Math 772 - Credit Risk and Interest Rate Modeling

M. R. Grasselli
Dept. of Mathematics and Statistics
McMaster University
Hamilton,ON, L8S 4K1

March 27, 2005

Contents

1 First Lecture **5**

1.1 Default-free Bond Markets 5
1.2 Rates 6
1.3 Differentials 8
1.4 No-arbitrage for Bond Markets 9
1.5 The change of Numeraire Technique 10

2 Second Lecture **13**

2.1 Pricing in Short Rate Models 13
2.2 Affine Models 15
2.3 The Vasicek Model 16
2.4 The Dothan Model 17
2.5 The Exponentiated Vasicek Model 18
2.6 The Cox–Ingersoll-Ross Model 19
2.7 Parameter Estimation and the Initial Term Structure 20

3 Third Lecture **21**

3.1 The Ho–Lee Model 21
3.2 The Hull–White Extended Vasicek Model 21
3.3 Deterministic–Shift Extensions 23
3.3.1 The CIR ++ Model 24
3.4 Forward Rate Models 25

4 Fourth Lecture **26**

4.1 Some interest rates derivatives 26
4.2 General Option Pricing Formulas 29

5 Fifth Lecture **32**

5.1 Default Events 32
5.2 Structural Models for Default Probabilities 34
5.2.1 The Merton Model (1974) 35
5.2.2 First passage models 36
5.2.3 Excursion Models 37

6 Sixth Lecture **38**

6.1 Implied Survival Probabilities 38
6.2 Credit Spreads 39

6.3 Properties of Structural Models 40
6.3.1 Credit Spreads 40
6.3.2 Risk Premium 41
6.3.3 Calibration 42

7 Seventh Lecture **44**
7.1 Reduced form models 44
7.1.1 Poisson processes 44
7.1.2 Inhomogeneous Poisson processes 45
7.2 Cox processes 46
7.2.1 Simulating the default time 48

8 Eighth Lecture **50**
8.1 Affine Intensity Models 50
8.1.1 CIR Intensities 50
8.1.2 Mean-reverting intensities with jumps 51
8.2 Risk-neutral intensity models 52
8.2.1 Valuation of defaultable bonds 52
8.2.2 Two-factor Gaussian models 53
8.2.3 From actual to risk neutral intensities 54

9 Ninth Lecture **55**
9.1 Credit Rating Models 55
9.1.1 Discrete-time Markov chain 55
9.1.2 Continuous-time Markov chain 57
9.2 Recovery Models 59
9.2.1 Zero recovery 59
9.2.2 Recovery of treasury 59
9.2.3 Recovery of par 60
9.2.4 Recovery of market value 60

10 Tenth Lecture **61**
10.1 The Binomial Expansion Technique 61
10.2 Default correlation in reduced-form models 62
10.2.1 Doubly stochastic models 62
10.2.2 Joint default events 63
10.2.3 Infectious defaults 64
10.3 Default correlation in structural models 65

11 Eleventh Lecture **68**
11.1 Definition of Copula Functions 68
11.2 Fréchet bounds 69
11.3 Tail dependence 70
11.4 Archimedean copulas 72
11.5 An example of default modeling with copulas 73

12 Twelfth Lecture **75**
12.1 Credit Default Swaps 75

1 First Lecture

Unless otherwise stated, we consider throughout these notes a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)$ satisfying the usual conditions of right-continuity and completeness. In what follows, W_t denotes a standard P–Brownian motion of unspecified dimension.

1.1 Default-free Bond Markets

A zero-coupon bond is a contract that pays the holder one unit of currency at its maturity time.

Definition 1.1 *Let P_{tT} denote the price at time $t \leq T$ of a default-free* zero-coupon bond *maturing at time T.*

We assume that there exits a frictionless market for such bonds for all maturities $T > 0$, with $P_{TT} = 1$ holding for all T. This is done for mathematical convenience, since real bond markets, albeit voluminous, are relatively illiquid and only offer bonds maturing at specific dates, typically ranging from one to 40 years (the majority being between 8 to 20 years). Zero-coupon bonds were introduced by the US Treasury in 1982 and subsequently made popular by local and municipal governments. We will assume these government backed securities to be free of default risk, as opposed to corporate bonds introduced later in these notes. They are, however, subject to volatility risk, since their prices are highly sensitive to fluctuations in interest rates.

As we will see shorlty, all the relevant interest rates can be deduced from the prices of zero-coupon bonds. Conversely, one can recover the prices of zero-coupon bonds from the observed rates quoted in the market. For further convenience, we assume that, for each fixed t, P_{tT} is differentiable with respect to T almost surely.

Definition 1.2 *The* time to maturity $(T-t)$ *is the amount of time, in years, from the present time to the maturity time $T > t$.*

As long as both T and t are expressed as real number, the difference above is unequivocal. However, if dates are represented in day/month/year format, then different day-counting convention result in different values for the time to maturity. In the sequel, we will largely ignore this issue, unless specific contracts force us to do otherwise. As a curiosity, the most popular day-counting convention are: (i) actual/365 (years with 365 days) (ii) actual/360 (years with 360 days), (iii) 30/360 (months with 30 days and years with 360 days).

1.2 Rates

We now proceed in defining interest rates in terms of zero-coupon bond prices. Consider the present date t and two future dates S and T with $t < S < T$.

Definition 1.3 *The* continuously compounded forward rate *for the period $(T - S)$ is the rate $R(t; S, T)$ satisfying*

$$e^{R(t;S,T)(T-S)} \frac{P_{tT}}{P_{tS}} = 1 \qquad \forall t < S < T, \tag{1}$$

That is, the unique rate which is compatible with prices of zero-coupon bonds at time t and continuously compounded interest being accrued deterministically from S to T.

Definition 1.4 *The* simply compounded forward rate *for the period $(T - S)$ is the rate $L(t; S, T)$ satisfying*

$$[1 + L(t; S, T)(T - S)] \frac{P_{tT}}{P_{tS}} = 1 \qquad \forall t < S < T, \tag{2}$$

that is, the unique rate which is compatible with prices of zero-coupon bonds at time t and simply compounded interest being accrued deterministically from S to T in proportion to the investment time.

If we set $t = S$ in the definitions above, we obtain the *continuously compounded yield* $R(t, T)$, defined by

$$e^{R(t,T)(T-t)} P_{tT} = 1, \qquad \forall t < T \tag{3}$$

and the *simply compounded yield* $L(t, T)$, defined by

$$[1 + L(t, T)(T - t)] P_{tT} = 1 \qquad \forall t < T. \tag{4}$$

We use the notation $L(t; S, T)$ above because the LIBOR (London Interbank Offered Rate), fixed daily in London, is the prime example of a simply compounded rate. Under the assumption of differentiability of the bond prices with respect to the maturity date, we obtain that the *instantaneous forward rate* $f(t,T)$ can be defined as

$$f_{tT} := \lim_{T \to S^+} L(t; S, T) = \lim_{T \to S^+} R(t; S, T) = -\frac{\partial \log P_{tT}}{\partial T}. \tag{5}$$

Similarly, we define the *instantaneous spot rate* r_t as

$$r_t := \lim_{T \to t^+} L(t, T) = \lim_{T \to t^+} R(t, T) \tag{6}$$

and it is easy to verify that $r_t = f_{tt}$. Moreover, we readily obtain that

$$P_{tT} = \exp\left(- \int_t^T f_{ts} ds \right). \tag{7}$$

Finally, using the short rate r_t, we can define the *money-market account* as the stochastic process satisfying

$$dC_t = r_t C_t dt, \qquad C_0 = 1 \tag{8}$$

that is, the value of a (locally) risk-less investment accruing interest at the short rate r_t. Consequently, we have that

$$C_t = \exp\left(\int_0^t r_s ds \right). \tag{9}$$

We then see that if we invested exactly C_t/C_T units of currency in the money-market account at time t, we would obtain one unit of currency at time T. An interesting question at this point is the relationship between the bond price P_{tT} and this "discount factor" C_t/C_T. Their difference resides in the fact that P_{tT} is the value of a contract and therefore must be known at time t, while C_t/C_T is a random quantity at t, which depend on the future evolution of the short rate process r_t. Therefore for deterministic interest rates we have that $P_{tT} = C_t/C_T$ and we will see that, for stochastic rates r_t, bond prices are expected values of the discount factor under an appropriate measure.

1.3 Differentials

We now investigate several formal relationships between the stochastic differential equations for bonds, short rates and forward rates. Let

$$\begin{aligned} dr_t &= a(t)dt + b(t)dW_t && (10) \\ \frac{dP_{tT}}{P_{tT}} &= M(t,T)dt + \Sigma(t,T)dW_t && (11) \\ df_{tT} &= \alpha(t,T)dt + \sigma(t,T)dW_t, && (12) \end{aligned}$$

where the last two equations should be interpreted as infinite-dimensional systems of SDE parametrized by the maturity date T. We assume enough regularity in the coefficient functions in order to perform all the formal operations in the sequel.

Proposition 1.5 *1. If P_{tT} satisfies* (10)*, then f_{tT} satisfies* (12) *with*

$$\begin{aligned} \alpha(t,T) &= \frac{\partial \Sigma(t,T)}{\partial T} - \frac{\partial M(t,T)}{\partial T}, \\ \sigma(t,T) &= -\frac{\partial \Sigma(t,T)}{\partial T}. \end{aligned}$$

2. If f_{tT} satisfies (12)*, then r_t satisfies* (10) *with*

$$\begin{aligned} a(t) &= \frac{\partial f(t,t)}{\partial T} + \alpha(t,t) \\ b(t) &= \sigma(t,t) \end{aligned}$$

3. If f_{tT} satisfies (12)*, then P_{tT} satisfies* (11) *with*

$$\begin{aligned} M(t,T) &= r_t - \int_t^T \alpha(t,s)ds + \frac{1}{2}\left\| \int_t^T \sigma(t,s)ds \right\|^2 \\ \Sigma(t,T) &= -\int_t^T \sigma(t,s)ds \end{aligned}$$

Proof:

1. For the first part, apply Itô's formula to $\log P_{tT}$, write in integral form and differentiate with respect to T.

2. Integrate (12), put $T = t$, use the fundamental theorem of calculus for the second variable in $\alpha(s,t)$ and $\sigma(s,t)$, change the order of integration and identify terms.

3. Write $P_{tT} = \exp Y(t,T)$ where $Y(t,T) = -\int_t^T f(t,s)ds$. Apply Itô's formula to it carefully to account for the double appearance of t. Then use the fundamental theorem of calculus and (12) to arrive at an expression for $dY(t,T)$ and finally exchange the order or integration.

1.4 No-arbitrage for Bond Markets

From its definition as a contract that pays one unit of currency at maturity date T, one is tempted to consider a bond as a contingent claim and try to derive its price using the no-arbitrage principles use for pricing options and other contracts in equity markets. For this purpose, let us quickly review the framework of derivative pricing.

Consider a finite time horizon T and an economy consisting of $d+1$ non-dividend paying traded securities whose prices are modeled by the $d+1$–dimensional adapted semimartingale $S_t = (C_t, S_t^1, \ldots, S_t^d)$. A *trading strategy* is a predictable S–integrable process $H_t = (\eta_t, H_t^1, \ldots, H_t^d)$. The *wealth* associated with a trading strategy H is given by $X_t^H = H_t S_t$ and the strategy is called *self-financing* if its wealth process satisfies

$$dX_t^H = H_t dS_t \tag{13}$$

An *arbitrage opportunity* is a self-financing trading strategy such that $X_0^H = 0$, $X_T^H \geq 0$ almost surely and $P(X_T^H > 0) > 0$. Under appropriate technical conditions on the price processes S_t and the allowed trading strategies H_t, the **First Fundamental Theorem of Arbitrage Pricing** says that the market is free from arbitrage opportunities if and only if there exists an equivalent martingale measure Q, that is, a measure with respect to which the discounted asset prices S_t^k/C_t are martingales.

A *contingent claim* for a certain maturity date T is an $\mathcal{F}_T$-measurable random variable B. It is said to be *replicable* if there exists a self-financing trading strategy such that $X_t^H = B$ in which case the law of one price dictates that its price at earlier times $t \leq T$ must be $\pi_t^B = X_t^H$.

The market is said to be complete if all (reasonably integrable) contingent claims are replicable. Under appropriate conditions, the **Second Fundamental Theorem of Arbitrage Pricing** says that for complete markets there exists at most one equivalent martingale measure Q_0. Because the discounted wealth of admissible self-financing portfolios, it follows that the discounted prices of contingent claims in complete markets are martingales with respect to Q_0. In incomplete markets which are free from arbitrage opportunities, there exits more than one (possibly infinitely many) equivalent martingale measures Q. By the same argument as before, replicable contingent claims in these markets will have discounted prices (given by $\frac{\pi_t^B}{C_t} = \frac{X_t^H}{C_t}$), which are martingales under any of the equivalent martingale measures Q. But what about a non-replicable claim ?

By definition, these claims have an effective impact in the economy, in the sense that their presence cannot be replicated by the assets which are already being traded. That is, once a non-replicable claim is written and start being traded, it must be considered as a new asset and arbitrage opportunities might arise if its price is not consistent with the previously existent assets. It follows from the FTAP that such arbitrage opportunities will not arise if and only if, for all $0 \leq t \leq T$, we have

$$\pi_t^B = E_t^Q \left[e^{-\int_t^T r_s ds} B \right], \tag{14}$$

for some equivalent martingale measure Q. In particular

$$P_{tT} = E_t^Q \left[e^{-\int_t^T r_s ds} \right]. \tag{15}$$

1.5 The change of Numeraire Technique

We now describe a general technique for simplifying the integral involved in calculating the prices of derivatives through expressions such as (14) which is particularly useful for interest rate derivatives. First let us define a *numeraire* N_t as any P–a.s strictly positive traded asset and prove the following intuitively clear *invariance lemma*.

Lemma 1.6 *A trading strategy is self-financing in terms of the traded assets with prices $S_t = (C_t, S^1, \ldots, S_t^d)$ if and only if it is self-financing in terms of the normalized asset prices S_t/N_t.*

Proof: Using Itô's formula for both the normalized wealth X_t^H/N_t and the normalized asset prices S_t/N_T, it is easy verify that

$$dX_t^H = H_t dS_t$$

if and only if

$$d\left(\frac{X_t^H}{N_t}\right) = H_t d\left(\frac{S_t}{N_t}\right).$$

Since the concepts of arbitrage, replicability and completeness are all expressed in terms of self-financing trading strategies, this lemma shows that they are invariant with respect to a change of numeraire.

Theorem 1.7 *Let Q be an equivalent martingale measure for the market described in section 1.4 (that is, such that the discounted price of any traded asset is a Q–martingale). Let N_t be an arbitrary numeraire. Then the price of any traded asset Z_t normalized by N_t is a martingale with respect to the measure Q^N defined by*

$$\frac{dQ^N}{dQ} = \frac{N_T C_0}{N_0 C_T}. \tag{16}$$

Proof: The outline of the proof for this proposition is as follows. Since Z_t is a traded asset, it must be a Q–martingale. Therefore

$$\begin{aligned} \frac{Z_0}{N_0} &= \frac{Z_0}{C_0}\frac{C_0}{N_0} = E^Q\left[\frac{Z_T}{C_T}\frac{C_0}{N_0}\right] \\ &= E^Q\left[\frac{Z_T}{N_T}\frac{dQ^N}{dQ}\right] = E^{Q^N}\left[\frac{Z_T}{N_T}\right] \end{aligned}$$

and the general formula

$$\frac{Z_t}{N_t} = E_t^{Q^N}\left[\frac{Z_T}{N_T}\right] \tag{17}$$

follows from Bayes's rule for conditional expectations.

As an application of this theorem, let us consider, for a fixed maturity T, the bond prices P_{tT} as a numeraire. Then $N_T = P_{TT} = 1$ and the price of a contingent claim B maturing at T reduces to

$$\pi_t^B = P_{tT} E_t^T[B], \tag{18}$$

where $E^T[\cdot]$ denotes expectations with respect to the *forward measure* Q^T defined by

$$\frac{dQ^T}{dQ} = \frac{\exp\left(-\int_0^T r_s ds\right)}{P_{0T}}. \tag{19}$$

In particular, under deterministic interest rates, the forward measure coincides with the equivalent martingale measure Q for all maturities.

Finally, we observe that Q^T is called the forward measure due to the following result.

Proposition 1.8 *The simply compounded forward rate for any period* $[S, T]$ *is a martingale under the* T*–forward measure. In particular*

$$L(t; S, T) = E_t^T[L(S, T)] \tag{20}$$

Proof: From the definition of the simply compounded forward rate for $[S, T]$ we have that

$$L(t; S, T)P_{tT} = \frac{1}{T - t}[P_{tS} - P_{tT}]$$

is the price at time t of a traded asset (why?). It then follows from the definition of Q^T that

$$L(t; S, T) = \frac{L(t; S, T)P_{tT}}{P_{tT}}$$

is a martingale under such measure.

