risk tool is the "Estimated Default Frequency" or EDF© credit measure: it is the probability that a firm will default within a given time horizon (KMV 2004), in particular, EDF has a term structure.

KMV has tested statistically as being a superior credit risk management tool than any of the more traditional ratings approaches, in particular, changes in EDF's lead changes in ratings. This is quite intuitive - ratings have very obvious time delays, whereas the EDF's are driven by equity prices, and hence react immediately to market sentiment. KMV claims that EDFs anticipate rating changes by six to eighteen months.

The equity value of a firm is represented as a call option on the assets of the firm, the strike is the value of "liabilities" which is equal to the short term liabilities (less than a year to maturity) plus half the value of the long term debt. The market (or asset) value of the firm is derived from the book-value of its obligations and the market value of its equity (using option-pricing theory).

The EDF measure is based on "distance to default": the distance between the current value of assets and the default point. It is defined to be the percentage drop required to bring a firm from its current asset value to the default point, normalized by the volatility of the asset value: the "distance to default", D is:

$$D = \frac{V - K}{\sigma_V} \tag{5.3}$$

An assumption of lognormally distributed returns would allow us to determine the probability of the firm value falling below D, which would then be the EDF. However, distribution assumptions are not made in the KMV model.

In conclusion, some strengths and weaknesses of KMV:

- Like its foundational model (Merton's model) it relies on stock prices and not issued debt instruments which are not always practically available. As Stephen Kealhofer (the K in KMV) explains, 'The problem is that the debt markets are just too thin to be a good source of information themselves, whereas the equity markets are continuous."
- The EDF credit measure is cardinal rather than ordinal - agency ratings are ordinal.
- KMV claim that their EDF measures consistantly outperform other methods and provide more accurate and timely results.
- KMV uses accounting data, which may be stale, or less appropriate in South Africa. It is definitely inappropriate for financial institutions because gearing seems to get lost. However, in their most recent release, KMV claim to have remodelled financial institutions in an appropriate way.
- Company levels of debt are often kept constant which is also inefficient.
- The approach is also quite narrowed in that only a fraction of the Credit-Risk information is provided; correlations of defaults and exposure and loss distributions are omitted.

5.4 Credit derivatives

5.4.1 Credit default swaps

In these products two parties enter into an agreement whereby one party pays the other a fixed (small) coupon for the specified life of the agreement. The other party makes no payments unless a specified credit event occurs on a 'reference asset'.

If such a credit event occurs, the swap terminates, typically with an accrued payment. The short party makes a 'large' payment to the long party - the insurance of the credit default swap. The payment is calculated as a function of the loss to the reference asset's market value following the credit event. The long party does not make any further spread payments.

Alternatively, the party that was long the credit default swap might be able to now sell the reference asset to the short party for some contractually specified sum (par, for example). The latter might be the case if the long party held the reference asset, of course they might only be speculating on the creditworthiness of the reference asset, in which case the former payoff would probably have been structured.

Let us look at a model for valuing these products. Let the product run from time 0 to time T. Let the notional principle be 1. Let π be the risk neutral probability of no credit event during the life of the swap. Then $\pi = 1 - \int_0^T h(t)\,dt$, where recall that $h(t)$ is the hazard rate. Let $u(t)$ be the present value of payments that are received on the swap in the time $[0,t]$, these payments being risk free, at the rate of R1 per annum. Let $A(t)$ be the accrued interest that has accrued at time t (this is a saw-tooth function, resetting to 0 every time a coupon payment is made) and let $e(t)$ be the risk free present value thereof.

Let the annual rate on the swap be w. The present value of the payments which occur periodically, until T or until a prior credit event, is

$$\int_0^T h(t)w(u(t) + e(t))\,dt + \pi w u(T)$$

If a default event occurs at time t, then the reference obligation has value $(1 + A(t))R$ where R is the risk neutral recovery rate on the defaulted asset. Hence, the risk neutral expected payoff on the credit default swap will be $1 - (1 + A(t))R = 1 - R - A(t)R$. So the value of the payoff is

$$\int_0^T h(t)Z(0,t)(1 - R - A(t)R)\,dt$$

Thus the value to the buyer of the credit protection is

$$V = \int_0^T h(t)[Z(0,t)(1 - R - A(t)R) - w(u(t) + e(t))]\,dt - \pi w u(T)$$

This is equal to 0 when

$$w = \frac{\int_0^T h(t)Z(0,t)(1 - R - A(t)R)\, dt}{\int_0^T h(t)(u(t) + e(t))\, dt + \pi u(T)} \tag{5.4}$$

This analysis has assumed that default events, interest rates, and recovery rates are mutually independent.

5.4.2 Total return swaps

In these products, the total return on a corporate bond is swapped for a return of IBOR plus a spread. By total return here, we mean all coupons and the change in value, which may be calculated once off at termination, or is periodically netted through the life of the transaction. If there is default on the corporate bond, the swap usually terminated and the change in value payment is made then.

