Math 774 - Credit Risk Modeling

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

January 3, 2010

Contents

1 Introduction **7**
- 1.1 Types of Financial Risk 7
- 1.2 The Nature of this Book 7

2 Bond Market Basics **9**
- 2.1 Default-free bonds and default-free interest rates 9
- 2.2 Defaultable bonds and credit spreads 13
- 2.3 Interest Rate and Credit Derivatives 14
  - 2.3.1 Bonds and floating rate notes 14
  - 2.3.2 Interest rate swaps 15
  - 2.3.3 Credit Default Swaps 16
  - 2.3.4 Options on bonds 18
- 2.4 Exponential Default Times 20
- 2.5 Exercises 20

3 Modeling of Interest Rates **23**
- 3.1 Differentials 23
- 3.2 One Factor Short Rate Models 25
  - 3.2.1 Market price of risk 25
  - 3.2.2 Affine Models 27
  - 3.2.3 The Vasicek Model 28
  - 3.2.4 The Dothan Model 29
  - 3.2.5 The Exponentiated Vasicek Model 30
  - 3.2.6 The Cox–Ingersoll-Ross Model 30
- 3.3 Fitting the initial term structure 31
  - 3.3.1 Parameter Estimation and the Initial Term Structure 31
  - 3.3.2 The Ho–Lee Model 32
  - 3.3.3 The Hull–White Extended Vasicek Model 33
  - 3.3.4 Deterministic–Shift Extensions 34
- 3.4 Forward Rate Models 35
- 3.5 General Option Pricing Formulas 36
- 3.6 Exercises 38

4 Structural Models of Credit Risk **41**
4.1 The Merton Model (1974) 41
4.2 Black-Cox model 44
4.3 KMV 45
4.4 Default Events and Bond Prices 47
4.4.1 Unconditional default probability 47
4.4.2 Conditional default probability 48
4.4.3 Implied Survival Probabilities and Credit Spreads 49
4.5 Exercises 50

5 Reduced Form Modelling **53**
5.1 Information Sets 53
5.2 Basic examples of reduced form models 54
5.2.1 Poisson processes 54
5.2.2 Inhomogeneous Poisson processes 55
5.2.3 Cox processes 56
5.3 Definition of reduced form models 58
5.4 Constructing reduced form models 61
5.4.1 Construction A 61
5.4.2 Construction B 62
5.4.3 Simulating the default time 62
5.5 Affine Intensity Models 63
5.5.1 CIR Intensities 63
5.5.2 Mean-reverting intensities with jumps 64
5.6 Risk-neutral and physical measures 64
5.7 Exercises 65

6 Further Reduced Form Modelling **69**
6.1 Recovery Models 69
6.1.1 Recovery of par 71
6.1.2 Recovery of treasury 71
6.1.3 Recovery of market value 72
6.1.4 Stochastic recovery models 72
6.2 Two-factor Gaussian models 73
6.3 General Two Factor Models 74
6.4 Credit Derivatives 75
6.4.1 Credit Default Swaps 75
6.5 Credit Rating Models 76
6.5.1 Discrete-time Markov chain 77
6.5.2 Continuous-time Markov chain 78
6.6 Exercises 80

7 Portfolio Credit Risk **85**
7.1 Credit Risk Basics 85
7.1.1 A Single Obligor 85
7.1.2 Multiple Obligors 86
7.1.3 Economic Capital 87
7.1.4 Market Risk 87
7.2 Modeling Dependent Defaults 88
7.3 The Binomial Expansion Technique 88
7.4 Default Correlation in Structural Models 89
7.4.1 CreditMetrics Approach 90
7.4.2 First Passage Models 91
7.4.3 Factor Models of Correlation 91
7.5 Default correlation in reduced-form models 93
7.5.1 Doubly stochastic models 93
7.5.2 Joint default events 95
7.5.3 Infectious defaults 95
7.6 Definition of Copula Functions 95
7.6.1 Fréchet bounds 96
7.6.2 Normal Copula and Student t-Copula 98
7.6.3 Tail dependence 99
7.6.4 Archimedean copulas 99
7.6.5 One factor normal copula default model 101
7.6.6 An example of default modeling with copulas [NOT COMPLETE] . 102
7.7 Exercises 103

8 CDOs and other basket credit products **105**
8.1 kth to default derivatives 105
8.1.1 First to Default Swap 105
8.1.2 mth-to-Default Swaps 107
8.1.3 First-m-to-Default Swap in Normal Copula Models 107
8.1.4 Nonhomogeneous Probabilities: Recursion Algorithm 108
8.1.5 Index Credit Default Swaps 108
8.2 Collateralized debt obligations 109
8.3 Pricing CDOs in Copula Models 112
8.3.1 Some specific CDOs 112
8.3.2 Pricing a Homogeneous CDO 113
8.3.3 Pricing a Nonhomogeneous CDO 114
8.3.4 Pricing CDOs with Varying Notionals 115
8.4 Exercises 115

A Mathematical Toolbox **117**
A.1 Itô diffusions 117
A.2 The Strong Markov Property and the Markov Generator 118

A.3 Itô Formula 119
A.4 The Feynman-Kac Formula 119
A.5 Kolmogorov Backward Equation 120
A.6 Ornstein-Uhlenbeck Processes 120
A.7 Girsanov Theorem 121
A.8 Arbitrage Pricing Theory 122
A.9 Change of Numeraire 123
A.10 First Passage Time for Brownian Motion with Drift 125
A.11 Multivariate Normal Distributions 126
A.12 Exercises 127

Chapter 1

Introduction

1.1 Types of Financial Risk

Duffie and Singleton [6] identify five categories of risk faced by financial institutions:

- Market risk: the risk of unexpected changes in prices;
- Credit risk: the risk of changes in value due to unexpected changes in credit quality, in particular if a counterparty defaults on one of their contractual obligations;
- Liquidity risk: the risk that costs of adjusting financial positions may increase substantially;
- Operational risk: the risk that fraud, errors or other operational failures lead to loss in value;
- Systemic risk: the risk of market wide illiquidity or chain reaction defaults.

This rough classification reflects the wide ranging character of risk as well as the methods industry uses to manage them. But clearly the classification is very blurry: for example, the risk that credit spreads will rise can be viewed as both a form of market risk and a form of credit risk.

In order to keep focus, this book will to a great extent restrict attention to market and credit risk, where the credit risk component will almost always refer to medium to large corporations. We will adopt the philosophy that because the drivers of credit risk are strongly correlated with drivers of market risk, it is important to deal with the joint nature of market and credit risk, and therefore the careful risk manager should not try to separate them.

1.2 The Nature of this Book

The primary aim of this book is to provide a coherent *mathematical* development of the theory that describes fixed income markets, both in the default-risk-free and the default-risky settings, that will allow the rigorous pricing, hedging and risk management of portfolios of government and corporate bonds. By extension, however, we must also address the modeling of credit derivatives, that is contingent claims written on credit risky underliers. To maintain mathematical clarity, we will confine discussion of points of industrial or financial background to "Remarks", separated from the main body of the text. These remarks are needed for a correct interpretation of the mathematical results, but not for the logical development. Questions concerning mathematical foundations or related mathematical issues that can be considered elementary or tangential to the main line will often be dealt with in Appendix A. Examples given in the text provide illustrations of the central mathematical notions, but are not necessary for the strict logical development. However, of central interest to mathematicians are the definitions, modeling hypotheses, and propositions: these we will highlight as best as we can in the body of the text.

Unless otherwise stated, we consider a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions of right-continuity and completeness. Usually P will have the meaning of the physical or historical measure; other relevant measures equivalent to P, for example the risk-neutral measure Q, will also be needed. The sigma-algebra $\mathcal{F}_t$ will usually have the meaning of the full market information available at time t. Asset price processes will always be assumed to be semimartingales[1], and our models to be free of arbitrage. In what follows, W_t denotes a standard P–Brownian motion of unspecified dimension.

[1]For a detailed discussion of this probabilistic setting, please see [21]

Chapter 2

Bond Market Basics

Bonds, the securitized version of loans, are the basic type of tradeable financial contract by which corporations and governments tap into the capital available from investors. While a varied array of such contracts are traded, the essential common feature of bonds is that the bond purchaser pays present cash for a fixed future cash flow. Equities (i.e. shares), on the other hand, finance the firm by conferring on the holder a fractional ownership of a corporation, and thus a fraction of the random future earnings of the firm.

While bond contracts are normally specific about the future cash flows they generate, those issued by corporations or the governments of developing countries carry a degree of risk that the contract provisions may be broken. This is the essence of what we mean by "default risk". Such is the nature of markets that at the moment a medium to large financial institution breaks the provisions of a particular bond contract by failing to make a single required payment, then at that time the market assumes that all other outstanding bond contracts written by that firm will fail. Thus failure of a firm to make a single contractual payment can be considered the trigger that causes the firm to suspend "normal operations" and enter a state of bankruptcy. This time is called the *default time* for the firm, and is represented mathematically by a *stopping time* τ, that is a random variable that takes values from the infinite interval $[0, \infty]$.

2.1 Default-free bonds and default-free interest rates

In contrast to corporate bonds, the bonds issued by sovereign governments of developed countries can for practical purposes be taken to be free of default risk. They are, however, still subject to volatility risk, since their prices are highly sensitive to fluctuations in interest rates. In this section, we focus on default-free bonds and the associated term structure of interest rates.

Definition 1. A *zero-coupon bond (or "zero")* is a contract that pays the holder one unit of currency at its maturity time T. Let $P_t(T) = P_t^T$ denote the price at time $t \le T$ of a default-free zero-coupon bond maturing at time T. The *time to maturity* $T - t$ is the amount of time, in years, from the present time to the maturity time $T \ge t$.

Here are the standing assumptions we make as we develop the modeling principles for default-free interest rates.

Assumption 1. We assume that there exists a frictionless[1] *arbitrage-free* market for bonds of all maturities $T > t$. They are default-risk-free, hence $P_T(T) = 1$ holds for all T. Finally, we assume that, for each fixed t, $P_t(T)$ is differentiable with respect to T almost surely.

The following two simple consequences of the no-arbitrage principle are proved as an exercise.

Proposition 2.1.1. *1. The contract that pays the value $P_S(T)$ at time S has the time $t < S$ value $P_t(T)$.*

2. If X is any $\mathcal{F}_t$ random variable, then a contract that pays X at a future date $T > t$ has the time t value $XP_t(T)$.

$P(t, T)$ is standard notation, but we prefer the notation $P_t(T)$, since it follows the convention that the subscript t denotes a stochastic process, and the bracketed variable T denotes a possibly continuous range of parameter values.

Remark 1. Zero-coupon bonds, also called strip bonds, while mathematically convenient, are a relatively recent invention, having been introduced by the US Treasury in 1982 and subsequently made popular by local and municipal governments. The vast majority of bonds mature at specific dates, typically ranging from one to 40 years (the majority being between 8 to 20 years), and pay coupons, typically at six month or one year intervals. Coupon bonds can be decomposed as sums of zeros. Short term government bonds are called treasury bills, and typically pay no coupon. Bond markets, albeit voluminous, are less liquid than equity markets, and are typically the domain of the institutional investor.

Remark 2. As long as both T and t are expressed as real numbers, the difference $T - t$ is unambiguous. However, if dates are represented in day/month/year format, then different day-counting conventions result in different values for the time to maturity. In what follows, we will largely ignore this cantankerous issue, unless specific contracts force us to do otherwise. As a curiosity, one finds in [13] that 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).

Different notions of interest rate are defined in terms of zero-coupon bond prices. Consider the present date t and two future dates S and T with $t < S < T$.

Definition 2. 1. The *continuously compounded forward rate* for the period $[S, T]$ is the rate $R_t(S, T)$ satisfying

$$e^{R_t(S,T)(T-S)} P_t(T) = P_t(S) \qquad \forall t < S < T, \tag{2.1}$$

[1] This means we assume the usual efficient market assumptions, such as no transaction costs, zero bid-ask spreads, small trades that do not “move the market”, unlimited short selling, etc.

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

2. The *simply compounded forward rate* for the period $[S,T]$ is the rate $L_t(S,T)$ satisfying
$$[1 + L_t(S,T)(T-S)]P_t(T) = P_t(S) \qquad \forall t < S < T, \tag{2.2}$$
that is, $L_t(S,T)$ is the unique rate that 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 $S = t$ in the above definitions, we obtain the *continuously compounded yield* $R_t(T)$, defined by
$$e^{R_t(T)(T-t)}P_t(T) = 1, \qquad \forall t < T \tag{2.3}$$
and the *simply compounded yield* $L_t(T)$, defined by
$$[1 + L_t(T)(T-t)]P_t(T) = 1 \qquad \forall t < T. \tag{2.4}$$
Thus, $L_t(T)(T-t)$ represents the amount of interest paid at time T ("in arrears", meaning at the end of a period) on a loan of \$1 made for the period (t,T).

Remark 3. We use the notation $L_t(S,T), L_t(T)$ above because LIBOR (London Interbank Offered Rates), fixed daily in London, are the prime examples of simply compounded rates. The British Bankers' Association publishes daily LIBOR values for a range of different currencies. Thus on the date t, in each currency, LIBOR rates are quoted which determine the simple interest to be charged for loans between banks for periods (t,T) with a range of maturity dates T, in other words, precisely the values $L_t(T)$. The actual determination of these values is obtained by averaging quotes for loans from a number of contributing banks. While no arbitrage implies $L_t(T)$ must satisfy (2.4), in practice, market inefficiencies, and a small degree of credit riskiness (historically at the level of 30 to 50 bps) lead to small inconsistencies. There are approaches to fixed income modeling that take LIBOR rates as the fundamental market variables, and derive the prices of zeros in terms of them by no arbitrage.

In the post 2007 market environment, LIBOR rates can no longer be taken as even approximately default-risk free. The so-called *TED spread*, the difference between the 3 month Eurodollar (LIBOR) rate and the 3 month T-bill rate, has recently blown out to as much as 475 bps, reflecting the extreme perception of counterparty risk between banks that has been experienced during the credit crunch.

The following example illustrates the fundamental nature of simple interest $L_S(T)$ when paid in arrears (i.e. at time T).

Example 1. If $t \leq S < T$, what is the time t value of a contract that pays $L_S(T)$ at time T?

Answer: By proposition 2.1.1, the time S value of the contract is $L_S(T)P_S(T)$. From the definition, $L_S(T)P_S(T) = (1 - P_S(T))/(T - S)$, meaning the contract is replicable by a portfolio strategy of bonds. Using proposition 2.1.1 twice more to value the portfolio on the right side, one finds that the contract has the time t value $(P_t(S) - P_t(T))/(T - S)$, which equals $L_t(S, T)P_t(T)$.

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

$$f_t(T) := \lim_{S \to T-} L_t(S, T) = \lim_{S \to T-} R_t(S, T) = -\frac{\partial \log P_t(T)}{\partial T}. \tag{2.5}$$

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

$$r_t := \lim_{T \to t+} L_t(T) = \lim_{T \to t+} R_t(T) \tag{2.6}$$

and it is easy to verify that $r_t = f_t(t)$. Moreover, we readily obtain that

$$P_t(T) = \exp\left(-\int_t^T f_t(s)ds\right). \tag{2.7}$$

Finally, we assume the existence of an important idealized asset defined using the short rate r_t. It has the meaning of a limiting case of the investment strategy that "rolls over" bonds of shorter and shorter maturity.

Assumption 2. There is a tradeable asset called the *money-market account*, defined as the stochastic process satisfying

$$dC_t = r_t C_t dt, \qquad C_0 = 1 \tag{2.8}$$

Thus, cash can be invested in an asset that is riskless over short time periods, and accrues interest at the short rate r_t.

We have that

$$C_t = \exp\left(\int_0^t r_s ds\right). \tag{2.9}$$

and thus we make the

Definition 3. The *stochastic discount factor* between two times $t < T$ is given by the formula

$$D(t, T) = C_t/C_T = \exp\left(-\int_t^T r_s ds\right), \tag{2.10}$$

and represents the amount at time t that is "equivalent" to one unit of currency payable at time T.

We can see that if we invested exactly $D(t,T)$ 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 $P_t(T)$ and the stochastic discount factor. Their difference resides in the fact that $P_t(T)$ is the value of a contract and therefore must be known at time t, while $D(t,T)$ is a random quantity at t depending on the evolution of the short rate process r over the future period (t,T). While for deterministic interest rates we have that $P_t(T) = D(t,T)$.

In general, however, for stochastic rates r_t, bond prices are expected values of the discount factor under an appropriate measure. This can be seen informally as follows. We know[2] that a given set of price processes $\{Y_t^I\}_{i=1,2,\ldots}$ (for non-dividend paying assets) will be free of arbitrage if and only if there is *some* measure $Q \sim P$ such that each discounted price process $C_t^{-1}Y_t^i$ is a Q-martingale. This means that for all $t < s$

$$C_t^{-1}Y_t^i = E^Q[C_s^{-1}Y_s^i|\mathcal{F}_t]$$

In particular, if $Y_t^i = P_t(T)$ then

$$P_t(T) = C^t E^Q[C_T^{-1}|\mathcal{F}_t] = E^Q[D(t,T)|\mathcal{F}_t].$$

2.2 Defaultable bonds and credit spreads

The term structure of risk free interest rates, in particular LIBOR rates, is a property of the entire economy of a developed country at a moment in time, and can be regarded as a market observable. By contrast, the term structure of credit risky bonds issued by corporations (or developing countries) depends on the specific nature of the issuer, most importantly whether the firm is *solvent* (i.e $\tau > t$) or *bankrupt* ($\tau \leq t$).

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

$$\begin{aligned} \bar{P}_T(T)\mathbf{1}_{\{\tau>T\}} &= \mathbf{1}_{\{\tau>T\}} = P_T(T)\mathbf{1}_{\{\tau>T\}}, \\ \bar{P}_T(T)\mathbf{1}_{\{\tau\leq T\}} &< \mathbf{1}_{\{\tau\leq T\}} = P_T(T)\mathbf{1}_{\{\tau\leq T\}}, \end{aligned} \tag{2.11}$$

the Law of One Price dictates that

$$\bar{P}_t(T)\mathbf{1}_{\{\tau>t\}} < P_t(T)$$

for all earlier times t (the inequality is strict as long as $P[\tau \leq T|\tau > t] > 0$). In parallel with default-free interest rates, if we assume that the firm's bonds exist for all maturities $T > t$ and that $\bar{P}_t(T) > 0$ is differentiable in T, then we can define the *default risky forward rate* $\bar{f}_t(T)$ by

$$\bar{P}_t(T) = e^{-\int_t^T \bar{f}_t(u)du} \tag{2.12}$$

[2] This is a version of the Fundamental Theorem of Asset Pricing discussed in the Appendix.

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, hence the difference $\bar{f}_t(s) - f_t(s)$ is non-negative almost surely. Such a difference is called a *credit spread*. For example, the *yield spread* (YS) and *forward spread* (FS) at time t for maturity $T > t$ are given by

$$\mathrm{YS}_t(T) = \frac{1}{T-t}\int_t^T \mathrm{FS}_t(s)ds = \frac{1}{T-t}\int_t^T (\bar{f}_t(s) - f_t(s))ds = \frac{1}{T-t}\log\left(\frac{P_t(T)}{\bar{P}_t(T)}\right), \quad (2.13)$$

One can also define the *simply compounded defaultable forward rate* or *defaultable LIBOR rate* $\bar{L}_t(S,T)$ for the period $[S,T]$ by

$$[1 + \bar{L}_t(S,T)(T-S)]\bar{P}_t(T) = \bar{P}_t(S).$$

2.3 Interest Rate and Credit Derivatives

In this section we review the simplest interest rate derivatives, and then describe several further examples of interest rate derivatives and provide simple arbitrage arguments to establish relations between them. The actual pricing of these derivatives requires the martingale methodology that will be addressed in a later chapter.

2.3.1 Bonds and floating rate notes

Recall that the simplest possible interest rate derivative is a *zero-coupon bond* with maturity T, whose price we denote by $P_t(T)$ or $\bar{P}_t(T)$. In this section we focus on the default-free case, but analogous products can be defined based on defaultable corporate bonds.

In practice, most traded bonds are *coupon-bearing bonds* that pay deterministic amounts $c = (c_1, \ldots, c_N)$ at specified times $\mathcal{T} = (T_1, \ldots, T_N)$. Such a bond can be replicated by a portfolio of zero-coupon bonds with maturities T_n, $n = 1, \ldots, N$ and notional amounts c_n and thus its price at time t is

$$P(t, c, \mathcal{T}) = \sum_{n=1}^{N} c_n P_t(T_n). \quad (2.14)$$

Remark 4. In North America, the standard convention for a coupon bond with maturity T, face value P and coupon rate c delivers half yearly coupons equal to $cP/2$ on dates $T - i/2, i = 1, 2, \ldots$ plus a final payment of the face value plus coupon $P(1 + c/2)$ at maturity. Usually the coupon rate c is chosen at the time of issue to be equal (or approximately equal) to the *par coupon rate*, that is, the value that makes the market value of the bond equal to the face value.

In an analogous contract, c_n is not deterministic, but is instead determined by the value of some floating interest rate prevailing for the periods between dates $\mathcal{T} = (T_0 = 0, T_1, \ldots, T_N)$. We assume that T_0 is either the issue date of the contract, or immediately following a contractual payment. The prototypical *floating-rate note* is defined by the payment stream

$$c_n = L_{T_{n-1}}(T_n)(T_n - T_{n-1})\mathcal{N}, \quad n = 1, \dots, N-1,$$

and $c_N = (1 + L_{T_{N-1}}(T_N)(T_N - T_{N-1}))\mathcal{N}$, where $L_{T_{n-1}}(T_n)$ is the simply compounded yield for the period $[T_{n-1}, T_n]$ and $\mathcal{N}$ denotes a fixed notional value. Recalling (2.4), we obtain that

$$c_n = \frac{\mathcal{N}}{P_{T_{n-1}}(T_n)} - \mathcal{N}, \quad n < N.$$

Now observe that at $t = 0$ we can replicate this payment at time 0 by buying $\mathcal{N}$ zero-coupon bonds with maturity T_{n-1} and selling $\mathcal{N}$ zero coupon bonds maturing at time T_n. Therefore the value at time t of the payment c_n is

$$\mathcal{N}(P_t(T_{n-1}) - P_t(T_n)),$$

implying that the value of the floating rate note at time 0 is

$$\mathcal{N}\left[\sum_{n=1}^{N}(P_0(T_{n-1}) - P_0(T_n)) + P_0(T_N)\right] = \mathcal{N}. \tag{2.15}$$

In practitioners' jargon this is expressed by saying that a floating-rate note "trades at par" immediately after any coupon is paid.

A small wrinkle arises in valuing the floating-rate note on a non-coupon date t: since the floating rate is set at the previous coupon date, not at time t, the bond trades near, but not at, par value.

A typical defaultable version of the floating rate note, called the *par floater* the coupon amount is the benchmark short-term rate (i.e. LIBOR) plus a constant spread s^{PF} to cover the extra credit risk, chosen such that the price is initially at par. Under the assumption of zero recovery on the bond at default, the payment stream is then

$$c_n = [L_{T_{n-1}}(T_n) + s^{PF}](T_n - T_{n-1})\mathcal{N}\mathbf{1}_{\{\tau > T_n\}}, \quad n = 1, \dots, N-1, \tag{2.16}$$

and $c_N = [1 + (L_{T_{N-1}}(T_N) + s^{PF})(T_N - T_{N-1})]\mathcal{N}\mathbf{1}_{\{\tau > T_N\}}$.

2.3.2 Interest rate swaps

The next important style of contract gives the holder a loan over a fixed period at a guaranteed rate, or equivalently swaps a fixed rate loan for a floating rate loan.

A simplest such swap is the *forward rate agreement*, which amounts to a loan with notional $\mathcal{N}$ for a single future period $[S, T]$ at a known simple interest rate K. In cash terms, the holder (the borrower) receives $\mathcal{N}$ at time S and repays $\mathcal{N}(1 + K(T - S))$ at time T. From the definition of the simply compounded instantaneous forward rate one can observe that this is equivalent to the contract where the holder receives $\mathcal{N}(1 +$

$L_S(T)(T-S))$ in exchange for $\mathcal{N}(1+K(T-S))$ at time T. The value to the holder of a forward rate agreement at time $t \leq S < T$ is

$$\mathcal{N}[P_t(S) - P_t(T)(1+K(T-S))] = \mathcal{N}P_t(T)(L_t(S,T)-K)(T-S).$$

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

$$K^* = L_t(S,T),$$

which serves as an alternative definition of the simply compounded forward rate $L_t(S,T)$.

A generalization of forward rate agreements for many periods is generically known as an *interest rate swap*. In such contracts, a payment stream based on a fixed simple interest rate K and a notional N is made at dates $\mathcal{T} = (T_1, \ldots, T_N)$ in return for a payment stream based on the same notional, the same periods, but a floating interest rate. At time T_n, for $n = 1, \ldots, N$, the cash flow is the same as that of a forward rate agreement for the period (T_{n-1}, T_n), that is

$$\mathcal{N}(L_{T_{n-1}}(T_n) - K)(T_n - T_{n-1}),$$

Using our previous result, the value at $t = 0$ of this cash flow is

$$\mathcal{N}(P_0(T_{n-1}) - P_0(T_n) - \mathcal{N}K(T_n - T_{n-1})P_t(T_n),$$

so that the total value of the interest rate swap at time 0 is

$$\mathrm{IRS}(\mathcal{N}, \mathcal{T}, K) = \mathcal{N}\left[1 - P_0(T_N) - K\sum_{n=1}^{N} P_0(T_n)(T_n - T_{n-1})\right]. \tag{2.17}$$

Similarly to a forward rate agreement, we can define the swap rate at time $t = 0$ as the rate K^* that makes this contract have value zero. That is

$$K^*(\mathcal{T}) = \frac{1 - P_0(T_N)}{\sum_{n=1}^{N} P_0(T_n)(T_n - T_{n-1})}, \tag{2.18}$$

which one observes is precisely the "par coupon rate" that makes the market value of a coupon bond equal to its face value.

2.3.3 Credit Default Swaps

Credit default swaps have emerged in recent years as the preferred contract for insuring lenders against default of their obligors. In the standard version, party A (the "insured") buys insurance from B (the " insurer") against default of a third party C (the "reference obligor"). Here default is defined in the contract, and typically involves the failure by C to make required payments on one of a specific set of similarly structured reference bonds. The protection buyer A pays the protection seller B a regular fee s_C, called the premium, at fixed intervals (typically quarterly) until either maturity T if no default happens, or default if $\tau \leq T$. If $\tau > T$, B doesnt have to pay anything, but if C defaults before the maturity, 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 value of the reference contract, where R_τ is the recovery rate prevailing at the default time τ. For instance, the default can be settled "physically" by the exchange between A and B of one of the specified reference bonds at its par value. Since these defaulted bonds can then be sold by B in the market at a fraction R_τ of their par value, the default payment nets to $(1 - R_\tau)$ times the notional amount. There is a risk of a "squeeze" if A does not already hold a reference bond: in this case, A is forced to purchase one of the reference bonds, which may be in short supply after default. For this reason, the insurance payment of CDS are now often *cash settled*: several independent dealers are asked to provide quotes on the defaulted bond, and the B pays A the difference between the average quoted value and the par bond value. In either way, the effect of a CDS is that if A owns assets associated with C, their default risk is completely transferred to B, while A still retains their market risk. Finally there is increasing use of a third mechanism for determining post default values of bonds by holding a credit auction. Figure 2.3.3 shows the results of all recent credit auctions held in the US.

The notional amount $\mathcal{N}$ for a typical CDS ranges from 1 million to several hundred millions of US dollars. The fair premium payment rate, called the CDS spread for the obligor C, is the value of s that makes the contract into a swap on the issue date, i.e. it makes the contract have initial value zero. It is quoted as an annualized rate on the notional, with the usual irritating basis points jargon, according to which $100bp = 1\%$. The maturity T of a CDS usually ranges from 1 to 10 years. The fees are arranged to be paid on specified dates $\mathcal{T} = \{0 < T_1, \ldots, T_K = T\}$.

Let us assume that the recovery rate has constant value R, which means that a defaultable T bond, pays

$$\mathbf{1}_{\{\tau > T\}} + R\mathbf{1}_{\{\tau \leq T\}} = R + (1 - R)\mathbf{1}_{\{\tau > T\}}. \tag{2.19}$$

at maturity, showing that a bond with recovery can be expressed as R times a default-free bond plus $1 - R$ times the zero-recovery defaultable bond (whose price we denote by $\overline{P}^0_t(T)$):

$$\overline{P}^R_t(T) = RP_t(T) + (1 - R)\overline{P}^0_t(T), \quad \overline{P}^0_t(T) = \frac{1}{1 - R}[\overline{P}^R_t(T_k) - RP_t(T_k)]. \tag{2.20}$$

As for the CDS itself, we suppose that if a default happens in the interval $(T_{k-1}, T_k]$, the default payment is settled physically at T_k, meaning the insurer B buys the defaulted bond from A for its par value. Then the *default payment* at time T_k is

$$(1 - R)\mathcal{N}(\mathbf{1}_{\{\tau > T_{k-1}\}} - \mathbf{1}_{\{\tau > T_k\}}) \tag{2.21}$$

The second term is the payoff of a $1 - R$ times a defaultable zero recovery zero coupon bond, and can be valued in terms of $\bar{P}_t(T_k)$ and $P_t(T_k)$ at any time $t < T_k$. The first term however, cannot be replicated by defaultable and default-free bonds: pricing this term at earlier times *requires a model*. Let $b_t^{(k)}$ denote the market value of the default payment at time $t < T_1$. The *premium payment* at time T_k

$$s(T_k - T_{k-1})\mathcal{N}\mathbf{1}_{\{\tau > T_k\}} \tag{2.22}$$

can be replicated by bonds of maturity T_k, and hence the market value at time $t < T_1$ is

$$a_t^{(k)} = s(T_k - T_{k-1})\mathcal{N}P_t^0(T) = \frac{s(T_k - T_{k-1})\mathcal{N}}{1-R}[\overline{P}_t(T_k) - RP_t(T_k)].$$

Here we neglect a so-called accrual term is often added to the premium for the period during which default occurred. The total value of the CDS at time t is

$$CDS(t, s, \mathcal{N}, \mathcal{T}) = \sum_{k=1}^{K}(b_t^{(k)} - a_t^{(k)}). \tag{2.23}$$

Remark 5. The best way to understand the basic CDS contract on a firm is to realize it as a swap of a coupon paying corporate bond for the comparable treasury bond. At the initiation date $t = 0$, the insurance buyer (the "insured") shorts the corporate bond and goes long the treasury, both with the notional amount $\mathcal{N}$; the insurance seller takes the reverse position. We suppose the bonds pay the "par coupon" rates $\bar{c}$ and c respectively on the same sequence of dates (suppose they are half yearly coupons): this means both bonds are worth $\mathcal{N}$ at time 0, and the contract is a swap (i.e. has zero value) at time 0. The only other contract detail to resolve is the default mechanism. We assume that at the time of default, the contract is unwound as follows: the insured buys back the defaulted bond (for its market-determined post default value) and sells the treasury bond (again, for the market value).

Some thought will then convince oneself that this swap contract replicates the payments of the basic CDS. When default does not happen over the lifetime of the contracts, the payments net to $s_c = \bar{c} - c$ on each coupon date. When default happens before maturity, the default settlement nets to the difference between the default free bond (whose value will be near $\mathcal{N}$) minus the post default value of the corporate bond (which can be written $R\mathcal{N}$ with the recovery fraction $R < 1$): this amounts to the insured collecting an insurance payment that covers the loss given default.