2 Second Lecture

2.1 Pricing in Short Rate Models

Let us consider the economy specified in 1.4 for the special case of $d = 0$, that is, where the only exogenously defined traded asset is the cash account

$$dC_t = r_t C_t dt \tag{21}$$

where the short rate of interest is the solution to

$$dr_t = a(t, r_t)dt + b(t, r_t)dW_t. \tag{22}$$

Assume further, in accordance with 1.1, that there exists an arbitrage–free market for zero-coupon bonds of all maturities $T > 0$. These are treated as derivatives written on the spot rate. Since there are fewer traded assets (besides the risk–free account) than sources of randomness, this market is incomplete. This implies that the zero-coupon bond prices, as well as any other derivative prices, are not uniquely given by arbitrage arguments alone. However, the absence of arbitrage opportunities imposes certain consistence relations on the possible bond prices, which can be derived as follows.

Suppose that the price of a zero-coupon bond with maturity T is given by $P_{tT} = p^T(t, r)$, where p^T is a smooth function of its two variables. It follows from Itö's formula that

$$dp^T = M^T p^T dt + \Sigma^T p^T dW. \tag{23}$$

where

$$M^T p^T = \partial_t p^T + a\partial_r p^T + \frac{1}{2} b^2 \partial^2_{rr} p^T \tag{24}$$

$$\Sigma^T p^T = b\partial_r p^T. \tag{25}$$

Consider now a different maturity date $S < T$, with the corresponding SDE for $P_{tS} = p^S(t, r)$, and suppose we construct a self–financing portfolio consisting of (H^S, H^T) units of the S–bonds and T–bonds, respectively. Then the wealth of the portfolio satisfies

$$\begin{aligned} dX^H &= H^S dp^S + H^T dp^T \\ &= \left[H^S p^S M^S + H^T p^T M^T)\right] dt + \left[H^S p^S \Sigma^S + H^T p^T \Sigma^T\right] dW \end{aligned}$$

Therefore, if we set

$$H^S p^S \Sigma^S + H^T p^T \Sigma^T = 0 \tag{26}$$

or portfolio will be (locally) risk–free. Therefore, in order to avoid arbitrage, its instantaneous rate of return must be the short rate of interest. This leads to

$$\frac{H^S p^S M^S + H^T p^T M^T}{H^S p^S + H^T p^T} = r$$

which upon using (26) and some algebra results in the relation

$$\frac{M^T - r_t}{\Sigma^T} = \frac{M^T - r_t}{\Sigma^T}. \tag{27}$$

We therefore conclude that there exists a process λ, called the *market price of risk*, such that

$$\lambda_t = \frac{M^T - r_t}{\Sigma^T} \tag{28}$$

holds for all t and for every maturity time T. Substituting the expressions (24) and (25) yields that arbitrage-free bond prices p^T satisfy the *term structure equation*

$$\partial_t p^T + [a(t,r) - \lambda(t,r)b(t,r)]\partial_r p^T + \frac{1}{2}b(t,r)^2 \partial^2_{rr} p^T - r p^T = 0, \tag{29}$$

subject to the boundary condition $p^T(T,r) = 1$. From the Feynman-Kac representation, we obtain that

$$p^T = E_t^{Q(\lambda)}[e^{-\int_t^T r_s ds}], \tag{30}$$

for a measure $Q(\lambda)$ with respect to which the dynamics of the short rate is

$$dr_t = [a(t) - \lambda_t b(t)]dt + b(t)dW_t^{Q(\lambda)}. \tag{31}$$

That is, using Girsanov's theorem, we see that the density of the pricing measure $Q(\lambda)$ with respect to the physical measure P is

$$\frac{dQ(\lambda)}{dP} = \exp\left(-\int_0^T \lambda_t dW_t - \frac{1}{2}\int_0^T \lambda_t^2 dt\right). \tag{32}$$

It is easy to generalize both the term structure equation (29) and the expectation formula (30) to incorporate general T-derivatives with pay-offs of the form $\Phi(r_T)$ for a deterministic function Φ.

To summarize, we see that, in order to solve the term structure equation for the prices of interest rate derivatives, we need to specify the P–dynamics of the short rate r_t *and* the market price of risk process $\lambda_t = \lambda(t, r_t)$. This in turn is equivalent to selecting one equivalent martingale measure $Q(\lambda)$ and the Q–dynamics, since interest rate derivatives can the be priced by expectations of their final pay-off with respect to $Q(\lambda)$.

2.2 Affine Models

We say that an interest rate model is *affine* if the zero-coupon bond prices can be written as

$$P_{tT} = \exp[A(t,T) + B(t,T)r_t], \tag{33}$$

for deterministic functions A and B. The following proposition establishes the existence of affine models by exhibiting a sufficient condition on the Q–dynamics for r_t.

Proposition 2.1 *Assume that the Q–dynamics for the short rate r_t is given by*

$$dr_t = a^Q(t, r_t)dt + b(t, r_t)dW_t$$

where the functions a^Q and b are of the form

$$\begin{aligned} a^Q(t,r) &= \kappa(t)r + \eta(t) & (34)\\ b(t,r) &= \sqrt{\gamma(t)r + \delta(t)}. & (35) \end{aligned}$$

Then the model is affine and the functions A and B satisfy the Ricatti equations

$$\begin{aligned} \frac{dB}{dt} &= -\kappa(t)B(t,T) - \frac{1}{2}\gamma B^2(t,T) + 1 & (36)\\ \frac{dA}{dt} &= \eta(t)B(t,T) - \frac{1}{2}\delta(t)B^2(t,T) & (37) \end{aligned}$$

with boundary conditions $B(T,T) = A(T,T) = 0$.

Proof: (i) Calculate the partial derivatives of P_{tT} in affine form and substitute into the term structure equation. (ii) Substitute the functional form for a^Q and b. (iii) Equate the coefficients of the r–term and the term independent of r to zero.

For a partial converse of this result, if we further assume that the Q–dynamics for r_t has time-homogeneous coefficients, then the interest rate model is affine if and only if a^Q and b^2 are themselves affine functions of r_t.

2.3 The Vasicek Model

This was the first one–factor model proposed in the literature in Vasicek (1977). We take the P–dynamics for the short rate of interest to be given by an Ornstein-Uhlenbeck process with constant coefficients, that is,

$$dr_t = \tilde{k}(\tilde{\theta} - r_t)dt + \sigma dW_t. \tag{38}$$

As we have seen above, the complete specification of an interest rate model additionally requires the choice of a market price of risk process. If we want to preserve the functional form for the dynamics of the short rate under the risk neutral measure Q, then we are led to a market price of risk of the form

$$\lambda(t) = \lambda r_t + c, \tag{39}$$

for constants λ and c. Therefore the Q–dynamics for the r_t is given by

$$dr_t = k(\theta - r_t)dt + \sigma dW_t^Q, \tag{40}$$

where $k = \tilde{k} + \lambda\sigma$ and $\theta = \frac{\tilde{k}\tilde{\theta} - \sigma c}{k + \lambda\sigma}$. The explicit solution of this linear SDE is easily found to be

$$r_t = r_0 e^{-kt} + \theta\left(1 - e^{-kt}\right) + \sigma \int_0^t e^{-k(t-s)} dW_s.$$

Therefore,

$$\begin{aligned} E^Q[r_t] &= r_0 e^{-kt} + \theta\left(1 - e^{-kt}\right) \\ \mathrm{Var}^Q[r_t] &= \frac{\sigma^2}{2k}\left[1 - e^{-2kt}\right]. \end{aligned}$$

We see that the Vasicek model gives rise to Gaussian mean-reverting interest rates with long term mean equal to θ and long term variance equal to $\sigma^2/2k$. The main drawback of this model is that it allows for interest rates to become negative. Observe also that the model is affine since $a^Q(t, r) = k\theta - kr$ and $b^2(t, r) = \sigma^2$, so that bond prices can be readily obtained.

Proposition 2.2 *In the Vasicek model, bond prices are given by* $P_{tT} = \exp[A(t,T) + B(t,T)r_t]$ *where*

$$B(t,T) = \frac{1}{k}\left[e^{-k(T-t)} - 1\right] \tag{41}$$

$$A(t,T) = \frac{1}{k^2}\left(\frac{1}{2}\sigma^2 - k^2\theta\right)[B(t,T) + T - t] - \frac{\sigma^2 B^2(t,T)}{4k} \tag{42}$$

Proof: (i) Obtain the Ricatti equations. (ii) Solve the easy linear equation for $B(t,T)$. (iii) Integrate the equation for $A(t,T)$ and substitute the expression obtained for $B(t,T)$.

Explicit formulas for the prices of options on bonds are known for this model (see Jamshidian (1989)).

2.4 The Dothan Model

In order to address the positivity of interest rates, Dothan (1978) introduced a lognormal model for interest rates in which the logarithm of the short rate follows a Brownian motion with constant drift. Let the P–dynamics of the short rate be given by

$$dr_t = \tilde{k} r_t dt + \sigma r_t dW_t, \tag{43}$$

with a market price of risk of the form $\lambda_t = \lambda$, so that its Q–dynamics is

$$dr_t = k r_t dt + \sigma r_t dW_t^Q, \tag{44}$$

with $k = (\tilde{k} - \lambda\sigma)$. It is again easy to see that the explicit solution for this SDE is

$$r_t = r_0 \exp\left[\left(k - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right],$$

so that

$$\begin{aligned} E^Q[r_t] &= r_0 e^{kt} \\ \mathrm{Var}^Q[r_t] &= r_0^2 e^{2kt}\left(e^{\sigma^2 t} - 1\right). \end{aligned}$$

Although this is positive, we can observe that it is mean-reverting if and only if $k < 0$ and that the mean-reversion level is necessarily zero. Observe also that the model is not affine since $b^2(t,r) = \sigma^2 r^2$. However, an explicit (albeit complicated) formula for the prices of zero–coupon bonds is available (see Dothan (1978)). No analytic formulas for options on bonds are available in this model. Moreover, it has the disadvantage that $E[C_t] = \infty$ whenever $k - \frac{1}{2}\sigma^2 > 0$, a complication common to all lognormal models for the short rate of interest.

2.5 The Exponentiated Vasicek Model

Another way of obtaining a lognormal model for interest rates is to suppose that th logarithm of the short rate follows an Ornstein–Uhlenbeck process. That is, we take the P–dynamics for r_t to be

$$dr_t = r_t(\tilde{\theta} - \tilde{k}\log r_t) + \sigma r_t dW, \tag{45}$$

for positive constants k and θ and take the market price of risk to be of the form $\lambda_t = \lambda \log r_t + c$. Then the Q–dynamics of the short rate is

$$dr_t = r_t(\theta - k\log r_t) + \sigma r_t dW \tag{46}$$

with $\theta = (\tilde{\theta} - \sigma c)$ and $k = (\tilde{k} + \lambda\sigma)$. The explicit solution to this SDE can be readily obtained from the solution to the Ornstein–Uhlenbeck SDE, namely

$$r_t = \exp\left[\log r_0 e^{-kt} + \frac{\theta - \sigma^2/2}{k}\left(1 - e^{-kt}\right) + \sigma\int_0^t e^{-k(t-s)}dW_s\right].$$

Moreover,

$$\begin{aligned} E^Q[r_t] &= \exp\left[\log r_0 e^{-kt} + \frac{\theta - \sigma^2/2}{k}\left(1 - e^{-kt}\right) + \frac{\sigma^2}{4k}\left(1 - e^{-2kt}\right)\right] \\ E^Q[r_t^2] &= \exp\left[2\log r_0 e^{-kt} + \frac{2\theta - \sigma^2}{k}\left(1 - e^{-kt}\right) + \frac{\sigma^2}{k}\left(1 - e^{-2kt}\right)\right]. \end{aligned}$$

We therefore see that the exponential Vasicek model mean-reverts to the long term average

$$\lim_{t\to\infty} E^Q[r_t] = \exp\left(\frac{\theta - \sigma^2/2}{k} + \frac{\sigma^2}{4k}\right)$$

with long term variance

$$\lim_{t\to\infty} \mathrm{Var}^Q[r_t] = \exp\left(\frac{2\theta - \sigma^2}{k} + \frac{\sigma^2}{2k}\right)\left[\exp\left(\frac{\sigma^2}{2k}\right) - 1\right].$$

The exponential Vasicek model is not an affine model and does not yield analytic expressions for either zero-coupon bonds or options on them.

2.6 The Cox–Ingersoll-Ross Model

The canonical choice for a positive mean-reverting affine models with time-homogeneous coefficients is the model for which the P–dynamics for the short rate is given by

$$dr_t = \tilde{k}(\tilde{\theta} - r_t)dt + \sigma\sqrt{r_t}dW_t, \tag{47}$$

for positive constants $\tilde{k}, \tilde{\theta}$ and σ satisfying the condition $4\tilde{k}\tilde{\theta} > \sigma^2$. In order to preserve the functional form of this model under the risk neutral measure, we take the market price of risk to be of the form $\lambda_t = \lambda\sqrt{r_t}$, so that its Q–dynamics is

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t^Q, \tag{48}$$

where $k = (\tilde{k} + \lambda\sigma)$ and $k\theta = \tilde{k}\tilde{\theta}$. With a bit of extra work one can deduce that r_t has a non-central chi-squared distribution from which one finds that

$$\begin{aligned} E^Q[r_t] &= r_0 e^{-kt} + \theta\left(1 - e^{-kt}\right) \\ \mathrm{Var}^Q[r_t] &= r_0\frac{\sigma^2}{k}\left(e^{-kt} - e^{-2kt}\right) + \theta\frac{\sigma^2}{2k}\left(1 - e^{-kt}\right)^2. \end{aligned}$$

Most importantly, since $a^Q(t,r) = k(\theta - r)$ and $b^2(t) = \sigma^2 r$, the CIR model is affine. The associated Ricatti equations for the functions $A(t,T)$ and $B(t,T)$ are not immediate to solve, but the following result follows for bond prices.

Proposition 2.3 *In the CIR model, bond prices are given by*

$$P_{tT} = \exp[A(t,T) + B(t,T)r_t]$$

where

$$A(t,T) = \frac{2k\theta}{\sigma^2}\log\left[\frac{2\gamma\exp[(k+\gamma)(T-t)/2]}{2\gamma + (k+\gamma)(\exp[(T-t)\gamma] - 1)}\right] \tag{49}$$

$$B(t,T) = \frac{2(1 - \exp[(T-t)\gamma])}{2\gamma + (k+\gamma)(\exp[(T-t)\gamma] - 1)} \tag{50}$$

and $\gamma^2 = k^2 + 2\sigma^2$.

2.7 Parameter Estimation and the Initial Term Structure

We have specified the short rate dynamics for the previous models under both the physical measure P and the risk neutral measure Q. This is because physical parameters can then be estimated from observations of historical interest rates (or some proxy for them), while risk neutral parameters can be estimated from the market data for derivatives, including bond prices themselves. Ideally, both sets of data will be used for robust estimation and significance tests.

The program of pricing interest rate derivatives can be schematically described as follows. After choosing a particular interest rate model, we obtain the theoretical expression for its initial term structure P_{0T} for all values of T. This depends on a vector parameters (for example (k, θ, σ) for the Vasicek model). We then collect data for the observed initial term structure P^*_{0T} and choose the parameter vector that best fits this empirical term structure. This set of estimated parameters is then used to calculate the prices of more complicated derivatives using either their term structure PDE or taking expectations under the equivalent martingale measure associated with the estimated parameters.

It is then clear that for all of the models discussed so far, which are specified by a finite number of parameters, the best fit described above for the initial term structure will never be a perfect fit, since this would involve solving infinitely many equations (one for each maturity date T) with finitely many unknowns (the parameter vector). Under the framework of short rate models, the only way to obtain a perfect fit for the observed initial term structure is to allow for models which depend on an infinite number of parameters. A natural way to do so is to extend the time-homogeneous models introduced above to models with the same functional form but time-dependent coefficients. Such extensions will be the subject of the next lecture.

3 Third Lecture

3.1 The Ho–Lee Model

In order to address the poor fit for the observed initial term structure P^*_{0T} obtained by the models described above, Ho and Lee (1986) proposed a model in which the initial term structure is given exogenously and evolves in time according to a binomial tree. Its continuous-time limit, derived by Dybvig (1988) and Jamshidian (1988) corresponds to is the following short rate Q–dynamics:

$$dr_t = \Theta(t)dt + \sigma dW_t^Q \tag{51}$$

where $\Theta(t)$ is a function determined by the initial term structure. This is done in the proof of the next proposition, which uses the fact that the model is obviously affine.

Proposition 3.1 *In the Ho–Lee model, bond prices are given by*

$$P_{tT} = \frac{P^*_{0T}}{P^*_{0t}} \exp\left[(T-t)f^*(0,t) - \frac{\sigma^2}{2}t(T-t)^2 - (T-t)r_t\right] \tag{52}$$

where $f^*(0,t)$ denotes the observed initial forward rates.

Proof: (i) Obtain the Ricatti equations and solve in terms of $\Theta(t)$. (ii) Match the initial forward rates obtaining

$$\Theta(t) = \frac{\partial f^*(0,t)}{\partial T} + \sigma^2 t.$$

(iii) Substitute back in the bond price formula.

