UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master's Degree in:

PHYSICS

CARBON DEFAULT SWAP:

FROM BROWNIAN MOTION TO CARBON RISK

Supervisor:

Prof. Giorgia Callegaro

Co-supervisor:

Prof. Antonio Trovato

Candidate:

Tommaso Zazzaron

Academic year 2023/2024

Contents

Introduction

The Brownian motion was first introduced by botanist Robert Brown who observed the random movement of pollen particles due to water molecules under a microscope. It was in the 1900s that the French mathematician Louis Bachelier applied the concept of Brownian motion to asset price behavior for the first time, and this led to Brownian motion becoming one of the most important fundamental of modern quantitative finance. In Bachelier's theory, price fluctuations observed over a small time period are independent of the current price along with historical behavior of price movements: they form a random pattern. The prices fluctuate everyday resulting from market forces like supply and demand, company valuation and earnings, and economic factors like inflation, liquidity, demographics of country and investors, political developments, etc. Market participants try to anticipate stock prices using all these factors and contribute to make price movements random by their trading activities as the financial and economics worlds are constantly changing. This led to the development of the Random Walk Hypothesis or Random Walk Theory, as it is known today in modern finance, which is a statistical phenomenon wherein stock prices move randomly. When the time step of a random walk is made infinitesimally small, the random walk becomes a Brownian motion which follows a Gaussian distribution: the random behavior of prices can be said to be represented by a normal distribution.
The Brownian motion plays a fundamental role also in pricing of derivatives and this is a consequence of their definition. A derivative can be defined as a financial instrument whose value depends on (or derives from) the values of other underlying variables, usually stocks. The Black-Scholes pricing model is a continuous time model widely used to build up their price. As we will see, derivatives can be used for hedging, speculation, and they also play a key role in transferring a wide range of risks in the economy from one entity to another. Through years, derivatives became effective financial instruments for managing credit risk, namely the risk that arises from the possibility that counterparties may default in repaying their debt. Before the introduction of these financial instruments, once a banks or other financial institutions had assumed a credit risk, they could just wait. Now they can manage their portfolios protecting themselves entering into credit derivative contracts. The most popular credit derivative is a Credit Default Swap, that is a contract that provides insurance against the risk of a default by a certain company.
There are many factors that can lead a company to default, one of the most relevant nowadays surely is the transformation required to achieve net-zero emissions targets. Changes in climate policy, technology, and market sentiment during the adjustment to a low-carbon economy could generate significant losses in firms' cash-flows and extreme lowering in their valuation, leading to an economic situation which could undermine their ability to repay debts and could lead to a high probability of default and credit risk. In order to understand the impact of these transformations we need measure firms' carbon risk exposure, which is encoded into their carbon emission data.

The purpose of this thesis is to highlight importance of the Brownian motion in Finance. The Brownian motion was developed in Physics to model the irregular random motion of a particle suspended in fluid under the impact of collisions with the molecules of the fluid. In the first part we will show the analogies between the motion of these so called "Brownian particles" and the trend of stock prices in financial market. We will see that, according to the Black-Scholes model, Brownian motion plays a crucial role in pricing financial stocks and derivatives. Then we will move on introducing the Merton model, in which the Brownian motion is used to estimate the default probability of a company. In the final part we will present a real study case in which derivatives can be implemented to manage firms' credit risk arising from carbon exposure.

In Chapter 1 we will introduce basic financial notions: interest rates and derivatives. In particular, we will present different definitions of interest rates, including the fundamental "risk-free rate", and their role in financial market to price bonds and investment projects. Subsequently we will give the definition of derivatives, with all the features needed to describe them. We will introduce different types of derivatives, such as futures and options. We will concentrate on the latter, defining their payoff and their fundamental properties.

In Chapter 2 we will introduce the concept of Brownian motion. Firstly, we will describe it from a stochastic point of view, paying attention to its implementation on the so called stochastic differential equation. We will present some mathematical features that will be useful, such as the famous Itô's Formula, and finally we will conclude with the Geometric Brownian Motion, a particular stochastic differential equation, which turns out to be the perfect tool for the description of stock prices. Furthermore, we will introduce Brownian motion in physics. We will present Einstein's theory for the probabilistic description of a sample of Brownian particles, and Langevin's implementation, aimed to describe the trajectory of a single Brownian particle starting from Newton's first law of dynamics. Finally, we will deal with Brownian motion in finance. Starting from the analogies between the motion of Brownian particles and the stock price trends, we will end up with the definition of random walk, which plays a fundamental role in modeling stock prices behaviour.

In Chapter 3 we will introduce the Black-Scholes model for derivatives pricing. After a brief introduction concerning the evolution of a discrete time model (binomial model) to a continuous time model (Black-Scholes model), we will deal with the question: "Does it exist a self-financing dynamic portfolio that can replicate the payoff of a derivative?". Solving this so called hedging problem, we will be able to obtain the Black–Scholes partial differential equation. Throughout a particular solution of it, we will deduce the Black–Scholes pricing formula, which will give us the price of option derivatives. Finally, we will present another way to obtain the Call option pricing formula, starting from the Fokker-Planck equation.

In Chapter 4 we will briefly see an application of the Black-Scholes pricing formula. We will introduce the "Greek Letters", defined starting from partial derivatives of the pricing formula, showing their role in managing the market risk of a company.

In Chapter 5 we will introduce credit risk, a risk that arises from the possibility that counterparties may default in repaying their debt. We will present three ways to estimate default probabilities: from historical data, from bond yield spreads, and the Merton model for default probabilities. In particular, we will concentrate on the latter. After some specifications on the definition of debt in case of default, we will show how the Black-Scholes pricing formula for Call options is a suitable tool for calculating the default probability of a company.

In Chapter 6 we will present credit derivatives: contracts where the payoff depends on the creditworthiness of one or more companies or countries. We will focus on the most famous credit derivatives, namely the credit default swaps, which are contracts that provide insurance against the risk of a default by a certain company. After a brief introduction aimed to describe their properties and uses, we will show how, starting from the default probability, is possible to achieve their the correct valuation under the no-arbitrage assumption. In particular, we will deal with all the theoretical calculations that lead to the definition of CDS spreads, namely the total amount of money paid per year to buy protection.

In Chapter 7 we will present a real study case in which, using CDS spreads, we will construct a carbon risk factor, and we will prove how carbon risk, estimated via firms' carbon emission, can affect firms' credit spread. We will show that all the transformations concerning the de-carbonization affect firms' portfolio enhancing their credit risk. We will appreciate that this phenomenon results to be larger for European firms than North American ones and varies across industries, suggesting that the market recognizes where and which sectors are more favorable to a low-carbon economy transition. Furthermore we will report that the effect of carbon risk on CDS spread in stronger during times of heightened attention to climate changes. In this final chapter, after an introductive part, we will show how to implement carbon risk into Merton model for credit risk, we will create a carbon risk factor starting from firms' CDS spreads, and we will introduce the quantile regression model in order to describe our framework. After exploiting this powerful mathematical tool, through some tables, we will have the tangible proof that carbon risk exposure is regional, sectoral, and depends also on climate change attention.

Chapter 1

Preliminary Financial Notions

References: [9] [1] [6]

In this first chapter we will introduce very briefly some financial features which will be fundamental for the treatment of all the arguments of this paper: **Interest Rates** and **Financial Derivatives**. In particular, we will introduce different types of interest rates and their compounding frequency, we will see their role in bond pricing and how they are related to bond yields. Finally we will give the definition of a derivative, showing their most important properties. In particular we will introduce *futures* and *options*, and we will deal with the very famous *call* and *put* options.

1.1 Interest Rates

The main role of interest rates is to establish how much should be the net cash-flow of a security at different points in time.
Interest rates tell us how to measure the monetary value of time, how to price fixed-income securities such as bonds, and how to value investment projects that are delivering cash-flows in the future.

1.1.1 Types of Interest Rates

In financial markets there are different type of real-world interest rates, the most known ones are:

- **Treasury Rates:** This is the rate on instruments issued by the government in its own currency. This represents the rate at which the government can borrow money for a certain period of time, and traditionally they are considered to be risk-free.
- **Interbank Rates:** This is the rate at which primary financial institutions can borrow money from each other for a certain period of time. The most famous one is the LIBOR Rate (London InterBank Offered Rate).
- **Overnight Rates:** Banks have to keep reserves at the central bank, depending on their assets and liabilities. Banks can adjust their deposits by borrowing and lending at the end of each day, at an overnight rate determined by the central bank.

There is another fundamental interest rate, which is not a real-world interest rate, called **Risk-free Interest Rate**, and it will be crucial for the whole analysis presented in this paper.
The risk-free interest rate plays a key role in the valuation of derivatives, but it is also applicable

in situations involving credit risk, namely the risk that the borrower fails to repay interests and principal amount to the lender. This because it is considered as independent of the risk preference of the market participants.
Does the risk free rate exist? Surely not in reality, but some market rates are usually used as proxies for a risk free rate:

- Treasury Rate: The rate at which a government bond is emitted.
- LIBOR Rate: No longer assumed to be risk-free after the financial crisis.
- Overnight Indexed Swap Rate: The most widely adopted proxy nowadays.

We will discuss in more details the risk-free rate in Chapter 3, when we will introduce some pricing models.

