:s the most comprehensive treatment of the theoretical concepts and mod- of quantitative risk management. Whether you are a financial risk analyst, student of quantitative finance, *Quantitative Risk Management* gives you the eed to solve real-world problems.

atest advances in the field, *Quantitative Risk Management* covers the methods ıd operational risk modelling. It places standard industry approaches on a and explores key concepts such as loss distributions, risk measures and risk ocation principles. The book's methodology draws on diverse quantitative thematical finance and statistics to econometrics and actuarial mathemat- : throughout is the need to satisfactorily address extreme outcomes and the isk drivers. Proven in the classroom, the book also covers advanced topics s.

l expanded to reflect developments in the field since the financial crisis

chapters to facilitate teaching and learning

ed coverage of Solvency II and insurance risk management and extended lit risk, including counterparty credit risk and CDO pricing

napter on market risk and new material on risk measures and risk aggregation

PRAISE FOR THE PREVIOUS EDITION:

Management can be highly recommended to anyone looking for an excel- most important techniques and tools used in this rapidly growing field."

ES, *Risk*

Management is highly recommended for financial regulators. The statis- atical tools facilitate a better understanding of the strengths and weak- ange of advanced risk-management concepts and models, while the focus enhances the publication's value to banking and insurance supervisors."

ESTEIN, *Financial Regulator*

npendium of the statistical arrows that should be in any quantitative risk It includes extensive discussion of dynamic volatility models, extreme ulas and credit risk. Academics, PhD students and quantitative practi- ıany new and useful results in this important volume."

IGLE III, 2003 Nobel Laureate in Economic Sciences, Michael Armellino Professor ıt of Financial Services at New York University's Stern School of Business

EIL is professor of actuarial mathematics and statistics at Heriot-Watt University **ER FREY** is professor of mathematics and finance at the Vienna University of Eco- ;. **PAUL EMBRECHTS** is professor of mathematics at the Swiss Federal Institute of

PRINCETON SERIES IN FINANCE
Darrell Duffie and Stephen Schaefer, Series Editors

N

Not for sale in South Asia
Jacket art © Michael D. Brown/Shutterstock.

ISBN: 978-0-691-16627-8
9 780691 166278 90000

MCNEIL
FREY
EMBRECHTS

QUANTITATIVE RISK MANAGEMENT

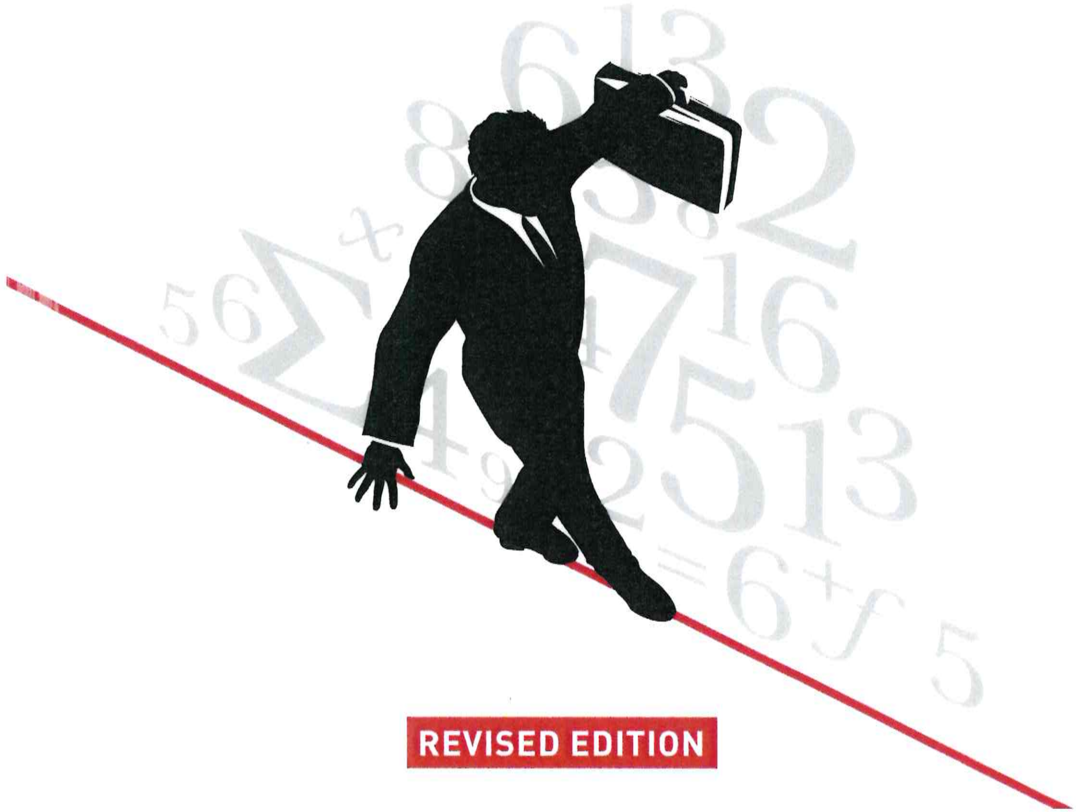
PRINCETON

PRINCETON SERIES IN FINANCE

QUANTITATIVE RISK MANAGEMENT

CONCEPTS, TECHNIQUES AND TOOLS

Alexander J. McNeil, Rüdiger Frey and Paul Embrechts



REVISED EDITION

10
Credit Risk

Credit risk is the risk of a loss arising from the failure of a counterparty to honour its contractual obligations. This subsumes both default risk (the risk of losses due to the default of a borrower or a trading partner) and downgrade risk (the risk of losses caused by a deterioration in the credit quality of a counterparty that translates into a downgrading in some rating system). Credit risk is omnipresent in the portfolio of a typical financial institution. To begin with, the lending and corporate bond portfolios are obviously affected by credit risk. Perhaps less obviously, credit risk accompanies any over-the-counter (OTC, i.e. non-exchange-guaranteed) derivative transaction such as a swap, because the default of one of the parties involved may substantially affect the actual pay-off of the transaction. Moreover, there is a specialized market for credit derivatives, such as credit default swaps, in which financial institutions are active players. Credit risk therefore relates to the core activities of most banks. It is also highly relevant to insurance companies, who are exposed to substantial credit risk in their investment portfolios and counterparty default risk in their reinsurance treaties.

The management of credit risk at financial institutions involves a range of tasks. To begin with, an enterprise needs to determine the capital it should hold to absorb losses due to credit risk, for both regulatory and economic capital purposes. It also needs to manage the credit risk on its balance sheet. This involves ensuring that portfolios of credit-risky instruments are well diversified and that portfolios are optimized according to risk–return considerations. The risk profile of the portfolio can also be improved by hedging risk concentrations with credit derivatives or by transferring risk to investors through securitization. Moreover, institutions need to manage their portfolio of traded credit derivatives. This involves the tasks of pricing, hedging and managing collateral for such trades. Finally, financial institutions need to control the counterparty credit risk in their trades and contracts with other institutions. In fact, in the aftermath of the 2007–9 financial crisis, counterparty risk management became one of the most important issues for financial institutions.

With these tasks in mind we have split our treatment of credit risk into four chapters. In the present chapter we establish the foundations for the analysis of credit risk. We introduce the most common credit-risky instruments (Section 10.1), discuss various measures of credit quality (Section 10.2) and present models for the credit risk of a single firm (Sections 10.3–10.6). Moreover, we study basic single-name credit derivatives such as credit default swaps.

Chapters 11 and 12 are concerned with portfolio models, and the crucial issue of dependence between defaults comes to the fore. Chapter 11 treats one-period models with a view to capital adequacy and credit risk management issues for portfolios of largely non-traded assets. Chapter 12 deals with properties of portfolio credit derivatives such as collateralized debt obligations; moreover, we discuss the valuation of these products in standard copula models. Finally, Chapter 17 is concerned with more advanced fully dynamic models of portfolio credit risk. This chapter is also the natural place for a detailed discussion of counterparty credit risk because a proper analysis requires dynamic multivariate credit risk models.

Credit risk models can be divided into *structural* or *firm-value models* on the one hand and *reduced-form models* on the other. Broadly speaking, in a structural model default occurs when a stochastic variable (or, in dynamic models, a stochastic process), generally representing an asset value, falls below a threshold, generally representing liabilities. In reduced-form models the precise mechanism leading to default is left unspecified and the default time of a firm is modelled as a non-negative rv, whose distribution typically depends on economic covariables. In this chapter we treat structural models in Section 10.3, simple reduced-form models with deterministic hazard rates in Section 10.4 and more advanced reduced-form models in Sections 10.5 and 10.6.

10.1 Credit-Risky Instruments

In this section we give an overview of the universe of credit-risky instruments, starting with the simplest examples of loans and bonds. We include discussion of the counterparty credit risk in OTC derivatives trades and we also describe some of the more common modern credit derivative products. In what follows we often use the generic term *obligor* for the borrower, bond issuer, trading partner or counterparty to whom there is a credit exposure. The name stems from the fact that in all cases the obligor has a contractual obligation to make certain payments under certain conditions.

10.1.1 Loans

Loans are the oldest credit-risky "instruments" and come in a myriad of forms. It is common to categorize them according to the type of obligor into *retail* loans (to individuals and small or medium-sized companies), *corporate* loans (to larger companies), *interbank* loans and *sovereign* loans (to governments). In each of these categories there are likely to be a number of different lending products. For example, retail customers may borrow money from a bank using mortgages against property, credit cards and overdrafts.

The common feature of most loans is that a sum of money, known as the *principal*, is advanced to the borrower for a particular term in exchange for a series of defined interest payments, which may be at fixed or floating interest rates. At the end of the term the borrower is required to pay back the principal.

A useful distinction to make is between *secured* and *unsecured* lending. If a loan is secured, the borrower has pledged an asset as collateral for the loan. A prime example is a mortgage, where the collateral is a property. In the event that the borrower is unable to fulfill its obligation to make interest payments or repay the principal, a situation that is termed *default*, the lender may take possession of the asset. In this way the loss in the event of default may be partly mitigated and money may be recovered by selling the asset. In an unsecured loan the lender has no such claim on a collateral asset and recoveries in the event of default may be a smaller fraction of the so-called *exposure*, which is the value of the outstanding principal and interest payments.

Unlike bonds, which are publicly traded securities, loans are private agreements between the borrower and the lender. Hence there is a wide variety of different loan contracts with different legal features. This makes loans difficult to value under fair-value principles. Book value is commonly used, and where fair-value approaches are applied these mostly fall under the heading of level 3 valuation (see Section 2.2.2).

10.1.2 Bonds

Bonds are publicly traded securities issued by companies and governments that allow the issuer to raise funding on financial markets. Bonds issued by companies are called *corporate bonds* and bonds issued by governments are known as *treasuries*, *sovereign bonds* or, particularly in the UK, *gilts* (gilt-edged securities).

The structure of the pay-offs is akin to that of a loan. The security commits the bond issuer (borrower) to make a series of interest payments to the bond buyer (lender) and pay back the principal at a fixed maturity. The interest payments, or coupons, may be *fixed* at the issuance of the bond (so-called fixed-coupon bonds). Alternatively, there are also bonds where the interest payments vary with market rates (so-called *floating-rate notes*). The reference for the floating rate is often LIBOR (the London Interbank Offered Rate). There are also *convertible bonds*, which allow the purchaser to convert them into shares of the issuing company at predetermined time points. These typically offer lower rates than conventional corporate bonds because the investor is being offered the option to participate in the future growth of the company.

A bondholder is subject to a number of risks, particularly interest-rate risk, spread risk and default risk. As for loans, default risk is the risk that promised coupon and principal payments are not made. Historically, government bonds issued by developed countries have been considered to be default free; for obvious reasons, after the European debt crisis of 2010–12, this notion was called into question.

Spread risk is a form of market risk that refers to changes in *credit spreads*. The credit spread of a defaultable bond measures the difference in the yield of the bond and the yield of an equivalent default-free bond (see Section 10.3.2 for a formal definition of credit spreads). An increase in the spread of a bond means that the market value of the bond falls, which is generally interpreted as indicating that the financial markets perceive an increased default risk for the bond.

10.1.3 Derivative Contracts Subject to Counterparty Risk

A significant proportion of all derivative transactions is carried out over the counter, and there is no central clearing counterparty such as an organized exchange that guarantees the fulfilment of the contractual obligations. These trades are therefore subject to the risk that one of the contracting parties defaults during the transaction, thus affecting the cash flows that are actually received by the other party. This risk, known as *counterparty credit risk*, received a lot of attention during the financial crisis of 2007–9, as some of the institutions heavily involved in derivative transactions experienced worsening credit quality or—in the case of Lehman Brothers—even a default event. Counterparty credit risk management is now a key issue for all financial institutions and is the focus of many new regulatory developments.

In order to illustrate the challenges in measuring and managing counterparty credit risk, we consider the example of an interest swap. This is a contract where two parties A and B agree to exchange a series of interest payments on a given nominal amount of money for a given period. For concreteness assume that A receives payments at a fixed interest rate and makes floating payments at a rate equal to the three-month LIBOR.

Suppose now that A defaults at time τ_A before the maturity of the contract, so that the contract is settled at that date. The consequences will depend on the value of the remaining interest payments at that point in time. If interest rates have risen relative to their value at inception of the contract, the fixed interest payments have decreased in value so that the value of the swap contract has increased for B. Since A is no longer able to fulfill its obligations, its default constitutes a loss for B; the exact size of the loss will depend on the term structure of interest rates at the default time τ_A. On the other hand, if interest rates have fallen relative to their value at $t = 0$, the fixed payments have increased in value so that the swap has a negative value for B. At settlement, B will still have to pay the value of the contract into the bankruptcy pool, so that there is no upside for B in A's default. If B defaults first, the situation is reversed: falling rates lead to a counterparty-risk-related loss for A. This simple example illustrates two important points: the size of the counterparty credit exposure is not known a priori, and it is not even clear who has the credit exposure.

The management of counterparty risk raises a number of issues. First, counterparty risk has to be taken into account in pricing and valuation. This has led to various forms of credit value adjustment (CVA). Second, counterparty risk needs to be controlled using risk-mitigation techniques such as netting and collateralization. Under a netting agreement, the value of all derivatives transactions between A and B is computed and only the aggregated value is subject to counterparty risk; since offsetting transactions cancel each other out, this has the potential to reduce counterparty risk substantially. Under a collateralization agreement, the parties exchange collateral (cash and securities) that serve as a pledge for the receiver. The value of the collateral is adjusted dynamically to reflect changes in the value of the underlying transactions.

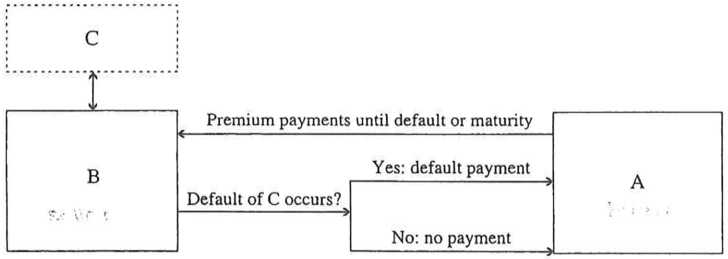


Figure 10.1. The basic structure of a CDS. Firm C is the reference entity, firm A is the protection buyer, and firm B is the protection seller.

The proper assessment of counterparty risk requires a joint modelling of the default times of the two counterparties (A and B) and of the price dynamics of the underlying derivative contract. For that reason we defer the detailed discussion of this topic to Chapter 17.

10.1.4 Credit Default Swaps and Related Credit Derivatives

Credit derivatives are securities that are primarily used for the hedging and trading of credit risk. In contrast to the products considered so far, the promised pay-off of a credit derivative is related to credit events affecting one or more firms. Major participants in the market for credit derivatives are banks, insurance companies and investment funds. Retail banks are typically net buyers of protection against credit events; other investors such as hedge funds and investment banks often act as both sellers and buyers of credit protection.

Credit default swaps. Credit default swaps (CDSs) are the workhorses of the credit derivatives market, and the market for CDSs written on larger corporations is fairly liquid; some numbers on the size of the market are given in Notes and Comments. The basic structure of a CDS is depicted in Figure 10.1. A CDS is a contract between two parties, the *protection buyer* and the *protection seller*. The pay-offs are related to the default of a reference entity (a financial firm or sovereign issuing bonds).

If the reference entity experiences a default event before the maturity date T of the contract, the protection seller makes a default payment to the protection buyer, which mimics the loss due to the default of a bond issued by the reference entity (the reference asset); this part of a CDS is called the *default payment leg*. In this way the protection buyer has acquired financial protection against the loss on the reference asset he would incur in case of a default. As compensation, the protection buyer makes periodic premium payments (typically quarterly or semiannually) to the protection seller (the *premium payment leg*); after the default of the reference entity, premium payments stop. There is no initial payment. The premium payments are quoted in the form of an annualized percentage x^* of the notional value of the reference asset; x^* is termed the (fair or market-quoted) *CDS spread*. For a mathematical description of the payments, see Section 10.4.4.

There are a number of technical and legal issues in the specification of a CDS. In particular, the parties have to agree on the precise definition of a default event and on a procedure to determine the size of the default payment in case a default event of the reference entity occurs. Due to the efforts of bodies such as the International Swaps and Derivatives Association (ISDA), some standardization of these issues has taken place.

Investors enter into CDS contracts for various reasons. To begin with, bond investors with a large credit exposure to the reference entity may buy CDS protection to insure themselves against losses due to the default of a bond. This may be easier than reducing the original bond position because CDS markets are often more liquid than bond markets. Moreover, CDS positions are quickly settled.

CDS contracts are also held for speculative reasons. In particular, so-called *naked* CDS positions, where the protection buyer does not own the bond, are often assumed by investors who are speculating on the widening of the credit spread of the reference entity. These positions are similar to short-selling bonds issued by the reference entity. Note that, in contrast to insurance, there is no requirement for the protection buyer to have *insurable interest*, that is, to actually own a bond issued by the reference entity. The speculative motive for holding CDSs is at least as important as the insurance motive.

There has been some debate about the risks of the CDS market, particularly with respect to the large volume of naked positions and whether or not these should be limited. By taking naked CDS positions speculators can depress the prices of the bonds issued by the reference entity so that default becomes a self-fulfilling prophecy. The debate about the pros and cons of limiting naked CDS positions is akin to the debate about the pros and cons of limiting short selling on equity markets.

A CDS is traded over the counter and is not guaranteed by a clearing house. A CDS position can therefore be subject to a substantial amount of counterparty risk, particularly if a trade is backed by insufficient collateral. A case in point arose during the credit crisis when AIG, which had sold many protection positions, had to be bailed out by the US government to prevent the systemic consequences of allowing it to default on its CDS contracts. There is concern that CDS markets have created a new form of dependency across financial institutions so that the default of one large (systemically important) institution could create a cascade of defaults across the financial sector due to counterparty risk.

On the other hand, CDSs are useful risk-management tools. Because of the liquidity of CDS markets, CDSs are the natural underlying security for many more complex credit derivatives. Models for pricing portfolio-related credit derivatives are usually calibrated to quoted CDS spreads. With improved collateral management in CDS markets it has been argued that the potential for CDS markets to create large-scale default contagion has been substantially reduced (see Notes and Comments).

Credit-linked notes. A credit-linked note is a combination of a credit derivative and a coupon bond that is sold as a fixed package. The coupon payments (and sometimes also the repayment of the principal) are reduced if a third party (the reference entity) experiences a default event during the lifetime of the contract, so the buyer of a credit-linked note is providing credit protection to the issuer of the note.

Credit-linked notes are issued essentially for two reasons. First, from a legal point of view, a credit-linked note is treated as a fixed-income investment, so that investors who are unable to enter into a transaction involving credit derivatives directly (such as life insurance companies) may nonetheless sell credit protection by buying credit-linked notes. Second, an investor buying a credit-linked note pays the price up front, so that the protection buyer (the issuer of the credit-linked note) is protected against losses caused by the default of the protection seller.

10.1.5 PD, LGD and EAD

Regardless of whether we make a loan, buy a defaultable bond, engage in an OTC derivatives transaction, or act as protection seller in a CDS, the risk of a credit loss is affected by three, generally related, quantities: the *exposure at default* (EAD), the *probability of default* (PD) and the *loss given default* (LGD) or, equivalently, the size of the recovery in the event of default. They are key inputs to the Basel formula in the internal-ratings-based (IRB) approach to determining capital requirements for credit-risky portfolios, so it is important to consider them.

Exposure at default. If we make a loan or buy a bond, our exposure is relatively easy to determine, since it is mainly the principal that is at stake. However, there is some additional uncertainty about the value of the interest payments that could be lost. A further source of exposure uncertainty is due to the widespread use of credit lines. Essentially, a credit line is a ceiling up to which a corporate client can borrow money at given terms, and it is up to the borrower to decide which part of the credit line he actually wants to use. For OTC derivatives, the counterparty risk exposure is even more difficult to quantify, since it is a stochastic variable depending on the unknown time at which a counterparty defaults and the evolution of the value of the derivative up to that point; a case in point is the example of an interest rate swap discussed in Subsection 10.1.3.

In practice, the concept that is used to describe exposure is exposure at default or EAD, which recognizes that the exposure for many instruments will depend on the exact default time. In counterparty credit risk the use of collateral can also reduce the exposure and thus mitigate losses.

Probability of default. When measuring the risk of losses over a fixed time horizon, e.g. one year, we are particularly concerned with estimating the probability that obligors default by the time horizon, a quantity known to practitioners as the probability of default, or PD. The PD is related to the credit quality of an obligor, and Sections 10.2 and 10.3 discuss some of the models that are used to quantify default risk. For instruments where the loss is dependent on the exact timing of default, e.g. OTC derivatives with counterparty risk, the risk of default is described by the whole distribution of possible default times and not just the probability of default by a fixed horizon.

Loss given default. In the event of default, it is unlikely that the entire exposure is lost. For example, when a mortgage holder defaults on a residential mortgage, and there is no realistic possibility of restructuring the debt, the lender can sell the property (the collateral asset) and the proceeds from the sale will make good some of the lost principal. Similarly, when a bond issuer goes into administration, the bondholders join the group of creditors who will be partly recompensed for their losses by the sale of the firm's assets.

Practitioners use the term loss given default, or LGD, to describe the proportion of the exposure that is actually lost in the event of default, or its converse, the recovery, to describe the amount of the exposure that can be recovered through debt restructuring and asset sales.

Dependence of these quantities. It is important to realize that EAD, PD and LGD are dependent quantities. While it is common to attempt to model them in terms of independent random variables, it is unrealistic to do so. For example, in a period of financial distress, when PDs are high, the asset values of firms are depressed and firms are defaulting, recoveries are likely to be correspondingly low, so that there is positive dependence between PDs and LGDs. This will be discussed further in 11.2.3.

Notes and Comments

For further reading on loans and loan pricing we refer to Benzschawel, Dagraca and Fok (2010). For an overview of bonds see Sharpe, Alexander and Bailey (1999).

To get an idea of the size of the CDS market, note that the nominal value (gross notional amount) of the market stood at approximately \$60 trillion by the end of 2007, before coming down to a still considerable amount of approximately \$25 trillion by the end of 2012. In 2013 the net notional amount was of the order of \$2 trillion. For comparison, by the end of 2012 world GDP stood at roughly \$80 trillion. To give an example of the size of the speculative market in CDSs, Cont (2010) reports that "when it filed for bankruptcy on September 14, 2008, Lehman Brothers had \$155 billion of outstanding debt, but more than \$400 billion notional value of CDS contracts had been written with Lehman as reference entity". A good discussion of the role of such credit derivatives in the credit crisis is given in Stulz (2010). The effect of improved collateral management for CDSs on the risk of large-scale contagion in CDS markets is addressed in Brunnermeier, Clerc and Scheicher (2013).

In this brief introduction we have discussed a few essential features of credit derivatives but have omitted the rather involved regulatory, legal and accounting issues related to these instruments. Readers interested in these topics are referred to the paper collections edited by Gregory (2003) and Perraudin (2004), in which pricing issues are also discussed. An excellent treatment of credit derivatives at textbook level is Schönbucher (2003). For a discussion of credit derivatives from the viewpoint of financial engineering we refer to Neftci (2008).

10.2 Measuring Credit Quality

There are various ways of quantifying the credit quality or, equivalently, the default risk of obligors but, broadly speaking, these approaches may be divided into two philosophies. On the one hand, credit quality can be described by a credit rating or credit score that is based on empirical data describing the borrowing and repayment history of the obligor, or of similar obligors. On the other hand, for obligors whose equity is traded on financial markets, prices can be used to infer the market's view of the credit quality of the obligor. This section is devoted to the first philosophy, and market-implied measures of credit quality are treated in the context of structural models in Section 10.3.