2.3.4 Options on bonds

The contracts considered so far all have an important theoretical feature that they are replicable by a strategy that trades in the underlying bonds on the payment dates, or equivalently a static bond portfolio constructed at time 0. This implies their prices depend on the initial term structure $\{P_0(T)\}_{T>0}$ but are otherwise independent of the interest rate model chosen. This key simplification is not true of the option contracts we now consider.

First let us introduce options on zero-coupon bonds. A European *call option* with strike price $\mathcal{K}$ and maturity date S on an underlying T-bond with $T > S$ is defined by the pay-off at time S

$$(P_S(T) - \mathcal{K})^+.$$

Its value at time $t < S < T$ is denoted by $c(t, S, \mathcal{K}, T)$. Similarly, a 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

$$(\mathcal{K} - P_S(T))^+.$$

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 of such a caplet is therefore buying protection against an increase in the floating rates above the cap rate. Using (2.4), this pay-off can be expressed as

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

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

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

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

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

A *swaption* gives the holder the option to enter into a specific interest rate swap contract on a certain date T.

2.4 Exponential Default Times

We have just seen that while simple derivatives like the interest rate swap can be replicated by bonds, and hence priced without a model, the more complicated derivatives, including the credit default swap and bond options, require a model to price them consistently. Model building requires the machinery described in the section on Arbitrage Pricing Theory in Appendix A.8. In this section we show how all of these securities can easily be priced in the particularly simple default model where τ is taken to be an exponential random variable. More specifically, we assume

Assumption 3 (Exponential Default Model)**.** Under the risk-neutral measure Q, τ, the default time of the firm, is taken to be an exponential random variable with parameter λ, independent of the interest rate process r_t. We also assume defaultable securities pay a constant recovery value R at the time of default.

For example, assuming $R = 0$, we can see that a zero coupon zero recovery defaultable bond has price

$$\bar{P}_t(T)\mathbf{1}_{\{\tau>t\}} = E_t^Q[e^{-\int_t^T r_s ds}\mathbf{1}_{\{\tau>T\}}] \quad (2.24)$$

$$= P_t(T)e^{-\lambda(T-t)}\mathbf{1}_{\{\tau>t\}} \quad (2.25)$$

implying that the credit spread is simply a constant $\bar{f}_t(s) = \lambda$. Other prices of other credit derivatives are left as exercises.

2.5 Exercises

Exercise 1. Prove proposition 2.1.1.

Exercise 2. If $t < S < T$, what is the time t value of a contract that pays $L_t(S,T)$ at time T? What is the value of this contract at time $s < t$?

Exercise 3. Consider a contract that pays the holder \$1 at time S, but requires the holder to make a payment of K at time T. Given a term structure of bond prices at time $t < S \wedge T$ (i.e. the bond prices $P_t(U), U > 0$, compute the fair value of the contract at time t. Let $\tilde{K}$ be the value of K that makes the contract a "swap", that is a contract with value 0. Relate $\tilde{K}$ to the forward interest rates $R_t(S,T)$ and $L_t(S,T)$ when $S > T$ and when $T < S$.

Exercise 4. Let Δt be a fixed time interval, and N a positive integer. Consider a contract that pays the holder a sequence of payments of the amounts $L_{t_{i-1}}(t_i)$ on the dates $t_i = i\Delta t, i = 1, 2, \ldots, N$. Find the fair value of the contract at time 0.

Exercise 5. An *interest rate swap* can be defined for any increasing sequence of times $t_0 = 0 < t_1 < t_2 < \cdots < t_N = T$ as follows. The contract "swaps" interest payments at a fixed simple rate K for interest payments paid at the natural floating (random) rate. That is, on the dates $t_i, i = 1, 2, \ldots, N$, the "fixed leg" pays $(t_i - t_{i-1})K$, while the "floating leg" pays $(t_i - t_{i-1})L_{t_{i-1}}(t_i)$. Find the fair value V_0 of the swap at time 0 (Hint: use the previous assignment to value the floating leg, and create a strategy of bond trades that replicates the fixed leg.) Compute the value of K that makes the swap have $V_0 = 0$.

Exercise 6. Show that under the T-forward measure, $f_t(T)$ is a martingale. What is the price at time $s \le t$ of a contract that delivers the value $f_t(T)$ at time T? (You will need to assume some regularity to make this argument work.)

Exercise 7. Consider the simplest possible credit model:

- Time consists of one period that starts "now" and ends one year "later.
- The default-free and credit-risky debt mature at the end of one year.
- The tenor for the coupon is one year. Let c denote the coupon rate and r the default free simple interest rate.
- The bond defaults, with risk-neutral probability π. If it defaults, the bondholder recovers the fraction $\rho < 1$ of the promised payment of interest and repayment of principal.

1. Assuming that the risky bond trades today at par, compute the fair coupon rate c in the one-period binomial model implied by these statements.
2. Compute the fair credit spread $\log(1 + c) - \log(1 + r)$.
3. In the two-period/two year binomial model, compute today's price of an option to purchase a one year credit risky bond one year from now for \$0.50.
4. Compute the value of the premium and default legs of a 4 year CDS for protection on the 4 year default risky bond with annual coupons c.

Exercise 8. Compute the following contracts in the exponential default model with parameter λ and constant interest rate r.

1. Find a formula for the spread of a par floater with the following specifications: maturity 5 years, 6 month coupons,
2. Find the price of a defaultable T-bond option with maturity $S < T$, strike K. Express the answer in terms of the price of a similarly structured default free bond.
3. Compute the fair value of both the default and premium legs of a CDS with maturity 5 years, a quarterly payment period, where the premium is paid at the annual rate X per unit notional. Then find a formula for the fair spread $\hat{X}$ that makes the value of the contract have zero value at time zero. The default insurance is paid as follows: at the end of a period in which default has occurred, A and B exchange the defaulted bond for its par value.

Date	Name	Final price as a percentage of par
2005-06-14	Collins & Aikman - Senior	43.625
2005-06-23	Collins & Aikman - Subordinated	6.375
2005-10-11	Northwest Airlines	28
2005-10-11	Delta Airlines	18
2005-11-04	Delphi Corporation	63.375
2006-01-17	Calpine Corporation	19.125
2006-03-31	Dana Corporation	75
2006-11-28	Dura - Senior	24.125
2006-11-28	Dura - Subordinated	3.5
2007-10-23	Movie Gallery	91.5
2008-02-19	Quebecor	41.25
2008-10-02	Tembec Inc	83
2008-10-06	Fannie Mae - Senior	91.51
2008-10-06	Fannie Mae - Subordinated	99.9
2008-10-06	Freddie Mac - Senior	94
2008-10-06	Freddie Mac - Subordinated	98
2008-10-10	Lehman Brothers	8.625
2008-10-23	Washington Mutual	57
2008-11-04	Landsbanki - Senior	1.25
2008-11-04	Landsbanki - Subordinated	0.125
2008-11-05	Glitnir - Senior	3
2008-11-05	Glitnir - Subordinated	0.125
2008-11-06	Kaupthing - Senior	6.625
2008-11-06	Kaupthing - Subordinated	2.375
2008-12-09	Masonite [4] - LCDS	52.5
2008-12-17	Hawaiian Telcom - LCDS	40.125
2009-01-06	Tribune - CDS	1.5
2009-01-06	Tribune - LCDS	23.75
2009-01-14	Republic of Ecuador	31.375
2009-02-03	Millennium America Inc	7.125
2009-02-03	Lyondell - CDS	15.5
2009-02-03	Lyondell - LCDS	20.75
2009-02-03	EquiStar	27.5
2009-02-05	Sanitec [5] - 1st Lien	33.5
2009-02-05	Sanitec [6] - 2nd Lien	4.0
2009-02-09	British Vita [7]	TBA

Figure 2.1: All post-default credit auctions for the period 06/2005 to 02/2009. Source: http://www.communicatorinc.com/information/affiliations/fixings.html .

Chapter 3

Modeling of Interest Rates

In the first chapter we learned that many important fixed income derivatives such as forward rate agreements and swaps, both default free and defaultable, can be priced by model independent formulas involving the prices of underlying zero coupon bonds. However, options and other more sophisticated contracts depend on more modeling assumptions. In this chapter, we briefly cover the theory of option pricing in models of the default free interest rate.

3.1 Differentials

First we investigate several formal relationships between bonds, short rates and forward rates, under the assumption that they satisfy stochastic differential equations in a Brownian filtration. These relationships show the connections between the three basic approaches to fixed income modeling.

Let

$$dr_t = a_t dt + b_t dW_t \tag{3.1}$$

$$\frac{dP_t(T)}{P_t(T)} = M_t(T)dt + \Sigma_t(T)dW_t \tag{3.2}$$

$$df_t(T) = \alpha_t(T)dt + \sigma_t(T)dW_t, \tag{3.3}$$

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

Proposition 3.1.1. *1. If $P_t(T)$ satisfies* (3.2)*, then $f_t(T)$ satisfies* (3.3) *with*

$$\alpha_t(T) = \Sigma_t(T)\frac{\partial \Sigma_t(T)}{\partial T} - \frac{\partial M_t(T)}{\partial T},$$

$$\sigma_t(T) = -\frac{\partial \Sigma_t(T)}{\partial T}.$$

2. *If* $f_t(T)$ *satisfies* (3.3)*, then* r_t *satisfies* (3.1) *with*

$$\begin{aligned} a(t) &= \frac{\partial f_t(T)}{\partial T}\Big|_{T=t} + \alpha_t(t) \\ b(t) &= \sigma_t(t) \end{aligned}$$

3. *If* $f_t(T)$ *satisfies* (3.3)*, then* $P_t(T)$ *satisfies* (3.2) *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 $\log P_t(T)$, write in integral form and differentiate with respect to T.

2. For any $t \leq T$, integration of (3.3) gives

$$f_t(T) = f_0(T) + \int_0^t [\alpha_s(T)ds + \sigma_s(T)dW_s] \tag{3.4}$$

$$\partial f_t(T) = \partial f_0(T) + \int_0^t [\partial\alpha_s(T)ds + \partial\sigma_s(T)dW_s]. \tag{3.5}$$

The equations

$$\alpha_s(T) = \alpha_s(s) + \int_s^T \partial\alpha_s(u)du, \quad \sigma_s(T) = \sigma_s(s) + \int_s^T \partial\sigma_s(u)du \tag{3.6}$$

plus $f_0(T) = r_0 + \int_0^T \partial f_0(u)du$ can be inserted into (3.4) with $T = t$, and after interchanging the order of integrations this leads to

$$r_t = r_0 + \int_0^t [\alpha_s(s)ds + \sigma_s(s)dW_s] \tag{3.7}$$

$$+ \int_0^t \left(\partial f_0(u) + \int_0^u [\partial\alpha_s(u)ds + \partial\sigma_s(u)dW_s]\right) du \tag{3.8}$$

We use (3.5) to write this as

$$r_t = r_0 + \int_0^t [\partial f_s(s)ds + \alpha_s(s)ds + \sigma_s(s)dW_s]$$

which is the required result.

3. Write $P_t(T) = \exp Y_t(T)$ 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 (3.3) to arrive at an expression for $dY_t(T)$ and finally the stochastic Fubini theorem to exchange the order of integration.

□

The proposition does not address the question of how modeling the spot rate r_t leads to prices $P_t(T)$ or the forward rate $f_t(T)$. It turns out there is one missing ingredient to be added to the equation (3.1) to fully specify the model. That is the topic of the next section.

3.2 One Factor Short Rate Models

In this section, we consider an economy as specified in A.8 for the special case of $d = 0$. Thus there is only one exogenously defined traded asset, namely the cash account

$$dC_t = r_t C_t dt \tag{3.9}$$

with the short rate of interest solving

$$dr_t = a(t, r_t)dt + b(t, r_t)dW_t. \tag{3.10}$$

We assume the deterministic functions a, b satisfy the usual Lipschitz and boundedness conditions that guarantee existence and uniqueness of solutions of the stochastic differential equation.

In this arbitrage–free market, zero-coupon bonds of all maturities $T > 0$ are treated as endogenous derivatives written on the single factor, namely the spot rate r_t. 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 consistency relations on the possible bond prices, which can be derived as follows.

3.2.1 Market price of risk

The spot rate, modeled here as an Itô diffusion, is therefore Markovian, and it follows that the price of a zero-coupon bond with maturity T is given by $P_t(T) = p^T(t, r_t)$ where p^T is a smooth function of its two variables. From Itô's formula

$$dP_t(T) = M_t(T)P_t(T)dt + \Sigma_t(T)P_t(T)dW_t. \tag{3.11}$$

where $M_t(T) = M^T(t, r_t), \Sigma_t(T) = \Sigma^T(t, r_t)$ with

$$M^T(t,r)p^T(t,r) = \partial_t p^T + a\partial_r p^T + \frac{1}{2}b^2\partial^2_{rr}p^T \tag{3.12}$$

$$\Sigma^T(t,r)p^T(t,r) = b\partial_r p^T. \tag{3.13}$$

Consider also a different maturity date $S < T$, with the corresponding SDE for $P_t(S) = p^S(t, r_t)$, 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 $X = H^S P(S) + H^T P(T)$ satisfies

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

If we set

$$H^S p^S \Sigma^S + H^T p^T \Sigma^T = 0 \tag{3.14}$$

our portfolio will be (locally) risk–free, and absence of arbitrage then implies that 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 (3.14) and some algebra results in the relation

$$\frac{M^T - r_t}{\Sigma^T} = \frac{M^S - r_t}{\Sigma^S}. \tag{3.15}$$

Since we have separated the S and T dependencies, this quantity is a function of t and r_t alone, and we conclude that there exists a process λ such that

$$\lambda_t = \lambda(t, r_t) = \frac{M^T - r_t}{\Sigma^T} \tag{3.16}$$

holds for all t and for every maturity time T. The quantity λ is the instantaneous return on a bond in excess of the spot rate, per unit of bond volatility. It is independent of bond maturity, and is called the *market price of interest rate risk*.

Substitution of the expressions (3.12) and (3.13) into the equation $M^T = r + \lambda\Sigma^T$ yields that arbitrage-free bond prices p^T satisfy the *term structure equation*

$$\partial_t p^T + [a(t,r) - \lambda(t,r)b(t,r)]\partial_r p^T + \frac{1}{2}b(t,r)^2\partial^2_{rr}p^T - rp^T = 0, \tag{3.17}$$

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

$$p^T = E_t^{Q(\lambda)}[e^{-\int_t^T r_s ds}], \tag{3.18}$$

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

$$dr_t = [a(t, r_t) - \lambda(t, r_t)b(t, r_t)]dt + b(t, r_t)dW_t^{Q(\lambda)}. \tag{3.19}$$

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

$$\frac{dQ(\lambda)}{dP} = \exp\left(-\int_0^T \lambda(t, r_t)dW_t - \frac{1}{2}\int_0^T \lambda(t, r_t)^2 dt\right). \tag{3.20}$$

It is easy to generalize both the term structure equation (3.17) and the expectation formula (3.18) to incorporate general T-derivatives with pay-offs of the form $\Phi(r_T)$ for a deterministic function Φ.

To summarize this section, we see that a complete specification of a one-factor spot rate model amounts to specifying both dynamics of the short rate r_t as an Itô diffusion under P *and* the market price of risk process $\lambda_t = \lambda(t, r_t)$. This is equivalent to selecting one equivalent martingale measure $Q(\lambda)$ of the form (3.20) and an Itô diffusion for the Q–dynamics of r. Either way, interest rate derivatives can then be priced by expectations of their final pay-off with respect to $Q(\lambda)$.

3.2.2 Affine Models

We say that a one factor short rate model is *affine* if the zero-coupon bond prices can be written as

$$P_t(T) = \exp[A(t,T) + B(t,T)r_t], \tag{3.21}$$

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 3.2.1. *Assume that the Q–dynamics for the short rate r_t is given by*

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

with a^Q and b of the form

$$\begin{aligned} a^Q(t,r) &= \kappa(t)r + \eta(t) && (3.22) \\ b(t,r) &= \sqrt{\gamma(t)r + \delta(t)} && (3.23) \end{aligned}$$

for some deterministic functions $\kappa, \eta, \gamma, \delta$. Then the model is affine and the functions A and B satisfy the Riccati equations

$$\begin{aligned} \frac{dB}{dt} &= -\kappa(t)B - \frac{1}{2}\gamma(t)B^2 + 1 && (3.24) \\ \frac{dA}{dt} &= -\eta(t)B - \frac{1}{2}\delta(t)B^2 && (3.25) \end{aligned}$$

for $0 \le t < T$, with boundary conditions $B(T,T) = A(T,T) = 0$.

Proof: (i) Calculate the partial derivatives of $P_t(T)$ in affine form and substitute into the term structure equation (3.17). (ii) Substitute the functional form for a^Q and b. (iii) Equate the coefficients of both 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.

We also note that under the same condition that the Q–dynamics for r_t has time-homogeneous coefficients, A and B are functions of the single variable $T-t$. The continuously compounded yield is then of the form

$$R_t(T) = -\frac{A(T-t)}{T-t} - \frac{B(T-t)}{T-t} r_t$$

which means that the possible yield curves generated by a time-homogeneous one factor affine model are simply the function $-\frac{A(T-t)}{T-t}$ plus a random multiple of the function $-\frac{B(T-t)}{T-t}$. It is also important for calibration purposes that the dependence of $R_t(T)$ on r_t is linear.

The affine property of bond prices implies only conditions on the dynamics of r under Q, but very often in affine modeling we suppose additionally that the market price of risk λ is such that the P dynamics is affine as well.

3.2.3 The Vasicek Model

The first one–factor model proposed in the literature was introduced in Vasicek (1977) who assumed that the P–dynamics for the short rate of interest is that of an Ornstein-Uhlenbeck process with constant coefficients, that is,

$$dr_t = \tilde{k}(\tilde{\theta} - r_t)dt + \sigma dW_t. \tag{3.26}$$

As we have seen above, the complete specification of an interest rate model also 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

$$\lambda(t) = \lambda(t, r_t) = ar_t + b, \tag{3.27}$$

for constants a, b. Then the Q–dynamics for the r_t is given by

$$dr_t = k(\theta - r_t)dt + \sigma dW_t^Q, \tag{3.28}$$

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

$$r_t = r_0 e^{-kt} + \theta\left(1 - e^{-kt}\right) + \sigma \int_0^t e^{-k(t-s)} dW_s^Q. \tag{3.29}$$

Therefore,

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

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 $\sigma^2/2k$. 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 3.2.2. *In the Vasicek model, bond prices are given by* $P_t(T) = \exp[A(t,T) + B(t,T)r_t]$ *where*

$$B(t,T) = \frac{1}{k}\left[e^{-k(T-t)} - 1\right] \tag{3.30}$$

$$A(t,T) = \frac{1}{k}\left(\frac{1}{2}\sigma^2 - k^2\theta\right)[B(t,T) + u(T-t)] - \frac{\sigma^2 B^2(t,T)}{4k} \tag{3.31}$$

Proof: (i) Obtain the Riccati 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 $B(t,T)$. □

Bond prices in the Vasicek model are thus very easy to compute: its main drawback is that it allows for negative interest rates. Explicit formulas for the prices of options on bonds are known for this model (see Jamshidian (1989)).

3.2.4 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

$$dr_t = \tilde{k} r_t dt + \sigma r_t dW_t, \tag{3.32}$$

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

$$dr_t = k r_t dt + \sigma r_t dW_t^Q, \tag{3.33}$$

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

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

so that

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

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. Finally, the model has the pathology, common to all lognormal models for the short rate of interest, that $E[C_t] = \infty$ whenever $k - \frac{1}{2}\sigma^2 > 0$. To understand this fact, note that for small Δt one can write $C_t \sim e^{(r_0 + r_{\Delta t})/(2\Delta t)}$ where $Y = r_{\Delta t}/(2\Delta t)$ is a log normal random variable. Therefore $E[C_t] \sim e^{r_0/(2\Delta t)} E[e^Y] = \infty$.

3.2.5 The Exponentiated Vasicek Model

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

$$dr_t = r_t(\tilde{\theta} - \tilde{k}\log r_t)dt + \sigma r_t dW_t, \tag{3.34}$$

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

$$dr_t = r_t(\theta - k\log r_t)dt + \sigma r_t dW_t^Q \tag{3.35}$$

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

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

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.

3.2.6 The Cox–Ingersoll-Ross Model

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

$$dr_t = \tilde{k}(\tilde{\theta} - r_t)dt + \sigma\sqrt{r_t}dW_t, \tag{3.36}$$

for positive constants $\tilde{k}, \tilde{\theta}$ and σ satisfying the condition $2\tilde{k}\tilde{\theta} \geq \sigma^2$. 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 $\lambda_t = a\sqrt{r_t} + b/\sqrt{r_t}$, so that its Q–dynamics is

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t^Q, \tag{3.37}$$

where $k = \tilde{k} + a\sigma$ and $k\theta = \tilde{k}\tilde{\theta} - b\sigma$. 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. Solving the associated Riccati equations for the functions $A(t,T)$ and $B(t,T)$ is lengthy, but straightforward, and leads to the following formula for bond prices.

Proposition 3.2.3. *In the CIR model, bond prices are given by*

$$P_t(T) = \exp[A(t,T) + B(t,T)r_t]$$

where

$$A(t,T) = \frac{2k\theta}{\sigma^2}\log\left[\frac{2\gamma\exp[(k+\gamma)(T-t)/2]}{2\gamma + (k+\gamma)(\exp[(T-t)\gamma] - 1)}\right] \tag{3.38}$$

$$B(t,T) = \frac{2(1 - \exp[(T-t)\gamma])}{2\gamma + (k+\gamma)(\exp[(T-t)\gamma] - 1)} \tag{3.39}$$

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

3.3 Fitting the initial term structure

3.3.1 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 $P_0(T)$ 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 $P_0^*(T)$ 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.

3.3.2 The Ho–Lee Model

In order to address the poor fit for the observed initial term structure $P_0^*(T)$ 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 the following short rate Q–dynamics:

$$dr_t = \Theta(t)dt + \sigma dW_t^Q \tag{3.40}$$

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 3.3.1. *In the Ho–Lee model, bond prices are given by*

$$P_t(T) = \frac{P_0^*(T)}{P_{0t}^*} \exp\left[(T-t)f^*(0,t) - \frac{\sigma^2}{2}t(T-t)^2 - (T-t)r_t\right] \tag{3.41}$$

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

Proof: (i) Obtain the Riccati 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.

3.3.3 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

$$dr_t = \kappa(\Theta(t) - r_t)dt + \sigma dW_t^Q. \tag{3.42}$$

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

Proposition 3.3.2. *In the Hull–White extended Vasicek model, bond prices are given by*

$$P_t(T) = \frac{P_0^*(T)}{P_0^*(t)} \exp\left[-B(t,T)f^*(0,t) - \frac{\sigma^2}{4\kappa}B^2(t,T)(1-e^{-2\kappa t}) + B(t,T)r_t\right]. \tag{3.43}$$

where

$$B(t,T) = \frac{1}{\kappa}(e^{-\kappa(T-t)} - 1) \tag{3.44}$$

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

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

$$\kappa\Theta(t) = \frac{\partial f^*(0,t)}{\partial T} + \kappa f^*(0,t) + \frac{\sigma^2}{2\kappa}(1-e^{-2\kappa t}). \tag{3.45}$$

(iii) Substitute back in the bond price formula.

You will be asked to integrate (3.42) 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

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

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

3.3.4 Deterministic–Shift Extensions

A similar extension was proposed by Hull and White (1990) for the CIR model, that is, by taking the coefficients in (3.36) to be time–dependent. This extension, however, does not lead to general analytic expressions for bond and options prices, since the associated Riccati 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

$$dx_t = \mu(x_t)dt + \sigma(x_t)dW_t^Q \tag{3.46}$$

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

$$r_t = x_t + \phi(t), \tag{3.47}$$

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 3.2.1 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

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

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

$$\phi(t) = f^*(0,t) + \frac{\partial \log F(0,t,x_0)}{\partial T}. \tag{3.48}$$

Inserting this back in the expression for bond prices leads to

$$P_t(T) = \frac{P_0^*(T)F(0,t,x_0)}{P_{0t}^*F(0,T,x_0)}F(t,T,r_t - \phi(t)). \tag{3.49}$$

Furthermore, whenever the original model has analytic expressions for the price of bond options, a similar argument leads to tractable expressions for these prices 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 extension discussed before.

The CIR ++ Model

Let the reference model be given by the Q–dynamics

$$dx_t = k(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t \tag{3.50}$$

with $4k\theta > \sigma^2$ and put $r_t = x_t + \phi(t)$. From proposition 3.2.3 we know that

$$F^{CIR}(0,T,x_0) = \left(\frac{2\gamma e^{(k+\gamma)T/2}}{2\gamma + (k+\gamma)(e^{T\gamma}-1)}\right)^{\frac{2k\theta}{\sigma^2}} \exp\left[\frac{2(1-e^{T\gamma})x_0}{2\gamma + (k+\gamma)(e^{T\gamma}-1)}\right], \tag{3.51}$$

where $\gamma = \sqrt{k + 2\sigma^2}$. According to (3.48) leads to

$$\begin{aligned} \phi^{CIR}(t) &= f^*(0,t) - \frac{2k\theta(e^{t\gamma}-1)}{2\gamma + (k+\gamma)(e^{t\gamma}-1)} \\ &+ x_0\frac{4\gamma^2 e^{t\gamma}}{[2\gamma + (k+\gamma)(e^{t\gamma}-1)]^2} \end{aligned} \tag{3.52}$$

and we can use (3.49) 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 (3.52) 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) [10] 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

$$df_t(T) = \alpha(t,T)dt + \sigma(t,T)dW_t, \tag{3.53}$$

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 $P_t(T)$. 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 3.4.1. *If the bond market is free from arbitrage, then there exists a process λ_t with the property that*

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds - \sigma(t,T)\lambda_t. \tag{3.54}$$

Proof: (i) Use the third part of proposition 3.1.1 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 (3.16) for the market price of risk.

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

$$df_t(T) = \alpha^Q(t,T)dt + \sigma^Q(t,T)dW_t^Q, \tag{3.55}$$

we see that the HJM drift condition reduces to

$$\alpha^Q(t,T) = \sigma^Q(t,T)\int_t^T \sigma^Q(t,s)ds. \tag{3.56}$$

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

3.5 General Option Pricing Formulas

Consider the market of section A.8 with $d = 1$ and a stochastic interest rate, that is

$$\begin{aligned} dS_t &= \mu(t,S_t)S_t + \sigma(t,S_t)S_t dW_t^1 \\ dr_t &= a(t,r_t)dt + b(t,r_t)[\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2] \\ dC_t &= r_t C_t dt, \end{aligned} \tag{3.57}$$

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 traded asset), there exists a measure Q^N such that the prices of any other traded assets in units of the numeraire N_t are martingales with respect to Q^N. In particular, the price of a derivative B is given by

$$\pi(t) = N_t E_t^{Q^N}\left[\frac{B}{N_T}\right] \tag{3.58}$$

In this context, let Q^S and Q^T denote the measures obtained using (A.23) for the numeraires S_t and $P_t(T)$, 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 (3.58) to each term in the pay-off above separately, obtaining the initial price of the call option

$$c_0 = S_0 Q^S(S_T \geq K) - KP_0(T)Q^T(S_T \geq K). \tag{3.59}$$

In order to calculate the expectations above, let us define the process $Z_t = S_t/P_t(T)$. Since this is the price of a traded asset in terms of the numeraire $P_t(T)$, 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

$$Z_T = \frac{S_0}{P_0(T)} \exp\left(-\frac{1}{2}\int_0^T |\sigma_{S,T}(t)|^2 dt + \int_0^T \sigma_{S,T}(t) dW^T\right).$$

Similarly, define $Y_t = Z_t^{-1} = P_t(T)/S_t$, 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

$$Y_T = \frac{P_0(T)}{S_0} \exp\left(-\frac{1}{2}\int_0^T |\sigma_{S,T}(t)|^2 dt - \int_0^T \sigma_{S,T}(t) dW^S\right).$$

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 (3.59), we obtain the following generalization of the Black–Scholes option price formula

$$\begin{aligned} c_0 &= S_0 Q^S(Y_T \leq 1/K) - K P_0(T) Q^T(Z_T \geq K) \\ &= S_0 N[d_1] - K P_0(T) N[d_2], \end{aligned} \tag{3.60}$$

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

$$d_1 = d_2 + \sqrt{\mathrm{Var}(T)} \tag{3.61}$$

$$d_2 = \frac{\log\left(\frac{S_0}{K P_0(T)}\right) - \frac{1}{2}\mathrm{Var}(T)}{\sqrt{\mathrm{Var}(T)}}. \tag{3.62}$$

Consider now a call option with strike price K and exercise date T_1 on an underlying bond $P_t(T_2)$ with $T_1 < T_2$. We can use the formulation above to price this contract by replacing S_t by $P_t(T_2)$. All we need to check is if the volatility parameter of the process $Z_t = P_t(T_2)/P_t(T_1)$ is deterministic. For this purpose let us assume that the interest rate follows the Hull-White extended Vasicek model (3.42) with the function $\Theta(t)$ given by (3.45). Then, this being an affine model, bond prices of any maturity are given by

$$P_t(T) = \exp[A(t,T) + B(t,T) r_t],$$

where, in particular, $B(t,T)$ is given by (3.44). 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

$$\sigma^Z(t) = [B(t,T_2) - B(t,T_1)]\sigma = \frac{\sigma}{\kappa} e^{\kappa t}\left[e^{-\kappa T_1} - e^{-\kappa T_2}\right], \tag{3.63}$$

which is indeed deterministic.

Applying (3.60) for this option we obtain

$$c_0^{HW}(t,T_1,K,T_2) = P_0(T_2)N[d_1] - KP_0(T_1)N[d_2], \tag{3.64}$$

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

$$d_1 = d_2 + \sqrt{\Sigma^2} \tag{3.65}$$

$$d_2 = \frac{\log\left(\frac{P_0(T_2)}{KP_0(T_1)}\right) - \frac{1}{2}\Sigma^2}{\sqrt{\Sigma^2}}. \tag{3.66}$$

$$\Sigma^2 = \frac{\sigma^2}{2\kappa^3}\left[1 - e^{-2\kappa T_1}\right]\left[1 - e^{-\kappa(T_2 - T_1)}\right]^2. \tag{3.67}$$

3.6 Exercises

Exercise 9. Complete the proof of Proposition 3.2.1.

Exercise 10. Prove (3.29). (Hint: find an integrating factor $C(t)$ such that $d(C(t)r_t) = C(t)\sigma dW_t^Q$.)

Exercise 11. Prove the formulas (3.30) that give the price of zero coupon bonds in the Vasicek model.

Exercise 12.
1. Use (3.21) to write down a formula for $f_t(T)$ as a function of r_t, t, T that involves the functions A, B.
2. Use this formula, and the Riccati DEs (3.25), (3.24) to compute the limit $\lim_{T\downarrow t} f_t(T)$.
3. Show that in the Vasicek model, $f_t(T)$ is Gaussian. Compute the mean and variance of $f_t(T)$.

Exercise 13. In the Vasicek model, find the distribution of the spot rate r_T under the T-forward measure. Using the T-forward measure, find a formula for the price of a maturity T, strike K European call option on the spot rate r_T. Combine this formula with (3.30) to derive the formula for the bond call option that pays $(P_T(S) - K)^+$, where $T < S$.