1.1.2 Measurement of Interest Rates

Let us suppose that we are investing a certain amount of money in a bank account for one year, we could ask ourselves how much money we will earn at the end of the year. The answer depends on two factors:

- **Interest Rate:** Denoted by R, it is expressed in % and it is referred to a period of 1 year.
- **Compounding Frequency:** Denoted by m, describes how many times per year interests are computed.

Let us make some examples to better understand what has been said.
Example 1. Suppose that we are investing an amount $\$A$ for 1 year at a rate $R = 8\%$ with *annual* compounding frequency ($m = 1$). Then, if a time $t = 0$ we have an amount A, at time $t = 1$ we will have an amount $A(1 + 0.08)$.
Example 2. Suppose that we are investing an amount $\$A$ for 1 year at a rate $R = 8\%$ with *semiannual* compounding frequency ($m = 2$). Then, if a time $t = 0$ we have an amount A, at time $t = 1$ we will have an amount $A(1 + \frac{0.08}{2})^2$.

To generalize our result, suppose that we are investing an amount $\$A$ for T years at a rate R with a compounding frequency m, then the terminal value of the investment at the maturity will be given by the following formula:

$$A\left(1 + \frac{R}{m}\right)^{mT} \tag{1.1}$$

It's interesting to notice that this function is an increasing function of m, this because we are receiving interests on the interests previously credited.

What happens if we let $m \to \infty$? We have the so called *continuous compounding*. The continuous compounding is almost identical to a daily compounding, in fact if we insert $m = 365$ into the previous formula we get a good approximation for $m \to \infty$.
Remembering the expression of the Taylor expansion for the exponential, we obtain the following relation:

$$\lim_{m\to+\infty}\left(1 + \frac{R}{m}\right)^{mT} = e^{RT} \tag{1.2}$$

For mathematical tractability, from now on we will only use continuously compounded rates.

1.1.3 Valuation of Bonds

In the previous subsection we have learnt how to compute the wealth generated by investing some money in a bank account, now we may ask how to do the reverse operation, namely which is the value today of future cash-flows.
In order to do this we introduce the **discounting** operation. Suppose that a certain amount of money $\$A$ is received in T years at a continuously compounded rate R. The value today of the future cash-flow A is:

$$Ae^{-T \cdot R} \tag{1.3}$$

The reason is simple: if we invest $\$Ae^{-T \cdot R}$ today, in T years we will get $\$A$.
Notice that in the previous formula we have assumed the total absence of risk, in the sense that the rate R is not going to change over the time, so the future cash-flow is completely known: we are using the risk-free rate.
Now we are going to exploit the discounting technique to price bonds.
Let us define firstly the *T-year zero rate* $R_0(T)$ as the interest rate which is applied to an investment that starts today and ends in T years from now, without any intermediate payments.
In our financial market bonds take two forms:

- **Zero-coupon Bonds:** They are securities that pay one unit of money at the maturity T without intermediate payments. The price of a zero-coupon bond is given by:

$$P_0 = e^{-T \cdot R_0(T)} \tag{1.4}$$

  where $P_0(T)$ represents the value today (at time $t = 0$) of a single cash-flow of the principal amount (set equal to 1\$ by convention) occurring at maturity (at time $t = T$).

- **Coupon Bonds:** They are securities that pay one unit of money at the maturity T and a fraction $\frac{c}{m}$ every $\frac{1}{m}$ year. The variable c is called coupon rate, while m is the payment frequency. The price of a coupon bond is given by:

$$P_0(T, c, m) = \frac{c}{m} \sum_{k=1}^{mT} e^{-\frac{k}{m} \cdot R_0(\frac{k}{m})} + e^{-T \cdot R_0(T)} \tag{1.5}$$

  where $P_0(T, c, m)$ represents the value today (at time $t = 0$) of the sum of:

  1. *A series of coupon payments.* The amount of each coupon is constant and equal to $\frac{c}{m}$. In total, there are mT coupon payments (m coupons every year), each taking place at time $t = \frac{k}{m}$ (measured in years) and discounted with the corresponding rate $R_0(\frac{k}{m})$.
  2. *A cash flow of the principal amount*, by convention set equal to 1, occurring at the maturity (at time $t = T$).

Another important feature when discussing about bonds is the **bond yield**.
In (1.5) we showed how to compute coupon bonds prices, but we may ask if does exist a single indicator which measures the return on investing in a bond.
We define the **bond yield** as the **single** discount rate that, if applied to **all** the cash-flows of the bond, gives a bond value equal to its market price. We call it y, and defining $P_0(T, c, m)$ the coupon bond price it is the solution to the equation:

$$\frac{c}{m} \sum_{k=1}^{mT} e^{-\frac{k}{m} \cdot y} + e^{-T \cdot y} = P_0(T, c, m) \tag{1.6}$$

For the sake of completeness we give the definition of Net Present Value which will be useful in the following chapters.
Suppose that we want to compute the value today of an investment project which generates cash-flows $(c_1...c_n)$ at dates $(T_1...T_n)$ and requires an initial capital, an investment K, today. We define the **Net Present Value** as:

$$NPV = PV - \text{Investment cost} \tag{1.7}$$

where PV is the present value of future cash-flows and can be computed similarly to the price of a coupon bond:

$$PV = \sum_{i=1}^{n} c_i e^{-T_i \cdot R_0(T_i)} \tag{1.8}$$

Therefore, in the end we have:

$$NPV = \sum_{i=1}^{n} c_i e^{-T_i \cdot R_0(T_i)} - K \tag{1.9}$$

and the investment project is said to be profitable if $NPV > 0$.

1.2 Financial Derivatives

Derivatives will be the link between all the chapters of this thesis. They represent a widely used class of securities implied as hedging instruments as well as investment instruments.
A derivative is an asset whose payoff is defined based on another asset called *underlying*, in most of the cases, but not always, the underlying is a traded security: a stock, a market index, a foreign currency and so on and so forth.
Derivatives are completely described starting from the following fundamental features:

- **Maturity:** the termination date of the derivative contract.
- **Underlying:** the asset from which the derivative gets its definition.
- **Payoff:** the cash-flow, typically delivered at the maturity, generated by the derivative and expressed as a function of the value of the underlying.
- **Price:** the market value of the derivative at a certain time t.

We will denote by S_t the price of a traded asset, the underlying, at time t, for $t \geq 0$.
In this section we will only introduce Futures and Options which are most common derivatives in the financial market.

1.2.1 Futures

Futures are standardized contract to buy or sell something at a predetermined price for delivery at a specified time in the future. Futures are available in an **Exchange-Traded market**, accessible to everybody, which differs from **Over-the-Counter (OTC) market**, which consists of bilateral transactions between two entities (typically large banks, investment founds, corporations). The asset transacted is usually a commodity or financial instrument. The predetermined price of the contract is known as the forward price or delivery price. Upon signing the contract there is no exchange of money, the specified time in the future when delivery and payment occur is known as

the delivery date.
A Future contract with maturity T and delivery price K is a contract with the following payoff at T:

$$\Pi_T^{fut.}(K) = S_T - K \tag{1.10}$$

where S_T in the value of the underlying at the maturity T. As one can see the interpretation is that we enter into an agreement to buy the underlying at a future date T for the predetermined price K.
We now define the **Forward Price** $F_t(T)$ at time t for one unit of underlying with delivery at T as the value of K such that $\Pi_T^{fut.}(K) = 0$. Under absence of arbitrage, which means a strategy which generates money with no risk, it holds that:

$$F_t(T) = S_t e^{(T-t)R_t(T-t)} \tag{1.11}$$

where $R_t(T-t)$ is the zero rate prevailing in the market at time t for an investment length $(T-t)$. Under absence of arbitrage, it is easy to demonstrate that the price of a future contract can be written as:

$$\Pi_t^{fut.}(K) = e^{-(T-t)R_t(T-t)}(F_t(T) - K) = S_t - Ke^{-(T-t)R_t(T-t)} \tag{1.12}$$

which is nothing else than the discounted value of the payoff. **Note.** The forward price is actually defined as the price of *Forward Contracts*. Like futures, forward contracts are agreements to buy the underlying at a future date for a predetermined price, but they are typically traded over-the-counter.

1.2.2 Options

An option is a contract which gives to its owner, the holder, the right, but not the obligation, to buy or sell a specific quantity of an underlying asset at a specified strike price $K > 0$ on or before a specified date, depending on the style of the option. An option is completely specified by the following features:

- **Exercise date:** the time at which you decide whether to exercise the right to buy or sell the underlying.
- **Payoff:** the cash-flow at the exercise date.
- **Strike price:** a fixed quantity, predetermined in the contract of the option, which is part of the option's payoff.
- **Option's price:** the market price of the option prior to exercise. It can never be negative since the option gives a right and not an obligation.

While Futures are called linear derivatives, because their payoff is a linear function of the value of the underlying, options are called non-linear derivatives.
Options can be divided into two big sub-groups:

- **European options:** the exercise date coincides with the maturity T, so the option's payoff is always paid only at the maturity.
- **American options:** the exercise date τ can be freely chosen, as long as it is before the maturity ($\tau \leq T$). If the option is exercised at τ than its payoff will be a function of the underlying price at τ. They are always more valuable than their European version, because you can always choose the exercise date.

In this thesis we will deal only with European options.
The most widely used options are called (European) **Call** and **Put** options.