Credit ratings and credit scores fulfill a similar function—they both allow us to order obligors according to their credit risk and map that risk to an estimate of default probability. Credit ratings tend to be expressed on an ordered categorical scale, whereas credit scores are often expressed in terms of points on a metric scale. The task of rating obligors, particularly large corporates or sovereigns, is often outsourced to a rating agency such as Moody's or Standard & Poor's (S&P); proprietary rating systems internal to a financial institution can also be used. In the S&P rating system there are seven pre-default rating categories, labelled AAA, AA, A, BBB, BB, B, CCC, with AAA being the highest rating and CCC the lowest rating; Moody's uses nine pre-default rating categories and these are labelled Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C. A finer alpha-numeric system is also used by both agencies.

Credit scores are traditionally used for retail customers and are based on so-called scorecards that banks develop through extensive statistical analyses of historical data. The basic idea is that default risk is modelled as a function of demographic, behavioural and financial covariates that describe the obligor. Using techniques such as logistic regression these covariates are weighted and combined into a score.

10.2.1 Credit Rating Migration

In the credit-migration approach each firm is assigned to a credit-rating category at any given time point. The probability of moving from one credit rating to another over a given risk horizon (typically one year) is then specified. Transition probabilities are typically presented in the form of a matrix; an example from Moody's is presented in Table 10.1. Transition matrices are estimated from historical default data, and standard statistical methods used for this purpose are discussed in Section 10.2.2

In the credit-migration approach we assume that the current credit rating completely determines the default probability, so that this probability can be read from the transition matrix. For instance, if we use the transition matrix presented in Table 10.1, we obtain a one-year default probability for an A-rated company of 0.06%, whereas the default probability of a Caa-rated company is 13.3%. In practice, a correction to the figures in Table 10.1 would probably be undertaken to account for rating withdrawals: that is, transitions to the WR state. The simplest correction would be to divide the first nine probabilities in each row of the table by one minus the final probability in that row; this implicitly assumes that the act of rating withdrawal contains no information about the likelihood of upgrade, downgrade or default of an obligor.

Table 10.1. Probabilities of migrating from one rating quality to another within one year. "WR" represents the proportion of firms that were no longer rated at the end of the year, for various reasons including takeover by another company. Source: Ou (2013, Exhibit 26).

Initial rating	Rating at year-end (%): Aaa	Aa	A	Baa	Ba	B	Caa	Ca–C	Default	WR
Aaa	87.20	8.20	0.63	0.00	0.03	0.00	0.00	0.00	0.00	3.93
Aa	0.91	84.57	8.43	0.49	0.06	0.02	0.01	0.00	0.02	5.48
A	0.06	2.48	86.07	5.47	0.57	0.11	0.03	0.00	0.06	5.13
Baa	0.039	0.17	4.11	84.84	4.05	7.55	1.63	0.02	0.17	5.65
Ba	0.01	0.05	0.35	5.52	75.75	7.22	0.58	0.07	1.06	9.39
B	0.01	0.03	0.11	0.32	4.58	73.53	5.81	0.59	3.85	11.16
Caa	0.01	0.02	0.02	0.12	0.38	8.70	61.71	3.72	13.34	12.00
Ca–C	0.00	0.00	0.00	0.00	0.40	2.03	9.38	35.46	37.93	14.80

Table 10.2. Average cumulative default rates (%). Source: Ou (2013, Exhibit 33).

Initial rating	Term: 1	2	3	4	5	10	15
Aaa	0.00	0.01	0.01	0.04	0.11	0.50	0.93
Aa	0.02	0.07	0.14	0.26	0.38	0.92	1.75
A	0.06	0.20	0.41	0.63	0.87	2.48	4.26
Baa	0.18	0.50	0.89	1.37	1.88	4.70	8.62
Ba	1.11	3.08	5.42	7.93	10.18	19.70	29.17
B	4.05	9.60	15.22	20.13	24.61	41.94	52.22
Caa–C	16.45	27.87	36.91	44.13	50.37	69.48	79.18

Rating agencies also produce cumulative default probabilities over larger time horizons. In Table 10.2 we reproduce Moody's cumulative default probabilities for companies with a given current credit rating. For instance, according to this table the probability that a company whose current credit rating is Baa defaults within the next four years is 1.37%. These cumulative default probabilities have been estimated directly from default data. Alternative estimates of multi-year default probabilities can be inferred from one-year transition matrices, as explained in more detail in the next section.

Remark 10.1 (accounting for business cycles). It is a well-established empirical fact that default rates tend to vary with the state of the economy, being high during recessions and low during periods of economic expansion (see Figure 10.2 for an illustration). Transition rates as estimated by S&P or Moody's, on the other hand, are historical averages over longer time horizons covering several business cycles. For instance, the transition rates in Table 10.1 have been estimated from rating-migration data over the period 1970–2012. Moreover, rating agencies focus on the

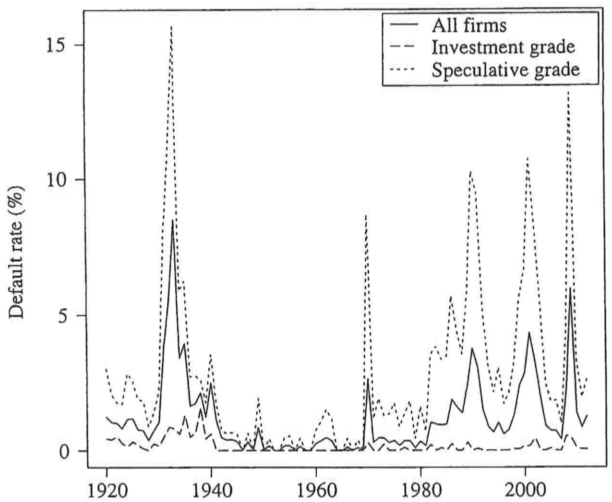


Figure 10.2. Moody's annual default rates from 1920 to 2012. Source for data: Ou (2013, Exhibit 30).

average credit quality "through the business cycle" when attributing a credit rating to a particular firm. The default probabilities from the credit-migration approach are therefore estimates for the average default probability, independent of the current economic environment. In some situations we are interested in "point-in-time" estimates of default probabilities reflecting the current macroeconomic environment, such as in the pricing of a short-term loan. In these situations adjustments to the long-term average default probabilities from the credit-migration approach can be made; for instance, we could use equity prices as an additional source of information, as is done in the public-firm EDF (expected default frequency) model discussed in Section 10.3.3.

10.2.2 Rating Transitions as a Markov Chain

Let (R_t) denote a discrete-time stochastic process defined at times $t = 0, 1, \dots$ that takes values in $S = \{0, 1, \dots, n\}$. The set S defines rating states of increasing creditworthiness, with 0 representing default. (R_t) models the evolution of an obligor's rating over time.

We will assume that (R_t) is a *Markov chain*. This means that conditional transition probabilities satisfy the Markov property

$$P(R_t = k \mid R_0 = r_0, R_1 = r_1, \dots, R_{t-1} = j) = P(R_t = k \mid R_{t-1} = j)$$

for all $t \geqslant 1$ and all $j, r_0, r_1, \dots, r_{t-2}, k \in S$. In words, the conditional probabilities of rating transitions, given an obligor's rating history, depend only on the previous rating $R_{t-1} = j$ at the last time point and not on the more distant history of how the obligor arrived at a rating state j at time $t-1$.

The Markov assumption for rating migrations has been criticized; there is evidence for both *momentum* and *stickiness* in empirical rating histories (see Lando and Skodeberg 2002). Momentum is the phenomenon by which obligors who have been recently downgraded to a particular rating are more likely to experience further downgrades than obligors who have had the same rating for a long time. Stickiness is the converse phenomenon by which rating agencies are initially hesitant to downgrade obligors until the evidence for credit deterioration is overwhelming. But despite these issues, the Markov chain assumption is very widely made, because it leads to tractable models with a well-understood theory and to natural estimators for transition probabilities.

The Markov chain is *stationary* if

$$P(R_t = k \mid R_{t-1} = j) = P(R_1 = k \mid R_0 = j)$$

for all $t \geqslant 1$ and all rating states j and k. In this case we can define the transition matrix to be the $(n+1) \times (n+1)$ matrix $\boldsymbol{P} = (p_{jk})$ with elements $p_{jk} = P(R_t = k \mid R_{t-1} = j)$ for any $t \geqslant 1$. Simple conditional probability arguments can be used to derive the *Chapman–Kolmogorov equations*, which say that for any $t \geqslant 2$, and any $j, k \in S$,

$$\begin{aligned} P(R_t = k \mid R_{t-2} = j) &= \sum_{l \in S} P(R_t = k \mid R_{t-1} = l) P(R_{t-1} = l \mid R_{t-2} = j) \\ &= \sum_{l \in S} p_{lk} p_{jl}. \end{aligned}$$

An implication of this is that the matrix of transition probabilities over two time steps is given by $\boldsymbol{P}^2 = \boldsymbol{P} \times \boldsymbol{P}$. Similarly, the matrix of transition probabilities over T time periods is $\boldsymbol{P}^T$. It is, however, not clear how we would compute a matrix of transition probabilities for a fraction of a time period. In fact, this requires the notion of a Markov chain in *continuous time*, which is discussed below.

We now turn to the problem of estimating $\boldsymbol{P}$. Suppose we observe, or are given information about, the ratings of companies at the time points $0, 1, \dots, T$. This information usually relates to a fluctuating population or cohort of companies, with only a few having complete rating histories throughout $[0, T]$: new companies may be added to the cohort at any time; some companies may default and leave the cohort; others may have their rating withdrawn. In the latter case we will assume that the withdrawal of rating occurs independently of the default or rating-migration risk of the company (which may not be true).

For $t = 0, \dots, T-1$ and $j \in S \setminus \{0\}$, let N_{tj} denote the number of companies that are rated j at time t and for which a rating is available at time $t+1$; let N_{tjk} denote the subset of those companies that are rated k at time $t+1$. Under the discrete-time, homogeneous Markovian assumption, independent multinomial experiments effectively take place at each time t. In each experiment the N_{tj} companies rated j can be thought of as being randomly allocated to the ratings $k \in S$ according to probabilities p_{jk} that satisfy $\sum_{k=0}^{n} p_{jk} = 1$.

In this framework the likelihood is given by

$$L((p_{jk}); (N_{tj}), (N_{tjk})) = \prod_{t=0}^{T-1}\left(\prod_{j=1}^{n}\left(N_{tj}!\prod_{k=0}^{n}\frac{p_{jk}^{N_{tjk}}}{N_{tjk}!}\right)\right),$$

and if this is maximized subject to the constraints that $\sum_{k=0}^{n} p_{jk} = 1$ for $j = 1, \dots, n$, we obtain the maximum likelihood estimator

$$\hat{p}_{jk} = \frac{\sum_{t=0}^{T-1} N_{tjk}}{\sum_{t=0}^{T-1} N_{tj}}. \qquad (10.1)$$

There are a number of drawbacks to modelling rating transitions as a discrete-time Markov chain. In practice, rating changes tend to take place on irregularly spaced dates. While such data can be approximated by a regularly spaced time series (or panel) of, say, yearly, quarterly or monthly ratings, there will be a loss of information in doing so. The discrete-time model described above would ignore any information about intermediate transitions taking place between two times t and $t+1$. For example, if an obligor is downgraded from A to BBB to BB over the course of the period $[t, t+1]$, this obligor will simply be recorded as migrating from A to BB and the information about transitions from A to BBB and BBB to BB will not be recorded. Moreover, the estimation procedure for a discrete-time chain tends to result in sparse estimates of transition matrices with quite a lot of zero entries. For example, if no transitions between AAA and default within a single time period are observed, then the probability of such a transition will be estimated to be zero. However, in reality such a transition is possible, if unlikely, and so its estimated probability of occurrence should not be zero.

It is thus more satisfactory to model rating transitions as a phenomenon in continuous time. In this case, transition probabilities are not modelled directly but are instead given in terms of transition rates. Intuitively, the relationship between transition rates and transition probabilities can be described as follows.

Assume that over any small time step of duration δt the probability of a transition from rating j to k is given approximately by $\lambda_{jk}\delta t$ for some constant $\lambda_{jk} > 0$, which is the *transition rate* between rating j and rating k. The probability of staying at rating j is given by $1 - \sum_{k \neq j} \lambda_{jk}\delta t$. If we define a matrix Λ to have off-diagonal entries λ_{jk} and diagonal entries $-\sum_{k\neq j}\lambda_{jk}$, we can summarize the implied transition probabilities for the small time step δt in the matrix $(I_{n+1} + \Lambda\delta t)$. We now consider transitions in the period $[0, t]$ and denote the corresponding matrix of transition probabilities by $\boldsymbol{P}(t)$. If we divide the time period into N small time steps of size $\delta t = t/N$ for N large, the matrix of transition probabilities can be approximated by

$$\boldsymbol{P}(t) \approx \left(I_{n+1} + \frac{\Lambda t}{N}\right)^N,$$

which converges, as $N \to \infty$, to the so-called matrix exponential of Λt:

$$\boldsymbol{P}(t) = \mathrm{e}^{\Lambda t}.$$

This formulation gives us a method of computing transition probabilities for any time horizon t in terms of the matrix Λ, the so-called *generator matrix*.

A Markov chain in continuous time with generator matrix Λ can be constructed in the following way. An obligor remains in rating state j for an exponentially distributed amount of time with parameter

$$\lambda_{jj} = \sum_{k\neq j}\lambda_{jk},$$

i.e. minus the diagonal element of the generator matrix. When a transition takes place the new rating is determined by a multinomial experiment in which the probability of a transition from state j to state k is given by $\lambda_{jk}/\lambda_{jj}$.

This construction also leads to natural estimators for the matrix Λ. Since λ_{jk} is the instantaneous rate of migrating from j to k, we can estimate it by

$$\hat{\lambda}_{jk} = \frac{N_{jk}(T)}{\int_0^T Y_j(t)\,\mathrm{d}t}, \qquad (10.2)$$

where $N_{jk}(T)$ is the total number of observed transitions from j to k over the time period $[0, T]$ and $Y_i(t)$ is the number of obligors with rating j at time t; the denominator therefore represents the total time spent in state j by all the companies in the data set. Note that this is the continuous-time analogue of the maximum likelihood estimator in (10.1); it is not surprising, therefore, that (10.2) can be shown to be the maximum likelihood estimator for the transition intensities of a homogenous continuous-time Markov chain.

Notes and Comments

There is a large literature on credit scoring, and useful starter references are Thomas (2009) and Hand and Henley (1997). In addition to the well-known commercial rating agencies there are now open rating systems. One example is the Credit Research Initiative at the Risk Management Institute of the National University of Singapore (see www.rmicri.org).

An alternative discussion of models based on rating migration is given in Chapters 7 and 8 of Crouhy, Galai and Mark (2001). Statistical approaches to the estimation of rating-transition matrices are discussed in Hu, Kiesel and Perraudin (2002) and Lando and Skodeberg (2002). The latter paper also shows that there is momentum in rating-transition data, which contradicts the assumption that rating transitions form a Markov chain. An example of an industry model based on credit ratings is CreditMetrics: see RiskMetrics Group (1997).

The literature on the statistical properties of rating transitions is surveyed extensively in Chapter 4 of Duffie and Singleton (2003). The maximum likelihood estimator of the infinitesimal generator of a continuous-time Markov chain is formally derived in Albert (1962). For further information on Markov chains we refer to the standard textbook by Norris (1997).

10.3 Structural Models of Default

In structural or firm-value models of default one postulates a mechanism for the default of a firm in terms of the relationship between its assets and liabilities. Typically, default occurs whenever a stochastic variable (or in dynamic models a stochastic process) generally representing an asset value falls below a threshold representing liabilities. The kind of thinking embodied in these models has been very influential in the analysis of credit risk and in the development of industry solutions, so that this is a natural starting point for a discussion of credit risk models. We begin with a detailed analysis of the seminal model of Merton (1974) (in Sections 10.3.1 and 10.3.2). Industry implementations of structural models are discussed in Section 10.3.3.

From now on we denote a generic stochastic process in continuous time by (X_t); the value of the process at time $t \geqslant 0$ is given by the rv X_t.

10.3.1 *The Merton Model*

The model proposed in Merton (1974) is the prototype of all firm-value models. Consider a firm whose asset value follows some stochastic process (V_t). The firm finances itself by *equity* (i.e. by issuing shares) and by *debt*. In Merton's model, debt consists of zero-coupon bonds with common maturity T; the nominal value of debt at maturity is given by the constant B. Moreover, it is assumed that the firm cannot pay out dividends or issue new debt.

The values at time t of equity and debt are denoted by S_t and B_t. Default occurs if the firm misses a payment to its debtholders, which in the Merton model can occur only at the maturity T of the bonds. At T we have to distinguish between two cases.

(i) $V_T > B$: the value of the firm's assets exceeds the nominal value of the liabilities. In that case the debtholders (the owners of the zero-coupon bonds) receive B, the shareholders receive the residual value $S_T = V_T - B$, and there is no default.

(ii) $V_T \leqslant B$: the value of the firm's assets is less than its liabilities and the firm cannot meet its financial obligations. In that case shareholders have no interest in providing new equity capital, as these funds would go immediately to the bondholders. They therefore let the firm go into default. Control over the firm's assets is passed on to the bondholders, who liquidate the firm and distribute the proceeds among themselves. Shareholders pay and receive nothing, so that we have $B_T = V_T$, $S_T = 0$.

Summarizing, we have the relationships

$$S_T = \max(V_T - B, 0) = (V_T - B)^+, \tag{10.3}$$
$$B_T = \min(V_T, B) = B - (B - V_T)^+. \tag{10.4}$$

Equation (10.3) implies that the value of the firm's equity at time T equals the pay-off of a European call option on V_T, while (10.4) implies that the value of the firm's debt at maturity equals the nominal value of the liabilities minus the pay-off of a European put option on V_T with exercise price equal to B.

This model is of course a stylized description of default. In reality, the structure of a company's debt is much more complex, so that default can occur on many different dates. Moreover, under modern bankruptcy code, default does not automatically imply bankruptcy, i.e. liquidation of a firm. Nonetheless, Merton's model is a useful starting point for modelling credit risk and for pricing securities subject to default.

Remark 10.2. The option interpretation of equity and debt is useful in explaining potential conflicts of interest between the shareholders and debtholders of a company. It is well known that, all other things being equal, the value of an option increases if the volatility of the underlying security is increased. Shareholders therefore have an interest in the firm taking on risky projects. Bondholders, on the other hand, have a short position in a put option on the firm's assets and would therefore like to see the volatility of the asset value reduced.

In the Merton model it is assumed that under the real-world or physical probability measure P the process (V_t) follows a diffusion model (known as the Black–Scholes model or geometric Brownian motion) of the form

$$\mathrm{d}V_t = \mu_V V_t \,\mathrm{d}t + \sigma_V V_t \,\mathrm{d}W_t \tag{10.5}$$

for constants $\mu_V \in \mathbb{R}$ (the drift of the asset value process), $\sigma_V > 0$ (the volatility of the asset value process), and a standard Brownian motion (W_t). Equation (10.5) can be solved explicitly, and it can be shown that

$$V_T = V_0 \exp((\mu_V - \tfrac{1}{2}\sigma_V^2)T + \sigma_V W_T).$$

Since $W_T \sim N(0, T)$, it follows that $\ln V_T \sim N(\ln V_0 + (\mu_V - \frac{1}{2}\sigma_V^2)T, \sigma_V^2 T)$. Under the dynamics (10.5), the default probability of the firm is readily computed. We have

$$P(V_T \leqslant B) = P(\ln V_T \leqslant \ln B) = \Phi\left(\frac{\ln(B/V_0) - (\mu_V - \frac{1}{2}\sigma_V^2)T}{\sigma_V\sqrt{T}}\right). \tag{10.6}$$

It may be deduced from (10.6) that the default probability is increasing in B, decreasing in V_0 and μ_V and, for $V_0 > B$, increasing in σ_V. All these properties are perfectly in line with economic intuition.

Figure 10.3 shows two simulated trajectories for the asset value process (V_t) for values $V_0 = 1$, $\mu_V = 0.03$ and $\sigma_V = 0.25$. Assuming that $B = 0.85$ and $T = 1$, one path is a default path, terminating at a value $V_T < B$, while the other is a non-default path.

10.3.2 *Pricing in Merton's Model*

In the context of Merton's model one can price securities whose pay-off depends on the value V_T of the firm's assets at T. Prime examples are the firm's debt, or, equivalently, the zero-coupon bonds issued by the firm, and the firm's equity. In our analysis of pricing in the context of the Merton model we make use of a few basic

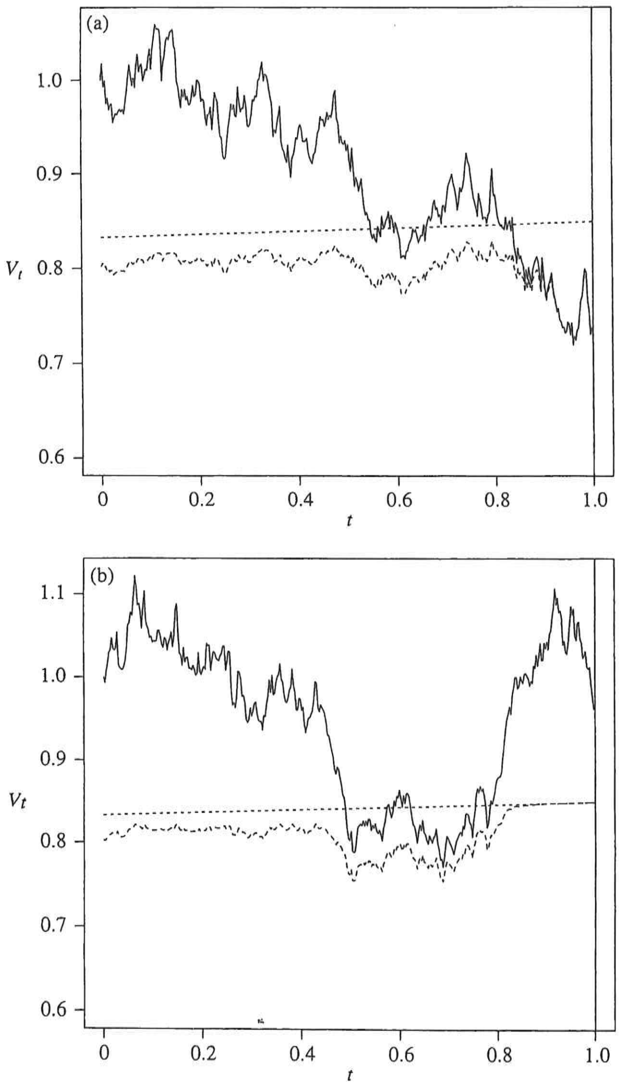


Figure 10.3. Illustration of (a) a default path and (b) a non-default path in Merton's model. The solid lines show simulated one-year trajectories for the asset value process (V_t) starting at $V_0 = 1$ with parameters $\mu_V = 0.03$ and $\sigma_V = 0.25$. Assuming that the debt has face value $B = 0.85$ and maturity $T = 1$ and that the interest rate is $r = 0.02$, the dotted curve shows the value of default-free debt $(Bp_0(t, T))$ while the dashed line shows the evolution of the company's debt B_t according to formula (10.12). The difference between the asset value V_t and the debt B_t is the value of equity S_t.

concepts from financial mathematics and stochastic calculus; references to useful texts in financial mathematics are given in Notes and Comments.

We make the following assumptions.

Assumption 10.3.

(i) *The risk-free interest rate is deterministic and equal to* $r \geqslant 0$.

(ii) *The firm's asset-value process* (V_t) *is independent of the way the firm is financed, and in particular it is independent of the debt level* B.

(iii) *The asset value* (V_t) *can be traded on a frictionless market, and the asset-value dynamics are given by the geometric Brownian motion (10.5).*

These assumptions merit some comments. First, the independence of (V_t) from the financial structure of the firm is questionable, because a very high debt level, and hence a high default probability, may adversely affect the ability of a firm to generate business, hence affecting the value of its assets. This is a special case of the indirect bankruptcy costs discussed in Section 1.4.2. Second, while there are many firms with traded equity, the value of the assets of a firm is usually neither completely observable nor traded. We come back to this issue in Section 10.3.3. For an example where (iii) holds, think of an investment company or trust that invests in liquidly traded securities and uses debt financing to leverage its position. In that case V_t corresponds to the value of the investment portfolio at time t, and this portfolio consists of traded securities by assumption.