Exercise 14. The spot rate r_t is an OU process in the Vasicek model, and thus tends to a steady state distribution for large times.

1. Find the distribution of $r_\infty := \lim_{T\to\infty} r_T$.
2. Find also the steady state bivariate distribution $(r_\infty, r_\infty^{-h}) = \lim_{T\to\infty}(r_T, r_{T-h})$ for any value of the lag variable $h \geq 0$. (For the second part you should review the properties of multivariate normal distributions in the Appendix).

Exercise 15. Consider the deterministic-shift extension to the Vasicek model. Show that this is equivalent to the Hull-White extension, and find the relation between $\phi(t)$ and $\Theta(t)$.

Exercise 16. Let X_t, Y_t be two independent CIR processes, with parameters (k, θ_x, σ) and $k, \theta_y, \sigma)$ respectively. Show that $Z_t = X_t + Y_t$ is also CIR, with parameters $(k, \theta_x + \theta_y, \sigma)$. (Hint: you will need to use the Lévy theorem on Brownian motion)

Exercise 17. Consider the affine short rate model as defined in Proposition 3.2.1. Show that the function

$$G^T(t, r; c_1, c_2) = E_t^Q[e^{-c_1 \int_t^T r_s ds - c_2 r_T}]$$

is given by $G^T(t, r) = \exp[A(t, T) + B(t, T)r]$ where A, B now satisfy the Riccati equations

$$\frac{dB}{dt} = -\kappa(t)B - \frac{1}{2}\gamma(t)B^2 + c_1 \tag{3.68}$$

$$\frac{dA}{dt} = -\eta(t)B - \frac{1}{2}\delta(t)B^2 \tag{3.69}$$

with boundary conditions $A(T, T) = 0, B(T, T) = -c_2$.

Exercise 18. Solve the Riccati equations (3.68) from the previous exercise for the Vasicek model to show the "master formula"

$$F^{VAS}(T, r; c_1, c_2) := E_r[e^{-c_1 \int_0^T r_s ds - c_2 r_T}] = \exp[A(T) + B(T)r] \tag{3.70}$$

$$B(T) = \frac{c_1}{k}(e^{-kT} - 1) - c_2 e^{-kT}$$

$$A(T) = \frac{1}{k^2}\left(\frac{1}{2}\sigma^2 - k^2\theta\right)[B(T) + c_1 T] - \frac{\sigma^2 B^2(T)}{4k}$$

Exercise 19. Consider the d-dimensional OU process X_t that solves the SDE

$$dX_t = -\alpha X_t dt + \sigma dW_t \tag{3.71}$$

for scalars α, σ (W is d-dimensional too). Show that $r_t = |X_t|^2$ solves

$$dr_t = 2\alpha\left(\frac{\sigma^2 d}{2\alpha} - r_t\right)dt + 2\sigma\sqrt{r_t}d\tilde{W}_t$$

for another one-dimensional Brownian motion $\tilde{W}$, and hence is a CIR process. (Hint: you will need to use the Lévy Theorem on characterizing Brownian motion.)

Exercise 20. From a textbook on probability (or a reliable on-line source), review the definition of the gamma distribution. Use the previous exercise for integer d to write the probability density of the random variable r_t in the CIR model, as a gamma distribution whose parameters are determined by the parameters $\alpha, \theta = \sigma^2 d/2\alpha, \sigma, r_0$. This distribution is also called a non-central chi-squared distribution. Extend the formula to d non-integer to obtain the general result (one can argue that the distribution must depend smoothly on d).

Exercise 21.

1. Prove Proposition 3.2.3. Give a formula for the continuously compounded yield $R_0(T)$ in both Vasicek and CIR models. Using MATLAB, or similar, plot this function (called the yield curve) over the period $T \in [0, 30]$ when the CIR parameters are chosen to be $\kappa = 1.5, \theta = 0.03, \sigma = 2, r_0 = 0.03$, and Vasicek parameters chosen to give a "reasonable" fit (for example, match moments). Comment on the differences between the curves.

2. Write a MATLAB script to compute the par coupon rate for a bond of maturity $T = N\Delta t$. Use this to plot the rate for $\Delta t = 1/2, N = 1, 2, \ldots, 60$

Chapter 4

Structural Models of 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*. Operationally, for medium to large cap firms, default is normally triggered by a failure of the firm to meet its debt servicing obligations, which usually quickly leads to bankruptcy proceedings, such as Chapter 11 in the U.S. Thus default is considered a rare and singular event after which the firm ceases to operate as a viable concern, and which results in *large* financial losses to *some* security holders. With some flexible thinking, this view of credit risk also extends to sovereign bonds issued by countries with a non-negligible risk of default, such as those of developing countries.

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. These models 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.

4.1 The Merton Model (1974)

The Merton model takes an overly simple debt structure, and assumes that the total value A_t of a firm's assets follows a geometric Brownian motion under the physical measure

$$dA_t = \mu A_t dt + \sigma A_t dW_t, \quad A_0 > 0, \tag{4.1}$$

where μ is the mean rate of return on the assets and σ is the asset volatility. We also need further assumptions: there are no bankruptcy charges, meaning the liquidation value equals the firm value; the debt and equity are frictionless tradeable assets.

Large and medium cap firms are funded by shares ("equity") and bonds ("debt"). The Merton model assumes that debt consists of 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 remainder is distributed among shareholders. However, if

$A_T < K$ then default is deemed to occur: the bondholders exercise a debt covenant giving them the right to liquidate the firm and receive the liquidation value (equal to the total firm value since there are no bankruptcy costs) in lieu of the debt. Shareholders receive nothing in this case, but by the principle of limited liability are not required to inject any additional funds to pay for the debt.

From these simple observations, we see that shareholders have a cash flow at T equal to

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

and so equity can be viewed as a European call option on the firm's assets. On the other hand, the bondholder receives $\min(A_T, K)$. Moreover, the physical probability of default at time T, measured at time t, is

$$P_t[\tau = T] = P_t[A_T \leq K] = N[-d_2^P]$$

where $d_2^P = \frac{\log(A_t/K)+(\mu-\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$.

The value E_t at earlier times $t < T$ can be derived using the classic martingale argument (see exercise 24 for an alternative derivation). Assuming one can trade the firm value A_t, we note that $e^{-rt}A_t$ is a martingale under the risk-neutral measure Q with market price of risk $\phi = (\mu - r)/\sigma$ and Radon-Nikodym derivative

$$\frac{dQ}{dP} = \exp\left(\phi W_T - \phi^2 T\right). \tag{4.2}$$

Then we find the standard Black-Scholes call option formula

$$E_t = E^Q[e^{-r(T-t)}(A_T - K)^+] = \mathrm{BSCall}(A_t, K, r, \sigma, T-t) \tag{4.3}$$

$$= A_t N[d_1] - e^{-r(T-t)} K N[d_2] \tag{4.4}$$

where

$$d_1 = \frac{\log(A_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = \frac{\log(A_t/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \tag{4.5}$$

Bond holders, on the other hand, receive

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

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 a European put option. Of course the fundamental identity of accounting holds:

$$A_t = E_t + D_t,$$

and all three assets are discounted risk neutral martingales. A zero coupon defaultable bond with face value 1 and maturity T will have the price $\bar{P}_t(T) = D_t/K$, and has the yield spread

$$YS_t(T) = \frac{1}{T-t}\log\frac{Ke^{-r(T-t)}}{D_t} = -\frac{1}{T-t}\log\left(e^{r(T-t)}\frac{A_t}{K}(1 - N[d_1]) + N[d_2].\right) \tag{4.6}$$

Despite being derived from a debt structured with a single maturity T, this equation is often interpreted as giving a function of T: it is imagined that if an additional bond of small face value with a different maturity were issued by the firm, it would also be priced according to (4.6). The qualitative behaviour of this term structure is that credit spreads start at zero for $T = 0$, increase sharply to a maximum, and then decrease either to zero at large times if $r - \sigma^2/2 \leq 0$ or a positive value if $r - \sigma^2/2 > 0$. This is in accordance with the diffusive character of the model. For very short maturity times, the asset price diffusion will almost surely never cross the default barrier. The probability density of default then increases for longer maturities but starts 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 can easily avoid such discrepancies.

The previously obtained formula for the physical default probability (that is under the measure P) can be used to calculate risk neutral default probability provided we replace μ by r. Thus one finds that

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

and as long as $\phi > 0$ we see that market implied (i.e. risk neutral) survival probabilities are always less than historical ones.

In the event of default, the bondholder receives only a fraction A_T/K, called the *recovery fraction*, of the bond principal K: the fractional loss $(K - A_T)/K$ is called the *loss given default* or LGD. As you will see in an exercise, the probability distribution of LGD can be computed explicitly in the Merton model.

Note that equity value increases with the firm's volatility (since its payoff is convex in the underlier), so shareholders are generally inclined to press for riskier positions to be taken by their managers. The opposite is true for bondholders. So-called "agency problems" relate to the contradictory aims of shareholders, bondholders and other "stakeholders".

Structural models like Merton's model depend on the unobserved variable A_t. On the other hand, for publicly traded companies, the share price (and hence the total equity) is closely observed in the market. The usual approach to obtaining an estimate for the firm's asset values A_t and volatility σ in Merton's model uses the Black-Scholes formula for a call option, that is,

$$E_t = \text{BSCall}(A_t, K, r, \sigma, T - t), \tag{4.7}$$

where K and T are determined by the firm's debt structure. We obtain a second equation by equating the equity volatility to the coefficient of the Brownian term obtained by applying Itô's formula to (4.7), namely,

$$\sigma A_t \frac{\partial \text{BSCall}}{\partial A} = \sigma^E E_t. \tag{4.8}$$

The Merton model is only a starting point for studying credit risk, and is obviously far from realistic:

- The non-stationary structure of the debt that leads to the termination of operations on a fixed date, and default can only happen on that date. Geske [8] extended the Merton model to the case of bonds of different maturities.
- It is incorrect to assume that the firm value is tradeable. In fact, the firm value and its parameters is not even directly observed.
- Interest rates should certainly be taken to be stochastic: this is not a serious drawback, and its generalization was included in Merton's original paper.
- The short end of the yield spread curve in calibrated versions of the Merton model typically remains essentially zero for months, in strong contradiction with observations.

The so-called first passage models extend the Merton framework by allowing default to happen at intermediate times.

4.2 Black-Cox model

The simplest *first passage model* again takes a firm with asset value given by (4.1) and outstanding debt with face value K at maturity T. However, instead of admitting only the possibility of default at maturity time T, Black and Cox (1976) [3] postulated that default occurs at the first time that the firm's asset value drops below a certain time-dependent barrier $K(t)$. This can be explained by the right of bondholders to exercise a "safety covenant" that allows them to liquidate the firm if at any time its value drops below the specified threshold $K(t)$. Thus, the default time is given by

$$\tau = \inf\{t > 0 : A_t < K(t)\} \tag{4.9}$$

For the choice of the time dependent barrier, observe that if $K(t) > K$ then bondholders are always completely covered, which is certainly unrealistic. On the other hand, one should clearly have $K_T \leq K$ for a consistent definition of default. One natural, but certainly not the only, choice is to take an increasing time-dependent barrier $K(t) = K_0 e^{kt}, K_0 \leq K e^{-kT}$.

The first passage time to the default barrier can now be reduced to the first passage time for Brownian motion with drift. Observing that

$$\{A_t < K(t)\} = \{W_t + \sigma^{-1}(r - \sigma^2/2 - k)t < \sigma^{-1}\log(K_0/A_0)\},$$

we obtain that the risk neutral probability of default occurring before time $t \leq T$ is then given by

$$Q[0 \leq \tau < t] = Q\left[\min_{s \leq t}(A_s/K(s)) \leq 1\}\right] = Q\left[\min_{s \leq t} X_s \leq \sigma^{-1}\log\left(\frac{K_0}{A_0}\right)\right] \tag{4.10}$$

where $X_t = W_t + mt, m = \sigma^{-1}(r - \sigma^2/2 - k)$. This is a classic problem of probability we discuss in Appendix A, whose solution is given by

$$\begin{aligned} Q[\min_{s \leq t} X_t \leq d] &= 1 - \mathrm{FP}(-d; -m, t) \\ \mathrm{FP}(d; m, t) &:= N\left[\frac{d - mt}{\sqrt{t}}\right] - e^{2md} N\left[\frac{-d - mt}{\sqrt{t}}\right], \quad d \geq 0 \end{aligned} \tag{4.11}$$

Thus we obtain the formula

$$Q[0 \leq \tau < t] = 1 - \mathrm{FP}(-d; -m, t) \tag{4.12}$$

with $m = \sigma^{-1}(r - \sigma^2/2 - k)$ and $d = \sigma^{-1}\log(K_0/A_0) < 0$.

The pay-off for equity holders at maturity is

$$(A_T - K)^+ \mathbf{1}_{\{\min_{s \leq T} X_s > d\}} = (e^{kT} A_0 e^{\sigma X_T} - K)^+ \mathbf{1}_{\{\min_{s \leq T} X_s > d\}}. \tag{4.13}$$

This is equivalent to the payoff of a down-and-out call option, and can be priced by closed form expressions found for example in (Merton 74) [19]. The equity in the Black-Cox model is smaller than the share value obtained in the Merton model, and is not monotone in the volatility. In the event of default, the pay-off for debt holders is $A_\tau = K(\tau)$ at the time of default, and the fair "recovery value" can be computed by integrating $K(s)$, discounted, with respect to the risk-neutral PDF for the time of default. The value of the bond at time t prior to default is therefore

$$D_t = \int_t^T e^{r(t-s)} K(s)(-\partial_s \mathrm{FP}(-d_t; -m, s - t))ds + e^{-r(T-t)} K\mathrm{FP}(-d_t; -m, T - t) \tag{4.14}$$

where $d_t = \sigma^{-1}\log(K(t)/A_t)$. Computation of the integral can be done explicitly, as you will be asked to do in an exercise.

4.3 KMV

KMV[1] is the name given to a particularly successful practical implementation of structural credit modeling. It is instructive to see what assumptions they make in order to produce commercially acceptable credit methods. The main difficulty, as in all structural models, is in assigning dynamics to the firm value, which is an unobserved process. We outline the main points, including the key point where KMV diverges from a strict structural model.

[1] The firm KMV is named after Kealhofer, McQuown and Vasicek, the founders of the company in 2002. It has since been sold to Moody's.

- Default trigger: The question is to determine a level for K from the structure of a firm's debt (which in practise consists of bond issues of different maturity, coupon rate, seniority and features such as convertibility). In a nutshell, KMV puts the value somewhere between the face value of short term debt, and the face value of the total debt, arguing that the firm will always have to service short term debt, but can be more flexible in servicing the long term debt. Typically, the trigger is given by the full short term debt plus half the long term debt.

- Value of firm: Rather than try to estimate the firm value directly from detailed balance sheet data (a procedure highly sensitive to assumptions and method) KMV infers A_t from the value of debt (which is taken from balance sheet) and equity.

- Equity: KMV takes the safest course, defining it to be the market capitalization (current share value times the number of shares outstanding). Since E_t is observed in the market, it is used to infer the value of A_t using the Black-Scholes call option formula:
$$E_t = BSCall(A_t, T-t, r, \sigma, K) \tag{4.15}$$
The maturity date T is not so clear, but should represent the approximate time scale of the debt[2]. A further relation between the key unknowns, A_t and σ and the key observables E_t, σ_E is obtained by taking the stochastic term from the Itô formula applied to (4.15), giving
$$\sigma_E E_t = \sigma A_t \frac{\partial BSCall}{\partial A} \tag{4.16}$$
From a time series of equity data and the associated estimate for σ_E, the two equations (4.15), (4.16) then lead to the time series for A and an estimate for σ.

From the data on a given date, KMV computes the key credit score DD_t for the firm at that time, called *distance to default*. Roughly speaking it gives the amount by which $\log A_t$ exceeds $\log K$ measured in standard deviations σ of the one year PDF $\log A_{t+1}$. That is
$$\mathrm{DD}_t = \frac{\log(A_t/K)}{\sigma} \tag{4.17}$$
By a strict structural interpretation, *EDF, the expected default frequency*, meaning the probability of observing the firm to default within one year, ought to equal the normal probability $\mathrm{EDF}_t = N(\mathrm{DD}_t)$. KMV however, breaks the model at this point, and instead relies on its large database of historical defaults to map DD to EDF by a proprietary function $\mathrm{EDF} = f(\mathrm{DD})$. $f(\mathrm{DD})$ gives the actual fraction of all firms with the given DD that have been observed to default within one year.

Studies such as Duffie et al [4] indicate that the distance to default DD_t is a reasonable firm-specific dynamic (defined by current observations of the firm) quantity that correlates strongly with credit spreads and observed historical default frequency.

[2]For example, one could take it to be the *duration* of the debt.

4.4 Default Events and Bond Prices

We have learned from structural models, particularly the Black-Cox model, that two factors are critical in determining the value of defaultable bonds, namely the probability of default (PD) and the loss given default (LGD). In this section we explore some general properties of the probability of default, and see what these imply for defaultable bonds.

Recall that the *default time* τ is a *stopping time*, that is, a random variable $\tau : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ such that $\{\tau \le t\} \in \mathcal{F}_t$, for every $t \ge 0$. In other words, a random time τ is a stopping time if the cádlág[3] stochastic process

$$H_t(\omega) = \mathbf{1}_{\{\tau \le t\}}(\omega) = \begin{cases} 1, & \text{if } \tau(\omega) \le t \\ 0, & \text{otherwise} \end{cases} \tag{4.18}$$

is adapted to the filtration $\mathcal{F}_t$. For default times, H_t is known as the *default indicator process*, and $H_t^c = 1 - H_t$ is the *survival indicator process.*

We say that a stopping time $\tau > 0$ is *predictable* if there is an *announcing sequence* of stopping times $\tau_1 \le \tau_2 \le \cdots$ 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 H_t is predictable. For the record, 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 $\hat{\tau}$. It can be shown that every stopping time can be expressed as the minimum of a predictable and a totally inaccessible stopping time.

Example 2. For any adapted process X_t, we may define stopping times $\tau_d = \inf\{t | X_t \ge d\}$. If X is an Itô diffusion (and therefore has continuous paths a.s.), this is a predictable stopping time: one takes any sequence $\{d_n\}_{n=1}^{\infty}$ that increases to d, and then $\{\tau_{d_n}\}$ is an announcing sequence. Conversely, if X_t is a pure jump process (for example a Poisson process), then τ_d is totally inaccessible.

4.4.1 Unconditional default probability

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

$$P[\tau > t] = 1 - P[\tau \le t] = 1 - E[\mathbf{1}_{\{\tau \le t\}}]. \tag{4.19}$$

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

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

[3] for "continueux á droite, limite á gauche", or "right continuous, left limit".

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 $s \le t$ years is

$$P[\tau > t|\tau > s] = \frac{P[\{\tau > t\} \cap \{\tau > s\}]}{P[\tau > s]} = \frac{P[\tau > t]}{P[\tau > s]}, \tag{4.20}$$

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

$$P[s \le \tau \le t|\tau > s] = 1 - P[\tau > t|\tau > s] = 1 - \frac{P[\tau > t]}{P[\tau > s]}. \tag{4.21}$$

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

$$h(t) = -\frac{\partial \log P[\tau > t]}{\partial t}. \tag{4.22}$$

It then follows that

$$P[\tau > t|\tau > s] = e^{-\int_s^t h(u)du}. \tag{4.23}$$

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 for a short time interval $[t, t+\Delta t]$,

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

4.4.2 Conditional default probability

The above probabilities are all derived from $P[\tau > t]$, that is, conditionally on the constant information set available at time 0. More generally, one can focus on $\mathrm{SP}_s(t) := P[\tau > t|\mathcal{F}_s]$, that is, the *survival probability* in t years conditioned on *all* the information available at time $s \le t$. If we assume positivity and differentiability in t, then this can be written as

$$\mathrm{SP}_s(t) = H_s^c e^{-\int_s^t h_s(u)du}, \tag{4.24}$$

where

$$h_s(t) = -\frac{\partial \log \mathrm{SP}_s(t)}{\partial t}. \tag{4.25}$$

We define $h_s(t)$ to be the *forward default rate process* given all the information up to time s: it clearly has the initial values $h_0(t) = h(t)$.

The indicator process H_t defined in (4.18) is a submartingale. Moreover, since H can be shown to be of class $\mathcal{D}$, it follows from the Doob-Meyer decomposition[4] that there exists a unique nondecreasing predictable process Λ_t, called the *compensator*, such that $H_t - \Lambda_t$ is a martingale. Since defaults happen only once, in general we know that

$$\Lambda_t = \Lambda_{\tau \wedge t}.$$

[4] Please see [21] for a fundamental discussion of such matters.

In some cases the compensator can be written as

$$\Lambda_t = \int_0^t \lambda_s ds \tag{4.26}$$

for a non-negative, progressively measurable process λ_t. In this case the process λ_t is called the *default intensity*. In other cases, such as Black-Cox models, the compensator does not admit a default intensity (intuitively, the instantaneous probability of default is either 0 or ∞).

Under suitable technical conditions it can be shown that

$$\lambda_t = h_t(t). \tag{4.27}$$

Therefore, while the forward default rate function $h(t)$ gives the instantaneous rate of default conditioned only on survival up to t, the default intensity λ_t 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 (4.25) and the associated intensity by (4.27). 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 straightforward, and the same is true for going from intensities to survival probabilities.

4.4.3 Implied Survival Probabilities and Credit Spreads

We now investigate how the prices of defaultable zero-coupon bonds can be used to infer *risk-neutral* default probabilities. As we will see later, PD and LGD become entangled in the bond pricing formula, and for that reason we will assume in this section that bonds pay nothing in the event of default ("zero-recovery"). We also assume in this section that interest rates r_t and the default time τ are *independent* under the risk-neutral measure Q.

Let $\bar{P}_t(T)\mathbf{1}_{\{\tau>t\}}$ be the price at time $t \leq T$ of a defaultable zero-coupon bond issued by a certain firm with maturity T and face value equal to one unit of currency. Then, since we assume bonds pay zero recovery, we then know that

$$\bar{P}_t(T)\mathbf{1}_{\{\tau>t\}} = E_t^Q\left[e^{-\int_t^T r_s ds}\mathbf{1}_{\{\tau>T\}}\right] \tag{4.28}$$

and since r_t and τ are *independent* this becomes

$$\bar{P}_t(T)\mathbf{1}_{\{\tau>t\}} = P_t(T)Q[\tau > T|\mathcal{F}_t].$$

Therefore, as long as $\tau > t$ the risk-neutral survival probability is given by

$$Q[\tau > T|\mathcal{F}_t] = \frac{\bar{P}_t(T)}{P_t(T)} = \exp[-\int_t^T (\bar{f}_t(s) - f_t(s))ds] \tag{4.29}$$

Comparing with the definition of the yield spread $YS_t(T)$ and the forward default rate $h_t(T)$, we see that

$$YS_t(T) = \frac{1}{T-t}\int_t^T h_t(s)ds, \quad h_t(s) = \bar{f}_t(s) - f_t(s).$$

Thus 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 what follows, even though the interpretation relies on the above two assumptions, (4.29) will be called the *implied survival probability*, emphasizing the fact that it is derived from market prices and associated to the risk-neutral measure Q.

4.5 Exercises

In the following exercises, we take the asset value process with parameters $\mu = r = 0.05$ and $\sigma = 0.20$.

Exercise 22. The *leverage ratio* or *debt/equity ratio* of a firm is defined to be the total debt of the firm divided by the equity $L_t = D_t/E_t$. A highly leveraged firm has a lot of debt which might indicate a danger of default, and certainly a susceptibility to rising interest rates. Consider the Merton model and use MATLAB or similar to plot the value of the debt D_0 as a function of debt maturity T for the following leverage levels: $10, 3, 1, 1/3, 1/10$. Thinking of $D_0(T)$ as the price of a zero coupon defaultable bond, compute the *credit spread* (defaultable bond yield minus default free bond yield) as a function over T, again for the same leverage levels.

Exercise 23. Find a closed formula for the price of the debt in the Black-Cox model. With $k = 0$, use MATLAB or similar to plot the value of the debt D_0 as a function of debt maturity T for the following leverage levels: 100, 30, 10, 3, 1. Discuss the differences between the two models.

Exercise 24. (Alternative derivation of the Merton model) Suppose $dA_t = A_t[\mu dt + \sigma dW_t]$ under the physical measure, plus the other assumptions of the Merton model. Suppose further that debt and equity are tradeable assets that satisfy $A_t = D_t + E_t$ and follow processes $D_t = D(t, A_t), E_t = E(t, A_t)$ for differentiable functions. By considering a locally risk-free self-financing portfolio of bonds and equity (which by necessity will earn the risk-free rate of return), prove directly that both D, E satisfy the Black-Scholes equation

$$\partial_t f + \frac{1}{2}\sigma^2 A^2 \partial_A^2 f + rA\partial_A f - rf = 0$$

Exercise 25. In the Merton model with the number of shares N constant, the stock price is $S_t = E_t/N$. Find a formula for the stock volatility as a function of time to maturity and current equity, i.e. $\sigma_t^{(S)} = f(T-t, S_t)$. Find the stochastic differential equation for S_t under the risk-neutral measure. Is S_t Markovian? Compute the asymptotics of stock volatility when the leverage ratio tends to ∞. What does this say about the behaviour of the stock price?

Exercise 26. For any $s < T$, compute the probability density function $p_\tau(t|\tau > s), t > s$ for the time to default τ in the Black-Cox model. What is the behaviour of p_τ as $t \to s+$? Does this model have a default intensity?

Exercise 27. *Loss given default* LGD is defined to be the fraction of the bond principal that is not returned to the bondholder at default. Compute the probability distribution of LGD in the Merton model. Plot the pdf with $T = 5$ for the leverage levels 100, 30, 10, 3, 1, and compute the mean and standard deviation in each case.

Chapter 5

Reduced Form Modelling

Reduced form models, also known as hazard rate models or intensity-based models form an approach to default complementary to the structural models. In structural models, default was directly linked to the value of the firm, and in the simplest versions, default times are predictable in the filtration available to traders. In contrast, reduced form models make the assumption that default is always a surprise, that is, a totally inaccessible stopping time. The firm value is not modeled, but rather attention is focussed on the instantaneous conditional probability of default. From the long list of works on reduced form models, we mention some of the pioneering papers: [15, 14, 16, 5].

5.1 Information Sets

Let us consider for the moment the physical probability setting: later we will transfer all results to equivalent statements under the risk-neutral measure. The stochastic process $H_t = \mathbf{1}_{\{\tau \le t\}}$ is a submartingale in its *natural filtration* $(\mathcal{H}_t)_{t\ge 0}$, and assuming suitable differentiability, we have seen there is a positive function $h(t)$ such that

$$H_t - \int_0^t h(s)(1-H_s)ds$$

is an $\mathcal{H}$ martingale. When there is a "market filtration" $(\mathcal{G}_t)_{t\ge 0}$ (which roughly speaking includes all information other than the fact of default or survival), one may have a $\mathcal{G}_t$-adapted process λ_t such that

$$H_t - \int_0^t \lambda_s(1-H_s)ds$$

is a martingale under the "full filtration"

$$\mathcal{F}_t := \mathcal{H}_t \vee \mathcal{G}_t.$$

Note how the possibility of further defaults is "switched off" at the time of default. The martingale condition implies that

$$P[t < \tau \le t + dt | \mathcal{F}_t] = \lambda_t(1-H_t)dt + o(dt)$$

for small dt, and captures the idea in reduced form modeling that defaults are unpredictable "totally inaccessible" stopping times. This of course contradicts a key feature of a basic structural model like Black-Cox: in that model $\mathcal{F}_t = \mathcal{G}_t$, and the compensator of H_t equals H_t itself, and does not admit an "intensity" λ.

5.2 Basic examples of reduced form models

Reduced form modelling makes a key simplifying assumption on the relation between the timing of defaults (encoded in the filtration $\mathcal{H}$) and the market filtration $\mathcal{G}$. Essentially, one assumes that the fact of a default or not at a given time t has no impact on the evolution of the market filtration beyond t. Before we formalize this notion in section 5.3, we will try to develop an intuitive idea by first investigating some important examples of basic reduced form models, each of increasing generality.

To begin, note that because we are assuming a default event does not influence the default rate (other than "switching it off"), one can imagine the process H continuing after default according to the same rate λ_t. This leads to the idea that $H_t = N_{t\wedge\tau}$ where N_t is a *counting process*, that is a non-decreasing, integer-valued process with $N_0 = 0$. Then the default time is defined to be

$$\tau = \inf\{t | N_t > 0\}. \tag{5.1}$$

We now focus on cases when the counting process N_t admits an intensity λ_t, that is

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

and $N_t - \Lambda_t^N$ is a martingale.

5.2.1 Poisson processes

The *Poisson process* N_t with parameter $\lambda > 0$, 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 $0 \le s < t$ we have

$$P[N_t - N_s = k] = \frac{(t-s)^k \lambda^k}{k!} e^{-(t-s)\lambda} \tag{5.2}$$

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 of its history up to t. We also see from (5.2) 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, one can easily show that the waiting time between two jumps is an exponentially distributed random variable with parameter λ. In particular, if we use (5.1) to define default as the arrival time of the first jump of N_t, then the expected default time is $1/\lambda$ and the probability of survival after t years is

$$P[\tau > t] = E[N_t = 0] = e^{-\lambda t}. \tag{5.3}$$

Therefore, from the definition of the forward default rate function (4.22) we obtain

$$h(t) = -\frac{\partial \log P[\tau > t]}{\partial t} = \lambda \tag{5.4}$$

so that

$$P[\tau > t | \tau > s] = e^{-\int_s^t h(u)du} = e^{-\lambda(t-s)}. \tag{5.5}$$

The Poisson model is thus seen to be equivalent to the Exponential default model of §2.4.

5.2.2 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 practice, the hazard rate changes in time, since survival up to different time horizons lead to different probabilities of default over the next small time interval. In order to obtain more realistic term structures of default probabilities, we can 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

$$P[N_t - N_s = k] = \frac{1}{k!}\left(\int_s^t \lambda(u)du\right)^k \exp\left(-\int_s^t \lambda(u)du\right), \tag{5.6}$$

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

$$P[\tau > t|\tau > s] = e^{-\int_s^t \lambda(u)du}. \tag{5.7}$$

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

$$h(t) = \lambda(t), \tag{5.8}$$

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

In the above Poisson models, the "intensity" λ is deterministic and thus gives a reduced form model with $\mathcal{F}_t = \mathcal{H}_t$ and $\mathcal{G}_t$ the trivial filtration. 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 needs 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.

5.2.3 Cox processes

Let us suppose that all the market information available in the economy, except for the default times, is expressed through the "market filtration" $\mathcal{G}_t$. For example, $\mathcal{G}_t$ 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 $\mathcal{G}_t$. Assume further that there exists a non-negative process λ_t which is also adapted to $\mathcal{G}_t$ that plays the role of a stochastic intensity, generally correlated with the different components of the driving process X_t.

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

$$\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t^N. \tag{5.9}$$

Definition 4. The point process N_t is a *Cox process* if, conditioned on the background information $\mathcal{G}_t$ available at time t, N_t is an inhomogeneous Poisson process with a time-varying intensity $\lambda(s) = \lambda_s$, for $0 \le s \le t$. That is

$$P[N_t - N_s = k|\mathcal{F}_s \vee \mathcal{G}_t] = \frac{(\Lambda_t - \Lambda_s)^k}{k!}e^{-(\Lambda_t - \Lambda_s)} \tag{5.10}$$

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

In other words, each realization of the process λ up to time t determines the local jump probabilities for the process N up to time 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 market 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 market information only). Finally, the doubly stochastic assumption proves to be inadequate for modeling the arrival of common credit events 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.