Call Option A Call option with maturity T and strike price K gives the right to buy one unit of the underlying at maturity T at a predetermined and fixed price K. Its payoff is given by:

$$\Pi_T^{Call} = (S_T - K)^+ = \begin{cases} S_T - K & \text{if } S_T > K \\ 0 & \text{otherwise} \end{cases} \tag{1.13}$$

Put Option A Put option with maturity T and strike price K gives the right to sell one unit of the underlying at maturity T at a predetermined and fixed price K. Its payoff is given by:

$$\Pi_T^{Put} = (K - S_T)^+ = \begin{cases} K - S_T & \text{if } K > S_T \\ 0 & \text{otherwise} \end{cases} \tag{1.14}$$

One of the principal objectives of quantitative finance is the pricing of options. In Chapter 3 we will present a couple of models proposed to solve this problem, namely the multi-period binomial model and the more important Black-Scholes-Merton model. Both these models are constructed by assuming the *absence of arbitrage* principle.

Definition 1. ***(Arbitrage opportunity)****: An arbitrage opportunity is defined as a strategy with:*

1. *zero cost at the initial time;*
2. *no risk of losses at any time;*
3. *strictly positive probability of having strictly positive profits at some future time.*

The **no-arbitrage principle** states the financial market does not admit arbitrage opportunity. We end this section giving two useful lemmas that will be used later on: the **Law of One Price** and the **Put-Call Parity**.

Lemma 1. ***(Law of One Price)****: Let X and Y be two random variables representing the values at some maturity T of two portfolios and let us denote by Π_t^X and Π_t^Y their prices at time $t \in (0, T)$. Suppose that $\mathbb{P}(X = Y) = 1$. If the no-arbitrage principle holds, then:*

$$\Pi_t^X = \Pi_t^Y \quad \textit{for all} \quad t \in [0, T] \tag{1.15}$$

Hence, if two portfolios yield the same payoff at maturity, then their market values should always coincide.

Lemma 2. ***(Put-Call Parity)****: Let $K > 0$ and $T > 0$. If the no-arbitrage principle holds, then:*

$$\Pi_t^{Call}(K) - \Pi_t^{Put}(K) = S_t - K P_t(T) \quad \textit{for all} \quad t \in [0, T] \tag{1.16}$$

where S_t denotes the value of the underlying at time t and $P_t(T) = e^{-(T-t)R_t(T-t)}$ is a discount factor from time t to T.

The Put-Call parity relation is valid only for European options.

Chapter 2

Brownian Motion

References: [6] [3] [5] [12] [7] [11] [13] [15]

In this second chapter we will introduce the concept of Brownian motion. In particular, we will deal with its different definitions, from a stochastic, physical, and financial point of view.

2.1 Probability Spaces

Definition 2. ***(Probability space)*** *A probability space is a triplet* $(\Omega, \mathcal{F}, \mathbb{P})$*, where:*

- Ω *is the set of all the possible states;*
- $\mathcal{F}$ *is a* σ*-algebra on* Ω*;*
- $\mathbb{P}$ *is a probability measure on* $(\Omega, \mathcal{F})$*, i.e. a function* $\mathbb{P} : \mathcal{F} \to [0, 1]$.

Let us give an interpretation fore these objects: an element $\omega \in \Omega$ represents a specific state of the world; $\mathcal{F}$ contains all the relevant events for a certain event $A \in \mathcal{F}$, $\mathbb{P}(\mathbb{A})$ measures the probability that event A occurs.
Given a Borel sigma-algebra $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$, a **Random Variable** X is a measurable function from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, namely a mapping X: $\Omega \to \mathbb{R}$ such that for every set $B \in \mathcal{B}(\mathbb{R})$, we have $X^{-1}(B) \in \mathcal{F}$. A random variable is therefore a function which maps the whole structure of the probability space on $\mathbb{R}$.

Definition 3. ***(σ-algebra)*** *A* σ*-algebra* $\mathcal{F}$ *on* Ω *is a family of subsets of* Ω *such that:*

- $\Omega \in \mathcal{F}$*;*
- *if* $A \in \mathcal{F}$*, than* $A^c \in \mathcal{F}$*;*
- *if* $(A_n)_{n \in N}$ *is a countable family of elements of* $\mathcal{F}$*, than* $\cup_{n \in N} A_n \in \mathcal{F}$*.*

Basically, a σ-algebra represents the "information" we have (or we are interested in) on the random experiment whose possible results are collected in Ω. If we can study the event A, then we can also study the event A^c. If we can study each individual event A_n for some $n \in N$, then we can also study the event that at least one of A_n occurs.

Definition 4. ***(Probability measure)*** *On* $(\Omega, \mathcal{F}, \mathbb{P})$*, a probability measure* $\mathbb{P}$ *is a function* $\mathbb{P} : \mathcal{F} \to [0, 1]$ *such that* $\mathbb{P}(\Omega) = 1$ *and, if* $(A_n)_{n \in N}$ *is a family of disjoint events belonging to* $\mathcal{F}$*, than:*

$$\mathbb{P}\left(\bigcup_n A_n\right) = \sum_n \mathbb{P}(A_n) \tag{2.1}$$

2.2 Stochastic Processes

Definition 5. ***Stochastic process*** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. A* ***Stochastic Process*** $X = (X_t)_{t\in[0,T]}$ *is a family of random variables* X_t*:* $\Omega \to \mathbb{R}$*, indexed by* $t \in [0,T]$*.*

Let us give an interpretation:

- $X = (X_t)_{t\in[0,T]}$ describes the random evolution of a phenomenon over time;
- X_t is the value of the process X at time t, for $t \in [0,T]$;
- for each $\omega \in \Omega$, the map $t \to X_t(\omega)$ denotes the trajectory (or path) of X associated to a specific state of the world ω.

If $[0,T] \subset \mathbb{N}$ we will say that X is a discrete time process, otherwise we will say that it is a continuous time process.
From now on we shall work in continuous time, so the time horizon is T (for $T > 0$), and the possible time points are all the real numbers $t \in [0,T]$.

Definition 6. ***(Filtration)*** *Let* Ω *be the set of the states of the world. A* ***Filtration*** $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ *is an increasing family of* σ*-algebras on* Ω*, so that* $\mathcal{F}_s \subseteq \mathcal{F}_t$*, for all* $0 \leq s \leq t \leq T$*.*

$\mathbb{F}$ represents the information flow of the market, as time passes by one collects more and more information (from the market prices and other sources).

- For a random variable ζ and $t \in [0,T]$, the conditional expectation
$$\mathbb{E}[\zeta|\mathcal{F}_t] \tag{2.2}$$
represents the expectation of ζ given the information available at date t
- We say that ζ is **independent** of $\mathcal{F}_t$ whenever the information collected up to time t is useless to forecast the value of ζ. In this case:
$$\mathbb{E}[\zeta|\mathcal{F}_t] = \mathbb{E}[\zeta] \tag{2.3}$$
- If ζ is $\mathcal{F}_t$**-measurable** then:
$$\mathbb{E}[\zeta|\mathcal{F}_t] = \zeta \tag{2.4}$$
- For any time $0 \leq s \leq t \leq T$, we have the **tower property**:
$$\mathbb{E}[\mathbb{E}[\zeta|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[\zeta|\mathcal{F}_s]|\mathcal{F}_t] = \mathbb{E}[\zeta|\mathcal{F}_s] \tag{2.5}$$

Definition 7. ***(Adaptability)*** *A stochastic process* $(X_t)_{t\in[0,T]}$ *is* ***adapted*** *to the filtration* $(F_t)_{t\in[0,T]}$ *if every* X_t *is* F_t*-measurable.*

A stochastic process is adapted to the filtration if at each time the available information is as much as the information obtained observing the process up to that time.

Definition 8. ***(Martingale)*** *On a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, endowed with a filtration* $\mathbb{F}$*, an adapted stochastic process* $X = (X_t)_{t\in[0,T]}$ *is a* ***Martingale*** *if it is adapted with respect to the filtration and:*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \quad \textit{for all} \quad 0 \leq s \leq t \leq T \tag{2.6}$$

A martingale is a stochastic process that on average remains constant. Let us make an example in order to understand it better: consider a man who is playing head or tail with a coin, let us suppose that he will earn 1£ if head comes out, while he will lose 1£ otherwise. Let $X_0, X_1, X_2, ...$ the total amount of money owned by the man respectively before the the first toss (X_0), after the first toss (X_1) and so on and so forth.
The expectation value of X_n, after n tosses, will be X_0, that is to say the total amount of money initially owned by the man. But if we discover that after m tosses the man owns x£, then the most reasonable amount of money he will own after n tosses (with $n > m$ will be x. This is a **martingale**.

Definition 9. ***(Supermartingale and Submartingale)*** *Let* $(X_n)_{n\geq 0}$ *be an adapted stochastic process. We say that* X *is a:*

- *supermartingale if* $\mathbb{E}[X_t|F_s] \leq X_s$ *for all* $t \geq s$.
- *submartingale if* $\mathbb{E}[X_t|F_s] \geq X_s$ *for all* $t \geq s$.

Definition 10. ***(Markov process)*** *On a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, endowed with a filtration* $\mathbb{F}$*, an adapted stochastic process* $X = (X_t)_{t\in[0,T]}$ *is a* ***Markov Process*** *if:*

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|X_s] \quad \textit{for all} \quad 0 \leq s \leq t \leq T. \tag{2.7}$$

A **Markov process** is a stochastic process in which the transition probability to a specific state of the system depends only on the immediately preceding state, and not on how we get into this preceding state. The future evolution of a Markov process depends only on the current value X_s of the process and not on its past values (X_u with $u < s$).