Pricing of equity and debt. Consider a claim on the asset value of the firm with maturity T and pay-off $h(V_T)$, such as the firm's equity and debt in (10.3) and (10.4). Under Assumption 10.3, the fair value $f(t, V_t)$ of this claim at time $t \leqslant T$ can be computed using the risk-neutral pricing rule as the expectation of the discounted pay-off under the risk-neutral measure Q, that is,

$$f(t, V_t) = E^Q(\mathrm{e}^{-r(T-t)} h(V_T) \mid \mathcal{F}_t). \qquad (10.7)$$

According to (10.3), the firm's equity corresponds to a European call on (V_t) with exercise price B and maturity T. The risk-neutral value of equity obtained from (10.7) is therefore given simply by the Black–Scholes price C^{BS} of a European call. This yields

$$S_t = C^{\mathrm{BS}}(t, V_t; r, \sigma_V, B, T) := V_t \Phi(d_{t,1}) - B\mathrm{e}^{-r(T-t)}\Phi(d_{t,2}), \qquad (10.8)$$

where the arguments are given by

$$d_{t,1} = \frac{\ln V_t - \ln B + (r + \frac{1}{2}\sigma_V^2)(T-t)}{\sigma_V \sqrt{T-t}}, \qquad d_{t,2} = d_{t,1} - \sigma_V\sqrt{T-t}. \qquad (10.9)$$

Next we turn to the valuation of the risky debt issued by the firm. Since we assumed a constant interest rate r, the price at $t \leqslant T$ of a default-free zero-coupon bond with maturity T and a face value of one equals $p_0(t, T) = \mathrm{e}^{-r(T-t)}$. According to (10.4) we have

$$B_t = Bp_0(t, T) - P^{\mathrm{BS}}(t, V_t; r, \sigma_V, B, T), \qquad (10.10)$$

where $P^{BS}(t, V; r, \sigma_V, B, T)$ denotes the Black–Scholes price of a European put with strike B, maturity T on (V_t) for given interest rate r, and volatility σ_V. It is well known that

$$P^{BS}(t, V_t; r, \sigma_V, B, T) = Be^{-r(T-t)}\Phi(-d_{t,2}) - V_t\Phi(-d_{t,1}), \tag{10.11}$$

with $d_{t,1}$ and $d_{t,2}$ as in (10.9). Combining (10.10) and (10.11) we get

$$B_t = p_0(t, T)B\Phi(d_{t,2}) + V_t\Phi(-d_{t,1}). \tag{10.12}$$

Lines showing the evolution of B_t as a function of the evolution of V_t under the assumption that $r = 0.02$ have been added to Figure 10.3. The difference between the asset value V_t and the debt B_t is the value of equity S_t; note how the value of equity is essentially negligible for $t > 0.8$ in the default path.

Volatility of the firm's equity. It is interesting to compute the volatility of the equity of the firm under Assumption 10.3. To this end we define the quantity

$$\nu(t, V_t) := \frac{V_t C_V^{BS}(t, V_t)}{C^{BS}(t, V_t)}.$$

In the context of option pricing this is known as the *elasticity* of a European call with respect to the price of the underlying security. In our context it measures the percentage change in the value of equity per percentage change in the value of the underlying assets.

If we apply Itô's formula to $S_t = C^{BS}(t, V_t; r, \sigma_V, B, T)$ we obtain

$$dS_t = \sigma_V C_V^{BS}(t, V_t)V_t\, dW_t + (C_t^{BS}(t, V_t) + \mu_V C_V^{BS}(t, V_t)V_t + \tfrac{1}{2}\sigma_V^2 V_t^2 C_{VV}^{BS})\, dt.$$

Using the definition of the elasticity ν, we may rewrite the dW_t term in the form

$$\sigma_V C_V^{BS}(t, V_t)V_t\, dW_t = \sigma_V \nu(t, V_t)C^{BS}(t, V_t)\, dW_t,$$

from which we conclude that the volatility of the firm's equity at time t is a function $\sigma_S(t, V_t)$ of time and of the current asset value V_t that takes the form

$$\sigma_S(t, V_t) = \nu(t, V_t)\sigma_V. \tag{10.13}$$

The volatility of the firm's equity is therefore greater than σ_V, since the elasticity of a European call is always greater than one.

Risk-neutral and physical default probabilities. Next we compare physical and risk-neutral default probabilities in Merton's model. It is a basic result from financial mathematics that under the risk-neutral measure Q the process (V_t) satisfies the stochastic differential equation (SDE) $dV_t = rV_t\, dt + \sigma_V V_t\, d\tilde{W}_t$ for a standard Q-Brownian motion $\tilde{W}$. Note how the drift μ_V in (10.5) has been replaced by the risk-free interest rate r. The risk-neutral default probability is therefore given by the formula (10.6), evaluated with $\mu_V = r$:

$$q = Q(V_T \leqslant B) = \Phi\left(\frac{\ln B - \ln V_0 - (r - \frac{1}{2}\sigma_V^2)T}{\sigma_V\sqrt{T}}\right).$$

Comparing this with the physical default probability $p = P(V_T \leqslant B)$ as given in (10.6) we obtain the relationship

$$q = \Phi\left(\Phi^{-1}(p) + \frac{\mu_V - r}{\sigma_V}\sqrt{T}\right). \tag{10.14}$$

The correction term $(\mu_V - r)/\sigma_V$ equals the *Sharpe ratio* of V (a popular measure of the risk premium earned by the firm). The transition formula (10.14) is sometimes applied in practice to go from physical to risk-neutral default probabilities. Note, however, that (10.14) is supported by theoretical arguments only in the narrow context of the Merton model.

Credit spread. We may use (10.12) to infer the credit spread $c(t, T)$ implied by Merton's model. The credit spread measures the difference between the continuously compounded yield to maturity of a defaultable zero-coupon bond $p_1(t, T)$ and that of a default-free zero-coupon bond $p_0(t, T)$. It is defined by

$$c(t, T) = \frac{-1}{T-t}(\ln p_1(t, T) - \ln p_0(T, t)) = \frac{-1}{T-t}\ln\frac{p_1(t, T)}{p_0(t, T)}. \tag{10.15}$$

Throughout the book we use the convention that a zero-coupon bond has a nominal value equal to 1. In line with this convention we assume that the pay-off at T of a zero-coupon bond issued by the firm is given by $(1/B)B_T$, so that the price of such a bond at time $t \leqslant T$ is given by $p_1(t, T) = (1/B)B_t$. We therefore obtain

$$c(t, T) = \frac{-1}{T-t}\ln\left(\Phi(d_{t,2}) + \frac{V_t}{Bp_0(t, T)}\Phi(-d_{t,1})\right). \tag{10.16}$$

Since $d_{t,1}$ can be rewritten as

$$d_{t,1} = \frac{-\ln(Bp_0(t, T)/V_t) + \frac{1}{2}\sigma_V^2(T-t)}{\sigma_V\sqrt{T-t}},$$

and similarly for $d_{t,2}$, we conclude that, for a fixed time to maturity $T - t$, the spread $c(t, T)$ depends only on the volatility σ_V and on the ratio $d := Bp_0(t, T)/V_t$, which is the ratio of the discounted nominal value of the firm's debt to the value of the firm's assets and is hence a measure of the relative debt level or *leverage* of the firm. As the price of a European put (10.11) is increasing in the volatility, it follows from (10.10) that $c(t, T)$ is increasing in σ_V. In Figure 10.4 we plot the credit spread as a function of σ_V and of the time to maturity $\tau = T - t$.

Extensions. Merton's model is quite simplistic. Over the years this has given rise to a rich literature on firm-value models. We briefly comment on the most important research directions (bibliographic references are given in Notes and Comments). To begin with, the observation that, in reality, firms can default at essentially any time (and not only at a deterministic point in time T) has led to the development of so-called *first-passage-time models*. In this class of models default occurs when the asset-value process crosses a default threshold B for the first time; the threshold is usually interpreted as the average value of the liabilities. Formally, the default time τ is defined by $\tau = \inf\{t \geqslant 0\colon V_t \leqslant B\}$. Further technical developments

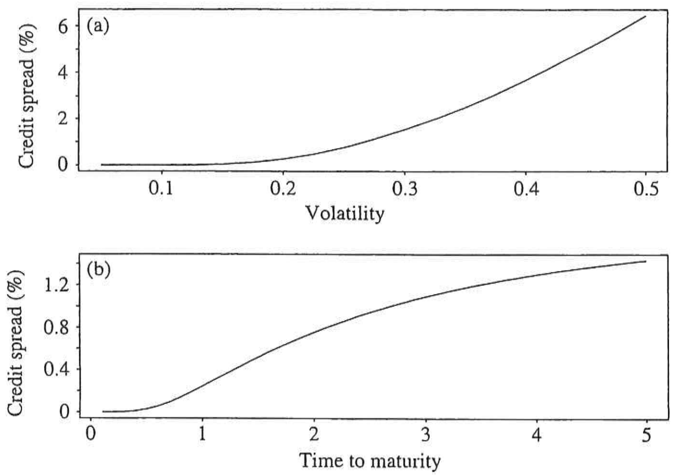


Figure 10.4. Credit spread $c(t, T)$ in per cent as a function of (a) the firm's volatility σ_V and (b) the time to maturity $\tau = T - t$ for fixed leverage measure $d = 0.6$ (in (a) $\tau = 2$ years; in (b) $\sigma_V = 0.25$). Note that, for a time to maturity smaller than approximately three months, the credit spread implied by Merton's model is basically equal to zero. This is not in line with most empirical studies of corporate bond spreads and has given rise to a number of extensions of Merton's model that are listed in Notes and Comments. We will see in Section 10.5.3 that reduced-form models lead to a more reasonable behaviour of short-term credit spreads.

include models with stochastic default-free interest rates and models where the asset-value process (V_t) is given by a diffusion with jumps.

Firm-value models with an *endogenous default threshold* are an interesting economic extension of Merton's model. Here the default boundary B is not fixed a priori by the modeller but is determined endogenously by strategic considerations of the shareholders. Finally, structural models with *incomplete information* on asset value and/or liabilities provide an important link between the structural and reduced-form approaches to credit risk modelling.

10.3.3 Structural Models in Practice: EDF and DD

There are a number of industry models that descend from the Merton model. An important example is the so-called *public-firm EDF model* that is maintained by Moody's Analytics. The acronym EDF stands for *expected default frequency*; this is a specific estimate of the physical default probability of a given firm over a one-year horizon. The methodology proposed by Moody's Analytics builds on earlier work by KMV (a private company named after its founders Kealhofer, McQuown and Vasicek) in the 1990s, and is also known as the KMV model. Our presentation of the public-firm EDF model is based on Crosbie and Bohn (2002) and Sun, Munves and Hamilton (2012). We concentrate on the main ideas, since detailed information about actual implementation and calibration procedures is proprietary and these procedures may change over time.

Overview. Recall that in the classic Merton model the one-year default probability of a given firm is given by the probability that the asset value in one year lies below the threshold B representing the overall liabilities of the firm. Under Assumption 10.3, the one-year default probability is a function of the current asset value V_0, the (annualized) drift μ_V and volatility σ_V of the asset-value process, and the threshold B; using (10.6) with $T = 1$ and recalling that $\Phi(d) = 1 - \Phi(-d)$ we infer that

$$\mathrm{EDF}_{\mathrm{Merton}} = 1 - \Phi\left(\frac{\ln V_0 - \ln B + (\mu_V - \frac{1}{2}\sigma_V^2)}{\sigma_V}\right). \tag{10.17}$$

In the public-firm EDF model a similar structure is assumed for the EDF; however, $1 - \Phi$ is replaced by some empirically estimated function, B is replaced by a new default threshold $\tilde{B}$ representing the structure of the firm's liabilities more closely, and the term $(\mu_V - \frac{1}{2}\sigma_V^2)$ in the numerator is sometimes omitted for expositional ease. Moreover, the current asset value V_0 and the asset volatility σ_V are inferred (or "backed out") from information about the firm's equity value.

Determination of the asset value and the asset volatility. Firm-value-based credit risk models are based on the *market value* V_0 of the firm's assets. This makes sense as the market value is a forward-looking measure that reflects investor expectations about the business prospects and future cash flows of the firm. Unfortunately, in contrast to the assumptions underlying Merton's model, in most cases there is no market for the assets of a firm, so that the asset value is not directly observable. Moreover, the market value can differ greatly from the value of a company as measured by accountancy rules (the so-called book value), so that accounting information and balance sheet data are of limited use in inferring the asset value V_0. For these reasons the public-firm EDF model relies on an indirect approach and infers values of V_t at different times t from the more easily observed values of a firm's equity S_t. This approach simultaneously provides estimates of V_0 and of the asset volatility σ_V. The latter estimate is needed since σ_V has a strong impact on default probabilities; all other things being equal, risky firms with a comparatively high asset volatility σ_V have a higher default probability than firms with a low asset volatility.

We explain the estimation approach in the context of the Merton model. Recall that under Assumption 10.3 we have that

$$S_t = C^{\mathrm{BS}}(t, V_t; r, \sigma_V, B, T). \tag{10.18}$$

Obviously, at a fixed point in time, $t = 0$ say, (10.18) is an equation with two unknowns, V_0 and σ_V. To overcome this difficulty, one may use an iterative procedure. In step (1), (10.18) with some initial estimate $\sigma_V^{(0)}$ is used to infer a time series of asset values $(V_t^{(0)})$ from equity values. Then a new volatility estimate $\sigma_V^{(1)}$ is estimated from this time series; a new time series $(V_t^{(1)})$ is then constructed using (10.18) with $\sigma_V^{(1)}$. This procedure is iterated n times, until the volatility estimates $\sigma_V^{(n-1)}$ and $\sigma_V^{(n)}$ generated in step $(n-1)$ and step (n) are sufficiently close.

In the public-firm EDF model, the capital structure of the firm is modelled in a more sophisticated manner than in Merton's model. There are several classes of liabilities, such as long- and short-term debt and convertible bonds, the model allows for intermediate cash payouts corresponding to coupons and dividends, and default can occur at any time. Moreover, the default point (the threshold value $\tilde{B}$ such that the company defaults if (V_t) falls below $\tilde{B}$) is determined from a more detailed analysis of the term structure of the firm's debt. The equity value is thus no longer given by (10.18) but by some different function $f(t, V_t, \sigma_V)$, which has to be computed numerically. The general idea of the approach used to estimate V_0 and σ_V is, however, exactly as described above.

Calculation of EDFs. In the Merton model, default occurs if the value of a firm's assets falls below the value of its liabilities. With lognormally distributed asset values, as implied for instance by Assumption 10.3, this leads to default probabilities of the form $\text{EDF}_{\text{Merton}}$ as in (10.17). This relationship between asset value and default probability may be too simplistic to be an accurate description of actual default probabilities. For instance, asset values are not necessarily lognormal but might follow a distribution with heavy tails and there might be payments due at an intermediate point in time causing default at that date.

For these reasons, in the public-firm EDF model a new state variable is introduced in an intermediate step. This is the so-called *distance-to-default* (DD), given by

$$\text{DD} := (\log V_0 - \log \tilde{B})/\sigma_V. \tag{10.19}$$

Here, $\tilde{B}$ represents the default threshold; in some versions of the model $\tilde{B}$ is modelled as the sum of the liabilities payable within one year and half of the longer-term debt. Sometimes practitioners call the distance-to-default the "number of standard deviations a company is away from its default threshold". Note that (10.19) is in fact an approximation of the argument of (10.17), since μ_V and σ_V^2 are usually small.

In the EDF methodology it is assumed that the distance-to-default *ranks* firms in the sense that firms with a higher DD exhibit a higher default probability. The functional relationship between DD and EDF is determined empirically; using a database of historical default events, the proportion of firms with DD in a given small range that default within a year is estimated. This proportion is the empirically estimated EDF. The DD-to-EDF mapping exhibits "heavy tails": for high-quality firms with a large DD the empirically estimated EDF is much higher than $\text{EDF}_{\text{Merton}}$ as given in (10.17). For instance, for a firm with a DD equal to 4 we find that $\text{EDF}_{\text{Merton}} \approx 0.003\%$, whereas the empirically estimated EDF equals 0.4%.

In Table 10.3 we illustrate the computation of the EDF for two different firms, Johnson & Johnson (a well-capitalized firm that operates in the relatively stable health care market) and RadioShack (a firm that is active in the highly volatile consumer electronics business). If we compare the numbers, we see that the EDF for Johnson & Johnson is close to zero whereas the EDF for RadioShack is quite high. This difference reflects the higher leverage of RadioShack and the riskiness of the underlying business, as reflected by the comparatively large asset volatility $\sigma_V = 24\%$. Indeed, on 11 September 2014, the *New York Times* reported that a bankruptcy filing for RadioShack could be near, suggesting that the EDF had good predictive power in this case.

Table 10.3. A summary of the public-firm EDF methodology. The example is taken from Sun, Munves and Hamilton (2012); it is concerned with the situation of Johnson & Johnson (J&J) and RadioShack as of April 2012. All quantities are in US dollars.

Variable	J&J	RadioShack	Notes
Market value of assets V_0	236 bn	1834 m	Determined from time series of equity prices
Asset volatility σ_V	11%	24%	
Default threshold $\tilde{B}$	39 bn	1042 m	Short-term liabilities and half of long-term liabilities
DD	16.4	2.3	Given by $(\log V_0 - \log \tilde{B})/\sigma_V$
EDF (one year)	0.01%	3.58%	Determined using empirical mapping between DD and EDF

10.3.4 Credit-Migration Models Revisited

Recall that in the credit-migration approach each firm is assigned to a credit-rating category at any given time point. There are a finite number of such ratings and they are ordered by credit quality and include the category of default. The probability of moving from one credit rating to another credit rating over the given risk horizon (typically one year) is then specified. In this section we explain how a migration model can be embedded in a firm-value model and thus be treated as a structural model. This will be useful in the discussion of portfolio versions of these models in Chapter 11. Moreover, we compare the public-firm EDF model and credit-migration approaches.

Credit-migration models as firm-value models. We consider a firm that has been assigned to some non-default rating category j at $t = 0$ and for which transition probabilities p_{jk}, $0 \leqslant k \leqslant n$, over the period $[0, T]$ are available on the basis of that rating. These express the probability that the firm belongs to rating class k at the time horizon T, given that it is in class j at $t = 0$. In particular, $p_{j,0}$ is the default probability of the firm over $[0, T]$.

Suppose that the asset-value process (V_t) of the firm follows the model given in (10.5), so that

$$V_T = V_0 \exp((\mu_V - \tfrac{1}{2}\sigma_V^2)T + \sigma_V W_T) \tag{10.20}$$

is lognormally distributed. We can now choose thresholds

$$0 = \tilde{d}_0 < \tilde{d}_1 < \cdots < \tilde{d}_n < \tilde{d}_{n+1} = \infty \tag{10.21}$$

such that $P(\tilde{d}_k < V_T \leqslant \tilde{d}_{k+1}) = p_{jk}$ for $k \in \{0, \ldots, n\}$. We have therefore translated the transition probabilities into a series of thresholds for an assumed asset-value process. The threshold $\tilde{d}_1$ is the default threshold, which in the Merton model of Section 10.3.1 was interpreted as the value of the firm's liabilities. The higher thresholds are the asset-value levels that mark the boundaries of higher rating categories. The firm-value model in which we have embedded the migration model can be summarized by saying that the firm belongs to rating class k at the time horizon T if and only if $\tilde{d}_k < V_T \leqslant \tilde{d}_{k+1}$.

The migration probabilities in the firm-value model obviously remain invariant under simultaneous strictly increasing transformations of V_T and the thresholds $\tilde{d}_j$. If we define

$$X_T := \frac{\ln V_T - \ln V_0 - (\mu_V - \frac{1}{2}\sigma_V^2)T}{\sigma_V \sqrt{T}}, \tag{10.22}$$

$$d_k := \frac{\ln \tilde{d}_k - \ln V_0 - (\mu_V - \frac{1}{2}\sigma_V^2)T}{\sigma_V \sqrt{T}}, \tag{10.23}$$

then we can also say that the firm belongs to rating class k at the time horizon T if and only if $d_k < X_T \leqslant d_{k+1}$. Observe that X_T is a standardized version of the *asset-value log-return* $\ln V_T - \ln V_0$, and we can easily verify that $X_T = W_T/\sqrt{T}$ so that it has a standard normal distribution. In this case the formulas for the thresholds are easily obtained and are $d_k = \Phi^{-1}(\sum_{l=0}^{k-1} p_{jl})$ for $k = 1, \dots, n$.

The public-firm EDF model and credit-migration approaches compared. The public-firm EDF model uses market data, most notably the current stock price, as inputs for the EDF computation. The EDF therefore reacts quickly to changes in the economic prospects of a firm, as these are reflected in the firm's share price and hence in the estimated distance-to-default. Moreover, EDFs are quite sensitive to the current macroeconomic environment. The distance-to-default is observed to rise in periods of economic expansion (essentially due to higher share prices reflecting better economic conditions) and to decrease in recession periods. Rating agencies, on the other hand, are typically slow in adjusting their credit ratings, so that the current rating does not always reflect the economic condition of a firm. This is particularly important if the credit quality of a firm deteriorates rapidly, as is typically the case with companies that are close to default. For instance, the investment bank Lehman Brothers had a fairly good rating (Aa or better) when it defaulted in September 2008. EDFs might therefore be better predictors of default probabilities over short time horizons.

On the other hand, the public-firm EDF model is quite sensitive to global over- and underreaction of equity markets. In particular, the bursting of a stock market bubble may lead to drastically increased EDFs, even if the economic outlook for a given corporation has not changed very much. This can lead to huge fluctuations in the amount of risk capital that is required to back a given credit portfolio. From this point of view the relative inertia of ratings-based models could be considered an advantage, as the ensuing risk capital requirements tend to be more stable over time.

Notes and Comments

There are many excellent texts, at varying technical levels, in which the basic mathematical finance results used in Section 10.3.2 can be found. Models in discrete time are discussed in Cox and Rubinstein (1985) and Jarrow and Turnbull (1999); excellent introductions to continuous-time models include Baxter and Rennie (1996), Björk (2004), Bingham and Kiesel (2004) and Shreve (2004b).

Lando (2004) gives a good overview of the rich literature on firm-value models. First-passage-time models have been considered by, among others, Black and Cox (1976) and, in a set-up with stochastic interest rates, Longstaff and Schwartz (1995). The problem of the unrealistically low credit spreads for small maturities $\tau = T - t$, which we pointed out in Figure 10.4, has also led to extensions of Merton's model. Partial remedies within the class of firm-value models include models with jumps in the firm value, as in Zhou (2001), time-varying default thresholds, as in Hull and White (2001), stochastic volatility models for the firm-value process with time-dependent dynamics, as in Overbeck and Schmidt (2005), and incomplete information on firm value or default threshold, as in Duffie and Lando (2001), Frey and Schmidt (2009) and Cetin (2012). Models with endogenous default thresholds have been considered by, among others, Leland (1994), Leland and Toft (1996) and Hilberink and Rogers (2002).

Duffie and Lando (2001) established a relationship between firm-value models and reduced-form models in continuous time. Essentially, they showed that, from the perspective of investors with *incomplete accounting information* (i.e. incomplete information about the assets or liabilities of a firm), a firm-value model becomes a reduced-form model. A less technical discussion of these issues can be found in Jarrow and Protter (2004).

The public-firm EDF model was first described in Crosbie and Bohn (2002); the model variant that is currently in use is described in Dwyer and Qu (2007) and Sun, Munves and Hamilton (2012).

10.4 Bond and CDS Pricing in Hazard Rate Models

Hazard rate models are the most basic reduced-form credit risk models and are therefore a natural starting point for our discussion of this model class. Moreover, hazard rate models are used as an input in the construction of the popular copula models for portfolio credit derivatives. For these reasons this section is devoted to bond and CDS pricing in hazard rate models. We begin by introducing the necessary mathematical background in Section 10.4.1. Since the pricing results that we present in this section rely on the concept of risk-neutral pricing and martingale modelling, we briefly review these notions in Section 10.4.2. The pricing of bonds and CDSs and some of the related empirical evidence is discussed in Sections 10.4.3, 10.4.4 and 10.4.5.

10.4.1 Hazard Rate Models

A hazard rate model is a model in which the distribution of the default time of an obligor is directly specified by a hazard function without modelling the mechanism by which default occurs.