From the definition, for $s < t$

$$E[N_t - \Lambda_t|\mathcal{F}_s] = E[E[N_t - \Lambda_t|\mathcal{F}_s \vee \mathcal{G}_t]|\mathcal{F}_s] = N_s - \Lambda_s$$

and hence 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 unconditional 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,

$$P[N_t - N_s = k] = E\left[\frac{1}{k!}\left(\int_s^t \lambda_u du\right)^k \exp\left(-\int_s^t \lambda_u du\right)\right]. \tag{5.11}$$

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

$$\frac{d}{dt}E[\mathbf{1}_{\tau_{\text{next}}>t}|\mathcal{F}_s] = E\left[\lambda(t)e^{-\int_s^t \lambda_u du}|\mathcal{G}_s\right].$$

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

$$P[\tau > t] = E\left[e^{-\int_0^t \lambda_s ds}\right]. \tag{5.12}$$

Therefore, from the definition of hazard rate given in (4.22) we obtain

$$h(t) = -\frac{\partial}{\partial t}\log E\left[e^{-\int_0^t \lambda_s ds}\right]. \tag{5.13}$$

More generally, conditioned on the information $\mathcal{F}_s$ available at time $s \leq t$, the probability of survival after t years is

$$P[\tau > t|\mathcal{F}_s] = H_s^c E\left[e^{-\int_s^t \lambda_u du}|\mathcal{G}_s\right]. \tag{5.14}$$

It then follows from definition (4.25) that the forward default rates are given by

$$h_s(t) = -H_s^c \frac{\partial}{\partial t} \log E\left[e^{-\int_s^t \lambda_u du} | \mathcal{G}_s\right]. \tag{5.15}$$

An easy but important consequence of (5.10) is the following observation:

$$E[H_t^c|\mathcal{G}_t] = E[H_t^c|\mathcal{F}_0 \vee \mathcal{G}_t] = e^{-\int_0^t \lambda_u du}. \tag{5.16}$$

That is, conditioned on observing only the market filtration $\mathcal{G}_t$ at time t, the probability that the firm has not defaulted yet is $e^{-\int_0^t \lambda_u du}$.

Credit insurance products can be thought of as products that pay a stochastic amount at the time of default. Let the amount paid at time τ be X_τ where X_t is a process adapted to the market filtration $\mathcal{G}$. The expected payoff for insurance over the period $[t, T]$ computed at time t is found by an intermediate conditioning, followed by integrating over the possible times of default:

$$\begin{aligned} E[X_\tau(H_t^c - H_T^c)|\mathcal{F}_t] &= E[E[X_\tau(H_t^c - H_T^c)|\mathcal{F}_t \vee \mathcal{G}_T]|\mathcal{F}_t] \\ = E[H_t^c \int_t^T X_s \lambda_s e^{-\int_t^s \lambda_u du} ds|\mathcal{F}_t] &= H_t^c \int_t^T E[X_s \lambda_s e^{-\int_t^s \lambda_u du} ds|\mathcal{G}_t] \end{aligned} \tag{5.17}$$

This intuitive argument, in particular the last step, is justified in the next section. Observe that the final formula is a $\mathcal{G}$ expectation: the reduction from $\mathcal{F}$ to $\mathcal{G}$ we see here is the essential property that gives the name to the whole modelling approach.

Notice that (5.14) is mathematically equivalent to (A.21) with the stochastic intensity playing the role of a stochastic short rate. Therefore the mathematical apparatus for calculating bond prices in default-free interest rate theory described in section 3.2 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.

5.3 Definition of reduced form models

The above discussion of Cox processes rested on a subtle form of independence between the counting process and the market filtration. To place the theory on firmer ground, we now give a technical definition of a general reduced form model, and to deduce its fundamental properties.

We suppose as before that the full filtration is $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{G}_t$ where $\mathcal{H}_t$ is the natural filtration of the default indicator process $H_t = \mathbf{1}_{\{\tau \le t\}}$ (or more generally the filtration of the default counting process N_t) and $\mathcal{G}_t$ denotes the "market filtration". We are interested in examples where τ is a $\mathcal{F}$ stopping time, but not a $\mathcal{G}$ stopping time.

The following two key properties, known to be equivalent to one another, capture the specific dependence structure of Cox processes:

Definition 5. The $\mathcal{H}$-condition means: Under P, $\mathcal{H}_t$ and $\mathcal{G}_\infty$ are independent conditioned on $\mathcal{G}_t$, that is, for any $\mathcal{H}_t$-measurable random variable X and $\mathcal{G}_\infty$ measurable random variable Y,

$$E[XY|\mathcal{G}_t] = E[X|\mathcal{G}_t]E[Y|\mathcal{G}_t].$$

Definition 6. The *martingale invariance property* means: any $\mathcal{G}_t$-martingale is also a $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{G}_t$-martingale.

Assuming these conditions will give us enough structure to define what we mean by "reduced form model" or "Cox process". Taken together, the $\mathcal{H}$-condition and martingale invariance property have several modelling implications: (i) the default event cannot cause any observable effect on the "market filtration"; (ii) in multifirm versions, the $\mathcal{H}$-condition forbids "contagion", namely the effect that the default of one firm influences the default intensity of another firm. Therefore reduced form models exclude several effects that are often viewed as important in credit risk: this is one of the primary weaknesses pointed out by detractors.

Definition 7. A *reduced form model* of default is a model for which the default time is given by the random variable

$$\tau = \inf\{t > 0 : N_t > 0\}. \tag{5.18}$$

where N is a counting process with intensity. Furthermore, we assume that the intensity λ is $\mathcal{G}$-adapted, and the filtrations $\mathcal{H} = \sigma\{N\}, \mathcal{G}$ satisfy the $\mathcal{H}$-condition and martingale invariance property.

Thus, τ is the arrival time for the first jump of the counting process N_t, the default indicator process is $H_t = N_{t\wedge\tau}$ and the compensator of H is $\Lambda^H_t = \Lambda^N_{t\wedge\tau}$. Therefore, given the distribution for N_t, survival probabilities can be calculated as

$$P[\tau > T] = E[H^c_t] = P[N_T = 0],$$

where $H^c_t := 1 - H_t$ is the indicator function for survival.

The following proposition gives the basic analytical tools for reduced form models, allowing one to "reduce" default related computations from the full filtration $\mathcal{F}_t$ to default free computations in the market filtration $\mathcal{G}_t$.

Proposition 5.3.1. *In a reduced form model, for any $\mathcal{F}$-measurable random variable Y*

1.

$$E[H^c_tY|\mathcal{F}_t] = H^c_t\frac{E[H^c_tY|\mathcal{G}_t]}{E[H^c_t|\mathcal{G}_t]} \tag{5.19}$$

2.

$$E[H^c_t|\mathcal{G}_t] = e^{-\int_0^t \lambda_u du} \tag{5.20}$$

3. *If Y is $\mathcal{G}$-measurable, then*

$$E[H_t^c Y|\mathcal{F}_s] = H_s^c E[e^{-\int_s^t \lambda_u du} Y|\mathcal{G}_s], \quad s \leq t \tag{5.21}$$

4. *For any bounded $\mathcal{G}_t$ predictable process Y_t, and $s \leq t$,*

$$E[(H_s^c - H_t^c)Y_\tau|\mathcal{F}_s] = H_s^c \int_s^t E[Y_u e^{-\int_s^u \lambda_v dv} \lambda_u|\mathcal{G}_s]du \tag{5.22}$$

Remark 6. 1. Note that (5.19) implies that for any $\mathcal{F}_t$ random variable Y, there is an $\mathcal{G}_t$ random variable $\hat{Y}$ such that $H_t^c Y = H_t^c \hat{Y}$.

2. (5.20) and (5.22) are generalizations of facts (5.16) and (5.17) from the previous section.

3. By taking $Y = Y_s = e^{\int_0^s \lambda_u du}$ one can see from (5.21) that $H_t^c e^{-\int_0^t \lambda_u du}$ is an $\mathcal{F}$-martingale. This is an alternative way of thinking of the $\mathcal{F}$ compensator.

Proof: To prove (5.19), it is sufficient to prove

$$E[H_t^c Y E[H_t^c|\mathcal{G}_t]|\mathcal{F}_t] = H_t^c E[H_t^c Y|\mathcal{G}_t]$$

Recall that for any $\mathcal{F}$ random variable X, $E[X|\mathcal{F}_t]$ is the unique $\mathcal{F}_t$-random variable such that $\int_A E[X|\mathcal{F}_t]dP = \int_A X dP$ for all $A \in \mathcal{F}_t$. Let $C = \{\tau > t\}$ and note that for any set $A \in \mathcal{F}_t$ one can find $B \in \mathcal{G}_t$ such that $A \cap C = B \cap C$ (this statement is true because for any set $D \in \mathcal{H}_t$ either $D \supset C$ or $D \cap C = \emptyset$). Then

$$\begin{aligned} \int_A H_t^c Y E[H_t^c|\mathcal{G}_t]dP &= \int_{A\cap C} Y E[H_t^c|\mathcal{G}_t]dP = \int_{B\cap C} Y E[H_t^c|\mathcal{G}_t]dP \\ &= \int_B H_t^c Y E[H_t^c|\mathcal{G}_t]dP = \int_B E[H_t^c E[H_t^c Y|\mathcal{G}_t]|\mathcal{G}_t]dP \\ &= \int_B H_t^c E[H_t^c Y|\mathcal{G}_t]dP = \int_{B\cap C} E[H_t^c Y|\mathcal{G}_t]dP \\ &= \int_{A\cap C} E[H_t^c Y|\mathcal{G}_t]dP = \int_A H_t^c E[H_t^c Y|\mathcal{G}_t]dP \end{aligned} \tag{5.23}$$

which proves (5.19).

To prove (5.20), we define $\tilde{\Lambda}_t$ by $E[H_t^c|\mathcal{G}_t] = e^{-\tilde{\Lambda}_t}$ and show that $\int_0^t H_s^c d\tilde{\Lambda}_s$ must be the $\mathcal{F}$-compensator of H_t^c. First note that $\tilde{\Lambda}$ is nondecreasing: using the $\mathcal{H}$-condition on can show that for $s < t$, $e^{-\tilde{\Lambda}_s} - e^{-\tilde{\Lambda}_t} = E[(H_s^c - H_t^c)|\mathcal{G}_t] \geq 0$. By (5.19) and the fact $H_s^c H_t^c = H_t^c$ for $s \leq t$, one finds

$$E[H_t^c e^{\tilde{\Lambda}_t}|\mathcal{F}_s] = H_s^c e^{\tilde{\Lambda}_s} E[H_s^c H_t^c e^{\tilde{\Lambda}_t}|\mathcal{G}_s] = H_s^c e^{\tilde{\Lambda}_s} E[E[H_t^c e^{\tilde{\Lambda}_t}|\mathcal{G}_t]|\mathcal{G}_s] = H_s^c e^{\tilde{\Lambda}_s},$$

in other words, $H^c_t e^{\tilde\Lambda_t}$ is an $\mathcal{F}_t$-martingale. This in turn implies $H^c_t + \int_0^t H^c_s d\tilde\Lambda_s$ is a $\mathcal{F}_t$-martingale: we see this by using stochastic integration by parts to write

$$\begin{aligned} H^c_t - H^c_0 + \int_0^t H^c_s d\tilde\Lambda_s &= \int_0^t \left[d[H^c_s] + H^c_s e^{-\tilde\Lambda_s} d[e^{\tilde\Lambda_s}] \right] \\ &= \int_0^t e^{-\tilde\Lambda_s} d[H^c_s e^{\tilde\Lambda_s}]. \end{aligned}$$

By the uniqueness of the $\mathcal{F}$-compensator and the fact that both $\tilde\Lambda, \Lambda$ are $\mathcal{G}$-adapted, we find that $\tilde\Lambda_t = \Lambda_t$.

To show equation (5.21) note the fact $H^c_s H^c_t = H^c_t$ for $s \le t$ and use (5.19) and (5.20) to write:

$$E[H^c_t Y|\mathcal{F}_s] = E[H^c_s(H^c_t Y)|\mathcal{F}_s] = H^c_s e^{\Lambda_s} E[E[H^c_t Y|\mathcal{G}_t]|\mathcal{G}_s]$$

By the H-condition $E[H^c_t Y|\mathcal{G}_t] = E[H^c_t|\mathcal{G}_t]E[Y|\mathcal{G}_t] = E[e^{-\Lambda_t}Y|\mathcal{G}_t]$, and the result now follows.

We prove (5.22) for simple predictable processes $Y_t = \sum_{i=1}^N Y_{t_{i-1}} \mathbf{1}_{\{t_{i-1}<t\le t_i\}}$, where Y_{t_i} is $\mathcal{G}_{t_i}$-measurable and $s = t_0 \le t_1 \le \ldots t_N = t$:

$$\begin{aligned} E[(H^c_s - H^c_t)Y_\tau|\mathcal{F}_s] &= E[\sum_i Y_{t_{i-1}} \mathbf{1}_{\{t_{i-1}<\tau\le t_i\}}|\mathcal{F}_s] \\ &= H^c_s e^{\Lambda_s} \sum_i E[(e^{-\Lambda_{t_{i-1}}} - e^{-\Lambda_{t_i}})Y_{t_{i-1}}|\mathcal{G}_s] \\ &= H^c_s e^{\Lambda_s} \sum_i E[(\int_{t_{i-1}}^{t_i} e^{-\Lambda_u}\lambda_u du)Y_{t_{i-1}}|\mathcal{G}_s] \\ &= H^c_s \int_s^t E[Y_u e^{-\int_s^u \lambda_v dv}\lambda_u|\mathcal{G}_s]du \end{aligned} \tag{5.24}$$

Now we use approximation theory to extend to general processes Y_t. □

5.4 Constructing reduced form models

We now give two natural ways to construct a reduced form model, starting with a $\mathcal{G}_t$ adapted non-decreasing process $\Lambda_t = \int_0^t \lambda_s ds$.

5.4.1 Construction A

We let $Z_1, Z_2 \ldots,$ be a sequence of iid unit-mean exponential random variables that are independent of $\mathcal{G}$, and define recursively

$$\begin{aligned} \tau_1 &= \inf\{t|\Lambda_t \ge Z_1\} \\ \tau_n &= \inf\{t|\Lambda_t - \Lambda_{\tau_{n-1}} \ge Z_n\}, \quad n = 2,3,\ldots, \end{aligned} \tag{5.25}$$

Then we define $N_t = \sum_{k=1}^\infty \mathbf{1}_{\{\tau_k \le t\}}$. Verification of this construction is given as an exercise.

5.4.2 Construction B

Let $N_t^{(1)}$ be a Poisson process with $\lambda = 1$ that is independent of $\mathcal{G}$. Then the time changed process $N_t := N_{\Lambda_t}^{(1)}$ defines a reduced form model with $\mathcal{H}_t := \sigma(N_s, s \leq \Lambda_t)$ and $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{G}_t$. To prove this, note that

$$\begin{aligned} P[N_t - N_s = k | \mathcal{H}_s \vee \mathcal{G}_\infty] &= P[N_{\Lambda_t}^{(1)} - N_{\Lambda_s}^{(1)} = k | \mathcal{G}_{\Lambda_s}^{N^{(1)}} \vee \mathcal{G}_\infty] \\ &= \frac{(\Lambda_t - \Lambda_s)^k}{k!} e^{-(\Lambda_t - \Lambda_s)} \end{aligned} \tag{5.26}$$

5.4.3 Simulating the default time

As is well-known in elementary probability theory, given a non-decreasing cádlág (french for "left limit, right continuous") function $F : \mathbb{R} \to [0,1]$ such that

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1,$$

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

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

Therefore, if a model for default arrival is such that the survival probabilities $P[\tau > t]$ 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

$$P[\tau > t] = U.$$

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 up to time t, we have

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

as required.

5.5 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 for $s < t$ the survival probabilities can be written in the form

$$P[\tau > t|\mathcal{F}_s] = H_s^c \exp[A(s,t) + B(s,t) \cdot X_s] \tag{5.27}$$

for some coefficient functions $A(s,t)$ and $B(s,t) \in \mathbb{R}^n$. In section 3.2.2 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 driver and can present jumps or stochastic volatilities.

5.5.1 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

$$d\lambda_t = k(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t, \tag{5.28}$$

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

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

where

$$A(s,t) = \frac{2k\theta}{\sigma^2} \log\left[\frac{2\gamma \exp[(k+\gamma)(t-s)/2]}{2\gamma + (k+\gamma)(\exp[(t-s)\gamma] - 1)}\right] \tag{5.29}$$

$$B(s,t) = \frac{2(1 - \exp[(t-s)\gamma])}{2\gamma + (k+\gamma)(\exp[(t-s)\gamma] - 1)} \tag{5.30}$$

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 (5.14).

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.

5.5.2 Mean-reverting intensities with jumps

Suppose now that the intensity process follows the dynamics

$$d\lambda_t = k(\theta - \lambda_t)dt + dZ_t \tag{5.31}$$

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. The Markov generator for this process is

$$[\mathcal{A}f](x) = [k(\theta - x) - cJ]\partial_x f + c\int_0^\infty [f(x+y] - f(x)]e^{-y/J}dy/J. \tag{5.32}$$

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

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

where

$$\begin{aligned} A(s,t) &= -(\theta - cJ/k)\left((t-s) + B(s,t)\right) - \frac{c}{J+k}\left[Jt - \log\left(1 - B(s,t)J\right)\right] \\ B(s,t) &= -\frac{1-e^{-k(t-s)}}{k}. \end{aligned} \tag{5.33}$$

5.6 Risk-neutral and physical measures

In anticipation of the next chapter which deals with the pricing of defaultable bonds and credit derivatives, we now turn our attention to reduced form models under a risk-neutral measure Q. Given a reduced form model under the physical measure P, it does not necessarily follow that the model will be of reduced form under the risk–neutral measure Q: the doubly stochastic assumption needs to be independently stated for P and Q. Moreover, the intensities λ_t^P and λ_t^Q themselves can depend differently on the state variables of the model, and may also have different likelihoods for each path. Even in the situation where $\lambda_t^P = \lambda_t^Q$ we can still have that

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

is different from

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

The form of the Girsanov Theorem applicable to reduced form models asserts the existence of a measure change $\frac{dQ}{dP}|_{\mathcal{F}_T}$ for any adapted processes θ_t, η_t (satisfying appropriate technical conditions) such that if

$$W_t^Q = W_t + \int_0^t \theta_s ds, \quad \lambda_t^Q = \eta_t \lambda_t,$$

then N_t, the counting process with intensity λ under P, has intensity λ_t^Q under Q and W_t^Q is Q-Brownian motion.

We still define the default event as the first jump of the counting process N_t, but the point is that the random default time

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

has a different distribution under Q. Accordingly, we assume that the counting process N_t is such that

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

is 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,

$$Q[\tau > t|\mathcal{F}_s] = H_s^c E^Q \left[e^{-\int_0^t \lambda_u^Q du} |\mathcal{G}_s \right], \tag{5.34}$$

so the survival probability expressions we just derived for affine intensity models apply for risk-neutral probabilities as well. The *credit risk premium*, defined to be the ratio $\eta_t := \lambda_t^Q / \lambda_t^P$, is an object of interest when comparing risk neutral and statistical estimates of the default intensity.

In the next chapter, we will investigate the risk neutral pricing of a variety of contingent claims whose value depends significantly on the credit risk of a counterparty. The most basic of all such contracts is the zero-coupon zero-recovery defaultable bond, and in Exercise 29 we already prove the *Lando formula* for its price, in the most general reduced form model:

$$H_t^c \bar{P}_t(T) := E^Q[H_T^c e^{-\int_t^T r_s ds} |\mathcal{F}_t] = H_t^c E^Q[e^{-\int_t^T [r_s + \lambda_s^Q] ds} |\mathcal{G}_t]$$

5.7 Exercises

Exercise 28. Show from the definition of a Poisson process N_t with parameter λ that the time of the first jump of N_t is an exponential random variable with parameter λ. For each $n \geq 1$, compute the probability density function of the time of the nth jump $\tau^{(n)} := \inf\{t | N_t \geq n\}$. Using Matlab, plot on a single graph the 5 curves with $n = 1, 2, 3, 4, 5$ and $\lambda = 1$.

Exercise 29. Let the spot interest rate r_t be $\mathcal{G}_t$-adapted. Prove *Lando's formula* for the price of a zero-coupon, zero-recovery bond:

$$H_t^c \bar{P}_t(T) := E^Q[H_T^c e^{-\int_t^T r_s ds} |\mathcal{F}_t] = H_t^c E^Q[e^{-\int_t^T [r_s + \lambda_s] ds} |\mathcal{G}_t] \tag{5.35}$$

Exercise 30. Adapt Construction A to a construction of a reduced form model for the default indicator process H_t, starting with an $\mathcal{G}_t$ adapted intensity process λ_t. Verify all the details of the construction.

Exercise 31. The aim is to verify the equivalence of the $\mathcal{H}$-condition and the martingale invariance condition.

1. Show that the $\mathcal{H}$-condition implies the martingale invariance condition. (Hint: argue that it is enough to show that for any $\mathcal{G}_s$ set B and $\mathcal{H}_s$ set C, and $\mathcal{G}$-martingale M, $\int_{B\cap C} M_s dP = \int_{B\cap C} M_t dP$.)

2. Show that the martingale invariance condition implies the $\mathcal{H}$-condition. (Hint: argue that it is enough to show that for any $\mathcal{G}_s$ set B and $\mathcal{H}_s$ set C, and $\mathcal{G}_t$-random variable Y $\int_B \mathbf{1}_C Y dP = \int_B E[\mathbf{1}_C|\mathcal{G}_s]E[Y|\mathcal{G}_s]dP$.)

Exercise 32. Prove by induction on n that for Construction A, $P_n(t) := P[\tau_n \le t|\mathcal{G}_\infty] = \sum_{k=n}^{\infty} \frac{\Lambda_t^k}{k!} e^{-\Lambda_t}$. (Hint: for the inductive step $k = n$ assuming the formula for $n-1$, use the derivative of P_{n-1} and the fact the $\tau_n \le t$ implies $\tau_{n-1} \le t$.)

Exercise 33. Consider the following SDE for an intensity process:

$$d\lambda_t = (a - b\lambda_t)dt + c\sqrt{\lambda_t}dW_t \tag{5.36}$$

where $4a > \sigma^2$. Compute the conditional survival function $E[H_t^c|\mathcal{G}_s], s < t$ in the reduced form model with this intensity.

Exercise 34. Write a MATLAB program to draw Monte Carlo samples from the following distributions:

1. Poisson distribution with parameter $\lambda > 0$.

2. The exponential distribution with parameter λ.

3. The gamma distribution $\Gamma(k, \theta)$ with k an integer.

Exercise 35. Consider a stochastic interest model where λ and r are correlated linear combinations

$$\begin{aligned} r_t &= aX_t + bY_t \\ \lambda_t &= cX_t + dY_t \end{aligned} \tag{5.37}$$

of independent driving processes X, Y. We take the following default values of the parameters: $a = 1, c = d = .1, b = 0$, with X, Y both Vasicek processes with $(k_X, \theta_X, \sigma_X) = (.5, 0.05, .05)$ and $(k_Y, \theta_Y, \sigma_Y) = (.1, 0.1, .025)$. In addition, we assume a constant recovery fraction $R = 0.5$. Investigate the dependence of survival probabilities on r. (Hint: use the “master formula” (3.70).)

Exercise 36. Consider the jump model with intensity process given by (5.31), but with Poisson jumps of a single size J. Derive the generalized Riccati equations for the survival probability.

Exercise 37. Write a MATLAB Monte Carlo simulator that will generate a sample path of the mean-reverting exponential jump process of subsection 5.5.2.

Exercise 38. Consider the jump model with intensity process given by (5.31) with exponentially distributed jumps. Derive the generalized Riccati equations for the survival probability, and prove (5.33).

Exercise 39. Show that the $\mathcal{H}$-condition implies the following two statements:

1. For any $\mathcal{G}_\infty$ random variable X, $E[X|\mathcal{H}_t \vee \mathcal{G}_t] = E[X|\mathcal{G}_t]$.

2. Any $\mathcal{G}_s$ martingale M_s is a $\mathcal{F}_s$-martingale.

In fact, it can be shown that the second statement implies the $\mathcal{H}$-condition.

Exercise 40. Complete the proof of (5.21).

Exercise 41. Let Y be any $\mathcal{F}$ measurable random variable and suppose $t < s$. Show that there is a $\mathcal{G}_s$ random variable $\tilde{Y}$ such that $E[H^c_s XY|\mathcal{F}_t] = H^c_t E[H^c_s X\tilde{Y}|\mathcal{G}_t]$ for any $\mathcal{G}_s$ random variable X.

Exercise 42. Let $X_t = (X^1_t, \ldots, X^n_t)$ be a multidimensional Markov "factor" process with generator $\mathcal{L}$. Consider any reduced form model with intensity $\lambda_t = g(t, X_t)$ for a given positive function $g(t, x)$.

1. Suppose $Y = F(X_T)$ for some function F. Show directly using the Itô formula that the conditional expectation $E[Y|\mathcal{F}_t]$ has the form $H^c_t f(t, X_t)$ where the function f is the solution of the parabolic partial differential equation:

$$\begin{cases} \partial_t f(t,x) + \mathcal{L}[f](t,x) - g(t,x)f(t,x) = 0 \quad t < T \\ f(T,x) = F(x) \end{cases} \tag{5.38}$$

2. Let $Z_t = h(\tau, X_\tau)H_t$ for $t \leq T$. Show that $E[Z_T|\mathcal{F}_t]$ has the form $Z_t + H^c_t f(t, X_t)$ where the function f solves:

$$\begin{cases} \partial_t f(t,x) + \mathcal{L}[f](t,x) + g(t,x)(h(t,x) - f(t,x)) = 0 \quad t < T \\ f(T,x) = 0 \end{cases} \tag{5.39}$$

Exercise 43. For any $\mathcal{G}$-adapted process Z_t and $\mathcal{G}_T$-measurable X, let $Y = Z_\tau H_T + X_T H^c_T$ be a $\mathcal{F}_T$ measurable payoff. Write $E[H^c_t Y|\mathcal{F}_t]$ in terms of $\mathcal{G}$ expectations.

Exercise 44. (Counterparty risk) Suppose you have bought an option with maturity T from a default-risky counterparty. In a general reduced form model, show that the fair price for the claim at time t is

$$H^c_t \tilde{V}_t = H^c_t E^Q[e^{-\int_t^T (r_s + \lambda_s)ds} V_T|\mathcal{G}_t] \tag{5.40}$$

where V_T, the payment in the event the counterparty is solvent at time T, is $\mathcal{G}_T$ measurable.

Exercise 45. (Counterparty risk continued) Suppose instead, you have written a swap with maturity T from the same risky counterparty. In this case, one might assume that in the event that the counterparty defaults at time $\tau > t$, one still owes the negative part of the contract value. In a general reduced form model, show that the fair price for the claim at time t satisfies

$$\tilde{V}_t = H_t^c E^Q[e^{-\int_t^T (r_s+\lambda_s)ds} V_T | \mathcal{G}_t] + H_t^c \int_t^T E^Q[e^{-\int_t^u (r_s+\lambda_s)ds} (\tilde{V}_u)^- \lambda_u | \mathcal{G}_t] du. \tag{5.41}$$

Note that V_T, the payment in the event the counterparty is solvent at time T, is $\mathcal{G}_T$ measurable, and that $(x)^- = x \wedge 0 = \min(x, 0)$.

Chapter 6

Further Reduced Form Modelling

Credit derivatives in the general sense are contracts that are linked to an underlying credit-sensitive asset. This reference asset may in principle be any default risky instrument or family of instruments, but for the most fundamental credit derivatives we will discuss, the underlier will be a corporate bond, in the simplest instances a zero-coupon bond.

In our treatment of reduced form default models, we have seen in (5.35) that the price of a zero coupon defaultable bond with zero recovery is given by Lando's formula

$$\bar{P}_t(T) = E_t^Q[e^{\int_t^T (r_s+\lambda_s)ds}]. \tag{6.1}$$

Such a contract is principally of mathematical interest: in practice, all corporate bonds retain a non-negligible residual value after the firm defaults, through the process we shall call *recovery*. Before studying credit derivatives, we begin by looking at different methods for valuing the post default value of zero coupon bonds, leading to formulas for the price of bonds with recovery. Then we define and analyze a number of standard credit derivative contracts whose payoffs are defined in terms of these underlying bonds.

6.1 Recovery Models

So far, 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 risky assets to drop by a stochastic amount in the event of a default at time τ. We consider the doubly stochastic model of section 5.2.3 with a stochastic intensity λ_t and assume that the *recovery process* $\mathcal{R}_t$ is $\mathcal{G}_t$–adapted. Assuming that $\tau > 0$, the price at time zero of a defaultable zero coupon bond with maturity T is given by

$$\bar{P}_0^{\mathcal{R}}(T) = E^Q\left[e^{-\int_0^\tau r_s ds}\mathcal{R}_\tau \mathbf{1}_{\{\tau\le T\}} + e^{-\int_0^T r_s ds}\mathbf{1}_{\{\tau>T\}}\right]$$

By (5.21) and (5.22) applied with $s = 0$ this becomes

$$
\begin{aligned}
&E^Q\left[\int_0^T e^{-\int_0^u (r_s+\lambda_s)ds}\mathcal{R}_u\lambda_u du\right] + E^Q\left[e^{-\int_0^T (r_s+\lambda_s)ds}\right] \\
&= \int_0^T E^Q\left[\mathcal{R}_u\lambda_u e^{-\int_0^u (r_s+\lambda_s)ds}\right] du + E^Q\left[e^{-\int_0^T (r_s+\lambda_s)ds}\right] \qquad (6.2)
\end{aligned}
$$

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

Let us now consider an idealized coupon bond that pays a constant coupon rate $c > 0$ continuously in time (this is simply mathematical convenience). The price of such a coupon bond including recovery would then be

$$
\begin{aligned}
&E^Q\left[e^{-\int_0^\tau r_s ds}\mathcal{R}_\tau \mathbf{1}_{\{\tau\le T\}} + \int_0^T c e^{-\int_0^u r_s ds}\mathbf{1}_{\{\tau>u\}}du + e^{-\int_0^T r_s ds}\mathbf{1}_{\{\tau>T\}}\right] \\
&\qquad = \int_0^T E^Q\left[(\mathcal{R}_u\lambda_u + c)e^{-\int_0^u (r_s+\lambda_s)ds}\right] du + E^Q\left[e^{-\int_0^T (r_s+\lambda_s)ds}\right] \qquad (6.3)
\end{aligned}
$$

Remark 7. Before discussing various models of recovery, we review here a few basic facts concerning the nature of default. The default of a firm is sometimes marked by a failure of the firm to make a contracted debt repayment, either of a coupon or the principal of any bond. More typically, default is triggered by the initiation of bankruptcy proceedings, in the US by either the Chapter 7 or Chapter 11 proceedings. In broad terms, these result in the "liquidation value" or post-bankruptcy value L of the firm, a random amount normally somewhat less than the value K of the debt outstanding, being redistributed "fairly" amongst the debt holders. "Fairly" means in proportion to the value of their holdings for bonds of similar seniority. Holders of subordinated bonds receive value only after senior bondholders are paid in full. Equity holders typically receive nothing. The *recovery fraction* $R := L/K$ is thus a stochastic quantity. Some of the observable factors that influence the recovery fraction L/K of the firm are:

1. the state of the economy: Numerous studies have verified that recovery values will be lower in an economic recession, for example 2002 average recoveries were 25% while in 2006 they averaged 0%;

2. the composition of the assets of the particular firm: for example, software firms have little value in "hard assets" such as machinery and real estate, hence a low recovery fraction. In general one can say that different sectors of the economy tend to yield different relative recovery values;

3. pre-default credit quality: CCC firms tend to recover much less than BBB firms ("junk" versus "fallen angels")*.