2.3 What is a Brownian Motion

Definition 11. ***(Brownian motion)*** *On a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, endowed with a filtration* $\mathbb{F}$*, a stochastic process* $W = (W_t)_{t\in[0,T]}$ *starting from* $W_0 = 0$ *is a* ***Brownian Motion*** *if:*

- $(W_t - W_s)$ *is independent of* $\mathcal{F}_s$*, for all* $0 \leq s \leq t \leq T$*;*
- $(W_t - W_s) \sim N(0, t-s)$*, which means that* $W_t - W_s$ *has a Gaussian distribution with mean* 0 *and variance* $(t - s)$*, for all* $0 \leq s \leq t \leq T$*;*
- W *has continuous trajectories.*

A Brownian motion is also defined as ***Weirner process***

Basic properties of the Brownian motion, for any $0 \leq s \leq t \leq T$:

- $\mathbb{E}[W_t] = 0$;
- $Var[W_t] = t$ and consequently $Var[W_t - W_s] = t - s$;
- $Cov[W_s, W_t] = s$;
- If W is a Brownian motion, then W is a martingale;

- If W is a Brownian motion, then W is a Markov process;
- The trajectories $t \in [0, T] \rightarrow W_t$ of a Brownian motion are continuous functions of time, but not nowhere differentiable. This implies that $\frac{dW_t}{dt}$ is not well-defined (i.e. the Brownian motion is not differentiable).

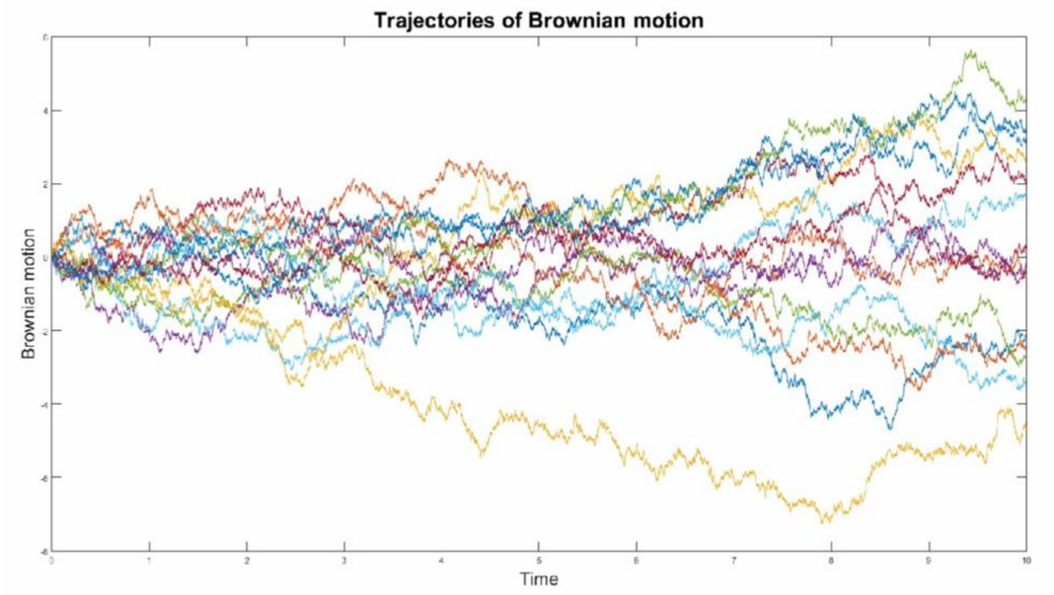


Figure 2.1: Trajectories of Brownian motion. [6]

2.4 Stochastic Differential Equation

We want to model asset prices as continuous time stochastic processes, and the most complete and elegant theory is obtained if we use diffusion processes and stochastic differential equations as our building blocks.
Let us take a stochastic process $(X_t)_{t\in[0,T]}$ with a starting position $X_0 = x \in \mathbb{R}$ which is assumed to be known. This process is called a **diffusion** if its evolution over a small amount of time Δt is driven by the following stochastic difference equation:

$$X_{t+\Delta} - X_t = \alpha(t, X_t)\Delta t + \beta(t, X_t)\sqrt{\Delta t} Z \tag{2.8}$$

where:

- $\alpha(t, X_t)$ is a locally deterministic velocity, called **drift term**.
- $\beta(t, X_t)$ is a locally deterministic **diffusion term**.
- $\sqrt{\Delta t} Z$ is a Gaussian disturbance term. It is a normally distributed variable proportional to $N(0, \Delta t)$ (remember that $Z \sim N(0, 1)$).

Exploiting the property of the Brownian motion according to which $(W_{t+\Delta t} - W_t) \sim N(0, \Delta t)$, we can rewrite the difference equation as:

$$X_{t+\Delta t} - X_t = \alpha(t, X_t)\Delta + \beta(t, X_t)(W_{t+\Delta t} - W_t) \tag{2.9}$$

What happens if we consider shorter and shorter time increments Δt? One could be tempted to divide by Δt and let $\Delta t \to 0$ obtaining:

$$\frac{dX_t}{dt} = \alpha(t, X_t) + \beta(t, X_t)\frac{dW_t}{dt} \tag{2.10}$$

But as we already said the derivative $\frac{dW_t}{dt}$ does not exist.
A possible way to solve the problem is rewriting the stochastic difference equation in terms of sums:

$$X_{t_n} = x + \sum_{k=0}^{n-1} \alpha(k, X_k)\Delta t + \sum_{k=0}^{n-1} \beta(k, X_k)\Delta W_{t_{k+1}} \tag{2.11}$$

where $t_n := n\Delta$, for $n = 0, 1, ..., \frac{T}{\Delta}$.
Now we can let $\Delta t \to 0$ and replace sums with integrals:

$$X_t = x + \int_0^t \alpha(u, X_u)\, du + \int_0^t \beta(u, X_u)\, dW_u \tag{2.12}$$

However we still have a problem: the integral $\int_0^t \beta(u, X_u)\, dW_u$ can not be defined in a standard way, because the trajectories $t \to W_t$ have infinite variation. In order to give a mathematical sense to the previous equation we have to redefine $\int_0^t \beta(u, X_u)\, dW_u$ as a **Stochastic Integral (Itô's Integral)**.

Definition 12. ***(Itô's Integral)*** *An* ***Itô's Integral*** *is an integral of the form:*

$$\int_0^t \beta_u\, dW_u \quad \textit{for} \quad t \in [0, T] \tag{2.13}$$

where:

- *The integrator $W = (W_t)_{t\in[0,T]}$ is a Brownian motion.*
- *The integrand $(\beta_t)_{t\in[0,T]}$ is a stochastic process such that $\mathbb{E}[\int_0^T \beta_u^2\, du] < \infty$.*

Let us see now some properties of the stochastic integral:

1. The stochastic integral $(\int_0^t \beta_u\, dW_u)_{t\in[0,T]}$ is a martingale
2. The following property, named Itô's isometry, holds:

$$\mathbb{E}\left[\left(\int_0^t \beta_u\, dW_u\right)^2\right] = \mathbb{E}\left[\int_0^t \beta_u^2\, du\right] \quad \text{for each} \quad t \in [0, T] \tag{2.14}$$

3. If β is deterministic, then $\int_0^t \beta_u\, dW_u$ has a normal distribution, namely:

$$\int_0^t \beta_u\, dW_u \sim N\left(0, \int_0^t \beta_u^2\, du\right) \quad \text{for each} \quad t \in [0, T] \tag{2.15}$$

Finally we are able to give a proper sense to the equation $X_t = x + \int_0^t \alpha(u, X_u)\, du + \int_0^t \beta(u, X_u)\, dW_u$, where we have understood that the second integral has to be interpreted as a **Itô's Integral**.
The latter is often written in a less cumbersome notation as:

$$dX_t = \alpha(t, X_t)dt + \beta(t, X_t)dW_t \quad \text{with} \quad X_0 = x \tag{2.16}$$

called **Stochastic Differential Equation (SDE)**.

From an intuitive point of view, the SDE is a much more natural object to consider than the corresponding integral expression, this because it gives us the infinitesimal dynamics of the stochastic process X_t.
Suppose now to have a stochastic process $X = (X_t)_{t\in[0,T]}$ that is a solution of the SDE. Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function of time and space; consider the the transformation $f(t, X_t)$, we could ask ourselves which is the SDE satisfied by $f(t, X_t)$.
The answer is provided by the following definition.

Definition 13. ***(Itô's Formula)*** *Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function of class $C^{1,2}$ and $X = (X_t)_{t\in[0,T]}$ a stochastic process solving the stochastic differential equation (2.16). Then $(f(t, X_t))_{t\in[0,T]}$ satisfies the following SDE:*

$$df(t,X_t) = \left(f_t(t,X_t) + f_x(t,X_t)\alpha(t,X_t) + \frac{1}{2} f_{xx}(t,X_t)\beta^2(t,X_t) \right) dt + f_x(t,X_t)\beta(t,X_t)dW_t \quad (2.17)$$

where:

$$f_t(t,X_t) = \frac{\partial f(t,X_t)}{\partial t} \quad f_x(t,X_t) = \frac{\partial f(t,X_t)}{\partial X_t} \quad f_{xx}(t,X_t) = \frac{\partial^2 f(t,X_t)}{\partial X_t^2} \quad (2.18)$$

The fact that $f \in C^{1,2}$ means that the partial derivatives are all well-defined and continuous functions.
The initial value of the function can be written as: $f(0, X_0) = f(0, x)$.