To set up a hazard rate model we consider a probability space $(\Omega, \mathcal{F}, P)$ and a random time τ defined on this space, i.e. an $\mathcal{F}$-measurable rv taking values in $[0, \infty]$. In economic terms, τ can be interpreted as the default time of some company. We denote the df of τ by $F(t) = P(\tau \leqslant t)$ and the tail or survival function by

$\bar{F}(t) = 1 - F(t)$; we assume that $P(\tau = 0) = F(0) = 0$, and that $\bar{F}(t) > 0$ for all $t < \infty$. We define the *jump* or *default indicator process* (Y_t) associated with τ by

$$Y_t = I_{\{\tau \leqslant t\}}, \quad t \geqslant 0. \tag{10.24}$$

Note that (Y_t) is a right-continuous process that jumps from 0 to 1 at the default time τ and that $1 - Y_t = I_{\{\tau > t\}}$ is the *survival indicator* of the firm.

Definition 10.4 (cumulative hazard function and hazard function). The function $\Gamma(t) := -\ln(\bar{F}(t))$ is called the *cumulative hazard function* of the random time τ. If F is absolutely continuous with density f, the function $\gamma(t) := f(t)/(1 - F(t)) = f(t)/\bar{F}(t)$ is called the *hazard function* of τ.

By definition we have $\bar{F}(t) = e^{-\Gamma(t)}$. If F has density f, we calculate that $\Gamma'(t) = f(t)/\bar{F}(t) = \gamma(t)$, so that we can represent the survival function of τ in terms of the hazard function by

$$\bar{F}(t) = \exp\left(-\int_0^t \gamma(s)\,ds\right). \tag{10.25}$$

The hazard function $\gamma(t)$ at a fixed time t gives the *hazard rate* at t, which can be interpreted as a measure of the instantaneous risk of default at t, given survival up to time t. In fact, for $h > 0$ we have $P(\tau \leqslant t+h \mid \tau > t) = (F(t+h) - F(t))/\bar{F}(t)$. Hence we obtain

$$\lim_{h \to 0} \frac{1}{h} P(\tau \leqslant t + h \mid \tau > t) = \frac{1}{\bar{F}(t)} \lim_{h \to 0} \frac{F(t+h) - F(t)}{h} = \gamma(t).$$

For illustrative purposes we determine the hazard function for the Weibull distribution. This is a popular distribution for survival times with df $F(t) = 1 - e^{-\lambda t^{\alpha}}$ for parameters $\lambda, \alpha > 0$. For $\alpha = 1$ the Weibull distribution reduces to the standard exponential distribution. Differentiation yields

$$f(t) = \lambda\alpha t^{\alpha-1} e^{-\lambda t^{\alpha}} \quad \text{and} \quad \gamma(t) = \lambda\alpha t^{\alpha-1}.$$

In particular, γ is decreasing in t if $\alpha < 1$ and increasing if $\alpha > 1$. For $\alpha = 1$ (exponential distribution) the hazard rate is time independent and equal to the parameter λ.

Filtrations and conditional expectations. In financial models, filtrations are used to model the information available to investors at various points in time. Formally, a *filtration* $(\mathcal{F}_t)$ on $(\Omega, \mathcal{F})$ is an increasing family $\{\mathcal{F}_t : t \geqslant 0\}$ of sub-σ-algebras of $\mathcal{F}$: $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$ for $0 \leqslant t \leqslant s < \infty$. The σ-algebra $\mathcal{F}_t$ represents the state of knowledge of an observer at time t, and $A \in \mathcal{F}_t$ is taken to mean that at time t the observer is able to determine if the event A has occurred.

In this section it is assumed that the only quantity that is observable for investors is the default event of the firm under consideration or, equivalently, the default indicator process (Y_t) associated with τ. The appropriate filtration is therefore given by $(\mathcal{H}_t)$ with

$$\mathcal{H}_t = \sigma(\{Y_u : u \leqslant t\}), \tag{10.26}$$

the *default history* up to and including time t. By definition, τ is an $(\mathcal{H}_t)$ stopping time, as $\{\tau \leqslant t\} = \{Y_t = 1\} \in \mathcal{H}_t$ for all $t \geqslant 0$; moreover, $(\mathcal{H}_t)$ is obviously the smallest filtration with this property.

In order to study bond and CDS pricing in hazard rate models we need to compute conditional expectations with respect to the σ-algebra $\mathcal{H}_t$. We begin our analysis of this issue with an auxiliary result on the structure of $\mathcal{H}_t$-measurable rvs. The result formalizes the fact that every $\mathcal{H}_t$-measurable rv can be expressed as a function of events related to the default history at t.

Lemma 10.5. *Every $\mathcal{H}_t$-measurable rv H is of the form $H = h(\tau)I_{\{\tau \leqslant t\}} + cI_{\{\tau > t\}}$ for a measurable function $h\colon [0, t] \to \mathbb{R}$ and some constant $c \in \mathbb{R}$.*

Proof. The σ-algebra $\mathcal{H}_t$ is generated by the events $\{Y_u = 1\} = \{\tau \leqslant u\}$, $u < t$, and $\{Y_t = 0\} = \{\tau > t\}$, and hence by the rvs $(\tau \wedge t) := \min\{\tau, t\}$ and $I_{\{\tau > t\}}$. This implies that any $\mathcal{H}_t$-measurable rv H can be written as $H = g(\tau \wedge t, I_{\{\tau > t\}})$ for some measurable function $g\colon [0, t] \times \{0, 1\} \to \mathbb{R}$. The claim follows if we define $h(u) := g(u, 0)$, $u \leqslant t$, and $c := g(t, 1)$. □

Lemma 10.6. *Let τ be a random time with jump indicator process $Y_t = I_{\{\tau \leqslant t\}}$ and natural filtration $(\mathcal{H}_t)$. Then, for any integrable rv X and any $t \geqslant 0$, we have*

$$E(I_{\{\tau > t\}}X \mid \mathcal{H}_t) = I_{\{\tau > t\}} \frac{E(I_{\{\tau > t\}}X)}{P(\tau > t)}. \tag{10.27}$$

Proof. Since $E(I_{\{\tau > t\}}X \mid \mathcal{H}_t)$ is $\mathcal{H}_t$-measurable and zero on $\{\tau \leqslant t\}$, we obtain from Lemma 10.5 that $E(I_{\{\tau > t\}}X \mid \mathcal{H}_t) = I_{\{\tau > t\}}c$ for some constant c. Taking expectations yields $E(I_{\{\tau > t\}}X) = cP(\tau > t)$ and hence $c = E(I_{\{\tau > t\}}X)/P(\tau > t)$. □

Lemma 10.6 can be used to determine conditional survival probabilities. Fix $t < T$ and consider the quantity $P(\tau > T \mid \mathcal{H}_t)$. Applying (10.27) with $X := I_{\{\tau > T\}}$ yields

$$P(\tau > T \mid \mathcal{H}_t) = E(X \mid \mathcal{H}_t) = E(I_{\{\tau > t\}}X \mid \mathcal{H}_t) = I_{\{\tau > t\}} \frac{\bar{F}(T)}{\bar{F}(t)}. \tag{10.28}$$

If τ admits the hazard function $\gamma(t)$, we get the important formula

$$P(\tau > T \mid \mathcal{H}_t) = I_{\{\tau > t\}} \exp\left(-\int_t^T \gamma(s)\,ds\right), \quad t < T. \tag{10.29}$$

The next proposition is concerned with stochastic process properties of the jump indicator process of a random time τ.

Proposition 10.7. *Let τ be a random time with absolutely continuous df F and hazard function γ. Then $M_t := Y_t - \int_0^t I_{\{\tau > u\}}\gamma(u)\,du$, $t \geqslant 0$, is an $(\mathcal{H}_t)$-martingale: that is, $E(M_s \mid \mathcal{H}_t) = M_t$ for all $0 \leqslant t \leqslant s < \infty$.*

In Section 10.5.1 we extend this result to doubly stochastic random times and discuss its financial and mathematical relevance.

Proof. Let $s > t$. We have to show that $E(M_s - M_t \mid \mathcal{H}_t) = 0$, i.e. that $E(Y_s - Y_t \mid \mathcal{H}_t) = E(\int_t^s \gamma(u) I_{\{u<\tau\}}\, du \mid \mathcal{H}_t)$. Using (10.28) we get

$$E(Y_s - Y_t \mid \mathcal{H}_t) = I_{\{\tau>t\}} P(\tau \leqslant s \mid \mathcal{H}_t) = I_{\{\tau>t\}}\left(1 - \frac{\bar F(s)}{\bar F(t)}\right)$$
$$= I_{\{\tau>t\}} \frac{\bar F(t) - \bar F(s)}{\bar F(t)}.$$

Note that $X := \int_t^s \gamma(u) I_{\{u<\tau\}}\, du$ is 0 on $\{\tau \leqslant t\}$, so $X = X I_{\{\tau>t\}}$. Hence we obtain from Lemma 10.6, the Fubini Theorem and the identity $\bar F'(t) = -f(t) = -\gamma(t)\bar F(t)$ that

$$E(X \mid \mathcal{H}_t) = I_{\{\tau>t\}} \frac{E(X)}{\bar F(t)} = I_{\{\tau>t\}} \frac{\int_t^s \gamma(u)\bar F(u)\, du}{\bar F(t)} = I_{\{\tau>t\}} \frac{\bar F(t) - \bar F(s)}{\bar F(t)},$$

and the result follows. □

10.4.2 Risk-Neutral Pricing Revisited

The remainder of Section 10.4 is devoted to an analysis of risk-neutral pricing results for credit products in hazard rate models. Risk-neutral pricing has become so popular that the conceptual underpinnings are often overlooked. A prime case in point is the mechanical use of the Gauss copula to price CDO tranches, a practice that led to well-documented problems during the 2007–9 financial crisis. It therefore seems appropriate to clarify the applicability and the limitations of risk-neutral pricing in the context of credit risk models.

Risk-neutral pricing. We build on the elementary discussion of risk-neutral valuation given in Section 2.2.2. In that section we considered a simple one-period default model for a defaultable zero-coupon bond with maturity T equal to one year and a deterministic recovery rate $1 - \delta$ equal to 60%. Moreover, we assumed that the real-world default probability was $p = 1\%$, the risk-free simple interest rate was $r_{0,1} = 5\%$, and the market price of the bond at $t = 0$ was $p_1(0, 1) = 0.941$.

Risk-neutral pricing is intimately linked to the notion of a risk-neutral measure. In general terms a risk-neutral measure is an artificial probability measure Q, equivalent to the historical measure P, such that the discounted prices of all traded securities are Q-martingales (fair bets). We have seen in Section 2.2.2 that in the simple one-period default model a risk-neutral measure Q is simply given by an artificial default probability q such that

$$p_1(0, 1) = (1.05)^{-1}((1 - q) \cdot 1 + q \cdot 0.6).$$

Obviously, q is uniquely determined by this equation and is given by $q = 0.03$. Note that in this example the risk-neutral default probability q is higher than the real-world default probability p. This reflects risk aversion on the part of investors and is typical for real markets; empirical evidence on the relationship between the physical and historical default probabilities will be presented in Section 10.4.5.

The *risk-neutral pricing rule* states that the price of a derivative security can be computed as the mathematical expectation of the discounted pay-off under a risk-neutral measure Q. In mathematical terms the price at time t of a derivative with pay-off H and maturity $T \geqslant t$ is thus given by

$$V_t^H = E^Q\left(\exp\left(-\int_t^T r_s\, ds\right) H \;\middle|\; \mathcal{F}_t\right), \tag{10.30}$$

where r_s denotes the continuously compounded default-free short rate of interest at time s and where the σ-algebra $\mathcal{F}_t$ represents the information available to investors at time t (see the discussion of filtrations in Section 10.4.1). Note that in one-period models, (10.30) reduces to the simpler expression $V_0^H = E^Q(H/(1 + r_{0,1}))$, where $r_{0,1}$ is the simple interest rate for the period.

There are two theoretical justifications for risk-neutral pricing. One argument is based on absence of arbitrage: according to the first fundamental theorem of asset pricing, a model for security prices is arbitrage free if and only if it admits at least one equivalent martingale measure Q. Hence, if a financial product is to be priced in accordance with no-arbitrage principles, its price must be given by the risk-neutral pricing formula for some risk-neutral measure Q. A second justification relies on hedging: in financial models it is often possible to replicate the pay-off of a financial product by (dynamic) trading in the available assets, and in a frictionless market the cost of carrying out such a hedge is given by the risk-neutral pricing rule.

Hedging and market completeness. Next we take a closer look at the concept of hedging. We work in the simple one-period default model that was introduced in the previous paragraph. Consider an investor, e.g. an investment bank, who plans to sell derivatives on the defaultable zero-coupon bond. For concreteness we consider a *default put option* with maturity date $T = 1$. This contract pays one unit if the bond defaults and zero otherwise; it can be thought of as a simplified version of a CDS. Obviously, the pay-off of the default put is unknown at date $t = 0$ and thus constitutes a risk for the investor. A possible strategy for dealing with this risk is to form a *hedging portfolio* in the defaultable bond and in cash that reduces the risk of selling the put: suppose that at time $t = 0$ we go short 2.5 units of the bond and hold $\frac{50}{21} \approx 2.38$ units of cash. At time $t = 1$ there are two possibilities for the value V_1 of this portfolio.

- Default occurs: in which case $V_1 = (-2.5) \cdot 0.6 + \frac{50}{21} \cdot 1.05 = 1$.
- No default: in which case $V_1 = (-2.5) \cdot 1 + \frac{50}{21} \cdot 1.05 = 0$.

In either case the value V_1 of the hedge portfolio equals the pay-off of the option and we have found a so-called *replicating strategy* for the option. In particular, by forming the replicating strategy the investor completely eliminates the risk from selling the option, and the *law of one price* dictates that the fair price at $t = 0$ of the option should equal the value of the hedge portfolio at $t = 0$ given by $V_0 = (-2.5) \cdot 0.941 + \frac{50}{21} \approx 0.0285$ (otherwise either the buyer or the seller could make some risk-free profit).

To construct the portfolio in this simple one-period, two-state setting we have to consider two linear equations. Denote by ξ_1 and ξ_2 the units of the defaultable bond and the amount of cash in our portfolio. At time $t = 1$ we must have

$$\xi_1 \cdot 0.6 + \xi_2 \cdot 1.05 = 1 \quad \text{(the default case)}, \tag{10.31}$$
$$\xi_1 \cdot 1.0 + \xi_2 \cdot 1.05 = 0 \quad \text{(the no-default case)}, \tag{10.32}$$

which leads to the above values of $\xi_1 = -2.5$ and $\xi_2 = \frac{50}{21}$. In mathematical finance a derivative security is called *attainable* if there is a replicating portfolio strategy in the underlying assets. The above argument shows that in the simple one-period default model with only two states every derivative security is attainable. Such models are termed *complete*.

The fair price of the default put (the initial value V_0 of the replicating portfolio) can alternatively be computed by the risk-neutral pricing rule. Recall that the risk-neutral default probability is given by $q = 0.03$. The risk-neutral pricing rule applied to the default put thus leads to a value of $(1.05)^{-1}(0.97 \cdot 0 + 0.03 \cdot 1) = 0.0285$, which is equal to V_0. This is, of course, not a lucky coincidence; a basic result from mathematical finance states that the fair price of any attainable claim can be computed as the expected value of the discounted pay-off under a risk-neutral measure. Armed with this result, we typically first compute the price (the expected value of the discounted pay-off under a risk-neutral measure) and then determine the replicating strategy. For this reason a lot of research focuses on the problem of computing prices. However, one should bear in mind that the economic justification for the risk-neutral pricing rule stems partially from the hedging argument, which applies only to attainable claims. This issue has, to a large extent, been neglected in the literature on the pricing of credit-risky securities. The next example illustrates some of the difficulties arising in *incomplete* markets, where most derivatives are not attainable.

Example 10.8 (a model with random recovery). As there is a substantial amount of randomness in real recovery rates, it is interesting to study the impact of random recovery rates on the validity of the above pricing arguments. We consider an extension of the basic one-period default model in which the loss given default may be either 30% or 50%. The price is assumed to be $p_1(0,1) = 0.941$ and the risk-free simple interest rate is assumed to be $r_{0,1} = 5\%$ as before. The evolution of the price $p_1(\cdot, 1)$ is illustrated in Figure 10.5. We leave the physical measure unspecified—we assume only that all three possible outcomes have strictly positive probability.

We begin our analysis of this model by determining the equivalent martingale measures. Let q_1 be the risk-neutral probability that default occurs and the LGD is 0.5, let q_2 be the risk-neutral probability that default occurs and the LGD is 0.3, and let $q_3 = 1 - q_1 - q_2$. It follows that q_1 and q_2 satisfy the equation

$$p_1(0,1) = 1.05^{-1}(q_1 \cdot 0.5 + q_2 \cdot 0.7 + (1 - q_1 - q_2) \cdot 1), \tag{10.33}$$

with the restrictions that $q_1 > 0$, $q_2 > 0$, $1 - q_1 - q_2 > 0$. Obviously, Q is no longer unique. It is easily seen from (10.33) that the set $\mathcal{Q}$ of equivalent martingale

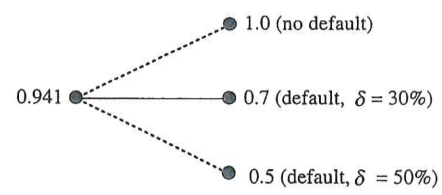


Figure 10.5. Evolution of the price $p_1(\cdot, 1)$ of the defaultable bond in Example 10.8.

measures is given by

$$\mathcal{Q} = \{\boldsymbol{q} \in \mathbb{R}^3 : q_1 \in (0, 0.024),\ q_2 = \tfrac{10}{3}(1 - 1.05 \cdot p_1(0,1) - 0.5 \cdot q_1),\ q_3 = 1 - (q_1 + q_2)\}. \tag{10.34}$$

It is interesting to look at the boundary cases. For $q_1 = 0$ we obtain $q_2 = 4\%$, $q_3 = 96\%$; this is the scenario where the risk-neutral default probability $q = q_1 + q_2$ is maximized. For $q_1 = 2.4\%$ we obtain $q_2 = 0$, $q_3 = 97.6\%$; this is the scenario where q is minimized. Note, however, that the measures $\boldsymbol{q}_0 := (0.024, 0, 0.976)$ and $\boldsymbol{q}_1 := (0, 0.04, 0.96)$ do not belong to $\mathcal{Q}$, as they are not equivalent to the physical measure P.

Consider a derivative security with pay-off H and maturity $T = 1$, such as the default put with pay-off $H = 0$ if $p_1(1,1) = 1$ (no default) and $H = 1$ otherwise. Every price of the form $H_0 = E^Q(1.05^{-1}H)$ for some $Q \in \mathcal{Q}$ is consistent with no arbitrage and will therefore be called an *admissible value* for the derivative. If $\mathcal{Q}$ contains more than one element, as in our case, there is typically more than one admissible value. For instance, we obtain for the default put option that

$$\inf_{Q \in \mathcal{Q}} E^Q\left(\frac{H}{1.05}\right) \approx 0.023 \quad \text{and} \quad \sup_{Q \in \mathcal{Q}} E^Q\left(\frac{H}{1.05}\right) \approx 0.038; \tag{10.35}$$

obviously, the infimum and supremum in (10.35) correspond to the measures $\boldsymbol{q}_0$ and $\boldsymbol{q}_1$, where q is minimized and maximized, respectively. This non-uniqueness of admissible values reflects the fact that in our three-state model the put is no longer attainable. In fact, the hedging portfolio (ξ_1, ξ_2) now has to solve the following three equations:

$$\left.\begin{aligned} \xi_1 \cdot 0.5 + \xi_2 \cdot 1.05 &= 1 \quad \text{(default, low recovery)},\\ \xi_1 \cdot 0.7 + \xi_2 \cdot 1.05 &= 1 \quad \text{(default, high recovery)},\\ \xi_1 \cdot 1 + \xi_2 \cdot 1.05 &= 0 \quad \text{(no default)}. \end{aligned}\right\} \tag{10.36}$$

It is immediately seen that the system (10.36) of three equations and only two unknowns has no solution, so that the default put is not attainable. This illustrates two fundamental results from modern mathematical finance: a claim with bounded pay-off is attainable if and only if the set of admissible values consists of a single number; an arbitrage-free market is complete if and only if there is exactly one equivalent martingale measure Q. The latter result is known as the *second fundamental theorem of asset pricing*.

Example 10.8 shows that in an incomplete market new issues arise; in particular, it is not obvious how to choose the correct price of a derivative security from the range of admissible values or how to deal with the risk incurred by selling a derivative security. This is unfortunate, as realistic models, which capture the dynamics of financial time series, are typically incomplete. In recent years a number of interesting concepts for the risk management of derivative securities in incomplete markets have been developed. These approaches typically propose mitigating the risk by an appropriate trading strategy and often suggest a pricing formula for the remaining risk. However, a discussion of this work is outside the scope of this book. A brief overview of the existing literature on hedging in (incomplete) credit markets is given in Notes and Comments.

Advantages and limitations of risk-neutral pricing. The risk-neutral pricing approach is a *relative pricing theory*, which explains prices of credit products in terms of observable prices of other securities. If properly applied, it leads to arbitrage-free prices of credit-risky securities, which are consistent with prices quoted in the market. These features make the risk-neutral pricing approach to credit risk the method of choice in an environment where credit risk is actively traded and, in particular, for valuing credit instruments when the market for related products is relatively liquid. On the other hand, since pricing models have to be calibrated to prices of traded credit instruments, they are difficult to apply when we lack sufficient market information. Moreover, in such cases prices quoted using an ad hoc choice of some risk-neutral measure are more or less plucked out of thin air.

This can be contrasted with the more traditional pricing methodology for loans and related credit products, where a loan is taken on the balance sheet if the spread earned on the loan is deemed by the lender to be a sufficient compensation for bearing the default risk of the loan and where the default risk is measured using the real-world measure and historical (default) data. Such an approach is well suited to situations where the market for related credit instruments is relatively illiquid and little or no price information is available; loans to medium or small businesses are a prime example. On the other hand, the traditional pricing methodology does not necessarily lead to prices that are consistent (in the sense of absence of arbitrage) across products or compatible with quoted market prices for credit instruments, so it is less suitable in a trading environment.

Martingale modelling. Recall that, according to the first fundamental theorem of asset pricing, a model for security prices is arbitrage free if and (essentially) only if it admits at least one equivalent martingale measure Q. Moreover, in a complete market, the only thing that matters for the pricing of derivative securities is the Q-dynamics of the traded underlying assets. When building a model for pricing derivatives it is therefore a natural shortcut to model the objects of interest—such as interest rates, default times and the price processes of traded bonds—directly, under some exogenously specified martingale measure Q. In the literature this approach is termed *martingale modelling*.

Martingale modelling is particularly convenient if the value H of the underlying assets at some maturity date T is exogenously given, as in the case of zero-coupon bonds. In that case the price of the underlying asset at time $t < T$ can be computed as the conditional expectation under Q of the discounted value at maturity via the risk-neutral pricing rule (10.30). Model parameters are then determined using the requirement that at time $t = 0$ the price of traded securities, as computed from the model using (10.30), should coincide with the price of those securities as observed in the market; this is known as *calibration* of the model to market data.

Martingale modelling ensures that the resulting model is arbitrage free, which is advantageous if one has to model the prices of many different securities simultaneously. The approach is therefore frequently adopted in default-free term structure models and in reduced-form models for credit-risky securities. Martingale modelling has two drawbacks. First, historical information is, to a large extent, useless in estimating model parameters, as these may change in the transition from the real-world measure to the equivalent martingale measure. Second, as illustrated in Example 10.8, realistic models for pricing credit derivatives are typically incomplete, so that one cannot eliminate all risk by dynamic hedging. In those situations one is interested in the distribution of the remaining risk under the physical measure P, so martingale modelling alone is not sufficient. In summary, the martingale-modelling approach is most suitable in situations where the market for underlying securities is relatively liquid. In that case we have sufficient price information to calibrate our models, and issues of market completeness become less relevant.

10.4.3 Bond Pricing

In this section we discuss the pricing of defaultable zero-coupon bonds in hazard rate models. Note that coupon-paying corporate bonds can be represented as a portfolio of zero-coupon bonds, so our analysis applies to coupon-paying bonds as well.

Recovery models. We begin with a survey of different models for the recovery of defaultable zero-coupon bonds. As in previous sections we denote the price at time t of a defaultable zero-coupon bond with maturity $T \geqslant t$ by $p_1(t, T)$; $p_0(t, T)$ denotes the price of the corresponding default-free zero-coupon bond. The face value of these bonds is always taken to be one. The following recovery models are frequently used in the literature.

(i) *Recovery of Treasury (RT).* The RT model was proposed by Jarrow and Turnbull (1995). Under RT, if default occurs at some point in time $\tau \leqslant T$, the owner of the defaulted bond receives $(1 - \delta_\tau)$ units of the default-free zero-coupon bond $p_0(\cdot, T)$ at time τ, where $\delta_\tau \in [0, 1]$ models the percentage loss given default. At maturity T the holder of the defaultable bond therefore receives the payment $I_{\{\tau > T\}} + (1 - \delta_\tau) I_{\{\tau \leqslant T\}}$.