Realistic stochastic modelling of the recovery fraction is a subject of current research.

Given the principle of a fair distribution of the liquidation value, the secondary question arises of how the firm's outstanding bonds, comprising various maturities, seniority and coupon rates, are to be compared and then valued after default. The various mechanisms we now discuss are in essence different mathematical answers to the question how to assign a fair value to each possible bond, given that the overall recovery fraction is R. In some circumstances, for example when the pre-default values of bonds are near par, the differences between methods will not be great. In other cases, for example, when coupon rates of different bonds are quite different, the different mechanisms lead to quite different results.

6.1.1 Recovery of par

Under this model, a coupon bond will pay a fraction $0 < R < 1$ of its principal at the time of default. Under recovery of par, the recovery process is $\mathcal{R}_t = R$ for $t \geq 0$. The same calculation as before leads to

$$\bar{P}_0^{RP}(T) = \int_0^T E^Q \left[(R\lambda_u + c) e^{-\int_0^u (r_s + \lambda_s) ds} \right] du + \bar{P}_0(T). \tag{6.4}$$

This simple approach might be very unfair: for example, at the time of default, a long maturity zero coupon bond is clearly worth less than a similar one of shorter maturity, but recovery of par would value them equally. However, recovery of par can be justified if the bonds in question are coupon bonds that trade close to par.

6.1.2 Recovery of treasury

A second mechanism is to consider a recovery of the form

$$\mathcal{R}_t = R\ P_t(T), \tag{6.5}$$

at the time of default of a zero coupon bond. In general, the recovery value is a constant fraction $0 < R < 1$ of the equivalent default-free bond, where "equivalent" means the bond with the same coupons and maturity. This partially corrects the time inconsistency of the recovery of par method, but still values long dated "junk bonds" (bonds with high default risk and hence high yield) higher than their "true value". Thus it works well for highly rated firms, but badly for junk bonds.

For a zero coupon bond with recovery of treasury, we have

$$\begin{aligned}
\bar{P}_0^{RT}(T) &= R \int_0^T E^Q \left[P_u(T) \lambda_u e^{-\int_0^u (r_s + \lambda_s) ds} \right] du + \bar{P}_0(T) \\
&= R E^Q \left[e^{-\int_0^T r_s ds} \int_0^T \lambda_u e^{-\int_0^u \lambda_s ds} du \right] + \bar{P}_0(T) \\
&= R E^Q \left[e^{-\int_0^T r_s ds} \left(1 - e^{-\int_0^T \lambda_s ds} \right) \right] + \bar{P}_0(T) \\
&= R P_0(T) + (1 - R) \bar{P}_0(T). \qquad (6.6)
\end{aligned}$$

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

The value of a coupon bond under the RT mechanism is given as Exercise 46.

6.1.3 Recovery of market value

The final model for recovery, introduced by [5], assumes that if default happens at time τ then each defaultable bond pays a fraction R of its pre-default market value. This corresponds to a recovery rate defined as

$$\mathcal{R}_t = R\bar{P}^{RMV}_{t-}(T), \tag{6.7}$$

where

$$\bar{P}^{RMV}_{t-}(T) = \lim_{s\nearrow t} \bar{P}^{RMV}_s(T) \tag{6.8}$$

and $\bar{P}^{RMV}_t(T)$ denotes the pre-default 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 more laborious than for other recovery models and is done in more generality in the next section. The resulting predefault price is

$$\bar{P}^{RMV}_0(T) = E^Q\left[e^{-\int_0^T (r_s+(1-R)\lambda_s)ds}\right] \tag{6.9}$$

6.1.4 Stochastic recovery models

All three of the above mechanisms can be extended to allow the recovery fraction R to be an $\mathcal{G}_t$ adapted process with values in $[0,1]$. Perhaps the most elegant result is the *generalized Lando formula*:

Theorem 6.1.1 (Generalized Lando Formula). *In any reduced form model with stochastic interest r_t, intensity λ_t, and stochastic recovery fraction R_t, all $\mathcal{G}_t$ adapted processes, the value of a zero coupon bond under recovery of market value is given by*

$$\bar{P}^{RMV}_t(T) = E^Q_t\left[e^{-\int_t^T (r_s+(1-R_s)\lambda_s)ds}\right] \tag{6.10}$$

Proof: The risk neutral valuation formula for the predefault value $V_t := P^{RMV}_t(T)$ of the bond can be rewritten using (5.22):

$$\begin{aligned} H^c_t V_t &= E^Q\left[H^c_T e^{-\int_t^T r_s ds} + [H^c_t - H^c_T]e^{-\int_t^\tau r_s ds} R_\tau V_\tau \middle| \mathcal{F}_t\right] \\ &= H^c_t E^Q\left[e^{-\int_t^T (r_s+\lambda_s)ds} + \int_t^T e^{-\int_t^u (r_s+\lambda_s)ds} R_u V_u \lambda_u du \middle| \mathcal{G}_t\right] \end{aligned} \tag{6.11}$$

Introduce the $\mathcal{G}_t$-martingale $M_t = E^Q[e^{-\int_0^T (r_s+\lambda_s)ds} + \int_0^T e^{-\int_0^u (r_s+\lambda_s)ds} R_u V_u \lambda_u du|\mathcal{G}_t]$ and rewrite the formula for V_t as

$$V_t = e^{\int_0^t (r_s+\lambda_s)ds}\left(M_t - \int_0^t e^{-\int_0^u (r_s+\lambda_s)ds} R_u V_u \lambda_u du\right)$$

By the Itô formula,

$$d\left(e^{-\int_0^t (r_s+(1-R_s)\lambda_s)ds} V_t\right) = e^{\int_0^t R_s\lambda_s ds} dM_t.$$

Since the right side is a martingale, we can integrate from $[t,T]$, take $\mathcal{G}_t$-conditional expectations, and find:

$$E^Q[e^{-\int_0^T (r_s+(1-R_s)\lambda_s)ds} V_T|\mathcal{G}_t] - e^{-\int_0^t (r_s+(1-R_s)\lambda_s)ds} V_t = 0,$$

which, since $V_T = 1$, yields the result. □

Remark 8. The combination $h_t^Q := (1-R_t)\lambda_t^Q$, sometimes called the "short spread", occurs throughout formulas when the RMV mechanism is assumed. It is not possible in the framework to separate R_t from λ_t. For this reason, in this framework one follows [5] and models h_t^Q directly.

6.2 Two-factor Gaussian models

As a concrete example of a calculation of prices for defaultable bonds according to (6.10), consider a reduced form model for which both the risk-free rate and the short spread process follow a Hull-White process, that is, let

$$dr_t = \kappa(\Theta(t) - r_t)dt + \sigma dW_t^Q. \tag{6.12}$$

and

$$dh_t^Q = \bar\kappa(\bar\Theta(t) - h_t^Q)dt + \bar\sigma dZ_t^Q, \tag{6.13}$$

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

$$dW_t^Q dZ_t^Q = \rho dt. \tag{6.14}$$

The function $\Theta(t)$ can be chosen to match the initial term structures of default-free bonds according to (3.45). We can now use proposition 3.3.2 to calculate the price of zero coupon bonds. Similarly, if for the moment we assume $h^Q = \lambda^Q$ (i.e. the default intensity solves (6.13)) risk-neutral survival probabilities are given by

$$Q[\tau > T|\mathcal{F}_t] = \exp[\bar A(t,T) + \bar B(t,T) h_t^Q] \tag{6.15}$$

where

$$\bar B(t,T) = \frac{1}{\bar\kappa}(e^{-\bar\kappa(T-t)} - 1) \tag{6.16}$$

and

$$\bar{A}(t,T) = -\frac{1}{2}\int_t^T \sigma^2 \bar{B}(s,T)^2 ds - \int_t^T \bar{B}(s,T)\bar{\kappa}\bar{\Theta}(s)ds. \tag{6.17}$$

More importantly, the price of a defaultable bond with recovery of market value in this model is given by

$$\bar{P}_t^{RMV}(T) = P_t(T)\exp[\tilde{A}(t,T) + \bar{B}(t,T)h_t^Q], \tag{6.18}$$

where

$$\tilde{A}(t,T) = -\frac{1}{2}\int_t^T \bar{B}(s,T)\left[\sigma^2\bar{B}(s,T) + 2\bar{\kappa}\bar{\Theta}(s) + \rho\bar{\sigma}\sigma B(s,T)\right] ds. \tag{6.19}$$

Equation (6.18) shows that in this simple model the yield spread is

$$YS_t(T) := \frac{1}{T-t}\int_t^T h_t(s)ds = \frac{-\tilde{A}(t,T) - \bar{B}(t,T)h_t^Q}{T-t}$$

and the forward default rate process is

$$h_t(s) = \dot{\tilde{A}}(t,T) + \dot{\bar{B}}(t,T)h_t^Q.$$

We see that credit spread curves, being affine in h_t^Q, are stochastic.

6.3 General Two Factor Models

Another way to combine r_t, λ_t was introduced in [5]. They write $[r,\lambda]$ as linear combinations of independent processes X, Y:

$$[r_t, \lambda_t]' = C[X_t, Y_t]' \tag{6.20}$$

for some constant matrix $C = [c_{11}\ c_{12}; c_{21}\ c_{22}]$. Let's suppose that $F^{(X)}(T, X_0, u) = E^Q_{X_0}[e^{-\int_0^T uX_s ds}]$ is known explicitly (for example, if X is Vasicek, then $F^{(X)}$ is given by (3.70). We will also need another function:

$$F^{(X,1)}(T, X_0, u) := E^Q_{X_0}[X_T e^{-\int_0^T uX_s ds}] = -u^{-1}\partial_T F^{(X)}(T, X_0, u)$$

Similarly we suppose $F^{(Y)}, F^{(Y,1)}$ are known explicitly.

The following proposition gives formulas for zero coupon bonds of various sorts:

Proposition 6.3.1. *Suppose r_t, λ_t are specified by (6.20). Suppose the recovery fraction R is constant. Let $[X_0, Y_0]' = C^{-1}[r_0, \lambda_0]'$. Then*

1. *The time 0 price of a default free zero coupon bond of maturity T is:*

$$P_0(T) = F^{(X)}(T, X_0, c_{11} + c_{21})$$

2. *The time* 0 *price of a defaultable zero recovery zero coupon bond of maturity* T *is:*

$$\bar{P}_0(T) = F^{(X)}(T, X_0, c_{11} + c_{21})F^{(Y)}(T, Y_0, c_{12} + c_{22})$$

3. *The time* 0 *price of a defaultable zero coupon bond of maturity* T *with recovery of par is:*

$$\begin{aligned}\bar{P}_0^{RP}(T) &= \bar{P}_0^0(T) + R\int_0^T \Big[c_{21}F^{(X,1)}(s, X_0, c_{11} + c_{21})F^{(Y)}(s, Y_0, c_{12} + c_{22}) \\ &\quad + c_{22}F^{(X)}(s, X_0, c_{11} + c_{21})F^{(Y,1)}(s, Y_0, c_{12} + c_{22})\Big] ds\end{aligned}$$

4. *The time* 0 *price of a defaultable zero coupon bond of maturity* T *with recovery of market value is:*

$$\bar{P}_0^{RMV}(T) = F^{(X)}(T, X_0, c_{11} + (1-R)c_{21})F^{(Y)}(T, Y_0, c_{12} + (1-R)c_{22})$$

The proof of these formulas can be left as an exercise.

6.4 Credit Derivatives

A primary motivation in designing credit derivatives is similar to the rationale behind insurance: credit derivatives can enable the transfer of credit risk from one party (the "buyer" of default insurance) to another party (the "seller" of insurance). In the early 1990's, this was often achieved by financial arrangements such as bond insurance whereby the issuer of the bond pays a bank an annual fee in return for the bank's promise to make interest and principal payments in the event the issuer defaults. An example of bond insurance is given in Exercise 56. More recently, the favoured contract is the *credit default swap* that we have already seen in section 2.3.3.

6.4.1 Credit Default Swaps

As we have seen, in this default insurance product the protection buyer A pays the protection seller B a regular premium $s_C\mathcal{N}\Delta t$, at fixed intervals, until $T \wedge \tau$, where τ denotes the time of default of the reference entity C and $\mathcal{N}$ denotes the "notional amount" of the C-bond specified as the reference security. If $\tau > T$, no insurance payment will be made, but if $\tau \le T$ B has to make a default payment to A. Let the payment dates be $t_k = k\Delta t, i = 1, 2, \ldots, N$ where the maturity date is $T = t_N = N\Delta t$.

By risk-neutral valuation, the value at time $t < t_1$ of the payment leg is

$$s_C\mathcal{N}\Delta t \sum_{k=1}^{N} E^Q[e^{\int_t^{t_k} r_s ds}\mathbf{1}_{\{\tau > t_k\}}]$$

which evaluates to $s_C \mathcal{N} \Delta t \sum_{k=1}^N \bar{P}_t(t_k)$. Typically there is an additional "accrual" payment for the fraction of the period in which default happens: a possible valuation formula is

$$s_C \mathcal{N} \sum_{k=1}^N E^Q[(\tau - t_{k-1}) e^{\int_0^{t_k} r_s ds} \mathbf{1}_{\{t_{k-1} < \tau \le t_i\}}],$$

representing the situation that the accrual payment is contracted to be made at the first payment date after default, and covers the period between the previous payment date and default. Note that with the accrual payment included, the total premium leg is a good approximation to the limiting case of continuous payments made at the rate of s_C per unit notional up to the time $T \wedge \tau$.

When the reference security is a coupon paying C-bond that trades close to par, the default settlement will be well modelled by the recovery of par mechanism with a stochastic recovery rate R_t. Under physical settlement, B pays A the full notional amount $\mathcal{N}$ in exchange for the defaulted bond, at the first payment date following default. The value of the defaulted coupon bond at the time of default will be $R_\tau \mathcal{N}$. Then the market value at time $t < t_1$ of a default payment at time t_k is

$$\begin{aligned} b_k &= \mathcal{N} E^Q \left[e^{-\int_t^{t_k} r_s ds} (1 - R_\tau) \mathbf{1}_{\{t_{k-1} < \tau \le t_k\}} | \mathcal{F}_t \right] \\ &= \mathcal{N} E^Q \left[e^{-\int_t^{t_k} r_s ds} \int_{t_{k-1}}^{t_k} (1 - R_u) \lambda_u^Q e^{-\int_t^u \lambda_s ds} du | \mathcal{G}_t \right] \qquad (6.21) \end{aligned}$$

Therefore the total value of the CDS at time t (neglecting the accrual term) is

$$CDS(t, s, \mathcal{N}, \mathcal{T}) = \sum_{k=1}^K (b_k - s_C \mathcal{N} \Delta t \bar{P}_t(t_k)). \qquad (6.22)$$

A useful mathematical idealization of the CDS contract supposes that the default premium is paid continuously in time at rate s until $T \wedge \tau$, while the default payment is made at the instant of default. We suppose now that at default τ, the defaulted bond is exchanged for the equivalent treasury bond, and moreover the post-default value of the bond equals R_τ times the equivalent treasury bond. We call this the recovery of treasury settlement mechanism. In this case, the value V_t at time $t < T \wedge \tau$ of the premium leg and W_t, the value of the default leg, are given by

$$\begin{aligned} V_t &= s \mathcal{N} E_t^Q \left[\int_t^T e^{-\int_t^u (r_s + \lambda_s) ds} du | \mathcal{G}_t \right] \\ W_t &= \mathcal{N} E_t^Q \left[\int_t^T e^{-\int_t^u r_s ds} (1 - R_u) \lambda_u e^{-\int_t^u \lambda_s ds} du | \mathcal{G}_t \right]. \qquad (6.23) \end{aligned}$$

6.5 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.

6.5.1 Discrete-time Markov chain

As a first step, one can consider *historical transition frequencies* published by a rating 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 (numbered zero to 7) represents a rating for the beginning of a year in order of increasing credit quality, each column represents the end-of-year rating, and the zeroth row and column correspond to default:

$$\Pi = \begin{bmatrix} 1.000 & .0000 & .0000 & .0000 & .0000 & .0000 & .0000 & .0000 \\ .2768 & .6236 & .0615 & .0285 & .0095 & .0000 & .0000 & .0000 \\ .0696 & .0329 & .8244 & .0643 & .0061 & .0022 & .0004 & .0001 \\ .0144 & .0058 & .0893 & .8241 & .0596 & .0059 & .0007 & .0003 \\ .0017 & .0008 & .0102 & .0582 & .8547 & .0701 & .0036 & .0006 \\ .0001 & .0001 & .0018 & .0069 & .0581 & .9028 & .0297 & .0006 \\ .0003 & .0000 & .0001 & .0011 & .0032 & .0925 & .8913 & .0114 \\ .0000 & .0000 & .0000 & .0003 & .0000 & .0106 & .0978 & .8914 \end{bmatrix} \qquad (6.24)$$

That is, the entry $\Pi_{75} = 0.0106$ gives the historical percentage of *Aaa* companies that have been downgraded to *A* rating during a one year period, averaged over the years 1980 to 2000. The entry $\Pi_{30} = 0.0144$ corresponds to the observed one-year default probability of 1.44% for *Ba* firms. Observe that the zeroes in the zeroth row indicate that no company can recover from the default state, which is thus said to be an *absorbing state*.

One can then model rating transitions over time as a discrete time *Markov chain* $Y_n \in \{0, 1, \ldots, K\}$ whose $K+1$ states are the possible rating classes (including default) at the end of year n 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

$$P[Y_n = j | Y_{n-1} = i] = \Pi_{ij}.$$

This simple and elegant model suffers from several pitfalls, mainly related to the fact that historical rating 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 statistically 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 extend the model in order to incorporate the influence of macro-economic 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, Y_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 states and 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.

6.5.2 Continuous-time Markov chain

We can obtain a continuous-time credit rating Markov chain model by assuming a process $Y_t \in \{0, 1, 2, \ldots, 7\}$ that undergoes transitions from state i to a state j with an intensity Λ_{ij}. That is, assume that, for an infinitesimal time interval dt, the transition probabilities to move from i to $j \neq i$ are

$$P[Y_{t+dt} = j|Y_t = i] \sim \Lambda_{ij}dt. \tag{6.25}$$

The matrix Λ is then called the *generator* for the continuous-time Markov chain Y_t and has the properties of a *stochastic matrix*. That is, the off-diagonals are non-negative

$$\Lambda_{ij} \geq 0, \qquad \text{for all } i \neq j \tag{6.26}$$

and the row sums are zero, implying

$$\Lambda_{ii} = -\sum_{j|j\neq i}\Lambda_{ij}. \tag{6.27}$$

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 exponential[1]

$$\Pi(s,t) = e^{\Lambda(t-s)}. \tag{6.28}$$

(Recall the definition $e^{tA} = \sum_{n=0}^{\infty} \frac{1}{n!}(tA)^n$.)

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

$$\Pi = \Pi(0,1) = e^{\Lambda}, \tag{6.29}$$

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. The following matrix

$$\mathcal{L}_Y = \begin{pmatrix} .000 & .000 & .000 & .000 & .000 & .000 & .000 & .000 \\ .286 & -.432 & .093 & .025 & .014 & .014 & .000 & .000 \\ .075 & .048 & -.193 & .057 & .007 & .003 & .002 & .000 \\ .027 & .014 & .118 & -.253 & .081 & .009 & .003 & .001 \\ .005 & .002 & .017 & .070 & -.171 & .071 & .005 & .001 \\ .001 & .000 & .005 & .011 & .069 & -.117 & .031 & .001 \\ .000 & .000 & .003 & .003 & .011 & .079 & -.104 & .009 \\ .000 & .000 & .000 & .003 & .002 & .008 & .102 & -.115 \end{pmatrix} \tag{6.30}$$

gives a reasonable fit to the given one year transition matrix (6.24).

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

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

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 (6.28). One special example for which this is true is if the generator can be written as

$$\Lambda(t) = B\mu(t)B^{-1}, \tag{6.31}$$

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

$$\Pi(s,t) = B\exp(\int_s^t \mu(u)du)B^{-1}. \tag{6.32}$$

[1] Recall that for any square matrix A, $B(t) = e^{tA} := \sum_{n=0}^{\infty}(n!)^{-1}(tA)^n$ is the solution of the matrix ODE $dB(t)/dt = AB(t), B(0) = I$.

As a final generalization, one should expect 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 (6.31) and taking the eigenvalues of Λ to be given as

$$\mu_j(t) = \alpha_j + \beta_j \cdot X_t, \tag{6.33}$$

for some multidimensional affine state process X_t[2]. One adopts a reduced form framework with $\mathcal{H} = \sigma(Y)$ and the market filtration assumed to satisfy the $\mathcal{H}$-condition. Then, conditional on a path for the $\mathcal{G}$ adapted process X_t, the transition probability matrix is given by (6.32). Therefore, given the value of the state process X_s, we have that this doubly stochastic transition model assigns the following probability for transition from rating i at time s to rating k at a later time t:

$$\begin{aligned} P[Y_t = k | Y_s = i, X_s] &= E_t\left[\sum_{j=0}^{K} \beta_{ijk} \exp\left(\int_s^t \mu_j(u)du\right)\right] \\ &= \sum_{j=0}^{K} \beta_{ijk} E_t\left[\exp\left(\int_s^t (\alpha_j + \beta_j \cdot X_u)du\right)\right] \\ &= \sum_{j=0}^{K} \beta_{ijk} \exp\left(A_j(s,t) + B_j(s,t) \cdot X_s\right), \end{aligned} \tag{6.34}$$

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

6.6 Exercises

Exercise 46. Consider the predefault value V_t of a defaultable coupon bond that pays at a continuous rate c until $\tau \wedge T$, plus the principal at maturity T if $\tau > T$. Find $\mathcal{G}_t$-conditional expectation formulas for V_t under the RT recovery mechanism, assuming the fractional recovery R_t is constant.

Exercise 47. Consider the three models for recovery in the case when interest rates are taken to be independent of the default intensity, and the recovery fraction R is constant. Reduce the formulas for $\bar{P}_t^{RP}(T), \bar{P}_t^{RT}(T), \bar{P}_t^{RMV}(T)$ to formulas involving $P_t(T), \bar{P}_t(T)$.

Exercise 48. Comparing the above formula for the recovery of market value mechanism to the zero-recovery formula, discuss the validity of the following statement: "In the general reduced form model, evaluation of any default risky security under the recovery of market value mechanism equals evaluation of the identical security under zero recovery, but with the stochastic hazard rate λ replaced by $(1 - R)\lambda$.

[2] This model has the technical defect that unless one is very careful in specifying X, the resulting generator Λ_t fails to satisfy the properties of a stochastic matrix.

The following exercises will be based on the model described in Exercise 35. We take the following default values of the parameters: $X \sim CIR(.5, 0.05, 1)$, $Y \sim CIR(.1, 0.1, .5)$. In addition, we assume a constant recovery fraction $R = 0.5$.

Exercise 49. Suppose a defaultable 20 year coupon bond pays half-yearly coupons at a rate chosen so that the bond trades at par on its issue date. Assuming the recovery of par mechanism, what should the value of this "defaultable par coupon rate" be? Compute the par coupon rate for a similar default free bond. Can you explain the difference? Compute the credit spread for a 20 year zero coupon defaultable bond.

Exercise 50. Compute the defaultable yield curve and credit spread curve for maturities $T \in [0, 30]$ years under the recovery of par method.

Exercise 51. Compute the defaultable yield curve and credit spread curve for maturities $T \in [0, 30]$ years under the recovery of treasury method.

Exercise 52. Compute the defaultable yield curve and credit spread curve for maturities $T \in [0, 30]$ years under the recovery of market value method.

Exercise 53. Describe and explain the differences you observe between the three recovery methods.

Exercise 54. Instead of a constant recovery, suppose that $R = R_t$ is the stochastic process given by $R_t = e^{-kr_t}$, where the constant k is chosen so that the unconditional mean of R_t is 0.5. Using the Lando formula (6.10), find an explicit formula for the price of a zero coupon bond under the recovery of market value. Use this formula to graph the defaultable yield curve and the credit spread curve, for maturities $T \in [0, 30]$ years. Compare to the previous graphs.

Exercise 55. Complete the proof of Theorem 6.1.1 by verifying all steps, and filling in the gaps.

Exercise 56. Bond insurance is a simple contract that repays the losses incurred to the insured party in the event that the reference bond defaults in the insurance period $[t, T]$. The insurer is usually a bank. Suppose a fixed premium p is paid at year end while the bond has not defaulted, and a final fractional premium payment $p(\tau - (k - 1))$ is paid at time k if default happens in the kth year. At the time of default, the insurer receives the defaulted bond and in return pays A the remaining coupons and principal. Assuming that fixed fractional recovery of treasury with $R < 1$ applies to each coupon and principal, compute the fair premium p in terms of observable bond prices.

Exercise 57. A *default put option* is a contract that pays the holder the total loss amount at the time τ of default of an underlying defaultable bond, provided $\tau \leq T$. Find explicit formulas for the fair premium at time 0 under the three different recovery mechanisms.

In the following exercises, use our standard reduced form model from Exercise 35.

Exercise 58. Show that under the recovery of treasury mechanism, and in the continuous time case, the price of the default leg equals the difference $P_0(T) - \bar{P}_0(T)$. Is this also true when the premium and default legs are paid at discrete times?

Exercise 59. Compute the CDS spread for a 10 year CDS with semiannual premiums, under the recovery of par assumption. Now compute the CDS spread for the same contract, under the recovery of treasury assumption. Assuming recovery of treasury, can you spot a relation between the CDS spread, the defaultable par coupon rate and the default-free par coupon rate?

Exercise 60. Describe and explain the differences you observe between the three recovery methods.

In the next two exercises, fix your own set of parameters for the Hull-White reduced form model in the risk neutral measure, taking $\Theta(t), \bar{\Theta}(t)$ both constants. Imagine these parameters are applicable to a BBB rated firm with a 50% recovery rate.

Exercise 61. Compute the default free and defaultable term structures over $[0, 30]$ years. (You may assume RMV).

Exercise 62. Verify equations (6.18), (6.19) in the formula for a defaultable bond in the two-factor Gaussian model.

Exercise 63. (Construction of continuous time Markov chains) Suppose the probabilities

$$\Pi_{ij}(s,t) = P[Y_t = j | Y_s = i], i, j = 0, \ldots, K$$

are well specified (i.e. $\Pi(s,t)$ is a $K+1 \times K+1$ stochastic matrix), and continuous in $s \le t$. Assume further that $\Pi(s,t) = \Pi(s,u)\Pi(u,t)$ for any $u \in [s,t]$ (and give a probabilistic interpretation of this property) and that for each s the limit

$$\Lambda(s) := \lim_{t \to s+} \frac{\Pi(s,t) - I}{t - s}$$

exists.

1. Show that $\Lambda(s)$ is a "generator" for all s.

2. Show that for each $s \le t$,

$$\lim_{h \to 0+} \frac{\Pi(s, t+h) - \Pi(s,t)}{h} = \Pi(s,t)\Lambda(t)$$

3. Assuming uniqueness of solutions of this matrix valued ODE, show that if there is a single matrix B such that for all s, $\Lambda(s) = B\mu(s)B^{-1}$ for a diagonal matrix $\mu(s)$, then

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

Exercise 64. Consider the continuous-time Markov chain model with historical generator matrix (6.30), but where the risk-neutral generator is stochastic:

$$\Lambda_t = \lambda_t \mathcal{L}_Y \tag{6.35}$$

with the scalar process λ_t a CIR(k, θ, σ) process conditionally independent of Y_t. Find explicit formulas for A_j, B_j so that the risk-neutral survival probability function for a firm with $Y_0 = i$ and initial value λ_0 is given by a sum over $k \neq 0$ of (6.34).

Chapter 7

Portfolio Credit Risk

7.1 Credit Risk Basics

7.1.1 A Single Obligor

Suppose you, as a senior analyst with a major bank, are called upon to assess a possible loan of \$100M to a long term client, a real estate developer. What are the key indicators of the quality of the loan? Ultimately one wishes to know that the interest charged on the loan reflects the risk involved, and even more importantly, that the risk profile of the loan is compatible with your existing credit portfolio.

You are concerned with the possibility of a large loss in the event that the client defaults over the next year. Let $\tilde{L}$ denote the loss random variable over one year. It depends fundamentally on three factors: default or not, the severity of the loss given default, and the exposure at default. We write:

- D, the event that the client defaults within one year. The "historical" or "physical" default probability is $\mathrm{DP} = E[\mathbf{1}_D]$;
- "severity" SEV denotes the random fraction of the total notional that is lost, given that D occurs; $\mathrm{LGD} = E[\mathrm{SEV}]$ is called "expected loss given default";
- "exposure at default" EAD (usually taken to be deterministic, but in reality it is stochastic too).

The fundamental relation is

$$\tilde{L} = \mathrm{EAD} \times \mathrm{SEV} \times \mathbf{1}_D. \tag{7.1}$$

Assuming independence (which of course is not really true), one has the expected loss

$$\mathrm{EL} := E[\tilde{L}] = \mathrm{EAD} \times \mathrm{LGD} \times \mathrm{DP}$$

Estimation of these quantities is of course a delicate art. DP may be obtained from rating class statistics produced by ratings agencies such as S & P, Moody's or Fitch.

Better yet, one should use a more dynamic method such as the Moodys-KMV distance-to-default technology, leading to DP = EDF, i.e. expected default frequency. For the retail sector (small loans), one uses "credit scoring" of some sort. As we have discussed in the context of modeling recovery in the reduced form setting, SEV $= 1 - R$ depends on such attributes as the seniority of the loan, the quality of the collateral, and the state of the economy at the time of default. EAD is in general also subject to uncertainty: for example, the obligor might have prior loans that should be accounted for. Also important are lines of credit: prior to default, the obligor is most likely to "max" out their lines of credit leading to a higher than expected EAD, and the prudent risk analyst must be receptive to such signals. In simple situations, however, we might take EAD $= \mathcal{N}$, the notional amount of the loan.

As a risk analyst, one is interested in more than simply the expected loss EL on the loan. For example, the standard deviation of the loss, assuming independence between SEV and D, is

$$\mathrm{UL} := \sqrt{\mathrm{Var}[\tilde{L}]} = \mathrm{EAD} \times \sqrt{\mathrm{Var}[\mathrm{SEV}] \times \mathrm{DP} + \mathrm{LGD}^2 \times \mathrm{DP} \times (1 - \mathrm{DP})}$$

Of course, the standard deviation of $\tilde{L}$ is not a good indicator of the "value at risk" VaR, but it nonetheless is treated by analysts as the definition of *unexpected loss* UL.