An explicit formula for call options on bonds under the Ho–Lee model is available and will be discussed later.

3.2 The Hull–White Extended Vasicek Model

Despite the attractive fact of being Gaussian, the Ho–Lee model has the obvious disadvantage of not being mean reverting. To remedy this feature, an extension of the Vasicek model with three time varying parameters was proposed by Hull and White (1990). This allowed not only for the matching of the initial term structure of interest rates, but also of the initial term structure of their volatilities. In what follows we present a simplified version of the model, with only one time varying coefficient. The Q—dynamics for the short rate is given by

$$dr_t = \kappa(\Theta(t) - r_t)dt + \sigma dW_t^Q. \tag{53}$$

Being an affine model, we can use the explicit expressions for the bond prices in order to determine the function Θ.

Proposition 3.2 *In the Hull–White extended Vasicek model, bond prices are given by*

$$P_{tT} = \frac{P_{0T}^*}{P_{0t}^*} \exp\left[-B(t,T)f^*(0,t) - \frac{\sigma^2}{4\kappa}B^2(t,T)(1-e^{-2\kappa t}) + B(t,T)r_t\right]. \tag{54}$$

where

$$B(t,T) = \frac{1}{\kappa}(e^{-\kappa(T-t)} - 1) \tag{55}$$

and $f^*(0,t)$ denotes the observed initial forward rates.

Proof: (i) Obtain the Ricatti equations and solve in terms of $\Theta(t)$. (ii) Match the initial forward rates obtaining

$$\kappa\Theta(t) = \frac{\partial f^*(0,t)}{\partial T} + \kappa f^*(0,t) + \frac{\sigma^2}{2\kappa}(1-e^{-2\kappa t}). \tag{56}$$

(iii) Substitute back in the bond price formula.

You will be asked to integrate (53) in the exercises and see that the solution is a Gaussian process. From there, it is easy to find that the mean and the variance of the short rate in this model are

$$\begin{aligned} E^Q[r_t] &= f^*(0,t) + \frac{\sigma^2}{2\kappa^2}(1-e^{-\kappa t})^2 + \frac{\sigma^2}{2\kappa^2}e^{-kt} \\ \text{Var}^Q[r_t] &= \frac{\sigma^2}{2\kappa}\left[1-e^{-2\kappa t}\right]. \end{aligned}$$

An explicit formula for call options on bonds under the Hull–White model is available and will be discussed later.

3.3 Deterministic–Shift Extensions

A similar extension was proposed by Hull and White (1990) for the CIR model, that is, by taking the coefficients in (47) to be time–dependent. This extension, however, does not lead to general analytic expressions for bond and options prices, since the associated Ricatti equations need to be solved numerically.

Alternatively, we now describe a general method due to Brigo and Mercurio (2001) that produces extensions of any time–homogeneous short rate model in a way that matches the observed initial term structure while preserving the analytic tractability of the original model. Let the original model have the time–homogeneous Q–dynamics

$$dx_t = \mu(x_t)dt + \sigma(x_t)dW_t^Q \tag{57}$$

and consider the function

$$F(t,T,x_t) = E^Q\left[e^{-\int_t^T x_s ds}\right].$$

The deterministic–shift extension consists of defining the instantaneous short rate as

$$r_t = x_t + \phi(t), \tag{58}$$

for a deterministic differentiable function $\phi(t)$, so that

$$dr_t = \left(\frac{d\phi}{dt} + \mu(r_t - \phi(t))\right)dt + \sigma(r_t - \phi(t))dW_t^Q.$$

It therefore follows from proposition 2.1 that, if the original model is affine, then so is the shifted model.

In order to determine the function $\phi(t)$, notice that bond prices are now given by

$$P_{tT} = E^Q\left[e^{-\int_t^T (x_s+\phi(s))ds}\right] = e^{-\int_t^T \phi(s)ds}F(t,T,x_t).$$

Denoting by $f^*(0,t)$ the observed initial term structure, we see that a perfect match is possible if the function ϕ is chosen to be

$$\phi(t) = f^*(0,t) + \frac{\partial \log F(0,t,x_0)}{\partial T}. \tag{59}$$

Inserting this back in the expression for bond prices leads to

$$P_{tT} = \frac{P^*_{0T}F(0,t,x_0)}{P^*_{0t}F(0,T,x_0)}F(t,T,r_t - \phi(t)). \tag{60}$$

Furthermore, whenever the original model has analytic expressions for the price of bond options, a similar argument leads to tractable expressions for these prices under in the shifted model as well.

As an example of this technique we consider the deterministic–shift extension of the CIR model. You will be asked to show in the exercises that the deterministic–shift extension of the Vasicek model is equivalent to the Hull–White extensions discussed before.

3.3.1 The CIR ++ Model

Let the reference model be given by the Q–dynamics

$$dx_t = k(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t \tag{61}$$

with $4k\theta > \sigma^2$ and put $r_t = x_t + \phi(t)$. From proposition 2.3 we know that

$$F^{CIR}(0,T,x_0) = \left(\frac{2\gamma e^{(k+\gamma)T/2}}{2\gamma + (k+\gamma)(e^{T\gamma}-1)}\right)^{\frac{2k\theta}{\sigma^2}} \exp\left[\frac{2(1-e^{T\gamma})x_0}{2\gamma + (k+\gamma)(e^{T\gamma}-1)}\right], \tag{62}$$

where $\gamma = \sqrt{k + 2\sigma^2}$. According to (59) leads to

$$\begin{aligned} \phi^{CIR}(t) &= f^*(0,t) - \frac{2k\theta(e^{t\gamma}-1)}{2\gamma + (k+\gamma)(e^{t\gamma}-1)} \\ &+ x_0 \frac{4\gamma^2 e^{t\gamma}}{[2\gamma + (k+\gamma)(e^{t\gamma}-1)]^2} \end{aligned} \tag{63}$$

and we can use (60) to obtain a closed-form expression for bond prices. Since the reference CIR model gives rise to analytic expressions for options on bonds, the same is also true for the CIR++ model.

We see from (63) that the positivity of interest rates in the CIR++ model depends not only on the parameters of the original model, but also on how they relate to the observed forward rates $f^*(0,t)$.

3.4 Forward Rate Models

Instead of modeling the short rate process r_t, Heath, Jarrow and Morton (1992) proposed to take the entire forward rate curve $f(t,T)$ as the infinite–dimensional state variable underlying the model. It is assumed that, for each fixed maturity date $T > 0$, the forward rate follows the P–dynamics

$$df_{tT} = \alpha(t,T)dt + \sigma(t,T)dW_t, \tag{64}$$

with $f(0,T) = f^*(0,T$, where $f^*(0,T)$ denotes the observed initial forward rate curve. The complete set of solutions to the equations above is therefore equivalent to specifying the entire term structure of bond prices P_{tT}. Since this leads to a market with many more assets than sources of randomness, we need to verify that the market is arbitrage free. This is secured by the HJM drift condition expressed in the next theorem.

Theorem 3.3 *If the bond market is free from arbitrage, then there exist a process λ_t with the property that*

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds - \sigma(t,T)\lambda_t. \tag{65}$$

Proof: (i) Use the third part of proposition 1.5 to obtain expressions for the $M(t,T)$ and $\Sigma(t,T)$ in terms of $\alpha(t,T)$ and $\sigma(t,T)$. Substitute this into the expression (28) for the market price of risk.

As a corollary, if the forward rates are modeled under the risk neutral measure through the Q–dynamics

$$df_{tT} = \alpha^Q(t,T)dt + \sigma^Q(t,T)dW_t, \tag{66}$$

we see that the HJM drift condition reduces to

$$\alpha^Q(t,T) = \sigma^Q(t,T)\int_t^T \sigma^Q(t,s)ds. \tag{67}$$

In the simplest example, one can take $\sigma(t,T) = \sigma$ to be constant. This to the Ho–Lee model discussed before.

4 Fourth Lecture

4.1 Some interest rates derivatives

In this section we describe several examples of interest rate derivatives and provide simple arbitrage arguments to establish relations between them. The actual pricing of these derivatives within the martingale methodology is going to be address in the next section.

Recall that the simplest possible interest rate derivative is a zero-coupon bond, whose price we denote by P_{tT}. In practice, most bonds trade are *coupon-bearing bonds*, that is, they pay specific amounts $c = (c_1, \ldots, c_K)$ at the times $\mathcal{T} = (T_1, \ldots, T_K)$. If the payments are deterministic it is clear that these bonds can be replicated by c_k units of zero-coupon bonds with maturities T_k, $k = 1, \ldots, K$. Therefore, their price at time t is given by

$$P(t, c, \mathcal{T}) = \sum_{k=1}^{K} c_k P_{tT_K}. \tag{68}$$

More generally, the coupon payments c_k are not deterministic, but rather specified by the value of a financial benchmark at the payment dates $\mathcal{T} = (T_1, \ldots, T_K)$. The most common of such bonds is a *floating-rate note*, for which the payment stream is given by

$$c_k = L(T_{k-1}, T_k)(T_k - T_{k-1})\mathcal{N}, \quad k = 2, \ldots, K-1,$$

and $c_K = \mathcal{N}$, where $L(T_{k-1}, T_k)$ is a simply compounded rate for the period $[T_{k-1}, T_k]$ (such as the LIBOR rate) and $\mathcal{N}$ denotes a fixed notional value. Recalling (4), we obtain that

$$c_k = \frac{\mathcal{N}}{P_{T_{k-1}T_k}} - \mathcal{N}.$$

Now observe that we can replicate the first term in the payment above by, at time t, buying $\mathcal{N}$ zero-coupon bonds with maturity T_{k-1}, then using the $\mathcal{N}$ units of currency obtained at time T_{k-1} to buy $\mathcal{N}/P_{T_{k-1}T_K}$ bonds with maturity T_k. Therefore the value at time $t < T_1$ of the payment c_k is

$$\mathcal{N}(P_{tT_{k-1}} - P_{tT_k}),$$

implying that the value of the floating rate note is

$$P(t,\mathcal{N},\mathcal{T}) = \mathcal{N}\left[P_{tT_K} + \sum_{k=2}^{K}(P_{tT_{k-1}} - P_{tT_k})\right] = \mathcal{N}P_{tT_1}. \tag{69}$$

In the practitioners' jargon this expressed by saying that a floating-rate note always "trades at par".

Our next example is a *forward rate agreement*, which gives its holder the payment of a fixed simple compounded interest rate L^* for the period $[S,T]$ on a notional $\mathcal{N}$ in exchange of the (stochastic) simply compounded spot rate $L(S,T)$ for the same period on the same notional. In other words, the cash flow at maturity T for the holder of a forward rate agreement is

$$\mathcal{N}[L^* - L(S,T)](T-S),$$

which at time $t < S < T$ has value

$$F(t,S,T,L^*,\mathcal{N}) = \mathcal{N}[P_{tT}L^*(T-S) - P_{tS} + P_{tT}]. \tag{70}$$

We then see that the fixed rate that makes this contract cost zero at time t is

$$L^* = \frac{P_{tS} - P_{tT}}{P_{tT}(T-S)} = L(t,S,T),$$

which serves as an alternative definition of the simply compounded forward rate $L(t,S,T)$. An entirely analogous definition exists for a forward rate agreement based on continuously compounded interest rates.

A generalization of forward rate agreements for many periods is what is known generically as an *interest rate swap*, whereby a payment stream based on a fixed *swap rate* L^* for a notional N is made at dates $\mathcal{T} = (T_2,\ldots,T_K)$ in return of a payment stream based on a floating rate for the notional and the same periods. At time T_k, for $k = 2,\ldots,K$, the cash flow for the holder of an interest rate swap is

$$c_k - \mathcal{N}L^*(T_k - T_{k-1}),$$

where c_k of a floating rate note considered before. Using our previous result, the value at $t < T_1$ of this cash flow is

$$\mathcal{N}(P_{tT_{k-1}} - P_{tT_k}) - \mathcal{N}L^*(T_k - T_{k-1})P_{tT_k},$$

so that the total value of the interest rate swap at time $t < T_1$ is

$$\mathrm{IRS}(t, \mathcal{N}, \mathcal{T}, L^*) = \mathcal{N}\left[P_{tT_1} - P_{tT_K} - L^* \sum_{k=2}^{K} P_{tT_k}(T_k - T_{k-1})\right]. \tag{71}$$

Similarly to a forward rate agreement, we can define the swap rate at time t as the rate that makes this contract cost zero, that is

$$L^*(t, \mathcal{T}) = \frac{P_{tT_1} - P_{tT_K}}{\sum_{k=2}^{K} P_{tT_k}(T_k - T_{k-1})}. \tag{72}$$

Next let us introduce options on zero-coupon bonds. An European *call option* with strike price $\mathcal{K}$ on an underlying T-bond is defined by the pay-off

$$(P_{ST} - \mathcal{K})^+$$

at the exercise date $S < T$. Its value at time $t < S < T$ is denoted by $c(t, S, \mathcal{K}, T)$. Similarly, an European *put option* with the same parameters has value at time $t < S < T$ denoted by $p(t, S, \mathcal{K}, T)$ and is defined by the pay-off

$$(\mathcal{K} - P_{ST})^+.$$

These basic vanilla options can be used to analyze more complicated interest rate derivatives. For example, a *caplet* for the interval $[S, T]$ with *cap rate* $\mathcal{R}$ on a notional $\mathcal{N}$ is defined as a contingent claim with a pay-off

$$\mathcal{N}(T - S))[L(S, T) - \mathcal{R}]^+$$

at time T. The holder such caplet is therefore buying protection against an increase in the floating rates above the cap rate. Using (4), this pay-off can be expressed as

$$\mathcal{N}\left(\frac{1 + \mathcal{R}(T - S)}{P_{ST}}\right)\left[\frac{1}{1 + \mathcal{R}(T - S)} - P_{ST}\right]^+,$$

which is therefore equivalent to $\mathcal{N}[1 + \mathcal{R}(T - S)]$ units of an European put option with strike $\mathcal{K} = \frac{1}{1+\mathcal{R}(T-S)}$ and exercise date S on the underlying T-bond.

A *cap* for the dates $\mathcal{T} = (T_1, \ldots, T_K)$ with notional $\mathcal{N}$ and cap rate $\mathcal{R}$ is defined as the sum of the caplets over the intervals $[T_{k-1}, T_k]$, $k = 2, \ldots, K$, with the same notional and cap rate. Therefore, the value of a cap at time $t < T_1$ is given by

$$\text{Cap}(t, \mathcal{T}, \mathcal{N}, \mathcal{R}) = \mathcal{N}\sum_{k=2}^{K}[1+\mathcal{R}(T_k-T_{k-1})]p\left(t, T_{k-1}, \frac{1}{1+\mathcal{R}(T_k-T_{k-1})}, T_k\right).$$

Similarly, a *floor* for the dates $\mathcal{T} = (T_1, \ldots, T_K)$ with notional $\mathcal{N}$ and *floor rate* $\mathcal{R}$ is defined as the sum of *floorlets* over the intervals $[T_{k-1}, T_k]$, $k = 2, \ldots, K$, each with a pay-off

$$\mathcal{N}(T_k - T_{k-1}))[\mathcal{R} - L(T_{k-1}, T_k)]^+$$

at times T_k. An analogous calculation then shows that each floorlet is equivalent to a call option with exercise date T_{k-1} on a bond with maturity T_k. Therefore the value of a floor at time $t < T_1$ is

$$\text{Flr}(t, \mathcal{T}, \mathcal{N}, \mathcal{R}) = \mathcal{N}\sum_{k=2}^{K}[1+\mathcal{R}(T_k-T_{k-1})]c\left(t, T_{k-1}, \frac{1}{1+\mathcal{R}(T_k-T_{k-1})}, T_k\right).$$

4.2 General Option Pricing Formulas

Consider the market of section 1.4 with $d = 1$ and a stochastic interest rate, that is

$$\begin{aligned} dS_t &= \mu(t, S_t)S_t + \sigma(t, S_t)S_t dW_t^1 \\ dr_t &= a(t, r_t)dt + b(t, r_t)[\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2 \\ dC_t &= r_t C_t dt, \end{aligned} \tag{73}$$

where $W_t = (W_t^1, W_t^2)$ is a standard two-dimensional P–Brownian motion.

Recall that given any numeraire N_t (i.e. any strictly positive trade asset), there exist a measure Q^N such that the prices of any other trade assets in terms of the numeraire N_t are martingales with respect to Q^N. In particular, the price of a derivative B is given by

$$\pi(t) = N_t E_t^{Q^N}\left[\frac{B}{N_T}\right] \tag{74}$$

In this context, let Q^S and Q^T denote the measures obtained using (16) for the numeraires S_t and P_{tT}, respectively.