(ii) *Recovery of Face Value (RF).* Under RF, if default occurs at $\tau \leqslant T$, the holder of the bond receives a (possibly random) recovery payment of size $(1 - \delta_\tau)$

immediately at the default time τ. Note that even with deterministic loss given default $\delta_\tau \equiv \delta$ and deterministic interest rates, the value at maturity of the recovery payment is random as it depends on the exact timing of default.

A further recovery model, the so-called *recovery of market value*, is considered in Section 10.5.3. In real markets, recovery is a complex issue with many legal and institutional features, and all recovery models put forward in the literature are at best a crude approximation of reality. The RF assumption comes closest to legal practice, as debt with the same seniority is assigned the same (fractional) recovery, independent of the maturity. On the other hand, for "extreme" parameter values (long maturities and high risk-free interest rates), RF may lead to negative credit spreads, as we will see in Section 10.6.3. Moreover, the RF model leads to slightly more involved pricing formulas for defaultable bonds than the RT model. Empirical evidence on recovery rates for loans and bonds is discussed in Section 11.2.3.

Bond pricing. Next we turn to pricing formulas for defaultable bonds. We use martingale modelling and work directly under some martingale measure Q. We assume that under Q the default time τ is a random time with deterministic risk-neutral hazard function $\gamma^Q(t)$. The information available to investors at time t is given by $\mathcal{H}_t = \sigma(\{Y_u : u \leqslant t\})$. We take interest rates and recovery rates to be deterministic; the percentage loss given default is denoted by $\delta \in (0,1)$, and the continuously compounded interest rate is denoted by $r(t) \geqslant 0$. Note that, in this setting, the price of the default-free zero-coupon bond with maturity $T \geqslant t$ equals

$$p_0(t,T) = \exp\left(-\int_t^T r(s)\,\mathrm{d}s\right).$$

This is the simplest type of model that can be calibrated to a given term structure of default-free interest rates and single-name credit spreads; generalizations allowing for stochastic interest rates, recovery rates and hazard rates will be discussed in Section 10.5.

The actual payments of a defaultable zero-coupon bond can be represented as a combination of a *survival claim* that pays one unit at the maturity date T and a recovery payment in case of default. The survival claim has pay-off $I_{\{\tau>T\}}$. Recall from (10.29) that

$$Q(\tau > T \mid \mathcal{H}_t) = I_{\{\tau>t\}} \exp\left(-\int_t^T \gamma^Q(s)\,\mathrm{d}s\right)$$

and define $R(t) = r(t) + \gamma^Q(t)$. The price of a survival claim at time t then equals

$$\begin{aligned} E^Q(p_0(t,T) I_{\{\tau>T\}} \mid \mathcal{H}_t) &= \exp\left(-\int_t^T r(s)\,\mathrm{d}s\right) Q(\tau > T \mid \mathcal{H}_t) \\ &= I_{\{\tau>t\}} \exp\left(-\int_t^T R(s)\,\mathrm{d}s\right). \end{aligned} \tag{10.37}$$

Note that for $\tau > t$, (10.37) can be viewed as the price of a default-free zero-coupon bond with adjusted interest rate $R(t) > r(t)$. A similar relationship between

defaultable and default-free bond prices can be established in many reduced-form credit risk models.

Under the RT model the value of the recovery payment at the maturity date T of the bond is $(1-\delta)I_{\{\tau\leqslant T\}} = (1-\delta) - (1-\delta)I_{\{\tau>T\}}$. Using (10.37), the value of the recovery payment at time $t < T$ is therefore

$$(1-\delta)p_0(t,T) - (1-\delta)I_{\{\tau>t\}} \exp\left(-\int_t^T (r(s)+\gamma^Q(s))\,\mathrm{d}s\right).$$

Under the RF hypothesis the recovery payment takes the form $(1-\delta)I_{\{\tau\leqslant T\}}$, where the payment occurs directly at time τ. Payments of this form will be referred to as *payment-at-default claims*). The value of the recovery payment at time $t \leqslant T$ therefore equals

$$E^Q\left((1-\delta)I_{\{t<\tau\leqslant T\}} \exp\left(-\int_t^\tau r(s)\,\mathrm{d}s\right) \,\Big|\, \mathcal{H}_t\right).$$

The evaluation of this expression is discussed in the following lemma.

Lemma 10.9. *Suppose that τ is a random time with hazard function $\gamma^Q(t)$, and let $R(t) = r(t) + \gamma^Q(t)$ as before. Then we have the identity*

$$\begin{aligned} &E^Q\left(I_{\{t<\tau\leqslant T\}} \exp\left(-\int_t^\tau r(s)\,\mathrm{d}s\right) \,\Big|\, \mathcal{H}_t\right) \\ &\qquad = I_{\{\tau>t\}} \int_t^T \gamma^Q(s) \exp\left(-\int_t^s R(u)\,\mathrm{d}u\right) \mathrm{d}s. \end{aligned}$$

Proof. Using Lemma 10.6 we get that

$$\begin{aligned} &E^Q\left(I_{\{t<\tau\leqslant T\}} \exp\left(-\int_t^\tau r(s)\,\mathrm{d}s\right) \,\Big|\, \mathcal{H}_t\right) \\ &\qquad = I_{\{\tau>t\}} \frac{E^Q(I_{\{t<\tau\leqslant T\}} \exp(-\int_t^\tau r(s)\,\mathrm{d}s))}{\exp(-\int_0^t \gamma^Q(s)\,\mathrm{d}s)}. \end{aligned} \tag{10.38}$$

Since τ has density

$$\gamma^Q(t) \exp\left(-\int_0^t \gamma^Q(s)\,\mathrm{d}s\right),$$

we have

$$\begin{aligned} &E^Q\left(I_{\{t<\tau\leqslant T\}} \exp\left(-\int_t^\tau r(s)\,\mathrm{d}s\right)\right) \\ &\qquad = \int_t^T \exp\left(-\int_t^s r(u)\,\mathrm{d}u\right) \gamma^Q(s) \exp\left(-\int_0^s \gamma^Q(u)\,\mathrm{d}u\right) \mathrm{d}s. \end{aligned}$$

Substitution of the right-hand side into equation (10.38) gives the result. □

10.4.4 CDS Pricing

The CDS market is among the most liquid markets for credit-risky securities, so the task of building a model using CDS spreads as input is frequently encountered in practice. In this section we therefore discuss CDS pricing and the calibration of hazard rate models to observed CDS spreads.

Pricing. We consider the following CDS contract. We take the notional to be one, so that percentage loss given default and absolute loss given default are the same. The premium payments are due at N points in time $0 < t_1 < \cdots < t_N$. If $\tau > t_k$, the protection buyer pays a premium of size $x^*(t_k - t_{k-1})$ at t_k, where x^* denotes the swap spread. After default, no further premium payments are made. If default occurs before the maturity date t_N of the swap, the protection seller makes a default payment of size δ to the buyer at the default time τ. In a standard CDS the protection buyer pays the protection seller at default the part of the premium that has accrued since the last regular premium payment date; here we ignore these accrued premium payments to simplify the exposition.

We use the same set-up as in the analysis of bond pricing in the previous section. As a first step we price the payments made by the protection buyer (the so-called premium payment leg of the swap) and the payments made by the protection seller (the default payment leg) separately, using a generic risk-neutral hazard function γ^Q and a generic spread x. The price of the premium payment leg at $t < t_N$ (the expected discounted value of the payments) is given by

$$\begin{aligned} V_t^{\text{prem}}(x; \gamma^Q) &= E^Q\Biggl(\sum_{k:\, t_k > t} \exp\biggl(-\int_t^{t_k} r(u)\,\mathrm{d}u\biggr) x(t_k - t_{k-1}) I_{\{t_k < \tau\}} \Biggm| \mathcal{H}_t \Biggr) \\ &= x \sum_{k:\, t_k > t} p_0(t, t_k)(t_k - t_{k-1}) Q(\tau > t_k \mid \mathcal{H}_t), \end{aligned} \tag{10.39}$$

which is easily computed using the formula

$$Q(\tau > t_k \mid \mathcal{H}_t) = I_{\{\tau > t\}} \exp\biggl(-\int_t^{t_k} \gamma^Q(s)\,\mathrm{d}s\biggr).$$

The default payment leg is a typical payment-at-default claim. Using Lemma 10.9 we obtain

$$\begin{aligned} V_t^{\text{def}}(\gamma^Q) &= E^Q\biggl(\exp\biggl(-\int_t^{\tau} r(u)\,\mathrm{d}u\biggr)\delta I_{\{t < \tau \leqslant t_N\}} \biggm| \mathcal{H}_t\biggr) \\ &= I_{\{t<\tau\}}\delta \int_t^{t_N} \gamma^Q(s) \exp\biggl(-\int_t^{s} (r(u) + \gamma^Q(u))\,\mathrm{d}u\biggr)\mathrm{d}s. \end{aligned} \tag{10.40}$$

According to market convention the CDS spread x_t^* quoted for the contract at time t (the so-called *fair CDS spread* x^*) is chosen such that the value of the contract is equal to zero. Hence x_t^* is defined by the equation $V_t^{\text{prem}}(x_t^*; \gamma^Q) = V_t^{\text{def}}(\gamma^Q)$, which gives

$$x_t^* = I_{\{t<\tau\}} \frac{\delta \int_t^{t_N} \gamma^Q(s) \exp(-\int_t^s (r(u) + \gamma^Q(u))\,\mathrm{d}u)\,\mathrm{d}s}{\sum_{k:\, t_k > t} p_0(t, t_k)(t_k - t_{k-1}) \exp(-\int_t^{t_k} \gamma^Q(s)\,\mathrm{d}s)}. \tag{10.41}$$

Obviously, x_t^* depends on the hazard function γ^Q, as V_t^{prem} and V_t^{def} depend on γ^Q.

Note that in the pricing argument we have neglected the issue of counterparty risk and, in particular, the possibility that the protection seller might default before the maturity of the CDS. A discussion of counterparty risk for CDS contracts is given in Section 17.2.

Calibration. Assume now that we observe spreads quoted in the market for one or more CDSs on the same reference entity. Under the martingale-modelling approach we have to calibrate our model to the available market information: that is, we have to determine the implied risk-neutral hazard function γ^Q, which ensures that the fair CDS spreads implied by the model equal the spreads that are quoted in the market.

Suppose that the market information at time $t = 0$ consists of the fair spread x^* of one CDS with maturity t_N; the risk-neutral hazard function γ^Q is constant, so that, for all $s \geqslant 0$, $\gamma^Q(s) = \bar{\gamma}^Q$ for some $\bar{\gamma}^Q > 0$, which we refer to as the risk-neutral hazard rate. It follows from (10.39) and (10.40) that the implied risk-neutral hazard rate $\bar{\gamma}^Q$ satisfies the equation

$$x^* \sum_{k=1}^{N} p_0(0, t_k)(t_k - t_{k-1})\mathrm{e}^{-\bar{\gamma}^Q t_k} = \delta\bar{\gamma}^Q \int_0^{t_N} p_0(0, t)\mathrm{e}^{-\bar{\gamma}^Q t}\,\mathrm{d}t. \tag{10.42}$$

Here, the left-hand side equals $V_0^{\text{prem}}(x^*, \bar{\gamma}^Q)$ and the right-hand side is obviously equal to $V_0^{\text{def}}(\bar{\gamma}^Q)$. There is a unique implied risk-neutral hazard rate solving equation (10.42). This may be seen by first noting that $V_0^{\text{prem}}(x^*, \bar{\gamma}^Q)$ is a decreasing function of $\bar{\gamma}^Q$ while $V_0^{\text{def}}(\bar{\gamma}^Q)$ is an increasing function of $\bar{\gamma}^Q$. Moreover, $V_0^{\text{def}}(0) = 0$, so the value of the premium payments exceeds the value of the default payment for small values of $\bar{\gamma}^Q$. On the other hand, as $\bar{\gamma}^Q$ tends to infinity, $V_0^{\text{prem}}(x^*, \bar{\gamma}^Q)$ converges to zero, so $V_0^{\text{prem}}(x^*, \bar{\gamma}^Q) < V_0^{\text{def}}(\bar{\gamma}^Q)$ for large values of $\bar{\gamma}^Q$.

If one observes spreads for several CDSs on the same reference entity but with different maturities, a time-independent risk-neutral hazard function is generally not sufficient to calibrate the model to the observed swap spreads. Instead one typically uses hazard functions $\gamma^Q(t)$ that are piecewise constant or piecewise linear. An exception occurs in the special case where (1) the spread curve is *flat* (i.e. all CDSs on the reference entity have the same spread x^*, independent of the maturity), (2) the risk-free interest rate is constant, and (3) the time points t_k are equally spaced ($t_k - t_{k-1} = \Delta t$ for all k). In that case the implied risk-neutral hazard rate $\bar{\gamma}^Q$ is the solution of equation (10.42) in the case where $N = 1$, that is, the solution of

$$x^* \Delta t p_0(0, \Delta t)\mathrm{e}^{-\bar{\gamma}^Q \Delta t} = \delta\bar{\gamma}^Q \int_0^{\Delta t} \mathrm{e}^{-rt}\mathrm{e}^{-\bar{\gamma}^Q t}\,\mathrm{d}t. \tag{10.43}$$

For Δt relatively small (quarterly or semiannual spread payments), a good approximation to the solution of (10.43) is given by $\bar{\gamma}^Q \approx x^*/\delta$, i.e. by the ratio of the fair swap spread and the percentage loss given default. This approximation is frequently used in practice.

Note, finally, that for most issuers the implied hazard rate is relatively small (of the order of a few percentage points). We therefore have the following approximation for the one-year default probability:

$$Q(\tau \leqslant 1) = 1 - \mathrm{e}^{-\bar{\gamma}^Q} \approx \bar{\gamma}^Q \approx x^*/\delta, \tag{10.44}$$

so the quantity x^*/δ can be viewed as a proxy for the risk-neutral one-year default probability.

10.4.5 *P* versus *Q*: Empirical Results

We have now assembled the necessary technical tools to discuss some of the empirical work on the relationship between physical and risk-neutral default probabilities. Understanding this relationship is important; it enables market participants to use information about historical default probabilities in pricing credit-risky securities. Conversely, it allows the use of market quotes for CDSs or defaultable bonds as additional inputs in determining historical default probabilities.

In most empirical studies risk-neutral default probabilities are estimated from credit-spread data for CDSs. By comparing these estimates with estimates of the physical default probability—obtained, for instance, from the public-firm EDF methodology introduced in Section 10.3.3—it is possible to gain some empirical evidence on the relationship between physical and risk-neutral default probabilities in real markets. An extensive empirical study along these lines is found in Berndt et al. (2008). The authors carried out a very detailed regression analysis of the observed spreads for five-year CDSs against five-year EDFs for a large pool of firms. The five-year EDF of a firm with publicly traded stock is an annualized estimate of the physical five-year default probability. The computation of EDFs is described in detail in Section 10.3.3, and annualization is a way of expressing EDFs for different time horizons on a common yearly scale.

Berndt et al. (2008) begin by estimating a linear model for the relationship between the observed swap spread $x^*_{t,i}$ of firm i at date t and the five-year EDF of that firm on the same day, labelled $\mathrm{EDF}_{t,i}$. The model takes the form

$$x^*_{t,i} = \alpha + \beta\,\mathrm{EDF}_{t,i} + \varepsilon_{t,i}, \quad (t,i) \in S, \tag{10.45}$$

where S denotes the set of all time points/firms for which there is an observable EDF–CDS pair. The model was fitted to a sample of 33 912 EDF–CDS observations for a large set of publicly traded US firms in the period December 2000 to December 2004. The estimated coefficients were given by $\alpha = 33$ bp and $\beta = 1.6$; the R^2 was 0.73.

Berndt et al. (2008) propose the following interpretation of this regression result. Their model implies that the fair swap spread x^* of a firm increases by approximately 16 basis points for every 10 basis point increase in the five-year EDF of that firm; neglecting the intercept, we thus have that $x^*_{t,i}/\mathrm{EDF}_{t,i} \approx 1.6$. Assuming a fixed loss given default δ, we may use the quantity $q_{t,i} = x^*_{t,i}/\delta$ as a proxy for the risk-neutral default probability of firm i at time t; moreover, $\mathrm{EDF}_{t,i}$ can be viewed as a proxy for the physical default probability of firm i at time t. The ratio of risk-neutral to historical default probabilities is therefore given approximately by

$$\frac{q_{t,i}}{p_{t,i}} \approx \frac{x^*_{t,i}}{\delta\,\mathrm{EDF}_{t,i}} \approx 1.6\delta^{-1}.$$

With $\delta = 0.75$ we obtain $q_{t,i}/p_{t,i} \approx 2.13$; higher recovery rates, i.e. smaller values of δ, would lead to an even higher estimate for $q_{t,i}/p_{t,i}$. The analysis of Berndt et al. (2008) clearly shows that physical and risk-neutral default probabilities can differ substantially, and care must be taken to distinguish between the two concepts.

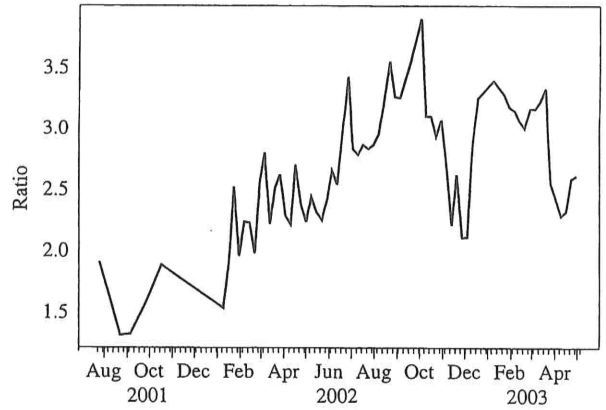


Figure 10.6. Ratio of one-year risk-neutral and historical default probabilities for Vintage Petroleum, as estimated by Berndt et al. (2008).

A careful inspection of the EDF–CDS relationship shows that the simple linear model (10.45) might not be appropriate for a number of reasons. First, the intercept of 33 basis points is implausible, as it would imply that even for a firm with historical default probability p close to zero the swap spread is still of the order of 30 basis points. Second, Berndt et al. (2008) found that the ratio x^*_i/EDF_i varies between industry sectors—reflecting different recovery rates for different industries—and over time, as is illustrated in Figure 10.6. Third, there seems to be some concavity in the relationship between swap spreads and EDFs; in particular, the ratio $x^*_{t,i}/\mathrm{EDF}_{t,i}$ is higher for high-quality firms with low EDF values than for low-quality firms. For these reasons the authors go on to consider more refined logarithmic regression models that fit the data significantly better.

Notes and Comments

Hazard rate models are a common tool in credit risk and survival analysis: see, for example, Bielecki and Rutkowski (2002) or, for a general introduction to survival analysis, the classical textbook by Cox and Oakes (1984). Further useful textbooks are Fleming and Harrington (2005), Marshall and Olkin (2007) and Aalen, Borgan and Gjessing (2010).

The fundamental theorems of asset pricing and the conceptual underpinnings of risk-neutral pricing are discussed in most textbooks on mathematical finance: see, for example, Duffie (2001), Björk (2004), Shreve (2004b) and Delbaen and Schachermayer (2006). The term martingale modelling was coined in Björk (2004) in the context of default-free short-rate models. In recent years a number of interesting approaches to the risk management of derivative securities in incomplete markets have been developed. *Quadratic hedging* approaches were first developed by Föllmer and Sondermann (1986) and Föllmer and Schweizer (1991); Schweizer (2001) is an excellent survey; *utility-based* approaches to pricing and hedging in incomplete markets are discussed in Delbaen et al. (2002) and Becherer (2004), and the latter paper explicitly considers applications of utility-based hedging strategies to credit risk models. Papers dealing with dynamic hedging and market incompleteness in credit risk models include Bielecki, Jeanblanc and Rutkowski (2004), Bielecki, Jeanblanc and Rutkowski (2007), Frey and Backhaus (2010) and Cont and Kan (2011).

A detailed analysis of CDS pricing can be found in many sources; a good reference is Schönbucher (2003). Theoretical results on the relationship between physical and risk-neutral default probabilities were obtained by Artzner and Delbaen (1995) and Jarrow, Lando and Yu (2005). In their paper, Berndt et al. (2008) go beyond the regression analysis presented in our text and estimate a full time-series model for the joint evolution of risk-neutral and actual default intensities. Further empirical studies of the relationship between actual and risk-neutral default probabilities include Fons (1994), Bohn (2000), Driessen (2005) and Huang and Huang (2012). These results largely corroborate the findings of Berndt et al. (2008).

10.5 Pricing with Stochastic Hazard Rates

In the models with deterministic hazard functions discussed in Section 10.4, the only risk factor affecting a defaultable bond or a CDS is default risk. Hence in these models credit spreads evolve deterministically prior to default, which is clearly unrealistic. Moreover, it is not possible to price options on defaultable bonds or CDSs in such models. In this section we consider models where the hazard function is replaced by a stochastic *hazard process*. In mathematical terms this leads to the notion of doubly stochastic random times, which is discussed in Section 10.5.1. In Section 10.5.2 we derive pricing formulas for certain building blocks that can be used to value many important credit-risky securities. Applications of these formulas are studied in Section 10.5.3.

10.5.1 Doubly Stochastic Random Times

We now consider a situation where additional information affecting the distribution of the random time τ is available. Formally, we represent this additional information by some filtration $(\mathcal{F}_t)$ on the underlying probability space $(\Omega, \mathcal{F}, P)$. In credit risk models this information is typically generated by some background process (Ψ_t) representing, for instance, the risk-free interest rate or various measures of economic activity, so that $\mathcal{F}_t = \sigma(\{\Psi_s : s \leqslant t\})$.

Consider some random time τ on $(\Omega, \mathcal{F}, P)$ with $P(\tau > 0) = 1$ and denote by $Y_t = I_{\{\tau \leqslant t\}}$ the associated jump indicator and by $(\mathcal{H}_t)$ the filtration generated by (Y_t) (see equation (10.26)). We introduce a new filtration $(\mathcal{G}_t)$ by

$$\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t, \quad t \geqslant 0, \tag{10.46}$$

meaning that $\mathcal{G}_t$ is the smallest σ-algebra that contains $\mathcal{F}_t$ and $\mathcal{H}_t$. We will frequently use the notation $(\mathcal{G}_t) = (\mathcal{F}_t) \vee (\mathcal{H}_t)$ below. The filtration $(\mathcal{G}_t)$ contains information about the background processes and the occurrence or non-occurrence of default up to time t, and thus typically corresponds to the information available

to investors. Obviously, τ is an $(\mathcal{H}_t)$ stopping time and hence also a $(\mathcal{G}_t)$ stopping time. Note, however, that we do not assume that τ is a stopping time with respect to the background filtration $(\mathcal{F}_t)$.

Doubly stochastic random times are a straightforward extension of the models considered in Section 10.4 to the present set-up with additional information.

Definition 10.10 (doubly stochastic random time). A random time τ is said to be doubly stochastic if there exists a positive $(\mathcal{F}_t)$-adapted process (γ_t) such that $\Gamma_t = \int_0^t \gamma_s \, ds$ is strictly increasing and finite for every $t > 0$ and such that, for all $t \geqslant 0$,

$$P(\tau > t \mid \mathcal{F}_\infty) = \exp\left(-\int_0^t \gamma_s \, ds\right). \tag{10.47}$$

In that case (γ_t) is referred to as the $(\mathcal{F}_t)$-*conditional hazard process* of τ.

In (10.47) $\mathcal{F}_\infty$ denotes the smallest σ-algebra that contains $\mathcal{F}_t$ for all $t \geqslant 0$: that is, $\mathcal{F}_\infty = \sigma(\bigcup_{t \geqslant 0} \mathcal{F}_t)$. Conditioning on $\mathcal{F}_\infty$ thus means that we know the past and future economic environment and in particular the entire trajectory $(\gamma_s(\omega))_{s \geqslant 0}$ of the hazard rate process. Hence (10.47) implies that, given the economic environment, τ is a random time with deterministic hazard function given by the mapping $s \mapsto \gamma_s(\omega)$. The term *doubly stochastic* obviously refers to the fact that the hazard rate at any time is itself a realization of a stochastic process. In the literature, doubly stochastic random times are also known as *conditional Poisson* or *Cox* random times. Note, finally, that (10.47) implies that $P(\tau \leqslant t \mid \mathcal{F}_\infty)$ is $\mathcal{F}_t$-measurable, so we have the equality

$$P(\tau \leqslant t \mid \mathcal{F}_\infty) = P(\tau \leqslant t \mid \mathcal{F}_t). \tag{10.48}$$

In the next lemma we give an explicit construction of doubly stochastic random times. This construction is very useful for simulation purposes.