2.4.1 Geometric Brownian Motion

We conclude this section introducing the **Geometric Brownian Motion**.
Geometric Brownian Motion is one of the fundamental building blocks for the modeling of asset prices. It can be viewed as the solution X to a linear ordinary differential equation with a stochastic term driven by a volatility coefficient:

$$dX_t = X_t \mu dt + X_t \sigma dW_t \quad (2.19)$$

where:

- μ is the **drift** coefficient;
- σ is the **volatility** coefficient.

Suppose the stochastic process $X = (X_t)_{t\in[0,T]}$ to be a positive solution to (2.20). In order to obtain the explicit form of X we can apply Itô's formula to the function $f = \log X_t$. By integrating and taking the exponential, we get:

$$X_t = xe^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad \text{for all} \quad t \in (0,T) \quad (2.20)$$

The Geometric Brownian Motion will be the starting point for the Black-Scholes-Merton model which will be presented in Chapter 3.

2.5 Brownian Motion in Physics

Part of probability theory is used to describing the macroscopic picture emerging in random systems defined as the result of microscopic random effects. Brownian motion is the macroscopic picture emerging from a particle moving randomly in d-dimensional space. On the microscopic level, at any time step, the particle receives a random displacement, caused for example by other particles hitting it, so that, if its position at time zero is S_0, its position at time T is given as:

$$S_t = S_0 + \sum_{i=1}^{t} X_i \tag{2.21}$$

where the displacements $X_1, X_2, X_3, ...$ are assumed to be independent, identically distributed random variables with values in $\mathbb{R}^3$. The process $\{S_t : t \geq 0\}$ is a **random walk**, and each displacement represents a well defined fraction of the path.
It can be demonstrated that, if they exist, only the mean and covariance of the displacements define the macroscopic picture: all random walks whose displacements have the same mean and covariance matrix give rise to same macroscopic process.
This effect is called *universality*, and the macroscopic process is often called a universal object.

2.5.1 Einstein's Theory

Einstein in 1905 introduced the first mathematical treatment describing the movement of Brownian particles. Rather than focusing on the trajectory of a single particle, Einstein introduced a probabilistic description valid for an ensemble of Brownian particles. The chaotic motion of a Brownian particle is the result of its collisions with the molecules of the surrounding fluid (remember that the Brownian particles are much bigger and heavier than the colliding molecules of the fluid). The superposition of these interactions produces an observable effect. The agitated motion of Brownian particle is then the result of random and rapid collisions due to density fluctuations in the fluid: it is a macroscopic manifestation of microscopic processes.
Firstly, Einstein introduced the concept of a *coarse-grained* description defined by a time τ such that different parts of the path separated by a time τ can be considered independent. Subsequently, he introduced a *probability density function* $f(\boldsymbol{\Delta})$ for the three-dimensional distance $\boldsymbol{\Delta} = (\Delta_x, \Delta_y, \Delta_z)$ travelled by the Brownian particle in a fixed time interval τ.
The only assumption he made about $f(\boldsymbol{\Delta})$ came from the fact that the collisions of the fluid molecules and the Brownian particle occur with the same probability in any direction. The absence of preferred directions translates to a symmetry condition for $f(\boldsymbol{\Delta})$:

$$f(\boldsymbol{\Delta}) = f(-\boldsymbol{\Delta}), \quad \forall \boldsymbol{\Delta} \in \mathbb{R}^3 \tag{2.22}$$

Finally, he considered an ensemble of N Brownian particles in a large enough system. Also, he focused on large spatial scales, much larger than the size of a Brownian particle, as a consequence we can define the particles density $n(\boldsymbol{x}, t)$ such that $n(\boldsymbol{x}, t)d\boldsymbol{x}$ is the number of particles in the interval spatial interval $(\boldsymbol{x}, \boldsymbol{x} + d\boldsymbol{x})$ at time t.
From the assumption that the parts of the trajectories separated by τ are statistically independent and the conservation of particle number, it follows that the number of particles at location $\boldsymbol{x}$ at time $t + \tau$ will be given by the number of particles at location $\boldsymbol{x} - \boldsymbol{\Delta}$ at time t multiplied by the probability that the particle jumps from $\boldsymbol{x} - \boldsymbol{\Delta}$ to $\boldsymbol{x}$ in a time step τ, which is $f(\boldsymbol{\Delta})$, and integrated over all the possible values of $\boldsymbol{\Delta}$:

$$n(\boldsymbol{x}, t + \tau) = \int_{\mathbb{R}^3} n(\boldsymbol{x} - \boldsymbol{\Delta}, t) f(\boldsymbol{\Delta}) d\boldsymbol{\Delta} \tag{2.23}$$

This represents the evolution equation for the number density $n(\boldsymbol{x}, t)$. It is a continuity equation expressing particle conservation: Brownian particles can neither be created, nor can they disappear

as a result of the collisions with the fluid molecules.
By Taylor expanding the above expression, namely:

$$n(\boldsymbol{x}, t+\tau) \sim n(\boldsymbol{x}, t) + \tau \frac{\partial n(\boldsymbol{x}, t)}{\partial t} \tag{2.24}$$

$$n(\boldsymbol{x} - \boldsymbol{\Delta}, t) \sim n(\boldsymbol{x}, t) - \boldsymbol{\Delta} \cdot \boldsymbol{\nabla} n(\boldsymbol{x}, t) + \frac{\boldsymbol{\Delta}^2}{2} \boldsymbol{\nabla}^2 n(\boldsymbol{x}, t) \tag{2.25}$$

and making use of the normalization of the pdf $f(\boldsymbol{\Delta})$ and the symmetry relation $f(\boldsymbol{\Delta}) = f(-\boldsymbol{\Delta})$, one can obtain the *diffusion equation* for the particles density:

$$\frac{\partial n(\boldsymbol{x}, t)}{\partial t} = D \boldsymbol{\nabla}^2 n(\boldsymbol{x}, t) \tag{2.26}$$

where D is called **Diffusion Coefficient** and it is given in terms of the pdf as:

$$D = \frac{1}{2\tau} \int_{\mathbb{R}^3} \boldsymbol{\Delta}^2 f(\boldsymbol{\Delta}) d\boldsymbol{\Delta} \tag{2.27}$$

To solve this partial differential equation, we need an initial condition. If we suppose that initially all Brownian particles are located at the origin, we thus have $n(\boldsymbol{x}, 0) = N\delta(\boldsymbol{x})$. Exploiting the Fourier transform one can obtain:

$$n(\boldsymbol{x}, t) = \frac{N}{(4\pi D t)^{\frac{3}{2}}} e^{-\frac{\boldsymbol{x}^2}{4Dt}} \tag{2.28}$$

which is a Gaussian function of the position $\boldsymbol{x}$.
From (2.28) Einstein was able to calculate the average value over many trajectories $\langle \boldsymbol{x}^2(t) \rangle$ of the squared displacement of a (*single*) Brownian particle at time t:

$$\left\langle \boldsymbol{x}^2(t) \right\rangle = \frac{1}{(4\pi D t)^{\frac{3}{2}}} \int |\boldsymbol{x}|^2 e^{-\frac{\boldsymbol{x}^2}{4Dt}} d\boldsymbol{x}^3 \tag{2.29}$$

Doing all the calculations, the latter gives:

$$\left\langle \boldsymbol{x}^2(t) \right\rangle = 6Dt \tag{2.30}$$

This prediction was successfully confirmed by experiments and contributed decisively to the acceptance of the atomic/molecular theory of matter.
If we consider a one-dimensional system with only one particle we will have:

$$n(x, t) = \frac{1}{\sqrt{4\pi D t}} e^{-\frac{x^2}{4Dt}} \tag{2.31}$$

and

$$\left\langle x^2(t) \right\rangle = 2Dt \tag{2.32}$$

2.5.2 Langevin's Theory

Einstein's approach was successful and incomplete at the same time, it couldn't yield, for instance, an explicit expression for the diffusion coefficient in terms of microscopic quantities.
Langevin initiated a different approach which, in some ways, can be considered complementary to the previous one. In his approach, Langevin focused on the trajectory of a single Brownian particle and wrote down Newton's first law of dynamics.
The trajectory of the Brownian particle is highly erratic and therefore its description would demand a particular kind of force. He considered Newton's first law with two types of forces acting on the Brownian particle:

- a **friction force $\boldsymbol{F_d}$**;
- a **fluctuating force $\boldsymbol{\zeta(t)}$**

Since the particle under consideration is much larger than the particles of the surrounding fluid, the collective effect of the interaction between the Brownian particle and the fluid's particles may be considered as a hydrodynamical frictional force. The friction force exerted by a fluid on a small sphere immersed in it can be determined from **Stokes' law**, which states that the drag force acting on a spherical particle of radius a is given by:

$$\boldsymbol{F}_d = M\frac{d\boldsymbol{v}}{dt} = -\gamma\boldsymbol{v}, \quad \gamma = 6\pi\eta a \tag{2.33}$$

where M is the mass of the Brownian particle and η is the fluid *viscosity* coefficient. Furthermore, Langevin made two more assumptions on the fluctuating force $\boldsymbol{\zeta(t)}$:

- it has *mean* equal to 0, namely:
$$\left\langle \boldsymbol{\zeta}(t) \right\rangle = 0 \tag{2.34}$$
meaning that collisions do not push the Brownian particle in any preferred direction;
- it is uncorrelated to the actual position of the Brownian particle, namely:
$$\left\langle \boldsymbol{x}\cdot\boldsymbol{\zeta}(t) \right\rangle = \langle \boldsymbol{x} \rangle \cdot \left\langle \boldsymbol{\zeta}(t) \right\rangle = 0 \tag{2.35}$$
meaning that the action of the molecules of fluid on the Brownian particle is the same no matter the location of the Brownian particle.