Consider first an equity call option with strike K and maturity T with pay-off written as

$$(S_T - K)^+ = (S_T - K)\mathbf{1}_{\{S_T \geq K\}},$$

where $\mathbf{1}_A$ denotes the indicator function of the set A. We can now apply (74) to each term in the pay-off above separately, obtaining the initial price of the call option

$$c_0 = S_0 Q^S(S_T \geq K) - K P_{0T} Q^T(S_T \geq K). \tag{75}$$

In order to calculate the expectations above, let us define the process $Z_t = S_t/P_{tT}$. Since this is the price of a trade asset in terms of the numeraire P_{tT}, it must be a positive martingale under the measure Q^T. We can therefore express its Q^T–dynamics as

$$dZ_t = Z_t \sigma_{S,T}(t) dW_t^T,$$

where W_t^T is a two-dimensional Q^T–Brownian motion. The solution to this equation is

$$Z_T = \frac{S_0}{P_{0T}} \exp\left(-\frac{1}{2}\int_0^T |\sigma_{S,T}(t)|^2 dt + \int_0^T \sigma_{S,T}(t) dW^T\right).$$

Similarly, define $Y_t = Z_t^{-1} = S_t/P_{tT}$, which must then be a positive martingale under the measure Q^S. It then follows from Itô's formula that

$$dY_t = -Y_t \sigma_{S,T}(t) dW_t^S,$$

for a two-dimensional Q^S–Brownian motion W_t^S, whose solution is

$$Y_T = \frac{P_{0T}}{S_0} \exp\left(-\frac{1}{2}\int_0^T |\sigma_{S,T}(t)|^2 dt + \int_0^T \sigma_{S,T}(t) dW^S\right).$$

If we now make the crucial assumption that the row vector $\sigma_{S,T}(t)$ is a *deterministic* function of time, then the stochastic integrals above are normally distributed under the respective measures, with variance

$$\mathrm{Var(T)} = \int_0^T |\sigma_{S,T}(t)|^2 dt.$$

Returning to (75), we obtain the following generalization of the Black–Scholes option price formula

$$\begin{aligned} c_0 &= S_0 Q^S(Y_T \leq 1/K) - K P_{0T} Q^T(Z_T \geq K) \\ &= S_0 N[d_1] - K P_{0T} N[d_2], \end{aligned} \tag{76}$$

where $N[\cdot]$ denotes the cumulative standard normal distribution and

$$d_1 = d_2 + \sqrt{\mathrm{Var}(T)} \tag{77}$$

$$d_2 = \frac{\log\left(\frac{S_0}{K P_{0T}}\right) - \frac{1}{2}\mathrm{Var}(T)}{\sqrt{\mathrm{Var}(T)}}. \tag{78}$$

Consider now a call option with strike price K and exercise date T_1 on an underlying bond P_{tT_2} with $T_1 < T_2$. We can use the formulation above to price this contract by identifying $S_t = P_{tT_2}$. All we need to check is if the volatility parameter of the process $Z_t = P_{tT_2}/P_{tT_1}$ is deterministic. For this purpose let us assume that the interest rate follows the Hull-White extended Vasicek model (53) with the function $\Theta(t)$ given by (56). Then, this being an affine model, bond prices of any maturity are given by

$$P_{tT} = \exp[A(t,T) + B(t,T) r_t],$$

where, in particular, $B(t,T)$ is given by (55). It then follows by Itô's formula that the Q–dynamics of the process Z_t is

$$dZ_t = Z(t)\mu^Z(t) + Z(t)\sigma^Z(t) dW_t^Q,$$

where

$$\sigma^Z(t) = [B(t,T_2) - B(t,T_1)]\sigma = \frac{\sigma}{\kappa} e^{\kappa t}\left[e^{-\kappa T_1} - e^{-\kappa T_2}\right], \tag{79}$$

which is indeed deterministic.

Applying (76) for this option we obtain

$$c_0^{HW}(t, T_1, K, T_2) = P_{0T_2} N[d_1] - K P_{0T_1} N[d_2], \tag{80}$$

where $N[\cdot]$ denotes the cumulative standard normal distribution and

$$d_1 = d_2 + \sqrt{\Sigma^2} \tag{81}$$

$$d_2 = \frac{\log\left(\frac{P_{0T_2}}{K P_{0T_1}}\right) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}}. \tag{82}$$

$$\Sigma^2 = \frac{\sigma^2}{2\kappa^3}\left[1 - e^{-2\kappa T_1}\right]\left[1 - e^{-\kappa(T_2 - T_1)}\right]^2. \tag{83}$$

5 Fifth Lecture

5.1 Default Events

We will now start the part of the course dealing with credit risk. Broadly speaking, credit risk concerns the possibility of financial losses due to changes in the credit quality of market participants. The most radical change in credit quality is a *default event*. As we will soon discover, the very definition of what constitutes a default event is model dependent, so it is relatively pointless to spend much time trying to put forward its precise properties. Suffice to have in mind the basic idea that a default event is a *rare* occurrence taking place at a *random* time and resulting in *large* financial losses to *some* sectors of the market.

Regardless of what definition is used for a default event, let us denote the *default time* by τ. The only mathematical structure assumed for τ is that it should be a *stopping time*, that is, a random variable $\tau : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ such that $\{\tau \leq t\} \in \mathcal{F}_t$, for every $t \geq 0$. In other words, a random time τ is a stopping time if the stochastic process

$$N_t(\omega) = \mathbf{1}_{\{\tau \leq t\}}(\omega) = \begin{cases} 1, & \text{if } \tau(\omega) \leq t \\ 0, & \text{otherwise} \end{cases} \tag{84}$$

is adapted. For default times, this is known as the *default indicator process.*

While we are on the subject, let us say that a stopping time $\tau > 0$ is *predictable* if there is an *announcing sequence* of stopping times $\tau_1 \leq \tau_2 \leq \cdots$ such that

$$\lim_{n\to\infty} \tau_n = \tau, \qquad \text{P-a.s}$$

If you have studied stochastic analysis you will recognize this as the statement that the indicator process N_t is predictable, but that is not necessary for the development of this notes. The opposite of a predictable stopping time is a *totally inaccessible* stopping time, that is, a stopping time τ such that

$$P[\tau = \hat{\tau} < \infty] = 0,$$

for any predictable stopping time $\hat{\tau}$. If you have studied stochastic analysis you will recognize this asIt can be shown that every stopping time can be decomposed into the sum of a predictable and a totally inaccessible stopping times.

Given a default time τ, define the *probability of survival* in t years is as

$$P[\tau > t] = 1 - P[\tau \leq t] = 1 - E[\mathbf{1}_{\{\tau \leq t\}}]. \tag{85}$$

Several other related quantities can be derived from this basic probability. For instance,

$$P[s \leq \tau \leq t] = P[\tau > s] - P[\tau > t]$$

is the unconditional probability of default occurring in the time interval $[s, t]$.

Using Bayes's rule for conditional probability, one can deduce that the probability of survival in t years conditioned on survival up to $s \leq t$ years is

$$P[\tau > t | \tau > s] = \frac{P[\{\tau > t\} \cap \{\tau > s\}]}{P[\tau > s]} = \frac{P[\tau > t]}{P[\tau > s]}, \tag{86}$$

since $\{\tau > t\} \subset \{\tau > s\}$. From this we can define the *forward default probability* for the interval $[s, t]$ as

$$P[s \leq \tau \leq t | \tau > s] = 1 - P[\tau > t | \tau > s] = 1 - \frac{P[\tau > t]}{P[\tau > s]}. \tag{87}$$

Assuming that $P[\tau > t]$ is strictly positive and differentiable in t, we define the *hazard rate* process as

$$h(t) = -\frac{\partial \log P[\tau > t]}{\partial t}. \tag{88}$$

It then follows that

$$P[\tau > t | \tau > s] = e^{-\int_s^t h(u)du}. \tag{89}$$

The forward default rate measures the instantaneous rate of arrival for a default event at time t conditioned on survival up to t. Indeed, if $h(t)$ is continuous we find that

$$h(t)\Delta t \approx P[t \leq \tau \leq t + \Delta t | \tau > t].$$

More generally, one can focused on $P[\tau > t | \mathcal{F}_s]$, that is, the survival probability in t years conditioned on all the information available at time $s \leq t$. If we assume positivity and differentiability in t, then this can be written as

$$P[\tau > t | \mathcal{F}_s] = e^{-\int_s^t f(s,u)du}, \tag{90}$$

where

$$f(s,t) = -\frac{\partial \log P[\tau > t|\mathcal{F}_s]}{\partial t} \quad (91)$$

is the *forward default rate* given all the information up to time s.

The indicator process N_t defined in (84) is clearly a submartingale. Moreover, it can be shown that it is of class $\mathcal{D}$, so that it follow from the Doob-Meyer decomposition that there exist an increasing predictable process Λ_t, called the *compensator*, such that $N_t - \Lambda_t$ is a martingale. If the compensator can be written as

$$\Lambda_t = \int_0^t \lambda(s)ds \quad (92)$$

for a non-negative, progressively measurable process $\lambda(t)$, then this process is called the *default intensity*. As we will see later, although all default indicators have a compensator, not all of them admit a density. This happens, for example, whenever the stopping time is predictable.

Under suitable technical conditions, it follows that

$$\lambda(t) = f(t,t) \quad (93)$$

Therefore, while the hazard rate $h(t)$ gives the instantaneous rate of default conditioned on survival up to t, the intensity measures the instantaneous rate of default conditioned on all the information available up to time t.

Finally, we observe that starting from a sufficiently regular family of survival probabilities one can obtain forward default rates by (91) and the associated intensity by (93). This is analogous to knowing a differentiable system of bond prices and then obtaining forward and spot interest rates from it. As we have seen, going in the opposite direction, that is from the spot interest r_t to bond prices, is not always straigtforward, and the exact same is true for going from intensities to survival probabilities.

5.2 Structural Models for Default Probabilities

Under structural models, a default event is deemed to occur for a firm when its assets reach a sufficiently low level compared to its liabilities. They require strong assumptions on the dynamics of the firm's asset, its debt and how its capital is structured. The main advantage of structural models is that they provide an intuitive picture, as well as an endogenous explanation, for default. We will discuss other advantages and some of their disadvantages in what follows.

5.2.1 The Merton Model (1974)

Assume that the total value A_t of a firm's assets follows a geometric Brownian motion

$$dA_t = (\mu - \delta)A_t + \sigma dW_t, \quad A_0 > 0, \tag{94}$$

where μ is the mean rate of return on the assets, δ is a proportional cash pay-out rate and σ is the asset volatility.

The firm is funded by shares and debt. Assume that the it has a single outstanding bond with face value K and maturity T. At maturity, if the total value of the assets is greater than the debt, the latter is paid in full and the remaining is distributed among shareholders. However, if $A_T < K$ then default is deemed to occur because bondholders exercise a debt covenants giving them the right to liquidate the firm and receive the liquidation value in lieu of the debt. Shareholders receive nothing in this case, but are not required to inject any additional funds to pay for the debt, in which is called limited liability. Therefore shareholders have a cash flow at T equal to

$$(A_T - K)^+,$$

so that equity can be view as an European call option on the firm's assets and its value E_t at earlier times $t < T$ can be calculated using the Black-Scholes formula. Note that equity value increases with the firm's volatility, so shareholders are generally inclined to press for riskier positions to be taken by their managers.

Bond holders, on the other hand, receive

$$\min(K, A_T) = K - (K - A)^+.$$

Therefore the value D_t for the debt at earlier times $t < T$ can be obtained as the value of a zero-coupon bond minus an European put option. It follows from the put-call parity relation that

$$A_t = E_t + D_t,$$

which is the fundamental identity of accounting (and also an instance of the Modigliani-Miller theorem).

Under this model, the default time is the random variable

$$\tau = \begin{cases} T, & \text{if } A_T < K \\ \infty, & \text{otherwise} . \end{cases}$$

Setting $m = \mu - \delta - \sigma^2/2$ we can readily obtain from Itô's formula that the probability of default at time T is given by

$$\begin{aligned} P[\tau = T] = P[A_T < K] &= P\left[A_0 \exp(mT + \sigma W_T) < K\right] \\ &= P\left[W_T < \frac{\log(D/A_0) - mT}{\sigma}\right] \\ &= N\left[\frac{\log(D/A_0) - mT}{\sigma\sqrt{T}}\right]. \end{aligned} \tag{95}$$

5.2.2 First passage models

Consider again a firm with asset value given by (94) and outstanding debt with face value K at maturity T. Instead of having the possibility of default only at maturity time T, Black and Cox (1976) postulated that default occurs at the first time that the the firm's asset value drop below a certain time-dependent barrier $K(t)$. That is, the default time is given by

$$\tau = \inf\{t > 0 : A_t < K_t\} \tag{96}$$

For the choice of the time dependent barrier, observe that is $K_t > K$ then bondholders are always completely covered, which is certainly unrealistic. On the other hand, one should clearly have $K_T = K$ for a consistent definition of default. A natural choice is to take the time-dependent barrier to be $K(t) = Ke^{-k(T-t)}$, that is, the face value discounted by a constant rate k.

Observing that

$$\{A_t < K(t)\} = \{(m-k)t + \sigma W_t < \log(K/A_0) - kT\},$$

we obtain that the probability of default occurring before time T is then given by

$$\begin{aligned} P[0 \leq \tau \leq T] &= P\left[\min_{t \leq T} A_t < K(t)\right] \\ &= P\left[\min_{t \leq T} \left[(m-k)t + \sigma W_t\right] < \log\left(\frac{K}{A_0}\right) - kT\right] \\ &= N\left(\frac{\log(K/A_0) - mT}{\sigma\sqrt{T}}\right) \\ &\quad + \left(\frac{Ke^{-kT}}{A_0}\right)^{\frac{2(m-k)}{\sigma^2}} N\left(\frac{\log(K/A_0) + (m-2k)T}{\sigma\sqrt{T}}\right) \end{aligned}$$

The pay-off for equity holders at maturity is $(A_T - K)^+ \mathbf{1}_{M_T > K(t)}$ where $M_t = \min_{s \leq t} A_s$ denotes the running minimum of the diffusion A_t. Its value is then given by the price of a down-and-out call option with a moving barrier $K(t)$, for which closed form expressions are available (Merton 76). This is smaller than the share value obtained in the Merton model, and is not monotone in the volatility. The pay-off for debt holders is $K - (K - A_T)^+ + (A_T - K)\mathbf{1}_{M_T \leq K(t)}$. Its value is then given by the price of a zero-coupon bond minus a vanilla put option plus a down-and-in call option. This is worth at least as much as the debt in Merton's model.

5.2.3 Excursion Models

Instead of setting the default time as the firm's asset hitting time for a barrier K_t, one can allow for the asset value to have "excursions" in the region under the barrier and set a default to occur if the time spent in that region is sufficiently long. For a concrete example, consider a fixed barrier K and define the total excursion time under the barrier as

$$\mathcal{T}(t) = \int_0^t \mathbf{1}_{\{A_s \leq K\}} ds. \tag{97}$$

Then, fixing a maximum excursion time $\mathcal{T}^* \geq 0$, the default time can be specified as $\tau = \min(\tau_1, \tau_2)$ where

$$\tau_1 = \inf\{t > 0 : \mathcal{T}(t) > \mathcal{T}^*\}$$

and

$$\tau_2 = \begin{cases} T, & \text{if } A_T < K \\ \infty, & \text{otherwise .} \end{cases}$$

Therefore, the probability of default is

$$P[0 < \tau \leq T] = 1 - P[\tau_1 > T, \tau_2 > T],$$

and can be calculate from the joint probability of $(A_t, \mathcal{T}(t))$. This is one of the explicit distribution known for functionals of Brownian motion (see Borodin and Salminen 96).

The pay-off for shareholders at maturity is $(A_T - K)^+ \mathbf{1}_{\{\tau_1 > T\}}$, so the equity value before maturity is an example of what is called occupation-time derivatives (Hugonnier 99) .

It is possible to modify the excursion-time functional in order to consider only the time spent under the barrier after the last crossing or to take into account the firm's cumulative shortfall.

6 Sixth Lecture

We will now step back from particular models and define some general concepts in credit risk, namely implied survival probabilities and credit spreads. Later in the course we will see how these concepts apply to different models, starting with structural models in the second half of this lecture.

6.1 Implied Survival Probabilities

Let $\bar{P}_{tT}\mathbf{1}_{\{\tau>t\}}$ be the price at time $t \leq T$ of a defaultable zero-coupon bond issued by a certain company with maturity T and face value equal to one unit of currency . Then clearly $\bar{P}_{tT} > 0$ denotes the price of this bond given that the company has survived up to time t. Since

$$\bar{P}_{TT}\mathbf{1}_{\{\tau>T\}} = \mathbf{1}_{\{\tau>T\}} \leq 1 = P_{TT},$$

the Law of one Price dictates that

$$\bar{P}_{tT}\mathbf{1}_{\{\tau>t\}} \leq P_{tT}$$

for all earlier times t.

Moreover, under a risk-neutral measure Q, we know that

$$\bar{P}_{tT}\mathbf{1}_{\{\tau>t\}} = E_t^Q\left[e^{-\int_t^T r_s ds}\mathbf{1}_{\{\tau>T\}}\right] \tag{98}$$

If we assume that r_t and τ are *independent* under the risk-neutral measure Q, then the last equation can be rewritten as

$$\bar{P}_{tT}\mathbf{1}_{\{\tau>t\}} = E_t^Q\left[e^{-\int_t^T r_s ds}\right]E_t^Q[\mathbf{1}_{\{\tau>T\}}].$$

Therefore, assuming that $\{\tau > t\}$, the risk-neutral survival probability is given by

$$Q[\tau > T|\mathcal{F}_t] = \frac{\bar{P}_{tT}}{P_{tT}}. \tag{99}$$

That is, under the independence assumption for r_t and τ, the term structure of risk-neutral survival probabilities is completely determined by the term structure of both defaultable and default-free zero-coupon bonds. In the sequel, these will be called *implied survival probability*, emphasizing the fact that they are derived from market prices and associated to the risk-neutral measure Q.