Lemma 10.11. *Let X be a standard exponentially distributed rv on $(\Omega, \mathcal{F}, P)$ independent of $\mathcal{F}_\infty$, i.e. $P(X \leqslant t \mid \mathcal{F}_\infty) = 1 - e^{-t}$ for all $t \geqslant 0$. Let (γ_t) be a positive $(\mathcal{F}_t)$-adapted process such that $\Gamma_t = \int_0^t \gamma_s \, ds$ is strictly increasing and finite for every $t > 0$. Define the random time τ by*

$$\tau := \Gamma^{\leftarrow}(X) = \inf\{t \geqslant 0 : \Gamma_t \geqslant X\}. \tag{10.49}$$

Then τ is doubly stochastic with $(\mathcal{F}_t)$-conditional hazard rate process (γ_t).

Proof. Note that by definition of τ it holds that $\{\tau > t\} = \{\Gamma_t < X\}$. Since Γ_t is $\mathcal{F}_\infty$-measurable and X is independent of $\mathcal{F}_\infty$, we obtain

$$P(\tau > t \mid \mathcal{F}_\infty) = P(\Gamma_t < X \mid \mathcal{F}_\infty) = e^{-\Gamma_t},$$

which proves the claim. □

Lemma 10.11 has the following converse.

Lemma 10.12. *Let τ be a doubly stochastic random time with $(\mathcal{F}_t)$-conditional hazard process (γ_t). Denote by $\Gamma_t = \int_0^t \gamma_s \, ds$ the $(\mathcal{F}_t)$-conditional cumulative hazard process of τ and set $X := \Gamma_\tau$. Then the rv X is standard exponentially distributed and independent of $\mathcal{F}_\infty$, and $\tau = \Gamma^{\leftarrow}(X)$ almost surely.*

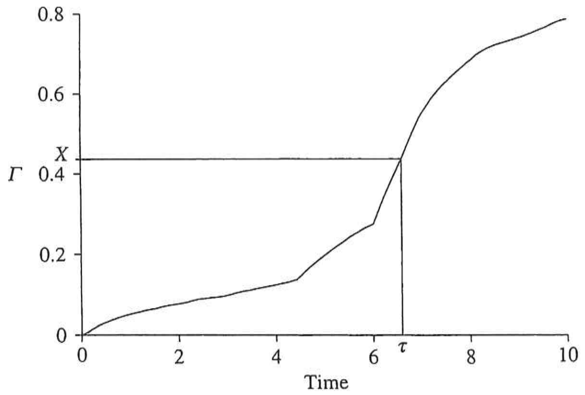


Figure 10.7. A graphical illustration of Algorithm 10.13: $X \approx 0.44$, $\tau \approx 6.59$.

Proof. Since (Γ_t) is strictly increasing by assumption, the relation $\tau = \Gamma^{\leftarrow}(X)$ is clear from the definition of X. To prove that X has the correct distribution we argue as follows:

$$P(X \leqslant t \mid \mathcal{F}_\infty) = P(\Gamma_\tau \leqslant t \mid \mathcal{F}_\infty) = P(\tau \leqslant \Gamma^{\leftarrow}(t) \mid \mathcal{F}_\infty).$$

Since τ is doubly stochastic, the last expression equals $1 - \exp(-\Gamma(\Gamma^{\leftarrow}(t))) = 1 - \mathrm{e}^{-t}$, as Γ is continuous and strictly increasing by assumption. This shows that X is independent of $\mathcal{F}_\infty$ and that it is standard exponentially distributed. □

Lemma 10.11 forms the basis for the following algorithm for the simulation of doubly stochastic random times.

Algorithm 10.13 (univariate threshold simulation).

(1) Generate a trajectory of the hazard process (γ_t). References for suitable simulation approaches are given in Notes and Comments.

(2) Generate a unit exponential rv X independent of (γ_t) (the threshold) and set $\tau = \Gamma^{\leftarrow}(X)$; this step is illustrated in Figure 10.7.

Moreover, Lemmas 10.11 and 10.12 provide an interesting interpretation of doubly stochastic random times in terms of *operational time*. For a given $(\mathcal{F}_t)$-adapted hazard process (γ_t), define a new timescale (operational time) by the associated cumulative hazard process $\Gamma_t = \int_0^t \gamma_s \,\mathrm{d}s$, so that c units of operational time correspond to $\Gamma^{\leftarrow}(c)$ units of real time. Take a standard exponential rv X independent of $\mathcal{F}_\infty$ and measure time in units of operational time. The associated calendar time $\tau := \Gamma^{\leftarrow}(X)$ is then doubly stochastic by Lemma 10.11. Conversely, by Lemma 10.12, if we take a doubly stochastic random time τ, the associated operational time $X := \Gamma_\tau$ is standard exponential, independent of $\mathcal{F}_\infty$. The notion of operational time plays an important role in insurance mathematics (see Section 13.2.7).

Intensity of doubly stochastic random times. We have seen in Proposition 10.7 that the jump indicator process (Y_t) can be turned into an $(\mathcal{H}_t)$-martingale if we subtract the process $\int_0^{t\wedge\tau} \gamma(s)\,\mathrm{d}s$, where $t \wedge \tau$ is a shorthand notation for $\min\{t, \tau\}$. We now generalize this result to doubly stochastic random times.

Proposition 10.14. *Let τ be a doubly stochastic random time with $(\mathcal{F}_t)$-conditional hazard process (γ_t). Then $M_t := Y_t - \int_0^{t\wedge\tau} \gamma_s \,\mathrm{d}s$ is a $(\mathcal{G}_t)$-martingale.*

Proof. Define a new artificial filtration $(\tilde{\mathcal{G}}_t)$ by $\tilde{\mathcal{G}}_t = \mathcal{F}_\infty \vee \mathcal{H}_t$, and note that $\tilde{\mathcal{G}}_0 = \mathcal{F}_\infty$ and $\mathcal{G}_t \subset \tilde{\mathcal{G}}_t$ for all t. As explained above, given $\mathcal{F}_\infty$, τ is a random time with deterministic hazard rate. Proposition 10.7 implies that $M_t := Y_t - \int_0^{t\wedge\tau} \gamma_s \,\mathrm{d}s$ is a martingale with respect to $(\tilde{\mathcal{G}}_t)$. Since (M_t) is $(\mathcal{G}_t)$-adapted and $\mathcal{G}_t \subset \tilde{\mathcal{G}}_t$, (M_t) is also a martingale with respect to $(\mathcal{G}_t)$. □

Finally, we relate Proposition 10.14 to the popular notion of the *intensity* of a random time.

Definition 10.15. Consider a filtration $(\mathcal{G}_t)$ and a random time τ with $(\mathcal{G}_t)$-adapted jump indicator process (Y_t). A non-negative $(\mathcal{G}_t)$-adapted process (λ_t) is called a *$(\mathcal{G}_t)$-intensity* process of the random time τ if $M_t := Y_t - \int_0^{t\wedge\tau} \lambda_s \,\mathrm{d}s$ is a $(\mathcal{G}_t)$-martingale.

In reduced-form credit risk models, (λ_t) is usually called the *default intensity* of the default time τ. It is well known that the intensity (λ_t) is uniquely defined on $\{t < \tau\}$. This is an immediate consequence of general results from stochastic calculus concerning the uniqueness of semimartingale decompositions (see, for example, Chapter 2 of Protter (2005)). Using the terminology of Definition 10.15, we may restate Proposition 10.14 in the following form: "the $(\mathcal{G}_t)$-intensity of a doubly stochastic random time τ is given by its $(\mathcal{F}_t)$-conditional hazard process (γ_t)". At this point a warning is in order: there are random times that admit an intensity in the sense of Definition 10.15 that are not doubly stochastic and for which the pricing formulas derived in Section 10.5.2 below do not hold.

Conditional expectations. Next we discuss the structure of conditional expectations with respect to the full-information σ-algebra $\mathcal{G}_t$; these results are crucial for the derivation of pricing formulas in models with doubly stochastic default times.

Proposition 10.16. *Let τ be an arbitrary random time (not necessarily doubly stochastic) such that $P(\tau > t \mid \mathcal{F}_t) > 0$ for all $t \geqslant 0$. We then have for every integrable rv X that*

$$E(I_{\{\tau>t\}} X \mid \mathcal{G}_t) = I_{\{\tau>t\}} \frac{E(I_{\{\tau>t\}} X \mid \mathcal{F}_t)}{P(\tau > t \mid \mathcal{F}_t)}.$$

Note that Proposition 10.16 allows us to replace certain conditional expectations with respect to $\mathcal{G}_t$ by conditional expectations with respect to the background information $\mathcal{F}_t$. The result is also known as the *Dellacherie formula*. In the special case where the background filtration is trivial, i.e. $\mathcal{F}_t = \{\emptyset, \Omega\}$ for all $t \geqslant 0$, Proposition 10.16 reduces to Lemma 10.6.

Proof. Standard measure-theoretic arguments show that for every $\mathcal{G}_t$-measurable rv X there is some $\mathcal{F}_t$-measurable rv $\tilde{X}$ such that $XI_{\{\tau>t\}} = \tilde{X}I_{\{\tau>t\}}$. This is quite intuitive since prior to default all information is generated by the background filtration $(\mathcal{F}_t)$; a formal proof is given in Section 5.1.1 of Bielecki and Rutkowski (2002). Now $E(I_{\{\tau>t\}}X \mid \mathcal{G}_t)$ is $\mathcal{G}_t$-measurable and zero on $\{\tau \leqslant t\}$. There is therefore an $\mathcal{F}_t$-measurable rv $\tilde{Z}$ such that $E(I_{\{\tau>t\}}X \mid \mathcal{G}_t) = I_{\{\tau>t\}}\tilde{Z}$. Taking conditional expectations with respect to $\mathcal{F}_t$ and noting that $\mathcal{F}_t \subset \mathcal{G}_t$ yields

$$E(I_{\{\tau>t\}}X \mid \mathcal{F}_t) = P(\tau > t \mid \mathcal{F}_t)\tilde{Z}.$$

Hence $\tilde{Z} = E(I_{\{\tau>t\}}X \mid \mathcal{F}_t)/P(\tau > t \mid \mathcal{F}_t)$, which proves the lemma. □

Corollary 10.17. *Let $T > t$ and assume that τ is doubly stochastic with hazard process (γ_t). If the rv $\tilde{X}$ is integrable and $\mathcal{F}_T$-measurable, we have*

$$E(I_{\{\tau>T\}}\tilde{X} \mid \mathcal{G}_t) = I_{\{\tau>t\}}E\bigg(\exp\bigg(-\int_t^T \gamma_s\,\mathrm{d}s\bigg)\tilde{X} \,\bigg|\, \mathcal{F}_t\bigg).$$

Proof. Let $X := I_{\{\tau>T\}}\tilde{X}$. Since $X = I_{\{\tau>t\}}X$ (as $T > t$), Proposition 10.16 yields

$$E(I_{\{\tau>T\}}\tilde{X} \mid \mathcal{G}_t) = E(I_{\{\tau>t\}}X \mid \mathcal{G}_t) = I_{\{\tau>t\}}\mathrm{e}^{\int_0^t \gamma_s\,\mathrm{d}s}E(I_{\{\tau>T\}}\tilde{X} \mid \mathcal{F}_t),$$

where we have used the fact that

$$P(\tau > t \mid \mathcal{F}_t) = \exp\bigg(-\int_0^t \gamma_s\,\mathrm{d}s\bigg).$$

Since $\tilde{X}$ is $\mathcal{F}_T$-measurable,

$$E(I_{\{\tau>T\}}\tilde{X} \mid \mathcal{F}_t) = E(\tilde{X}P(\tau > T \mid \mathcal{F}_T) \mid \mathcal{F}_t) = E\bigg(\tilde{X}\exp\bigg(-\int_0^T \gamma_s\,\mathrm{d}s\bigg) \,\bigg|\, \mathcal{F}_t\bigg),$$

and the result follows. □

Corollary 10.17 will be very useful for the pricing of various credit-risky securities in models with doubly stochastic default times. Moreover, the corollary implies that in the above setting γ_t gives a good approximation to the one-year default probability. This follows by setting $T = t + 1$ and $\tilde{X} = 1$ to obtain

$$P(\tau > t + 1 \mid \mathcal{G}_t) = I_{\{\tau>t\}}E\bigg(\exp\bigg(-\int_t^{t+1} \gamma_s\,\mathrm{d}s\bigg) \,\bigg|\, \mathcal{F}_t\bigg). \tag{10.50}$$

Now assume that $\tau > t$ and that the hazard rate remains relatively stable over the time interval $[t, t + 1]$. Under these assumptions the right-hand side of (10.50) is approximated reasonably well by $\mathrm{e}^{-\gamma_t}$ and, if γ_t is small, the one-year default probability satisfies

$$P(\tau \leqslant t + 1 \mid \mathcal{G}_t) \approx 1 - \mathrm{e}^{-\gamma_t} \approx \gamma_t. \tag{10.51}$$

10.5.2 *Pricing Formulas*

The main result of this section concerns the pricing of three types of contingent claims that can be used as building blocks for constructing the pay-off of many important credit-risky securities. We will show that, for a default time that is doubly stochastic, the computation of prices for these claims can be reduced to a pricing problem for a corresponding default-free claim if we adjust the interest rate and replace the default-free interest rate r_t by the sum $R_t = r_t + \gamma_t$ of the default-free interest rate and the hazard rate of the default time.

The model. We consider a firm whose default time is given by a doubly stochastic random time as in Section 10.5.1. The economic background filtration represents the information generated by an arbitrage-free and complete model for non-defaultable security prices. More precisely, let $(\Omega, \mathcal{F}, (\mathcal{F}_t), Q)$ denote a filtered probability space, where Q is the equivalent martingale measure. Prices of default-free securities such as default-free bonds and the default-free rate of interest (r_t) are $(\mathcal{F}_t)$-adapted processes; $B_t = \exp(\int_0^t r_s\,\mathrm{d}s)$ models the default-free savings account.

Let τ be the default time of some company under consideration and let $Y_t = I_{\{\tau\leqslant t\}}$ be the associated default indicator process. As before we set $\mathcal{H}_t = \sigma(\{Y_s : s \leqslant t\})$ and $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$; we assume that default is observable and that investors have access to the information contained in the background filtration $(\mathcal{F}_t)$, so that the information available to investors at time t is given by $\mathcal{G}_t$. We consider a market for credit products that is liquid enough that we may use the martingale-modelling approach, and we use Q as the pricing measure for defaultable securities. According to (10.30), the price at time t of an arbitrary, non-negative, $\mathcal{G}_T$-measurable contingent claim H is therefore given by

$$H_t = E^Q\bigg(\exp\bigg(-\int_t^T r_s\,\mathrm{d}s\bigg)H \,\bigg|\, \mathcal{G}_t\bigg). \tag{10.52}$$

Finally, we assume that, under Q, the default time τ is a doubly stochastic random time with background filtration $(\mathcal{F}_t)$ and hazard process (γ_t). This latter assumption is crucial for the results that follow.

Definition 10.18. We introduce the following building blocks.

(i) A *survival claim*, i.e. an $\mathcal{F}_T$-measurable promised payment X that is made at time T if there is no default; the actual payment of the survival claim equals $XI_{\{\tau>T\}}$.

(ii) A *risky dividend stream*. Here, we consider a promised dividend stream given by the $(\mathcal{F}_t)$-adapted rate process ν_s, $0 \leqslant s \leqslant T$. The payments of a risky dividend stream stop when default occurs, so that the actual payments of this building block are given by the dividend stream with rate $\nu_t I_{\{\tau>t\}}$, $0 \leqslant t \leqslant T$.

(iii) A *payment-at-default claim* of the form $Z_\tau I_{\{\tau\leqslant T\}}$, where $Z = (Z_t)_{t\geqslant 0}$ is an $(\mathcal{F}_t)$-adapted stochastic process and where Z_τ is short for $Z_{\tau(\omega)}(\omega)$. Note that the payment is made directly at τ, provided that $\tau \leqslant T$, where T is the maturity date of the claim.

Recall from Section 10.4.3 that defaultable bonds can be viewed as portfolios of survival claims and payment-at-default claims. Credit default swaps can also be written as a combination of these claims, as will be shown in Section 10.5.3. A further example is provided by option contracts that are subject to counterparty risk. For concreteness we consider a call option on some default-free security (S_t). Denote the exercise price by K and the maturity date by T and suppose that if the writer defaults at time $\tau \leqslant T$, then the owner of the option receives a fraction $(1-\delta_\tau)$ of the intrinsic value of the option at the time of default. This can be modelled as a combination of the survival claim $(S_T - K)^+ I_{\{\tau > T\}}$ and the payment-at-default claim $(1-\delta_\tau)(S_\tau - K)^+ I_{\{\tau \leqslant T\}}$.

Pricing results. In the following theorem we show that the pricing of the building blocks introduced in Definition 10.18 can be reduced to a pricing problem in a default-free security market model with investor information given by the background filtration $(\mathcal{F}_t)$ and with adjusted default-free interest rate.

Theorem 10.19. *Suppose that, under Q, τ is doubly stochastic with background filtration $(\mathcal{F}_t)$ and hazard process (γ_t). Define $R_s := r_s + \gamma_s$. Assume that the rvs $\exp(-\int_t^T r_s\,ds)|X|$, $\int_t^T |\nu_s| \exp(-\int_t^s r_u\,du)\,ds$ and $\int_t^T |Z_s\gamma_s| \exp(-\int_t^s R_u\,du)\,ds$ are all integrable with respect to Q. Then the following identities hold:*

$$E^Q\left(\exp\left(-\int_t^T r_s\,ds\right) I_{\{\tau>T\}} X \;\Big|\; \mathcal{G}_t\right) = I_{\{\tau>t\}} E^Q\left(\exp\left(-\int_t^T R_s\,ds\right) X \;\Big|\; \mathcal{F}_t\right), \tag{10.53}$$

$$E^Q\left(\int_t^T \nu_s I_{\{\tau>s\}} \exp\left(-\int_t^s r_u\,du\right) ds \;\Big|\; \mathcal{G}_t\right) = I_{\{\tau>t\}} E^Q\left(\int_t^T \nu_s \exp\left(-\int_t^s R_u\,du\right) ds \;\Big|\; \mathcal{F}_t\right), \tag{10.54}$$

$$E^Q\left(I_{\{t<\tau\leqslant T\}} \exp\left(-\int_t^\tau r_s\,ds\right) Z_\tau \;\Big|\; \mathcal{G}_t\right) = I_{\{\tau>t\}} E^Q\left(\int_t^T Z_s\gamma_s \exp\left(-\int_t^s R_u\,du\right) ds \;\Big|\; \mathcal{F}_t\right). \tag{10.55}$$

Proof. The integrability conditions ensure that all conditional expectations are well defined. We start with the pricing formula (10.53) for the vulnerable claim. Define the $\mathcal{F}_T$-measurable rv $\tilde{X} := \exp(-\int_t^T r_s\,ds)X$. Using Corollary 10.17 with $s = T$ and $\Gamma_t = \int_0^t \gamma_s\,ds$ we find that

$$E^Q(\tilde{X} I_{\{\tau>T\}} \mid \mathcal{G}_t) = I_{\{\tau>t\}} E^Q(\exp(-(\Gamma_T - \Gamma_t))\tilde{X} \mid \mathcal{F}_t).$$

Noting that $\Gamma_T - \Gamma_t = \int_t^T \gamma_s\,ds$ and using the definition of $\tilde{X}$, it follows that the right-hand side equals $I_{\{\tau>t\}} E^Q(\exp(-\int_t^T R_s\,ds)X \mid \mathcal{F}_t)$. The pricing formula (10.54) follows from (10.53) and the Fubini Theorem for conditional expectations. Finally,

we turn to (10.55). Lemma 10.16 implies that

$$E^Q\left(I_{\{\tau>t\}} \exp\left(-\int_t^\tau r_s\,ds\right) Z_\tau I_{\{\tau\leqslant T\}} \;\Big|\; \mathcal{G}_t\right) = I_{\{\tau>t\}} \frac{E^Q(I_{\{\tau>t\}} \exp(-\int_t^\tau r_s\,ds) Z_\tau I_{\{\tau\leqslant T\}} \mid \mathcal{F}_t)}{P(\tau>t \mid \mathcal{F}_t)}. \tag{10.56}$$

Now note that

$$P(\tau \leqslant t \mid \mathcal{F}_T) = 1 - \exp\left(-\int_0^t \gamma_s\,ds\right),$$

so the conditional density of τ given $\mathcal{F}_T$ equals $f_{\tau|\mathcal{F}_T}(t) = \gamma_t \exp(-\int_0^t \gamma_s\,ds)$. Hence

$$\begin{aligned} E^Q\left(I_{\{\tau>t\}} \exp\left(-\int_t^\tau r_s\,ds\right) Z_\tau I_{\{\tau\leqslant T\}} \;\Big|\; \mathcal{F}_T\right) &= \int_t^T \exp\left(-\int_t^s r_u\,du\right) Z_s\gamma_s \exp\left(-\int_0^s \gamma_u\,du\right) ds \\ &= \exp\left(-\int_0^t \gamma_u\,du\right) \int_t^T Z_s\gamma_s \exp\left(-\int_t^s R_u\,du\right) ds. \end{aligned}$$

Using iterated conditional expectations we obtain the formula

$$E^Q\left(I_{\{\tau>t\}} \exp\left(-\int_t^\tau r_s\,ds\right) Z_\tau I_{\{\tau\leqslant T\}} \;\Big|\; \mathcal{F}_t\right) = \exp\left(-\int_0^t \gamma_u\,du\right) E^Q\left(\int_t^T Z_s\gamma_s \exp\left(-\int_t^s R_u\,du\right) ds \;\Big|\; \mathcal{F}_t\right),$$

and the identity (10.55) follows from (10.56). □

10.5.3 Applications

Credit default swaps. We extend our analysis of Section 10.4.4 and discuss the pricing of CDSs in models where the default time is doubly stochastic. This allows us to incorporate stochastic interest rates, recovery rates and hazard rates into the analysis.

We quickly recall the form of the payments of the CDS contract. As in our previous analysis, the premium payments are due at N points in time $0 < t_1 < \cdots < t_N$; at a pre-default date t_k, the protection buyer pays a premium of size $x(t_k - t_{k-1})$, where x denotes the swap spread in percentage points (again we take the nominal of the swap to be one). If $\tau \leqslant t_N$, the protection seller makes a default payment of size δ_τ to the buyer at the default time τ, where the percentage loss given default is now a general $(\mathcal{F}_t)$-adapted process. Using Theorem 10.19, both legs of the swap can be priced. The regular premium payments constitute a sequence of survival claims.

Using (10.53) the fair price of the premium leg at $t = 0$ is

$$V^{\text{prem},1} = \sum_{k=1}^{N} E^Q\Big(\exp\Big(-\int_0^{t_k} r_u\,du\Big)x(t_k - t_{k-1})I_{\{t_k<\tau\}}\Big)$$
$$= x\sum_{k=1}^{N}(t_k - t_{k-1})E^Q\Big(\exp\Big(-\int_0^{t_k} R_u\,du\Big)\Big).$$

The default payment leg is a payment-at-default claim with $Z_s = \delta_s$ and maturity t_N, so its value is given by $V^{\text{def}} = E^Q(\int_0^{t_N} \delta_s\gamma_s \exp(-\int_0^s R_u\,du)\,ds)$. We have therefore reduced the pricing of credit default swaps to a pricing problem in the default-free world. Methods for solving this problem will be discussed in the next section.

Recovery of market value. Recovery of market value, abbreviated RM, is an alternative recovery model for defaultable bonds and other credit-risky securities that has been put forward by Duffie and Singleton (1999); its main virtue is the fact that it leads to particularly simple pricing formulas. Consider a claim whose payoff consists of the survival claim X and a recovery payment at the default time. Under the RM hypothesis it is assumed that this recovery payment is given by $(1-\delta_\tau)V_\tau I_{\{\tau\leqslant T\}}$, where the $(\mathcal{F}_t)$-adapted process $(\delta_t) \in (0, 1)$ gives the percentage loss given default of the claim and where the $(\mathcal{F}_t)$-adapted process (V_t) gives the pre-default value of the claim. Note that this is a recursive definition, as the pre-default value at time t also depends on the form of the recovery payments in the time period $(t, T]$. Nonetheless, the following result can be established.