Said so, the equation of motion of the Brownian particle becomes:

$$M\frac{d\boldsymbol{v}}{dt} = -6\pi\eta a\boldsymbol{v} + \boldsymbol{\zeta} \tag{2.36}$$

Multiplying both sides by $\boldsymbol{x}$, and knowing that $\boldsymbol{v} = \frac{d\boldsymbol{x}}{dt}$, with some algebra one gets:

$$\frac{m}{2}\frac{d^2\boldsymbol{x}^2}{dt^2} - m\left(\frac{d\boldsymbol{x}}{dt}\right)^2 = -3\pi\eta a\frac{d\boldsymbol{x}^2}{dt} + \boldsymbol{x}\cdot\boldsymbol{\zeta} \tag{2.37}$$

Taking the averages, and exploiting (2.34) (2.35), one gets:

$$\frac{m}{2}\frac{d^2\left\langle \boldsymbol{x}^2\right\rangle}{dt^2} = m\left\langle \boldsymbol{v}^2\right\rangle - 3\pi\eta a\frac{d\left\langle \boldsymbol{x}^2\right\rangle}{dt} \tag{2.38}$$

which is an equation for the average square position of the Brownian particle.
Note. *We neglected the time dependence in order to a have a sloppier notation.*
Finally, Langevin assumed that the system was in a regime in which thermal equilibrium between the Brownian particle and the surrounding fluid has been reached. In particular, this implies that, according to the equipartition theorem, the average kinetic energy of the Brownian particle is:

$$\left\langle \frac{m\boldsymbol{v}^2}{2} \right\rangle = \frac{3}{2}k_BT \tag{2.39}$$

where k_B is the Boltzmann's constant and T is the fluid temperature.
Solving the previous equation one can find that the *mean square displacement* is given by:

$$\left\langle \boldsymbol{x}^2\right\rangle = \frac{k_BT}{\pi\eta a}t \tag{2.40}$$

This is nothing but Einstein's diffusion law (2.30) and now we have an explicit expression for the diffusion coefficient in terms of other macroscopic variables:

$$D = \frac{k_B T}{6\pi\eta a} \tag{2.41}$$

This last equation is known as **Stokes-Einstein** relation.
Langevin's equation is an example of stochastic differential equation, and Langevin's random force $\boldsymbol{\zeta}$ is an illustration of a stochastic process. The solutions $\hat{\boldsymbol{x}}$ can be seen as the collection of all the possible trajectories of the Brownian particle.
The random force $\boldsymbol{\zeta}$ gives the effect of background noise due to the fluid on the Brownian particle. If we would neglect this force, the Langevin equation would become:

$$M\frac{d\boldsymbol{v}}{dt} = -\gamma\boldsymbol{v} \tag{2.42}$$

which has the familiar solution:

$$\boldsymbol{v}(t) = e^{-\frac{\gamma}{M}t}\boldsymbol{v}(0) \tag{2.43}$$

According to this, the velocity of the Brownian particle is predicted to decay to zero at long times. This cannot be true since in equilibrium we must have (in one dimension):

$$\left\langle \boldsymbol{v}^2 \right\rangle = \frac{k_B T}{M} \tag{2.44}$$

which is different from zero. The random force is therefore necessary to obtain the correct equilibrium. One can obtain an explicit formal solution of (2.34) as:

$$\boldsymbol{v}(t) = e^{-\frac{\gamma}{M}t}\boldsymbol{v}(0) + \frac{1}{M}\int_0^t e^{-\frac{\gamma}{M}(t-s)}\boldsymbol{\zeta}(\boldsymbol{s})ds \tag{2.45}$$

Now we rewrite (2.36) as:

$$d\boldsymbol{v} = -\frac{\gamma}{M}\boldsymbol{v}(t)dt + \frac{1}{M}d\boldsymbol{U}(t) \tag{2.46}$$

where

$$d\boldsymbol{U}(t) = \boldsymbol{\zeta}(t)dt \quad \longrightarrow \quad \boldsymbol{U}(t) = \boldsymbol{U}(0) + \int_0^t \boldsymbol{\zeta}(t)dt \tag{2.47}$$

Since a Brownian particle in a liquid solution undergoes many random collisions per second with the particles of the environment, we can assume that these numerous impacts destroy all correlations between what happens during the time interval $[t_i, t_{i-1}]$ and what happened before t_{i-1}. This implies that $\boldsymbol{U}(t)$ is a Markov process. From this discussion we also argue that the increments are independent, stationary and identically distributed with zero mean, therefore, it has all the requirements for being a Wiener process:

$$\boldsymbol{U}(t) = \boldsymbol{W}(t) \tag{2.48}$$

then, (2.45) becomes:

$$\boldsymbol{v}(t) = e^{-\frac{\gamma}{M}t}\boldsymbol{v}(0) + \frac{1}{M}\int_0^t e^{-\frac{\gamma}{M}(t-s)}d\boldsymbol{W}(s) \tag{2.49}$$

2.5.3 Random Walk

The most well-known example of a stochastic process is that of the random walk. Moreover, the continuum limit, on large scale, of a 3-dimensional random walk offers a good description of Brownian motion. It also represents a simple model to understand financial markets.
For simplicity we will focus only on the one-dimensional random walk. The probabilistic experiment can be represented as a repetition of coin tosses. Let us consider that the tossing takes place at given discrete times $\{0, \tau, 2\tau, ...\}$, to the outcome of this set of probabilistic experiments, we associate a one-dimensional function $x(t)$ which starts at $x(0) = 0$ and that moves to the right (left) at time $k\tau$ an amount $+a(-a)$ if the *k-th* result of the tossed coin was $+1(-1)$, namely *head(tail)*.
The random walk constitutes an example of a Markov process, as the probability of having a particular value of the position at time $(k+1)\tau$ depends only on the particle's location at time $k\tau$ and not on the way it got to this location.
Starting at $x = 0$ at time $t = 0$, the location of the random walker after tossing the coin n times is given by the number of steps n_+ of increasing x (”*heads*”) minus the number of steps of decreasing x, $n_- = n - n_+$ (”*tails*”), namely:

$$x(t) = x(n\tau) = (n_+ - n_-)a = (2n_+ - n)a \tag{2.50}$$

The probability of having n_+ ”*heads*” after n throws is given by a **binomial distribution**:

$$\mathbb{P}(n_+) = \binom{n}{n_+} p^{n_+}(1-p)^{n-n_+} \tag{2.51}$$

where p is the probability of having a ”*head*”, namely $p = \frac{1}{2}$.
Using the binomial distribution properties we have that:

- $\mathbb{E}[x(n\tau)] = 0$
- $Var[x(n\tau)] = \mathbb{E}[x^2(n\tau)] - (\mathbb{E}(x(n\tau)))^2 = \mathbb{E}[(x(n\tau))^2] = na^2$

It can be easily demonstrated that in the limit $n >> 1$ the binomial distribution can be approximated by a Gaussian distribution, consequently, if we take the continuum limit $n \to \infty$ the random walk process converges to the so-called **Brownian Motion $\boldsymbol{W(t)}$**, characterized by a Gaussian probability distribution function with *mean* zero and *variance* $2Dt$, namely

$$f(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \tag{2.52}$$

One can appreciate that the variance is:

$$Var[x(t)] = 2Dt = \mathbb{E}[x^2(n\tau)] \tag{2.53}$$

which is exactly the expression (2.32), solution of the diffusion equation in Einstein's theory of Brownian motion in one dimension.

2.6 Brownian Motion in Finance

Brownian motion, in particular Geometric Brownian Motion is a mathematical approach for stock price modelling. It is a stochastic process, which assumes that returns, profits and losses on the stock are independent and normally distributed.

2.6.1 From Physics to Finance

Financial market dynamics is rigorously studied via the generalized Langevin equation.
We recall that the Brownian motion was developed in statistical mechanics to model the irregular, random motion of a particle (usually called Brownian particle) suspended in fluid under the impact of collisions with the molecules of the fluid.
The situation in economics and finance is analogous: the price of an asset depends on many factors and for this reason it continuously changes. The analogy between the buffeting effect of molecules on a particle in the statistical mechanics context on the one hand, and that of this continuous flood of new price-sensitive information on the other, is highly suggestive. Despite these similarities, Brownian motion itself is inadequate for modelling prices, this because:

1. it attains negative levels;
2. one should think in terms of *return*, rather than prices themselves.