Assuming differentiability with respect to the maturity date, we can define the *implied forward default rate* by

$$f^Q(t,s) = -\frac{\partial \log Q[\tau > s|\mathcal{F}_t]}{\partial s}, \tag{100}$$

so that

$$Q[\tau > s|\mathcal{F}_t] = e^{-\int_t^s f(t,u)du}, \tag{101}$$

It is reasonable to assume that the prices of defaultable bonds show a sharper decrease as a function of maturity than do prices of default-free bonds. This corresponds to decreasing implied survival probabilities or, equivalently, positive implied forward default rates. Moreover, we expect default to eventually occur. Therefore, the term structure of implied survival probabilities as functions of the maturity date share the properties of the term structure of bond prices, namely, initial value equal to 1, decreasing and approaching zero at infinity.

6.2 Credit Spreads

Assume that $\{\tau > t\}$ and that interest rates are independent of default events. A credit spread is the difference between rates for defaultable bonds and the corresponding default-free bonds. For example, the yield spread for maturity T is given by

$$\mathrm{YS}(t,T) = \frac{1}{T-t}[\bar{R}(t,T) - R(t,T)] = \frac{1}{T-t}\log\left(\frac{P_{tT}}{\bar{P}_{tT}}\right), \tag{102}$$

where yields are defined by (3). Using (99) for implied survival probabilities and (101) for the implied forward default rate, we obtain that

$$\begin{aligned} \mathrm{YS}(t,T)(T-t) &= -\log(Q[\tau > T|\mathcal{F}_t]) \\ &= \int_t^T f^Q(t,s)ds \end{aligned} \tag{103}$$

Similarly, for the forward rate spread we find

$$\begin{aligned} \text{FS(t,T)} &= \bar{f}(t,T) - f(t,T) \\ &= \frac{\partial}{\partial T} \log \left(\frac{P_{tT}}{\bar{P}_{tT}} \right) \\ &= -\frac{\partial \log Q[\tau > T|\mathcal{F}_t]}{\partial T} \\ &= f^Q(t,T). \end{aligned} \tag{104}$$

In particular

$$Q[t < \tau \le t + \Delta t|\mathcal{F}_t] \approx \bar{r}_t - r_t =: \lambda^Q(t) \tag{105}$$

That is, credit spreads are a direct measure of implied default rates.

6.3 Properties of Structural Models

In what follows, we consider some general properties of structural models regarding credit spreads, risk premium and calibration issues. To facilitate the discussion, we will assume that the risk-free interest rate is a constant r. Extensions to stochastic interest rates will be consider in later sections.

6.3.1 Credit Spreads

In the classical Merton model, we have seen that the price at time zero of a defaultable bond with face vale K and maturity T is given by $D_0 = A_0 - E_0$, where the equity value E_0 is calculated as the price of a call option with maturity T and strike price K. Using (76) for option prices with stochastic interest rates, we find

$$D_0 = A_0 - A_0 N[d_1] + K P_{0T} N[d_2], \tag{106}$$

where d_1 and d_2 are given by (77) and (78). Therefore, the term structure for this firm's yield spread is

$$\begin{aligned} \text{YS}(0,T) &= -\frac{1}{T} \log \left(\frac{A_0(1 - N[d_1]) + K P_{0T} N[d_2]}{K P_{0T}} \right) \\ &= -\frac{1}{T} \log \left(\frac{A_0}{K} \frac{1}{P_{0T}} N[-d_1] + N[d_2] \right) \end{aligned} \tag{107}$$

To analyze the qualitative behaviour of this term structure for very short maturity times, let us assume that the risk-free interest rate is a constant r and that $A_0 > K$. Then one finds that credit spreads for this model start at zero for $T = 0$, then increase sharply to a maximum and start to decrease to a positive plateau. This is in accordance with the diffusion character of the model. For very short maturity times, the asset price diffusion will almost surely never cross the default barrier. The probability of default then increases for longer maturities but start to decrease again as the geometric Brownian motion drifts away from the barrier.

This behaviour is also observed in first passage and excursion models, except that spreads exhibit a faster decrease for longer maturities. It is at odds with empirical observations in two respects: (i) observed spreads remain positive even for small time horizons and (ii) tend to increase as the time horizon increases. The first feature follows from the fact that there is always a small probability of immediate default. The second is a consequence of greater uncertainty for longer time horizons. One of the main reasons to study reduced-form models is that, as we will see, they do not give rise to such discrepancies.

6.3.2 Risk Premium

For constant interest rate, the market involving equity and debt for a firm whose assets evolve according to (94) is complete. That is, the default risk on the bond can be completely hedge by taking positions on equity. We will discuss this kind of delta hedge later in connection with pricing derivatives under different credit risk model. Here we just want to observe that, as a consequence of market completeness, the equivalent martingale measure Q is unique, and can be found from the Radon-Nikodym derivative

$$\frac{dQ}{dP} = \exp\left(-\phi W_T - \phi^2 T\right), \tag{108}$$

where

$$\phi = \frac{\mu - r}{\sigma} \tag{109}$$

is the market price of default risk in the model. Therefore, the asset dynamics under the risk neutral measure is

$$dA_t = (r - \delta)A_t + \sigma dW_t^Q, \quad A_0 > 0, \tag{110}$$

where $dW^Q = dW + \phi dt$ is a standard Q–Brownian motion.

It follows that all the previously obtained formulas for historical default probabilities (that is under the measure P) can be used to calculate implied default probabilities (that is under the risk neutral measure Q) provided we replace μ by the risk-free rate r.

For example, for the Merton model one finds that

$$Q[\tau > T] = N\left(N^{-1}(P[\tau > T]) - \phi\sqrt{T}\right).$$

That is, as long as $\phi > 0$ we have that implied default probabilities are always higher that historical ones.

6.3.3 Calibration

One of the main drawbacks of structural models is that they depend on the unobserved variable A_t. In this section we survey the methods for estimating the model parameters from market data. For publicly trade firms, a readily available information is equity, which can be obtained by simply multiplying the number os shares by the stock price at any given time. Suppose that the total value of a firm's equity follows the dynamics

$$dE_t = \mu^E E_t dt + \sigma^E E_t dW_t, \tag{111}$$

for constants μ^E and σ^E. Given a time series for E_t, the equity volatility σ^E can be estimated either by implied volatility of option prices on the firm's stock or by the empirical standard deviation of the equity log-returns, while the expected growth rate μ^E can be estimated by its empirical growth rate. From this and from an estimate of the risk-free interest rate r obtained from Treasury bonds one can calculate the market price of risk as

$$\phi = \frac{\mu^E - r}{\sigma^E}.$$

In order to obtain an estimate for the firm's asset values A_t and volatility σ one need to use two equations relating them to equity values. For Merton's model, the first is Black-Scholes formula for a call option, that is,

$$E_t = c(t, A_t, K, \sigma, T), \tag{112}$$

where K and T are determined by the firm's debt structure. The second equation arises by equating the volatility term in (111) with the corresponding term obtained by applying Itô's formula to (112), namely,

$$\sigma \frac{\partial c}{\partial A} A_t = \sigma^E E_t. \tag{113}$$

Similar equations can be derived for first passage and excursion models, depending on the underlying pricing formula relating equity to assets.

Alternatively, one can use maximum likelihood methods to estimate σ and μ directly from the equity time series E^i. Once an estimate for σ is obtained in this way, it can be inserted back into a pricing formula such as (112) in order to produce estimates for the firm values A^i.

7 Seventh Lecture

7.1 Reduced form models

Let us define a counting process N_t as a non-decreasing, integer-valued process with $N_0 = 0$. As in section 5.1, we say that the counting process N_t admits an intensity λ_t if and only if

$$\Lambda_t = \int_0^t \lambda_s ds$$

and $N_t - \Lambda_t$ is a martingale. By a *reduced form model* of default arrival we mean a model for which the default time is given by the random variable

$$\tau = \inf\{t > 0 : N_t > 0\}. \tag{114}$$

That is, τ is the arrival time for the first jump of the counting process N_t. Therefore, given the distribution for N_t, survival probabilities can be calculated as

$$P[\tau > T] = P[N_T = 0].$$

We emphasize that all models in this section describe default events under the historical probability P. Risk-neutral reduced form models and their use for pricing purposes, as well as their relation to historical reduced form models through a risk premium, will be described in the next lecture.

7.1.1 Poisson processes

As a first example of an intensity model, let us consider a *Poisson process* N_t with parameter $\lambda > 0$, that is, a non-decreasing, integer-valued process starting at $N_0 = 0$ with independent and stationary increments which are Poisson distributed. More explicitly, for all $0 \leq s < t$ we have

$$P[N_t - N_s = k] = \frac{(t-s)^k \lambda^k}{k!} e^{-(t-s)\lambda} \tag{115}$$

It is clear from this definition that $E[N_t] = \lambda t$ and that $(N_t - \lambda t)$ is a martingale. Therefore the compensator for N_t is simply

$$\Lambda_t = \lambda t,$$

from which it follows that the constant λ is the intensity for the Poisson process.

The Poisson process has a number of important properties making it ubiquitous for modeling discrete events. Being Markovian, the occurrence of its next k jumps during any interval after time t is independent from its history up to t. We also see from (115) that the probability of one jump during a small interval of length Δt is approximately $\lambda \Delta t$ and that the probability for two or more jumps occurring at the same time is zero. Moreover, the waiting time between two jumps is an exponentially distributed random variable with parameter λ. In particular, if we use (114) to define default as the arrival for the first jump of N_t, then the expected default time is $1/\lambda$ and the probability of survival after t years is

$$P[\tau > t] = e^{-\lambda t}. \tag{116}$$

Therefore, from the definition of the hazard rate (88) we obtain

$$h(t) = -\frac{\partial \log P[\tau > t]}{\partial t} = \lambda \tag{117}$$

so that

$$P[\tau > t | \tau > s] = e^{-\int_s^t h(u)du} = e^{-\lambda(t-s)}. \tag{118}$$

7.1.2 Inhomogeneous Poisson processes

As we have just seen, modeling default as the arrival of the first jump of a Poisson process leads to a constant hazard rate. In pratice, the hazard rate changes in time, since survival up to different time horizons lead to different probability of default over the next small time interval. In order to obtain more realistic term structures of default probabilities, we are led to introduce a time-varying intensity $\lambda(t)$.

Let N_t then denote an *inhomogeneous Poisson process*, that is, a non-decreasing, integer-valued process starting at $N_0 = 0$ with independent increments satisfying

$$P[N_t - N_s = k] = \frac{1}{k!}\left(\int_s^t \lambda(u)du\right)^k \exp\left(-\int_s^t \lambda(u)du\right), \tag{119}$$

for some positive deterministic function $\lambda(t)$.

Observe that

$$N_t - \int_0^t \lambda(s)ds$$

is a martingale, that is,

$$\Lambda_t = \int_0^t \lambda(s)ds$$

is the compensator for N_t and the function $\lambda(t)$ is the intensity for the inhomogeneous Poisson process.

The properties of a Poisson process extend naturally to the inhomogeneous case. For instance, the probability of a jump over a small interval Δt is approximately given by $\lambda(t)\Delta t$. Furthermore, the waiting time between two jumps is a continuous random variable with density

$$\lambda(t)e^{-\int_0^t \lambda(s)ds}.$$

Therefore, defining default as the arrival of its first jump leads to the following survival probabilities

$$P[\tau > t|\tau > s] = e^{-\int_s^t \lambda(u)du}. \tag{120}$$

From this expression we obtain that, for the inhomogeneous Poisson process reduced form model,

$$h(t) = \lambda(t), \tag{121}$$

corresponding to a non-flat term structure for forward default rates, which can be calibrated to historical data.

In reality, survival up to time t is not the only relevant information in order to determine the probability of default for the next interval $[t, t + \Delta t]$. Other drivers, such as the credit rating and equity value of an obligor, or macroeconomic variables such as recession and business cycles, provide an additional flux of information that need to be incorporated when assigning default probabilities. Our next step in generalizing intensity based models is to allow for a stochastic intensity λ_t while retaining some of the desirable properties of Poisson processes.

7.2 Cox processes

Let us suppose that all the background information available in the economy, except for the default times, is expressed through the filtration $(\mathcal{F}_t)$. For example, $(\mathcal{F}_t)$ might be the filtration generated by a d–dimensional driving process X_t. We assume that all the default-free economic factors, including the risk-free interest rates, are adapted to $(\mathcal{F}_t)$. Assume further that there exists a non-negative process λ_t which is also adapted to $(\mathcal{F}_t)$ and plays the role of a stochastic intensity, generally correlated with the different components of the driving process X_t.

Next assume that $(\mathcal{F}_t^N)$ is the filtration generated by a point process N_t. The full filtration for the model is obtained as

$$(\mathcal{G}_t) = (\mathcal{F}_t) \vee (\mathcal{F}_t^N). \tag{122}$$

We say that the point process N_t is a *Cox process* if, conditioned on the background information $\mathcal{F}_t$ available at time t, N_t is an inhomogeneous Poisson process with a time-varying intensity $\lambda(s) = \lambda_s$, for $0 \leq s \leq t$. In other words, each realization of the process λ_t determines the local jump probabilities for the process N_t.

This definition is sometimes called the *doubly stochastic assumption*. Although very natural, it excludes several plausible situations. For example, the process N_t cannot be adapted with respect to the background information, nor can it be directly triggered by any of the driving processes (as is the case with structural models). It also excludes the possibility of N_t directly influencing the background processes (for instance, by causing a simultaneous jump in some of them, since this would reveal a jump in N_t by observing the background information only). Finally, the doubly stochastic assumption proves to be inadequate for modeling the arrival of common credit event when treating correlated obligors. Nevertheless, it provides a very convenient analytic framework for dealing with stochastic intensities, and should be used as a benchmark for more complicated models.

It follows from the definition that the compensator of a Cox process has exactly the form

$$\Lambda_t = \int_0^t \lambda_s ds,$$

which justifies calling λ_t its stochastic intensity.

To obtain the jump probabilities of a Cox process we average over realizations of this stochastic intensity, using the expression for inhomogeneous Poisson jump probabilities for each realization, that is,

$$P[N_t - N_s = k] = E\left[\frac{1}{k!}\left(\int_s^t \lambda_u du\right)^k \exp\left(-\int_s^t \lambda_u du\right)\right]. \tag{123}$$

Similarly, the waiting between each of its jumps is a continuous random variable with density

$$E\left[\lambda(t)e^{-\int_0^t \lambda(s)ds}\right].$$

Defining default as the arrival of its first jumps then leads to the following expression for survival probabilities:

$$P[\tau > t] = E\left[e^{-\int_0^t \lambda_s ds}\right]. \tag{124}$$

Therefore, from the definition of hazard rate given in (88) we obtain

$$h(t) = -\frac{\partial}{\partial t}\log E\left[e^{-\int_0^t \lambda_s ds}\right]. \tag{125}$$

More generally, conditioned on the background information available at time $s \leq t$, the probability of survival after t years is

$$P[\tau > t|\mathcal{F}_s] = E_s\left[e^{-\int_s^t \lambda_u du}\right]. \tag{126}$$

It then follows from definition (91) that the forward default rates are given by

$$f(s,t) = -\frac{\partial}{\partial t}\log E_s\left[e^{-\int_s^t \lambda_u du}\right]. \tag{127}$$

Notice that (126) is mathematically equivalent to (15) with the stochastic intensity playing the role of a stochastic interest rate. Therefore the mathematical apparatus for calculating bond prices in default-free interest rate theory described in sections 2 and 3 can be employed to calculate historical survival probabilities in reduced-form default models under the doubly stochastic assumption. In particular, we can model the intensity process λ_t as one of the convenient processes leading to affine term structures, such as the CIR process. We will see some examples of such intensity models in the next lecture.

7.2.1 Simulating the default time

As it is well-known in elementary probability theory, given a strictly decreasing function $F : \mathbb{R}^+ \to [0, 1]$ such that $F(0) = 0$ and

$$\lim_{t\to\infty} F(t) = 1,$$

we can construct a continuous random variable τ having F as its *cumulative distribution function* by setting $\tau = F^{-1}(U)$, for a uniformly distributed random variable $U : \Omega \to [0, 1]$. This is because

$$P[\tau \leq t] = P[F^{-1}(U) \leq t] = P[U \leq F(t)] = F(t).$$

Therefore, if a model for default arrival is such that the survival probabilities $1 - F(t) = P[\tau > t]$ can be easily inverted, we can obtain the correct distribution for the default time by simulating a uniform random variable U and setting τ as the solution to

$$1 - F(\tau) = U.$$

For Cox processes, an alternative method for simulating default times without the need to invert the term structure for survival probabilities is the *compensator simulation*, which is based on the numerical simulation of the compensator process

$$\Lambda_t = \int_0^t \lambda_s ds.$$

The method consists of simulating a unit-mean exponential random variable Z, *independently* of the intensity process λ_t, and setting the default time as

$$\tau = \inf\{t > 0 : \Lambda_t \geq Z\}.$$

Then, using the fact that $P[Z > z] = e^{-}z$, we have that, conditional on the path of λ_t up to time t, we have

$$P[\tau > t] = P[Z > \int_0^t \lambda_s ds] = e^{-\int_0^t \lambda_s ds}$$

as required.