7.1.2 Multiple Obligors

A loan portfolio will be a sum of many loans to a total of I different clients. We then need to compute the probabilistic nature of the total portfolio loss over one year. We suppose that the loans to the ith obligor are treated as one single loan, and proceed to combine them. The total one year portfolio loss random variable will be

$$\tilde{L}_{PF} = \sum_{i=1}^{I} \mathrm{EAD}_i \times \mathrm{SEV}_i \times \mathbf{1}_{D_i} \tag{7.2}$$

By the additivity of expectations, the expected loss is simply $\mathrm{EL}_{PF} = \sum_i \mathrm{EL}_i$. The *unexpected loss* $\mathrm{UL}_{PF} := \mathrm{Var}[\tilde{L}_{PF}]$ and the portfolio *value-at-risk* however, depend crucially on the *default correlations* between different obligors, and is the central difficulty in portfolio credit risk management.

We define the one-year default correlation between obligors $i \neq j$ to be

$$\rho_{ij} = \frac{E[\mathbf{1}_{D_i}\mathbf{1}_{D_j}] - \mathrm{DP}_i\mathrm{DP}_j}{\sqrt{\mathrm{DP}_i(1 - \mathrm{DP}_i)\mathrm{DP}_j(1 - \mathrm{DP}_j)}} \tag{7.3}$$

If one assumes that each SEV_i is deterministic, one has

$$\mathrm{UL}_{PF}^2 := \mathrm{Var}[\tilde{L}_{PF}] = \sum_{i,j}^{I} \mathrm{EAD}_i\mathrm{EAD}_j\mathrm{LGD}_i\mathrm{LGD}_j\rho_{ij}\mathrm{DP}_i \times (1 - \mathrm{DP}_i)\mathrm{DP}_j \times (1 - \mathrm{DP}_j) \tag{7.4}$$

Recall that the one-year portfolio value-at-risk VaR_{PF} with confidence level $\alpha < 1$ (typical values are $\alpha = 0.99$ or $\alpha = .9998$) is the α-quantile of the loss random variable $\tilde{L}_{PF}$:

$$\mathrm{VaR}_\alpha := \inf\{x|P[\tilde{L}_{PF} \leq x] > \alpha\} \tag{7.5}$$

This is highly dependent on the joint default characteristics of the I obligors, and determining it is computationally intensive. Ultimately one needs large scale Monte Carlo simulations. To get a feel for such risk numbers, one can try to match the loss distribution to a specific parametric family of distributions: here the beta distribution is a favorite choice.

7.1.3 Economic Capital

The most common way to quantify economic capital EC is by value at risk. For example, the Basel II Accord specifies that for credit portfolio risk,

$$\mathrm{EC} = \mathrm{VaR}_{0.9998} - \mathrm{EL}$$

where $\mathrm{VaR}_{0.9998}$ is computed within the so-called "internal ratings" framework. This means the bank is approved by the regulator to implement their own "in-house" risk management system.

Of course, from a statistical point of view, $\mathrm{VaR}_{0.9998}$ is *unreasonably* sensitive to assumptions about the tail of the loss distribution. Typically risk managers compute a VaR value with a reasonable assumption on the tail distribution and a reduced confidence level, and quote the result as a $\mathrm{VaR}_{0.9998}$ coming from a normal distribution when communicating with regulators and senior executives.

The rationale for reducing $\mathrm{VaR}_{0.9998}$ by the amount EL is that EL is typically charged to the client, along with management fees and the like. Also, EL is a simple sum of terms, unlike $\mathrm{VaR}_{0.9998}$ which is heavily sensitive to the treatment of default dependency.

7.1.4 Market Risk

In the above discussion, we have imagined that the market value of the loan in question in one year, if no default occurs, will be the full notional amount $\mathrm{EAD} = \mathcal{N}$. In reality, risk management requires estimation of the marked-to-market value of all securities. If the loan is a corporate bond, it will valued at any future date according the credit spreads prevailing on that date; the associated risk is sometimes called "spread risk", or if losses arise from rating transitions "migration risk". Industry practice is to separate $\tilde{L}$ into market risk and credit risk components. However, the theory we have developed in this course unifies market and credit risk, provided we value assets, including any loan in default, on a mark-to-market basis.

7.2 Modeling Dependent Defaults

In the remainder of this chapter we revert to our notation for default events, recovery and consider portfolios of corporate bonds. With some slight change of perspective, everything we discuss applies to more general loan portfolios involving consumer loans, credit cards, and mortgages.

We see that the central difficulty in credit risk management is to determine the joint probability of the defaults of many firms. The most dangerous scenarios that can affect a well-diversified portfolio of corporate bonds are typically those events one can call "correlated defaults".

We define two measures of default dependence

Definition 8. 1. The *default correlation function* $\rho(t)$ between two firms is

$$\rho(t) = \frac{P[\tau_1 \leq t, \tau_2 \leq t] - P[\tau_1 \leq t]P[\tau_2 \leq t]}{\sqrt{P[\tau_1 \leq t](1 - P[\tau_1 \leq t])P[\tau_2 \leq t](1 - P[\tau_2 \leq t])}} \tag{7.6}$$

It measures the degree of dependence between defaults of two firms over the period $[0, t]$.

2. The *default time correlation* between τ_1, τ_2 is:

$$\rho_{\tau_1,\tau_2} = \frac{E[\tau_1 \tau_2] - E[\tau_1]E[\tau_2]}{\sqrt{\mathrm{Var}[\tau_1]\mathrm{Var}[\tau_2]}} \tag{7.7}$$

Typical values of the one-year default correlation $\rho(1)$ between speculative grade firms (i.e below BBB rated) within an industry are 1% (technology) to 7% (banking).

In this chapter we initiate the study of multifirm defaults, and investigate three leading approaches to introducing dependence between credit events associated with different firms. In the following, we consider a credit portfolio consisting of exposure to I obligors, each labeled by an index $i \in \{1, \dots, I\}$.

7.3 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

$$P[X \leq x] = \sum_{m=0}^{n} \frac{I!}{m!(I-m)!} p^m (1-p)^{I-m}, \tag{7.8}$$

where $nfloor(x/(1-R)L)$ 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, Moodys 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

$$P[X \leq x] = \sum_{m=0}^{n} \frac{D!}{m!(D-m)!} p^m (1-p)^{D-m}, \tag{7.9}$$

where $n = floor(xD/(1-R)LI)$.

This method is not based on any theoretical model, and the parameter D used by Moodys is determined by a complicated recipe. 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.

Warning: This method seriously underestimates "tail risk" of large losses!

7.4 Default Correlation in Structural Models

Default correlation can be incorporated quite naturally into structural models by making the asset value dynamics of different firms correlated through time. We illustrate the main ideas with the example of two firms whose assets follow the dynamics

$$dA_t^i = A_t^i[(\mu^i - \delta^i)dt + \sigma^i dW_t^i], A_0^i > 0, i = 1, 2 \tag{7.10}$$

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 is

$$\begin{aligned} p(T_1, T_2) &= P[A_{T_1}^1 < K^1, A_{T_2}^2 < K^2] \\ &= \Phi_2\left(\frac{\rho T_1 \wedge T_2}{\sqrt{T_1 T_2}}, \frac{\log(K^1/A_0^1) - m^1 T_1}{\sigma^1 \sqrt{T_1}}, \frac{\log(K^2/A_0^2) - m^2 T_2}{\sigma^2 \sqrt{T_2}}\right), \end{aligned} \tag{7.11}$$

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

$$\Phi_2(\rho, a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)}\right) dxdy. \tag{7.12}$$

7.4.1 CreditMetrics Approach

CreditMetrics, an industry method developed by Moody's and described in [9], extends the Merton model by assuming that the credit rating of a firm is its basic measure of credit quality, and that changes in credit rating can be attributed to changes in the firm's asset value. One chooses a fixed time duration Δt, and a specific rating transition matrix $\Pi := \Pi_{\Delta t}$ similar in form to (6.29). Then for each possible initial rating class $k = 1, \dots, K$ one computes a set of marks $d_0^k = -\infty < d_1^k < \cdots < d_{K+1}^k := \infty$ such that the final rating is determined by the draw of a standard normal variate X^i:

$$\Pi_{kj} := P[Y_{\Delta t}^i = j | Y_0^i = k] = P[d_j^k < X^i \le d_{j+1}^k] = \Phi(d_{j+1}^k) - \Phi(d_j^k) \tag{7.13}$$

Roughly this can be interpreted by relating the asset return over the period to X^i:

$$\log(A_{\Delta t}^i / A_0^i) = (\mu^i - \delta^i - (\sigma^i)^2/2)\Delta t + \sigma^i \sqrt{\Delta t} X^i. \tag{7.14}$$

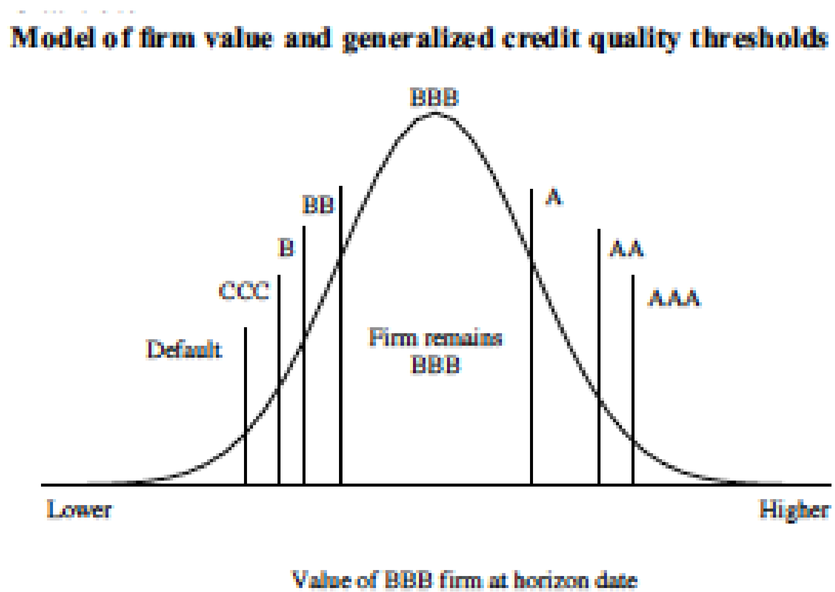

Figure 7.1: CreditMetrics mapping of the transition matrix to the normal distribution.

In the multifirm setting, one treats the vector of asset returns $(X^1, X^2, \dots, X^I)$ as a correlated multivariate normal $N(0, \Sigma)$ with $N(0, 1)$ marginals, and a specified correlation matrix Σ (a positive definite matrix with 1 in each diagonal entry). Several alternatives exist for determining Σ. One can assume a strong correlation between A^i and the firm's market capitalization (stock price times the number of shares), and thus simply take Σ to be exactly the correlation matrix determined by historically observed stock returns. The CreditMetrics people do this somewhat more precisely, by extending the method described in Section 4.3 to back out the asset return correlation matrix Σ from the observations of the multivariate stock returns.

Remark 9. The CreditMetrics method is widely used, but should be viewed with caution for credit risk management. Its weakness is of course its reliance on the multivariate normal distribution: the kinds of default dependence that it produces are of the tame "thin tailed" type that have proved to be utterly useless in explaining the 2007-2008 credit crunch.

7.4.2 First Passage Models

To illustrate this approach, we consider a two-firm 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

$$p(T_1, T_2) = \Psi_2\left(\frac{\rho T_1 \wedge T_2}{\sqrt{T_1 T_2}}, T_1, T_2, \log\frac{D_1}{A_0^1}, \log\frac{D_2}{A_0^2}\right), \tag{7.15}$$

where $\Psi_2(\rho, \cdot, \cdot, a, b)$ is the bivariate inverse Gaussian distribution function with correlation ρ and parameters a, b [26]. A similar result exists for an exponentially growing time-dependent barrier in terms of modified Bessel functions.

No general analytical first passage formulas are known for three or more firms, and one must resort to Monte Carlo simulations to compute.

7.4.3 Factor Models of Correlation

In all of these approaches, we can obtain a full range of Gaussian 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 $I \times I$ correlation 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*, also known as *Bernoulli mixture models* (see [18]). For example, in the CreditMetrics framework, a *one-factor model* corresponds to assuming that $X^i = (\sigma^i\sqrt{\Delta t})^{-1}(\log(A^i_{\Delta t}/A^i_0) - (\mu^i - \delta^i - (\sigma^i)^2/2)\Delta t) \sim N(0,1)$ is such that

$$X^i = \beta^i Z + \bar{\beta}^i \epsilon^i, \quad \bar{\beta}^i := \sqrt{1 - (\beta^i)^2} \tag{7.16}$$

where Z and ϵ^i are independent standard normal random variables. This has the interpretation that asset values for all firms are driven by one common factor Z plus firm specific (idiosyncratic) factors ϵ^i. Then the entire dependence structure is reduced to specifying the correlation parameters β^i. One can attempt to connect Z to some observed macroeconomic factor such as an economic activity indicator like GDP: there have been many studies of this type, for example [4] and the references therein.

To further investigate the implications of the one-factor model, let us assume that the credit state of individual firms depends only on their rating, so that $\beta^i = \beta$ for all firms.

Then the two-firm transition probabilities are computed by

$$\begin{aligned}
P[Y^1_{\Delta t} = j_1, Y^2_{\Delta t} = j_2 | Y^1_0 = k_1, Y^2_0 = k_2] \quad &:= \quad \Pi^{(2)}(j_1, j_2; k_1, k_2) \\
&= \int_{\mathbb{R}} \left[\Phi\Big(\frac{d^{k_1}_{j_1+1} - \beta z}{\bar\beta}\Big) - \Phi\Big(\frac{d^{k_1}_{j_1} - \beta z}{\bar\beta}\Big)\right] \left[\Phi\Big(\frac{d^{k_2}_{j_2+1} - \beta z}{\bar\beta}\Big) - \Phi\Big(\frac{d^{k_2}_{j_2} - \beta z}{\bar\beta}\Big)\right] \phi(z) dz \\
&= \Phi_2(\beta^2, d^{k_1}_{j_1+1}, d^{k_2}_{j_2+1}) - \Phi_2(\beta^2, d^{k_1}_{j_1}, d^{k_2}_{j_2+1}) - \Phi_2(\beta^2, d^{k_1}_{j_1+1}, d^{k_2}_{j_2}) + \Phi_2(\beta^2, d^{k_1}_{j_1}, d^{k_2}_{j_2}) \quad (7.17)
\end{aligned}$$

where Φ_2 is the bivariate standard normal distribution function.

The probability of observing up to n defaults in a set of I firms in rating k in time ΔT is

$$\sum_{m=0}^{n} \int_{-\infty}^{\infty} \frac{I!}{m!(I-m)!} p_k(z)^m (1 - p_k(z))^{I-m} \phi(z) dz, \tag{7.18}$$

where ϕ is the standard normal density and the probability of one firm's default conditional on $Z = z$ is

$$p_k(z) = \Phi\Big(\frac{d^k_1 - \beta z}{\bar\beta}\Big). \tag{7.19}$$

Most generally, one can determine the probability that exactly $n_1, \dots, n_K$ defaults occur from each of the K rating classes, given that there are initially $I^1, \dots, I^K$ in each rating class ($I = \sum_k I^k$). In the exercises we will show that

$$P[N_1 = n_1, \dots, N_K = n_K] = \int_{\mathbb{R}} \prod_{k=1}^{K} \left[\binom{I_k}{n_k} p_k(z)^{n_k} (1 - p_k(z))^{M_k - n_k} \right] \phi(z) dz \tag{7.20}$$

We can obtain more explicit results than (7.18) in the *large portfolio* approximation, that is, assuming that $I \to \infty$. Then, by the Law of Large Numbers, $p_k(z)$ given by (7.19) represents the conditional average of the fraction of a portfolio of I firms all with rating k that experiences default over the period $[0, \Delta T]$. If we normalize the exposure $L^i = 1$ and assume zero-recovery, then this fraction is exactly the portfolio default loss per firm. Its distribution is then

$$\begin{aligned}
F(x) \quad &:= \quad P[L \le x] = E[P[X \le x | Z]] \\
&= \int_{-\infty}^{\infty} P[X \le x | Z = z] \phi(z) dz \\
&= \int_{-\infty}^{\infty} \mathbf{1}_{\{p_k(z) \le x\}} \phi(z) dz \\
&= \Phi\left(\frac{d^k_1 - \bar\beta \Phi^{-1}(x)}{\beta} \right). \qquad (7.21)
\end{aligned}$$

A more accurate large portfolio approximation is obtained by invoking the Central Limit Theorem instead (in fact the DeMoivre-Laplace Limit Theorem for binomial random variables), which implies that the conditional distribution of $L = I^{-1} \sum_i \mathbf{1}_{Y^i_{\Delta t} = 0} \simeq$

$N(p_k(z), I^{-1}p_k(z)(1-p_k(z)))$, which leads to

$$P[L \leq x] \simeq \int_{-\infty}^{\infty} \Phi\left(\frac{x-p_k(z)}{\sqrt{p_k(z)(1-p_k(z))/I}}\right)\phi(z)dz \tag{7.22}$$

7.5 Default correlation in reduced-form models

Recall that in the reduced form models of Section 5.1, 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

$$N_t = \sum_{i=1}^{I} N_t^i, \tag{7.23}$$

then it follows that

$$\Lambda_t = \sum_{i=1}^{I}\int_0^t \lambda_s^i ds \tag{7.24}$$

is the compensator for N_t. However, it is not true in general that N_t is itself a counting process, since its jumps can have a magnitude greater than one in the presence of simultaneous defaults. 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.

7.5.1 Doubly stochastic models

As we have seen, the doubly stochastic assumption of Section 5.1 needs to be modified in the presence of several default events. Now, a *doubly stochastic model of correlated defaults* means a model whereby, conditioned on the background filtration $(\mathcal{G})_t$, 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 intensityλ_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 difficulties in obtaining realistic correlations for default events.

As an example, consider the joint default probability for obligors i, j for a fixed time horizon T. That is

$$\begin{aligned} p_{ij} &:= P[\tau_i \leq T, \tau_j \leq T] \\ &= E[\mathbf{1}_{\{\tau_i \leq T\}}\mathbf{1}_{\{\tau_j \leq T\}}] = E[E[\mathbf{1}_{\{\tau_i \leq T\}}\mathbf{1}_{\{\tau_j \leq T\}}|\mathcal{G}_T]] \\ &= E[(1-e^{-\int_0^T \lambda_s^i ds})(1-e^{-\int_0^T \lambda_s^j ds})] \qquad (7.25) \\ &= p_i + p_j + E[e^{-\int_0^T (\lambda_s^i + \lambda_s^j)ds}] - 1, \qquad (7.26) \end{aligned}$$

where $p_i = 1 - E[e^{-\int_0^T \lambda_s^i ds}]$. This achieves its maximum value for perfectly correlated intensities, that is $\lambda^i = \lambda^j$ in which case the correlation between default events is

$$\begin{aligned} \rho &= \frac{p_{ij} - p_i^2}{p_i(1-p_i)} = \frac{2p_i + E[e^{-2\int_0^T \lambda_s^i ds}] - 1 - p_i^2}{p_i(1-p_i)} \\ &= \frac{\mathrm{Var}[e^{-\int_0^T \lambda_s^i ds}]}{p_i(1-p_i)}. \end{aligned} \tag{7.27}$$

To estimate the order of magnitude of this correlation for intensities given by diffusion processes, let us assume that the integral $\int_0^T \lambda_s^i ds$ is normally distributed with mean μ and variance σ^2, both small. We then obtain that

$$\rho = \frac{1-p}{p}(e^{\sigma^2} - 1) \sim \frac{\sigma^2}{\mu}. \tag{7.28}$$

Usually this is too low to account for empirical data for correlated defaults. The only way to remedy this situation within the reduced form setting is to allow for joint large jumps in the intensity processes.

Consider the model of Section 5.5.2 with mean-reverting intensities with exponentially distributed jumps. Let us suppose for simplicity that all firms have identical characteristics, that is, the parameters k, θ, J, c are firm independent. One can create the possibility of strong default correlations by letting some jumps affect all firms, while other jumps are firm-specific. Thus we write

$$d\lambda_t^i = k(\theta - \lambda_t^i)dt + dZ^i + dZ \tag{7.29}$$

where the collection $\{Z^i, Z\}$ are independent compound Poisson processes with identically distributed jump sizes, but arriving at rates $c(1-\rho)$ and ρc respectively. In this picture, jumps that negatively affect all firms simultaneously arrive at rate ρc, while the idiosyncratic jumps arrive at the rate consistent with the total arrival rate of jumps to be c.

More generally, one can imagine that the common adverse changes in credit quality that arise from Z affect only a random collection of firms each time. One simple way to realize this idea is by supposing that a common jump affects a given firm with probability p. To be consistent with the previous idea, one changes the rate of Z to $\rho c/p$ and at the time of each jump of Z one introduces a set of independent $Bern(p)$ random variables X^i.Then the SDE for intensities becomes

$$d\lambda_t^i = k(\theta - \lambda_t^i)dt + dZ^i + X^i dZ \tag{7.30}$$

As p tends to 0, the model becomes one of independent defaults, but for $p = 1, \rho = 1$ defaults are not perfectly correlated. While in this case, the firms have identical intensities, conditional on the intensities defaults are independent.

7.5.2 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 process $\widehat{N}_t^j$ 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

$$\begin{aligned} \lambda_A &= \widehat{\lambda}_A + \widehat{\lambda}_{AB} \\ \lambda_B &= \widehat{\lambda}_B + \widehat{\lambda}_{AB}. \end{aligned} \tag{7.31}$$

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.

7.5.3 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.

7.6 Definition of Copula Functions

In all default correlation models presented so far, information about the credit quality of individual firms comes intertwined with the default dependence. Copula models of default have arisen as a way of separating the dependence structure of correlated defaults from 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

$$P[U \leq u] = P[F(\tau) \leq u] = P[\tau \leq F^{-1}(u)] = F(F^{-1}(u)) = u. \tag{7.32}$$

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

$$C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i, \text{for all } i = 1, \ldots, I, \tag{7.33}$$

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

The fundamental existence result concerning copula functions is

Theorem 7.6.1 (Sklar's theorem)**.** *Given any set of continuous random variables $\tau_1, \ldots, \tau_I$ with marginal distribution functions $F_1, \ldots, F_I$, there exists a unique copula function C such that their joint distribution*

$$F(t_1, \ldots, t_I) := P[\tau_1 \leq t_1, \ldots, \tau_n \leq t_n] \tag{7.34}$$

can be written as

$$F(t_1, \ldots, t_I) = C(F_1(t_1), \ldots, F_I(t_I)). \tag{7.35}$$

In other words, the copula function C in (7.35) is the joint distribution for the uniformly distributed random variables $F_1(\tau_1), \ldots, F_I(\tau_I)$. At first sight this theorem seems 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 $F_1, \ldots, F_I$.

Note: If we map $(\tau_1, \ldots, \tau_I) \to (Y_1, \ldots, Y_I)$ by $Y_i = h_i(\tau_i)$ for any continuous strictly increasing functions h_i, then $(\tau_1, \ldots, \tau_I)$ and $(Y_1, \ldots, Y_I)$ will have the identical copula.

7.6.1 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

$$C(u, v) = P[U \leq u, V \leq v] = P[U \leq u]P[V \leq v] = uv. \tag{7.36}$$

Conversely, a product copula leads to

$$\begin{aligned} F(t_1, t_2) &= P[\tau_1 \leq t_1, \tau_2 \leq t_2] = C(F_1(t_1), F_2(t_2)) = F_1(t_1)F_2(t_2) \\ &= P[\tau_1 \leq t_1]P[\tau_2 \leq t_2], \end{aligned} \tag{7.37}$$

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

$$C(u, v) = P[U \leq u, U \leq v] = \min(u, v). \tag{7.38}$$

Conversely, this copula leads to perfectly correlated defaults, since

$$\begin{aligned} P[\tau_1 \leq t_1, \tau_2 \leq t_2] &= C(F_1(t_1), F_2(t_2)) = \min(F_1(t_1), F_2(t_2)) \\ &= \min(P[\tau_1 \leq t_1, P[\tau_2 \leq t_2]), \end{aligned} \tag{7.39}$$

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

$$C(u, v) = P[U \leq u, 1 - U \leq v] = \max(u + v - 1, 0). \tag{7.40}$$

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

$$\begin{aligned} P[\tau_1 \leq t_1, \tau_2 \leq t_2] &= C(F_1(t_1), F_2(t_2)) = \max(F_1(t_1) + F_2(t_2) - 1, 0) \\ &= \max(P[\tau_1 \leq t_1 + P[\tau_2 \leq t_2] - 1, 0), \end{aligned} \tag{7.41}$$

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

$$\tau_2 = F_2^{-1}(1 - F_1(\tau_1)), \tag{7.42}$$

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

In general, the Fréchet bounds state that a two dimensional copula function is bounded by these two limiting cases, that is

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v). \tag{7.43}$$

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

7.6.2 Normal Copula and Student t-Copula

For another example of copula, consider the joint probability (7.4) for default times derived in Section 7.11 for the Merton structural model. From Sklar's theorem, we can rewrite it as

$$\begin{aligned} p(T_1,T_2) &= \Phi_2\left(\frac{\rho T_1\wedge T_2}{\sqrt{T_1T_2}}, \frac{\log(K^1/A^1_0)-m^1T_1}{\sigma^1\sqrt{T_1}}, \frac{\log(K^2/A^2_0)-m^2T_2}{\sigma^2\sqrt{T_2}}\right) \\ &= C^G_{\tilde\rho}(P[\tau_1\le T_1],P[\tau_2\le T_2]), \qquad (7.44) \\ \tilde\rho &= \frac{\rho T_1\wedge T_2}{\sqrt{T_1T_2}} \qquad (7.45) \end{aligned}$$

where $C^G_\rho(\cdot,\cdot)$ is the *normal copula* with correlation ρ:

$$C^G_\rho = \Phi_2(\rho, \Phi^{-1}(u), \Phi^{-1}(v)). \tag{7.46}$$

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

A *multidimensional normal copula* is defined analogously. We start with normally distributed random variables $[X_1,\ldots,X_I]$ with zero mean, unit variances and correlation matrix Σ. Then define the uniform random variables

$$U_i = \Phi\left(X_i-\mu_i\right), \tag{7.47}$$

and define the normal copula C_Σ to be their joint distribution. The joint distribution of $(\tau_1,\ldots,\tau_I)$ with marginals $F_1(t_1),\ldots,F_I(t_I)$ and the normal copula with correlation matrix Σ follows by taking

$$[\tau_1,\ldots,\tau_I] = [F_1^{-1}(\Phi(X_1)),\ldots,F_I^{-1}(\Phi(X_I))] \tag{7.48}$$

Consider now a χ^2 random variable Y with ν degrees of freedom and suppose that Y is independent of $X_1,\ldots,X_I$. Then each univariate random variable $T_i=\sqrt{\frac{\nu}{Y}}X_i$ has a Student t distribution with mean 0, variance 1 and ν degrees of freedom, with CDF t_ν. The $\mathbb{R}^I$ valued random variable $[T_1,\ldots,T_I]$ is then said to have a multivariate Student t distribution with ν degrees of freedom and correlation Σ. If we now define the uniform random variables

$$U_i = t_\nu\left(\sqrt{\frac{\nu}{Y}}X_i\right), \tag{7.49}$$

then their joint distribution function is called a *t-copula*.

7.6.3 Tail dependence

The main difference between a normal copula and a t-copula is the weight they 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*

$$\lim_{u\to 0}\frac{C(u,u)}{u} \tag{7.50}$$

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

$$\lim_{u\to 1}\frac{1+C(u,u)-2u}{1-u} \tag{7.51}$$

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 normal copula is tail independent, while the t-copula has both upper and lower dependence.

7.6.4 Archimedean copulas

Clearly the set of I-dimensional copula functions has way too much freedom for modeling choices. In practice, one tries to concentrate on parametric families of copulas, such as the normal 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 exists a decreasing, convex function $\psi : [0,1] \to \mathbb{R}^+$ with $\psi(1) = 0$ and $\psi(0) = \infty$ such that

$$C(t_1,\ldots,t_I) = \psi^{-1}(\psi(t_1)+\cdots+\psi(t_I)). \tag{7.52}$$

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 exchangeable, 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.

To construct the copula $C(t_1,\ldots,t_I)$ as the joint cumulative distribution function of random variables $\{U_j\}_{j=1}^I$, we write $U_j = \psi^{-1}(Z_j)$. Here $Z_j = -Y^{-1}\log Y_j$ where $\{Y, Y_1,\ldots,Y_M\}$ are mutually independent random variables. Y takes values in $\mathbb{R}^+$ and has distribution function $F(y)$ and Laplace transform $\Phi_Y(a) = \psi^{-1}(a)$. Each Y_j has a uniform $[0,1]$ distribution. One can then check that the marginals of U_j are uniform:

$$\begin{aligned} P[U_j \le u] &= \int_0^\infty P[Z_j \ge \psi(u)|Y=y]dF(y) \\ &= \int_0^\infty P[Y_j \le \exp(-y\psi(u))]dF(y) \\ &= \int_0^\infty \exp(-y\psi(u))dF(y) \\ &= \psi^{-1}(\psi(u)) = u. \end{aligned} \tag{7.53}$$

One can also check that the copula has the "arithmetic" form:

$$\begin{aligned} C(u_1,\ldots,u_I) &= \int_0^\infty \prod_j \left(P[Z_j \ge \psi(u_j)|Y=y]\right) dF(y) \\ &= \int_0^\infty \prod_j \left(P[Y_j \le \exp(-y\psi(u_j))]\right) dF(y) \\ &= \int_0^\infty \exp\bigl(-y\sum_j \psi(u_j)\bigr) dF(y) \\ &= \psi^{-1}\left(\psi(u_1)+\cdots+\psi(u_M)\right). \end{aligned} \tag{7.54}$$

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

$$\psi(t) = t^{-\theta} - 1, \theta > 0, \tag{7.55}$$

from which we can calculate both the lower tail dependence coefficient

$$\lim_{u\to 0} \frac{C(u,u)}{u} = \frac{1}{2^{1/\theta}} > 0 \tag{7.56}$$

and the upper tail dependence coefficient

$$\lim_{u\to 0} \frac{1+C(u,u)-2u}{1-u} = 0 \tag{7.57}$$

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

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

$$\psi(t) = (-\log t)^\theta, \theta \ge 1. \tag{7.58}$$

We then have

$$\lim_{u\to 0} \frac{C(u,u)}{u} = 0 \tag{7.59}$$

and the upper tail dependence coefficient

$$\lim_{u\to 0} \frac{1+C(u,u)-2u}{1-u} = 2 - 2^{1/\theta} > 0 \tag{7.60}$$

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

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

$$\psi(t) = -\log\left(\frac{e^{\theta t}-1}{e^{\theta}-1}\right), \quad \theta \neq 0. \tag{7.61}$$

We find that

$$\lim_{u\to 0}\frac{C(u,u)}{u} = 0 \tag{7.62}$$

and

$$\lim_{u\to 0}\frac{1+C(u,u)-2u}{1-u} = 0 \tag{7.63}$$

so that the Frank copula is tail independent.

7.6.5 One factor normal copula default model

In this model, the joint distribution of default times $\{\tau_j\}$ is specified by an arbitrary choice of marginal cumulative distribution function $F_j(t) := E[I(\tau_j \le t)]$ for each firm and the selection of a one factor normal copula to describe the correlation structure. The default times $\{\tau_j\}$ are defined in terms of a multidimensional normal random variable $\vec{Z} = (Z_1,\dots,Z_I)$ by $Z_j = H_j(\tau_j),\ j = 1,2,\dots,I$. The marginals of Z_j are assumed to be standard normals, and $H_j = \Phi^{-1}\circ F_j$ where Φ^{-1} denotes the inverse cumulative distribution function of the standard normal random variable. This guarantees that the marginal CDF of τ_j is F_j.