Proposition 10.20. *Suppose that, under Q, τ is doubly stochastic with hazard rate process (γ_t). Suppose, moreover, that X is integrable and that the RM assumption holds. Then the pre-default value process (V_t) is uniquely determined and is given by*

$$V_t = E^Q\Big(\exp\Big(-\int_t^T (r_s + \delta_s\gamma_s)\,ds\Big)X \;\Big|\; \mathcal{F}_t\Big), \quad 0 \leqslant t \leqslant T. \tag{10.57}$$

Note that for $\delta_t \equiv 1$ the claim is a standard survival claim; in that case, (10.57) reduces to the formula (10.53). On the other hand, for $\delta_t \equiv 0$ the claim is essentially default free; in that case, (10.57) reduces to the standard pricing formula for the claim X in a default-free security market model. For a proof of Proposition 10.20 we refer to the references given in Notes and Comments.

Credit spreads and hazard rates. With doubly stochastic default times the risk-neutral hazard process (γ_t) and the credit spread

$$c(t, T) = -\frac{1}{T-t}(\ln p_1(t, T) - \ln p_0(t, T))$$

of defaultable bonds are closely related. Analytic results are most easily derived for the instantaneous credit spread given by

$$c(t, t) = \lim_{T\to t} c(t, T) = -\frac{\partial}{\partial T}\Big|_{T=t}(\ln p_1(t, T) - \ln p_0(t, T)). \tag{10.58}$$

Assume that $\tau > t$, so that $p_1(t, t) = p_0(t, t) = 1$. We therefore obtain

$$\frac{\partial}{\partial T}\Big|_{T=t} \ln p_1(t, T) = \frac{\partial}{\partial T}\Big|_{T=t} p_1(t, T), \tag{10.59}$$

and similarly for $p_0(t, T)$. To compute the derivative in (10.59) we need to distinguish between the different recovery models. Under the RM hypothesis we can apply Proposition 10.20 with $X = 1$. Exchanging expectation and differentiation we obtain

$$-\frac{\partial}{\partial T}\Big|_{T=t} p_1(t, T) = -E^Q\Big(\frac{\partial}{\partial T}\Big|_{T=t} \exp\Big(-\int_t^T (r_s + \delta_s\gamma_s)\,ds\Big) \;\Big|\; \mathcal{F}_t\Big)$$
$$= r_t + \delta_t\gamma_t. \tag{10.60}$$

Applying (10.60) with $\delta_t \equiv 0$ yields

$$-\frac{\partial}{\partial T}\Big|_{T=t} p_0(t, T) = r_t,$$

so that $c(t, t) = \delta_t\gamma_t$, i.e. the instantaneous credit spread equals the product of the hazard rate and the percentage loss given default, which is quite intuitive from an economic point of view. Under the RF recovery model, $p_1(t, T)$ is given by the sum of the price of a survival claim $I_{\{\tau>T\}}$ and a payment at default of size $(1-\delta_\tau)$. Equation (10.60) with $\delta_t \equiv 1$ shows that the derivative with respect to T of the survival claim at $T = t$ is equal to $-(r_t + \gamma_t)$. For the recovery payment we get

$$\frac{\partial}{\partial T}\Big|_{T=t} E\Big(\int_t^T \gamma_s(1-\delta_s)\exp\Big(-\int_t^s R_u\,du\Big)\,ds \;\Big|\; \mathcal{F}_t\Big) = (1-\delta_t)\gamma_t.$$

Hence

$$-\frac{\partial}{\partial T}\Big|_{T=t} p_1(t, T) = r_t + \gamma_t - (1-\delta_t)\gamma_t = r_t + \delta_t\gamma_t,$$

so that $c_1(t, t)$ is again equal to $\delta_t\gamma_t$. An analogous computation shows that we also have $c_1(t, t) = \delta_t\gamma_t$ under RT. However, for $T - t > 0$, the credit spread corresponding to the different recovery models differs, as is illustrated in Section 10.6.3.

Notes and Comments

The material discussed in this section is based on many sources. We mention in particular the books by Lando (2004) and Bielecki and Rutkowski (2002). The text by Bielecki and Rutkowski is more technical than our presentation; among other things the authors discuss various probabilistic characterizations of doubly stochastic random times. The threshold-simulation approach for doubly stochastic random times requires the simulation of trajectories of the hazard process. An excellent source for simulation techniques for stochastic processes is Glasserman (2003).

More general reduced-form models where the default time τ is not doubly stochastic are discussed, for example, in Kusuoka (1999), Elliott, Jeanblanc and Yor (2000), Bélanger, Shreve and Wong (2004), Collin-Dufresne, Goldstein and Hugonnier (2004) and Blanchet-Scalliet and Jeanblanc (2004).

Theorem 10.19 is originally due to Lando (1998); related results were obtained by Jarrow and Turnbull (1995) and Jarrow, Lando and Turnbull (1997). Proposition 10.20 is due to Duffie and Singleton (1999); extensions are discussed in Becherer and Schweizer (2005). An excellent text for the overall mathematical background is Jeanblanc, Yor and Chesney (2009).

The analogy with default-free term structure models makes the reduced-form models with doubly stochastic default times relatively easy to apply. However, some care is required in interpreting the results and applying the *linear pricing rules* for corporate debt that the models imply. In particular, one must bear in mind that in these models the default intensity does not explicitly take into account the structure of a firm's outstanding risky debt. A formal analysis of the effect of debt structure on bond values is best carried out in the context of firm-value models, where the default is explicitly modelled in terms of fundamental economic quantities. A good discussion of these issues can be found in Chapter 2 of Lando (2004).

10.6 Affine Models

In order to apply the pricing formulas for doubly stochastic random times obtained in Theorem 10.19 we need effective ways to evaluate the conditional expectations on the right-hand side of equations (10.53), (10.54) and (10.55). In most models, where default is modelled by a doubly stochastic random time, (r_t) and (γ_t) are modelled as functions of some p-dimensional Markovian state variable process $(\boldsymbol{\Psi}_t)$ with state space given by the domain $D \subset \mathbb{R}^p$, so that the natural background filtration is given by $(\mathcal{F}_t) = \sigma(\{\boldsymbol{\Psi}_s : s \leqslant t\})$. Moreover, $R_t := r_t + \gamma_t$ is of the form $R_t = R(\boldsymbol{\Psi}_t)$ for some function $R\colon D \subseteq \mathbb{R}^p \to \mathbb{R}_+$. We therefore have to compute conditional expectations of the form

$$E\left(\exp\left(-\int_t^T R(\boldsymbol{\Psi}_s)\,\mathrm{d}s\right)g(\boldsymbol{\Psi}_T) + \int_t^T h(\boldsymbol{\Psi}_s)\exp\left(-\int_t^s R(\boldsymbol{\Psi}_u)\,\mathrm{d}u\right)\mathrm{d}s \;\Big|\; \mathcal{F}_t\right) \tag{10.61}$$

for generic functions $g, h\colon D \to \mathbb{R}_+$. Since $(\boldsymbol{\Psi}_t)$ is a Markov process, this conditional expectation is given by some function $f(t, \boldsymbol{\Psi}_t)$ of time and the current value $\boldsymbol{\Psi}_t$ of the state variable process. It is well known that under some additional regularity assumptions the function f can be computed as solution of a parabolic partial differential equation (PDE)—this is the celebrated *Feynman–Kac formula*. The Feynman–Kac formula provides a way to determine f using analytical or numerical techniques for PDEs. In particular, it is known that in the case where $(\boldsymbol{\Psi}_t)$ belongs to the class of *affine jump diffusions* (see below), R is an affine function, $g(\boldsymbol{\psi}) = \mathrm{e}^{\boldsymbol{u}'\boldsymbol{\psi}}$ for some $\boldsymbol{u} \in \mathbb{R}^p$ and $h \equiv 0$, the function f takes the form

$$f(t, \boldsymbol{\psi}) = \exp(\alpha(t, T) + \beta(t, T)'\boldsymbol{\psi}) \tag{10.62}$$

for deterministic functions $\alpha\colon [0, T] \to \mathbb{R}$ and $\beta\colon [0, T] \to R^p$; moreover, α and β are determined by a $(p+1)$-dimensional ordinary differential equation (ODE) system that is easily solved numerically. Models based on affine jump diffusions and an affine specification of R are therefore relatively easy to implement, which

explains their popularity in practice. A relationship of the form (10.62) is often termed an *affine term structure*, as it implies that continuously compounded yields of bonds at time t are affine functions of $\boldsymbol{\Psi}_t$.

In this section we discuss these results. We concentrate on the case where the state variable process is given by a one-dimensional diffusion; extensions to processes with jumps will be considered briefly at the end.

10.6.1 *Basic Results*

The PDE characterization of f. We assume that the state variable process (Ψ_t) is the unique solution of the SDE

$$\mathrm{d}\Psi_t = \mu(\Psi_t)\,\mathrm{d}t + \sigma(\Psi_t)\,\mathrm{d}W_t, \quad \Psi_0 = \psi \in D, \tag{10.63}$$

with state space given by the domain $D \subseteq \mathbb{R}$. Here, (W_t) is a standard, one-dimensional Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, and μ and σ are continuous functions from D to $\mathbb{R}$ and D to $\mathbb{R}_+$, respectively. The next result shows that the conditional expectation (10.61) can be computed as the solution of a parabolic PDE.

Lemma 10.21 (Feynman–Kac). *Consider generic functions $R, g, h\colon D \to \mathbb{R}_+$. Suppose that the function $f\colon [0, T] \times D \to \mathbb{R}$ is continuous, once continuously differentiable in t and twice continuously differentiable in ψ on $[0, T) \times D$, and that f solves the terminal-value problem*

$$\left.\begin{aligned} f_t + \mu(\psi)f_\psi + \tfrac{1}{2}\sigma^2(\psi)f_{\psi\psi} + h(\psi) &= R(\psi)f, \quad (t, \psi) \in [0, T) \times D, \\ f(T, \psi) &= g(\psi), \quad \psi \in D. \end{aligned}\right\} \tag{10.64}$$

If f is bounded or, more generally, if $\max_{0\leqslant t\leqslant T} f(t, \psi) \leqslant C(1 + \psi^2)$ for $\psi \in D$, then

$$E\left(\exp\left(-\int_t^T R(\Psi_s)\,\mathrm{d}s\right)g(\Psi_T) + \int_t^T h(\Psi_s)\exp\left(-\int_t^s R(\Psi_u)\,\mathrm{d}u\right)\mathrm{d}s \;\Big|\; \mathcal{F}_t\right) = f(t, \Psi_t). \tag{10.65}$$

The Feynman–Kac formula is a standard result of stochastic calculus and it is discussed in many textbooks on stochastic processes and financial mathematics, so we omit the proof (references are given in Notes and Comments).

Affine term structure. We begin with the case $h \equiv 0$; in financial terms this means that we concentrate on survival claims. The following assumption ensures that for $h \equiv 0$ the solution of the PDE (10.64), with terminal condition $g(\psi) = \mathrm{e}^{u\psi}$, $u\psi \leqslant 0$, for $\psi \in D$, is of the form (10.62), so that we have an affine term structure. Note that $g \equiv 1$ for $u = 0$; this is the appropriate terminal condition for pricing zero-coupon bonds.

fact substantial evidence for contagion effects. A good example is provided by the default of the investment bank Lehman Brothers in autumn 2008; the default event combined with the general nervousness caused by the worsening financial crisis sent credit spreads to unprecedentedly high levels.

We can also distinguish between *bottom-up* and *top-down* models of portfolio credit risk. This distinction relates to the quantities that are modelled and cuts through all types of reduced-form portfolio credit risk models including the copula models of Chapter 12.

The fundamental objects that are modelled in a bottom-up model are the default indicator processes of the individual firms in the portfolio under consideration; the dynamics of the portfolio loss are then derived from these. In this approach it is possible to price portfolio products consistently with observed single-name CDS spreads and to derive hedging strategies for portfolio products that use single-name CDSs as hedging instruments. These are obvious advantages of this model class. However, there are also some drawbacks related to tractability: in the bottom-up approach we have to keep track of all default-indicator processes and possibly also background processes driving the model. This typically leads to substantial computational challenges in pricing and model calibration, particularly if the portfolio size is fairly large.

In top-down models, on the other hand, the portfolio loss process is modelled directly, without reference to the individual default indicator processes. This obviously drastically reduces the dimensionality of the resulting models. It can be argued that top-down models are sufficient for the pricing of index derivatives, since the payoff of these contracts depends only on the portfolio loss. However, in this model class the information contained in single-name spreads is neglected for pricing purposes, and it is not possible to study the hedging of portfolio derivatives with single-name CDSs.

There is no obvious and universally valid answer to the question of which model class should be preferred; in Notes and Comments we provide a few references in which this issue is discussed further. In our own analysis we concentrate on bottom-up models.

Notes and Comments

The limitations of static copula models are discussed in a number of research papers; a particularly readable contribution is Shreve (2009). An interesting collection of papers that deal with portfolio credit risk models "after copulas" is found in Lipton and Rennie (2008). Dynamic hedging strategies for portfolio credit derivatives are studied by Frey and Backhaus (2010), Laurent, Cousin and Fermanian (2011) and Cont and Kan (2011), among others; an earlier contribution is Bielecki, Jeanblanc and Rutkowski (2004). A detailed mathematical analysis of hedging errors for equity and currency derivatives is given in El Karoui, Jeanblanc-Picqué and Shreve (1998).

There is a rich literature on models with interacting intensities. Bottom-up models are considered by Davis and Lo (2001), Jarrow and Yu (2001), Yu (2007), Frey and Backhaus (2008) and Herbertsson (2008). Top-down models with interacting intensities include the contributions by Arnsdorf and Halperin (2009) and Cont and Minca (2013). Moreover, there are top-down models where the dynamics of the whole "surface" of CDO tranche spreads—that is the dynamics of CDO spreads for all maturities and attachment points—are modelled directly: see, for example, Ehlers and Schönbucher (2009), Sidenius, Piterbarg and Andersen (2008) and Filipović, Overbeck and Schmidt (2011). The modelling philosophy of these three papers is akin to the well-known HJM models for the term structure of interest rates. A general discussion of the pros and cons of bottom-up and top-down models can be found in Bielecki, Crépey, and Jeanblanc (2010) (see also Giesecke, Goldberg and Ding 2011).

CAP. 17

Credit risk models with explicitly specified interactions between default intensities are conceptually close to network models and to models for interacting particle systems developed in statistical physics. Föllmer (1994) contains an inspiring discussion of the relevance of these ideas to financial modelling; the link to credit risk is explored by Giesecke and Weber (2004, 2006) and Horst (2007). Network models are frequently used for the study of systemic risk in financial networks, an issue that has become highly relevant in the aftermath of the financial crisis of 2007–9. Interesting contributions in this rapidly growing field include the papers by Eisenberg and Noe (2001), Elsinger, Lehar and Summer (2006), Gai and Kapadia (2010), Upper (2011) and Amini, Cont and Minca (2012).

There are some empirical papers on default contagion. Jarrow and Yu (2001) provide anecdotal evidence for counterparty-risk-related contagion in small portfolios. In contrast, Lando and Nielsen (2010) find no strong empirical evidence for default contagion in historical default patterns.

Other work has tested the impact of the default or spread widening of a given firm on the credit spreads or stock returns of other firms: see Collin-Dufresne, Goldstein and Helwege (2010) or Lang and Stulz (1992). The evidence in favour of this type of default contagion is quite strong. For instance, Collin-Dufresne, Goldstein and Helwege (2010) found that, even after controlling for other macroeconomic variables influencing bond returns, the returns of large corporate bond indices in months where one or more large firms experienced a significant widening in credit spreads (above 200 basis points) were significantly lower than the returns of these indices in other months; this is clear evidence supporting contagion.

17.2 Counterparty Credit Risk Management

A substantial proportion of all derivative transactions are carried out OTC, so that counterparty credit risk is a key issue for financial institutions. The management of counterparty risk poses a number of challenges. To begin with, a financial institution needs to measure (in close to real time) its counterparty risk exposure to its various trading partners. Moreover, counterparty risk needs to be taken into account in the pricing of derivative contracts, which leads to the issue of computing credit value adjustments. Finally, financial institutions and major corporations apply various risk-mitigation strategies in order to control and reduce their counterparty risk exposure. In particular, many OTC derivative transactions are now *collateralized*.

Consider a derivative transaction such as a CDS contract between two parties—the protection seller S and the protection buyer B—and suppose that the deal is collateralized. If the value of the CDS is negative for, say, S, then S passes cash or securities (the collateral) to B. If S defaults before the maturity of the underlying CDS and if the value of the CDS at the default time τ_S of S is positive for B, the protection buyer is permitted to liquidate the collateral in order to reduce the loss due to the default of S; excess collateral must be returned to S. Most collateralization agreements are symmetric so that the roles of S and B can be exchanged when the value of the underlying CDS changes its sign.

In this section we study quantitative aspects of counterparty risk management. In Section 17.2.1 we introduce the general form of credit value adjustments for uncollateralized derivative transactions and we discuss various simplifications that are used in practice. In Section 17.2.2 we consider the case of collateralized transactions. For concreteness, we discuss value adjustments and collateralization strategies for a single-name CDS, but our arguments apply to other contracts as well.

17.2.1 *Uncollateralized Value Adjustments for a CDS*

We begin with an analysis of the form of credit value adjustments for an uncollateralized single-name CDS contract on some reference entity R. We work on a filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$, where Q denotes the risk-neutral measure used for pricing derivatives and where the filtration $(\mathcal{G}_t)$ represents the information available to investors. Our notation is as follows: the default times of the protection seller S, the protection buyer B and the reference entity R are denoted by the $(\mathcal{G}_t)$ stopping times τ_S, τ_B and τ_R; δ^R, δ^S and δ^B are the losses given default (LGDs) of the contracting parties; $T_1 = \min\{\tau_R, \tau_S, \tau_B\}$ denotes the first default time; $\xi_1 \in \{R, S, B\}$ gives the identity of the firm that defaults first. We assume that δ^R, δ^S and δ^B are constant; for a discussion of the calibration of these parameters in the context of counterparty credit risk, we refer to Gregory (2012).

The CDS contract referencing R has premium payment dates $t_1 < \cdots < t_N = T$, where t_1 is greater than the current time t and a fixed spread x. The default-free short rate is given by the $(\mathcal{G}_t)$-adapted process (r_t). In discounting future cash flows it will be convenient to use the abbreviation

$$D(s_1, s_2) = \exp\left(-\int_{s_1}^{s_2} r_u \, du\right), \quad 0 \leqslant s_1 \leqslant s_2 \leqslant T.$$

The promised cash flow of a protection buyer position in the CDS between two time points $s_1 < s_2$, discounted back to time s_1, will be denoted by $\Pi(s_1, s_2)$. Ignoring for simplicity accrued premium payments, we therefore have

$$\Pi(s_1, s_2) = \int_{s_1}^{s_2} D(s_1, u)\delta^R \, dY_{R,u} - x \sum_{t_n \in (s_1, s_2]} D(s_1, t_n)(1 - Y_{R,t_n}). \quad (17.2)$$

The first term on the right-hand side of (17.2) represents the discounted default payment, and the second term corresponds to the discounted premium payment.

The value at some stopping time $\tau \geqslant t$ of the promised cash-flow stream for B is then given by

$$V_\tau := E^Q(\Pi(\tau, T) \mid \mathcal{G}_\tau); \quad (17.3)$$

sometimes we will call (V_t) the counterparty-risk-free CDS price. The discounted cash flows that are made or received by B over the period $(s_1, s_2]$ (the real cash flows) are denoted by $\Pi^{\text{real}}(s_1, s_2)$. Note that Π and Π^{real} are in general different as S or B might default before the maturity date T of the transaction.

In order to describe Π^{real} we distinguish the following scenarios.

- If $T_1 > T$ or if $T_1 \leqslant T$ and $\xi_1 = R$, that is if both S and B survive until the maturity date of the CDS, the actual and promised cash-flow streams coincide, so that $\Pi^{\text{real}}(\cdot, T) = \Pi(\cdot, T)$.
- Consider next the scenario where $T_1 < T$ and $\xi_1 = S$, that is the protection seller defaults first and this occurs before the maturity date of the CDS. In that case, prior to T_1, actual and promised cash flows coincide. At T_1, if the counterparty-risk-free CDS price V_{T_1} is positive, B is entitled to charge a *close-out amount* to S in order to settle the contract. Following the literature we assume that this close-out amount is given by V_{T_1}. However, S is typically unable to pay the close-out amount in full, and B receives only a recovery payment of size $(1 - \delta^S)V_{T_1}$. If, on the other hand, V_{T_1} is negative, B has to pay the full amount $|V_{T_1}|$ to S. Using the notation $x^+ = \max\{x, 0\}$ and $x^- = -\min\{x, 0\}$, the actual cash flows on the set $\{T_1 < T\} \cap \{\xi_1 = S\}$ are given by
$$\Pi^{\text{real}}(t, T) = \Pi(t, T_1) + D(t, T_1)((1 - \delta^S)V_{T_1}^+ - V_{T_1}^-). \quad (17.4)$$
- Finally, consider the scenario where $T_1 < T$ and $\xi_1 = B$. If $V_{T_1} > 0$, S has to pay the full amount V_{T_1} to B; if $V_{T_1} < 0$, the protection buyer makes a recovery payment of size $(1 - \delta^B)|V_{T_1}|$ to S. Thus, on the set $\{T_1 < T\} \cap \{\xi_1 = B\}$ we have
$$\Pi^{\text{real}}(t, T) = \Pi(t, T_1) + D(t, T_1)(V_{T_1}^+ - (1 - \delta^B)V_{T_1}^-). \quad (17.5)$$

The correct value of the protection-buyer position in the CDS in the presence of counterparty risk is given by $E^Q(\Pi^{\text{real}}(t, T) \mid \mathcal{G}_t)$. For $t < T_1$ the difference

$$\text{BCVA}_t := E^Q(\Pi(t, T) \mid \mathcal{G}_t) - E^Q(\Pi^{\text{real}}(t, T) \mid \mathcal{G}_t) \quad (17.6)$$

is known as the *bilateral credit value adjustment* (BCVA) at time t. Note that $E^Q(\Pi^{\text{real}}(t, T) \mid \mathcal{G}_t) = V_t - \text{BCVA}_t$: that is, BCVA_t is the adjustment that needs to be made to the counterparty-risk-free CDS price in order to obtain the value of the cash-flow stream Π^{real}. The term *bilateral* refers to the fact that the value adjustment takes the possibility of the default of both contracting parties, B and S, into account. By definition, the bilateral value adjustment is symmetric in the sense that the value adjustment computed from the viewpoint of the protection seller at time

t is given by $-\mathrm{BCVA}_t$; this is obvious since the cash-flow stream received by S is exactly the negative of the cash-flow stream received by B. This contrasts with so-called unilateral value adjustments where each party neglects the possibility of its own default in computing the adjustment to the value of the CDS.

The next proposition gives a more succinct expression for the BCVA.

Proposition 17.1. *For $t < T_1$ we have that* $\mathrm{BCVA}_t = \mathrm{CVA}_t - \mathrm{DVA}_t$, *where*

$$\mathrm{CVA}_t = E^Q(I_{\{T_1 \leqslant T\}} I_{\{\xi_1 = S\}} D(t, T_1) \delta^S V_{T_1}^+ \mid \mathcal{G}_t), \tag{17.7}$$

$$\mathrm{DVA}_t = E^Q(I_{\{T_1 \leqslant T\}} I_{\{\xi_1 = B\}} D(t, T_1) \delta^B V_{T_1}^- \mid \mathcal{G}_t). \tag{17.8}$$

Comments. The CVA in (17.7) reflects the potential loss incurred by B due to a premature default of S; the *debt value adjustment*, or DVA, in (17.8) reflects the potential loss incurred by S due to a premature default of B. A similar formula obviously holds for other products; the only part that needs to be adapted is the definition of the promised cash-flow stream in (17.2).

Accounting rules require that both CVA and DVA have to be taken into account if an instrument is valued via mark-to-market accounting techniques. Note, however, that the use of DVA is somewhat controversial for the following reason: a decrease in the credit quality of B leads to an increase in the probability that B defaults first and hence to a larger DVA term. If both CVA and DVA are taken into account in the valuation of the CDS, this would be reported as a profit for B. It is not clear, however, how B could turn this accounting profit into an actual cash flow for its shareholders.

Proposition 17.1 shows that the problem of computing the BCVA amounts to computing the price of a call option and a put option on (V_t) with strike $K = 0$ and random maturity date T_1. The computation of the value adjustments is therefore more involved than the pricing of the CDS itself, and a dynamic portfolio credit model is needed to compute the value adjustment in a consistent way. The actual computation of value adjustments depends on the structure of the underlying credit model. For further information we refer to Sections 17.3.3 and 17.4.4.