8 Eighth Lecture

8.1 Affine Intensity Models

In analogy with interest rate theory, one can specify intensity processes of the form

$$\lambda_t = a + b \cdot X_t$$

where a and $b = (b_1, \ldots, b_n)$ are positive constants and $X_t = (X_t^1, \ldots, X_t^n)$ is a multidimensional Markov "factor" process. One then say that the model is *affine* if survival probabilities can be written in the form

$$P[\tau > t|\mathcal{F}_s] = \exp[A(s,t) + B(s,t)\lambda_s] \tag{128}$$

for some coefficient functions $A(s,t)$ and $B(s,t)$. In section 2 we have seen examples of one factor affine models where the underlying factor was the interest rate r_t itself and was specified as an Itô diffusion process. Several generalizations of this set up are possible, in which the underlying factors correspond to more than one economic drivers and can present jumps or stochastic volatilities.

8.1.1 CIR Intensities

As a first example of an affine model, let us consider a one factor model where the intensity process λ_t following a CIR dynamics

$$d\lambda_t = k(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t, \tag{129}$$

for positive constants k, θ and σ satisfying the condition $4k\theta > \sigma^2$. As usual, the parameters k and θ represent the long-term average and the rate of mean reversion for λ_t, while σ is a volatility coefficient.

Borrowing from the work we have already done for interest rate models, we have that survival probabilities in the CIR intensity model have the form

$$P[\tau > t|\mathcal{F}_s] = \exp[A(s,t) + B(s,t)\lambda_s]$$

where

$$A(s,t) = \frac{2k\theta}{\sigma^2}\log\left[\frac{2\gamma\exp[(k+\gamma)(t-s)/2]}{2\gamma + (k+\gamma)(\exp[(t-s)\gamma] - 1)}\right] \tag{130}$$

$$B(s,t) = \frac{2(1-\exp[(t-s)\gamma])}{2\gamma + (k+\gamma)(\exp[(t-s)\gamma] - 1)} \tag{131}$$

and $\gamma^2 = k^2 + 2\sigma^2$.

It is interesting to notice from this formula that survival probabilities increase if we increase the volatility parameter, while keeping all other parameters fixed. In other words, forward default rates decrease as the volatility in the intensity process increases. This is a consequence of Jensen's inequality and expression (126).

The effects of volatility on survival probabilities and forward rates are compensated, on the other hand, by the rate of mean reversion. Higher values of k mean that λ_t stays close to its long-term average θ. This has the effect of bringing the forward rate close to a long-term level as well. Conversely, smaller values of k accentuate the impact of the volatility in λ_t, leading to higher survival probabilities and smaller forward rates.

8.1.2 Mean-reverting intensities with jumps

Suppose now that the intensity process follows the dynamics

$$d\lambda_t = k(\theta - \lambda_t)dt + dZ_t \tag{132}$$

where $Z_t = J_t - cJt$ is a compensated compound Poisson process. That is, J_t is a pure jump process with independently distributed jumps at Poisson arrival times with intensity c and independent jump sizes drawn from an exponential distribution with mean J.

One can then prove, using generalized Ricatti equations, that survival probabilities in this model have the form

$$P[\tau > t|\mathcal{F}_s] = \exp[A(s,t) + B(s,t)\lambda_s]$$

where

$$\begin{aligned} A(s,t) &= -(\theta - cJ/k)\left((t-s) + B(s,t)\right) - \frac{c}{J+k}\left[Jt - \log\left(1 - B(s,t)J\right)\right] \\ B(s,t) &= -\frac{1 - e^{-k(t-s)}}{k}. \end{aligned} \tag{133}$$

8.2 Risk-neutral intensity models

We now turn our attention to reduced form models under a risk-neutral measure Q. We still define the default event as the first jump of a counting process N_t. The only difference is that the random default time

$$\tau = \inf\{t > 0 : N_t > 0\}$$

now has a different distribution under Q. Accordingly, to obtain risk-neutral survival probabilities we need to consider the distribution of the counting process N_t under the risk neutral measure Q, since

$$Q[\tau > t] = Q[N_t = 0]. \tag{134}$$

As before, we assume that the counting process N_t admits an intensity λ_t^Q, in the sense that

$$N_t - \int_0^t \lambda_s^Q ds$$

is now a Q–martingale.

All the previous definitions and examples, including the doubly stochastic assumption, have risk-neutral analogues in terms of the risk-neutral intensity λ_t^Q. In particular,

$$Q[\tau > t|\mathcal{F}_s] = E_t^Q\left[e^{-\int_s^t \lambda_u^Q du}\right], \tag{135}$$

so the survival probability expressions we just derived for affine intensity models apply for risk-neutral probabilities as well.

8.2.1 Valuation of defaultable bonds

The main reason for considering risk-neutral intensity models is that we can make use of risk-neutral valuation techniques to price credit derivatives. For example, we have seen that the price at time t of a defaultable bond with maturity T, given that default hasn't occurred up to t, is

$$\bar{P}_{tT} = E_t^Q\left[e^{-\int_t^T r_s ds}\mathbf{1}_{\{\tau > T\}}\right] \tag{136}$$

If both the interest rate r and the risk-neutral intensity λ^Q are constant, the expression above reduces to

$$\bar{P}_{tT} = e^{-r(T-t)}Q[\tau > T|\tau > t] = e^{-(r+\lambda^Q)(T-t)}, \tag{137}$$

where we have used the risk-neutral analogue of (118).

More generally, under a risk-neutral doubly-stochastic assumption, we have that

$$\bar{P}_{tT} = E_t^Q\left[e^{-\int_t^T (r_s+\lambda_s)ds}\right], \tag{138}$$

a result known as *Lando's formula.* This formula accommodates correlation between r_t and τ, therefore generalizing (99).

8.2.2 Two-factor Gaussian models

As a concrete example of a calculation of prices for defaultable bonds according to (138), consider a model for which both the risk-free rate and the intensity process follow a Hull-White process, that is, let

$$dr_t = \kappa(\Theta(t) - r_t)dt + \sigma dW_t^Q. \tag{139}$$

and

$$d\lambda_t^Q = \bar{\kappa}(\bar{\Theta}(t) - \lambda_t^Q)dt + \bar{\sigma}dZ_t^Q, \tag{140}$$

where W^Q and Z^Q are correlated Q–Brownian motions with

$$dW_t^Q dZ_t^Q = \rho dt. \tag{141}$$

The function $\Theta(t)$ can be chosen to match the initial term structures of defautl-free bonds according to (56). We can now use proposition 3.2 to calculate the price of zero coupon bonds. Similarly, risk-neutral survival probabilities are given by

$$Q[\tau > T|\mathcal{F}_t] = \exp[\bar{A}(t,T) + \bar{B}(t,T)\lambda_t^Q] \tag{142}$$

where

$$\bar{B}(t,T) = \frac{1}{\bar{\kappa}}(e^{-\bar{\kappa}(T-t)} - 1) \tag{143}$$

and

$$\bar{A}(t,T) = \frac{1}{2}\int_t^T \sigma^2 \bar{B}(t,s)^2 ds - \int_t^T \bar{B}(t,s)\bar{\kappa}\bar{\Theta}(s)ds. \tag{144}$$

More importantly, by evaluating (138) in the T-forward measure, the price of a defaultable bond in this model is given by

$$\bar{P}_{tT} = P_{tT}\exp[\tilde{A}(t,T) + \bar{B}(t,T)\lambda_t^Q], \tag{145}$$

where

$$\tilde{A}(t,T) = \frac{1}{2}\int_t^T \sigma^2 \bar{B}(t,s)^2 ds - \int_t^T \bar{B}(t,s)[\bar{\kappa}\bar{\Theta}(s) - \rho\bar{\sigma}\sigma B(t,s)]ds. \tag{146}$$

8.2.3 From actual to risk neutral intensities

Given an intensity model under the physical measure P, it does not necessarily follow that there exist a risk–neutral intensity model with the same properties. For instance, the doubly stochastic assumption needs to be independently stated for P and Q. Moreover, the intensities λ_t and λ_t^Q themselves can depend differently on the state variables of the model, as well as have different likelihood for each path. Even in the situation where $\lambda_t = \lambda_t^Q$ we can still have that

$$P[\tau > T|\mathcal{F}_t] = E_t\left[e^{-\int_t^T \lambda_s ds}\right]$$

is different from

$$Q[\tau > T|\mathcal{F}_t] = E_t^Q\left[e^{-\int_t^T \lambda_s^Q ds}\right].$$

In practical implementations, one might assume a simple functional relation such as

$$\lambda_t = \Psi\lambda_T^Q \tag{147}$$

and try to match the scaling factor Ψ to empirical data, using both historical default probabilities and market prices for defaultable bonds.

More generally, in the case of multidimensional affine models with a state vector $X_t = (X_t^1, \ldots, X_t^n)$, we can set

$$\lambda_t = \Psi_0\lambda_t^Q + \Psi \cdot X \tag{148}$$

and try to estimate the parameters Ψ_0 and $\Psi = (\Psi_1, \ldots, \Psi_n)$ from empirical data.

9 Ninth Lecture

9.1 Credit Rating Models

Rating agencies such as Moody's, Standard & Poor's and Fitch classify companies according to *credit rating classes*. These ratings are then made publicly available and are used by market participants as an indicator of the credit quality of a given company. For instance, most investors are regulated to only buy corporate bonds from companies of specified ratings. Some bonds contain contractual specifications that depend explicitly on the credit rating of the issuer, such as an increase in coupon payments if the issuer happens to be downgraded. In more extreme circumstances, a downgrade can be interpreted as a default event itself, triggering the exercise of several credit derivatives. For all these reasons, a reasonable understanding of credit rating and rating transitions is a necessary ingredient of any credit risk model.

9.1.1 Discrete-time Markov chain

As a first step, one can considered *historical transition frequencies* published by an agency. For example, the following matrix corresponds to Moody's all-corporate average transition frequencies for 1980 to 2000 for the rating classes *Aaa, Aa, A, Baa, Ba, B* and *Caa-C*. Each row represents a rating for the beginning of a year, while each column represents the end-of-year rating, with the last row and column corresponding to default:

$$\Pi = \begin{bmatrix} 89.14 & 9.78 & 1.06 & 0.00 & 0.03 & 0.00 & 0.00 & 0.00 \\ 1.14 & 89.13 & 9.25 & 0.32 & 0.11 & 0.01 & 0.00 & 0.03 \\ 0.06 & 2.97 & 90.28 & 5.81 & 0.69 & 0.18 & 0.01 & 0.01 \\ 0.06 & 0.36 & 7.01 & 85.47 & 5.82 & 1.02 & 0.08 & 0.17 \\ 0.03 & 0.07 & 0.59 & 5.96 & 82.41 & 8.93 & 0.58 & 1.44 \\ 0.01 & 0.04 & 0.22 & 0.61 & 6.43 & 82.44 & 3.29 & 6.96 \\ 0.00 & 0.00 & 0.00 & 0.95 & 2.85 & 6.15 & 62.36 & 27.68 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 100.00 \end{bmatrix} \tag{149}$$

That is, the entry $\Pi_{13} = 1.06$ gives the percentage of *Aaa* companies that have been downgraded to *A* rating during a one year period, averaged over the years 1980 to 2000. Observed that the last row indicates that no company can recover from default, which is deemed to be an *absorbing state*.

One can then model rating transitions over time as a *Markov chain* C_n whose K states are the rating classes and having Π as its transition probabilities. More explicitly, this says that the P–probability that a firm will have a rate j at the end of the year depends only on its beginning-of-the-year rate i and is given by Π_{ij}. Albeit elegant, this model suffers from several pitfalls, mainly related to the fact that historical transition frequencies do not take into account all available information. For example, historical frequencies for low-probability events are based on a small number of observations, leading to estimated probabilities which are not significant. The Markov chain assumption also ignores empirically observed *momentum* and *aging effects* in rating transitions, that is, a higher upgrade/downgrade probability for firms that have been upgraded/downgraded in the previous year than for firms which remained in the same class for longer. Both phenomena are manifestations of the more general property that firms of the same rating exhibit different (and time dependent) credit qualities. Nevertheless, due to its mathematical tractability, we explore the Markov chain model as a first approximation for the mechanism of rating transition.

We can extended the model in order to incorporate the influence of macroeconomic factors, such as business cycles, by taking the yearly transition matrix to be of the form $\Pi(X_n)$, where the state variable X_n is assumed to be itself a Markov chain with a finite number of states. We then make the *doubly stochastic* assumption that, conditioned on a path for the process X, the probability of making a transition from rate i at time k to rate j at time $m > n$ is given by the product

$$\Pi(X_n)\Pi(X_{n+1})\cdots\Pi(X_m).$$

More generally, we can view the pair (X_n, C_n) itself as a Markov chain. For example, if X represents business cycles and can take three values, corresponding to peak, normal and recession periods, then a model with seven rating classes plus a default state leads to a Markov chain with $3 \times 7 = 21$ non-absorbing 3 absorbing states, for which the $21 \times 21 + 3 \times 21 = 504$ one-period transition probabilities required. The doubly stochastic assumption simplifies this to the specification of a 3×3 transition matrix p_{xy} for the process X_t and three different 7×8 transition matrices $\Pi^(x)$, so that the transition probability from (x, i) to (y, j) is given by $p_{xy}\Pi_{ij}(x)$. That is, it requires only 156 parameters.

9.1.2 Continuous-time Markov chain

We can obtain a continuous-time credit rating model by assuming that the transition from a state i to a state j occurs with an intensity Λ_{ij}. That is, assume that, for an infinitesimal time interval Δ_t, the probability of starting in a state i at time t and ending in a state j at time $t+\Delta_t$ is approximately equal to $\Lambda_{ij}\Delta t$. The matrix Λ is then called the *generator* for the continuous-time Markov chain and has the properties that

$$\Lambda_{ij} \geq 0, \qquad \text{for all } i \neq j \tag{150}$$

and

$$\Lambda_{ii} = -\sum_{i \neq j} \Lambda ij. \tag{151}$$

The transition probability of starting in a state i at time s and ending in a state j at any later time t is then given as the ij entry of the matrix

$$\Pi(s,t) = e^{\Lambda(t-s)}. \tag{152}$$

The one-year transition matrix from the previous session can then be calculated as

$$\Pi = \Pi(0,1) = e^{\Lambda}, \tag{153}$$

from which one can calibrate the generator Λ to historical data. It is not generally true, however, that there will exist a generator for any given a one-year transition matrix. Moreover, the generator is not uniquely determined by the one-year probabilities. Therefore, distinct generators can be compatible with the same one-year transition matrix, but will assign different transition probabilities for time intervals other than one year.

For time-varying generators, one obtains that the transition matrix satisfies the differential equation

$$\frac{\partial \Pi(s,t)}{\partial t} = -\Lambda \Pi(s,t).$$

For the special case where the matrices Λ_s and Λ_t commute for all $t \neq s$, this differential equation has solution of the form $\Pi(s,t) = \exp(\int_s^t \Lambda(u)du)$, which generalizes (152). One special example for which this is true is if the generator can be written as

$$\Lambda(t) = B\mu(t)B^{-1}, \tag{154}$$

where $\mu(t)$ is the diagonal matrix whose entries are the eigenvalues of Λ and B is an orthogonal matrix of eigenvectors. it then follows that the transition probability matrix has the form

$$\Pi(s,t) = B\exp(\int_s^t \mu(u)du)B^{-1}. \tag{155}$$

As a final generalization, one should expected the generator to vary stochastically in time, in order to respond to new information in the market. Lando (1998) proposes a concrete model for stochastic generators by assuming the diagonalized form (154) and taking the eigenvalues of Λ to be given as

$$\mu_j(t) = \alpha_j + \beta_j \cdot X_t, \tag{156}$$

for some multidimensional affine state process X_t. Then, conditional on a path for the process X_t, the transition probability matrix is given by (155). Therefore, given the value of the state process X_s, we have that this doubly stochastic transition model assigns the following probability for from rate i at s to rate k at a later time t:

$$\begin{aligned}
\Pi(s,t,X_s) &= E_t\left[\sum_{j=1}^{K+1}\beta_{ijk}\exp\left(\int_s^t \mu_j(u)du\right)\right] \\
&= \sum_{k=1}^{K+1}\beta_{ijk}E_t\left[\exp\left(\int_s^t(\alpha_j+\beta_j\cdot X_u)du\right)\right] \\
&= \sum_{k=1}^{K+1}\beta_{ijk}\exp\left(A_j(s,t)+B_j(s,t)X_s\right),
\end{aligned} \tag{157}$$

where $\beta_{ijk} = B_{ij}B^{-1}_{jk}$ and coefficients $A_j(s,t)$ and $B_j(s,t)$ satisfying the usual Ricatti equations for affine processes.

