Credit Risk Models: Theory, Applications and Implementation

Master's Thesis in Financial Economics

Ulf Nore

Credit Risk Models:
Theory, Applications and Implementation

By:
Ulf Nore

Advisor:
Steinar Ekern

NHH, Spring 2010

Contents

List of Figures

Introduction

Background

Recalling the credit crisis as one of the events defining of the last decade, it is no wonder credit risk modeling has become one of the central research areas in modern finance. Leading up to and even more after the aforementioned credit crisis, there has been much debate on the need for regulating credit derivatives, an asset class by many viewed as important for understanding the background for the crisis. A natural extension of this debate is that of the value of the models used for valuing an risk managing such instruments, in particular with respect to the quality of credit ratings. The same models have also seen applications in banking capital regulation, another area where previously held beliefs have been challenged by these events.

Overview

The focus of this thesis is the two main classes of credit risk models that appear in the academic literature and are used by practitioners in financial institutions and credit rating agencies. There is no "industry standard" pricing model for credit derivatives or risk management, in the manner of the Black-Scholes Model for stock options. I will therefore cover qualitatively some of the variation in the field. Because of the limited scope of this thesis, the focus of this presentation is on the basic principles and methods, which are presented in a detailed and more formal way. I will also outline how the basic models can be extended.

The first two chapters provide an introduction to these two model frameworks, known as reduced form and structural models, respectively. Structural models build more or less directly on option pricing theory, and make specific assumptions on the causal relationship between structural variables such as asset values, debt level, interest rate on the one hand and credit events on the other, viewing a credit event mainly as an endogenous event - an event that is explained *inside* the models by other variables. Reduced form models, on the other hand, see defaults as exogenous. No causal relationships are assumed, we are only trying to obtain a probabilistic model based on available market data and certain assumptions about the data generating processes. A brief discussion on some simple techniques for model calibration is also included.

Chapter 4 illustrates some of the theoretical concepts developed in the preceding chapters by applications to derivatives pricing. Finally, the appendices contain a brief overview over some concepts in valuation theory and stochastic modeling that are used throughout the thesis, as well as the simulation methods used in model implementation.

Methods

In the field of credit risk research, there are numerous articles and books containing analytical results for highly sophisticated models. While recognizing the practical usefulness of such contributions, I believe there are certain important advantages to focusing on a numerical approach.

The constraints related to computational costs that used to be the main problem with numerical techniques have become less important due to the exponential growth in computing power. Secondly, it can often be a simpler modelling task to implement a numerical approach than to search for analytical solutions for many complex problems, and it is often sufficient with a select set of numerical methods for tackling many problems. Analytical approaches on the other hand, often require considerable mathematical ingenuity and sophistication that may be beyond many practitioners. Furthermore, a simple numerical model can often easily be extend to more complex cases without modifying core parts of the program.

Acknowledgements

I wish to thank my advisor, professor Steinar Ekern, not only for his guidance and advice that has been invaluable for my work with this thesis, but also for his teaching in financial theory and derivatives pricing at NHH that stirred my interest in the fields of mathematical and theoretical finance, hereunder the methods and problems I discuss in this thesis.

Chapter 1

Credit Risk - Empirical Data and Some Notes on Modeling

> "Credit default swaps (CDSs) have proved to be one of the most successful financial innovations of the 1990s."
>
> Hull and White (2003)

1.1 Background

Financial activities create wealth whenever they lead to a more productive allocation of capital and risk between the agents in the economy. For many agents, financial institutions in particular, the handling of *credit risk* – i.e. the risk of a borrower being, totally or partially, unable to repay a loan – is an issue of utmost importance. Until quite recently, managing credit risk has been difficult due to the low liquidity of debt securities, so that agents have been unable to reduce their exposure to such risk, either by selling debt instruments or taking offsetting positions in other instruments. While traditional debt instruments, such as corporate bonds, obviously are *credit derivatives*, they also have an embedded interest risk element, which make them less ideal for trading and transferring credit risk.

In the last two decades, the way financial institutions handle credit risk has been altered in a fundamental way by the introduction of modern credit derivatives, the most important being the *credit default swap* or CDS. The CDS is a simple instrument that for a periodic payment guarantees protection against the credit risk of a reference entity, usually in terms of some predefined cash settlement between the issuer and the buyer of

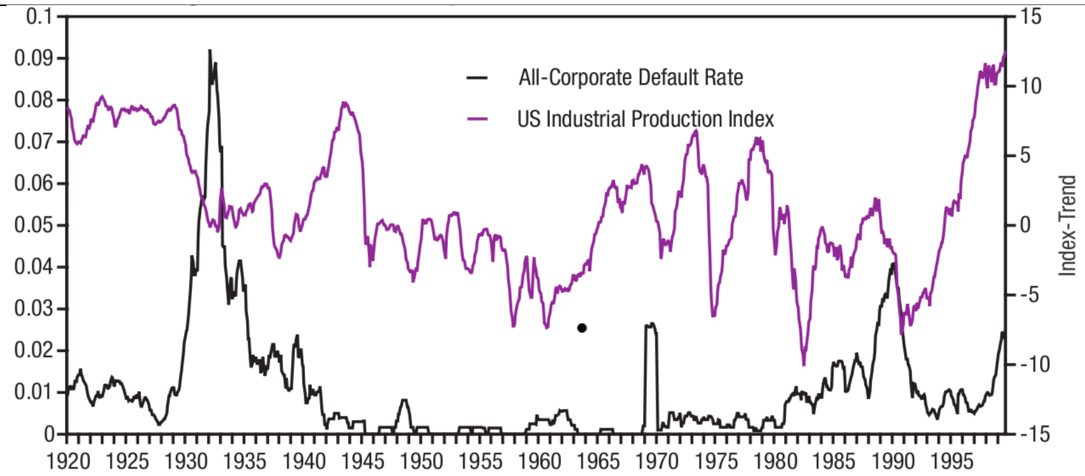


Figure 1.1: Historical default rates. Source: Moody's (2000).

credit protection in event of default, wholly or partially covering the loss caused by the credit event. An institution having a large credit exposure to some particular entity can therefore use a CDS to neutralize this position. Furthermore, it is of course unnecessary to actually hold the underlying bonds in order to obtain a certain risk profile; trading in CDS's alone is sufficient, as these instruments can be issued independently of whether or not the bonds are actually issued.

1.2 Data Sources and Some Empirical Facts About Credit Risk

From Figure 1.1, where the historical over-all US corporate default rates are plotted as a time series together with the US Industrial Production Index (a measure of economic growth), we get a few impressions of some properties of default rates. Though fairly weak ($-.14$), there is a correlation between the IP index and default rates. Strong economic growth tends to go hand in hand with low default rates, though there has been a variying pattern with resepect to whether a weakening of the economy preceeds or follows an increase in default rates.

Another concept of key interest in credit risk modeling in addition to default rates is *default severity*, often referred to as loss given default, usually a percentage of outstanding principal. Moody's (2000) have compiled similar data for this quantity, and it exhibits similar time series properties. On average, recovery rates are low near the bottom of business cycle contractions and high after periods of strong economic growth.

The cyclical nature of credit risk that is apparent from Figure 1.1 is also reminiscent of the problem of default correlation or *clustering*, the fact that one default tends to be followed by others. We can explain such causality by considering the dependence between firms in a supply chain; if a major buyer shuts down production, the suppliers are also more likely to default. We can also think of how similar firms depend on the same macroeconomic factors such as fuel prices, and aggregate demand, etc., and particular risk factors such as trends or hypes.

1.2.1 Ratings Data and the Estimation of Default Probabilities

As defaults are infrequent low-probability events, empirical data on default probabilities and interdependences are hard to compile. Of course, for a firm that has not defaulted, we cannot directly *observe* its default probability as this is an event that only occurs once. Hence, we need to come up with some estimates of these probabilities based on data available for similar firms, or imply them from market prices using some pricing model.

	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79
Default	0	0	0	0	0	0	0	100.00

Figure 1.2: One year transition matrix of Standard and Poor's credit ratings for the period 1981-1996. Source: CreditMetrics.

One common method for estimating such probabilities is using data published by rating agencies such as Standard & Poor. Consider Figure 1.2 where a one year *transition matrix* of credit ratings is given. Entry a_{ij} in the table gives the probability of a firm going from rating i to rating j over the course of one year. There are several things to note about such data. We see that the *default* state is absorbing; once a firm has defaulted, it will never live again, and the probability of transition from default to any other rating is consequently zero. Furthermore, the transition probabilities are *physical* probabilities. This should be quite obvious as they are estimated from actual historical data. They will therefore generally differ from the risk neutral default probabilities that can be implied from market prices. This method is discussed in Sections 2.6.1 and 2.6.2. Appendix A explains the distinction between risk neutral and physical probabilities.

Using this matrix it is simple to compute the n-year probability matrix. If $\mathbf{T}_1$ denotes the one year transition matrix, then the two year transition matrix is given by $\mathbf{T}_2 = \mathbf{T}_1 \cdot \mathbf{T}_1$. To see that this holds consider the probability of starting in state AAA, and being in state AAA after two years. This is the probability of staying in AAA two years in a row plus the probability of going from AAA to AA the first year and back to AAA the second, and so forth:

$$p^2(AAA|AAA) = p^2_{AAA,AAA} + p_{AAA,AA}p_{AA,AAA} + p_{AAA,A}p_{A,AAA} + ... + p_{AAA,CCC}p_{CCC,AAA}$$

In the same manner, we can find the n-year transition probability matrix as $\mathbf{T}_n = \mathbf{T}_1^n$. Considering only the rightmost column of the matrices $\{\mathbf{T}_1, \mathbf{T}_2, ..., \mathbf{T}_n,\}$ we have estimates of the physical default probabilities for a firm of a given rating, for any time horizon. The cumulative density function following from this method is plotted in Figure 1.3.

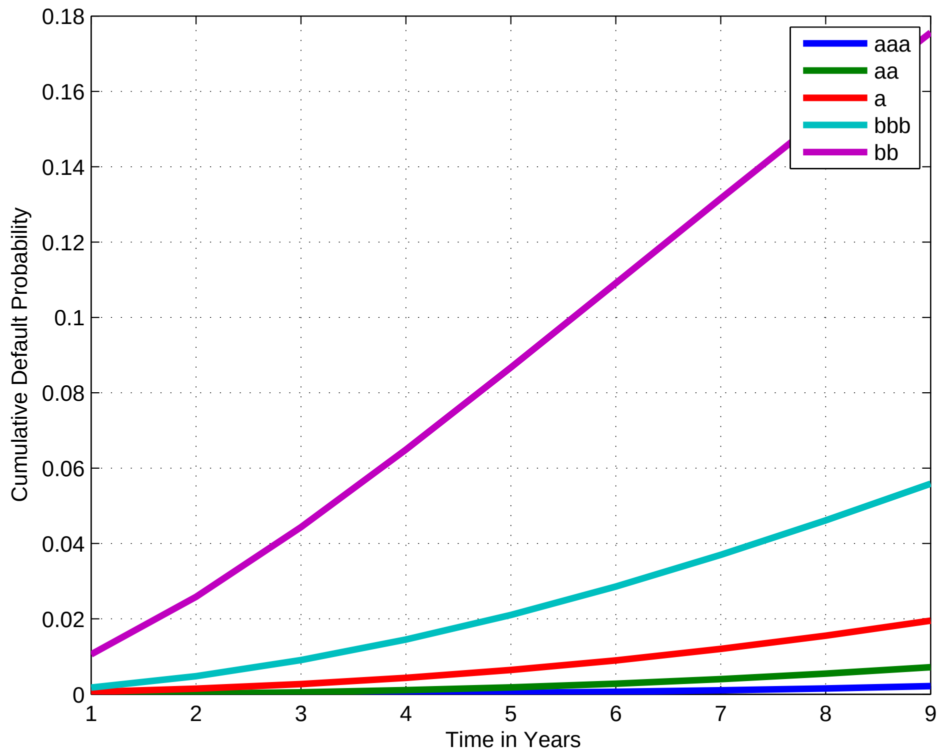


Figure 1.3: Physical cumulative default probabilities for some ratings classes from Figure 1.2.

Problems with Ratings Data

There are several reasons why probabilities implied from market data using models is preferable to ratings data for the pricing applications:

- Rating agencies react slower than the market in anticipation of future credit quality. The most striking example is the recent credit crisis where the sub-prime mortgage backed securities defaulted with a triple-A status.
- Firm specific information contained in market prices is ignored; the default probabilities inferred from ratings data are averages over a potentially very heterogeneous group of firms that are likely exposed to very different risk factors.
- The probabilities are physical, and can therefore not be used directly as input to the valuation models as they usually are stated.

1.2.2 Credit Derivatives Markets

As we have seen, there are good arguments for that ratings data may not be the best data source for estimating default probabilities. Often, a better alternative is to use market data. There are three important markets from which we can infer credit risk information using the modeling tools discussed later. These are the equity, bond and credit derivatives markets. This thesis explores some methods for implying credit risk information from the securities traded in these markets.

Obviously, the quality of such information depends crucially on the liquidity and the transparency of the financial markets. If market participants are uninformed with respect to the assets that are traded, the market prices do not reflect actual values or probabilities and is therefore worthless. Likewise, if markets are illiquid, market prices may not reflect actual asset values. The latter is often a problem with using bond prices which is why credit derivative prices are often preferred in estimating default probabilities.

Furthermore, credit default swap rates are usually quoted for a larger number of maturities than bonds which means a finer credit curve[1]. This approach is illustrated in Chapter 2.

1.3 Modeling Credit Risk

The two classes of models presented here can be seen as representing two different "traditions". Structural models are straightforward extensions of classical option pricing theory, and was indeed one of the first applications of this theory outside contingent claims valuation (see for instance Merton (1974) and Black and Cox (1976)). They rely explicitly on a theory on the causal relationship between asset prices[2] and bankruptcy.

[1] The *credit curve* is common term describing the term structure of default probabilities.

[2] Or, in more advanced cases such as Goldstein et al. (2001), the relationship between cash flows, interest rates etc. and bankruptcy.

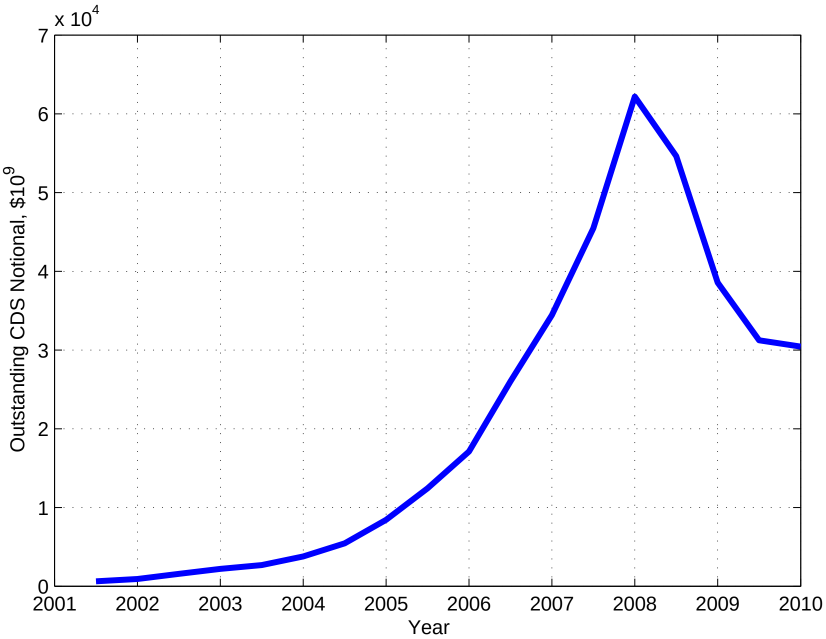


Figure 1.4: Outstanding CDS notional. Source: ISDA (http://www.isda.org/statistics/).

From a theoretical point of view, structural models are for many reasons the "preferred" framework , as they not only provide a causal relationship between the structural variables of the firm and the default probabilities, but also a coherent framework for valuing *any* claim on the firm's assets. Furthermore, they can be extended in many directions incorporating, among other things, endogenous capital structure changes, so that there is an interdependence between asset values and capital structure decisions. Such models, which in the literature is referred to as *dynamic capital structure models*, do not appear often practice in credit derivatives valuation as they are harder to calibrate than the simpler *static capital structure* models considered here. Therefore, it is assumed throughout the discussion on structural models in this thesis that the capital structure irrelevance assumption[3] holds.

As pointed out by Vasicek (1984), modern structural credit risk models are purely quantitative, and is therefore radically different from "traditional" methods for asset pricing and credit valuation that relies on the analyst's knowledge of a firm's operations to project future cash flows under various scenarios. However, the data used in the traditional method, both about the firm and the markets in which it operates is presumably public information. Assuming a certain degree of market efficiency, this information will already be reflected in the prices of the firm's assets as reflected by debt and equity values.

Reduced form models represent an approach based on reliability theory that is similar

[3]The asset value is independent of the financial structure of an entity.

to modeling in insurance and operations management. They do not model causal relationships between structural variables, rather use default probabilities as inferred from market prices. A default is considered "similar" to the occurrence of an event triggering a payment from an insurance company, or in the operations management case, the breakdown of a particular machine that is part of a production process. There may be several causes behind such a breakdown; it may be due to human failing or have some technical cause. In the model however, these are seen as random events occurring according to some process. From a modeling perspective, we are interested in determining this process than the actual causality.

What separates the credit risk setting from the operations management setting is the role of interdependence. Default time interdependence is a major risk factor that must be accounted for in credit portfolio valuation and risk assessment. While in a production process, simultaneous breakdowns may be preferable so that a total maintenance can be performed, a large number of defaults occurring over a short period of time is clearly problematic for a financial institution with a limited cash flow and capital reserve. Another important problem is that in many practical problems the credit portfolio may contain a large number of assets, so that in order to "scale down" the problem in such a way that we can make qualitative sense of the data, some reduction of dimensionality is necessary. This topic is central throughout this thesis and as we will see, many different methods are proposed in the literature. One standard method is to assume correlation arises through the individual assets' dependence on a set of systemic risk factors.

1.4 Evaluating Models

This thesis presents the two fundamental classes of credit risk models as well as some of the several extensions of these models that have been proposed. From a practical point of view, it is necessary to have some criteria by which these models are evaluated depending on their application.

Based on the nature of defaults suggested by empirical studies such as Moody's (2000), we can specify requirements a model should be able to reproduce with respect to key quantities like default rate correlations and default probabilities. As demonstrated in the CDO example in 4.4, multi-name credit derivative values are extremely sensitive to default correlations[4]. Furthermore, the analyst implementing the model is faced with several important constraints such as:

- *Scarcity of data.* Data on defaults is limited in many respects. One may not have sufficiently long time series available or there may be changes in the data generating

[4]As discussed in Hull (2007), this is quite clear from the cash flow mechanics of these instruments.

processes[5] so that older observations are no longer valid. Hence, a model with few parameters to estimate is tractable due to the uncertainty in the estimates.

- Time constraints in implementing, testing and calibrating the models. A simple numerical model is often simpler to verify against an analytic base case.

The last point shows that there is an important trade-off between the richness of the model and the time spent on implementing and maintaining it. The focus here is therefore the basic cases of the models that are treated thoroughly in a quantitative manner and implemented numerically. Extending these is usually a quite straightforward issue of adding more "bells and whistles" to the fundamentals.

[5] Such shifts may be caused for instance be caused by regulatory changes.

Chapter 2

Reduced Form Credit Risk Models

This chapter provides an introduction to the theory behind one of the two standard classes of credit risk models often referred to as *reduced form credit risk models*. According to Hull et al. (2006), this class of models is largely the industry standard in credit derivative modeling, primarily because they are easy to fit to observed market prices.

This chapter also pays some attention to different methods for correlation modeling that are also used later on for structural models. Particular attention is paid to the so-called *copula approach* that provides a technically efficient method for implementing the multivariate distribution of a set of assets given the marginal distributions and estimates of correlations.

2.1 Single Credit Framework

Consider a single defaultable security and let τ denote its survival time as measured from $t = 0$. On the filtered probability space[1] $(\mathbb{P}, \mathcal{F}, \Omega)$. Here $\mathbb{P}$ denotes the risk neutral probability measure. τ is a stopping time (a random variable) with respect to the filtration $\mathcal{F}_t$ that represents the accumulated market information available at time t.

We are now interested in a framework in which probabilistic statements about τ can be made. Therefore let $F(T) = \mathbb{P}(\tau \leq T)$ be the cumulative distribution function (cdf) of the default time, ie. the probability that the time of default occurs before a particular time T. An equivalent statement is the *survival function* $S(T) = 1 - F(T)$ which is the probability that a security does *not* default prior time T. Closely related to $F(T)$ is the probability density function (pdf) $f(t) = \frac{dF(t)}{dt}$ that can be interpreted as the default probability on an infinitesimally small time interval around some point in time t.

[1] See Appendix A for some background and references on this terminology.

2.1.1 A Binomial Model of Credit Risk

As an illustration, consider a bond with face value 100, and 6% annual coupon rate paid annually until maturity in year 3. Let $q = \lambda \Delta t = .08$ be the conditional *risk neutral* default probability on an interval of length $\Delta t = 1$, i.e. the probability of default occurring during the interval, conditioned on survival up to the start of the interval. Further assume that if default occurs, the value recovered is constant $R = 40$ paid at the end of the year as illustrated in Figure 2.1.

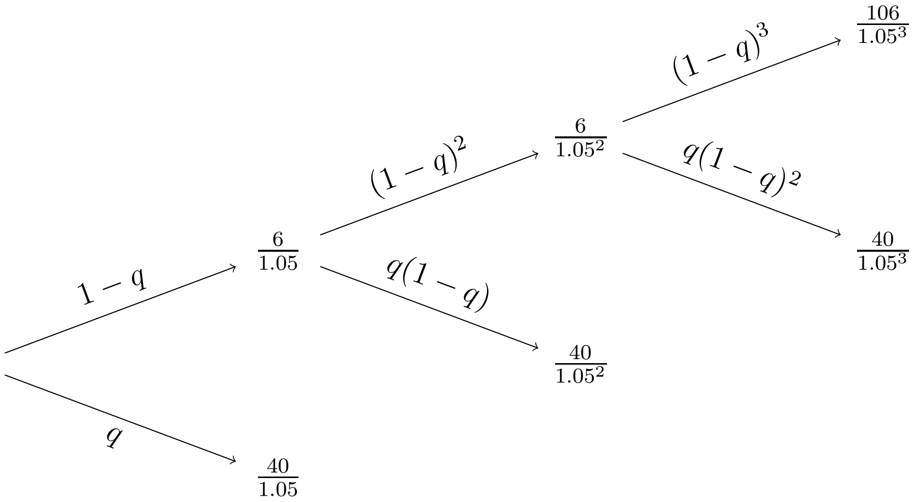


Figure 2.1: Binomial tree illustrating discrete time default process.