One can allow for both of these by using **Geometric Brownian Motion**, rather than the ordinary Brownian motion. Recalling some definitions, introducing the standard Brownian motion as $W(t)$, we define the Brownian motion with drift as:

$$dX(t) = \alpha dt + \beta dW(t) \tag{2.54}$$

and the GBM as:

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t) \quad \text{or} \quad \frac{dX(t)}{X(t)} = \mu dt + \sigma dW(t) \tag{2.55}$$

Let us see how to interpret it: $X(t)$ is the stock price at time t, $dX(t) = X(t + dt) - X(t)$ is the change in $X(t)$ over a small time interval dt, $\frac{dX(t)}{X(t)}$ is the gain per unit of value, namely the *return*. The latter is a sum of two components:

- a deterministic component μdt, equivalent to investing risk-free money in a bank account at interest-rate $\mu > 0$, called the underlying return rate for the stock;
- a random noise component $\sigma dW(t)$, with volatility parameter $\sigma > 0$ and driving Brownian motion $W(t)$, which models the market uncertainty. Note that $dW(t)$ is a Brownian increment, so is normally distributed. Therefore returns are normally distributed.

2.6.2 Stock Price Behaviour

In financial markets the dynamics of stock prices are reflected by uncertain movements of their value over time. One possible reason for the random behavior of the asset price is the efficient market hypothesis which states two things:

1. The past history of a stock price is fully reflected in present prices.
2. The markets respond immediately to any new information about the stock.

These two assumptions imply that the stock price, that is a stochastic process, is a Markov process. This means that the expected future value of a stock depends only on its current price. Predictions remain uncertain and may be only expressed in term of probability distribution.
If one looks to the price behaviour of stocks, it will be easy to see that it shows the same behaviour as the Brownian motion. Consequently, some properties of the stock price process can be derived from those of the Brownian motion process.

The **Random Walk** is the first step in understanding the Brownian motion. A random walk is a formalization of the intuitive idea of taking successive casual steps. The simplest random walk is a path constructed such that for an integer $n > 0$ we define the random walk process as $W_n(t)$, with $t \in [0, n]$. The random walk presents the following features:

- The initial value of the process is $W_n(0) = 0$
- The layer spacing between two successive jumps is equal to $\frac{1}{n}$
- The "up" and "down" jumps are equal and of size $\frac{1}{\sqrt{n}}$, with equal probability.

Basically, if we consider a sequence of independent binomial variable X_i taking values $+1$ or -1 with equal probability $\frac{1}{2}$, then the value of the random walk at the *i-th* step is defined as:

$$W_n\left(\frac{i}{n}\right) = W_n\left(\frac{i-1}{n}\right) + \frac{X_i}{\sqrt{n}} \quad \text{for all} \quad i \in [0, n] \tag{2.56}$$

In the following figure we show the first two steps of a random walk $W_n(t)$.

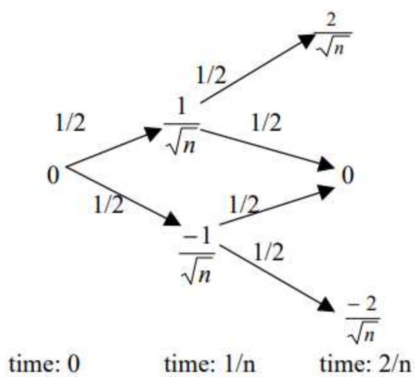


Figure 2.2: First two steps of the random walk. [5]

One can easily calculate the expectation value and the variance of the random walk at each step, for example for the first step we will have:

$$\mathbb{E}\left[W_n\left(\frac{1}{n}\right)\right] = 0.5 \cdot \frac{1}{\sqrt{n}} + 0.5 \cdot \frac{-1}{\sqrt{n}} = 0 \tag{2.57}$$

$$Var\left[W_n\left(\frac{1}{n}\right)\right] = \mathbb{E}\left[W_n^2\left(\frac{1}{n}\right)\right] - \left(\mathbb{E}\left[W_n\left(\frac{1}{n}\right)\right]\right)^2 = \ldots = \frac{1}{n} \tag{2.58}$$

The same can be done for the second step, namely:

$$\mathbb{E}\left[W_n\left(\frac{2}{n}\right)\right] = 0.25 \cdot \frac{2}{\sqrt{n}} + 0.5 \cdot 0 + 0.25 \cdot \frac{-2}{\sqrt{n}} = 0 \tag{2.59}$$

$$Var\left[W_n\left(\frac{2}{n}\right)\right] = \ldots = \frac{2}{n} \tag{2.60}$$

But what does the random walk look like if n gets larger and larger, or equivalently, the time intervals become smaller and smaller?
It can be proved that, according to central limit theorem, for n large the random walk process $W_i(t)$

tends to a normal distribution with mean zero and variance t, namely $N(0,t)$. The mathematical model used to describe random movements it this scaling limit is the so called Brownian motion. Hence, the random walk tends to a Brownian motion when the number of steps in the random path is very large, or equivalently, the time intervals go to zero.

The generalized random walk, also called *Brownian motion with drift*, is a stochastic process $X(t)$ such that:

$$X(t) = \alpha t + \beta W(t) \tag{2.61}$$

where t represents the time, and $W(t)$ represents a random walk in the limit in which the number of steps tend to infinity. One can easily appreciate that:

$$\mathbb{E}[X(t)] = \mathbb{E}[\alpha t] + \mathbb{E}[\beta W(t)] = \alpha t + \beta \mathbb{E}[W(t)] = \alpha t \tag{2.62}$$

$$Var[X(t)] = \mathbb{E}[X^2(t)] - (\mathbb{E}[X(t)])^2 = \ldots = t \tag{2.63}$$

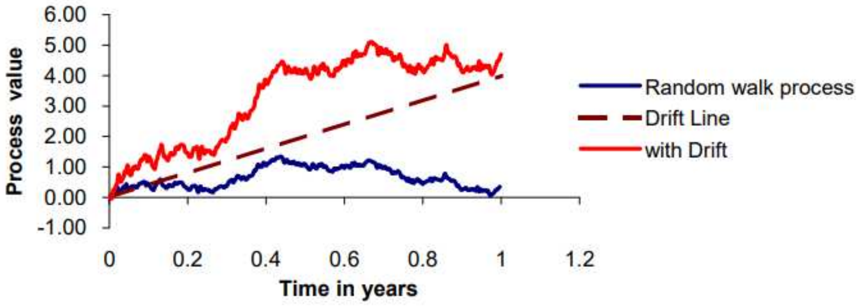


Figure 2.3: Generalized Brownian motion process with **positive** drift. [5]

Figure 2.4: S&P 500 stock price between 2018 and 2021. [5]

The figure 2.3 shows the generalized Brownian process given by (2.59) with its components. It consists of a random walk process with a drift, and the result is an increasing process if the drift term is positive or a decreasing one if it is negative.
The figure 2.4 shows the S&P 500 stock price trend from 2018 to 2021. S&P 500 is an index referring to a basket of 500 stock titles representing the 80% of the whole market capitalization.

If we compare the two figures above, we can easily conclude that the two processes show the same behavior in time. This statement is the first step towards the mathematical modelling of stock prices.

2.6.3 Stock Price Modeling

In the previous subsection we presented similarities between the Brownian motion and the stock price process. However there is a misgiving about Brownian motion as a global model for stock behavior. This because of the property of the Brownian motion which states that the process is normally distributed with mean equal zero, meaning that stock prices could have negative values, and this is in general not true. One way to solve this problem is to model the stock price as a sum of a deterministic function of time and a Brownian motion.
In the real world of financial markets, investors and financial analysts are generally more interested in the profit or loss of the stock over a period of time i.e. the *increasing* or *decreasing* of the price, than in the price itself, this will be our starting point.
Note. *In this subsection we will change the notation. We will write $X(t)$ as X_t in order to highlight the fact that X is a stochastic process.*

Like a Brownian particle being hit in its Brownian motion, stock prices deviate from a steady state as a result of being jolted by trades i.e. ask and bid in financial markets. If we consider the stock price as a stochastic process S_t at time t, then the *return* or relative change in its price during the next period of time dt can be decomposed into:

- A predictable and deterministic part that is the expected return from the stock during a period of time dt. This term is the equal to:

  $$\mu S_t dt \tag{2.64}$$

  where μ is called **drift** term.

- A stochastic and unexpected part, which reflects the random changes in stock price during the time interval dt, as response to external effects such as unexpected news on the stock. A reasonable assumption is to take this contribution as proportional to the stock, namely:

  $$\sigma S_t dW_t \tag{2.65}$$

  where σ is called **volatility** and W_t is the random walk.