9.2 Recovery Models

Up to this point, our discussion of prices for defaultable bonds has implicitly assumed a zero recovery rate in the event of default, that is, we only considered the case for which the value of the bond drops to zero when default occurs. In this section we generalize our framework by allowing the pay-off of credit derivatives to drop by a stochastic factor ϕ_τ in the event of a default at time τ. We considered the doubly stochastic model of section 7.2 with a stochastic intensity λ_t and assume that the *recovery rate* R_t is $\mathcal{G}_\tau$–measurable. Assuming that $\tau > 0$, the price at time zero of a defaultable bond with maturity T is given by

$$\begin{aligned}
\bar{P}_{0T} &= E^Q\left[e^{-\int_0^\tau r_s ds} R_\tau \mathbf{1}_{\{\tau \leq T\}} + e^{-\int_0^T r_s ds}\mathbf{1}_{\{\tau > T\}}\right] \\
&= E^Q\left[E^Q\left[e^{-\int_0^\tau r_s ds} R_\tau \mathbf{1}_{\{\tau \leq T\}}\middle|\mathcal{F}_T\right]\right] + E^Q\left[e^{-\int_0^T r_s ds}\mathbf{1}_{\{\tau > T\}}\right] \\
&= E^Q\left[\int_0^T e^{-\int_0^u (r_s+\lambda_s)ds} R_u \lambda_u du\right] + E^Q\left[e^{-\int_0^T (r_s+\lambda_s)ds}\right] \\
&= \int_0^T E^Q\left[R_u\lambda_u e^{-\int_0^u (r_s+\lambda_s)ds}\right] du + E^Q\left[e^{-\int_0^T (r_s+\lambda_s)ds}\right] \qquad (158)
\end{aligned}$$

that is, the sum of a recovery term and a survival term. In the following subsections we show how to calculate this expression under different models for the recovery rate.

9.2.1 Zero recovery

This corresponds to $R_t = 0$ for all $t \geq 0$, so that the expression above reduces to Lando's formula for the survival term

$$\bar{P}_{0T} = E^Q\left[e^{\int_0^T (r_s+\lambda_s)ds}\right]. \qquad (159)$$

Although unrealistic, the price of a zero recovery bond has theoretical importance and often enters more complicated expressions under more general recovery frameworks.

9.2.2 Recovery of treasury

A first generalization is to consider a recovery of the form

$$R_t = RP_{tT}, \qquad (160)$$

that is, a constant fraction $0 < R < 1$ of the equivalent default-free bond. We then have

$$\begin{aligned}
\bar{P}_{0T}^{RT} &= R\int_0^T E^Q\left[P_{uT}\lambda_u e^{-\int_0^u (r_s+\lambda_s)ds}\right]du + \bar{P}_{0T} \\
&= RE^Q\left[e^{-\int_0^T r_s ds}\int_0^T \lambda_u e^{-\int_0^u \lambda_s ds}du\right] + \bar{P}_{0T} \\
&= RE^Q\left[e^{-\int_0^T r_s ds}\left(1 - e^{-\int_0^T \lambda_s ds}\right)\right] + \bar{P}_{0T} \\
&= RP_{0T} + (1-R)\bar{P}_{0T}. \qquad (161)
\end{aligned}$$

In other words, an RT (recovery of treasury) bond can be priced as the sum of R units of a default-free bonds plus $(1-R)$ units of a ZR (zero recovery) bond, a result that could have been deduced from the pay-off structure alone.

9.2.3 Recovery of par

Under this model, a bond will pay a fraction $0 < R < 1$ of its promised pay-off if a default happens. That is, the recovery rate is $R_t = R$ for $t \geq 0$. The same calculation as before leads to

$$\bar{P}_{0T}^{RP} = R\int_0^T E^Q\left[\lambda_u e^{-\int_0^u (r_s+\lambda_s)ds}\right]du + \bar{P}_{0T}. \qquad (162)$$

9.2.4 Recovery of market value

The next model for recovery is to assume that if default happens at time τ then a defaultable bond pays a fraction R of its pre-default market value. This corresponds to a recovery rate defined as

$$R_t = R\bar{P}^{RMV}(t-), \qquad (163)$$

where

$$\bar{P}^{RMV}(t-) = \lim_{s\nearrow t}\bar{P}_{sT}^{RMV} \qquad (164)$$

and $\bar{P}_{tT}^{RMV}$ denotes the price at time $t < \tau$ of an RMV (recovery of market value) defaultable bond with maturity T.

The derivation of an RMV defaultable bond given the recovery rate above is quite more laborious than for other recovery models and will be omitted in this notes. The resulting price, assuming that $\tau > 0$, is

$$\bar{P}_{0T}^{RMV} = E_t^Q\left[e^{-\int_t^T (r_s+(1-R)\lambda_s)ds}\right] \qquad (165)$$

10 Tenth Lecture

We investigate in this lecture several ways to introduce correlations between credit events associated with different firms. In the following, we consider a credit portfolio consisting of exposure to I obligors, labeled by an index $i \in \{1, \ldots, I\}$.

10.1 The Binomial Expansion Technique

As a first step, let us suppose that default events are *independent* across obligors and happen with probability p over a fixed time horizon T. Assume further that our portfolio has equal exposure L to each obligor with identical recovery rates R. Then if n denotes the number of defaults occurring before T, the loss due to default for this portfolio during the period $[0, T]$ will be simply $X = n(1-R)L$. Therefore, its distribution is completely determined by the distribution of the number of defaults. Under our assumptions, this is in turn given by the binomial distribution

$$P[X \leq x] = \sum_{m=0}^{n} \frac{I!}{n!(I-n)!} p^n (1-p)^{I-n}, \tag{166}$$

where n is the rounded-down integer part of $x/(1-R)L$.

This familiar distribution is contradicted by empirical data. In order to relax the independence assumption, Moody's proposes to group the obligors into $1 \leq D \leq I$ classes. Within each class, obligors are assumed to be completely dependent, and can then be treated as a single firm with exposure LI/D, while the different classes are deemed to be fully independent. For example, with $D = I$ we recover full independence between obligors, while for $D = 1$ we have complete dependence. The parameter D is called the *diversity score* of the portfolio, and the intermediate cases between full independence and complete dependence can now be analyzed for varying values of D.

For a fixed D, the distribution of defaults can be calculated as a binomial distribution. Therefore, the BET loss distribution for a diversity score D is given by

$$P[X \leq x] = \sum_{m=0}^{n} \frac{D!}{m!(D-m)!} p^m (1-p)^{D-m}, \tag{167}$$

where $n = \left[\frac{xD}{(1-R)LI}\right]$.

This method is not based on any theoretical model, and the parameter D is determined by a complicated recipe used by Moody's. Moreover, its loss resolution is limited to intervals of size I/D, which are often too coarse. In order to have a deeper understanding of default correlation, we must resort to our previous models for single-name default. Nevertheless, the BET method became something of a market standard to which more detailed models should be compared.

10.2 Default correlation in reduced-form models

Recall that in intensity-based models, the default of obligor i occurs at the first jump of a point process N_t^i with intensity λ_t^i. If we simply consider the combined process

$$N_t = \sum_{i=1}^{I} N_t^i,$$

then it follows that

$$\Lambda_t = \sum_{i=1}^{I} \int_0^t \lambda_s^i ds$$

is the compensator for N_t. However, it is not true in general that N_t is itself a point process, since its jumps can have a magnitude greater than one. Therefore, in the presence of simultaneous defaults, we cannot assert that the sum of the individual intensities will be an intensity for N_t in the sense described in section 7.1. Moreover, even if simultaneous defaults are not allowed in the model, starting with Cox processes N_t^i does not guarantee that N_t is itself a Cox process with intensity $\lambda_t^1 + \ldots \lambda_t^I$. In the next sections, we consider more specific assumptions on the intensity processes in order to appropriately model the possibility of joint defaults.

10.2.1 Doubly stochastic models

As we have seen, the doubly stochastic assumption of section 7.2 needs to be modified in the presence of several default events. By a doubly stochastic model of correlated defaults we mean a model whereby, conditioned on the background filtration $(\mathcal{F})_t$, which in particular contains the realization of the multidimensional intensity process $(\lambda_t^1, \ldots, \lambda_t^I)$, the processes N^i are *independent* inhomogeneous Poisson processes, each with intensity λ_t^i.

That is, the only interdependence between default events occurs through the correlations prevailing between the intensity processes. Once the intensities are revealed, the remaining stochastic processes triggering default are independent. Although this assumption seems a natural generalization of the single-name doubly stochastic setup, it leads to nonrealistic correlations for default events. As an example, consider the joint default probability for obligors i, j for a fixed time horizon T. That is

$$\begin{aligned} p_{ij} := P[\tau_i \leq T, \tau_j \leq T] &= E\left[\mathbf{1}_{\{\tau_i \leq T\}}\mathbf{1}_{\{\tau_j \leq T\}}\right] \\ &= E\left[E\left[\mathbf{1}_{\{\tau_i \leq T\}}\mathbf{1}_{\{\tau_j \leq T\}}|\lambda^i, \lambda^j\right]\right] \\ &= E\left[(1 - e^{-\int_0^T \lambda_s^i ds})(1 - e^{-\int_0^T \lambda_s^j ds})\right] \\ &= p_i + p_j + E\left[e^{-\int_0^T (\lambda_s^i + \lambda_s^j) ds}\right] - 1, \end{aligned}$$

where $p_i = P[\tau_i \leq T]$. This achieves its maximum value for perfectly correlated intensities, that is $\lambda^i = \lambda^j$. In this case we have that the correlation between the default events is

$$\begin{aligned} \rho = \frac{p_{ij} - p_i^2}{p_i(1 - p_i)} &= \frac{2p_i + E\left[e^{-2\int_0^T \lambda_s^i ds}\right] - 1 - p_i^2}{p_i(1 - p_i)} \\ &= \frac{\mathrm{Var}\left[e^{-2\int_0^T \lambda_s^i ds}\right]}{p_i(1 - p_i)}. \end{aligned}$$

To estimate the order of magnitude of this correlation for intensities given by diffusion processes, let us assume that the integral in the exponent is normally distributed with mean μ and variance σ^2. We then obtain that

$$\rho = \frac{p}{1 - p}(e^{\sigma^2} - 1),$$

that is, the default correlation is *at most* of the same of the default probabilities. Most often, this is too low to account for empirical data for correlated defaults. The only way to remedy this situation is to allow for joint large jumps in the intensity processes.

10.2.2 Joint default events

As an alternative to the doubly stochastic framework above, Duffie and Singleton (1998) suggest a model in which there are J credit events, each triggered by the first jump of a Cox $\widehat{N}_t^j$ with intensity $\widehat{\lambda}_t^j$. Each of the events consist of multiple defaults, for example, in the case of two obligors A and B the credit events are default for A, default for B and simultaneous default for A and B, with the respective intensities $\widehat{\lambda}_A$, $\widehat{\lambda}_B$ and $\widehat{\lambda}_{AB}$. For consistency with single-names intensities, we must have

$$\begin{aligned}\lambda_A &= \widehat{\lambda}_A + \widehat{\lambda}_{AB} \\ \lambda_B &= \widehat{\lambda}_B + \widehat{\lambda}_{AB}\end{aligned}$$

In principle, for I firms, one needs to consider all the 2^I possible subsets of I, leading to unmanageable complexity. One could try to consider only the number of defaults in a certain credit event, making no distinction between the actual firms involved. Besides being too restrictive on the possible dependency structure, this assumption fails to capture the fact that, over a non-infinitesimal time horizon, the intensity of a k–default event must also depend on the occurrence of events with a smaller number of defaults.

More fundamentally, this approach leads to joint defaults happening most likely at the same time, possibly involving a large number of firms, but without affecting the surviving firms. Amongst other things, this precludes the emergence of crises, that is, periods during which default intensities are higher for all firms.

10.2.3 Infectious defaults

The idea of a large default event spreading its influence to surviving firms is expressed in the models proposed by Davis and Lo (2001) and Jarrow and Yu (2001). Here we consider obligors with initially identical intensities $\lambda_0^{(i)} = \lambda$, which get uniformly increased by a *risk enhancement* factor $a \geq 1$ at each default and thereafter decay to λ after an exponentially distributed time with parameter μ. The main drawback of this intuitively appealing model is that its joint distribution of default over a time horizon T is hard to calculate. Moreover, the resulting joint process for default indicators is not a Cox process, since we can obtain information about the default times by observing the joint intensities alone.

10.3 Default correlation in structural models

Unlike reduced-form models, default correlation can be incorporated quite naturally into structural models. This is achieved by making the asset value dynamics correlated through time. We illustrate the main ideas with the example of two firms whose assets follow the dynamics

$$dA_t^i = (\mu^i - \delta^i)A_t^i + \sigma^i dW_t^i, \quad A_0^i > 0, i = 1,2 \tag{168}$$

where W^1 and W^2 are correlated Brownian motions with constant correlation ρ. Then for the classical Merton model, where default occurs at time T if $A_T^i < K^i$, we obtain that the joint default probability

$$\begin{aligned} p(T_1, T_2) &= P[A_{T_1}^1 < K_1, A_{T_2}^2 < K^2] \\ &= \Phi_2\left(\rho, \frac{\log(K^1/A_0^1) - m^1 T_1}{\sigma^1\sqrt{T_1}}, \frac{\log(K^2/A_0^2) - m^2 T_2}{\sigma^2\sqrt{T_2}}\right), \end{aligned} \tag{169}$$

where $m^i = \mu^1 - \delta^1 - (\sigma^i)^2/2$ and $\Phi_2(\rho, \cdot, \cdot)$ is the bivariate standard normal distribution function

$$\Phi_2(\rho, a, b) = \int_{-\infty}^{a}\int_{-\infty}^{b} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)}\right) dxdy.$$

If we instead consider a first-passage model with barriers D_1 and D_2, then the joint probability for firm 1 to default before time T_1 and firm 2 to default before time T_2 is given in closed form by

$$p(T_1, T_2) = \Psi_2\left(\rho, T_1, T_2, \log\frac{D_1}{A_0^1}, \log\frac{D_2}{A_0^2}\right), \tag{170}$$

where $\Psi_2(\rho, \cdot, \cdot, a, b)$ is the bivariate inverse Gaussian distribution function with correlation ρ and parameters a, b. A similar result exist for an exponentially growing time-dependent barrier in terms of modified Bessel functions.

In all of these approaches, we obtain can obtain a full range of correlations between default events. The results carry over in principle for a larger number of firms, although the number of parameters gets quickly out of hand. We would need to specify a full $I \times I$ variance-covariance matrix in order to obtain the entire dependence structure for I firms. This is still small compared to the 2^I correlated default events which would have to be considered if we were not dealing with Gaussian random variables, but it is nevertheless a paralyzing task if we have, say, 100 obligors.

Instead, one simplifies the picture by introducing *factor models*. For example, after appropriately normalizing the constants appearing in the in the Merton framework, a *one-factor* model corresponds to assuming that $V_t^i = \log A_t^i$ is such that

$$V_T^i = \rho Y + \sqrt{1-\rho^2}\varepsilon_i, \tag{171}$$

where Y and ε^i are independent standard normal random variables. This has the interpretation that asset values are driven by one common factor Y plus idiosyncratic factor ε. Then the entire dependence structure is reduced to specifying the correlation parameter ρ. The important feature of the model is that, conditional on the realization of the systematic random variable Y, all the companies are identical and independent. As before, we say that default occurs for firm i at time T if $V_T^i < \widehat{K}^i$. The default threshold $\widehat{K}^i$ can be calibrated to individual default probabilities $p_i = P[\tau_i = T]$ by setting

$$\widehat{K}^i = \Phi^{-1}(p_i).$$

To further investigate the implications of the one-factor model, let us assume that all the individual default probabilities are the same, so that $\widehat{K}^i = \widehat{K}$ and $p_i = p$ for all firms. Then conditioned on a value of the systematic random variable Y, the probability of default for each company is given by

$$p(y) = P[V_T^i < \widehat{K} | Y = y] = \Phi\left(\frac{\widehat{K} - \rho y}{\sqrt{1-\rho^2}}\right), \tag{172}$$

where Φ is the standard normal distribution function. Therefore, the probability of observing up to n defaults occurring by time T is

$$\sum_{m=1}^{n} \int_{-\infty}^{\infty} \frac{I!}{m!(I-m)!} p(y)^m (1-p(y))^{I-m} \phi(y) dy, \tag{173}$$

where ϕ is the standard normal density.