As discussed in Appendix A, we can value this risky bond by discounting the expected cash flows by the risk free interest rate, here assumed to be 5% with *discrete* compounding. The example is summarized in Figure 2.2. The probabilities in row 2 and 4 are the cumulative survival and annual default probabilities, respectively. To arrive at the results here, the conditional probability of default in a particular year is the probability of surviving up to that year times the probability of defaulting in that particular year.

Year	**1**	**2**	**3**
Cash flow, survival	6	6	106
Cumulative Probability	0.9200	0.8464	0.7787
Cash flow, default	40	40	40
Annual Probability	0.0800	0.0736	0.0677
Expected Cash Flow	8.7200	8.0224	85.2494
Discounted Cash Flow	8.3048	7.2766	73.6416
Expected NPV			88.2230

Figure 2.2: Pricing in the binomial model.

In comparison, the present value of a risk free bond with the same cash flow structure is $6 \cdot 1.05^{-1} + 6 \cdot 1.05^{-2} + 106 \cdot 1.05^{-3} = 102.7232$, so the risk premium on the risky bond is 13.5003. It is assumed throughout this thesis that credit risk is the *only* risk factor. In reality, such a price difference is usually explained in terms of other, additional risk factors, liquidity risk being the most important.

Mixed Probability Binomial Models

In many valuation problems, the binomial model is an excellent tool; its primary advantage being its technical simplicity and intuitive nature. It is the among the simplest derivative pricing models to understand, explain and implement numerically, yet powerful enough to to replicate the results from simulation models in many cases given a sufficiently small step size.

The key problem with this model as it is formulated above is that it does not account for dependence between default times which is, as mentioned in the introduction, one of the most important risk factors that any credit risk model must handle well if it is to be applied to portfolio modeling. One common extension of the binomial model is to randomize the default probability q to mimic dependence between the binomial trees representing the various firms in the portfolio.

While such binomial models are used in practice, the next sections, take a different approach to modeling correlation that uses a continuous time framework.

2.1.2 The Hazard Rate Function

A key quantity of interest[2] is the instantaneous default probability conditional on survival up to a certain point in time t. This probability is often referred to as the hazard rate function. It is defined as the limit of the probability of survival on an interval $(t, t + \Delta t)$, given $\tau > t$, as Δt approaches zero:

Definition 2.1.1. ***Hazard Rate Function***

Let $F(t)$ be the cumulative distribution function of the default time t and $f(t)$ its derivative, then the hazard rate function $\lambda(\tau)$ is defined as:

$$\lambda(\tau) = \lim_{\Delta t \to 0} \mathbb{P}[t < \tau < t + \Delta t | \tau > t] = \frac{f(\tau)}{1 - F(\tau)} = \frac{f(\tau)}{S(\tau)} \tag{2.1.1}$$

The last equality can be seen by writing out the probabilities as integrals and applying the fundamental theorem of calculus to the numerator and recognizing the denominator as $1 - F(t)$.

[2] This is because it specifies the default generating process in this model framework.

Note that we are yet to specify the functional form of F, f and λ as we have so far only dealt with them abstractly. In the the example in Section 2.1.1, λ is assumed constant and $F(t)$ is on the form:

$$F(n\Delta t) = \lambda\Delta t + (1-\lambda\Delta t)\lambda\Delta t + (1-\lambda\Delta t)^2\lambda\Delta t + ... + (1-\lambda\Delta t)^n\lambda\Delta t$$

Here $\lambda\Delta t$ is the probability of defaulting on an interval of length Δt. In the next section we consider a model where λ acts as the parameter in a continuous default time distribution.

2.1.3 The Poisson/Cox Process

As initially noted, we want to provide some model of defaults as the occurrence of a discrete and rare event without, as in the structural models considering the underlying economic processes driving these events. A simple example of a process satisfying these requirements is the *Poisson* process $N(t)$ which is a continuous time, discrete space counting process. We want to define the default of asset i as the first jump of the process $N_i(t)$. The interdependence between the firms in the portfolio is given by the correlation structure of a set of Poisson processes.

Walpole et al. (2007) defines the Poisson process in terms of three key properties:

Definition 2.1.2. ***Poisson Process***

Let I be the indicator function associated with the stopping time τ. The Poisson process is a function $F:\Omega\rightarrow\mathbb{N}^+$ mapping the sample space to the set of positive integers such that:

$$N(t) = \sum_{i=1}^{n} I_{\tau_i \le t} \tag{2.1.2}$$

satisfying the following properties:

1. The Markov property *or "*memorylessness*": the number of events occurring on a time interval $[t_0, t_1]$ is independent of the number of events occurring on any other disjoint time interval $[T_1, T_2]$.*[3]

2. *The probability of an event occurring on a particular time interval is proportional to the length of the interval.*

3. *The probability of more than one event occurring an an infinitesimal time interval is negligible.*

[3] In particular, any event occurring on a time interval starting at t is independent of $\mathcal{F}_t$ (here: the set of information revealed to the market (historical default data)).

Some Properties of the Poisson Distribution

Two important consequences of this definition are:

- The probability distribution of N(t) is the Poisson distribution, that is, the probability of exactly k events occurring on a time interval of length τ is then given by the probability mass function of the Poisson distribution:

$$F(T,k) = \mathbb{P}[N(t+T) - N(t) = k] = \frac{e^{-\lambda T}(\lambda T)^k}{k!} \tag{2.1.3}$$

- In particular we see that the probability that no defaults occur on a given time interval is given by:

$$F(T,0) = \mathbb{P}[N(t+T) - N(t) = 0] = e^{-\lambda T} \tag{2.1.4}$$

  That is, the probability distribution of the waiting time until the first occurrence is an exponential distribution with parameter λ[4].

The last point above is important as we interpret the time τ_1 of first jump as the time of default. The time to default (or survival time) is therefore exponentially distributed with a mean $\frac{1}{\lambda}$ and variance $\frac{1}{\lambda^2}$. Note that we could also start with the assumption that time to default is exponentially distributed, and then arrive at the above definition of the Poisson process.

We can show the latter by considering a discrete setting where $\lambda(t)h$ denotes the probability of surviving on an interval $[t, t+h]$ conditional on no previous default. The cumulative probability of surviving up to time t is $p_s(t)$. It follows that:

$$p_s(t+h) - p_s(t) = -\lambda(t)p_s(t)h$$

Taking the limit as $h \to 0$:

$$\frac{dV}{dt} = -\lambda(t)p_s(t)$$

which has the solution:

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

We say that $N(t)$ is a counting or "jump" process. We interpret the time τ of the occurrence of the first "jump" of the process $N(\tau)$ as default.

[4]For notational simplicity, λ is assumed constant here.

The Poisson process is entirely specified by a single parameter λ, the hazard rate, often referred to as the process' *intensity*, which is as the name indicates, a measure of the frequency of events occurring. The Poisson process, or as it is sometimes called, the (time) homogeneous Poisson process is a particular case of the more general Cox process, where $\lambda(t) = \lambda$ is a constant. Later on, $\lambda(t)$ is defined in terms of a stochastic differential equation so as to allow for random variations in default intensities.

The default of a single credit is in this framework given as the first jump of the Poisson process which is the *first passage time to* $N(t) = 1$, τ defined similarly to a default in the Black-Cox model:

$$\tau = \inf\{t \in \mathbb{R}^+ | N(t) = 1\} \tag{2.1.5}$$

The Credit Curve

The notion of a *term structure of default intensities* or, more colloquially, *credit curve* is occurring frequently in the literature on credit risk. Similarly to the yield curve in interest rate modeling, expressing the yield on a short interval $[t, t + dt]$, the credit curve is the instantaneous default probability or hazard rate on a short interval. The credit curve does of course contain precisely the same information as the survival or default time distributions.

The Cox Process

The above Poisson model can be generalized to allowing for a time varying and even stochastic default intensity. This type of process is referred to as a Cox process or a non-homogeneous Poisson process. For instance, we could allow $\lambda = \lambda(t)$ to be given by the following stochastic differential equation (SDE):

$$d\lambda(t, \lambda(t)) = \mu(t, \lambda(t))dt + \sigma(t, \lambda(t))dW(t) \tag{2.1.6}$$

where $W(t)$ is the standard univariate Wiener process defined in Appendix A. We can think of the process driving this as the "state of the economy", where $\lambda(t)$ will be inversely related to state variables such as GDP growth, credit spreads and so forth. One approach to mimic the cyclicality apparent in actual default data is to use a mean-reverting SDE, such as the Ornstein-Uhlenbeck process defined in Appendix A.

From the instantaneous default probability it is a simple matter to derive an expression for the probability of a security surviving on a time interval $[t, T]$ conditional on no prior default as the "sum" of all the instantaneous default probabilities:

$$p_s(t,T) = \mathbb{P}[\tau > T | \tau > t] = \mathbb{E}\left[\exp\left(-\int_t^T \lambda(s)ds\right) \middle| \mathcal{F}(t)\right] \tag{2.1.7}$$

The probability of default occurring on the same interval is denoted $p_d(t,T)$:

$$p_d(t,T) = 1 - p_s(t,T) \tag{2.1.8}$$

These integrals are not necessarily simple or even possible to evaluate analytically. This depends on the functional form of λ. However, simple numerical methods often do a good job approximating them.

In the homogeneous case (constant λ), the survival probability can be simplified:

$$p_s(t,T) = e^{-\lambda(T-t)} \tag{2.1.9}$$

and likewise the cumulative default probability:

$$p_d(t,T) = 1 - e^{-\lambda(T-t)} \tag{2.1.10}$$

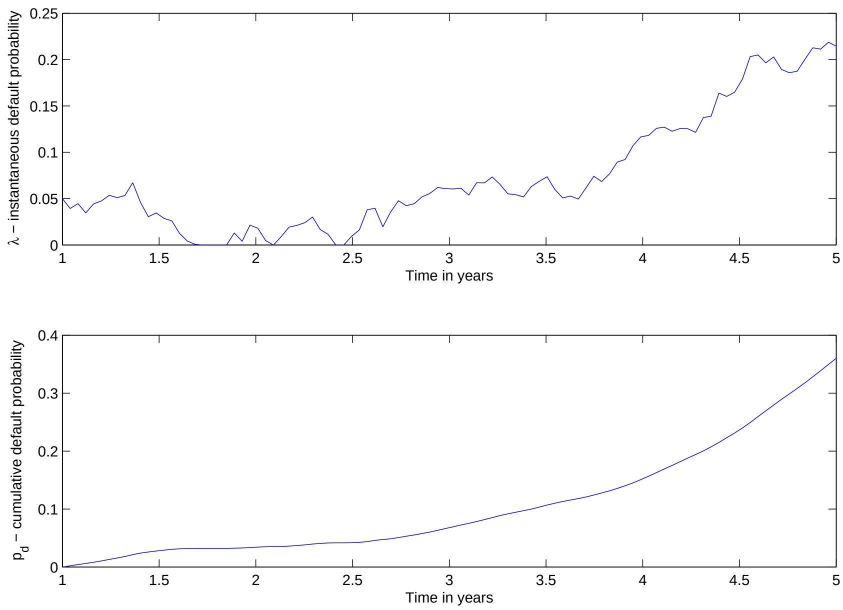


Figure 2.3: Hazard rate as Ornstein-Uhlenbeck process and corresponding default time cdf.

Figure 2.3 illustrates the relationship between hazard rates and the cumulative default probability. Here the hazard rate function is given by the stochastic differential equation[5] $d\lambda_t = \alpha(\lambda_0 - \lambda_t)dt + \sigma dW_t$. As before, the survival probability is $p_s(0,t) = \exp[-\int_0^t \lambda(s)ds]$. Conversely, the cumulative default probability $p_d(0,t) = 1 - p_s(0,t)$.

The top figure shows a particular trajectory for the mean-reverting default intensity process. To compute the integral behind the second figure, the midpoint method for numerical integration[6] is used. Note how the cdf below is flat in the times where the default intensity is low and steep later on when λ is high. For a simulation model, it is necessary to simulate a large number of trajectories for λ.

Summary

To conclude the discussion here, we restate some key points that are central to the simulation algorithms later on. With a constant hazard rate λ, time to default is characterized by an exponential distribution. The properties of this distribution is summarized below.

- Cumulative probability distribution of defaulting prior to t: $F(t) = 1 - e^{-\lambda t}$.
- Corresponding probability density function $f(t) = \lambda e^{-\lambda t}$.
- Mean survival time: $1/\lambda$ and variance: $1/\lambda^2$.

2.2 Cash Flow Pricing in a Reduced Form Model

From the above framework it is possible to work out formulas pricing risky cash flows using its default probability and an interest rate model. Consider first the simple case of finding the time t value of a defaultable zero coupon bond $G(t,T)$ paying a unit cash flow at time T contingent on survival and nothing otherwise[7].

Letting $P(t,T)$ denote the *risk-free* discount function we have the following which is a direct application of the risk neutral pricing framework described earlier[8] for a defaultable zero coupon bond:

$$G(t,T) = \mathbb{E}[P(t,T)|\mathcal{F}(t)] = P(t,T)p_s(t,T) = e^{-\int_t^T (r(s)+\lambda(s))ds} \tag{2.2.1}$$

When both the hazard and interest rates are stochastic processes, there is a resemblance between the above pricing equation and the bond pricing expressions found in

[5] There is a very important problem to note about using this particular process as a model for default intensities; namely that it is *not* strictly non-negative, clearly at odds with the definition of the hazard rate as a probability.

[6] See Cheney and Kincaid (2007).

[7] This assumption will be relaxed later on. In the most general case the fraction lost to bankruptcy cost $\alpha(t)$ is specified as a stochastic process.

[8] Under the standard assumptions of arbitrage free markets, the same results hold for almost any process for asset values.

multi-factor interest rate models[9]. In the case of constant default intensity and interest rates we get a very simple pricing equation:

$$G(t,T) = e^{-(r+\lambda)(T-t)}$$

From these equations, it is reasonable to interpret λ as a *risk premium*. Using these equations, any other defaultable security can be priced similarly to the above zero coupon bond.

2.2.1 Recovery Rates

The above example is clearly stylized as it assumes that recovery rates are zero; either there is a unit cash flow at time T or there is no cash flow. This is of course unrealistic, and as in the structural models of Chapter 3, we can introduce a recovery value proportionate to the face value of the bond.

This approach is known as recovery of face value (RFV), and is perhaps the simplest possible approach, in particular when the fraction recovered is constant. More advanced models may apply recovery of market value or model the fraction recovered as a stochastic process. Hull (2006) discusses a number of different models of recovery rates with references to the literature.

Let α denote the fraction recovered, τ the stopping time indicating default, the value of a defaultable zero coupon bond with unit face value is now given as:

$$G(t,T) = \mathbb{E}[P(t,T) + \alpha P(t,\tau)|\mathcal{F}(t)] \tag{2.2.2}$$

While a closed form expression can be derived for the above expectation, I will only consider an intuitive numerical method of evaluating the integrals using a midpoint approximation and compute the expectations by Monte Carlo simulation.

2.3 Default Correlation and Model Implementation

Now that a reduced model of default probability and single entity or asset pricing has been established, the key problem still remains, namely specifying dependence or association structure between default times. While the primary question of interest is the correlations between default times, it is important to stress that it is not the only. In more advanced models we are also interested in the relationship between variables such

[9]Even though there is a well-established theory on multi-factor interest rate models, working out an analytic expression in the most general case with correlated rates is non-trivial.

as default, recovery, interest rates, etc. In this thesis, the main concern is default time dependence.

Before we can start implementing a model, an appropriate measure of interdependence must be chosen. Whereas this is a relatively simple matter in terms of structural models, where it is one usually can settle with the correlation $< dA_1, dA_2 >$ between two Itô processes (see Shreve (2004) for rigorous definition), there are several approaches to modeling asset price interdependence in reduced form models. As discussed in Li (2000) and Elizalde (2005a), one could choose the standard Pearson correlation coefficient given, in the bivariate case, as:

$$\rho_{XY} = \frac{cov[X,Y]}{\sigma_X \sigma_Y}$$

Translating this into our framework of defaultable securities, we can let $1_A(t)$ and $1_B(t)$ denote two indicator random variables taking on the value one if entity A or B, respectively, have defaulted by time t. Letting $p_A(t)$ be the probability that A defaults prior to time t:

$$var(1_i) = p_A(t)(1 - p_A(t))$$

and:

$$cov[1_i, 1_j] = p_{ij} - p_i p_j$$

we get the following:

$$\rho_{XY} = \frac{p_{AB} - p_A p_B}{\sqrt{p_A p_B (1 - p_A)(1 - p_B)}} \tag{2.3.1}$$

For a particular class of multivariate distributions, known as *elliptical* distributions, which includes the important Gaussian distribution, the correlation coefficient (or more generally, the correlation matrix) fully determines the dependence structure. However, it can be problematic due to its linearity which means that we can have a fully deterministic relationship between two variables yet zero correlation. A simple illustration is if $X \sim \Phi(0,1)$ and Y is an even function of X, for instance, $Y = X^2$. Obviously, this is problematic, as we want a zero correlation coefficient to signify that there is no association between the variables. This is a key problem that is discussed later in the section on copulas. Suffice it to say for now that the Pearson correlation measure remains important in this analysis, in particular as an input to the copula models.

2.3.1 Simulating Defaults – The Inversion Method

We have now covered sufficient detail to develop a simple simulation algorithm when we know the functional form and parameters of $\lambda(t)$ as well as the correlation matrix $\mathbf{\Sigma}$.

Let X be a random variable and F be some associated cumulative distribution function function (cdf). Since F is a non-decreasing function, it has an inverse F^{-1}:

$$F^{-1}(q) = \inf\{x : F_X(x) \geq q\} \tag{2.3.2}$$

From the definition of the cdf and the properties of the uniform distribution, the following important relationship that is central to the simulation algorithms applied to reduced form models follows. Let $\mathcal{U}$ be a uniform random variable on the interval $[0, 1]$. Then we have the following relationship:

$$\begin{aligned}
\mathbb{P}[X \leq x] &= \mathbb{P}[F^{-1}(\mathcal{U}) \leq x] && (2.3.3)\\
&= \mathbb{P}[F(F^{-1}(\mathcal{U})) \leq F(x)] && (2.3.4)\\
&= \mathbb{P}[\mathcal{U} \leq F(x)] && (2.3.5)\\
&= F(x) && (2.3.6)
\end{aligned}$$

The first equality uses the fact that $X = F^{-1}(\mathcal{U})$. To see this, consider the particular case of default times. Now the domain of F is $\mathbb{R}^+$ and its range is $[0, 1]$ (by the definition of a probability). The inverse F^{-1}, therefore, must transform elements in $[0, 1]$ onto $\mathbb{R}^+$ according to the cdf. The last equality above follows from the property of the uniform distribution on $[0, 1]$, that $P(\mathcal{U} < u) = u$.

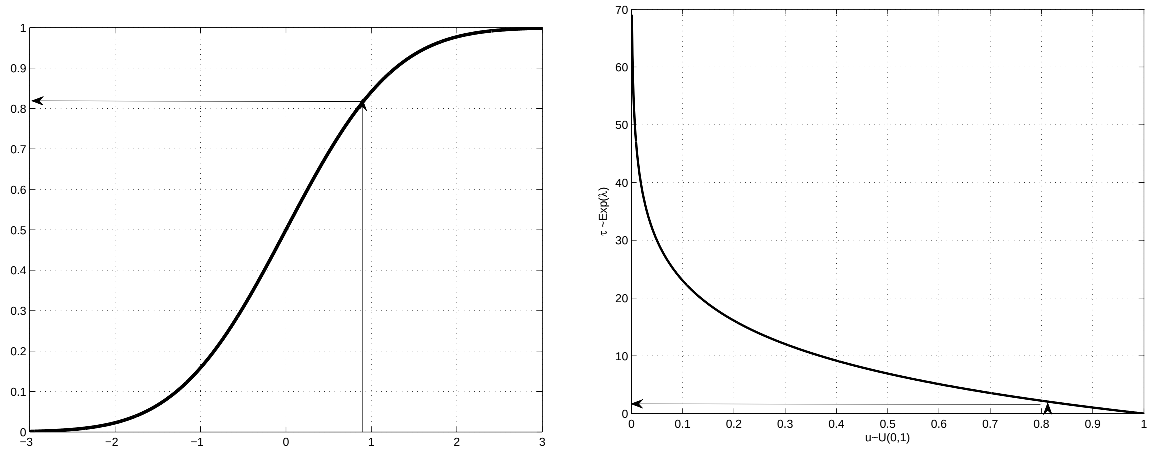
So we have that X and $F_X^{-1}(\mathcal{U})$ have the same cdf. Thus random variables with any given cdf can be simulated by drawing uniform random variables and applying the inverse cdf. This algorithm is known as the *inversion method*[10].

For example: we can generate two correlated uniform random vectors $[\mathcal{U}_1, \mathcal{U}_2]$. Assuming asset 1 has a t_5 distributed returns while asset 2 is normally distributed, we set $X_1 = t_5^{-1}(\mathcal{U}_1)$ and $X_2 = \Phi^{-1}(\mathcal{U}_1)$. Using this we let the above cdf $F(t) = e^{-\lambda t}$ be the survival function, ie. probability of no default prior to time x. The inverse of this function is:

$$T = -\frac{\ln(p_s)}{\lambda}$$

Since p_s is a probability we can generate default times by simulating a set of $[0, 1]$ uniform random variates $\{u_1, u_2, ..., u_n\}$ and transforming them by the formula: $T_i = -\frac{\ln(u_i)}{\lambda}$. This method is discussed further in Section 4.1.

[10] As an aside, the inversion method can be very useful when simulating a portfolio of assets where the individual assets have different (marginal) probability distributions. For example, if we assume two assets A and B have normally and t_5 distributed returns, we can generate two uniform random vectors $\{u_1, u_2\}$ and let the return vectors be $R_A = \Phi^{-1}(u_1)$ and $R_B = t_5^{-1}(u_2)$.



(a) Transforming $\Phi(0,1)$ random variates to uniform (0,1) variates.

(b) Transforming uniform variates to exponential variates with parameter λ.

Figure 2.4: The inversion method.