Proof of Proposition 17.1. For $t < T_1 \leqslant s$ it holds that

$$D(t, s) = D(t, T_1) D(T_1, s).$$

Hence, on the set $\{T_1 \leqslant T\}$ we may write $\Pi(t, T) = \Pi(t, T_1) + D(t, T_1)\Pi(T_1, T)$. This yields

$$\Pi(t, T) - \Pi^{\mathrm{real}}(t, T) = I_{\{T_1 \leqslant T\}} D(t, T_1)(\Pi(T_1, T) - I_{\{\xi_1 = S\}}((1 - \delta^S) V_{T_1}^+ - V_{T_1}^-) - I_{\{\xi_1 = B\}}(V_{T_1}^+ - (1 - \delta^B) V_{T_1}^-)).$$

By iterated conditional expectations it follows that

$$E^Q(\Pi(t, T) - \Pi^{\mathrm{real}}(t, T)) = E^Q(E^Q(\Pi(t, T) - \Pi^{\mathrm{real}}(t, T) \mid \mathcal{G}_{T_1})). \tag{17.9}$$

We concentrate on the inner conditional expectation. Since $D(t, T_1)$ and the events $\{T_1 \leqslant T\}$, $\{\xi_1 = S\}$ and $\{\xi_1 = B\}$ are $\mathcal{G}_{T_1}$-measurable, we obtain

$$E^Q(\Pi(t, T) - \Pi^{\mathrm{real}}(t, T) \mid \mathcal{G}_{T_1}) \tag{17.10a}$$

$$= I_{\{T_1 \leqslant T\}} I_{\{\xi_1 = S\}} D(t, T_1) E^Q(\Pi(T_1, T) - ((1 - \delta^S) V_{T_1}^+ - V_{T_1}^-) \mid \mathcal{G}_{T_1}) \tag{17.10b}$$

$$+ I_{\{T_1 \leqslant T\}} I_{\{\xi_1 = B\}} D(t, T_1) E^Q(\Pi(T_1, T) - (V_{T_1}^+ - (1 - \delta^B) V_{T_1}^-) \mid \mathcal{G}_{T_1}). \tag{17.10c}$$

Now, by the definition of (V_t) we have that $E^Q(\Pi(T_1, T) \mid \mathcal{G}_{T_1}) = V_{T_1}$. Moreover, we use the decomposition $V_{T_1} = V_{T_1}^+ - V_{T_1}^-$, where $V_{T_1}^+$ and $V_{T_1}^-$ are $\mathcal{G}_{T_1}$ measurable. Hence (17.10b) equals $I_{\{T_1 \leqslant T\}} I_{\{\xi_1 = S\}} D(t, T_1) \delta^S V_{T_1}^+$ and, similarly, (17.10c) equals $-I_{\{T_1 \leqslant T\}} I_{\{\xi_1 = B\}} \delta^B V_{T_1}^-$. Putting these together, (17.10a) is equal to

$$I_{\{T_1 \leqslant T\}} D(t, T_1)(I_{\{\xi_1 = S\}} \delta^S V_{T_1}^+ - I_{\{\xi_1 = B\}} D(t, T_1) \delta^B V_{T_1}^-),$$

and substituting this into (17.9) gives the result. □

Simplified value adjustments and wrong-way risk. In order to simplify the computation of value adjustments, it is often assumed that $(Y_{t,S})$, $(Y_{t,B})$ and the counterparty-risk free CDS price (V_t) are independent stochastic processes and that the risk-free interest rate is deterministic. We now explain how the value adjustment formulas (17.7) and (17.8) simplify under this independence assumption. For simplicity we consider the case $t = 0$. Denote by $\bar{F}_S(t)$, $\bar{F}_B(t)$, $f_S(t)$ and $f_B(t)$ the survival functions and densities of τ_S and τ_B. Since $\{\xi_1 = S\} = \{\tau_S < \tau_B\} \cap \{\tau_S < \tau_R\}$ and since $V_{T_1} = V_{\tau_S}$ on $\{\xi_1 = S\}$, we obtain

$$\begin{aligned} \mathrm{CVA} = \mathrm{CVA}_0 &= E^Q(I_{\{\tau_S \leqslant T\}} I_{\{\tau_S < \tau_B\}} I_{\{\tau_S < \tau_R\}} D(0, \tau_S) \delta^S V_{\tau_S}^+) \\ &= \delta^S \int_0^T E^Q(I_{\{\tau_S < \tau_B\}} D(0, \tau_S) V_{\tau_S}^+ \mid \tau_S = t) f_S(t)\, dt. \end{aligned}$$

In the last line we have used the fact that δ^S is deterministic and we have used the identity $V_s \equiv 0$ on $\{\tau_R \leqslant s\}$, which allows the indicator $I_{\{\tau_S < \tau_R\}}$ to be dropped. The independence of the processes $(Y_{t,S})$, $(Y_{t,B})$, (V_t) and the fact that interest rates are deterministic imply that

$$\begin{aligned} E^Q(I_{\{\tau_S < \tau_B\}} D(0, \tau_S) V_{\tau_S}^+ \mid \tau_S = t) &= E^Q(I_{\{t < \tau_B\}} D(0, t) V_t^+) \\ &= \bar{F}_B(t) D(0, t) E^Q(V_t^+). \end{aligned}$$

Hence $\mathrm{CVA}^{\mathrm{indep}}$, the credit value adjustment at $t = 0$ under the independence assumption, is given by

$$\mathrm{CVA}^{\mathrm{indep}} = \mathrm{CVA}_0^{\mathrm{indep}} = \delta^S \int_0^T \bar{F}_B(t) D(0, t) E^Q(V_t^+) f_S(t)\, dt, \tag{17.11}$$

and, by a similar argument, the debt value adjustment under independence is

$$\mathrm{DVA}^{\mathrm{indep}} = \mathrm{DVA}_0^{\mathrm{indep}} = \delta^B \int_0^T \bar{F}_S(t) D(0, t) E^Q(V_t^-) f_B(t)\, dt. \tag{17.12}$$

Note that formulas (17.11) and (17.12) are much easier to evaluate than the "correct" expressions (17.7) and (17.8). In particular, we only need to determine the marginal distribution of τ_S and τ_B and the so-called *expected exposures* $E^Q(V_t^+)$ and $E^Q(V_t^-)$. On the other hand, the assumption that (V_t), $(Y_{t,S})$ and $(Y_{t,B})$ are independent is often difficult to justify, and the simplified adjustments can be misleading. Consider, for instance, the case where S, R and B are financial institutions and suppose that $T_1 < T$ and $\xi_1 = S$. In that case it is quite likely that τ_S falls in a time period where financial institutions face adverse conditions so that the credit spread of the reference entity at τ_S and, hence, the market value V_{τ_S} of the CDS referencing R are comparatively high. We therefore expect that

$$E^Q(V_{\tau_S}^+ \mid \tau_S = t) > E^Q(V_t^+),$$

so that CVA > CVA$^{\text{indep}}$. Similarly, we expect that $E^Q(V_{\tau_B}^- \mid \tau_B = t) < E^Q(V_t^-)$, so that DVA < DVA$^{\text{indep}}$. The aggregate effect would be that

$$\text{BCVA} > \text{BCVA}^{\text{indep}}$$

in that case. Some numerical results that support this intuition are given in Section 17.4.4. The phenomenon whereby the conditional expected exposure given the default of the counterparty is higher than the unconditional expected exposure is a typical example of an unfavourable dependence between the size of an exposure and the credit quality of the counterparty. In counterparty risk management, such an unfavourable dependence is known as *wrong-way risk* (since the exposure to a counterparty and the credit quality of that party evolve in the "wrong way"). Wrong-way risk is an important issue in counterparty risk management (see, for example, Chapter 15 of Gregory (2012)).

Unilateral credit value adjustments. In a unilateral value adjustment each party neglects the possibility of its own default. The unilateral value adjustment for the protection buyer B is therefore obtained from the formula for the bilateral value adjustment by assuming that $Q(\xi_1 = B) = 0$. This gives

$$\text{UCVA}_t = E^Q(I_{\{\tau_S<T\}} D(t, \tau_S)\delta^S V_{\tau_S}^+ \mid \mathcal{G}_t).$$

An analogous formula holds for the unilateral value adjustment of the protection seller. Unilateral value adjustments avoid the problem that a worsening credit spread of a financial institution leads to an accounting profit. On the other hand, if B and S use unilateral adjustments, they might come to different conclusions about the value of a given deal.

Netting. A further issue that arises in practice is *netting*. Under a legally enforceable netting agreement the market value of all CDS transactions between B and S at T_1 is computed and only the aggregate value is subject to bankruptcy procedures. In particular, perfectly offsetting transactions cancel each other out. Netting can reduce counterparty risk substantially, so netting agreements are widely used in practice. On the other hand, netting substantially increases the computational complexity of CVA and DVA computations, as we now explain. Suppose that there are N transactions between B and S that fall under a netting agreement and let these be indexed by $n \in \{1, \dots, N\}$. Denote by $(V_{t,n})$ the market value from the point of view of B of the nth transaction. A similar argument to the one used in the proof of Proposition 17.1 implies that

$$\text{CVA}_t = E^Q\biggl(I_{\{T_1<T\}} I_{\{\xi_1=S\}} D(t, T_1)\delta^S \biggl(\sum_{n=1}^{N} V_{T_1,n}\biggr)^+ \biggm| \mathcal{G}_t\biggr),$$

and a similar formula applies to the debt value adjustment. Hence in the presence of netting agreements the computation of value adjustments amounts to the pricing of an option on the sum of the market value of all transactions covered by the netting agreement. In the case of CDS contracts each would typically refer to a different reference entity, so we have to consider $n + 2$ different default times. This is in general a much more difficult problem than pricing the options individually.

17.2.2 *Collateralized Value Adjustments for a CDS*

In this section we introduce popular collateralization strategies and analyse qualitatively the impact of collateralization on credit value adjustments for a CDS. To keep things simple we assume that the collateral is posted in the form of cash and earns the risk-free rate of interest. Many collateralization arrangements used in practice are of this form, but arrangements where other securities are used as collateral can also be found.

Details of the collateralization arrangement for an OTC CDS transaction are fixed in the credit support annex of the transaction. Roughly speaking, the procedure works as follows. At $t_0 = 0$ a collateral account is opened. Let C_t denote the cash balance in the account at time t. Here $C_t > 0$ means that S has posted the collateral and that B is the collateral taker, whereas $C_t < 0$ means that B has posted the collateral and that S is the collateral taker. The collateral position is updated at discrete time points $t_1, \dots, t_N \leqslant T$. At t_1 the collateral taker pays interest on the collateral, and the cash balance C_{t_1} is adjusted in reaction to changes in the price of the underlying CDS over $(t_0, t_1]$. This procedure continues up to the maturity of the CDS or until the first default occurs. If $T_1 > T$ or if $T_1 < T$ and $\xi_1 = R$, the collateral account is closed at the "natural end" of the contract, so $C_t \equiv 0$ for $t \geqslant T_1 \wedge T$. If there is an early default of B or S—that is, if $T_1 \leqslant T$ and $\xi_1 \in \{B, S\}$—the collateral is used to reduce the loss of the collateral taker and any remaining collateral is returned.

An issue arising in this context is *rehypothecation*. The collateral taker typically has unrestricted access to the posted collateral; in particular, the funds can be used as collateral in other OTC derivative transactions. It is therefore possible to have a situation in which the collateral taker defaults and a part of the collateral that should be returned is missing. To keep things simple we ignore this issue and assume that, in the case of a default of the collateral taker, the collateral is always returned in full to the other party. We refer to Notes and Comments for references regarding credit value adjustments in the presence of rehypothecation.

Collateralization strategies. We describe the cash balance in the collateral account by a $(\mathcal{G}_t)$-adapted process (C_t), the so-called collateralization strategy. For convenience we allow for strategies where the collateral account is changed continuously and not just at predetermined rebalancing dates. Recall that the counterparty-risk-free CDS price is denoted by (V_t). In practice, most collateralization arrangements take the form of a threshold collateralization strategy. Formally, a *threshold collateralization strategy* with *thresholds* M_1, $M_2 \geqslant 0$, labelled $(C_t^{M_1,M_2})$ for $0 \leqslant t \leqslant T_1 \wedge T$, is given by

$$C_t^{\gamma,M_1,M_2} = (V_t^+ - M_1) I_{\{V_t^+ > M_1\}} - (V_t^- - M_2) I_{\{V_t^- > M_2\}}. \tag{17.13}$$

Under this strategy collateral is posted if V_t^+ (the exposure of B) exceeds the threshold M_1 or if V_t^- (the exposure of S) exceeds the threshold M_2. A threshold strategy is used if B and S want to protect themselves against severe counterparty-risk-related losses, while accepting the possibility of smaller losses in order to simplify the practical management of the collateralization process. For $M_1 = M_2 = 0$ we obtain the special case of *market-value collateralization* with

$$C_t^{\text{market}} = V_t, \quad 0 \leqslant t \leqslant T_1 \wedge T. \tag{17.14}$$

Collateralized value adjustment. Value adjustments for collateralized CDS contracts are largely analogous to the uncollateralized case, so we keep our presentation short. The bilateral collateralized value adjustment BCCVA is the difference between the collateralized credit value adjustment CCVA and the collateralized debt value adjustment CDVA. As before, the CCVA gives the value of the potential loss for B due to an early default of S, whereas the CDVA gives the value of the potential loss for S due to an early default of B.

In order to describe these potential losses we have to consider the payments at an early default. Note that no additional collateral is posted at or after the default of B or S. The amount of collateral available for the settlement of the contract is therefore given by C_{T_1-} (the amount of collateral that has been posted immediately prior to T_1). This distinction matters if the close-out amount (V_t) jumps at T_1, for instance due to contagion effects, or if there is some delay between the last adjustment of the collateral account and the settlement of the positions. We begin with the scenario where the protection seller defaults first. We have to distinguish the cases $V_{T_1} > 0$ and $V_{T_1} < 0$.

- Suppose that $V_{T_1} \geqslant 0$ and that the protection buyer is the collateral taker, that is $C_{T_1-} \geqslant 0$. In that case the collateral is used to reduce the loss of the protection buyer and excess collateral is returned. If C_{T_1-} is smaller than V_{T_1}, the protection buyer incurs a counterparty-risk-related loss of size $\delta^S(V_{T_1} - C_{T_1-})$; if $C_{T_1-} \geqslant V_{T_1}$, the amount of collateral is sufficient to protect B from losses due to counterparty risk. If S is the collateral taker, i.e. if $C_{T_1-} \leqslant 0$, there is no available collateral to protect B and he suffers a loss of size $\delta^S V_{T_1}$.

- Suppose that $V_{T_1} \leqslant 0$. In that case B has no exposure to S, so he does not suffer a loss related to counterparty risk (the fact that he has incurred a loss due to the decrease in the counterparty-risk-free CDS price is irrelevant for the computation of value adjustments for counterparty risk).

Summarizing, the counterparty-risk-related loss of B is given by

$$I_{\{T_1<T\}} I_{\{\xi_1=S\}} \delta^S (V_{T_1}^+ - C_{T_1-}^+)^+.$$

Similarly, S suffers a loss in the scenario where $V_{T_1} \leqslant 0$ and where there is insufficient collateral to settle the contract in full, that is, for $V_{T_1}^- > C_{T_1}^-$. The counterparty-risk-related loss of S is thus given by

$$I_{\{T_1<T\}} I_{\{\xi_1=B\}} \delta^B (V_{T_1}^- - C_{T_1-}^-)^+.$$

Thus BCCVA_t, the *bilateral collateralized credit value adjustment* at time t, is given by

$$\text{BCCVA}_t = \text{CCVA}_t - \text{CDVA}_t, \tag{17.15}$$

where

$$\text{CCVA}_t := E(I_{\{T_1<T\}} I_{\{\xi_1=S\}} D(t,T_1) \delta^S (V_{T_1}^+ - C_{T_1-}^+)^+ \mid \mathcal{G}_t),$$
$$\text{CDVA}_t := E(I_{\{T_1<T\}} I_{\{\xi_1=B\}} D(t,T_1) \delta^B (V_{T_1}^- - C_{T_1-}^-)^+ \mid \mathcal{G}_t).$$

Without collateralization, i.e. for $C_t \equiv 0$, formula (17.15) reduces to the simpler result of Proposition 17.1.

Performance of market-value collateralization. The sum $\text{CCVA}_t + \text{CDVA}_t$ gives the value in t of the entire counterparty-risk-related loss, and it can therefore be viewed as a measure of the performance of a given collateralization strategy. Here we make the following immediate observation: suppose that market-value collateralization with $C_t^{\text{market}} = V_t$ is used and that the market value of the CDS does not jump at T_1, that is, $V_{T_1} = V_{T_1-}$ almost surely. In that case the formulas for CCVA_t and CDVA_t in (17.15) show that

$$\text{CCVA}_t = \text{CDVA}_t = 0, \quad t \leqslant T_1,$$

so that market-value collateralization works perfectly. If, on the other hand, $|\Delta V_{T_1}| = |V_{T_1} - V_{T_1-}|$ is comparatively large, the performance of market-value collateralization will be not so good. Some numerical results supporting this observation will be presented in Section 17.4.4.

Notes and Comments

The literature on counterparty risk management is growing rapidly, leading to a proliferation of valuation-adjustment acronyms (CVA, DVA, FVA and others). A detailed introduction can be found in the textbooks by Gregory (2012), Cesari et al. (2009) and Brigo, Morini and Pallavicini (2013). A non-technical introduction to the computation of value adjustments for counterparty risk is given in the papers by Hull and White (2012, 2013).

The derivation of the bilateral credit value adjustments in Propositions 17.1 and (17.15) is based on the papers by Brigo and Chourdakis (2009) and Brigo, Capponi and Pallavicini (2014) (see also Frey and Rösler 2014). The last two papers also consider the case of rehypothecation and discuss the actual computation of value adjustments in various portfolio credit risk models. Credit value adjustments in structural credit risk models are studied in Lipton and Sepp (2009). A very general technical analysis of value adjustments is given in Crépey (2012a,b).

17.3 Conditionally Independent Default Times

In this section we discuss models with conditionally independent default times. We begin with a discussion of general mathematical properties; applications and specific examples from the literature are considered in Sections 17.3.2 and 17.3.3.

17.3.1 *Definition and Mathematical Properties*

Throughout we consider a portfolio of m obligors with default times τ_i and default indicators $Y_{t,i} = I_{\{\tau_i \leqslant t\}}$, $1 \leqslant i \leqslant m$, on a probability space $(\Omega, \mathcal{F}, P)$. The ordered default times are denoted by $0 = T_0 < T_1 < \cdots < T_m$, and $\xi_n \in \{1, \dots, m\}$ gives the identity of the firm defaulting at time T_n. We introduce the filtrations $(\mathcal{H}_t^i)$, $1 \leqslant i \leqslant m$, and $(\mathcal{H}_t)$ defined by

$$\mathcal{H}_t^i = \sigma(\{Y_{s,i} : s \leqslant t\}) \quad \text{and} \quad \mathcal{H}_t = \mathcal{H}_t^i \vee \cdots \vee \mathcal{H}_t^m. \tag{17.16}$$

$(\mathcal{H}_t^i)$ is the filtration generated by the default observation for obligor i alone; $(\mathcal{H}_t)$ is the filtration generated by default observations for all obligors. Often $(\mathcal{H}_t)$ is called the *default history* of the portfolio or the *internal filtration* generated by the default times $\tau_1, \dots, \tau_m$. The definition of conditionally independent default times is a straightforward multivariate extension of the notion of doubly stochastic default times from Section 10.5.1. In particular, the distribution of the default times is affected by additional information on top of the default history $(\mathcal{H}_t)$. Formally, we represent this information by a filtration $(\mathcal{F}_t)$ on the underlying probability space. Typically, $(\mathcal{F}_t)$ is generated by some observable background process. The information available to investors is given by the filtration $(\mathcal{G}_t) = (\mathcal{F}_t) \vee (\mathcal{H}_t)$ (see also (10.46)).

Definition 17.2. The default times $\tau_1, \dots, \tau_m$ are conditionally independent doubly stochastic random times if there are positive, $(\mathcal{F}_t)$-adapted processes $(\gamma_{t,i})$, $1 \leqslant i \leqslant m$, with $\Gamma_{t,i} = \int_0^t \gamma_{s,i}\,\mathrm{d}s$ strictly increasing and finite for every $t > 0$, such that

$$P(\tau_1 > t_1, \dots, \tau_m > t_m \mid \mathcal{F}_\infty) = \prod_{i=1}^m \exp\biggl(-\int_0^{t_i} \gamma_{s,i}\,\mathrm{d}s\biggr). \tag{17.17}$$

Note that the definition implies that each of the τ_i is a doubly stochastic random time with conditional hazard process $(\gamma_{t,i})$ in the sense of Definition 10.10 and that the rvs $\tau_1, \dots, \tau_m$ are conditionally independent given $\mathcal{F}_\infty$.

Construction and simulation via thresholds. Lemma 17.3 extends Lemma 10.11.

Lemma 17.3. *Let $(\gamma_{t,1}), \dots, (\gamma_{t,m})$ be positive, $(\mathcal{F}_t)$-adapted processes such that $\Gamma_{t,i} := \int_0^t \gamma_{s,i}\,\mathrm{d}s$ is strictly increasing and finite for any $t > 0$. Let $X = (X_1, \dots, X_m)'$ be a vector of independent, standard exponentially distributed rvs independent of $\mathcal{F}_\infty$. Define*

$$\tau_i = \Gamma_i^{\leftarrow}(X_i) = \inf\{t \geqslant 0 : \Gamma_{t,i} \geqslant X_i\}.$$

Then $\tau_1, \dots, \tau_m$ are conditionally independent doubly stochastic random times with hazard processes $(\gamma_{t,i})$, $1 \leqslant i \leqslant m$.

Proof. By the definition of τ_i we have $\tau_i > t \iff X_i > \Gamma_{t,i}$. The rvs $\Gamma_{t_i,i}$ are now measurable with respect to $\mathcal{F}_\infty$, whereas the X_i are mutually independent, independent of $\mathcal{F}_\infty$ and standard exponentially distributed. We therefore infer that

$$\begin{aligned} P(\tau_1 > t_1, \dots, \tau_m > t_m \mid \mathcal{F}_\infty) &= P(X_1 > \Gamma_{t_1,1}, \dots, X_m > \Gamma_{t_m,m} \mid \mathcal{F}_\infty) \\ &= \prod_{i=1}^m P(X_i > \Gamma_{t_i,i} \mid \mathcal{F}_\infty) \\ &= \prod_{i=1}^m \mathrm{e}^{-\Gamma_{t_i,i}}, \end{aligned} \tag{17.18}$$

which shows that the τ_i satisfy the conditions of Definition 17.2. □

Lemma 17.3 is the basis for the following simulation algorithm.

Algorithm 17.4 (multivariate threshold simulation).

(1) Generate trajectories for the hazard processes $(\gamma_{t,i})$ for $i = 1, \dots, m$. The same techniques as in the univariate case can be used here.

(2) Generate a vector X of independent standard exponentially distributed rvs (the threshold vector) and set $\tau_i = \Gamma_i^{\leftarrow}(X_i)$, $1 \leqslant i \leqslant m$.

As in the univariate case (see Lemma 10.12), Lemma 17.3 has a converse.

Lemma 17.5. *Let $\tau_1, \dots, \tau_m$ be conditionally independent doubly stochastic random times with $(\mathcal{F}_t)$-conditional hazard processes $(\gamma_{t,i})$. Define a random vector X by setting $X_i = \Gamma_i(\tau_i)$, $1 \leqslant i \leqslant m$. Then X is a vector of independent, standard exponentially distributed rvs that is independent of $\mathcal{F}_\infty$, and $\tau_i = \Gamma_i^{\leftarrow}(X_i)$ almost surely.*

Proof. For $t_1, \dots, t_m \geqslant 0$ the conditional independence of the τ_i implies that

$$P(\Gamma_1(\tau_1) \leqslant t_1, \dots, \Gamma_m(\tau_m) \leqslant t_m \mid \mathcal{F}_\infty) = \prod_{i=1}^m P(\Gamma_i(\tau_i) \leqslant t_i \mid \mathcal{F}_\infty).$$

Moreover, similar reasoning to the univariate case implies that

$$P(\Gamma_i(\tau_i) \leqslant t_i \mid \mathcal{F}_\infty) = P(\tau_i \leqslant \Gamma_i^{\leftarrow}(t_i) \mid \mathcal{F}_\infty) = 1 - \mathrm{e}^{-t_i},$$

which proves that X has the claimed properties. □