The joint distribution function of default is given by

$$P[\tau_j \le t_j, j=1,\dots,I] = P\big[Z_j \le H_j(t_j), j=1,\dots,I\big].$$

The one factor normal copula arises by taking $\vec{Z}$ to have mean zero and covariance matrix

$$E[Z_iZ_j] = \begin{cases} 1 & i=j \\ a_ia_j & i\neq j \end{cases}$$

with the values a_j in $[-1,1]$.

If we let $\{Y,Y_1,\dots,Y_I\}$ be iid standard normal random variables, then the random variables $Z_j = a_jY + \sqrt{1-a_j^2}Y_j, j=1,\dots,I$ have the required joint distribution. Then, since the Z's are independent conditionally on Y, one can compute that:

$$\begin{aligned} P[\tau_j \le t_j, j=1,\dots,I] &= \int_{\mathbb{R}} P\big[a_jY + \sqrt{1-a_j^2}Y_j \le z_j, j=1,\dots,I \mid Y=y\big]\phi(y)dy && (7.64)\\ &= \int_{\mathbb{R}}\prod_j P\big[Y_j \le (z_j - a_jy)/\sqrt{1-a_j^2}\big]\phi(y)dy \\ &= \int_{\mathbb{R}}\prod_j \Phi\Big((H_j(t_j)-a_jy)/\sqrt{1-a_j^2}\Big)\phi(y)dy, && (7.65)\end{aligned}$$

Here Φ,ϕ are the CDF and PDF respectively of the standard normal random variable.

7.6.6 An example of default modeling with copulas [NOT COMPLETE]

Let us start with a static model, that is, one for default or survival over a fixed time interval $[0, T]$. As before, consider I obligors and denote by p_i the probability that the ith firm 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

$$U_i \leq 1 - p_i. \tag{7.66}$$

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

$$u_i = \begin{cases} 1 - p_i & \text{if } i \in I_S \\ 1 & \text{otherwise} \end{cases} \tag{7.67}$$

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

$$P(I_S, (I_D \cup \{j\})) = P(I_S, I_D) - P(I_S \cup \{j\}, I_D). \tag{7.68}$$

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 then 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

$$P_i(0, \tau_i) = U_i. \tag{7.69}$$

[NOT COMPLETE]

7.7 Exercises

Exercise 65. Let two firms be treated by the Merton model with $T = 1$, where in addition to the usual parameters in the physical measure, the underlying Brownian motions have correlation

$$dW_t^1 dW_t^2 = \rho dt.$$

Find a formula for the one year default correlation $\rho(1)$ as a function of the underlying parameters. Take parameters $\mu = 0.05, \delta = 0, \sigma = 0.2, A_0/K = 1.2$ and plot this as a function of $\rho \in [-1, 1]$.

Exercise 66. Suppose two firms are modelled in the reduced form method with physical intensities λ_t^i governed by Vasicek dynamics. In case A $\lambda_t^1 = \lambda_t^2$, while in case B λ_t^1, λ_t^2 are identical in distribution, but independent. For case A, derive a formula for the default correlation function $\rho(t)$. Using MATLAB, plot ρ versus $t \in [0, 30]$ (take parameters $(\lambda_0 = 0.01, k = 1, \theta = 0.01, \sigma = 0.002)$. Now extend to I obligors, and compute the value of UL for one year, as a function of the total number of obligors, in both case A and B. Assume $\mathrm{SEV}_i = .5$, and that each loan has the same notional and the total loan portfolio has $\mathcal{N} = 1$. Plot UL for both cases for $i = 1, 2, \ldots, 100$.

Exercise 67. Follow the CreditMetrics method to compute the marks d_j^k that fit the one year rating transition matrix (6.24).

Exercise 68. In this problem, use the definitions found in Section 7.6.2.

1. Compute the moment generating function $M(u) := E[e^{uY}]$ of the χ^2 random variable Y

2. Use this result to compute the PDF of the I-dimensional Student t random variable $T = [T_1, \ldots, T_I]$ with $\nu > 0$ degrees of freedom and correlation matrix Σ.

3. Compute the marginal PDF for T_1.

4. Write a Monte Carlo simulator that generates a Student t random vector for any input values of ν, Σ, I. Produce a scatterplot of 1000 points for $\nu = 6$, correlation $\rho = .3$ and $I = 2$.

Exercise 69. Prove (7.20).

Chapter 8

CDOs and other basket credit products

We conclude this course with several examples of basket credit derivatives and their prices under the different models introduced so far. Many of these derivatives generalize the CDS to multiple firms, and have the following fundamental structure: a protection buyer A, a protection seller B and a set of reference obligors $C_1, \ldots, C_I$ (or securities issued by these obligors). Their purpose is to transfer all or part of the obligors default risk from A to B. In return, B receives some financial compensation from A, be it in the form of an upfront fee, a payment stream or the possibility of high returns at the end of the contract.

8.1 kth to default derivatives

In this class of products, the reference securities are usually corporate bonds, and the events that trigger insurance payments are their defaults. They are a cheaper and less cumbersome alternative than buying default insurance on each obligor separately. The so-called "first-to-default" contract makes a single payment when the first of I bonds defaults, while the "mth-to-default" pays only when the mth of the set of reference obligors defaults. These contracts may be too little insurance, so a more general product called the "first-m-to-default" swaps makes payments on each of the first m defaults. Finally, one can find products that pay for defaults in a "layer" or "tranche", that we can call the "m-to-n" default swap.

8.1.1 First to Default Swap

The time of the first default τ^* is given by

$$\tau^* = \inf\{t | N_t > 0\} \tag{8.1}$$

where $N_t = \sum_i \mathbf{1}_{\{\tau^i \leq t\}}$. Let us suppose that the premium payments of $C(t_n - t_{n-1})$ are made at the dates $t_n, n = 1, 2, \ldots, N$, up until the time τ^*. The default (insurance) payment of a random amount h depending on which is the defaulting obligor, happens at the time τ^*. The risk-neutral pricing formula then gives the price of the two sides of the contract:

$$\begin{aligned} V^p &= CE^Q\left[\sum_{n=1}^{N} e^{-\int_0^{t_n} r_s ds}(t_n - t_{n-1})\prod_{i=1}^{I}\mathbf{1}_{\{\tau^i > t_n\}}\right] \\ V^d &= E^Q\left[\mathbf{1}_{\{\tau^*\le T\}} e^{-\int_0^{\tau^*} r_s ds}\left(\sum_{i=1}^{I} h(i,\tau)\mathbf{1}_{\{\tau^i=\tau^*\}}\right)\right] \end{aligned} \tag{8.2}$$

The swap rate is the value C such that $V^p = V^d$.

We now consider these formulas in the multifirm reduced form setting in which the individual obligors have $\mathcal{G}$-adapted default intensities λ^i_t. Then, for example, the survival probability for τ^* to exceed t is given by

$$P[\mathbf{1}_{\{\tau^*>t\}}] = P[N_t = 0] = E[P[N_t = 0|\mathcal{G}_\infty]] = E\left[\prod_i e^{-\int_0^t \lambda^i_s ds}\right] = E[e^{-\int_0^t \lambda^*_s ds}] \tag{8.3}$$

where $\lambda^*_t = \sum_i \lambda^i_s$ represents the total intensity for one firm to default. Similarly, the premium leg can be seen to have price

$$V^p = CE^Q\left[\sum_{n=1}^{N} e^{-\int_0^{t_n}(r_s+\lambda^*_s)ds}(t_n - t_{n-1})\right] \tag{8.4}$$

The default leg is somewhat trickier, since to determine the payment we may need to identify the defaulting obligor. Suppose that the default payment in the event that the first obligor to default is firm i is an $\mathcal{G}_{\tau^*}$-measurable random variable $h_{\tau^*}(i)$. We notice that the basic probability needed to compute such payoffs is the probability density for τ^* to be t while the defaulting obligor is i, and is given by

$$\frac{d}{dt}P[\tau^*\le t, \tau^* = \tau^i] = E[\lambda^i_t e^{-\int_0^t \lambda^*_s ds}] \tag{8.5}$$

The default leg thus has price

$$V^d = E^Q\left[\int_0^T e^{-\int_0^t (r_s+\lambda^*_s)ds}\left(\sum_i h_t(i)\lambda^i_t\right)dt\right] \tag{8.6}$$

In the important special case of *homogeneous default payments* where $h_t(i)$ is independent of i, the default leg simplifies considerably to

$$V^d = E^Q[\int_0^T e^{-\int_0^t (r_s+\lambda^*_s)ds} h_t\lambda^*_t dt] \tag{8.7}$$

It might seem straightforward to compute such expectations in our typical multifirm affine setting, but we will see that difficulties will tend to spin out of control as the number of obligors gets larger than 50 or so. The copula models were developed as a way to simplify these computations.

8.1.2 mth-to-Default Swaps

A moment's thought tells us that in the reduced form setting, the mth-to-default contract will not be solved simply by studying the distributional properties of the joint counting process N_t, because the event $N_t > 1$ may happen when $N_t^i > 1$ for a single firm i as well as if more than one firm defaults. An alternative approach can be made to work by decomposing the mth-to-default contract into first-to-default contracts. For example, the second-to-default contract on three firms (A, B, C) can be decomposed as three first-to-default contracts on pairs $(A, B), (A, C), (B, C)$, minus two times the first-to-default contract on the triple (A, B, C). The general decomposition formula for the mth-to-default contract is known, but the computational intensity makes it intractable for large scale contracts when m, I are large.

8.1.3 First-m-to-Default Swap in Normal Copula Models

Let us now consider pricing this product in the one-factor normal copula model of Section 7.6.5 when the default payments do not depend on which firms default. The premium payment at time t_n is assumed to be $(m - N_{t_n})^+(t_n - t_{n-1})$ while each default payment is assumed to occur at the first payment date following the default in question. For simplicity we take the interest rate to be constant r, and suppose the zero recovery yield curve $Y_0^i(T)$ for firm i is known at time 0. Then, from

$$Y_0^i(T) = -T^{-1} \log(e^{-rT}(1 - PD^i(T)) \tag{8.8}$$

we can determine $F^i(t) := PD^i(t) = 1 - e^{(r - Y_0^i(t))T}$, the marginal risk neutral default probability for firm i.

In the one-factor normal copula model, we can generalize (7.65) and find the probability for exactly k defaults to occur by time t from a given set of obligors:

$$\begin{aligned} S(k, \mathcal{I}, t) &:= P[\text{exactly } k \text{ defaults from the set of obligors } \mathcal{I} = \{1, \dots, I\} \text{ at } t] \\ &= \sum_{\sigma \subset \mathcal{I}, |\sigma| = k} \int_{\mathbb{R}} \prod_{j \in \sigma} \Phi\left(\frac{z_j(t) - \beta_j y}{\bar{\beta}_j}\right) \prod_{j \notin \sigma} \left(1 - \Phi\left(\frac{z_j(t) - \beta_j y}{\bar{\beta}_j}\right)\right) \phi(y) dy \end{aligned} \tag{8.9}$$

where $z_j(t) := \Phi^{-1}(F_j(t))$ and the sum over σ denotes the sum over the possible subsets of the defaulted obligors.

Under the constant interest rate assumption, the premium leg of the mth-to-default swap has value

$$\begin{aligned} V^p &= \sum_{n=1}^{N} C(t_n - t_{n-1}) e^{-rt_n} E[(m - N_{t_n})^+] \\ E[(m - N_{t_n})^+] &= \sum_{k=0}^{m-1} (m - k) S(k, \mathcal{I}, t_n)) \end{aligned} \tag{8.10}$$

The default payment at time t_n can be seen to be $(m - N_{t_{n-1}})^+ - (m - N_{t_n})^+$, and hence the default leg is valued by

$$\begin{aligned} V^d &= \sum_{n=1}^{N} e^{-rt_n} E[(m - N_{t_{n-1}})^+ - (m - N_{t_n})^+] \\ &= \sum_{n=1}^{N} e^{-rt_n} [\sum_{k<m} (m-k) \left[S(k, \mathcal{I}, t_n - 1) - \sum_{k<m} S(k, \mathcal{I}, t_n) \right]] \end{aligned} \tag{8.11}$$

8.1.4 Nonhomogeneous Probabilities: Recursion Algorithm

The following procedure eliminates inefficiencies in computing the combinatorial sums in equations (8.9) and (8.11) in the previous section. Fix t and for $i = 1, 2, \ldots, I$ and $k \le i$ let

$$S(k, i) = \text{Probability exactly } k \text{ firms default from first } i \text{ firms}$$

The following recursion algorithm solves the problem in $O(M^2)$ flops:

1. Initialize
$$S(k, 1) = \begin{cases} 1 - p_1(t) & k = 0, \\ p_1(t) & k = 1, \\ 0 & k \neq 0, 1 \end{cases}$$
2. $S(k, 2) = S(k, 1)(1 - p_2(t)) + S(k-1, 1)p_2(t), \quad k = 0, 1, \ldots;$
3. For $2 < i \le M$,
$$S(k, i) = S(k, i-1)(1 - p_i(t)) + S(k-1, i-1)p_i(t), \quad k = 0, 1, \ldots.$$

8.1.5 Index Credit Default Swaps

The most important examples of index CDSs are the series of CDX NA IG contracts, where the underlying names are a specific set of 125 investment grade North American firms, and the iTraxx Europe contracts written on a set of 125 European firms. They are available with 3, 5, 7 and 10 year maturity. The default leg of an index CDS pays the insured party the losses given default of any firm, until maturity. In exchange, a premium is paid periodically, usually quarterly in arrears. Since firms in default are replaced in the basket, the premium does not change after a default occurs. Moreover, assuming that the replacement firm has similar characteristics the firm it replaces, the situation is as if the ith firm may default any number of times.

Index CDSs are now highly liquid with low bid/ask spreads, and provide perhaps the most important indicator of the overall state of the credit markets. At any time, the quoted market spread on an index CDS (i.e the premium that makes the contract have zero value) can be viewed as the "average" credit spread among the constituent firms, and assuming a fixed recovery rate, provides a direct measure of their risk-neutral default probability over the duration.

8.2 Collateralized debt obligations

CDO is the generic name given to a whole host of products, such as CBO/CLO (collateralized bond or loan obligations) and other asset-backed securities (ABS), that repackage a collection of similarly structured securities such as bonds, loans or CDS, into "tranches" ordered by seniority. These tranches are sold to investors who receive a fair portion of the cashflows generated by the underlying securities. More specifically, each tranche is characterized by an attachment level a and a detachment level b: the tranche absorbs no losses until the losses in the collateral pool hit a (as a percentage of total notional); it absorbs all losses between a and b; and finally no losses in excess of b. This practice of decomposing large portfolios of assets into tranches has become known as "securitization", and is playing an ever increasing role in the finance industry.

A simple example of securitization could be initiated by a bank as an attempt at "regulatory arbitrage" or "balance sheet arbitrage". This is a means to reduce the economic capital (risk capital) required to satisfy the government's risk regulators (in Canada this would be the Office of the Superintendent of Financial Institutions, OSFI), by moving assets off the bank's balance sheet. Suppose in our example that the bank wishes to move a portion of a \$500M portfolio of 100 corporate bonds "off balance sheet". To this aim, they would create a "special purpose vehicle" or SPV, to which they transfer legal ownership of the bundle of 100 bonds. This bundle is decomposed into "tranches", ordered by seniority, and all but the equity tranche sold to other investors with a promise of a fair share of the bond interest. To reduce "moral hazard" connected with the possibility of poor management of the portfolio, the SPV itself retains the 10% "equity tranche", meaning the tranche that absorbs the first 10% of portfolio losses due to default amongst the firms. As viewed by the regulators, the bank, now owners of the SPV, might be seen to retain only \$50M in potential losses, rather than \$500M, and the risk capital assessment to the bank reduced accordingly. On the other hand, the true risk of the equity tranche is almost equal to the risk of the original bond portfolio (since there is only a slight possibility that the portfolio loss will exceed 10%). The Basel II Capital Accord, currently in its implementation stage worldwide, has been to a large extent driven to close regulatory credit risk loopholes such as this.

A CDO such as this that involves the actual ownership of the underlying securities is called a *cash CDO*. Increasingly popular are the so-called *synthetic CDOs*, which are derivatives written on the cash flows of a portfolio of CDSs written on I firms. In their basic forms, synthetic and cash CDOs generate equivalent cash flows, and can be priced by the same methods.

Issues that affect the pricing of bonds and CDSs must all be addressed when pricing CDOs. Thus, in principle, one should choose between structural and reduced-form approaches, a stochastic interest rate model or a fixed interest rate, which recovery mechanism is most appropriate. For our present purposes however, we will focus on the issues particular to CDOs and simplify other aspects by assuming:

Assumption 4. We will assume

1. a constant rate of interest r;

2. all defaultable bonds and CDSs will be priced assuming the recovery of par mechanism with a fixed fractional recovery rate R;

3. Premium payments are made quarterly in arrears, at dates $t_k, k = 1, \ldots, K$. Accrual payments will be ignored unless otherwise stated.

4. Defaults may occur anytime, and default payments are made at the instant of default.

A CDO is best understood schematically as composed of two types of basic contingent claims whose cash flows depend on the default losses of the underlying collateral pool. We suppose for the moment that the collateral pool is a portfolio of corporate bonds of similar maturity issued by a number of firms. The cash flows are analogous to insurance and premium payments paid to insure against losses from default events that impact the bond portfolio. We define the following quantities:

- the maturity date T of the contract;

- I reference credits with notional amounts (par value) of $N_i, i = 1, 2, \ldots, I$;

- the default time τ^i of the ith credit, a $\mathcal{F}_t$ stopping time;

- the loss $l_i = (1 - R)N_i/N$ caused by default of the ith credit as a fraction of the total notional $N = \sum_i N_i$;

- the cumulative portfolio loss $L(t) = \sum_i l_i \mathbf{1}_{\{\tau^i \leq t\}}$ up to time t as a fraction of the total notional.

- the attachment/detachment points $b_0 = 0 < b_1 < b_2 < \ldots, b_T = 1$. Two standard sets are $0, 0.03, 0.07, 0.10, 0.15, 0.30$ for the CDX index CDOs, and $0, 0.03, 0.06, 0.09, 0.12, 0.22$ for the I-Traxx index CDOs. These points define nonoverlapping tranches $[b_{i-1}, b_i]$.

- the *tranche function* $T^{(1)}(b|t) := E^Q[(b - L(t))^+]$.

The *super-senior tranche* is the highest tranche $[b_{T-1}, 1]$, that is with detachment point 1. The *equity tranche* is the most junior tranche, $[0, b_1]$, with 0 as its attachment point. Intermediate tranches are called *mezzanine tranches.* We also consider so-called *base tranches* $[0, b_i], i = 1, 2, \ldots$. A general tranche decomposes as the difference between base tranches, so first we consider the CDO cash flows for base tranches.

The writer (i.e. the *seller of default insurance*) of one unit of a *base default leg with detachment level* b, that is the default leg of a base tranche, pays the holder (the *buyer of default insurance*) all default losses up to a certain level $0 < b \leq 1$ that occur over the period $[0, T]$, the maturity of the contract. Specifically, this means that at the default time $\tau^i \leq T$ of any firm, the parties exchange cash of the amount

$$\left(b - L(\tau^i-)\right)^+ - \left(b - L(\tau^i)\right)^+ ,$$

where $L(\tau^i-)$ denotes the left limit of the cádlág process L_t.

The writer (i.e. the buyer of default insurance) of one unit of a *base premium leg with detachment level* b, that is the premium leg of a base tranche, pays the holder (the insurer) on prespecified payment dates $t_k < t_K = T$, $k = 1, \dots, K$ an amount jointly proportional to the year fraction $t_k - t_{k-1}$ and the residual portfolio value below b. In practice there is also an extra amount, called the *accrual term*, to account for defaults between payment dates: this we ignore in this discussion. Thus the cash exchanged on date t_k is

$$(t_k - t_{k-1})(b - L(t_k))^+.$$

These cash flows are nicely summarized by the following risk neutral valuation formulas:

1. The fair price of the base default leg with detachment level b is
$$W(b) = -E^Q[\int_0^T e^{-rt} d(b - L(t))^+]; \tag{8.12}$$

2. The fair price of the base premium leg with detachment level b is
$$V(b) = \sum_{k=1}^{K} E^Q[e^{-rt_k}(t_k - t_{k-1})(b - L(t))^+] = \sum_{k=1}^{K} e^{-rt_k}(t_k - t_{k-1}) T^{(1)}(b|t_k) \tag{8.13}$$

Notice that the formula for $V(b)$ is an approximation to the Riemann integral $\int_0^T e^{-rt} T^{(1)}(b|t) dt$: the accrual terms usually added can be thought of as corrections to the Riemann sum approximation. To understand the formula for $W(b)$, note that $(b - L(t))^+$ is a non-increasing function, and hence its integrals can be constructed path-by-path. One can show that

$$\lim_{h \to 0} \frac{1}{h} E^Q[(b - L(t+h))^+ - (b - L(t))^+] = \frac{d}{dt} T^{(1)}(b|t) \tag{8.14}$$

and thus

$$W(b) = -\int_0^T e^{-rt} \frac{d}{dt} T^{(1)}(b|t) dt \tag{8.15}$$

Finally, it is often convenient to eliminate the time derivative by integration by parts:

$$W(b) = b - e^{-rT} T^{(1)}(b|T) - \int_0^T r e^{-rt} T^{(1)}(b|t) dt \tag{8.16}$$

Some thought about computational issues suggests that the time integrals can be done efficiently enough by Simpson's rule, and the true computational roadblocks will come in the accurate and efficient computation of the tranche function $T^{(1)}(b|t)$ (and possibly its time derivative) over the full range of parameters $t \in [0, T], b \in [0, 1]$, for portfolios with a large number M of possibly nonhomogeneous reference credits. This indeed will prove to be the main challenge for any numerical algorithm.

General tranche legs are now defined in terms of the base tranche legs. The *default* $[b, b']$*–tranche leg* pays the amount

$$(b' - L(\tau_j-))^+ - (b' - L(\tau_j))^+ - (b - L(\tau_j-))^+ - (b - L(\tau_j))^+$$

at each default time $\tau_j \leq T$ and its fair price is therefore $W(b') - W(b)$. The *premium* $[b, b']$*–tranche leg* pays periodic amounts

$$(t_{k+1} - t_k)\left((b' - L(t_k))^+ - (b - L(t_k))^+\right)$$

and hence the fair price for the continuous time premium leg is $V(b') - V(b)$. The $[b, b']$–*tranche spread* is that multiplier $X^{[b,b']}$ of the premium tranche leg that solves the balance equation

$$X\left((V(b') - V(b)\right) = W(b') - W(b). \tag{8.17}$$

8.3 Pricing CDOs in Copula Models

In this section we consider the computational issues in implementing CDO pricing in copula models, generalizing the discussion of Section 7.6.5. We suppose in all cases that the market implied marginal cumulative default distribution for each firm is $F^{(i)}(t) := E^Q[\tau^i \leq t]$, determined by the market prices of that firm's bonds and CDSs.

8.3.1 Some specific CDOs

1. (Fully homogeneous portfolio) Consider the case of a cash CDO written on a portfolio of I bonds, each with the same notional $\mathcal{N}$, and each with a constant hazard rate λ for default so that $F^{(i)}(t) = 1 - e^{-\lambda t}$. Let the premiums be paid quarterly in arrears. Consider the attachment points $3\%, 7\%, 10\%, 15\%, 30\%$, standard for the CDX index CDOs. This is the so-called *standard model* for CDOs.

2. (Homogeneous notionals, varying default characteristics) Here we take the same portfolio as in the standard model, but suppose instead that the hazard rates $\lambda_m, m = 1, \ldots, I$ vary amongst the firms.

3. (Varying notionals, varying default characteristics) Here we allow both the notional amounts $\mathcal{N}_m$ and hazard rates λ_m to vary amongst the firms.

4. (Homogeneous notionals, general defaultable yield curves) Here we take each firm with the same notional $\mathcal{N}$, but suppose each firm has a marginal default distribution $F^{(i)}(t)$ determined from that firm's defaultable yield curve.

We now investigate what is involved in computing CDO prices in some specific copula models. The following lists some simple choices we will consider.

1. One factor normal copula model with homogeneous weights $a_i = a \in [0,1]$.
2. One factor normal copula model with nonhomogeneous weights $a_i \in [-1,1]$.
3. The Clayton copula model with parameter θ.

8.3.2 Pricing a Homogeneous CDO

We first consider a fully homogeneous portfolio in the homogeneous one-factor normal copula model. In this case, $L(t) = \frac{1-R}{I}N(t)$ where $N(t) = \sum_i \mathbf{1}_{\{\tau^i \le t\}}$ is the number of defaults by time t. An easy calculation shows that

$$\begin{aligned} T^{(1)}(b|t) &= \sum_{n=0}^{n_{\max}} \left(b - \frac{(1-R)n}{I}\right) P[N(t)=n], \qquad n_{\max} = \left\lfloor \frac{Ib}{1-R} \right\rfloor \quad (8.18) \\ P[N(t)=n] &= \int_{-\infty}^{\infty} \frac{I!}{(I-n)!n!} \; p(t,y)^n (1-p(t,y))^{I-n} \phi(y)dy \\ p(t,y) &= \Phi\left(\frac{\Phi^{-1}(1-e^{-\lambda t}) - ay}{\sqrt{1-a^2}}\right) \quad (8.19) \end{aligned}$$

By plugging $T^{(1)}$ into (8.16) and (8.13) one finds model formulas for $W(b), V(b)$ and hence the spreads $X^{[b,b']}$.

Before proceeding, note that this computation can be reexpressed as

$$\begin{aligned} T^{(1)}(b|t) &= \int_{-\infty}^{\infty} T^{(1)}(b|t,y)\phi(y)dy \quad (8.20) \\ T^{(1)}(b|t,y) &= E^Q[(b-L(t))^+|Y=y] = \sum_{n=0}^{n_{\max}} \left(b - \frac{(1-R)n}{I}\right)^+ P[N(t)=n|y] \\ P[N(t)=n|y] &:= \frac{I!}{(I-n)!n!} \; p(t,y)^n (1-p(t,y))^{I-n} \quad (8.21) \end{aligned}$$

This shows clearly the usefulness of the conditional independence when conditioned on the underlying factor Y.

The Large I Approximation: It was first observed by Vasicek [23] that the conditional binomial distribution of $L(t)|_{Y=y}$ implied in (8.20) can be speedily approximated by a normal distribution with mean $m(t,y) := (1-R)p(t,y)$ and variance $V(t,y) := I^{-1}(1-R)^2 p(t,y)(1-p(t,y))$ when I is large:

$$T^{(1)}(b|t,y) \sim \int_{-\infty}^{\infty} (b - m(t,y) - \sqrt{V(t,y)}z)^+ \phi(z)dz \quad (8.22)$$

which has a closed formula.

The Standard Model: In our discussion of index credit default swaps in Section 8.1.5, we have seen that practitioners view the par credit premium CDS as the average recovery adjusted credit spread amongst the constituent names of the index. It is used to fix a value of λ via the formula $CDS = (1-R)\lambda$, with R taken to be fixed at 40%. This specification of the one-factor normal copula model is called the *standard model* for index CDOs, and is used by traders in much the same way the Black-Scholes formula is used by option traders, that is as a reference frame for quoting prices of index CDOs. In the standard model, the key parameter of interest is the so-called *base correlation*.

Base Correlation: Let the model formulas for the base legs be given as functions W, V of the two parameters b, ρ. Now suppose that the market observed spreads on the index CDO tranches are $\hat{X}^{[b_{i-1},b_i]}$. For each attachment point b_j, one defines the base correlations $\rho_1 := \rho(b_1), \rho_2, \ldots$ as the solution of the following set of equations:

$$\begin{aligned} \hat{X}^{[0,b_1]}V(b_1,\rho_1) &= W(b_1,\rho_1) \\ \hat{X}^{[b_{i-1},b_i]}(V(b_i,\rho_i) - V(b_{i-1,\rho_{i-1}})) &= W(b_i,\rho_i) - W(b_{i-1},\rho_{i-1}), \quad i = 2,3,\ldots \end{aligned} \tag{8.23}$$

In much the same way that traders quote derivative prices in terms of implied volatility, and develop an intuition for the market by plotting implied volatility smiles, CDO traders quote prices of index tranches in terms of the implied base correlation. The plot of base correlation as a function of b gives a snapshot of the CDO market. It is another case of a wrong formula being used with wrong parameters to quote the right (market) prices.

The standard model owes its popularity to its simplicity, rather than its accuracy, and the use of base correlation as a metric for CDO pricing can be seen to be very dangerous in some situations.

One known deficiency of the standard model stems from its lack of tail dependency, which leads to a tendency to undervalue the premium to be paid on the most senior tranche. The Clayton copula, with its lower tail dependency, has more tendency to produce clustering of defaults, leading to higher prices for the premium leg.

8.3.3 Pricing a Nonhomogeneous CDO

The first extension of the standard model must be to allow individual firms to have their individual default distribution. We first consider specification 2, with firms having equal notionals and different constant hazard rates $\lambda_i > 0$. To compute the tranche function in the one-factor normal copula model with nonhomogeneous correlations, we again have (8.18), but now

$$\begin{aligned} P[N(t)=n] &= \sum_{\Sigma\in\{1,\ldots,I\},|\Sigma|=n} \int_{-\infty}^{\infty} \prod_{i\in\Sigma} p(t,y,i) \prod_{i\notin\Sigma} (1-p(t,y,i))\phi(y)dy \\ p(t,y,i) &= \Phi\left(\frac{\Phi^{-1}(1-e^{-\lambda_i t}) - a_i y}{\sqrt{1-a_i^2}}\right) \end{aligned} \tag{8.24}$$

For each value of t, one computes $P[N(t) = n]$ using the recursion algorithm of Section 8.1.4. Assuming one uses Simpson's rule with K time steps, the total algorithm runs in $O(I^2K)$ flops.

More generally, in specification 4, each firm's marginal risk neutral default probability is $F_i(t)$, and one has the same formula for $P[N(t) = n]$ with

$$p(t, y, i) = \Phi\left(\frac{\Phi^{-1}(F_i(t)) - a_i y}{\sqrt{1 - a_i^2}}\right)$$

When M is large enough, we again have an alternative approximation formula to the exact recursion algorithm by approximating the conditional loss distribution $L(t)|_{Y=y}$ by a normal random variable with mean $m(t, y) = I^{-1}\sum_i(1 - R)p(t, y, i)$ and variance $V(t, y) = I^{-2}\sum_i(1 - R)^2 p(t, y, i)(1 - p(t, y, i))$.

8.3.4 Pricing CDOs with Varying Notionals

Most standardized CDOs are designed with all firms having identical notionals, and priced under the assumption of a fixed fractional recovery rate R, which implies that the fractional loss given default l_i is independent of the firm. On the other hand, much of the CDO business is concerned with so-called *bespoke* CDOs, that are over-the-counter contracts tailor-made for an individual client ("bespoke" is the term to describe an expensive tailor-made suit one can find in Saville Row in London), and bespoke CDOs will generally have unequal notionals.