2.4 Conditionally Independent Defaults

We now turn to the first technique for dealing with correlation modeling. The core idea behind conditionally independent defaults - CID-models, is that defaults are independent conditioned on the realization of a set of systemic factors that determine the hazard rate. Such factors may be GDP, the short interest rate[11], credit spreads, etc. To illustrate the technique we let $\lambda(t)$ be a stochastic process. Firm i is assumed to default at time τ given by:

$$\tau = \inf\left\{t : \int_0^t \lambda(t)dt \geq E_i\right\} \tag{2.4.1}$$

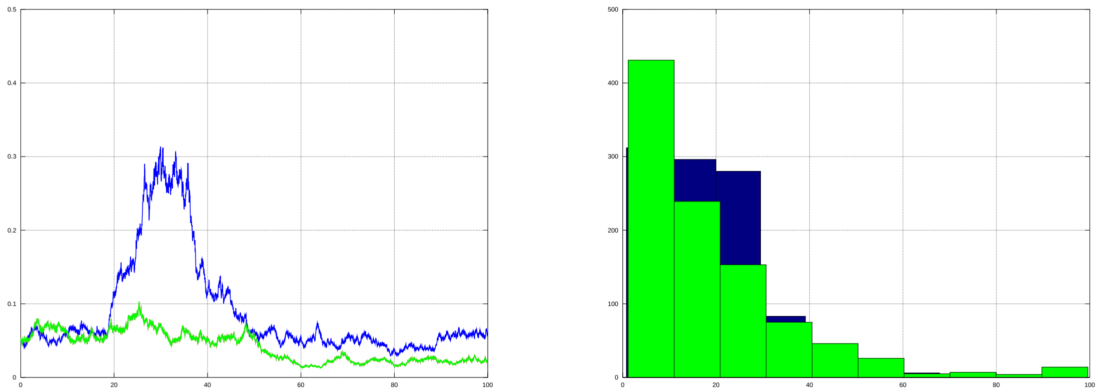
Where E_i is an unitary exponentially distributed random variable ($E_i \sim e^{Z_{0,1}}$), and E_i and E_j are independent for $i \neq j$.

Illustration

Most authors, such as Duffee (1999) use rather complicated models to determine λ relying on multi-factor techniques from term structure modeling. To illustrate, we consider a simplified model, where the hazard rate is a zero drift geometric Brownian motion with constant volatility:

$$d\lambda(t, \lambda(t)) = \lambda(t)\sigma dW(t)$$

[11] Duffee (1999) proposes a model on the form $\lambda_i(t) = \lambda_i^*(t) + \alpha s_1(t) + \beta s_2(t)$ where the s_i are factors inferred from a two-factor model of the short rate.



(a) Two sample paths of hazard rate process.

(b) Default time simulation histograms.

Figure 2.5: CID Simulation

The default criterion for firm i is still as given in Equation 2.4.1. We then have that:

$$\mathbb{P}[\tau_i > t] = exp\left(-\int_0^t \lambda(t)dt\right) \tag{2.4.2}$$

so that $1_{t>\tau_i}$ is a Cox, or doubly stochastic Poisson process. The simulation algorithm is summarized below:

1. Generate one path of λ and approximate the integral in Equation 2.4.1.
2. Generate N exponential random variates and determine the time of default according to Equation 2.4.1.
3. Repeat step 1 and 2 above.

Two sample pats and default time histograms are plotted in Figure 2.5.

2.5 Copula Functions

2.5.1 Definition and Some Central Properties

A popular method for correlation modeling in the reduced form framework is the copula method, a method that uses a transformation of a set of marginal distributions to create a joint distribution. This section will present the fundamentals of copula theory and some particular copula functions illustrating the basic concept as well as the breadth of models available. The next section shows how it can be applied to pricing problems using simulation in a reduced form model.

Several good references on copula theory and its applications in financial modeling are available, hereunder Nelsen (1999) and Li (2000). A comprehensive article on the measuring and modeling of correlated risks is Wang (1998). Elizalde 2005a contains a comprehensive list of references to further articles on this field. Finally, many software packages and financial algorithms libraries such as MATLAB and QuantLib contain routines for copula models that are comprehensively documented.

We start by a definition:

Definition 2.5.1. *Copula*

A n-dimensional copula *is defined as the joint cumulative density function* $C : [0,1]^n \to [0,1]$ *of a uniformly distributed random vector* $\mathbf{U} \in \mathbb{R}^n$*:*

$$C(u_1, u_2, ..., u_n, \Sigma) = \mathbb{P}\{\mathcal{U}_1 \leq u_1, ..., \mathcal{U}_N \leq u_n\} \tag{2.5.1}$$

A copula is therefore a multivariate distribution function with uniformly distributed marginals. An important result in the theory of copulas states that the marginal distributions and the dependence between the set of variables can be separated. Firstly, we can use copulas to link a set of marginal distributions to a joint distribution:

$$\begin{aligned} C(F_1(x_1), ..F_n(x_n) &= \mathbb{P}[\mathcal{U}_1 \leq F_1(x_1), ..., \mathcal{U}_n \leq F_n(x_n)] && (2.5.2) \\ &= \mathbb{P}[F_1^{-1}(\mathcal{U}_1) \leq x_1, ..., F_n^{-1}(\mathcal{U}_n) \leq x_n] && (2.5.3) \\ &= \mathbb{P}[X_1 \leq x_1, ..., X_n \leq x_n] && (2.5.4) \\ &= F(x_1, ..., x_n, \mathbf{\Sigma}) && (2.5.5) \end{aligned}$$

For instance, in the bivariate case with X and Y random variables with marginal cdfs F_X and F_Y: $C(x, 1) = \mathbb{P}[\mathcal{U} \leq x, \mathcal{U} \leq 1] = x$.

The following theorem, first proven by Sklar, shows the the converse also holds; any multivariate distribution function can, under certain technical assumptions be written as a copula.

Theorem 2.5.2. (Sklar) *Let G be an n-dimensional distribution function with continuous marginals $F_1, ..., F_n$. Then there exists an n-dimensional copula C such that:*

$$G(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) \tag{2.5.6}$$

If we consider two bivariate uniform random variables on $[0, 1]$, X and Y, with the copula function $C(x, y, \rho) = \mathbb{P}(X < x, Y < y|\rho)$, we observe that:

- $C(x,1,\rho) = \mathbb{P}(X < x, Y < 1|\rho) = \mathbb{P}(X < x) = x$, ie. we can obtain the of a variable X by evaluating the copula when all other parameters are 1.
- If X and Y are independent, then $C(x,y,\rho) = \mathbb{P}(X < x)\mathbb{P}(Y < y) = xy$.
- With perfect correlation, $C(x,y,\rho) = \mathbb{P}(X < x)\mathbb{P}(Y < y) = \min(x,y)$

Why Use Copula Models?

While the theory of copulas may perhaps seem unnecessarily complex at first sight, the key point to the above discussion about what a copula actually *does*, namely creating a multivariate joint distribution that is consistent with the specified marginal distributions of the systemic and idiosyncratic factors. While we have certain "simple" multivariate distributions that can be used to generate multivariate data such as a default times of a portfolio, this set is limited. Furthermore, most simple methods impose restrictions that are important in practice, the most important being that the marginals must have the same univariate distribution. For example, the multivariate Gaussian distribution has univariate Gaussian marginals.

For instance, consider the correlation structure that will be used much later on in the discussion on structural models. Let X_i be the random variable that determines the time of default for firm i. It is a function of a systemic risk factor Y and an idiosyncratic risk factor ϵ_i where Y and ϵ_i are independent:

$$X_i = \rho_i Y_i + \sqrt{1-\rho_i^2}\,\epsilon_i$$

Now, the choice of marginal distribution for Y and ϵ_i will determine the copula *uniquely*. If for instance both Y and ϵ_i are standard normally distributed, a Gaussian copula will result. For any other choice of distributions, a different copula is the result.

To summarize, what is tractable about the copula approach is that it provides simple method to specify a multivariate joint distribution for *any* set of marginal distributions.

2.5.2 Some Classes of Copula Functions

For the purpose of this thesis we consider three copula functions that appear frequently in the financial literature in general, and particularly in that on reduced form models - normal, t- and mixed normal copulas. These are under no circumstances the only ones available, but they are comparatively simple to estimate and implement with standard software. Furthermore, the basic properties of these distributions are well known from fundamental probability theory. For further discussion on copula models see for instance Li (2000) and Elizalde (2005a) and sources cited therein.

Definition 2.5.3. ***Normal Copula***

Let Φ^N denote the N-dimensional normal cumulative distribution function, the N-dimensional normal or Gaussian copula C^N is given by:

$$C^N(u_1, u_2, ..., u_N) = \Phi^N(\Phi^{-1}(u_1), \Phi^{-1}(u_2), ..., \Phi^{-1}(u_N), \Sigma) \tag{2.5.7}$$

As a particular example we note the bivariate normal copula given by:

$$C^2(u_1, u_2) = \Phi^2(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \rho)$$

In a similar fashion to that above, we can define the N dimensional Student t copula with v degrees of freedom.

Definition 2.5.4. ***Student t Copula***

Let t_v^N denote the student t cumulative distribution function with v degrees of freedom. Then the N dimensional t-copula C^t is defined by:

$$C^t(u_1, u_2, ..., u_N) = t_v^N(t_v^{-1}(u_1), t_v^{-1}(u_2), ..., t_v^{-1}(u_N)) \tag{2.5.8}$$

Typically, for financial applications, v is chosen to a low number such as 5 or 3 producing a fat tailed distribution (higher risk of extreme losses and gains). As the number of degrees of freedom gets very high, the distribution converges to a normal distribution.

Finally, we consider two copulas that are somewhat different from the two previous. The first approach follows from the two last properties of copulas at the end of Section 2.5.1, that $C(x, y, 1) = \min(x, y)$ and $C(x, y, 0) = xy$. Consider next a weighted combination of these two functions ρ be the weight assigned to the first. We consider the bivariate normal case:

Definition 2.5.5. ***Mixed Bivariate Copula***

Let (x, y) be a set of random variables that are independent. A copula is then given by $C_1 = xy$. Let (v, w) be two perfectly correlated random variables. Another copula is then given by $C_2 = \min(x, y)$. If $0 < \rho \leq 1$,

$$C(u, v) = (1 - \rho)uv + \rho \min(u, v) = (1 - \rho)C_1 + \rho C_2 \tag{2.5.9}$$

defines a mixed bivariate copula.

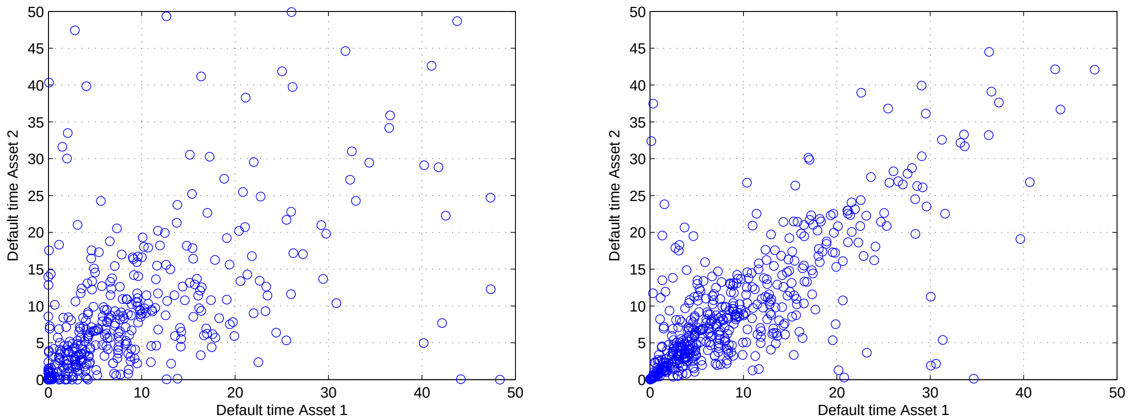
Finally, as an illustration of the breadth of copula functions available as alternatives to the more common normal and t-copulas, we consider a type of copula that is not determined by the standard correlation coefficient.

Definition 2.5.6. ***Clayton Copula***

Let u and v be uniform random variables on $[0, 1]$ and $0 < \theta < \infty$ be a constant. The function $C(u, v)$ defines a bivariate Clayton copula if:

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}} \tag{2.5.10}$$

The parameter θ is here a parameter determining the dependence between the two variables, where $\theta = 0$ means independent marginals. Contrary to the copulas above, the Clayton copula does not allow for negative correlation. However, as Trivedi and Zimmer (2005) states, it exhibits strong left tail dependence which makes it an appropriate model for credit risk. This type of dependence is important in credit risk modeling; once one firm defaults, it has consequences for other firms it is doing business with.



(a) Normal copula. (b) t_1-copula.

Figure 2.6: Default times simulation with two different copula functions.

Figure 2.6 illustrate default times generated using the inversion method from a bivariate normal copula versus default times from a t_1-copula. It is apparent that the normal copula yield much more scattered default times than the t-copula that exhibits more of a default clustering.

Figures 2.7-2.8 are plots of the random variates from bivariate copulas themselves. Notice the difference between the Gaussian and the t-copula; while the first tend to scatter the observations more, the t-copula gives a "clearer" pattern. For a correlation coefficient of .8, the band formed in the t-copula example is much slimmer than in the Gaussian case.

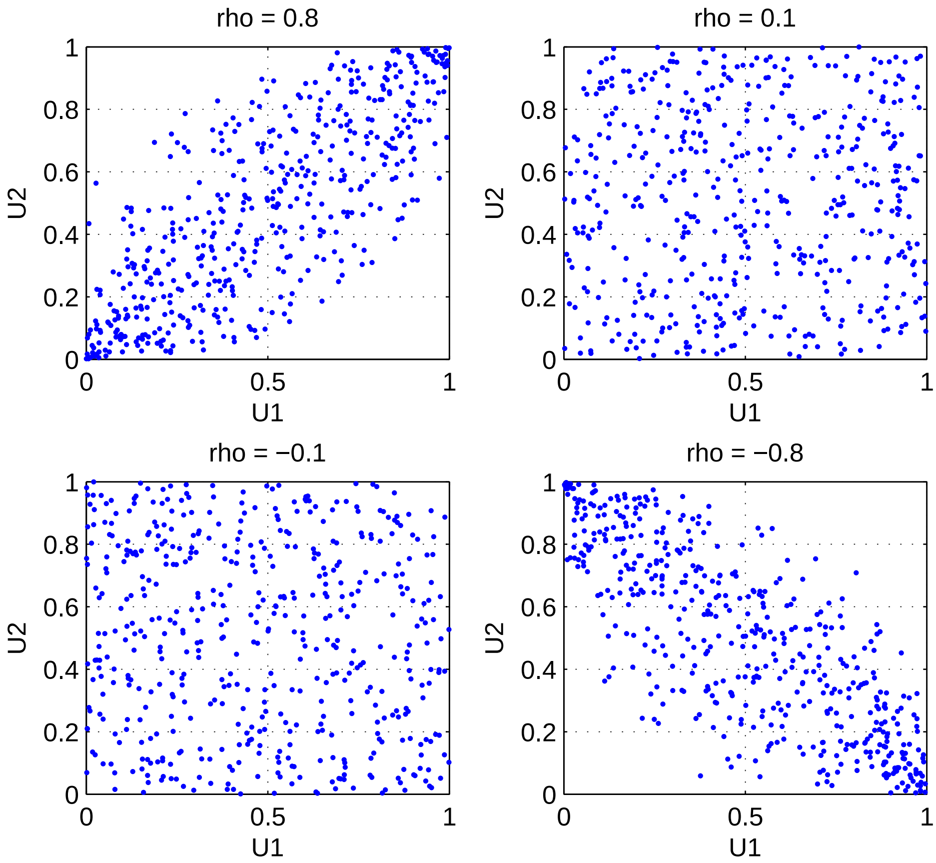


Figure 2.7: Bivariate normal copula simulation.

2.5.3 Relationship Between Input and Copula Correlations

While the dependence between the random variates generated by the copula model is determined by the $\rho_{i,j}$ that are inputs to the model, these $\rho_{i,j}$ do not measure the correlation between the variates generated by the copula. To see this, consider a bivariate Gaussian copula. Once we have generated the two correlated standard normal vectors Z_1 and Z_2, the inverse standard normal cdf transforms these vectors to make them a copula $\{U_1, U_2\}$.

Now the linear correlation $\rho(Z_1, Z_2)$ between Z_1 and Z_2 is clearly *not* the same as the correlation between U_1 and U_2, as a non-linear transformation has been applied. In the general case, there is no simple relationship between the before and after correlations. For this reason, *rank correlations*[12] such as Spearman's ρ or Kendall's τ are often used instead as these are invariant under any monotonic transformation.

The definitions of these rank correlation coefficients, are somewhat more technical

[12]Informally stated, rank correlations measure the degree to which values of the same magnitude of a set of random variables are associated (occur at the same time).

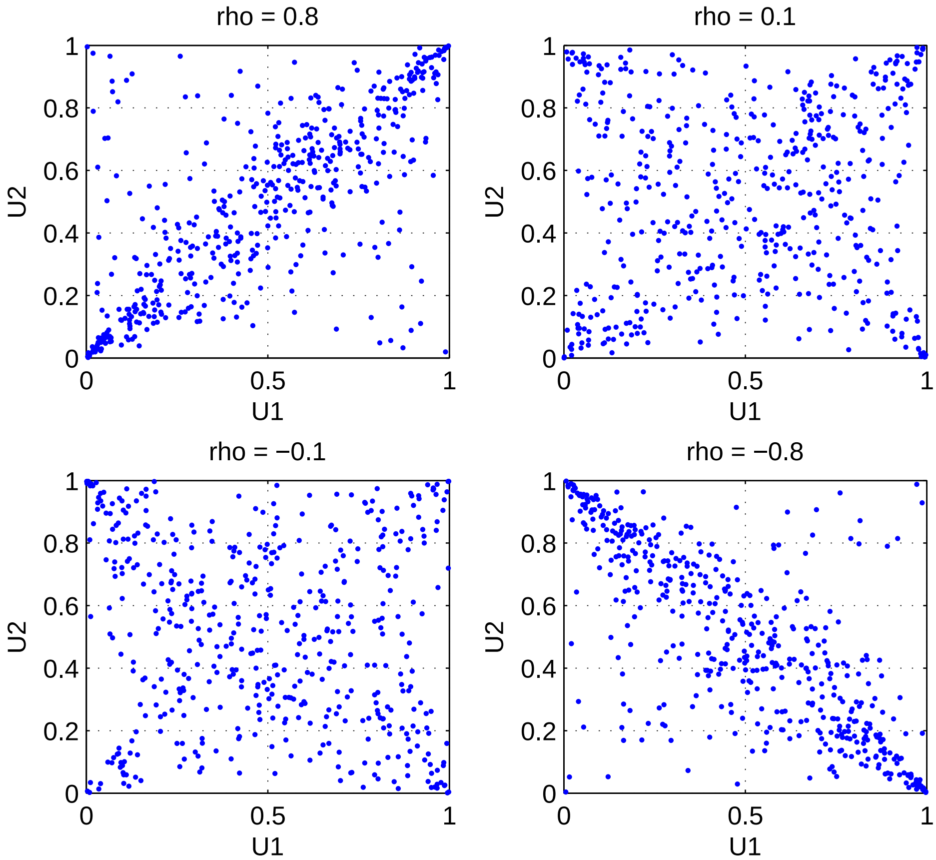


Figure 2.8: Bivariate t-copula simulation with one degree of freedom.

than the Pearson coefficient, but they are already implemented in most software packages. Like the Pearson correlation coefficient, ρ, these measures lie on the interval $[-1, 1]$ with zero for independence and 1 or -1 for deterministically related data.

In the Gaussian copula case the relationship between τ and ρ is given as $T = \frac{2\arcsin(R)}{\pi}$ where T is the matrix of Kendall's τ and R the linear correlation matrix[13].

Consider the default time simulation for two assets in Figure 2.6. The input correlation matrix (Pearson) and the corresponding T-matrix using the above formula are:

$$R_{input} = \begin{bmatrix} 1 & .8 \\ .8 & 1 \end{bmatrix}, T_{input} = \begin{bmatrix} 1 & .5903 \\ .5903 & 1 \end{bmatrix}$$

If we measure the copula correlation:

$$R_{NC} = \begin{bmatrix} 1 & .7202 \\ .7202 & 1 \end{bmatrix}, T_{NC} = \begin{bmatrix} 1 & 0.5913 \\ 0.5913 & 1.0000 \end{bmatrix}$$

[13] Cf. Nelsen (1999) and Trivedi and Zimmer (2005) for this related formulas for other copulas and definitions of correlation coefficients.

From the above, we see that while the Pearson matrix changes, the Kendall rank coefficient from the simulation remains quite the same (only subject to some minor perturbations due to the fact that only 500 variates are simulated for each vector).

2.5.4 Simulation Algorithms for Copulas

Sampling from the Normal Copula

The following procedure has been applied in the generation of random variables from the Gaussian copula with correlation matrix Σ:

1. Find a decomposition matrix M so that $\Sigma = MM'$.[14]
2. Draw an n-dimensional vector $\mathbf{v} = [v_1, ..., v_n]$ of standard normal variates[15].
3. Let $\mathbf{v}^* = \mathbf{v}M$. $\mathbf{v}^*$ is now a vector of correlated standard normal variates.
4. Transform $\mathbf{v}^*$ into a uniformly distributed random vector by applying the standard normal cdf: $\mathbf{u} = \Phi(\mathbf{v}^*)$.

Sampling from the Student t Copula with v Degrees of Freedom

Similar to the above, we state a simulation algorithm for the Student t copula:

1. Find a decomposition matrix M of the correlation matrix $\mathbf{\Sigma}$ so that $\Sigma = MM'$.
2. Generate an n-dimensional vector $\mathbf{v} = [v_1, ..., v_n]$ of standard normal variates.
3. Generate an independent χ_v^2 random variable s.
4. Let $\mathbf{v}^* = \mathbf{v}M$. $\mathbf{v}^*$ is now a vector of correlated standard normal variates.
5. Let $\mathbf{x} = \mathbf{v}^* \sqrt{v/s}$.
6. Transform $\mathbf{v}^*$ into a uniformly distributed random vector by applying the t distribution cdf with v degrees of freedom: $\mathbf{u} = t_v(\mathbf{v}^*)$.

Now we have provided sufficient details on the theory behind the reduced form models and the copula approach to implement a simple pricing model from scratch; later chapters will treat more security specific issues as well as parameter estimation and model calibration.

[14] There are several algorithms for doing this; most linear algebra packages have at least one implementation. For symmetric positive definite matrices such as the correlation matrix with positive entries, an efficient method is the Cholesky decomposition. For details see for instance Cheney and Kinciad (2007).

[15] Most software packages and programming languages such as Excel and MATLAB contain methods for generating standard normal variates. If such functions are not available, a uniform variates $\{u_i\}$ can be transformed to a distribution specified by the cdf $F(x)$ by $x_i = F^{-1}(u_i)$.

2.6 Estimating Default Probabilities and Calibrating a Reduced Form Model

The discussion so far has focused on the theory behind reduced form models and how they can be implemented. To round up the discussion on these models, some notes on parameter estimation is included.

2.6.1 Using a Single Bond Price - Constant Default Probability

Consider first a simple example where the default probability is assumed constant. We now consider a single risky bond with maturity 3 years. The market yield to maturity is 7.0% with coupon rate 6.0% paid semi-annually. With a face value of 100, this corresponds to semi-annual coupons of 3. The risk free rate is assumed to be 5%. The amount recovered given default is constant equal to 40.