Suppose an investor wishes to obtain exposure to a corporate (the reference asset). One way is to borrow the money required to buy the corporate bond. Another is to pay LIBOR and receive the total return of the reference asset in such a swap. The counterparty will be the bank that they might otherwise have borrowed the money from. The bank receives IBOR plus the spread, while the investor receives full exposure to the reference asset. The advantage to the bank is that their exposure to the investor is much reduced.

5.4.3 Credit spread options

This is an option whose payoff is a function of the spread between the bonds of the reference borrower and risk free bonds. The tenor of the bonds will be specified in the contract.

5.4.4 Collateralised debt obligations

Collateralised debt obligations are securitised interests in pools of assets. These assets are usually loans (then, a CLO), bonds (then, a CBO), mortgages (then, a CMO) or in reality any debt instruments.

Investors bear the credit risk of the party which is repaying the debt. Typically multiple tranches are issued, and these tranches are categorized as senior, mezzanine, and equity, according to their degree of credit risk. If there are defaults or the CDO underperforms in some contractually specified way, then the equity holders suffer first. Eventually the equity holders get nothing, and then the mezzanine holders suffer. Such a structure is known as a waterfall: the repayments are first made to the senior tranche, if there are payments remaining, these cascade over the waterfall into the mezzanine, etc.

Senior and mezzanine tranches are typically rated. The senior tranche can be rated very highly - typically at least A, and often at the highest rating. Of course, the senior tranches offer the lowest return.

5.4.5 Credit in convertible bonds

We won't look at this here because the impact of stochasticity of interest rates in the pricing of these instruments is very small. See (West 2004*a*).

Bibliography

Abramowitz, M. & Stegun, I. A. (1974), *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, Dover.

Black, F. (1976), 'The pricing of commodity contracts', *Journal of Financial Economics* **3**, 167–179.

Black, F., Derman, E. & Toy, W. (1990), 'A one-factor model of interest rates and its application to treasury bond options', *Financial Analysts Journal* pp. 33–39.

Black, F. & Karasinski, P. (1991), 'Bond and option pricing when short rates are lognormal', *Financial Analysts Journal* pp. 52–59.

Brace, A., Gątarek, D. & Musiela, M. (1997), 'The market model of interest rate dynamics', *Mathematical Finance* **7**(2), 127–155.

Brennan, M. J. & Schwartz, E. S. (1979), 'A continuous time approach to the pricing of bonds', *Journal of Banking and Finance* **31**, 33–155.

Brennan, M. J. & Schwartz, E. S. (1982), 'An equilibrium model of bond pricing and a test of market efficiency', *Journal of Financial and Quantitative Analysis* **17**, 301–329.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**, 385–407.

Dothan, M. (1978), 'On the term structure of interest rates', *Journal of Financial Economics* **6**, 59–69.

Hagan, P. & West, G. (2004), 'Interpolation methods for yield curves'. In preparation.

Ho, T. S. Y. & Lee, S.-B. (1986), 'Term structure movements and pricing interest rate contingent claims', *Journal of Finance* **41**.

Hull, J. (2002), *Options, Futures, and Other Derivatives*, fifth edn, Prentice Hall.

Hull, J. & White, A. (1990), 'Pricing interest rate derivative securities', *Review of Financial Studies* **3**.

International Swaps and Derivatives Association (25 November 1998), 'EMU and market conventions: Recent developments'.
*http://www.isda.org/c_and_a/pdf/mktc1198.pdf

Jamshidian, F. (1989), 'An exact bond option pricing formula', *Journal of Finance* **44**, 205–209.

J.P.Morgan & Reuters (1997), *CreditMetrics - Technical Document*, first edn, J.P.Morgan and Reuters, New York.
*http://www.riskmetrics.com

KMV (2004), 'Moody's kmv'.
*http://www.moodyskmv.com/

Longstaff, F. & Schwartz, E. (1992), 'Interest rate volatility and the term structure: A two-factor general equilibrium model', *Journal of Finance* **47**.

Rebonato, R. (1998), *Interest-Rate Option Models*, second edn, John Wiley and Sons Ltd.

Rendelman, R. & Bartter, B. (1980), 'The pricing of options on debt securities', *Journal of Financial and Quantitative Analysis* **15**.

Skora, R. K. (2000), 'Modeling credit risk'.
*http://www.fenews.com/fen16/creditrisk.html

Svoboda, S. (2002), An investigation of various interest rate models and their calibration in the South African market, Master's thesis, University of the Witwatersrand.
*http://www.cam.wits.ac.za/mfinance/research.html

Svoboda, S. (2003), *Interest Rate Modelling*, Finance and Capital Markets Series, Palgrave Macmillan.

Vasicek, O. A. (1977), 'An equilibrium characterisation of the term structure', *Journal of Financial Economics* **15**.

West, G. (2004*a*), 'Convertible bond pricing offered by RiskWorX'.
*http://www.riskworx.com/pdfs/cbsummary2.pdf

West, G. (2004*b*), 'The Mathematics of South African Financial Markets and Instruments'. Lecture notes, Honours in Mathematics of Finance, University of the Witwatersrand, Johannesburg.
*http://www.cam.wits.ac.za/mfinance/graeme/safm.pdf

Wilmott, P. (2000), *Paul Wilmott on Quantitative Finance*, John Wiley and Sons Ltd.