Some thought will convince one that no straightforward extension of the recursion algorithm of Section 8.1.4 will now work, and exact computation of the conditional loss distribution is an intractible analytical problem for I larger than about 20. Instead one is forced to adopt one of a variety of approximation techniques. [1] consider the approximation where the possible loss amounts are taken on a fine grid, and develop Monte Carlo methods for this approximation. Several authors have considered large I approximations based on the Poisson distribution [12], the normal approximation and the saddlepoint approximation [2, 25].

8.4 Exercises

Exercise 70. Prove (8.5).

Exercise 71. Read the articles [1, 17, 11, 24]

Exercise 72. Compute the fair CDS rate for an index CDS on I firms with varying constant hazard rates λ^i. Treat the interest rate r, recovery rate R, maturity T as parameters. Discuss why the CDS rate should not be very sensitive to a stochastic interest rate, justifying treating r as constant (you should first verify numerically that the CDS rate is indeed insensitive to r).

Exercise 73. Implement the standard homogeneous CDO model for finite I, and investigate the accuracy of the large I approximation. Take parameters $R = 0.40, \lambda = 0.01$ and the standard CDX tranche attachment points. Tabulate the exact and approximate tranche spreads for $I = 20, 50, 100, 200$ and $\rho = 0, .2, .5$.

Exercise 74. Implement the Clayton copula model for both the homogeneous and nonhomogeneous CDO. Compare the exact tranche spread values to those computed in the large I approximation. Pick a set of parameters for this model that give a good fit to the standard model you implemented in the previous exercise. See if you can identify where the Clayton model differs most significantly from the standard model (I suggest you plot the tranche functions and see where they differ most.)

Exercise 75. Develop an algorithm to compute the base correlations ρ_i from observed market data consisting of the index CDS rate and the tranche spreads. Use the standard model, and be sure to match the CDS and CDO maturities.

Exercise 76. This exercise is really an open ended project that aims to develop a detailed practical understanding of the CDX NA IG index products.

1. Go to the Trading Room at the Business School and find out what data is available both from Reuters and Bloomberg on this index. Be sure to include both the index CDS and the CDO index tranches. If possible, download daily data on at least one date, preferably a number of dates.

2. Plot the base correlation curves for the data you have.

3. Interpolate the base correlation curve to compute the spread for the "bespoke" CDO tranche $[0.04, 0.08]$ with maturity $T = 6$.

Exercise 77. (This exercise should hopefully be done based on data from the same date as the previous exercise). An important ingredient in CDO pricing is obtaining the risk neutral CDF $F^i(t)$ for each of the constituent firms.

1. Go to the Trading Room at the Business School and find CDS data (all available maturities on a single date) on 4 individual firms from the CDX NA IG portfolio. Select one typical firm from each rating class AAA, AA, A, BBB. Obtain the index CDS rates for the same dates.

2. Based on the recovery of par formulas, use the CDS data together with interpolation (I suggest cubic splines), to construct an implied CDF for the default time of each firm. Denote these functions by $F^{AAA}, F^{AA}, F^A, F^{BBB}$.

3. Now assume that each constituent firm from the CDX index has the default CDF corresponding to its current rating class. Price the index CDOs using these marginal default distributions in the nonhomogeneous one factor normal copula model. Use the base correlations from the previous exercise to price each tranche.

4. Compare the model prices just computed to the observed CDO prices for that date.

Appendix A

Mathematical Toolbox

The technical results that follow form the underpinning of the modeling done in these notes. It is important to have a strong facility with them.

A.1 Itô diffusions

A one-dimensional *Itô diffusion* is a stochastic process $X_t, t \geq 0$ that satisfies an SDE of the form

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \tag{A.1}$$

for deterministic functions a, b.

The following[1] is a typical existence and uniqueness theorl2506em for Itô SDEs taking values in n-dimensional Euclidean space $\mathbb{R}^n$ and driven by an m-dimensional Brownian motion W; the proof may be found in [20][§5.2].

Let $T > 0$, and let

$$a \;:\; \mathbb{R}^n \times [0, T] \to \mathbb{R}^n \tag{A.2}$$

$$b \;:\; \mathbb{R}^n \times [0, T] \to \mathbb{R}^{n \times m}; \tag{A.3}$$

be measurable functions for which there exist constants C and D such that

$$\big|a(x,t)\big| + \big|b(x,t)\big| \;\leq\; C\big(1 + |x|\big) \tag{A.4}$$

$$\big|a(x,t) - a(y,t)\big| + \big|b(x,t) - b(y,t)\big| \;\leq\; D|x - y|; \tag{A.5}$$

for all $t \in [0, T]$ and all x and $y \in \mathbb{R}^n$, where

$$|b|^2 = \sum_{i,j=1}^{n} |b_{ij}|^2. \tag{A.6}$$

[1]This section is taken from Wikipedia:
`http://en.wikipedia.org/wiki/Stochastic differential equation.`

Let Z be a random variable that is independent of the σ-algebra generated by $W_s, s \geq 0$, and with finite second moment:

$$\mathbb{E}\left[|W|^2\right] < +\infty. \tag{A.7}$$

Then the stochastic differential equation/initial value problem

$$\begin{aligned} \mathrm{d}X_t &= a(X_t, t)\,\mathrm{d}t + b(X_t, t)\,\mathrm{d}W_t \text{ for } t \in [0, T]; && \text{(A.8)} \\ X_0 &= Z; && \text{(A.9)} \end{aligned}$$

has an almost surely unique t-continuous solution $(t, \omega) \mapsto X_t(\omega)$ such that X is adapted to the filtration generated by Z and $W_s, s \leq t$, and

$$\mathbb{E}\left[\int_0^T |X_t|^2\,\mathrm{d}t\right] < +\infty. \tag{A.10}$$

A.2 The Strong Markov Property and the Markov Generator

A time homogeneous Itô diffusion X_t subject to an SDE of the form

$$dX_t = a(X_t)dt + b(X_t)dW_t$$

satisfies the strong Markov property that says that the dynamics of X to the future of any *stopping time* may depend on the value of X at the stopping time, but not on the values of X to the past of that time.

Theorem A.2.1 (Strong Markov Property)**.** *Let f be a bounded Borel function on $\mathbb{R}^n$, τ a stopping time w.r.t $\mathcal{F}_t$ with $\tau < \infty$ a.s. Then for all $h \geq 0$*

$$E_x[f(X_{\tau+h})|\mathcal{F}_\tau] = E_{X_\tau}[f(X_h)].$$

Definition 9. The *Markov generator* of the process X_t is defined by

$$\mathcal{L}f(x) = \lim_{t \to 0+} \frac{E_x[f(X_t)] - f(x)}{t}$$

for functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit exists at x. For C_0^2 functions f it is given by

$$\mathcal{L}f(x) = \sum_i a_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{ij}(bb')_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}$$

A.3 Itô Formula

In its simplest form, Itô's formula states that for an Itô process

$$dX_t = \mu_t\, dt + \sigma_t\, dW_t$$

and any twice continuously differentiable function f on the real numbers, then $f(X)$ is also an Itô process satisfying

$$\begin{aligned} df(X_t) &= f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\sigma_t^2\, dt \\ &= f'(X_t)\sigma_t\, dW_t + \left(f'(X_t)\mu_t + \frac{1}{2} f''(X_t)\sigma_t^2 \right) dt. \end{aligned}$$

Or, for any function $f(t,x)$ with continuous first-order partial derivatives and bounded second-order partial derivatives

$$\dot{f}(t,x) = \frac{\partial f(t,x)}{\partial t}, \quad f'(t,x) = \frac{\partial f(t,x)}{\partial x}, \quad f''(t,x) = \frac{\partial^2 f(t,x)}{\partial x^2}.$$

then:

$$\begin{aligned} df(t,X_t) &= \dot{f}(t,X_t)\, dt + f'(t,X_t)\, dX_t + \frac{1}{2} f''(t,X_t)\sigma_t^2\, dt, \\ &= \dot{f}(t,X_t)\, dt + f'(t,X_t)(\mu_t\, dt + \sigma_t\, dW_t) + \frac{1}{2} f''(t,X_t)\sigma_t^2\, dt, \quad \text{rearranging terms} \\ &= \left(\dot{f}(t,X_t) + \mu_t f'(t,X_t) + \frac{\sigma_t^2}{2} f''(t,X_t) \right) dt + f'(t,X_t)\sigma_t\, dW_t \end{aligned}$$

More generally, Itô's formula applies for any continuous d-dimensional semimartingale $X = (X^1, X^2, \ldots, X^d)$, and twice continuously differentiable and real valued function f on $\mathbb{R}^d$, then $f(X)$ is an Itô process satisfying

$$df(X_t) = \sum_{i=1}^{d} f_{,i}(X_t)\, dX_t^i + \frac{1}{2} \sum_{i,j=1}^{d} f_{,ij}(X_t)\, d[X^i, X^j]_t.$$

In this expression, the term f,i represents the partial derivative of $f(x)$ with respect to x^i, and $[X^i, X^j]$ is the quadratic covariation process of X^i and X^j.

A.4 The Feynman-Kac Formula

For any sufficiently integrable functions $F(x), \phi(t,x)$ (bounded is sufficient), and time $T > 0$, the Markov property implies the existence of a function $f(t,x), t \in [0,T], x \in \mathbb{R}$ such that

$$f(t,X_t) = E[F(X_T) e^{\int_t^T \phi(s,X_s) ds} | \mathcal{F}_t] \tag{A.11}$$

The Feynman-Kac formula states that f is the solution of the parabolic partial differential equation:

$$\begin{cases} \partial_t f(t,x) + \mathcal{L}[f](t,x) + \phi(t,x) f(t,x) = 0 & t < T \\ f(T,x) = F(x) \end{cases} \tag{A.12}$$

The validity of this formula can be seen as follows: since the process $e^{\int_0^t \phi(s,X_s)ds} f(t,X_t)$ must be a martingale, the zero-drift condition and Itô's formula lead to the PDE.

A.5 Kolmogorov Backward Equation

For any times $0 \le t \le T$, the conditional probability density function of $X_t|_{\mathcal{F}_s}$,

$$p(t,x;T,y) := \partial_y P[X_T \le y | X_t = x] \tag{A.13}$$

satisfies the PDE

$$\begin{cases} \partial_t p(t,x;s,y) + \mathcal{L}[p](t,x;s,y) = 0 & t < s \\ p(s,x;s,y) = \delta(x-y) \end{cases} \tag{A.14}$$

This result is simply giving the fundamental solution of the Feynman-Kac equation with $\phi = 0$. For any smooth function F

$$f(t,x) := E[F(X_T)|X_t = x] = \int_{\mathbb{R}} F(y) p(t,x;T,y) dy \tag{A.15}$$

solves the FK PDE, which then implies (A.14).

A.6 Ornstein-Uhlenbeck Processes

A one-dimensional OU process is a special form of Itô process:

$$dX_t = (a - bX_t)dt + c dW_t, \qquad a, b, c \ \text{ constants} \tag{A.16}$$

This SDE can be solved explicitly as follows. First note that

$$d\left(e^{bt}(X_t - a/b)\right) = c e^{bt} dW_t$$

can be integrated from $[S,T]$ giving

$$e^{bT}(X_T - a/b) = e^{bS}(X_S - a/b) + c \int_S^T e^{bt} dW_t$$

or equivalently

$$X_T = e^{-b(T-S)} X_S + (1 - e^{-b(T-S)}) a/b + c \int_S^T e^{-b(T-t)} dW_t$$

Thus the marginals are normally distributed with

$$\begin{aligned} E[X_t|\mathcal{F}_s] &= e^{-b(t-s)}X_s + (1-e^{-b(t-s)})a/b \qquad (A.17)\\ \mathrm{Var}(X_t|\mathcal{F}_s) &= E[(c\int_s^t e^{-b(t-u)}dW_u)^2|\mathcal{F}_s] = c^2\int_s^t e^{-2b(t-u)}du \\ &= \frac{c^2}{2b}(1-e^{-2b(t-s)}) \qquad (A.18)\end{aligned}$$

where in the last line we use the Itô Isometry. The above discussion extends easily to the general multi-dimensional OU process

$$dX_t = (A + BX_t)dt + CdW_t$$

where X, A, B, C, W are matrices of size $(d,1), (d,1), (d,d), (d,d), (d,1)$ respectively.

A.7 Girsanov Theorem

Here is a statement of the theorem:

Theorem A.7.1 (Cameron-Martin-Girsanov Theorem)**.** *Let ϕ_t be a process adapted to the natural filtration of the Wiener process $\{\mathcal{F}_t^W\}$ that satisfies the Novikov condition:*[2] *For each $t > 0$*

$$E[e^{\frac{1}{2}\int_0^t \phi_s^2 ds}] < \infty.$$

Then

$$Z_t = \exp[\int_0^t \phi_s dW_s - \frac{1}{2}\int_0^t \phi_s^2 ds],$$

is a martingale that satisfies $E[Z_t] = 1$. A probability measure Q equivalent to P can be defined by the Radon-Nikodym derivative

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = Z_t$$

and the process

$$W_t^Q = W_t - \int_0^t \phi_s ds$$

is a Brownian motion on the probability space $\{\Omega, \mathcal{F}, \mathcal{F}_t^W, Q\}$.

[2]The theorem applies in greater generality: strengthening Girsanov's theorem is a cottage industry in probability theory.

A.8 Arbitrage Pricing Theory

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[3] $S_t = (S_t^0 = C_t, S_t^1, \ldots, S_t^d)$. An *admissible 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[4] $X_t^H = H_t' S_t$ and the strategy is called *self-financing* if its wealth process satisfies

$$dX_t^H = H_t' dS_t \tag{A.19}$$

As always, we assume the nonexistence of arbitrage, where 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, equivalent to P, 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 H_t such that $X_T^H = B$ (almost surely) in which case the law of one price dictates that the price of the claim 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 the equivalent martingale measure Q_0 is unique. Because the discounted wealth of admissible self-financing portfolios are always Q_0 martingales, it follows that the discounted prices of contingent claims in complete markets are martingales with respect to Q_0. In *incomplete markets*, there exist more than one (in fact 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, non-replicable claims have an effective impact in the economy, in the sense that their presence cannot be replicated by the assets that are already being traded. That is, once a non-replicable claim is written and starts 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

$$\pi_t^B = E_t^Q \left[e^{-\int_t^T r_s ds} B \right], \tag{A.20}$$

for some equivalent martingale measure Q. In particular, the price of a zero coupon bond can be written

[3]The general theory of semimartingales is summarized in the textbook [21].

[4]Here, A' denotes the transpose of a matrix A and we use matrix multiplication.

$$P_t(T) = E_t^Q \left[e^{-\int_t^T r_s ds} \right]. \tag{A.21}$$

A.9 Change of Numeraire

We now describe a general technique, particularly useful in interest rate theory, that transforms integrals involved in calculating the prices of derivatives through expressions such as (A.20).

Definition 10. A *numeraire* N_t is any P-a.s strictly positive traded asset. The price of an asset *normalized by the numeraire* N has the form X_t/N_t. The standard example of a numeraire is the money-market C_t, in terms of which normalization is discounting.

The concepts of arbitrage, replicability and completeness are all expressed in terms of self-financing trading strategies, and the following *invariance lemma* shows that the self-financing condition is invariant with respect to a change of numeraire.

Lemma A.9.1. *Let N_t be any P-a.s. strictly positive process (possibly, but not necessarily a numeraire). 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: For semimartingales X_t, Y_t, the generalized Itô product formula can be written

$$d(X_t Y_t) = Y_{t-} dX_t + X_{t-} dY_t + d[X,Y]_t$$

where $[X,Y]_t$ is the quadratic variation process[5] We also note that if X is a stochastic integral $X = H \circ S$ (meaning $dX_t = H_t' dS_t$), then $[X,Y] = H \circ [S,Y]$. Supposing now that $X_t^H = H_t' S_t$, these facts imply that

$$d\left(\frac{X_t^H}{N_t}\right) - H_t' d\left(\frac{S_t}{N_t}\right) = (dX_t^H - H_t' dS_t)/N_t. \tag{A.22}$$

which proves the result. □

Theorem A.9.2. *Let Q be an equivalent martingale measure for the market described in section A.8 (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 X_t normalized by N_t is a martingale with respect to the measure Q^N defined by the Radon-Nikodym derivative*

$$\frac{dQ^N}{dQ} := Z := \frac{N_T C_0}{N_0 C_T}. \tag{A.23}$$

[5] In the general case, processes are not necessarily continuous but rather càdlàg, and we write Y_{t-} to denote the left limit of Y at t.

Proof: The following facts about measure changes are needed in the proof. If $P \sim Q$ where $dP/dQ := Z$ is $\mathcal{F}-$measurable, and $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{F}$, then for any $\mathcal{F}$ random variable X

$$\begin{aligned} E^P[X] &= E^Q[XZ] &\text{(A.24)}\\ E^P[X|\mathcal{H}] &= (E^Q[Z|\mathcal{H}])^{-1}E^Q[XZ|\mathcal{H}] &\text{(A.25)}\end{aligned}$$

Since X_t/C_t and N_t/C_t are both discounted traded assets, they must both be Q–martingales. Therefore

$$\begin{aligned} \frac{X_t}{N_t} &= \left(\frac{X_tC_0}{C_tN_0}\right)\left(\frac{C_tN_0}{C_0N_t}\right)\\ &= E_t^Q\left[\frac{X_TC_0}{C_TN_0}\right]\left(E_t^Q\left[\frac{C_0N_T}{C_TN_0}\right]\right)^{-1}\\ &= E_t^Q\left[\frac{X_T}{N_T}Z\right](E_t^Q[Z])^{-1} = E_t^{Q^N}\left[\frac{X_T}{N_T}\right]\end{aligned}$$

□

As an application of this theorem, let us consider, for a fixed maturity T, the bond prices $P_t(T)$ as a numeraire. Then $N_T = P_T(T) = 1$ and the price of a contingent claim B maturing at T reduces to

$$\pi_t^B = P_t(T)E_t^T[B], \qquad \text{(A.26)}$$

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

$$\frac{dQ^T}{dQ} = \frac{P_T(T)C_0}{P_0(T)C_T} = \frac{\exp\left(-\int_0^T r_s ds\right)}{P_0(T)}. \qquad \text{(A.27)}$$

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

The following result is a reformulation of Exercise 1 and motivates calling Q^T the T-forward measure.

Proposition A.9.3. *The simply compounded forward rate for any period $[S,T]$ is a martingale under the T–forward measure:*

$$L_t(S,T) = E_t^T[L_S(T)] \qquad \text{(A.28)}$$

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

$$L_t(S,T)P_t(T) = \frac{1}{T-S}[P_t(S) - P_t(T)]$$

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

$$L_t(S,T) = \frac{1}{T-S}\left[\frac{P_t(S)}{P_t(T)} - 1\right]$$

is a Q^T-martingale. □

A.10 First Passage Time for Brownian Motion with Drift

We have seen that many models in credit risk are solvable in terms of the first passage time for Brownian motion with drift. This beautiful result from classic stochastic analysis is proved many places, for example [22], [7]. We present here the key ideas, and leave the details for the interested reader. Let

$$\overline{M}_t^\mu = \max_{s\le t} X_t, \quad \underline{M}_t^\mu = \min_{s\le t} X_s \tag{A.29}$$

denote the maximum and minimum processes for a BM $X_t = W_t + \mu t$ with constant drift $\mu \in \mathbb{R}$. The joint probability distributions for $(\overline{M}_t^\mu, X_t)$ and $(\underline{M}_t^\mu, X_t)$ are all expressible in terms of the following important function

$$\mathrm{FP}(a,b;\mu,t) := N\left[\frac{a-\mu t}{\sqrt{t}}\right] - e^{2\mu b} N\left[\frac{a-2b-\mu t}{\sqrt{t}}\right], \quad -\infty < a \le b < \infty, t \ge 0 \tag{A.30}$$

Proposition A.10.1. *The following formulas hold for all $a \in \mathbb{R}, b \ge 0$:*

1. $P[\overline{M}_t^\mu \le b, X_t \le a] = FP(a\wedge b, b; \mu, t)$.
2. $P[\underline{M}_t^\mu \ge -|b|, X_t \ge a] = FP(-a\wedge b, b; -\mu, t)$.
3. $P[\overline{M}_t^\mu \le b] = FP(b; \mu, t) := FP(b, b; \mu, t)$.
4. $P[\underline{M}_t^\mu \ge -|b|] = FP(b; -\mu, t)$.

Outline Proof: The first formula implies the remaining formulas simply by changing the sign of the Brownian motion and using the fact that $\underline{M}_t^\mu \ge W_t + \mu t \le \overline{M}_t^\mu$ with probability one. Proof of the first formula has two steps. First it is proved for $\mu = 0$ by the reflection principle that states: for $a \le b$ and $\mu = 0$,

$$\begin{aligned} P[\overline{M}_t^0 \le b, W_t \le a] &= P[W_t \le a] - P[\overline{M}_t^0 \ge b, W_t \le a] \\ &= P[W_t \le a] - P[\overline{M}_t^0 \le b, W_t \ge 2b - a] \\ &= P[W_t \le a] - P[W_t \ge 2b - a] \\ &= N[a/\sqrt{t}] - N[(a-2b)/\sqrt{t}] := \mathrm{FP}(a\wedge b, b; 0, t) \end{aligned} \tag{A.31}$$

Then the result is extended to the case $\mu \ne 0$ by use of the Girsanov Theorem for the change of measure with Radon-Nikodym derivative

$$Z := \left(\frac{dP^{(\mu)}}{dP}\right)\Bigg|_{\mathcal{F}_t} = e^{-\mu W_t - \frac{1}{2}\mu^2 t}$$

The theorem implies that $\underline{M}_t^\mu, X_t$ under P^μ are distributed identically to $\underline{M}_t^0, W_t$ under P. Thus

$$\begin{aligned} E[\mathbf{1}_{\{\overline{M}_t^\mu \le b\}}\mathbf{1}_{\{X_t\le a\}}] &= E^\mu[Z^{-1}\mathbf{1}_{\{\overline{M}_t^\mu \le b\}}\mathbf{1}_{\{X_t\le a\}}] \\ &= E^\mu[e^{\mu(X_t-\mu t)+\frac{1}{2}\mu^2 t}\mathbf{1}_{\{\overline{M}_t^\mu \le b\}}\mathbf{1}_{\{X_t\le a\}}] \\ &= \int_{-\infty}^a e^{\mu(y+\mu t)-\frac{1}{2}\mu^2 t}\left[\int_{y\vee 0}^b [\partial_x\partial_y F(y,x;0,t)]dx\right]dy \\ &= \int_{-\infty}^a e^{\mu(y+\mu t)-\frac{1}{2}\mu^2 t}\frac{1}{\sqrt{t}}\left[\phi(-|y|/\sqrt{t})-\phi(y-2b/\sqrt{t})\right]dy \end{aligned} \tag{A.32}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}\exp[-x^2/2]$. Completing the square in the exponent then leads to computable integrals, and the desired formula. □

A.11 Multivariate Normal Distributions

In this section, Σ is a symmetric $I \times I$ matrix. Σ is said to be *positive definite* if $x'\Sigma x > 0$ for all $x \in \mathbb{R}^I \setminus 0$.

The following statements are equivalent:

1. Σ is positive definite;
2. Σ admits a *Cholesky decomposition*, that is there is a non-singular square matrix σ such that $\Sigma = \sigma\sigma'$;
3. Σ has I positive eigenvalues (counting multiplicity).

Definition 11 (Multivariate Normal)**.** The $\mathbb{R}^I$-valued random variable $X = (X_1,\dots,X_I) \sim N(0,\Sigma)$ for Σ positive definite has the joint probability density function

$$\phi_X(x) = \frac{1}{(2\pi)^{I/2}(\det\Sigma)^{1/2}}e^{-x\Sigma^{-1}x'/2}, x = (x_1,\dots,x_I). \tag{A.33}$$

For any $\mu = (\mu_1,\dots,\mu_I)$, $X+\mu \sim N(\mu,\Sigma)$.

One can show that

1. The jth marginal X_j of $X \sim N(0,\Sigma)$ is $N(0,\Sigma_{jj})$ distributed.
2. if $X \sim N(0,\Sigma)$ then the variance-covariance matrix of X is the matrix with components
$$\mathrm{Cov}(X_i,X_j) := E[X_iX_j] - E[X_i]E[X_j] = \Sigma_{ij}$$
3. if $Y \sim N(0,I)$ then $X := \sigma Y \sim N(0,\Sigma)$;
4. If $Y \sim N(0,\Sigma), X \sim N(0,\Gamma)$ are independent, then $X+Y \sim N(0,\Sigma+\Gamma)$.
5. If $Y \sim N(0,\Sigma)$, $X \sim N(0,1)$ are independent and $\alpha \in \mathbb{R}^I$ then $Y+\alpha X \sim N(0,\Sigma+\alpha\alpha')$.

A.12 Exercises

Exercise 78. (Quick derivation of Black-Scholes call option formula) We wish to evaluate the expectation $E^Q[e^{-rT}(S_T-K)\mathbf{1}_{\{S_T>K\}}]$, based on the model with $S_t = S_0 e^{\sigma W_t+(r-\sigma^2/2)t}$.

1. Show that the second term is simply $-Ke^{-rT}N[d_2]$ where $d_2 = N^{-1}[Q[S_T > K]]$.
2. Apply the Girsanov theorem with measure change $d\tilde{Q}/dQ = Z_T = S_T/E^Q[S_T]$ to show that the first term equals $S_0 N[d_1]$ where $d_1 = N^{-1}[\tilde{Q}[S_T > K]]$.

Exercise 79. Let $X^\mu_t, \bar{W}^\mu_t$ be drifting BM and its running maximum. Compute the following structure functions for $b > 0$:

1. $$E[\mathbf{1}_{\{\bar{W}^\mu_t \le b\}} e^{-uX^\mu_t}]$$
2. $$E[\mathbf{1}_{\{\bar{W}^\mu_t \le b\}} e^{-uX^\mu_t} | \mathcal{F}_s], \quad s \le t$$
3. $$E[\mathbf{1}_{\{\bar{W}^\mu_t \ge b\}} \delta(x + X^\mu_t - b)]$$

Exercise 80. A down and out call option with strike K, maturity T and barrier $B < K$ is valued by the formula $E^Q[e^{-rT}(S_T - K)\mathbf{1}_{\{S_T>K\}}\mathbf{1}_{\{\min_{s\le T} S_t>B\}}]$. Follow the steps of the previous exercise to prove that this is given in terms of the first passage function (A.30) by

$$S_0 \mathrm{FP}(\frac{1}{\sigma}\log(S_0/K), \frac{1}{\sigma}\log(S_0/B), -(r+\sigma^2/2)T/\sigma, T)$$
$$-Ke^{-rT}\mathrm{FP}(\frac{1}{\sigma}\log(S_0/K), \frac{1}{\sigma}\log(S_0/B), -(r-\sigma^2/2)T/\sigma, T) \quad \text{(A.34)}$$

Bibliography

[1] L. Andersen, J. Sidenius, and S. Basu. All your hedges in one basket. *Risk*, 16:67–72, 2003.

[2] A. Antonov, S. Mechkov, and T. Misirpashaev. Analytical techniques for synthetic CDOs and credit default risk measures. Numerix Preprint www.defaultrisk.com/pp_crdrv_77.htm, 2005.

[3] F. Black and J. C. Cox. Valuing corporate securities. *J. Finance*, 31:351–367, 1976.

[4] D. Duffie, L. Saita, and K. Wang. Multi-period corporate default prediction with stochastic covariates. Working paper, 2006.

[5] D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.

[6] D. Duffie and K. Singleton. *Credit risk: pricing, measurement and management.* Princeton University Press, 2003.

[7] R. J. Elliott and P. E. Kopp. *Mathematics of Financial Markets.* Springer-Verlag, Berlin Heidelberg New York, 1999.

[8] R. Geske. The valuation of corporate liabilities as compound options. *J. Fin. Quant. Anal.*, 12:541–552, 1977.

[9] G. M. Gupton, C. C. Finger, and M. Bhatia. Creditmetrics technical document. *J.P. Morgan & Co. Incorporated*, pages 1–230, 1997.

[10] D. C. Heath, R. A. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates. *Econometrica*, 60:77–105, 1992.

[11] J. Hull and A. White. Valuation of a CDO and an nth to default CDS without Monte Carlo simulation. *Jour. of Derivatives*, pages 8–23, 2004.

[12] I. Iscoe and A. Kreinin. Pricing and valuation of synthetic CDOs. Working Paper www.defaultrisk.com/pp_model_60.htm, 2004.

[13] J. James and N. Webber. *Interest Rate Modelling.* J. Wiley, Chichester, 2000.

[14] R. Jarrow, D. Lando, and S. Turnbull. A Markov model for the term structure of credit risk spreads. *Review of Financial Studies*, pages 481–523, 1997.

[15] R. A. Jarrow and S. M. Turnbull. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50:53–85, 1995.

[16] D. Lando. On Cox processes and risky bonds. *Review of Derivatives Research*, 2:99–120, 1998.

[17] J.-P. Laurent and J. Gregory. Basket default swaps, CDOs and factor copulas. *Jour. of Risk*, page 21, 2003.

[18] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 2005.

[19] R. C. Merton. On the pricing of corporate debt: the risk structure of interest rates. *J. Finance*, 29:449–470, 1974.

[20] B. Øksendal. *Stochastic Differential Equations, Fifth Edition*. Springer-Verlag Berlin, 1998.

[21] P. Protter. *Stochastic Integration and Differential Equations*. Springer Verlag, Berlin Heidelberg New York, 1990.

[22] J.M. Steele. *Stochastic Calculus and Financial Applications*. Springer-Verlag, Berlin Heidelberg New York, 2000.

[23] Oldrich Vasicek. Probability of loss on loan portfolio. KMV Working Paper `www.defaultrisk.com/pp_model_60.htm`, 1987.

[24] D. Wang, S. T. Rachev, and F. J. Fabozzi. Pricing tranches of a CDO and a CDS index: Recent advances and future research. working paper available at, 2006.

[25] J. P. Yang, T. R. Hurd, and X. P. Zhang. Saddlepoint approximation method for pricing CDOs. *Jour. Comp. Finance*, pages 1–20, 2006.

[26] C. Zhou. The term structure of credit spreads with jump risk. *J. Bus. and Finan.*, 25:2015–2040, 2001.

Index

$\mathcal{H}$-condition, 59

admissible trading strategy, 122
arbitrage, 122

bankruptcy, 13
bond
  zero-coupon, 9

compensator, 48
complete market, 122
contingent claim, 122
counting process, 54
credit default swap, 16
credit risk premium, 65
credit spread, 14
CreditMetrics, 90

default correlation function , 88
default intensity, 49
default time, 9
default time correlation, 88
distance to default, 46
diversity score, 89

EDF, 46

first passage model, 44
floating rate note, 15
forward default rate process, 48
forward rate
  continuously compounded, 10
  default risky, 13
  instantaneous, 12
  simply compounded, 11
forward rate agreement, 15
forward spread, 14
Fundamental Theorem of Arbitrage Pricing
  First, 122
  Second, 122

generalized Lando formula, 72

leverage ratio, 50
LIBOR, 11
loss given default, 43

Markov generator, 118
martingale invariance property, 59
money-market account, 12

numeraire, 123

par floater, 15
Poisson process, 54

recovery fraction, 70
replicable claim, 122

self-financing, 122
solvency, 13
survival probability, 48
swaption, 19

time to maturity, 9

yield
  continuously compounded, 11
  simply compounded, 11
yield spread, 14