The price of the risky bond is then:

$$\sum_{t=1}^{5} 3e^{.07t/2} + 103e^{.07 \cdot 3} = 97.01$$

and the risk free bond price is similarly:

$$\sum_{t=1}^{5} 3e^{.05t/2} + 103e^{.05 \cdot 3} = 105.58$$

If we assume that the default probability is constant equal to Q over the horizon, we get the following:

Time	Cash flow	RF DF	Risky DF	Value RF Bond	LGD	PV exp. Loss
0	3	1	1	105.58		
0.5	3	0.98	0.97	105.17	65.17	63.57 Q
1	3	0.95	0.93	104.76	64.76	61.60 Q
1.5	3	0.93	0.9	104.34	64.34	59.69 Q
2	3	0.9	0.87	103.90	63.90	57.82 Q
2.5	3	0.88	0.84	103.46	63.46	56.00 Q
3	103	0.86	0.81	103.00	63.00	54.22 Q

Figure 2.9: Estimating default probability from a single bond price.

Figure 2.10 illustrate the computation of the risk neutral default probability Q. Columns 3 and 4 contain the risk free and risky discount factors, respectively. Column 5 is the value of the risk free bond at time t. Column 6 is the loss given default at time t given as the difference between the risk free bond value and the loss given default

(40). The final column gives the discounted loss given default times the risk neutral default probability.

Summing the expected loss from the last column above and equating it to the price difference between the risky and risk free bond we get $352.9Q = 105.58 - 97.01$ or $Q = .0243$ assuming as before the risk premium is entirely constituted by credit risk. Here Q represents the probability of the bond issuer defaulting on any half year interval conditioned on previous survival.

2.6.2 Using a Set of Bond Prices

A weakness of the above procedure is that it gives a flat term structure of credit risk (constant default probability). However, if we have a set of bonds with different maturities for the same entity, we can use a *bootstrap* procedure to estimate default probabilities for shorter time intervals.

From the previous section we see that a bond with maturity in one year paying either the full face value of 100 at maturity or some recovery amount R, the price difference d_1 between this risky bond (whose value is $(100(1 - q_1) + Rq_1)e^{-r}$ and a similar zero coupon bond (with value $100e^{-r}$) is given by:

$$d_1 = (100 - R)e^{-r} q_1$$

where q_1 denotes the default probability during the first year. We can now solve this equation to obtain the one year default probability as before.

Next consider a bond with a coupon payment in one year from now and maturity in two years. The price difference between this and a two year zero coupon bond is:

$$d_2 = (100 - R)e^{-r \cdot 1} q_1 + (1 - R)e^{-r \cdot 2} q_2 = d_1 + (100 - R)e^{-r \cdot 2} q_2$$

And generally:

$$d_n = d_1 + d_2 + ... + (100 - R)e^{-r \cdot n} q_n$$

Using this iterative procedure (that can of course be modified to handle coupon bonds) coupled with the method from the preceding section, the credit curve can be constructed from an arbitrarily large set of corporate bonds.

Illustration

Consider a set of risky and risk free bonds with maturities from one to five years, all with coupon rates of 6,0% and 40 recovery value in the case of default. Further assume that defaults can occur only at coupon payment dates and that the yields for the risky and risk free bonds are as given by columns 2 and 3 below. The prices implied by these yields follow from column 4 and 5 (these are computed as the expected value of future cash flows discounted by the respective yields). For each price, a table similar to that in Figure 2.10 needs to be compiled.

The annual default probabilities computed by the above method follow in column 6. For instance, the probability of the firm defaulting during year 3, conditioned on survival up to the start of the year is .0342. The cdf following from the data in this example is plotted in Figure 2.11.

Maturity	Risky YTM	RF YTM	RF Price	Risky Price	Price Difference	Default prob
1	3.75%	3.00%	102.87	102.10	0.77	0.0109
2	3.98%	3.25%	105.40	103.89	1.51	0.0354
3	4.15%	3.50%	107.61	105.46	2.15	0.0342
4	4.77%	4.00%	109.28	105.46	3.83	0.0651
5	5.14%	4.25%	110.67	106.99	3.68	0.0642

Figure 2.10: Estimating default probability from a single bond price.

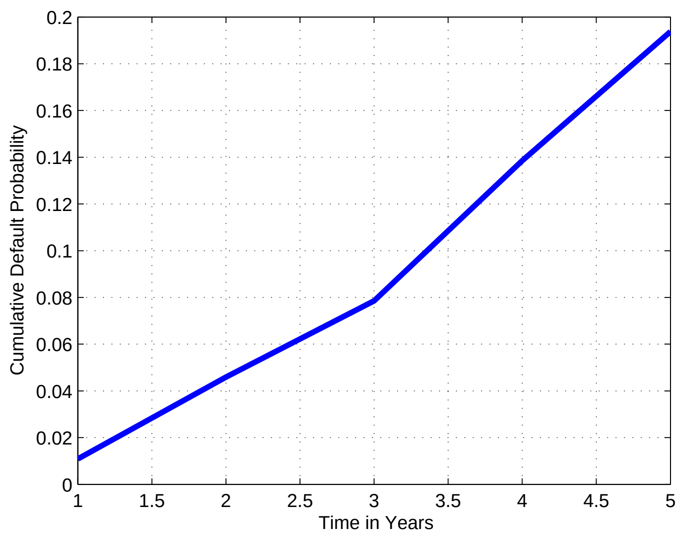


Figure 2.11: Cumulative probabilty function estimated using a set of bond prices.

2.6.3 Using CDS Prices

There are two important problems with the method previously suggested:

- It is relatively burdensome computationally (large amounts of data must be compiled and processed).

- In many cases, there aren't sufficiently many maturities available for corporate bonds.

Therefore, is may often be more convenient to estimate default probabilities from CDS spreads when these are available.

Section 4.2 discusses the CDS instrument in greater detail. Here we'll just state the pricing equation for a simple CDS structure that ignores premium accrual:

$$\sum_{i=1}^{N} \delta e^{-rt_i} = \mathbb{E}\left[L(\tau)e^{-r\tau}\right] \tag{2.6.1}$$

where δ is the fixed leg premium and $L(\tau)$ is the loss given default at time τ (assumed to be paid immediately upon default). Now we want to use the known CDS spreads[16] to compute the default probabilities using the models developed so far. We consider a simplified example:

Illustration

Consider CDS on a bond with face value 100 maturing in 2 years. Defaults can occur at times $\{0.5, 1.5\}$, and CDS premia δ are paid at the end of each year. Assume the risk free interest rate is 3% and the annual default probability is constant $Q(t) = q$ for $t \in \{1, 2\}$. Amount recovered given default is constant $R(\tau) = 60$.

t	Default Prob.	Risk Free DF	Expected payments	Expected NPV
0.5	q	0.9851	$40q$	$0.9851 \cdot 40q$
1.5	$(1-q)q$	0.9560	$40(1-q)q$	$0.9560 \cdot 40(1-q)q$

Figure 2.12: Floating leg cash flows.

t	Survival Prob.	Risk Free DF	Expected payments	Expected NPV
1	$1-q$	0.9704	$\delta(1-q)$	$0.9704\delta(1-q)$
2	$(1-q)^2$	0.9417	$\delta(1-q)^2$	$0.9417\delta(1-q)^2$

Figure 2.13: Fixed Leg cash flows.

From the above table we see that we can recover the average annual risk neutral default probability q by equating the floating and fixed leg cash flows and finding the roots of the second degree polynomial (ie. finding the *break-even* default probability):

$$0.9851 \cdot 40q + 0.9560 \cdot 40(1-q)q = 0.9704\delta(1-q) + 0.9417\delta(1-q)^2$$

[16] In reality there is a *bid-ask* spread on CDS's. It is therefore common in practice to use an arithmetic average of these, or *mid-market* CDS spread.

The left hand side are the floating leg cash flows from Figure 2.12, the right hand side the fixed cash flows from Figure 2.13. For $\delta = 2$, $q = .0469$.

To estimate a more realistic term structure of default probabilities; multiple CDS spreads can be used similarly to the multiple bond prices example in Section 2.6.2.

2.6.4 Physical or Risk Neutral Probabilities?

So far we have exclusively dealt with risk neutral default probabilities. To see why this must be so, consider the examples in Sections 2.6.1 and 2.6.2. Here we use the risk free interest rate as a discount factor. This means that expected losses must be computed with respect to an equivalent *risk neutral* probability measure, so the probabilities estimated from the bond prices are the *risk neutral* default probabilities.

From this discussion it is quite clear that the hazard rates (λ) are dependent on which probability measure is used, and care must be taken when switching from a valuation to a risk management perspective as a change of measure means shifting the probability distribution. Hull (2006) pp. 488-489 discusses some of the reasons for the difference between the risk neutral and physical default probabilities.

2.6.5 Calibrating Copula Models and the Relationship to Structural Models

The simplest way to estimate the correlation parameters in a copula model is to use estimates of asset, or yet simpler, equity price correlations as a proxy. These can be computed readily from time series using built-in functions in any standard software package such as Excel. As the causal relationship between asset prices and default probabilities is somewhat weak, another and perhaps more suitable proxy is the correlation between CDS spreads (see Section 4.2 for discussion on CDS's) as the changes in these spreads can be interpreted as changes in markets assessments on the likelihood of default. The advantage of the latter is that the correlations are risk neutral.

If the asset price correlation approach is taken, the normal copula approach can be shown to yield a correlation structure equivalent to the Merton model discussed in the next chapter if asset prices follow a geometric Brownian motion. If we let q_A and q_B denote the risk neutral one-year default probabilities for assets A and B, we obtain z_A and z_B so that:

$$\begin{aligned} q_A &= \Phi(z_A) \\ q_B &= \Phi(z_B) \end{aligned} \tag{2.6.2}$$

The joint probability of both these assets defaulting is given by:

$$\mathbb{P}(Z_A < z_A, Z_B < z_B) = \Phi^2(z_A, z_B, \rho)$$

In the bivariate normal copula, the same probability looks like:

$$\mathbb{P}(\tau_A < 1, \tau_B < 1) = C((F_A(1)), (F_B(1)), \rho) = \Phi^2(\Phi^{-1}(F_A(1)), \Phi^{-1}(F_B(1)), \rho) \quad (2.6.3)$$

Noting that the probability of default during year one is:

$$q_i = \mathbb{P}(\tau_i < 1) = F_i(1)$$

So (2.6.3) is really:

$$\Phi^2(\Phi^{-1}(q_A), \Phi^{-1}(q_B), \rho)$$

But by (2.6.2), $\Phi^{-1}(q_i) = Z_i$ which completes the argument.

Using Rank Correlations

In Section 2.5.3, we argued that contrary to the Pearson correlation coefficient, rank correlations are invariant under transformations such as copula functions, so rank correlation may be a more attractive measure. Hence, the rank correlation produced by the copula model should be the same as for the data it is estimated from. In the Gaussian (and some other simple cases), we have simple formulas relating the two measures. If, for example we have measured the input data rank correlation and want to use a Gaussian copula, we can use the relationship $T = \frac{2\arcsin(R)}{\pi}$ to find the correlation matrix R to use in the simulation algorithm. For a general copula where we don't have such a relationship, it can be obtained numerically using simulation. This procedure gets somewhat more complicated if the individual correlations are allowed to vary.

Chapter 3

Structural Credit Risk Models

Structural models are the second of the two most widely used approaches to credit risk modeling. The core idea is the realization of the "option-like" feature of corporate (or sovereign) securities due to limited liability. The application of option pricing theory to problems of capital structure and valuation of corporate security was stated by Merton (1974) in one of the earliest papers on the Black-Scholes-Merton model.

3.1 The Merton Model

As an illustration of the concept, we consider first a simple example close to Merton's model. Let $A(t)$ denote the value of a firm's assets at time t, and assume it follows the following stochastic process under the *physical* probability measure $\mathbb{P}$:

$$dA(t)/A(t) = \mu dt + \sigma dW(t) \tag{3.1.1}$$

where μ and σ are constants.

Under the standard assumptions[1] of arbitrage-free markets, we can restate the same process under the *risk-neutral* probability measure $\tilde{\mathbb{P}}$:

$$dA(t)/A(t) = rdt + \sigma d\tilde{W}(t) \tag{3.1.2}$$

We now want a precise mathematical formulation of the two claims on the total value of this firm, equity and debt. We will therefore make certain simplifying assumptions that allow for a convenient analytical treatment. First we define the debt value at time t, $D(t,T)$ as a cash flow D at time T where $D \geq 0$. Here D denotes the face value of the debt.

[1] Hereunder that the firm pays no dividend cash flows. See for instance Goldstein et al. (2001) or Duffie (2001) for a discussion on these topics.

Equity $E(t,T)$ is defined as the value of an *option*; the holder can choose either to receive residual $E = A(T) - D$ or zero at the maturity of the debt. If the equity holder chooses the zero cash flow, debt holders receive the entire firm value $A(T)$. In this interpretation of the model, $A(T) < D$ corresponds to default. To summarize:

$$E(T) = \begin{cases} 0 & \text{if } A(T) < D \\ A(T) - D & \text{if } A(T) \geq D \end{cases} = \max(A(T) - D, 0)$$

$$D(T) = \begin{cases} A(T) & \text{if } A(T) < D \\ D & \text{if } A(T) \geq D \end{cases} = \min(A(T), D) = A(T) - max(A(T) - D, 0)$$

Under standard assumptions, equity is of course precisely equivalent to a European call option and debt is a portfolio with a long position in the asset and a short call option. Therefore the claims can be valued by a straight-forward application of the Black-Scholes-Merton formula[2]:

$$E(t) = A_t \Phi(d_1) - e^{-r(T-t)} D \Phi(d_2)] \tag{3.1.3}$$

with:

$$d_1 = \frac{\ln(\frac{A_t}{D}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{3.1.4}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{3.1.5}$$

The debt value can then be found as the residual $D(t) = A(t) - E(t)$[3].

$\Phi(-d_2)$ in the above formula can be interpreted as the *risk neutral* probability of default similar to the probability of a European call being out of the money at the time of exercise, or $A_T < D_T$. We can show this using the properties of the Wiener process, writing out the condition for default using the closed form expression for the asset price process under $\tilde{\mathbb{P}}$:

$$A_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(\tilde{W}_T - \tilde{W}_t)} = A_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma\sqrt{T-t}Z_{0,1}} < D \tag{3.1.6}$$

as $\tilde{W}_T - \tilde{W}_t \sim \Phi(0, \sqrt{T-t})$, which can be rewritten in terms of a standard normal random variable $Z_{0,1}$. Rearranging we get the default probability as a standard normal

[2]Usually, the value of these claims are known as they are interpreted as the observable market prices of debt and equity. We can then use the model to imply the asset volatility σ as in Section **??**.

[3]Throughout this thesis, we assume that capital structure does not affect asset values. More advanced structural models (for instance Goldstein et al. (2001) relax this assumption. These models are less commonly used in derivatives pricing however.

probability:

$$\tilde{\mathbb{P}}[A_T \leq D] = \tilde{\mathbb{P}}[Z_{0,1} \leq z_0] = \tilde{\mathbb{P}}\Big[Z_{0,1} \leq \frac{\ln(\frac{A_t}{D}) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\Big] = \Phi(-d_2) \qquad (3.1.7)$$

3.2 Extending the Merton Model

The above model is extremely simplistic with regard to the assumptions on capital structure. Several criticisms may arise:

- Defaults do not only occur at maturity T.
- We need to allow for coupon payments.
- The process $A(t)$ is hard to estimate as it is not a traded asset.
- We may need to account for portfolios of several assets $A_i(t)$ that are correlated.
- and so forth..

To fix the first point above means considering equity as an American derivative. Generally, there are no closed form solutions to the problem of valuing finite maturity American derivatives, so either analytic approximations[4] or numerical methods are required. The latter is discussed in Appendix B.

If we are content with the assumption that defaults can only occur at maturity, the Merton model can be extended in such a way that we can study asset portfolios. This is done in the next section. The so-called Black-Cox model, where defaults occur once asset values fall below a certain threshold, is treated in the final section of this chapter.

3.3 Correlations in the Merton Model

Next we consider an approach to modeling asset portfolios in the Merton model that was first proposed by Vasicek (1987). Given a portfolio of assets indexed $n = 1, \ldots, N$, we denote the probability of firm i defaulting prior to the time T of its debt maturing, $p_{d,i}$, which as before is given by $\Phi(-d_{2,i})$. It is further assumed for simplicity that all firms are equal, having identical capital structures and following identical Wiener processes under the risk neutral measure $\tilde{\mathbb{P}}$:

[4]Such as assuming infinite maturity.

$$dA_t^i/A_t^i = rdt + \sigma d\tilde{W}_t^i \tag{3.3.1}$$

The approaches to modeling portfolios of correlated assets are more or less straight forward extensions of either the single credit framework discussed so far or the Black-Cox model in the next section. Several methods for modeling the interdependencies between credit events are proposed:

1. Directly modeling *correlated processes*; ie. letting $\tilde{W}_t^1, .., \tilde{W}_t^N$ be a set of brownian motions with $d\tilde{W}_t^i d\tilde{W}_t^j = \rho_{i,j}$.

2. *Factor models*: Letting the Z_t^i be the random variable (or state variable) determining the time of default for asset i, Z_t^i is now a function of a set of common factors $\{X_1, ..., X_n\}$ affecting, to a greater or lesser extent, all assets in the portfolio, in addition to an idiosyncratic risk factor. We can think of several macroeconomic variables that can be used as proxies in common factor models: interest rates, GDP (or proxies such as stock indices), CDS spreads and so forth.

3. *Contagion models* is another commonly proposed method in the literature. Instead of simply letting the asset values determine the timing of the default, we could also imagine the default thresholds as correlated random variables[5].

   Technically, such *contagion effects* can be implemented using indicator variables for certain events affecting credit risk. Theoretically, several mechanisms can be triggered, for instance, a higher default threshold (that eventually returns to normal after some time), increased volatility, asset price jumps and so forth.

3.3.1 Common Factor Models

The default probability of a single firm i is given as a function of a standard normal variable X_i. Therefore we can specify the correlation structure between the firms in the portfolio in terms of the X_i's. One approach is to use so-called common factor models where we introduce a set of factors Y_j, $j = 1, .., M$, and let X_i be a weighted sum of the *systemic* risk factors Y_i and a *idiosyncratic* risk factor ϵ_i where Y_j and ϵ_i are independent. For simplicity, we consider only the single factor case:

$$X_i = \sqrt{\rho_i} Y + \sqrt{1 - \rho_i}\, \epsilon_i \tag{3.3.2}$$

[5] For instance, the default threshold could be increased in a financial crisis where short term financing is hard to obtain, or if a major firm in a supply chain defaults, the probability of default increases for firms further up the chain.

Here Y is a global or common risk factor affecting all firms in the portfolio to various degrees measured by the correlation coefficient ρ_i that varies between the different entities thus forming the structure of correlations between firms as the ϵ_i's are independent. We can therefore think of the terms $\sqrt{\rho_i}$ and $\sqrt{1-\rho_i}$ as firm i's exposure to the systemic and idiosyncratic risk factors

We now make two simplifying assumptions for the sake of easier notation:

- All firms have the same correlation ρ with the common factor.
- All firms have the same default probability p_d (this follows from equal dynamics and capital structure for all firms).

Consider now the conditional default probability:

$$
\begin{aligned}
p_{d,Y} &= \mathbb{P}[X_i < x|Y] \\
&= \mathbb{P}[\sqrt{\rho}Y + \sqrt{1-\rho_i}\,\epsilon_i < x|Y] \\
&= \mathbb{P}[\sqrt{1-\rho}\,\epsilon_i < x - \sqrt{\rho}Y|Y] \\
&= \mathbb{P}\Big[\epsilon_i < \frac{x-\sqrt{\rho}Y}{\sqrt{1-\rho}}\Big|Y\Big] \\
&= \Phi\Big[\frac{x-\sqrt{\rho}Y}{\sqrt{1-\rho}}\Big]
\end{aligned}
$$

The last equality above follows from the independence between the systemic and idiosyncratic risk factors. If we have some estimate of the default probability of each firm, that is, the individual default probabilities, $p_i = p$, we can find the default thresholds $-d_2$ from equation 3.1.7; $p = \mathbb{P}[A_T \leq D] = \Phi(-d_2)$ which gives $d_2 = \Phi^{-1}(p)$, so that:

$$
p_{d,Y} = \Phi\Big[\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1-\rho}}\Big] \tag{3.3.3}
$$

Equation 3.3.3 is an expression for the default probability of a firm conditional on the realization of the common factor Y expressed in terms of the common correlation coefficient between each firm and the systemic risk factor and the default probability of the firm, both of which are possible to estimate from market data.

3.3.2 Portfolio Loss Rates

We can now represent losses on a portfolio level by introducing a set of indicator variables I_n, $n = 1, ..., N$ for each of the N firms representing whether a firm has defaulted.

A counting variable $S = \sum_{n=1}^{N} I_n$ denotes the number of firms in the portfolio having defaulted. The default rate can therefore be given as $R = S/N$. With the above framework, we are now interested in the distribution of R for instance given by its unconditional cumulative distribution function $F(r, p, \rho) = \mathbb{P}[R \leq r]$. Several approaches are available, here divided in two. The LHP model (Vasicek (1987)) that makes a number of simplifying assumptions so that closed form valuation formulas are available for many of the instruments considered and secondly, models with more realistic assumptions that often require numerical approaches:

1. **Large Homogeneous Portfolio** Assume a large number of credits in the portfolio ($N \to \infty$). R will now converge to p (the individual, unconditional default probabilities) as defaults are independent conditioned on the common factor, so $F = \Phi(x)$.

2. **Finite or Heterogeneous Portfolios** The most widely traded CDO instruments today are CDO indices such as the *iTraxx* and the *CDS NA IG* which are single tranche CDOs on an equally weighted portfolio of 125 reference entities. This is probably a too small number of firms for the law of large numbers to guarantee accurate pricing.

In the case of a heterogeneous portfolio, the reference entities have different characteristics such as default probabilities (above assumed to be equal), unequal weights in the portfolio, different recovery rates and so forth. Under the first assumption it is possible to derive analytical pricing expressions involving binomial probabilities, whereas under the second this becomes much more complicated.