We can obtain even more explicit results in the *large portfolio* approximation, that is, assuming that $I \to \infty$. Then $p(y)$ given by (172) represents, on average, the fraction of the portfolio that experiences default over the period $[0, T]$. If we normalize the exposure $L^i = 1$ and assume zero-recovery, then this fraction is exactly the portfolio loss in default. Its distribution is then

$$\begin{aligned} F(x) &= := P[X \leq x] = E\left[P[X \leq x|Y]\right] \\ &= \int_{-\infty}^{\infty} P[X \leq x|Y = y]\phi(y)dy \\ &= \int_{-\infty}^{\infty} \mathbf{1}_{\{p(y) \leq x}\phi(y)dy \\ &= \Phi\left(\frac{\sqrt{1-\rho^2}\Phi^{-1}(x)}{\rho} - \Phi^{-1}(p)\right). \end{aligned} \tag{174}$$

11 Eleventh Lecture

11.1 Definition of Copula Functions

In all models presented so far for default correlation, information about the credit quality of individual firms comes intertwined with the default dependence. Copula models arise as a way of separating the dependence structure of correlated defaults form their marginal distributions. One can then easily calibrate a general default dependence to any given set of individual credit spreads.

Recall that, if $\tau : \Omega \to \mathbb{R}^+$ is a continuous random variable with cumulative distribution $F(t)$, then $U = F(\tau)$ will be a uniform random variable, since

$$P[U \leq u] = P[F(\tau) \leq u] = P[\tau \leq F^{-1}(u)] = F(F^{-1}(u)) = u.$$

We have already used this fact for simulating a default time τ when the survival probability $P[\tau > t] = 1 - F(t)$ could be easily inverted. The idea of copula models is to extend this observation to multidimensional processes.

We say that $C : [0,1]^I \to [0,1]$ is a *copula function* if

$$C(1,\ldots,1,u_i,1,\ldots,1) = u_i, \qquad \text{for all } i = 1,\ldots,I, \tag{175}$$

and there exit random variables $U_1,\ldots,U_I$ taking values in $[0,1]$ such that C is their distribution function.

The fundamental result concerning copula functions is Sklar's theorem. It states that, given any set of continuous random variables $\tau_1,\ldots,\tau_I$ with marginal distribution functions $F_1,\ldots,F_I$, there exist a unique copula function C such that their joint distribution

$$F(t_1,\ldots,t_I) := P[\tau_1 \leq t_1,\ldots,\tau_n \leq t_n]$$

can be written as

$$F(t_1,\ldots,t_I) = C(F_1(t_1),\ldots,F_I(t_I)). \tag{176}$$

In other words, the copula function C in (176) is the joint distribution for the uniformly distributed random variables $F_1(\tau_1),\ldots,F_I(\tau_I)$.

At first sight this theorem seem like an innocuous rewriting of the joint distribution F for the original random variables $\tau_1,\ldots,\tau_I$ in terms of the joint distribution C of the secondary random variables $F_1(\tau_1), \ldots, F_I(\tau_I)$. The essential point, however, is that these secondary random variables are uniformly distributed regardless of the distribution for the original random variables $\tau_1, \ldots, \tau_I$. One can then take them as the primitives of the model and concentrate on their correlation structure alone. Therefore, when specifying default correlation with copula functions, the modeler is free to express his views on the dependence structure alone by specifying a copula function C, while his views on individual default events are separately specified by the marginal distributions $F_1, \ldots, F_I$.

11.2 Fréchet bounds

To give some concrete examples, let us concentrate to the case of two obligors. Then independent defaults lead to independent random variables $F_1(\tau_1)$ and $F_2(\tau_2)$, so that the copula function must be the joint distribution of two independent uniform random variables U, V, that is

$$C(u,v) = P[U \leq U, V \leq v] = P[U \leq u]P[V \leq v] = uv. \tag{177}$$

Conversely, a product copula leads to

$$\begin{aligned} P[\tau_1 \leq t_1, \tau_2 \leq t_2] = F(t_1,t_2) &= C(F_1(t_1), F_2(t_2)) \\ &= F_1(t_1)F_2(t_2) \\ &= P[\tau_1 \leq t_1]P[\tau_2 \leq t_2], \end{aligned}$$

which is the condition for independent defaults.

Similarly, perfectly correlated defaults correspond to perfectly correlated random variables $F_1(\tau_1)$ and $F_2(\tau_2)$, so the copula function must be the joint distribution for two identical uniform random variables, that is

$$C(u,v) = P[U \leq u, U \leq v] = \min(u,v). \tag{178}$$

Conversely, this copula leads to perfectly correlated defaults, since

$$\begin{aligned} P[\tau_1 \leq t_1, \tau_2 \leq t_2] &= C(F_1(t_1), F_2(t_2)) \\ &= \min\left(F_1(t_1), F_2(t_2)\right) \\ &= \min\left(P[\tau_1 \leq t_1, P[\tau_2 \leq t_2]\right), \end{aligned}$$

which implies that either $\{\tau_1 \leq t_1\} \subset \{\tau_2 \leq t_2\}$ or $\{\tau_2 \leq t_2\} \subset \{\tau_1 \leq t_1\}$. Moreover, since

$$\tau_2 = F_2^{-1}(F_1(\tau_1)),$$

we have that the default time of one firm is an increasing function of the default time of the other. In the special case where $F_1 = F_2$, both defaults occur at exactly the same time.

In the same vein, perfectly anti-correlated default events correspond to perfectly anti-correlated random variables $F_1(\tau_1)$ and $F_2(\tau_2)$, so that the copula function must be the joint distribution for the uniform random variables U and $1-U$, that is

$$C(u,v) = P[U \leq u, 1-U \leq v] = \max(u+v-1,0). \tag{179}$$

Conversely, this copula leads to perfectly anti-correlated defaults, since

$$\begin{aligned} P[\tau_1 \leq t_1, \tau_2 \leq t_2] &= C(F_1(t_1), F_2(t_2)) \\ &= \max\left(F_1(t_1) + F_2(t_2) - 1, 0\right) \\ &= \max\left(P[\tau_1 \leq t_1 + P[\tau_2 \leq t_2] - 1, 0\right), \end{aligned}$$

implying either $\{\tau_1 \leq t_1\} \cap \{\tau_2 \leq t_2\} = \emptyset$ or $\{\tau_1 \leq t_1\}^c \cap \{\tau_2 \leq t_2\}^c = \emptyset$. Moreover, we have that

$$\tau_2 = F_2^{-1}(1 - F_1(\tau_1)),$$

so the default time of one firm is a decreasing function of the default time of the other.

In general, it can be shown that a two dimensional copula function is bounded by these two limiting cases, called the *Fréchet bounds*, that is

$$\max(u+v-1,0) \leq C(u,v) \leq \min(u,v). \tag{180}$$

Multidimensional copulas satisfy similar bounds, with similar interpretations in term of maximally correlated and anti-correlated random variables.

11.3 Tail dependence

For another example of copula, consider the joint probability (169) for default times derived in section 10.3 for the Merton structural model. From Sklra's theorem, we can rewrite it as

$$\begin{aligned} p(T_1, T_2) &= \Phi_2\left(\rho, \frac{\log(K^1/A_0^1) - m^1 T_1}{\sigma^1\sqrt{T_1}}, \frac{\log(K^2/A_0^2) - m^2 T_2}{\sigma^2\sqrt{T_2}}\right) \\ &= C_\rho^G(P[\tau_1 \leq T_1], P[\tau_2 \leq T_2]), \end{aligned}$$

where $C_\rho^G(\cdot,\cdot)$ is the *Gaussian copula* with correlation ρ:

$$C_\rho^G = \Phi_2(\rho, \Phi^{-1}(u), \Phi^{-1}(v)). \tag{181}$$

This illustrates the essential point about copulas: although a Gaussian copula is derived from a bivariate Gaussian random variable with Gaussian marginal, it can now be used for random variables with arbitrary marginals. For example, we could have exponentially distributed random variables and then imposed a Gaussian dependence structure by choosing the Gaussian copula to model their joint distribution.

A multidimensional Gaussian copula is defined analogously. We start with normally distributed random variables $X_1, \ldots, X_I$ with means $\mu_1, \ldots, \mu_I$, variances $\sigma_1, \ldots, \sigma_I$ and correlation matrix R. Then define the uniform random variables

$$U_i = \Phi\left(\frac{X_i - \mu_i}{\sigma_i}\right), \tag{182}$$

and set C to be their joint distribution.

Consider now a χ^2 random variable Y with ν degrees of freedom and suppose that Y is independent of $X_1, \ldots, X_I$. Define the uniform random variables

$$U_i = t_\nu\left(\frac{\sqrt{\nu}}{\sqrt{Y}} X_i\right), \tag{183}$$

where $t_\nu(\cdot)$ denotes the Student-t distribution with ν degrees of freedom. Then their joint distribution function is called a *t-copula*. The main difference between a Gaussian and a t-copula is the weight their assign to correlated events happening in the tails of the marginal distributions.

To formalize this concept, let C be a bivariate copula for continuous underlying random variables. We say that C is *lower tail dependent* if

$$\lim_{u \to 0} \frac{C(u,u)}{u} \tag{184}$$

exists and takes a value in $(0, 1]$. Analogously, we say that C is *upper tail dependent* if

$$\lim_{u \to 1} \frac{1 + C(u,u) - 2u}{1 - u} \tag{185}$$

exists and takes a value in $(0, 1]$. A lower tail dependent copula tends to generate low values in all marginals simultaneously, while an upper tail dependent copula tends to generate high values in all marginals simultaneously.

Therefore, one kind of copula is relevant to model multiple default scenarios in short time intervals while the other is relevant to model the probability of joint defaults in the indefinite future. One can verify that the Gaussian copula is tail independent, while the t-copula presents both upper and lower dependence.

11.4 Archimedean copulas

In general, there as as many I–dimensional copula functions as there are random variables on $[0,1]^I$. Clearly this gives way too much freedom for modeling choices. In practice, one tries to concentrate on parametric families of copulas, such as the Gaussian and t-copulas presented earlier. The purpose of this section is to introduce another popular class of copula functions.

We say that a copula function C is *Archimedean* if there exist a function $\psi : [0,1] \to \mathbb{R}^+$ with $\psi(1) = 0$ and $\phi(0) = \infty$ such that

$$C(t_1, \dots, t_I) = \phi^{-1}\left(\phi(t_1) + \cdots + \phi(t_I)\right). \tag{186}$$

The function ψ is then called the *generator* of the copula.

Observe that, from the point of view of an Archimedean copula, the random variables are interchangeable, in the sense that the correlation between any two of the underlying random variables does not depend on the identity of the random variables. This is particularly useful for modeling large homogeneous portfolios.

Archimedean copulas can produce a variety of tail dependence. For example, the *Clayton* copula has generator

$$\psi(t) = t^{-\theta} - 1, \qquad \theta > 0, \tag{187}$$

from which we can calculate both the lower tail dependence coefficient

$$\lim_{u \to 0} \frac{C(u,u)}{u} = \frac{1}{2^{1/\theta}} > 0$$

and the upper tail dependence coefficient

$$\lim_{u \to 1} \frac{1 + C(u,u) - 2u}{1-u} = 0.$$

Therefore the Clayton copula is lower tail dependent but upper tail independent.

For the opposite case, consider the *Gumbel* copula, whose generator is

$$\psi(t) = (-\log t)^{\theta}, \qquad \theta \geq 1. \tag{188}$$

We then have

$$\lim_{u \to 0} \frac{C(u,u)}{u} = 0$$

and

$$\lim_{u \to 1} \frac{1 + C(u,u) - 2u}{1-u} = 2 - 2^{1/\theta} > 0.$$

Therefore the Gumbel copula is lower tail independent but upper tail dependent.

For a final example, consider the *Frank* copula, with generator

$$\psi(t) = -\log\left(\frac{e^{\theta t} - 1}{e^{\theta} - 1}\right), \qquad \theta \neq 0. \tag{189}$$

We find that

$$\lim_{u \to 0} \frac{C(u,u)}{u} = 0$$

and

$$\lim_{u \to 1} \frac{1 + C(u,u) - 2u}{1-u} = 2 - 2^{1/\theta} = 0,$$

so that the Frank copula is tail independent.

11.5 An example of default modeling with copulas

Let us start with a static model, that is, one for default or survival over o fixed time interval $[0,T]$. As before, consider I obligors and denote by p_i the probability that the i–th one defaults before T. Given a copula function C we can implement this input data by drawing uniform random variables $U_1, \ldots, U_n$ with distribution C and say that the i-th obligor survived if and only if

$$U_i \leq 1 - p_i.$$

From this we can obtain obtain several joint survival probabilities. For instance, the probability of survival of a given set $I_S \subset \{1, \ldots, I\}$ of obligors, with obligors in I_S^c either surviving or defaulting, is given by $C(u_1, \ldots, u_I)$ where

$$u_i = \begin{cases} 1 - p_i & \text{if } i \in I_S \\ 1 & \text{otherwise} \end{cases}$$

In particular, the probability of no default is $C(1-p_1,\ldots,1-p_I)$, the probability of survival of the first k obligors is $C(p_1,\ldots,p_k,1\ldots,1)$.

The combination of survival and default probabilities is more delicate. Let us denote by $P(I_S,I_D)$ the probability that obligors in the set I_S survive *and* obligors in the set I_D default, where we assume that $I_S\cap I_D=\emptyset$ but not necessarily that $I_S\cup I_D=\{1,\ldots,I\}$. Note that we are silent about obligors who are neither in I_S nor in I_D, they might survive or default. Then, for any $j\notin I_S\cup I_D$ we have

$$P(I_S,(I_D\cup\{j\}))=P(I_S,I_D)-P(I_S\cup\{j\},I_D). \tag{190}$$

Therefore, to calculate the probability of an event with exactly n defaults we need to perform 2^n recursions and be able to efficiently calculate the copula for any dimension less than I. Archimedean copulas than present an analytically feasible framework for these calculations. As before, more explicit formulas can be obtained in a large portfolio approximation.

To introduce some dynamics in the model, suppose that we know the term structure of survival probabilities $P_i(0,t)$ for the different obligors. Given a copula function, we can implement these initial data by drawing uniform random variables $U_1,\ldots,U_n$ with distribution C and setting the default times $\tau_1,\ldots,\tau_n$ to satisfy

$$P_i(0,\tau_i)=U_i. \tag{191}$$

12 Twelfth Lecture

We conclude this course with several examples of credit derivatives and their prices under the different model introduced so far. Most of these derivatives share the following fundamental structure: a protection buyer **A**, a protection seller **B** and a set of reference obligors $\mathbf{C}_1, \ldots, \mathbf{C}_I$. Their purpose is to transfer all or part of the obligor's default risk from **A** to **B**. In return, **B** receives some financial compensation from **A**, be it in the form of a upfront fee, a payment stream or the possibility of high returns.

12.1 Credit Default Swaps

In this basic derivative, the protection buyer **A** pays the protection seller B a regular fee s_C at fixed intervals until maturity if no default happens, in which case **B** doesn't have to pay anything. If **C** defaults before the maturity of the CDS, then **B** has to make a default payment. The default payment is specified in the contract, but typically nets to $(1-R_\tau)$ times the notional of the contract, where R_τ is the recovery rate prevailing at the default time τ. For instance, it can correspond to the physical purchase of **C**-bonds from **A** at par, which can then be sold back in the market by a fraction R_τ of its face value. In this way, if **A** owns assets associated with **C**, their default risk is completely transfered to **B**, while **A** still retains their market risk.

The notional amount $\mathcal{N}$ for a typical CDS ranges from 1 million to several hundred millions of US dollars. The fee payment is quoted as an annualized rate on the notional, with the usual irritating "basis points" jargon, according to which 100 bp = 1%. By the price of a CDS what we actually mean is the determination of this rate, also called the CDS *spread* for the obligor **C**. The maturity T of a CDS usually ranges from 1 to 10 years. The fees are arranged to be paid are specified dates $\mathcal{T} = \{0 < T_1, \ldots, T_K = T\}$.

We can now price a CDS in terms of more fundamental objects. Let us assume that the recovery rate has constant risk-neutral expected value R and that, if a default happens in the interval $(T_{k-1}, T_k]$, the default payment is made at T_k. Then the market value at time $t < T_1$ of a default payment at time T_k is

$$\begin{aligned} b_k &= (1-R)\mathcal{N} E^Q \left[e^{-\int_t^{T_k} r_s ds} \left(\mathbf{1}_{\{\tau > T_{k-1}\}} - \mathbf{1}_{\{\tau > T_k\}} \right) \right] \\ &= (1-R)\mathcal{N} \left(E^Q \left[e^{-\int_t^{T_k} r_s ds} \mathbf{1}_{\{\tau > T_{k-1}\}} \right] - \overline{P}_{tT_k} \right) \end{aligned}$$

and the market value at time $t < T_1$ of a fee payment at time T_k is

$$\begin{aligned} a_k &= s(T_k - T_{k-1})\mathcal{N} E^Q \left[e^{\int_t^{T_k} r_s ds} \mathbf{1}_{\{\tau > T_k\}} \right] \\ &= s(T_k - T_{k-1})\mathcal{N} \overline{P}_{tT_k} \end{aligned}$$

Therefore the total value of the CDS at time t is

$$\text{CDS}(t, s, \mathcal{N}, \mathcal{T}) = \sum_{k=1}^{K} (b_k - a_k). \tag{192}$$

As usual, the spread s is chosen so that the initial value of the CDS is zero.