3.3.3 Default Rates in a Portfolio

Equation 3.3.3 gives the default probability contingent on the realization of the common factor Y as:

$$p_{d,Y} = \Phi\Big[\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1-\rho}}\Big]$$

Next, we are interested in the probability distribution of defaults in a portfolio *contingent* on Y. This depends only on the idiosyncratic risk factors that are independent. So if X is a random variable denoting the number of defaults in a portfolio of N assets, this random variable would have a binomial distribution characterized by the individual default probabilities $p_{d,Y}$. The probability of exactly $x \leq N$ defaults occurring is therefore:

$$\mathbb{P}(X = x|Y = y) = \binom{N}{x}(p_{d,Y})^x(1 - p_{d,Y})^{N-x} \tag{3.3.4}$$

The unconditional probability of exactly x defaults is therefore a probability weighted "sum" over all the possible realizations of Y.

$$\mathbb{P}(X = x) = \int_{y\in\omega} \binom{N}{x}(p_{d,Y})^x(1 - p_{d,Y})^{N-x} f(y)dy \tag{3.3.5}$$

If the systemic factor is normally distributed, then $f(y) = \phi(y) = \frac{d}{dy}\Phi(y)$, so we get the cumulative distribution function $F(m) = \mathbb{P}(X \leq m)$ as:

$$F(x) = \mathbb{P}(X \leq m) = \sum_{x=0}^{m} \binom{N}{x} \int_{y=-\infty}^{\infty} (p_{d,Y})^x(1 - p_{d,Y})^{N-x}\phi(y)dy \tag{3.3.6}$$

Illustration

The expression 3.3.6 can be evaluated numerically using standard software packages. Appendix C contains a script implementing it using a simple numerical integration scheme in MATLAB.

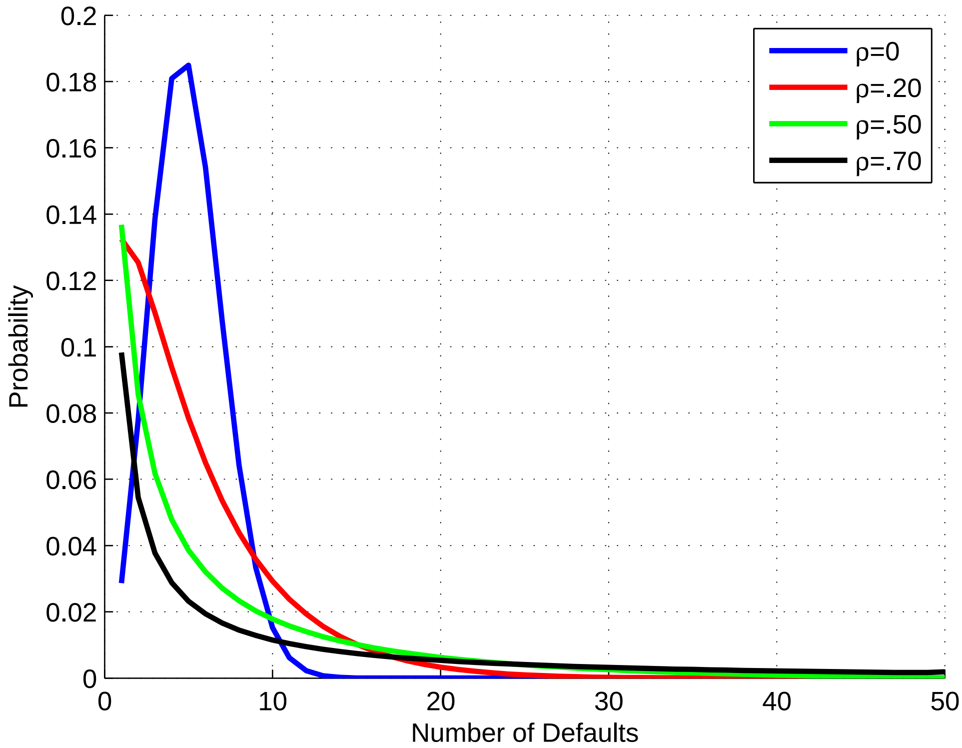


Figure 3.1: The density function for the number of asset defaults.

As an illustration we let $p = .1$ be the one year individual default probabilities that is equal over the portfolio of $N = 50$ assets. Figure 3.1 shows the density function for the number of defaults in this portfolio given various levels of the correlation coefficient ρ. We notice that for a large ρ the right tail of the distribution is fatter, much as one would expect.

Using this model, for instance as implemented in Appendix C, it is simple to value credit derivatives written on a portfolio of credit derivatives. It is also applicable for computing risk measures such as value at risk for the underlying loan portfolio.

There is one computational issue with this model however, namely that it uses a "direct" method for computing the binomial coefficients. This may become computationally burdensome and even imprecise under some implementations for large portfolios. Hull and White (2004) discusses some alternative methods for evaluating such probabilities.

3.4 The Black-Cox Model

A more "advanced" structural model was introduced by Black and Cox (1976) that allows for defaults prior to maturity. My presentation, largely based on Hull, Predescu and White (2006) is somewhat truncated and informal as the mathematics of this model is quite advanced and beyond the scope of this thesis. Rather, I will attempt to give a intuitive presentation aimed at a simulation-based implementation.

Next we consider the case where defaults can occur at any time $t \leq T$ prior to maturity. As before, the dynamics of the asset prices are given by Equation 3.1.2, ie. a geometric Brownian motion with constant drift r and volatility σ under the risk neutral measure. Time of default is now defined as the first time τ the asset value A_t passes below some threshold K, often referred to as the *first passage time* to K:

$$\tau = \inf\{\tau : t_0 \leq \tau \leq T | A_\tau < K\} \tag{3.4.1}$$

As in the Merton model, we want to express the default threshold at any point in time t in terms of a standard normally distributed random variable $X(t)$ rather than the log-normal asset price. This involves expressing the above inequality in terms of $X(t)$ by taking the logarithms as shown previously. This gives the threshold for $X(t)$:

$$K^*(t) = \frac{\ln K - \ln A_0 - (r - \sigma^2/2)t}{\sigma}$$

where K is a firm specific constant determining the probability of default. Letting

$$\beta = \frac{\ln K - \ln A_0}{\sigma}$$

and

$$\gamma = -\frac{r - \sigma^2/2}{\sigma}$$

we obtain the default threshold as a linear equation in time: $K^*(t) = \beta + \gamma t$. Assuming a non-stochastic default threshold $K(t)$, the that the probability of default on a time interval $[t, T]$ is given as[6]:

$$\mathbb{P}[t \leq \tau \leq T] = \Phi\left(d_1\right) + e^{2(X(t)-\beta-\gamma t)\gamma}\Phi\left(d_2\right) \tag{3.4.2}$$

with

$$d_1 = \frac{\beta + \gamma \cdot (T - t) - X(t)}{\sqrt(T - t)}$$

$$d_2 = \frac{\beta + \gamma \cdot (2t - T) - X(t)}{\sqrt(T - t)}$$

Among the main drawbacks of first passage models is the analytical complexity. The derivation of the above formula is technical and lengthy and therefore omitted here. Though complex, there are several conditions under which it is possible to derive closed form expressions for the values of corporate securities.

Among the simplest of these is debt is rolled over infinitely. Here, corporate securities are equivalent to perpetual American options. In this case, the partial differential equation describing the option price reduces to a second degree differential equation in one variable which gives a simple algebraic solution. For an example of such analysis, see Leland (1994), and Goldstein et al. (2001), and references cited therein.

3.4.1 Specification and Solution Method

As mentioned, a numerical approach is often necessary for first passage models. This method takes a simple approach by drawing a set of standard normal variates $[X_{t_1}, X_{t_2}, ..., X_{t_n}]$ that are compared to the default threshold $K^*(t) = \beta + \gamma t$. For correlation modeling, we have the usual options such as copulas, common factors and so on. Finally, it is worthwhile to note that there are numerical methods by which the model parameters can be determined so as to match the default probabilities on an interval $[t_k, t_{k+1}]$ in this discrete approximation to Equation 3.4.2[7].

Another approach to the American option valuation problem developed by Longstaff and Schwartz (2001), the least squares Monte Carlo method (LSMC), is explained in the appendix. This method takes uses a dynamic programming technique involving the

[6] A similar derivation, in the context of American option valuation is found in Shreve (2004) chapter 9.

[7] Cf. Hull (2001).

method of least squares to determine the optimal exercise time for each simulated price path. Note that the optimality criterion implies a model different from the above model with a time dependent threshold that is suboptimal. Optimal exercise by equity holders implies different prices for corporate securities as discussed in Leland (1994).

This method is simple to implement and has many advantages when it comes to multi-factor models. The disadvantage of this method is that to determine the optimal exercise price for each path requires a large number of function evaluations, so that precise computation is very time consuming when there are many factors.

Illustration

Let us consider a simple example to see how the Black-Cox model can be implemented in valuing a multi-name credit derivative problem. We let $\mathbf{t} = [.25, .50, .75]$ be the possible default times. To value a credit derivative, we need to simulate credit events for a portfolio of correlated assets at three different points in time. Asset i defaults at time .25 if $X_i(.25) < K^*(.25)$. We consider a simple model where the dynamics of the asset price is $dX_i(t) = \rho dY(t) + \sqrt{1-\rho^2}\, d\epsilon_i(t)$.

So we are considering a single-factor model. As we have seen, the default threshold determines the individual default probabilities. The procedure is therefore similar to the Merton model, only it is repeated for the three time-steps; for each realization of $Y(t)$ generate $\epsilon_1(t), ..., \epsilon_N(t)$ for the N assets in the portfolio. $Y(t)$ will now represent the systemic, or market risk factor and $\epsilon_i(t)$ the idiosyncratic factors for each firm.

Next, compute and compare the $X_i(t)$ to $K(t)$, and determine the time of default as the infimum over t of the set $X_i(t) < K(t)$ as defined in Equation 3.4.1.

The default threshold is on the form $K^*(t) = \beta + \gamma t$. We now obtain the discretized asset price process as:

$$X_i(t) = X_i(t-1) + \rho Y(t) + \sqrt{1-\rho^2}\,\epsilon_i(t)$$

$$\epsilon_i(t) = \epsilon_i(t-1) + \sqrt{\Delta t}\, Z_{0,1}$$

$$Y(t) = Y(t-1) + \sqrt{\Delta t}\, Z_{0,1}$$

Figure 3.2 illustrates ten correlated paths from this simulation scheme. We can think of this as one realization of a portfolio of 10 assets. Each simulation of the ten assets require one simulation of the systemic factor. To value the portfolio, we clearly need to repeat this process a large number of times. The thin blue line marks the default threshold. The time of default for an asset is the earliest time its path passes below the threshold line marked by circles.

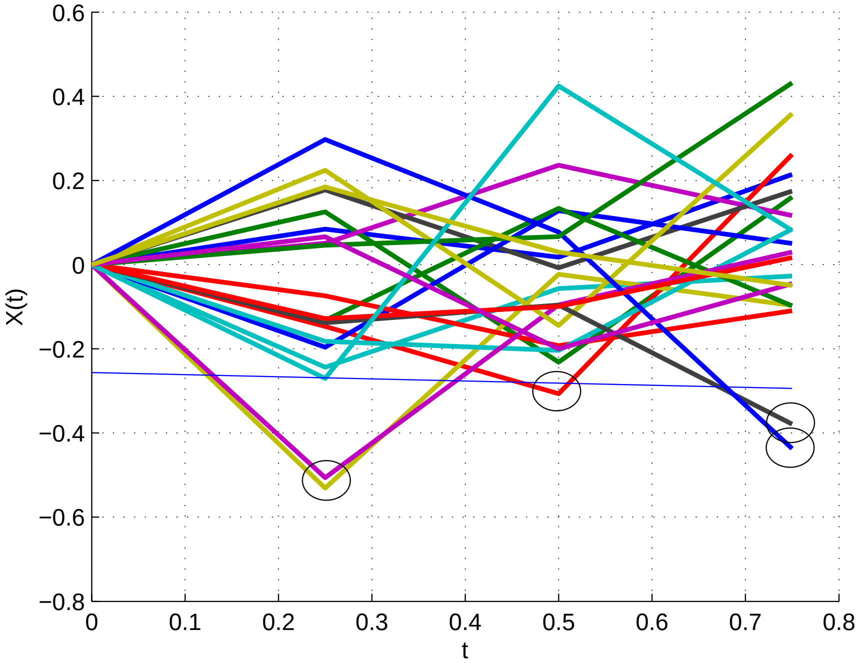


Figure 3.2: Sample paths from Black-Cox simulation.

3.5 Alternative Structural Models

In this section we have considered some classical examples of structural models where the primitive state variable is the asset price which follows a geometric Brownian motion process, time of liquidation equals the time of default and interest and recovery rates are constant. Many extensions have been proposed:

- *EBIT models*: rather than using the asset price as the primitive variable, some authors such as Goldstein et al. (2001) have proposed to let earnings before interest and taxes be a stochastic process in turn determining, equity, debt and asset prices. Now the current asset price endogenously given as the discounted expected value of future EBIT flows, which among other things allow for a more complex relationship between interest rates and asset prices.

- *Stochastic interest rates.* Means a more realistic term structure (discount function)[8] thus better fitting of the models to actual data. Longstaff and Schwartz (1995) develops this type of model using the Vasicek term structure model for interest rate risk where the firm value and the interest rate follow correlated Brownian motions.

[8]In the examples considered here, the discount function is on the form $B(t) = e^{-rt}$ where r is a constant. It is well known that this is not a realistic term structure.

- *Stochastic recovery rates.* These are likely correlated between assets; in recessions when the number of defaults are up, recovery rates can also be expected to be lower.

- *Liquidation process models*: Finally, rather than assuming that liquidation takes place at the same time τ that asset values hit the lower threshold, this triggers a negotiation process (dependent on the future asset price path) that can either end in the firm being liquidated or it can continue its operations. This is similar to debt negotiation under Chapter 11 under the US bankruptcy code. An example of this type of model with dynamic capital structure is in Goldstein et al. (2001).

Finally, it is worthwhile to explicitly some of the limitations behind the models as presented in chapters 2 and 3. The focus of my presentation is on credit risk modeling. Issues of taxation, liquidity (market) risk, counter-party risk and so forth are therefore ignored. Ignoring liquidity risk for bonds may be problematic in estimating default probabilities if the price difference between risk free and risky bonds are assumed to be purely a credit risk premium.

3.6 Calibrating Structural Models

The structural models introduced so far use firm values as the primitive state variables. As this is not directly a traded asset, we need a method for estimating the parameters of the equation $dA_t/A_t = \mu dt + \sigma dW_t$. The simplest approach when equity E_t is a traded asset is to assume it is a function of the asset value: $E_t = f(t, A_t)$ with dynamics given by the PDE:

$$dE_t/E_t = rdt + \sigma dW_t \tag{3.6.1}$$

We can then apply Ito's formula:

$$dE_t = \left(\frac{\partial f}{\partial t} + rA_t\frac{\partial f}{\partial A_t} + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}A_t\sigma^2\right)dt + \frac{\partial f}{\partial A_t}A_t\sigma dW_t \tag{3.6.2}$$

Equating the Brownian motion terms from the two preceding equations we obtain:

$$\sigma_E E_t = \frac{\partial f}{\partial A_t}A_t\sigma_A = \Phi(d_1)A_t\sigma_A \tag{3.6.3}$$

Matching the observed $\hat{E}_t$ with the theoretical equity values $f(A_t, t)$, a system of two equations in A_t and σ_A is obtained.

An even simpler approach to this problem is to compile times series data on traded equity and debt values as well as estimates of non-traded securities such as bank loans.

Now the problem can be solved using standard time series techniques such as the GARCH model.

Chapter 4

Applications and Examples

The two classes of models developed so far have a wide array of applications in finance. Here I will concentrate on two of them, valuation of credit derivatives and risk management. This section explains the valuation of three very common credit derivative products and provides explicit mathematical formulas that can be used in valuing these along with the numerical models explained earlier.

As my primary approach is numerical, many valuation expressions are given in terms of expected values of stopping times or their associated indicator variables. These expressions convenient when translating the methods to some programming language for implementation using Monte Carlo methods. Under some assumptions these values can be stated analytically by evaluating the proper integrals, but that is beyond the scope of this thesis.

This chapter is primarily intended to demonstrate how the models presented in the chapters 2 and 3 coupled with the numerical methods discussed in Appendix B can be applied to valuing some "vanilla" credit derivatives. The valuation formulas are simplified versions of similar formulas found for example in Hull (2006), Hull and White (2005), O'Kane and Turnbull (2003), Elizalde (2005). The formulas used in this thesis ignore certain issues related to day counting, accrued premia and so forth. These assumptions are not conflict with the pricing theory; it is the cash flow structure of the instruments in the examples that are simplified.

All expectations and probabilities in this chapter are with respect to the *risk neutral* measure. To provide a transparent introduction to these instruments, we first consider three instruments with a very simple cash flow structure, paying a unit cash flow at default, thus abstracting from issues such as recovery rates, coupon payments, etc. As we will see later on, the conclusions drawn in the instruments also hold up when we introduce a slightly more advanced cash flow structure.

4.1 Binary Credit Derivatives

A binary credit default swap (CDS) is a relatively simple instrument to value as the payoff is not a variable amount, but a constant cash flow that is paid if the underlying asset (called the reference entity) defaults. This means that we only need to worry about issues of probability, and the cash flow given default can without loss of generality be set to 1. We consider two variations on this instrument; first a binary CDS written on a single credit and secondly a basket binary CDS.

4.1.1 Single Credit Binary CDS

A single credit binary CDS with maturity T is an instrument paying $1 \cdot e^{-r\tau}$ if time of default $\tau < T$. We are therefore interested in the time to default distribution for the reference entity. In a reduced form model, this is exponentially distributed with hazard rate $\lambda(t)$ and pdf $f(t) = \lambda(t)e^{-\lambda(t)}$, so we can state the CDS value as:

$$\begin{aligned} V_{CDS} &= \int_0^T e^{-rt} f(t) dt \\ &= \int_0^T \lambda(t) e^{-(r+\lambda(t))t} dt \end{aligned}$$

In the case of constant λ we get the simpler expression:

$$\begin{aligned} V_{CDS} &= \int_0^T \lambda e^{-(r+\lambda)t} dt \\ &= \frac{-\lambda}{r+\lambda} \left[e^{-(r+\lambda)t} \right]_0^T \\ &= \frac{\lambda}{r+\lambda} \left(1 - e^{-(r+\lambda)T} \right) \end{aligned}$$

If we consider a one year contract with risk free rate $r = .05$ and $\lambda = .10$, the price $V_{CDS} = 0.0928613$. This can provide a benchmark for the simulation program in Appendix C.2. With 50,000 simulations using this program, we get $V_{CDS} = 0.09276$ with a standard error $1.27e - 04$, which seems reasonable; the answer is right to the fourth digit when rounded off. Figure 4.1 shows a sample path of convergence for this problem for the number of simulations varying from 1,000 to 200,000 with the absolute pricing error[1] along the ordinate axis.

[1] Analytical price minus simulated price.

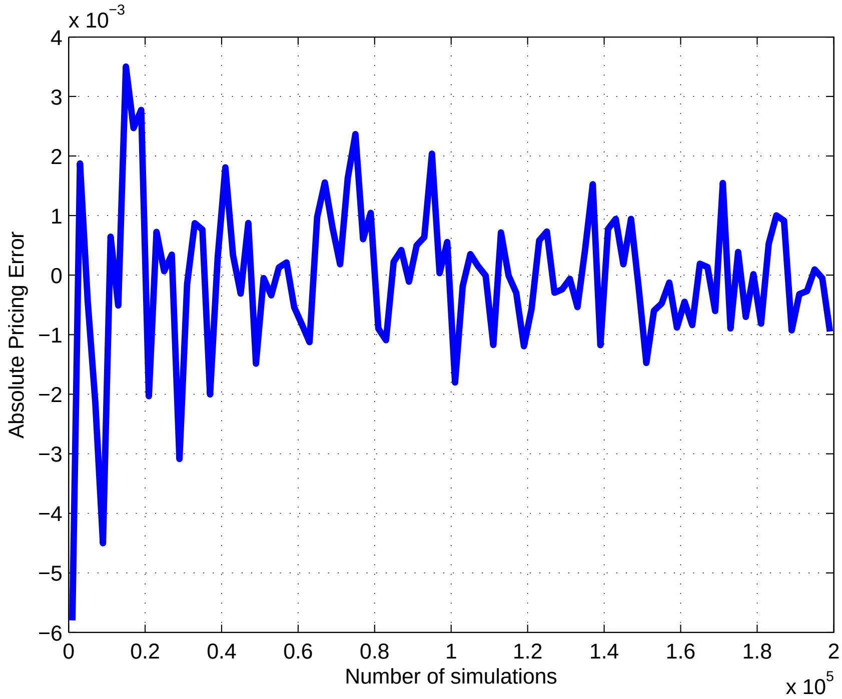


Figure 4.1: The convergence of the reduced form simulation model for a simple single asset pricing problem.

4.1.2 Counterparty Risk

Consider the same CDS, now relaxing the assumption that the issuing entity is risk free. We now wish to account for this counter party risk. There are three possible outcomes of the transaction:

- Neither entities default - zero cash flow from the CDS.
- The reference entity defaults - unit cash flow from the CDS.
- The issuing entity defaults prior to the reference entity - zero cash flow from the CDS.

Figure 4.2 illustrates the relationship between the one year CDS price and the counter-party hazard rate and the correlation between the counter-party and the buyer of the protection contract. As we would expect, with zero correlation and counter-party hazard rate, the CDS price equals the price from the last example. Furthermore, as the correlation increases, the impact of the hazard rate on the contract value increases too. All of these results seem quite intuitive.

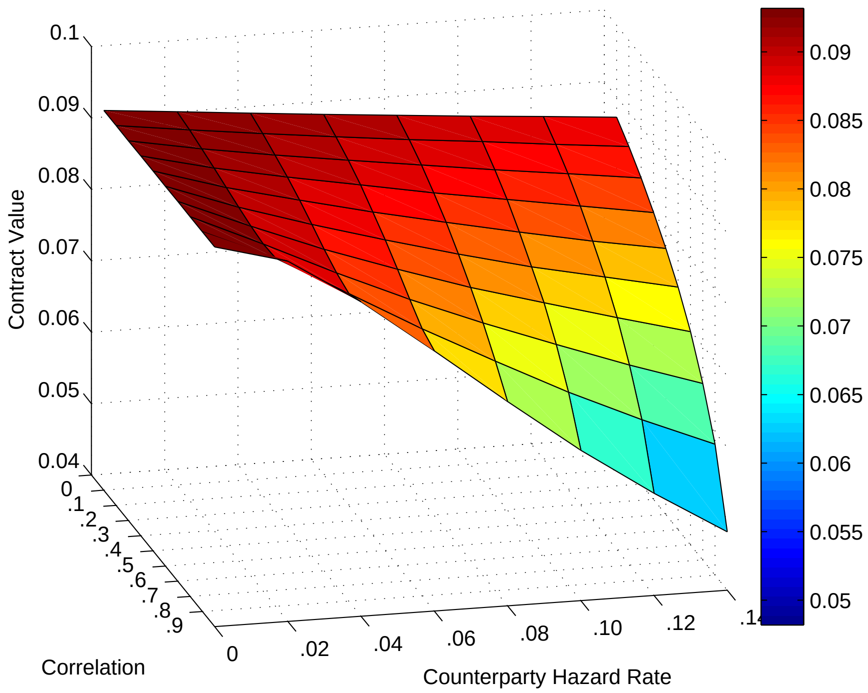


Figure 4.2: Value of a one year CDS with counterparty risk.

4.1.3 Binary Basket CDS

In the previous section, we demonstrated a simulation based reduced form model and did an informal comparison with the analytic result for a very simple pricing problem. Next we consider a more complex problem with a binary basket[2] CDS that pays a unit cash flow at the time of default for the first five entities in a portfolio of 100 assets. Figure 4.3 contains the copula simulation section of the MATLAB implementation of this model from Appendix C. It will provide an illustration of a simulation algorithm using a Gaussian copula.

```
vMat=randn(noAssets,noSims);
vMat=Rho*chol(Rho)';
uMat=normcdf(vMat);
defTimes=-log(uMat)/lambda;
```

Figure 4.3: Code from Copula Example.

The first line generates a matrix `vMat` of independent $\Phi(0, 1)$ random variates whose dimensions are the number of assets in the portfolios times the number of simulations

[2]This is similar to the first n of N to default securities considered later on in this chapter.

of each portfolio. Then the matrix is multiplied by the lower Cholesky matrix (see Appendix B) of the correlation matrix `Rho` to create a correlated set of random $\Phi(0,1)$ matrix. The last two lines apply the normal cdf to generate the copula `uMat` before the inverse exponential cdf is applied to generate the times to default `defTimes` that are now exponentially distributed according to the correlation matrix `Rho`. This is the central part of the simulation algorithm for the Gaussian copula model which is the same for valuing any instrument in a Gaussian copula reduced form model. The only parts of the program that need to be modified to value different instruments are more of an cash flow "accounting" nature.

Illustration I - Gaussian Copula

To illustrate the technique, we consider two binary basket CDS's on a portfolio of 100 assets; one that pays a unit cash flow for each of first five reference entities to default if this occurs prior to maturity in one year and one that provides credit protection for the first 20 assets. As before, $\lambda = .10$ and the risk free rate $r = .05$.

Figure 4.4 illustrates the relationship between price asset correlation. Note that prices here are quoted in absolute values; ie. with zero correlation, the price of the first five basket is about 4.8804. This refers to the expected NPV of cash flow to the buyer of protection is 4.8804.

The results are as expected; the higher the correlation, the lower the value of the basket. If defaults are completely independent, we would expect 10 defaults on the average over the course of one year with $\lambda = .10$ and portfolio size 100. If correlation increase, we expect to see defaults that are less scattered, and more clustering will occur. In the extreme case, if assets are perfectly correlated, we will on average see all assets default during the first year in 1 out of 10 simulations.

Illustration II - Comparing Copula Models

To see the effect of the choice of copula model, we can compare the Gaussian to the t-copula for various degrees of freedom holding the other parameters of the model constant.

Figure 4.5 shows the computed basket values for 20,000 simulations in the t-copula model for varying number of degrees of freedom. In comparison, the normal copula gives the value 4.3740 for the same basket. While the numbers are somewhat imprecise due to the low number of simulations (the standard error is around .010), the general picture is clear. A model with a low number of degrees of freedom gives a higher cost of protection (contract value) for an otherwise similar portfolio. When the number of degrees of freedom goes to infinity the t-distribution, as is well known from mathematical

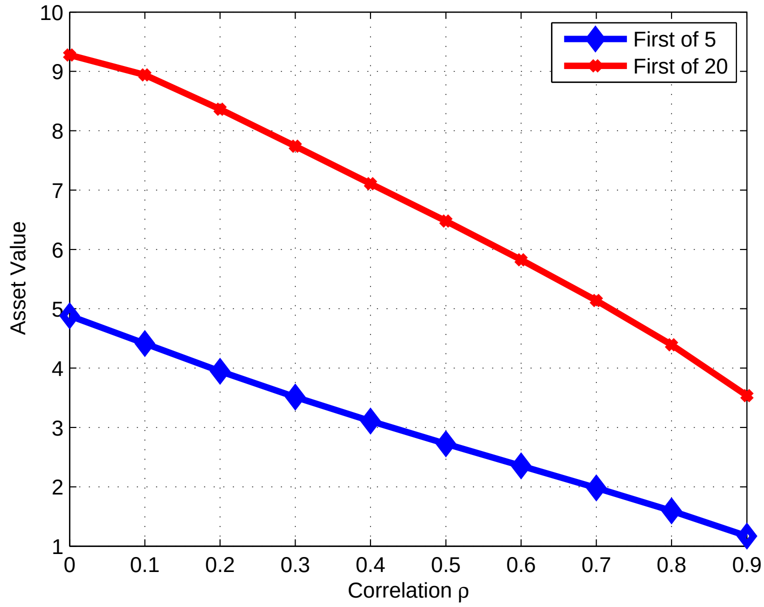


Figure 4.4: Prices of a first of 5 and first of 20 to default binary basket CDS as a function of asset correlation.

Degrees of Freedom	1	2	3	4	5	6	7	8	9
Basket Value	2.11	2.87	3.27	3.51	3.67	3.78	3.85	3.92	3.96

Figure 4.5: Basket values for different degrees of freedom in a copula model.

statistics, converges to the standard normal distribution.

Now, how are we to interpret these results? In fact they are quite similar to the analysis of the impact correlation in the previous example. In this application we can view the number of degrees of freedom as a measure of how "scattered" the data is. A high number of degrees of freedom means the data is more concentrated around the mean such as is the case for the normal distribution. For a low df. number, the tails of the distribution are fatter, and we see more clustering.

The value of a small basket, which is similar to an CDO equity tranche (cf. Section 4.4), is positively related to the default correlation of the underlying assets. More clustering (higher correlations) increases the probability of no defaults as well. In the case of independence, the average long-run portfolio default rate is the accumulated hazard rate on the time interval considered.

4.1.4 Binary Collateralized Debt Obligation

The concept of a CDO is presented in some detail in Section 4.4. Here we define a binary CDO tranche on a portfolio of 100 reference entities as a security that pays 1 unit cash flow for each default occurring from the n'th asset to default to the N'th asset to default. We use the senior tranche as an illustration. This tranche pays a unit cash flow for each default from the 30th asset to the last asset to default. We use precisely the same script as before and analyze the impacts of asset correlation and hazard rate on the price of a one year contract.

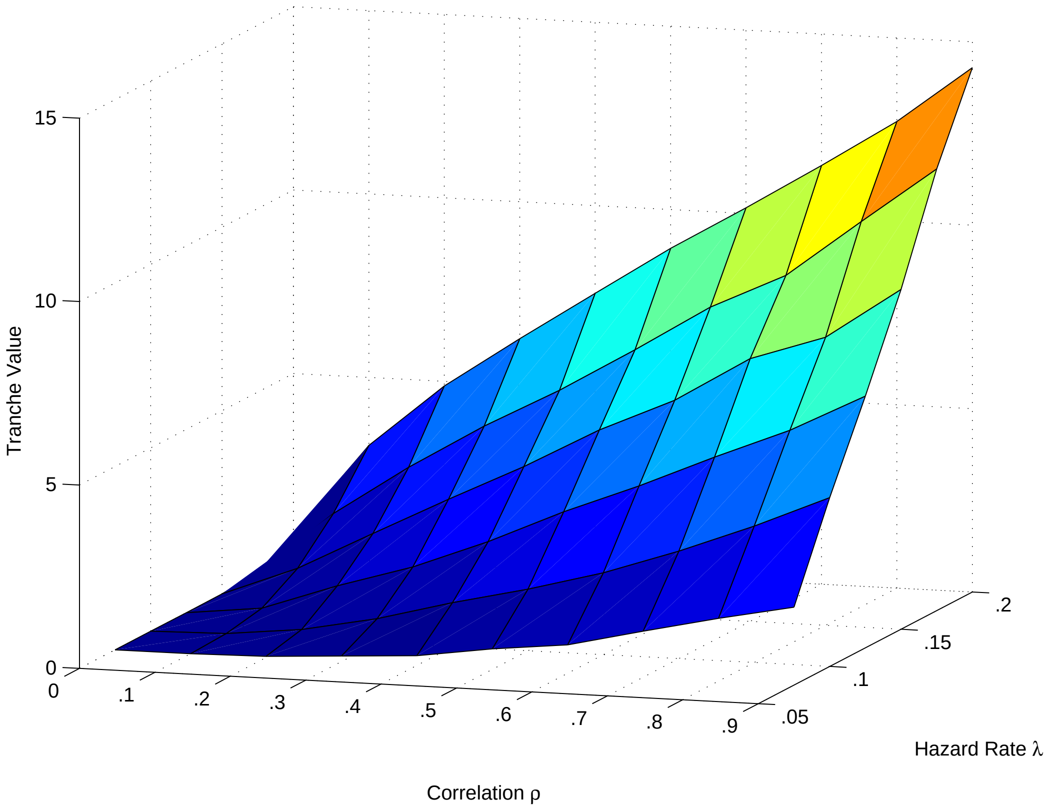


Figure 4.6: Senior CDO tranche values agains ρ and λ.

Figure 4.6 and the corresponding table in Figure 4.7 shows how the senior CDO tranche value varies with the correlation coefficient ρ and the hazard rate λ. By the term *tranche value*, I mean the cost of protection for the particular tranche. A higher tranche value would therefore mean a lower value of the underlying assets, as these would be more exposed to loss. This is a somewhat unconventional price quotation; the standard being in basis points per quarter. The advantage of this quotation is that the number is an estimate of the expected loss in a particular tranche which can be used directly for instance in risk management and is simpler to verify.

As we would expect, a higher default probability increases the cost of the contract thus increasing the value of the protection contract. We also see that the cost of protection for a senior tranche increases when asset correlation in the reference portfolio is very high as this increases the probability of large default clusters where the default rate is sufficiently high to affect the senior tranche.

An important point here is how the sensitivity to asset correlation varies with the hazard rate. The cost of protection as a function of correlation increases at a steeper rate when the hazard rate is high compared to when it is low. This seems intuitively reasonable; for a portfolio of high credit quality assets, a default cluster occurs very rarely, even with high correlations. For a portfolio of junk bonds, on the other hand, a large default cluster is very likely to occur when correlations are high.

λ,ρ	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
.05	0	0.000	0.033	0.152	0.266	0.561	0.776	1.253	1.710	2.123
.10	0	0.043	0.261	0.662	1.204	1.677	2.230	2.928	3.714	4.599
.15	0	0.230	0.943	1.565	2.361	3.279	4.091	4.985	5.822	6.858
.20	0.0052	0.809	1.862	2.902	3.897	5.004	5.919	7.160	7.850	9.261
.25	0.0897	1.793	3.162	4.389	5.473	6.684	7.951	8.921	10.497	12.038
.30	0.6172	3.145	4.870	6.267	7.609	8.941	10.151	11.416	12.721	14.296

Figure 4.7: Binary CDO prices for various hazard rates and correlation coefficients.

In the sections to come we will expand upon these examples by introducing slightly more realistic assumptions on cash flow structures and recovery values. We will see, however, that many of the conclusions drawn in the simpler examples we have seen so far still hold.

4.2 Credit Default Swaps

A credit default swap is a contract between two parties, where where one party pays a fixed leg in return for credit protection (also known as a floating leg) against the default (or more generally, the occurrence of *credits events*) of a third party reference entity, corporate or sovereign.

The fixed leg is usually paid until the occurrence of a credit event or the end of the CDS's life. For the CDS to break even, the expected NPV of the two cash flows must equate.

$$\text{Protection Buyer} \xleftarrow{\;L(\tau) \text{ if } \tau < T\;} \qquad \xrightarrow[\text{Fixed premium } \delta]{} \text{Protection Seller}$$

Figure 4.8: Cash flow structure of a CDS.

Let $L(\tau)$ denote the amount lost given default at time τ, $p_s(t_i)$ the risk neutral probability of no default occurring up to time t_i, r the assumed constant interest rate, δ the fixed leg payments and $[t_0, t_1, ...t_N]$ the premium payment dates, the break-even equation takes the following form:

$$\sum_{i=1}^{N} \delta p_s(t_i) e^{-rt_i} = \mathbb{E}\left[L(\tau)e^{-r\tau}\right] \tag{4.2.1}$$

Defining $P_{floating}$ as the value of the protection payments (the right hand side of the above equation), the fixed leg premium is given by:

$$\delta = \frac{P_{floating}}{\sum_{i=1}^{N} p_s(t_i)e^{-rt_i}} \tag{4.2.2}$$

The procedure for valuation is therefore to generate N independent default times according to the model of choice (either reduced or structural) and approximate the expected time to default and expected floating payment by the averages.

For single CDS instruments the whole concept of correlation is irrelevant which greatly simplifies the valuation algorithm[3].

4.2.1 Numerical Example - Valuing a CDS Using a Reduced Form Model

Consider a CDS on a bond with a 6% coupon rate paid semi-annually, face value 100 and maturity 3 years. Let the risk free rate $r = .05$ and $\lambda = .10$. The loss given default is assumed to be constant equal to 40.

Assuming that fixed leg premia are paid semi-annually, this valuation model gives a protection leg value of approximately 10.35 and a cds spread of 0.0445. A MATLAB script implementing this example is included in Appendix C.

4.3 Basket Credit Default Swaps

Next consider a contract where an agent holds a portfolio of K assets wishes to buy a partial protection against credit risk, that is the first $\kappa < K$ assets to default[4].

By considering two extreme cases of asset interdependence, we can note some properties about the pricing of these assets. First, if assets are perfectly correlated, either all or no assets default, with probabilities equal to the individual default and survival probability, respectively. This means that the insurance would be the same for each firm, ie. the contract price a linear function of the number n of firms protected. With perfect independence between defaults, insurance against the first default will be more expensive than the second and so forth, as the probability of m defaults is smaller than that of

[3]This is actually not quite the case. This example ignores the problem of counter-party risk, ie. the risk of the protection seller defaulting. To properly account for this, we need to know the correlation between the reference entity and the counter-party

[4]This is also sometimes referred to as an "first n of N to default CDS".

$m + 1$ defaults.

Assume that the loss given default $L(\tau)$ is equal across firms, $p_s(t_i)$ the cumulative survival probability up to date t_i and $[t_0, t_1, ..., t_N]$ are the premium payment dates as before. The pricing equation by equating the expected value of the fixed leg (representing the fair value of the protection contract) and the floating leg (under the risk neutral measure):

$$\sum_{i=1}^{N} \delta e^{-rt_i} p_s(t_i) = \sum_{k=1}^{K} \mathbb{E}\left[L(\tau_k) e^{-r\tau_k}\right] \tag{4.3.1}$$

Denote the right hand side of the previous equation as $P_{floating}$, the basket premium δ is:

$$\delta = \frac{P_{floating}}{\sum_{i=1}^{N} e^{-rt_i} p_s(t_i)} \tag{4.3.2}$$

As we are now interested in credit events affecting a set of assets, we need to account for default correlation as well. As we shall see in the next section, this type of contract is actually similar to a CDO equity tranche.

4.4 Collateralized Debt Obligations

Collateralized debt obligations function in a manner similar to that of the basket instruments discussed above. The holder of a portfolio seeks to buy protection against losses due to default. But rather than buying protection for the first n assets, protection is now bought for all assets in the portfolio. However, for several reasons[5], the portfolio is sold in different *tranches* or slices, rather than as a whole. The buyer of a tranche acts as a seller of protection, and receives a fixed cash flow for protection against losses within a particular percentage range of the portfolio face value.

Tranche	**Lower Bound**	**Upper Bound**
Senior	.15	1.0
Mezzanine 2	.12	.15
Mezzanine 1	.08	.12
Junior	.03	.08
Equity	0	.03

Figure 4.9: A simple CDO structure.

As an illustration, consider the equity tranche from the CDO in Figure 4.9 that

[5] A discussion on the rationale for tranching is found in Duffie (2007).

provides protection against the first 3% of losses in the portfolio. If the total portfolio loss is 1.5%, this amounts to a loss of 50% in the equity tranche, while the more senior tranches suffer no loss.

The loss in each tranche i, $L_i(t)$, ie. the floating leg cash flow from the holder of that tranche is a function of the upper and lower *detachment points* of the tranche, denoted K_L^i and K_U^i respectively, and the total losses on the portfolio, $L(t)$, all measured as percentages of total initial value, P.

$$L_i(t) = \begin{cases} 0 & \text{if } L(t) < K_L^i \\ (L(t) - K_L^i)P & \text{if } K_L^i \leq L(t) \leq K_U^i \\ (K_U^i - K_L^i)P & \text{if } L(t) > K_U^i \end{cases} \tag{4.4.1}$$

Let $\delta_i(t_j)$ be the fixed leg premium for tranche i. These are usually paid on a discrete set of dates t_j for $j \in \{1, 2, ..., T\}$. The fixed leg payments for tranche i can then be expressed as:

$$\sum_{j=1}^{T} e^{-rt_j} \delta_i(t_j)$$

The premium δ_i for the i'th tranche is chosen so that fixed leg payments equal the expected loss.

4.4.1 Numerical Example - Merton Model

To illustrate consider a simple example. A CDO is written on a portfolio of 100 risky zero-coupon bonds with a one year horizon. The risk free interest rate is assumed to be constant ($r = .03$). Each asset has a face value 100, and the recovery rate for each asset is constant 40% of face value. Each asset has a default probability of .07, and the asset price volatility, σ is 20%.

Appendix C contains a MATLAB script implementing this example.

Figure 4.9 on page 72 illustrates the expected values of the different tranches for various values of the correlation coefficient under the Merton model with 20,000 simulations as tabulated in Figure 4.10.

It is quite clear that for the given set of parameters (individual default probability of .07 and .20 annual volatility) that the equity tranche benefits from increased default correlation in terms of a decreasing expected loss rate. For the more senior tranches, the effect is opposite and more pronounced for the most senior tranche where expected loss rate is clearly a convex function of ρ. These results are very much as expected and are quite in line with what we see in the simpler binary CDO example. A low correlation means that the loss rate will be close to the individual default probability, so that the

ρ	**Tranche 1**	**Tranche 2**	**Tranche 3**	**Tranche 4**	**Tranche 5**
0	94.714	16.407	0.971	0.000	0.000
.1	39.152	3.003	1.525	0.145	0.00360
.2	23.169	3.084	2.930	0.795	0.0420
.3	16.573	3.690	3.960	1.860	0.184
.4	13.189	4.358	4.480	2.665	0.377
.5	11.041	5.143	5.380	3.730	0.709
.6	9.616	5.571	5.715	4.650	1.110
.7	8.769	6.028	6.235	5.345	1.629

Figure 4.10: Expected losses in CDO tranches under various asset correlations. N=20,000 simulations.

equity tranche will likely suffer some losses. If assets are perfectly correlated, it is an "all or nothing" scenario, which means that in some cases all assets will default thus making the more senior tranches equally exposed.

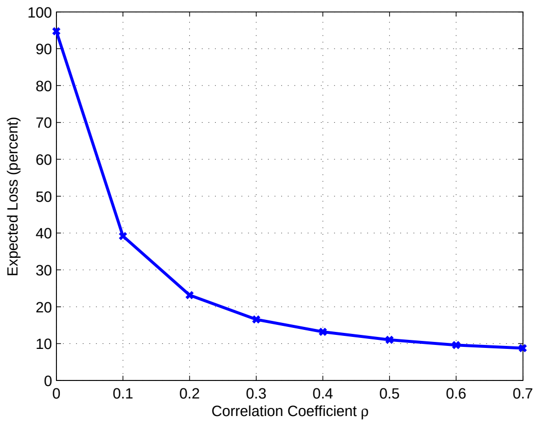
4.4.2 Correlation Trading

Because of the sensitivity of the CDO tranches to the underlying asset correlations, CDO tranches have been used for hedging and betting against correlations. To see how this is done, consider a trader who believes that the correlations implied by the market price of a particular CDO is too low. That would mean that the equity tranche is under-priced (as expected losses in this tranche are overestimated). Similarly, the market price of a more senior tranche is too high (as the price of these tranches depend negatively on asset correlations). An appropriate trading strategy in this situation would then be to take a short position in the senior tranche and a long in the equity tranche.

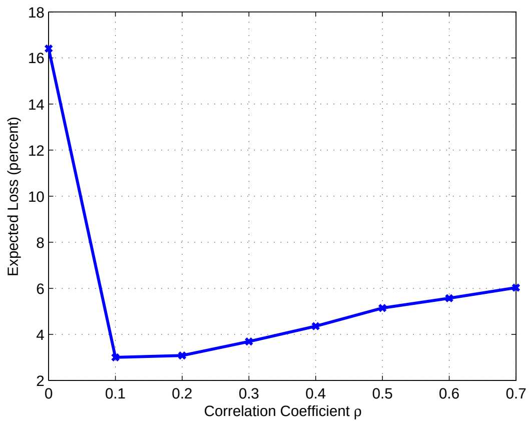
4.4.3 Implications for Risk Management

Clearly, another important aspect of the correlation issue pertains to risk management. As discussed previously, defaults occur in clusters; in terms of the models here, this means a time varying default correlation. If defaults under normal circumstances are largely uncorrelated but correlations increase during economic downturns, crises, and so forth, it can be argued that using average correlations for risk management purposes, at least in a constant correlation model, is negligent of an important risk factor. Intuitively, there is reason to believe that correlations and hazard rates are correlated. Figure 4.6 illustrates the impact of this on a senior CDO tranche; when both correlations and hazard rates increase, the value of a senior tranche plummets.

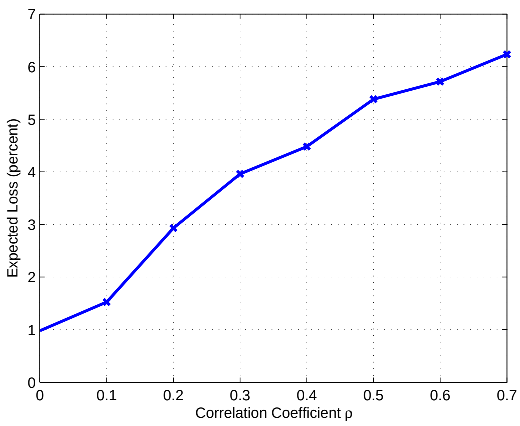
The parallels to the recent economic crisis is not hard to draw; such effects should be more pronounced for low quality credits, such as sub-prime mortgage loans. These are borrowers with low credit quality, with high leverage, low wages and highly exposed to unemployment in the case of a recession. In addition, when mortgages are issued with initial "teaser rates" that are subsequently increased, we can also expect a deterministic, time-dependent increase in hazard rates.



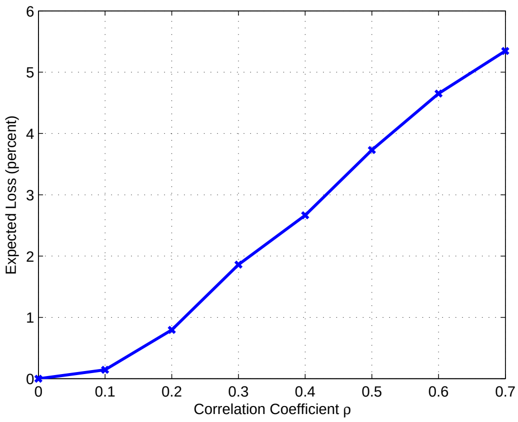
(a) Equity Tranche



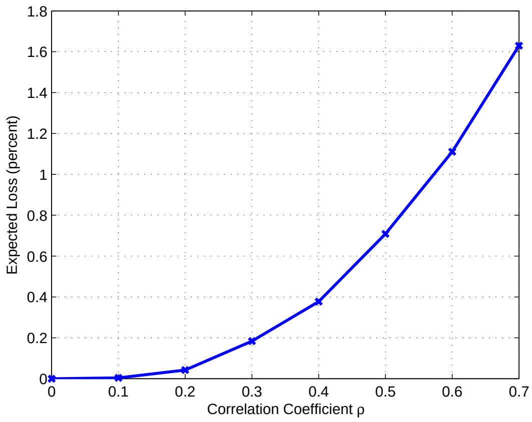
(b) Junior Tranche



(c) 1st Mezzanine Tranche



(d) 2nd Mezzanine Tranche



(e) Senior Tranche

Figure 4.9: Expected losses in a CDO as a function of asset correlation.

Chapter 5

Summary and Conclusion

Events in the financial world the last few years have made the importance of credit risk modeling and research abundantly clear. These models are critical to the valuation process for credit derivatives, an asset class that has seen an explosive growth in the past two decades, and are an intricate part of risk management systems for financial institutions.

This thesis has studied the two key classes of credit risk models that appear in the academic literature and in practice. Emphasis has been put on various methods of modeling correlation, issues of model implementation as well as estimation. Problems related to the specification and measurement of correlation in the various models have also been discussed. It has been argued that default probabilities vary over time, which in a reduced form mode framework corresponds a non-constant term structure of hazard rates. Three standard methods of estimating these term structures have been considered; a simple option theoretical approach using Itô's lemma, using historical data from rating agencies and implying the probabilities from the prices of risky assets (here bonds and credit swaps).

Chapter 4 has explored the properties of many of the models presented in Chapters 2 and 3 by examples of credit derivative valuation. The examples clearly illustrate the importance of hazard rates and correlation for asset values. A key issue in most any type of risk measurement and management is interdependence. In the context of credit risk, this means default clustering or default time correlation. It is important for valuing multi-name credit derivatives such as basket credit default swaps and CDOs, which are, as shown in several of the examples in this thesis, such as the CDO illustration in Section 4.4. Correlation can also be important when accounting for counter-party risk as illustrated in Section 4.1.

Finally, this thesis has also discussed the many possible extensions of the basic models that have been the focus of this presentation. As of now, while there are no industry standard credit risk models, some theoretically very impressive models have been proposed that incorporate a high level of detail, in particular within the structural class. Still, there will always be a certain trade-off between the simplicity, with respect to both implementation and estimation, and the degree of detail to be included. In conclusion, credit risk modeling is likely to remain an important field of research in the years to come, in academia and financial institutions alike, as there are still problems to be solved, both theoretically and empirically.

Appendix A

The Black-Scholes-Merton Framework

A.1 A Model of Uncertainty

This appendix is a brief overview over some of the mathematical assumptions and results fundamental to the models above. It is largely based on Harrison and Pliska (1980), Hull (2006), Shreve (2004) where the topics are treated more thoroughly.

To model the uncertainty faced by economic agents about the future *state* of quantities of interest such as asset prices or interest rate, we consider a *probability space* denoted $(\mathbb{P}, \mathcal{F}_t, \Omega)$ where $\mathcal{F}_t$ is a *filtration* on the sample space Ω. In this thesis, Ω is assumed to be a non-countable or continuous set.

The filtration $\mathcal{F}(t) \equiv \mathcal{F}_t$ is the model of the *set* of information about the market variables that is available to economic agents at time t. An important property to note is that $\mathcal{F}_t \subseteq \mathcal{F}_T \iff t \leq T$, that is that all previously revealed information is available; as we will see later on, the primary importance of this (in the models considered here) is for the statistical purpose of parameter estimation.

To illustrate this somewhat abstract concept, consider the binomial credit risk model of Section 2.1.1. Let's say we do not know the outcome ω in the final period $(T = 3)$. Without any information, we know still know that $\omega \neq \emptyset$ and $\omega \in \Omega$. We denote this by $\mathcal{F}_0 = \{\emptyset, \Omega\}$. After the first period, we know whether or not the firm has defaulted by $T = 1$, so we add an additional piece of information, $\omega(1)$, to the filtration, $\mathcal{F}_1 = \{\emptyset, \Omega\ , \omega(1)\}$. The process can be continued up to the final period preserving the relationship $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3$. This concept can be generalized to continuous probability spaces as in Shreve (2004) and Øksendal (2002).

We define asset or security prices $A_i(t)$ as functions of future uncertain cash flows; thus random variables measurable with respect to $\mathcal{F}(t)$. In a structural model, equity and debt are considered contingent claims on asset prices, that is, the value of the claims are functions of the A_i's.

A.2 The Wiener Process

So far, the discussion has centered on abstract notions of uncertainty and information, hence the next objective is to suggest a concrete model of the how uncertainty affects asset prices, so we introduce the notion of a *stochastic process*, ie. a collection of random variables $X_1, X_2, ..., X_T$, subscripts denoting the time dimension. A common classification of stochastic processes is between discrete and continuous time processes. A continuous process is defined at any point in time on an interval, whereas a discrete process is defined only for a particular set of times. Similarly, a distinction is made between continuous and discrete *range* processes, referring to the set possible values the process can attain.

A *standard one-dimensional Wiener process* denoted $W(t) \equiv W_t$, often referred to as a *Brownian motion*, is a particular stochastic process satisfying:

- $W(0) = 0$.
- W_t is *almost surely* continuous.
- Increments $W_T - W_t \sim \Phi(0, T-t)$ for $t < T$ are normally distributed, accumulating one unit of variance per time unit.
- Over non-overlapping intervals $[t_0, t_1]$ and $[t_2, t_3]$, increments $W(t_1) - W(t_0)$ and $W(t_3) - W(t_2)$ are independent.

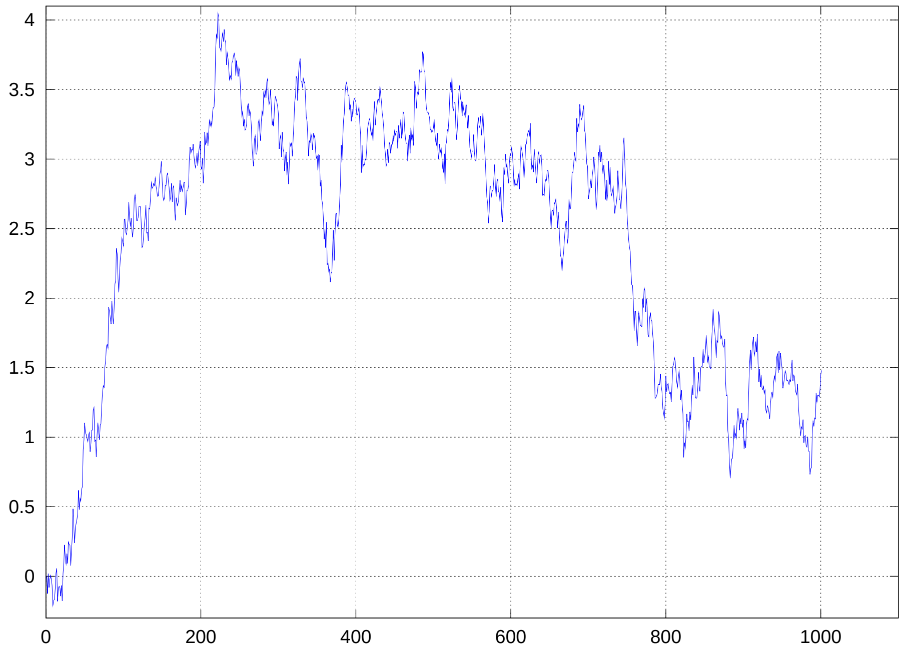


Figure A.1: Sample trajectory of a Brownian motion.

From the latter it is clear that the Wiener process is stationary. It can be shown that a symmetric random walk[1] process will have the Wiener process as its scaling limit; that is the random walk converges to the Wiener process as the time step becomes arbitrarily small.

A.2.1 Asset Price Dynamics

The Wiener process $W(t)$ does not satisfy our requirements for a model of asset price dynamics. Firstly, it can be negative with a strictly positive probability and secondly, we need a richer model of the drift and volatility of prices. The solution is to use a so-called Itô process, $X(t)$ described by the following stochastic differential equation:

$$dX(t) = \sigma(t, X(t))dW(t) + \mu(t, X(t))dt \tag{A.2.1}$$

The above equation states that the change in the quantity X is the sum of a deterministic part μ and a stochastic part containing a differential of the Wiener process. It is important to note that while the notation here is quite similar to that of classical calculus, the mathematical concepts differ as $W(t)$ is nowhere differentiable with respect to t. This also mean that the integral with respect to $W(t)$ is not a Riemann integral and that several standard techniques cannot be applied here. However, using a change-of-variable formula known as Itô's lemma (cf. section A.3.1), we can arrive at most of the results that required for the purposes here.

An important special cases of A.2.1 is the geometric Brownian motion (GBM) process with time constant coefficients given by:

$$\frac{dX(t)}{X(t)} = \sigma dW(t) + \mu dt \tag{A.2.2}$$

Now X has a constant drift and volatility as illustrated in figure **??** for coefficients for $\mu = .05$ and $\sigma \in \{.20, .25, .30, .35, .4\}$. Another important particular case is the mean reverting Ornstein-Uhlenbeck process (sometimes referred to as the Vasicek process in financial applications due to its appearance in the term structure model of the same name):

$$dX(t) = \sigma dW(t) + \alpha(\beta - X(t))dt \tag{A.2.3}$$

This process is mean reverting towards the level β where α is a factor measuring the speed of reversion.

[1] A *symmetric random walk process* is a discrete time, discrete range stochastic process that can either increment or decrement by 1 for each time step. Such processes can be used in simulation models, but usually generating continuous random variables are more efficient.

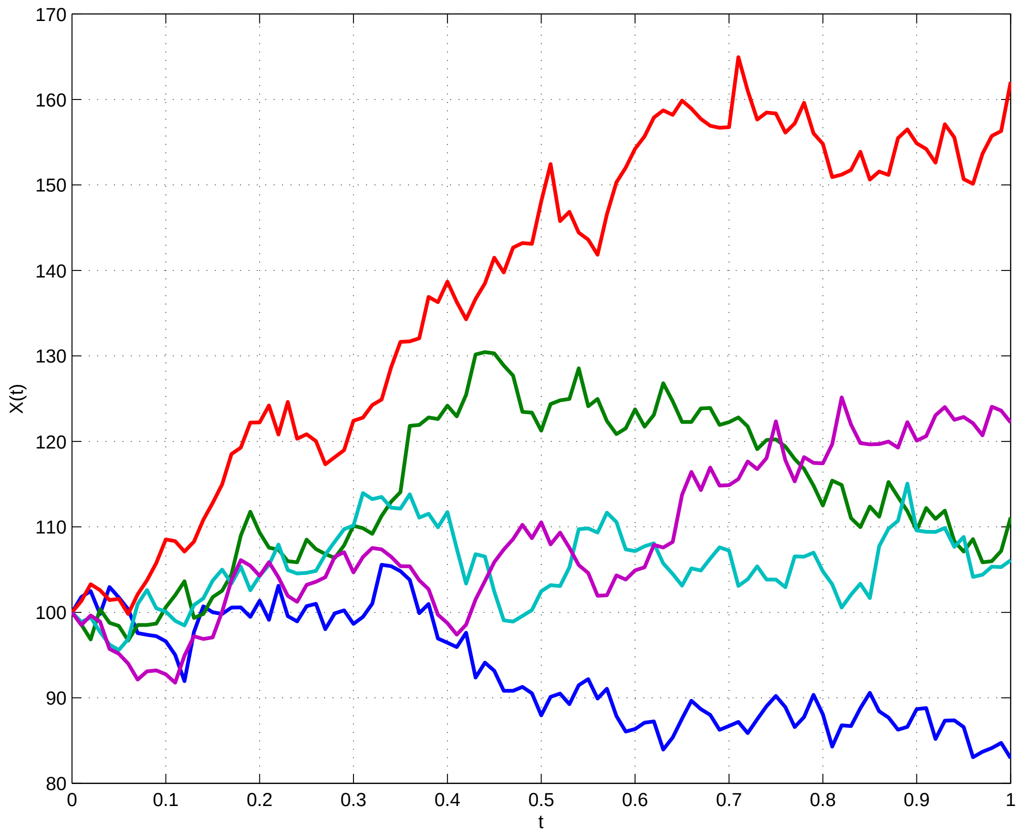


Figure A.2: Geometric Brownian Motion Paths

Before going into some of the standard results that we will need later, I will summarize some of the properties of the process given by equation A.2.1 that are important for its application to problems in finance.

- On a partition $P = \{0 \leq t_0 < t_1 < ... < t_n\}$, the increments $(X_0, X_{t_1} - X_{t_0}, ..., X_{t_n+1} - X_{t_n})$ are simultaneously independent. This means that asset returns are independent of past returns and the current asset price.

- $X(t)$ is continuous for all $t \geq 0$ and strictly positive for all$t \geq 0$ if $x_0 > 0$.

- When $\sigma(t) = \sigma$ and $\mu(t) = \mu$, that is, the parameters are constant over time, $X(t) = x_0 e^{\mu - 1/2\sigma^2 + \sigma W(t)}$, and since $W(t) \sim \phi(0, t)$, $X(t)$ is log-normally distributed; this is a simple exercise to show using Itô's lemma.

- Taking the expected value of this we find that $\mathbb{E}[X(t)|\mathcal{F}(0)] = x_0 e^{\mu - 1/2\sigma^2}$.

- Asset prices have the Markov property; which loosely means that the expectation when conditioning on the entire past history of prices equals the expectation when conditioning on the current price (ie. all relevant information is reflected in the current asset price).[2]

The above properties are not in conflict with financial theory and the efficient market hypothesis, and is therefore a coherent framework precluding arbitrage. There is a correspondence between the concept of markets that are efficient with respect to the information available at any point in time, and the randomness in the model. When all available information about causal factors relevant to the pricing of the assets in question, only future events can affect prices - and these of course, will appear random unless agents have some kind of foresight.

This is not to say that it replicates the observed (dynamic) behavior of asset prices. For instance, the assumption of normally distributed returns has long been criticized, as actual asset returns, in particular for longer time series, exhibit both excess kurtosis and skewness significantly different from that of the normal distribution. In "normal" circumstances, however, the normality assumption seems quite appropriate. By introducing "jumps", stochastic volatility and so forth, it is a simple matter to circumvent these problems[3].

A.3 Some Key Results and Assumptions

A.3.1 Itô's Lemma

Itô's Lemma is, a fundamental tool for studying functions of Wiener processes which is how derivatives are modeled in this framework.

Lemma A.3.1. ***(Itô)*** *Let $X(t)$ denote a Itô process, and $f(X(t),t)$ be a C_2 function[4] of X and C_1 of time t, then (letting subscripts denote partial derivatives) we have:*

$$\begin{aligned} df &= f_t dt + f_X dX + 1/2 f_{XX}\sigma^2 dt \\ &= \left(f_t + \mu f_X + \frac{\sigma^2}{2} f_{XX}\right) dt + f_X \sigma dW \end{aligned} \tag{A.3.1}$$

As simple illustrations of the above equation we can show that returns are normally distributed when prices follow a GBM. Letting $dS_t = \mu S_t dt + \sigma S_t dW_t$ be the equation for the asset price process, and $f(S) = \ln(S)$ so that $f_X = 1/X$ and $f_{XX} = -1/X^2$, the by Itô's lemma: $df = (\mu - 1/2\sigma^2)dt + \sigma^2 dW_t$. f(S) is here the instanvtaneous rate of return process on an infinitesimal time interval.

[2] In the notation established earlier: $\mathbb{E}[X(t)|\mathcal{F}(0)] = \mathbb{E}[X(t)|x_0]$.

[3] This applies in particular to simulation based models, where such modifications are much simpler than analytical models.

[4] A function is C_i if it has continuous i'th derivatives.

A.3.2 Arbitrage Free Pricing

The final concept explored in this chapter is the concept of *arbitrage*. We begin by the definition of a self-financing trading strategy:

Definition A.3.2. ***Self-financing strategy***
Let θ be (the cash flow from) a trading strategy, ie. some combination of securities:

$$\theta(t) = \omega_1(t)A_1(t) + \omega_2(t)A_2(t) + ... + \omega_n(t)A_n(t)$$

where ω_i denotes the number or portfolio weight of asset i. A trading strategy ω is said to be self-financing *if*

$$d\theta(t) = \omega_1(t)dA_1(t) + \omega_2(t)dA_2(t) + ... + \omega_n(t)dA_n(t)$$

Basically, this means that the only thing that can change over time is the values A_i of the assets and the allocation of wealth ω_i between assets. No cash is added to the portfolio or taken out from it.

Definition A.3.3. ***Arbitrage***
Let θ be a self financing trading strategy. θ is said to be an arbitrage *if $\theta(0) = 0$ and $\mathbb{P}(NPV(\theta(T)) > 0) = 1$ for $t \geq 0$.*

The above definition of an arbitrage is a trading strategy that has zero cost initially and is set up so that it yields a risk free, positive cash flow at some future time.

Under the assumption that agents can take any position in the set of traded assets, a model would be contradiction if it would allow for an arbitrage; if such a trading strategy would exist, it would be possible to take a position so as to obtain an infinite cash flow. This would clearly lead to a incoherent pricing framework. Therefore we wish to specify a market model that precludes arbitrage. To summarize, we have the following assumptions that underly the Black-Scholes-Merton model:

- Existence of a risk free asset.
- Trading, both of the underlying asset and the risk free, takes place in continuous time, ie. the asset price $A(t)$ can at any time be exchange for the same amount of money.
- Investors can take any position in any traded asset.
- Assets are perfectly divisible.
- Absence of arbitrage.

An implication of the above assumptions is that any derivative instrument on any traded asset can be hedged by constructing a portfolio of the primitive assets, whose price must equal the price of the derivative. A market in which any future cash flow can thusly be hedged is termed a *complete* market. Under all but certain technical conditions, this can be shown to be equivalent to the existence of a unique *risk neutral* probability measure.

Under this probability measure, expected cash flows are be valued by discounting at the risk free rate (rather than at a risk-adjusted rate). This key result is known as the *fundamental theorem of arbitrage-free pricing*. It follows from it that there exists a state price or a single price for an Arrow-Debreu claim[5] that can be used for discounting cash flows for any possible event or state future of the world. This notion plays a key role in reduced form credit risk models used in corporate finance (see for instance Leland (1994)).

The equivalent risk neutral measure $\widetilde{\mathbb{P}}$ is characterized by the following properties:

- Letting $\mathbb{P}$ denote the *actual* or *physical* probability measure, $\widetilde{\mathbb{P}}(\omega) = 0 \leftrightarrow \mathbb{P}(\omega) = 0$. This is the equivalence part; the two measures agree on which events have zero probability.

- The present value $g(t)$ of a claim to cash flow $g(T)$ at a future date T is given by the product of the *risk free* discount factor (or zero coupon bond price), $P(t, T)$ and the expected cash flow under $\widetilde{\mathbb{P}}$:

$$g(t) = P(t, T)\widetilde{\mathbb{E}}(g(T)) \tag{A.3.2}$$

  This key concept is also used extensively in the reduced form models developed later on as it does not require any particular assumptions on asset dynamics; it will also hold for lattice (bi- and multinomial) discrete time models as well.

- The above can also be stated in terms of the *pricing kernel of the economy* or *stochastic discount factor* which is technically the Radon-Nikodym derivative of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$:

For a complete discussion of the topics of this appendix, the reader is referred to the classic article on the subject by Harrison and Pliska (1981) or textbooks by Duffie (2001), Shreve (2004) or Hull(2006).

[5] An Arrow-Debreu state contingent claim is a financial instrument that pays a unit cash flow given the realization of a particular future state of the world. This is a theoretical concept that is comparable, but not equivalent to a digit option.

Appendix B

Monte Carlo Simulation

This section will briefly describe the numerical methods that are used to implement an extended version of the structural model developed earlier to allow features such as for finite maturity debt, more sophisticated structures of correlation and interest rate risk.

A central result in stochastic calculus, Girsanov's theorem, states that a stochastic differential equation have an equivalent representation as a partial differential equation (PDE). This is an important technique in financial theory; it is for instance the technique used by Black and Scholes to derive their option pricing formula. Furthermore it is often used for closed form solutions to structural models.

There are however certain problems with this approach. Both deriving and solving these equations is quite demanding. With respect to the solutions part, there are a handful of numerical methods available, but these are often difficult to implement, in particular when dealing with problems of higher dimensionality such as when dealing with multi-name credit derivatives. A much simpler and more intuitive approach for such problems is to use Monte Carlo simulation.

B.1 The Basic Concept

Monte Carlo simulation is a technique for approximating the solution y to a problem that can be stated on the form $y = \mathbb{E}[X]$ where X is some random quantity which means that the solution can be reached using artificial sampling experiments.

So while a stock option can (usually) be priced faster by solving the BSM PDE with the appropriate boundary conditions, this method provides little information about the distribution of returns which is of key interest in portfolio and risk management problems.

To illustrate the concept, consider a European call option $C(S_t, t, K) = \max(S_t - K, 0)$ on a geometric Brownian motion $S_t = S_0 \exp[(\mu - 1/2\sigma^2)t + \sigma W_t]$ with $W_t = \sqrt{t} Z_{0,1}$. The approach is then to simulate a vector of standard normal variates, compute the price

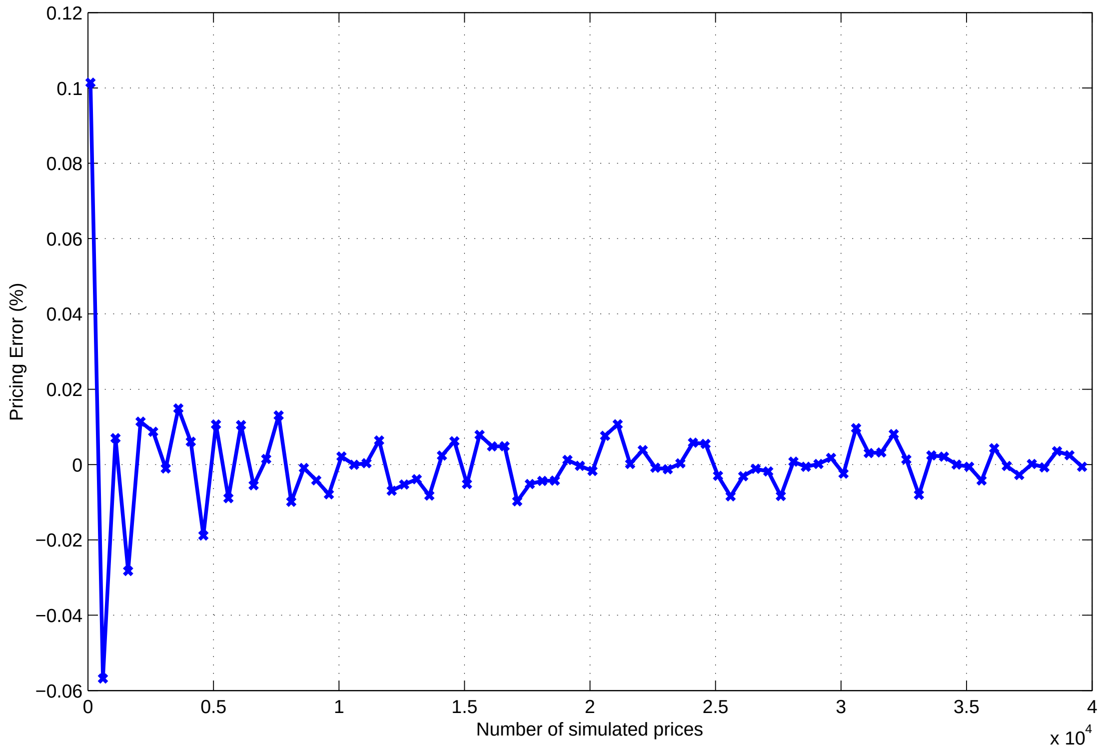


Figure B.1: Pricing error (relative to Black Scholes price) in a naive Monte Carlo method.

S_t^i for each realization i according to the above formula, then the option prices C^i for each realization. The estimate of the option price $V(C)$ is then given as the arithmetic average over the set of option prices:

$$V(C) = \frac{1}{n}\sum_{i=1}^{n} C_i \simeq \mathbb{E}[C] \tag{B.1.1}$$

B.1.1 Error Bounds and Rate of Convergence for Monte Carlo Algorithms

Monte Carlo algorithms differ in an important way from other simulation algoritms. As Monte Carlo methods rely on random sampling, it is not possible to give precise the error bounds as a function of the number of operations performed. Instead we have to rely on probabilistic statements about the error bounds based on what we know about the sampling distribution. Alternatively, if we are considering a particular implementation, it is of course possible to store the mean value (which is what we are usually interested in) and the deviations from the mean in each simulation for an error estimate.

Using the central limit theorem from probability theory, it is not hard to see that the standard Monte Carlo method has square root convergence, in the sense that in order to reduce the standard deviation of the solution, a quadrupling of the number of simulations is necessary. Square root convergence is considered slow, so usually variance reduction techniques such as sampling antithetic paths or low-discrepancy sequences are employed. For a discussion of such methods, see Brandimarte (2006).

B.1.2 Correlation in Monte Carlo Simulation

One of the attractive features of the Monte Carlo method is that it allows for a large number of ways, ranging from the simple and intuitive to the highly sophisticated, to treat the of covariation between the processes of interest.

The perhaps most obvious way of simulating correlated stochastic processes is to specify the variance-covariance matrix $\mathbf{\Sigma}$ of the assets in the portfolio whose entries σ_{ij} is the covariance between assets i and j, and σ_{ii} is the volatility from the SDE governing the dynamics of asset i. We note that $\mathbf{\Sigma}$ is positive definite, symmetric and diagonally dominant. This is important for an algorithm used later on.

Factor Models

Clearly, any given structure of interdependence between the assets of a portfolio can thusly be specified by a listing of all the σ_{ij}'s, but this is often inconvenient, and one would instead be inclined to explain correlation through a set of *systemic factors*. There are at least two good reasons for this; firstly, the assets may be of such a nature that it is hard to obtain a good estimate of the σ_{ij}, such as may be the case if the portfolio consists of non-traded assets. Secondly, the systemic factors often lend themselves to a meaningful economic interpretation as they can often be identified as interest rates, GDP, and so forth. Of course, this will also mean that the number of parameter estimates can be reduced.

In a general factor model, with Y_i denoting the realizations of the systemic factors, ρ_i, j a constant that gives the exposure of asset j to factor i, and ϵ_j an idiosyncratic factor, the realization of a random variable X_j is on the form:

$$X_j = \rho_1 Y_1 + \rho_2 Y_2 + ... + \rho_n Y_n + \epsilon_j \tag{B.1.2}$$

The next section considers some methods for determining the weights according to the correlations of an asset j to the set of factors $\{Y_i\}$.

Generating Correlated Paths

Correlated random variates cannot be generated directly from the variance-covariance matrix using most standard random number generators. We therefore develop a simple algorithm for solving this.

Let $\mathbf{\Sigma} = [\rho_{i,j}]$ be the $N \times N$ correlation matrix. For simplicity we assume that $\rho_{i,j} = \rho$ if $i \neq j$ and 1 otherwise. The *Cholesky decomposition*[1] of $\mathbf{\Sigma}$ is given by:

$$\mathbf{\Sigma} = U'U$$

Here U is an upper triangular matrix and U' denotes its transpose which is of course lower triangular. Consider a standard normal random vector $\mathbf{Z} \in \mathbb{R}^N$. We can transform this to a $\Phi(\mu, \mathbf{\Sigma})$ random vector X^* where μ denotes the mean vector, by the following procedure

$$X^* = \mu + U'\mathbf{Z} \tag{B.1.3}$$

B.2 Longstaff and Schwartz's Algorithm for American Options

The Longstaff and Schwartz method, sometimes referred to as least-squares Monte Carlo simulation (LSMC), can be thought of as a particular *dynamic programming* approach that simplifies pricing of American derivatives. The problem of valuing American derivatives is recurring in many applications of mathematical finance outside of stock option pricing hereunder structural credit risk models under some assumptions on the default threshold and real options valuation.

Dynamic programming involves the breaking up of a large problem into smaller subproblems for which we have simple solution methods. The key problem in American options valuation is determining the optimal exercise boundary. In the LSMC method, we determine the optimal course of action (exercise vs. continue) backwards (in time) along the set of price paths to determine the option value. In this respect the LSMC approach is similar to the the other numerical methods such as finite differences, and binomial and multinomial (lattice) methods.

What separates the LSMC method is the computation of the continuation value; whereas this is a trivial issue in the binomial model, a continuous time setting requires a more sophisticated approach. This is of course where the least squares part of the

[1]When a matrix A is symmetric positive definite, it can be shown that there is a unique matrix U satisfying $A = U'U$. There are other general decomposition algorithms that hold for matrices that are not positive definite. See Cheney and Kincaid (2007) for a background on the Cholesky and other decomposition methods.

algorithm comes in. When the price paths are simulated under the risk neutral measure $\widetilde{\mathbb{P}}$, we estimate the continuation value at each point in time t as a function of the set of state variables at time $t-1$ using the least squares method.

Illustration

To illustrate the above, consider an option on a single asset S and let

$$\mathbf{S}_{t-1} = [S_1(t-1), ..., S_n(t-1)]$$

be the vector of prices generated for the time $t-1$. Similarly, we denote the option values at time t

$$\mathbf{X}_t = [X_1(t), X_2(t), ...,_n (t)]$$

that may come either from the boundary conditions or the preceding step of this algorithm. Assume f is the vector valued function describing the relationship between option prices and the preceding prices of the underlying asset: $X(t) = f(S(t-1))$. For simplicity, assume is on the form $f(S) = 1 + S + S^2 + ... + S^m$. The following expression results:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} 1 & S_1 & S_1^2 & \cdots & S_1^m \\ 1 & S_2 & S_2^2 & \cdots & S_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & S_n & S_n^2 & \cdots & S_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \tag{B.2.1}$$

Or in shorthand vector notation: $\mathbf{X}_t = \mathbf{S}_{t-1}\mathbf{A}$ where $\mathbf{A}$ is the unknown vector of coefficients for the polynomial function f. If $m > n$ the above system is overdetermined. Due to perturbations in the data, the system is likely to be inconsistent. Hence, instead of trying to the system as stated above, we instead solve the corresponding system of *normal equations* for the least squares problem which always has a unique solution:

$$\mathbf{S}^T\mathbf{S}A = \mathbf{S}^T\mathbf{X}$$

There are routines for this in most software packages. The result is a vector of coefficients $\mathbf{A} = [a_1, ..., a_m]$ that are the coefficients in the function $f(S_t)$ that is the predictor function for the continuation value. So for each simulated price $S_i(t)$, the continuation value is computed as $f(S_i(t)) = a_0 + a_1 S_i(t) + ... + a_n(S_i(t))^m$.

Appendix C

MATLAB Code from Examples

This appendix includes MATLAB scripts implementing some of the models used in the examples. They are included to provide a different statement of the models presented here and are thus prototype models not illustrations of how these models would be implemented in practice. For some background on the implementation of pricing models, see Joshi (2008). Some of the scripts must be split into two separate m-files in the same folder for them to run.

C.1 Defaults Distribution

```
%% start defProb.m
function probs = defProb(N,p,rho)
% computes the probabilities of x defaults
% where x runs from 0 to N in a portfolio of N assets
    probs=zeros(N,1);
    yvals=-100:.01:100;
    condProbs=normcdf((norminv(p)-sqrt(rho)*yvals)/sqrt(1-rho));
    for x=1:N
        bin=nchoosek(N,x);
        integral=bin*trapz(yvals,(condProbs.^x).*((1-condProbs).^(N-x)).*
            normpdf(yvals));
        probs(x)=integral;
    end
end
%% end defProb.m
```

```
%% start portfRates.m
% gives a few plots from the previous function defProb
```

```
% with various rho's
clear all; close all;

format long
hold on;
grid on;
p=.10;
N=50;


rho=0;
plot(defProb(N,p,rho),'LineWidth',2);
rho=.2;
plot(defProb(N,p,rho),'-rx','LineWidth',2);
rho=.5;
plot(defProb(N,p,rho),'-go','LineWidth',2);
rho=.7;
plot(defProb(N,p,rho),'-kv','LineWidth',2);

legend('\rho=0','\rho=.20','\rho=.50','\rho=.70');
xlabel('Number of Defaults')
ylabel('Probability')

%% end portfRates.m
```

C.2 Binary CDS Example

```
%basket cds script
function price=basketCDS(rho)
    %randn('seed',0)
    format long;
    lambda=.1;
    T=1;
    steps=1000;
    r=.05;
    noAssets=100;
    noSims=1;
    basketLimit=20;

    tVec=linspace(0,T,steps);
    discount=exp(-r*tVec);
    Rho = repmat(rho,noAssets,noAssets);
```

```
    for i=1:noAssets
        Rho(i,i) = 1;
    end
    Rho = chol(Rho)';

    vMat=randn(noAssets,noSims);
    vMat=Rho*vMat;
    uMat=normcdf(vMat);
    defTimes=-log(uMat)/lambda;
    cfs=0;
    for i=1:noSims
        thisPath=defTimes(:,i);
        thisPath=sort((thisPath(thisPath<T)));
        thisPath=thisPath(1:min(length(thisPath),basketLimit));
        thisCF=sum(exp(-r*thisPath));
        cfs=cfs+thisCF;
    end
    price=cfs/noSims
end
```

C.3 CDS Example

```
% cds pricing script

T=3;            % maturity
timeSteps=1000;    % possible default dates
dt=T/timeSteps;
tVet=dt:dt:T;
r=.05;            % risk free rate
couponTimes=[.5,1,1.5,2,2.5,3]; % premium payments dates
lambda=.10;        % hazard rate
face=100;
lgd =40;          % loss given default
N=500000;         % number of simulations
defT=ones(N,1)*T;    % time of default in each simulation
                     % (set to maturity if no default occurs)


pd=1-exp(-lambda*tVet);
x=rand(N,1);
```

```
% compute default probabilities and price protection leg
protectionCF=zeros(N,1);
for i=1:N
    if x(i)<pd(end)
        t=find(x(i)<ps, 1 );
        defT(i)=t*dt;
        protectionCF(i)=lgd*exp(-r*t*dt);
    end
end

protection=mean(protectionCF)
ps=exp(-lambda*couponTimes);
discount=exp(-r*couponTimes);

% price fixed leg and get cds spread
fixed=protection/(sum(discount.*ps));
cdsSpread=2*fixed/face
```

C.4 Default Basket

```
% Simple CDO tranche/N-th to default CDS Gaussian copula pricing script.
% Returns floating leg cash flow for the tranche starting at 'attach' and
    ending
% at 'detach'.

randn('seed',0) % reset random variable generator

c=.05;              % coupon rate
r=.05;              % risk free rate
rho=.9;             % asset correlation
N=5;                % Number of firms
NoSims=10000;       % Number of simulated paths
lambda=.1;          % default intensity
attach=.0;          % attachment
detach=.03;         % detachment
T= 5;               % Time to maturity
dt=.5;              % Coupon dates
n=100;              % Notional principal
Rec=60;             % Recovery
lgd=n-Rec;          % Loss given default

tVec=0:dt:T;                        % vector of coupon dates
```

```
discount=exp(-r*tVec(2:end));    % risk neutral discount vector
allCTs=repmat(tVec,N,1);         % matrix of coupon dates

%correlation matrix, cholesky decomposition
rhoMat=rho*ones(N,N)+diag((1-rho*ones(N,1)));
rhoChol=chol(rhoMat)';

% initialize variables to hold 'accumulated' loss for each simulation
totFloat=0;

% simulate correlated default times
nCorr=normcdf(rhoChol*randn(N,NoSims)); % use cdf to create copula
tDef=-log(1-nCorr)/lambda; % use inverse exp. dist cdf to get default times



for i=1:NoSims
    thisPath=tDef(:,i); % get results from simulation i from tDef
    thisMod=repmat(thisPath,1,2*T+1); % matrix to compare
                          % default times to coupon dates
    lossMat=thisMod<allCTs; % binary matrix of default indicators

    % at each coupon date (vectors): %percentage lost
  % the percentage of losses taken by each tranche
  % absolute losses, and remaining tranche notional:
    pctLoss=sum(lgd*lossMat)/(n*N);
    trancheLossPct=max(pctLoss-attach,0)-max(pctLoss-detach,0);
    trancheLossAbs=trancheLossPct*n*N;
    notionalLeft=n*N*(detach-attach-trancheLossPct);

  % temporary calculations
    tempPctloss =trancheLossPct(2:end)-trancheLossPct(1:end-1);
    temp =n*N * tempPctloss;

  % add value for this simulation to total
    totFloat = totFloat + sum(discount.*temp);
end

% compute average
floating=totFloat/NoSims
```

C.5 CDO Example

```
%% start runTest.m
S0=100; pd=.07; mu=.05; sigma=.2; r=.05; T=1; N=100; rho=.7; recovery=.4;

noSims=20000;
trancheLoss0=test(S0,pd,mu,sigma,r,T,N,0,noSims,recovery);
trancheLoss1=test(S0,pd,mu,sigma,r,T,N,.1,noSims,recovery);
trancheLoss2=test(S0,pd,mu,sigma,r,T,N,.2,noSims,recovery);
trancheLoss3=test(S0,pd,mu,sigma,r,T,N,.3,noSims,recovery);
trancheLoss4=test(S0,pd,mu,sigma,r,T,N,.4,noSims,recovery);
trancheLoss5=test(S0,pd,mu,sigma,r,T,N,.5,noSims,recovery);
trancheLoss6=test(S0,pd,mu,sigma,r,T,N,.6,noSims,recovery);
trancheLoss7=test(S0,pd,mu,sigma,r,T,N,.7,noSims,recovery);

tls=[trancheLoss0;trancheLoss1;trancheLoss2;trancheLoss3;trancheLoss4;
    trancheLoss5;trancheLoss6;trancheLoss7]

for i=1:5
  figure(i); grid on;
  plot([0,.1,.2,.3,.4,.5,.6,.7],tls(:,i),'-x','linewidth',2);
  xlabel('Correlation Coefficient \rho');
  ylabel('Expected Loss (percent)');
end
%% end runTest.m
```

```
%% start test.m

function trancheLoss=test(S0,pd,mu,sigma,r,T,N,rho,noSims,recovery);

% recovery= present value of recovered face value

  detachments=[.03, .08, .12, .15, 1]*100; % detachment points (upper)
  k=norminv(pd);
  Y=randn(noSims,1);
  tranches=[3*ones(noSims,1) 8*ones(noSims,1) 12*ones(noSims,1) 15*ones(
      noSims,1) 100*ones(noSims,1)];

  for i=1:noSims
    epsilon=randn(N,1);
    noDefaults=0;
    for j=1:N
      x=sqrt(rho)*Y(i)+(1-sqrt(rho))*epsilon(j);
```

```
        if x<k
  noDefaults=noDefaults+1;
  end
  end % calculate defaults


  %% tranches:
  if noDefaults*(1-recovery) < detachments(1)
    tranches(i,1)=tranches(i,1)-noDefaults*(1-recovery);

  elseif noDefaults*(1-recovery) < detachments(2)
    tranches(i,2)=tranches(i,2)-noDefaults*(1-recovery)+detachments(1);
    tranches(i,1)=0;

  elseif noDefaults*(1-recovery)<detachments(3)
    tranches(i,3)=tranches(i,3)-noDefaults*(1-recovery)+detachments(2);
    tranches(i,1)=0;
    tranches(i,3)=0;

  elseif noDefaults*(1-recovery)<detachments(4)
    tranches(i,3)=tranches(i,3)-noDefaults*(1-recovery)+detachments(3);
    tranches(i,1)=0;
    tranches(i,2)=0;
    tranches(i,3)=0;

  elseif noDefaults*(1-recovery)<detachments(5)
    tranches(i,5)=tranches(i,5)-noDefaults*(1-recovery)+detachments(4);
    tranches(i,1)=0;
    tranches(i,2)=0;
    tranches(i,3)=0;
    tranches(i,4)=0;
    end

  end %for

  trancheLoss=(1-mean(tranches)./detachments)*100;
  % returns losses in pct of initial values
end

%%end test.m
```

Bibliography

[1] Brandimarte, P., 2006, Numerical Methods in Finance and Economics: A MATLAB-Based Introduction, Wiley

[2] Black, F., and J. C. Cox, 1976, Valuing Corporate Securities: Some Effects of Bond Indenture Provisions, *Journal of Finance*, 31, 351-367.

[3] Cheney, E., Kincaid, D., 2007, *Numerical Mathematics and Computing (6th Edition)*, Brooks/Cole

[4] Duffie, D., 2001, *Dynamic Asset Pricing Theory (3rd edition)*, Princeton

[5] Duffie, D., 2007, Innovations in Credit Risk Transfer: Implications for Financial Stability, Obtained from <http://www.bis.org/events/brunnen07/duffiepap.pdf>

[6] Elizalde, A., 2005a, Credit Risk Models I: Default Correlation in Intensity Models, Obtained from <http://abelelizalde.com/>.

[7] Elizalde, A., 2005b, Credit Risk Models II: Structural Models, Obtained from <http://abelelizalde.com/>.

[8] Elizalde, A., 2005c, Credit Risk Models IV: Understanding and Pricing CDOs, Obtained from <http://abelelizalde.com/>.

[9] Goldstein, R., Ju, N, and Leland, H., 2001, An EBIT-Based Model of Dynamic Capital Structure, *Journal of Business*, 483-572.

[10] Harrison, J.; Pliska, S., 1981, Martingales and Stochastic integrals in the theory of continuous trading, *Stochastic Processes and their Applications* 11, 215–260.

[11] Hull, J., 2006, *Options, Futures and Other Derivatives (6th Edition)*, Pearson Prentice Hall

[12] Hull, J., and White, A., 2004, Valuation of a CDO and an nth to Default CDS Without Monte Carlo Simulation, *Journal of Derivatives*, 8-23.

[13] Hull, J., Predescu, M. and White, A., 2006, The Valuation of Correlation-Dependent Credit Derivatives Using a Structural Model, Working Paper

[14] Hull, J., Nelken, I. and White, A., 2004, Merton's Model, Credit Risk and Volatility Skews, *Journal of Credit Risk*, 3-28

[15] Joshi, M. 2008, *C++ Design Patterns and Derivatives Pricing (2nd Edition)*, Cambridge University Press

[16] Leland, H., 1994, Corporate Debt Value, Bond Covenants, and Optimal Capital Structure., *Journal of Finance* 49, 1213-1252.

[17] Li, D., 2000, On Default Correlation: A Copula Function Approach, The RiskMetrics Group, Working Paper Number 99-07

[18] Longstaff, F. and Schwartz, E., 1995, A Simple Approach to Valuing Risky Fixed and Floating Rate Debt, *Journal of Finance*, 50, 789-819.

[19] Longstaff, F. and Schwartz, E., 2001, Valuing American Options by Simulation: A Simple Least-Squares Approach, *Review of Financial Studies*, 14, 113-147

[20] Keenan, S. C., Hamilton D. T. and Berthault, A., 2000, Historical Default Rates of Corporate Bond Issuers, 1920-1999, Moody's Investor Service Global Credit Risk Research. Obtained from: <http://www.moodyskmv.com/research/whitepaper/52453.pdf>

[21] Merton, R., 1974, On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance* 29, 449-470.

[22] Nelsen R. R., 1999, *An Introduction to Copulas*, Springer, New York.

[23] O'Kane, D. and Turnbull, S., 2003, Valuation of Credit Default Swaps, *Quantitative Credit Research Quarterly*, vol. 2003-Q1/Q2

[24] Ross, S., 2007, *Introduction to Probability Models (9th edition)*, Academic Press

[25] Shreve, S., 2004, *Stochastic Calculus for Finance II: Continuous Time Models (1st edition)*, Springer Finance

[26] Trivedi P.K., Zimmer D.M., 2005, Copula Modeling: An Introduction for Practitioners *Foundations and Trends in Econometrics*, 1, 1–111

[27] Vasicek, O., 1984, Credit Valuation, KMV Corporation.

[28] Vasicek, O., 1987, Probability of Loss on Loan Portfolio, KMV Corporation.

[29] Walpole, R., Myers, R., Myers, S. and Ye, K., 2006, *Probability and Statistics for Engineers and Scientists (8th edition)*, Pearson Education.

[30] Wang, S., 1998, Aggregation of Correlated Risk Portfolios: Models & Algorithms, *Proceedings of the Casualty Actuarial Society*, 85, 848-937.

[31] Øksendal, B., 2002, *Stochastic Differential Equations: An Introduction with Applications (4th edition)*, Springer.