This definition of the daily return leads to the stochastic differential equation for the stock price, which is nothing else than a Geometric Brownian Motion:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.66}$$

or

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{2.67}$$

which highlights the instantaneous rate of return on S_t.
The solution for the latter can be found by applying Itô's formula, let us remind it:
Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function of class $C^{1,2}$ and $S = (S_t)$ a stochastic process solving the stochastic differential equation (2.64). Then $f(t, S_t)$ satisfies the following SDE:

$$df(t,S_t) = \left(f_t(t,S_t) + f_x(t,S_t)\mu S_t + \frac{1}{2} f_{xx}(t,S_t)(\sigma S_t)^2 \right) dt + f_x(t,S_t)\sigma S_t dW_t \tag{2.68}$$

where:

$$f_t(t, S_t) = \frac{\partial f(t, S_t)}{\partial t} \quad f_x(t, S_t) = \frac{\partial f(t, XS_t)}{\partial S_t} \quad f_{xx}(t, S_t) = \frac{\partial^2 f(t, S_t)}{\partial S_t^2} \tag{2.69}$$

If we now consider the function $f(t, S_t) = \ln(S_t)$, then:

$$\frac{\partial f(t, S_t)}{\partial t} = 0 \qquad \frac{\partial f(t, XS_t)}{\partial S_t} = \frac{1}{S_t} \qquad \frac{\partial^2 f(t, S_t)}{\partial S_t^2} = -\frac{1}{S_t^2} \tag{2.70}$$

By inserting them into Itô's formula, one can get the solution for the GBM:

$$S_t = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t} \tag{2.71}$$

which can be adapted to any period of time.
It is easy to demonstrate that:

- $\mathbb{E}[S_t] = S_0 e^{\mu t}$
- $Var[S_t] = \mathbb{E}[S_t^2] - (\mathbb{E}[S_t])^2$

We could appreciate that in the expression of the stock price there are two fundamental parameters. These parameters are supposed to be known. If for some reason there is no available information about the drift and the volatility of the stock, they need to be estimated form historical data.
The volatility σ is a constant characteristic of a stock, usually expressed as an annual percentage. It gives an idea about the stability of stock price. Relatively high volatility means that the stock price varies continuously within relatively large interval. The value most usual method of measuring the stock volatility is the standard deviation of the price returns.
The rate of return μ is the gain (or loss) compared to the cost of an initial investment, typically expressed in the form of a percentage. When it is positive, it is considered a gain, and when the it is negative, it reflects a loss on the investment. This term will have a central role in the discussion of Chapter 3.

Chapter 3

The Black-Scholes-Merton Model

References: [9] [1] [6]

In the following section we will present the Black-Scholes-Merton model, which is a stochastic model for the dynamics of a financial market containing derivative investment instruments, using various assumptions.
The starting point will be to solve the so called hedging problem: *Does it exist a self-financing dynamic portfolio that can replicate the payoff of a European derivative?* Once solved it, we will be able to obtain the Black–Scholes partial differential equation (**Black-Scholes PDE**). Through a particular solution of the latter, we will deduce the **Black–Scholes pricing formula**, which will give us a theoretical estimate of the price of European-style options and will show us that the options have a unique price.

3.1 From Binomial Model to Black-Scholes-Merton Model

In this section we will discuss the multi-period binomial model, an easy discrete time pricing model which will lead us to the more complicated continuous time pricing model known as Black-Scholes-Merton model.
The model is assumed to have a discrete time structure of N periods. $t = 0$ represents the current time when we compute prices and construct portfolios, $t = N$ represents the end of the investment horizon, when we receive payoffs, called maturity.
In the Binomial model the market consists of a risky asset (a stock), and a risk-free asset (a bond):

- The **Risk-free asset** has value B_t, for $t \in (0, 1, ..., N)$, with:

$$B_t = (1 + r_f)^t \tag{3.1}$$

  where r_f is the risk-free rate over the period $[t-1, t]$, and it is assumed to be independent on time, than it is the same for all time intervals. The starting point is assumed to be $B_0 = 1$.

- The **Risky asset** has value S_t, for $t \in (0, 1, ..., N)$, with:

$$S_t = s \prod_{i=1}^{t} Z_i \tag{3.2}$$

where s is the price of the stock at time $t = 0$, and $Z_1, ..., Z_N$ is a family of identically distributed random variables such that:

$$Z_i = \begin{cases} u & \text{with probability } p_u \\ d & \text{with probability } p_d = 1 - p_u \end{cases} \tag{3.3}$$

Z_i represents the random return in the period $[i-1, i]$ for $i = 1, ..., N$:

1. $Z = u$ represents a move up.
2. $Z = d$ represents a move down.

Generally, we will have a certain number "n" of move up, and "$N - n$" move down.

Let us now introduce a *dynamic portfolio*.

Definition 14. ***(Dynamic portfolio)*** *A* ***Dynamic portfolio*** *is a couple $\theta = (\theta^B, \theta^S)$ such that:*

- $\theta_0 = (\theta_0^B, \theta_0^S)$ *is the portfolio initially created at time $t = 0$ which has value:*

$$V_0(\theta) = \theta_0^B B_0 + \theta_0^S s \tag{3.4}$$

- $\theta_t = (\theta_t^B, \theta_t^S)$ *is a stochastic process, which represents the portfolio rebalanced at time t, function of the random returns $Z_1, ..., Z_t$, whose value is:*

$$V_t(\theta) = \theta_t^B B_t + \theta_t^S S_t \tag{3.5}$$

Suppose that a dynamic portfolio is created at $t = 0$, its initial value, equal to its cost, is:

$$V_0(\theta) = \theta_0^B B_0 + \theta_0^S S_0 \tag{3.6}$$

At time $t = 1$ the portfolio yields a value equal to:

$$V_1(\theta) = \theta_0^B B_1 + \theta_0^S S_1 \tag{3.7}$$

At time $t = 1$ we are able to rebalance our portfolio depending on the realization of Z_1. The new portfolio requires a capital equal to:

$$V_{1^+}(\theta) = \theta_1^B B_1 + \theta_1^S S_1 \tag{3.8}$$

and it will lead, at $t = 2$, to a value:

$$V_2(\theta) = \theta_1^B B_2 + \theta_1^S S_2 \tag{3.9}$$

Definition 15. ***(Self-financing portfolio)*** *A portfolio is said* ***self-financing*** *if $V_1(\theta) = V_{1^+}(\theta)$. This means that the value of the previous portfolio (created at $t = 0$) is equal to the cost of setting up the new rebalanced portfolio at time $t = 1$.*
At generic time t, the self-financing condition can be written as:

$$\theta_{t-1}^B B_t + \theta_{t-1}^S S_t = \theta_t^B B_t + \theta_t^S S_t \tag{3.10}$$

Definition 16. ***(Arbitrage portfolio)*** *A dynamic self-financing portfolio θ is said to be an* ***arbitrage portfolio*** *if:*

- $V_0(\theta) = 0$.
- $\mathbb{P}(V_N(\theta) \geq 0) = 1$.
- $\mathbb{P}(V_N(\theta) > 0) > 0$.

No-arbitrage holds if arbitrage portfolios do not exist.
It is straightforward to demonstrate that in the multi-period binomial model, the no-arbitrage condition holds if and only if:

$$d < 1 + r_f < u \tag{3.11}$$

Let us now consider a derivative with a risky asset as the underlying, the payoff of the derivative can be represented as a function of the identically distributed random variables $g(Z_1, ..., Z_N)$, where $g : [u, d]^N \to \mathbb{R}$.
If the market is complete, like in the binomial case, it is always possible to find a dynamic self-financial portfolio such that, at time $t = N$, its value replicates the payoff of the derivative, namely:

$$V_N(\theta) = g(Z_1, ..., Z_N) \tag{3.12}$$

Our main goal it to find a price for this derivative at time $t = 0$.
By the **Law of One Price** and the **No-Arbitrage condition** one can easily conclude that its price $\Pi_0(g)$ at $t = 0$ is:

$$\Pi_0(g) = V_0(\theta) \tag{3.13}$$

Let us take as an example the single period binomial model, given a derivative with payoff $g(Z)$ we want to determine a self-financing portfolio $\theta = (\theta^B, \theta^S)$ such that $V_1(\theta) = g(Z)$, this is equivalent to solve the following system:

$$\begin{cases} \theta^B(1 + r_f) + \theta^S su = g(u) \\ \theta^B(1 + r_f) + \theta^S sd = g(d) \end{cases} \tag{3.14}$$

which has the following solution:

$$\theta^B = \frac{1}{1 + r_f} \frac{ug(d) - dg(u)}{u - d} \qquad \theta^S = \frac{1}{s} \frac{g(u) - g(d)}{u - d} \tag{3.15}$$

Resuming what we have seen, holding a portfolio $\theta = (\theta^B, \theta^S) = \left(\frac{1}{1+r_f} \frac{ug(d)-dg(u)}{u-d}, \frac{1}{s} \frac{g(u)-g(d)}{u-d} \right)$ leads to the same payoff $g(Z)$ of the derivative at $t = 1$. By now we do not know the price $\Pi_0(g)$ of the derivative at time $t = 0$, but we know the value $V_0(\theta)$ of the hedging portfolio at time $t = 0$. Under no-arbitrage and by the law of one price we must have that:

$$\Pi_0(g) = V_0(\theta) = \frac{1}{1 + r_f} \frac{ug(d) - dg(u)}{u - d} + \frac{1}{s} \frac{g(u) - g(d)}{u - d} \cdot s \tag{3.16}$$

With some basic algebra one can rewrite it as:

$$\Pi_0(g) = \frac{1}{1 + r_f} \left[\underbrace{\frac{(1 + r_f) - d}{u - d}}_{q_u} \cdot g(u) + \underbrace{\frac{u - (1 + r_f)}{u - d}}_{qd} \cdot g(d) \right] \tag{3.17}$$

Let us now focus on q_u and q_d.
Under no-arbitrage condition we know that $d < (1 + r_f) < u$, so $q_u, q_d > 0$. Furthermore it is easy

to see that $q_u + q_d = 1$.
We could interpret q_u and q_d as new probabilities for $\{Z = u\}$ and $\{Z = d\}$ under a new probability measure $\mathbb{Q}$ called **Risk-neutral probability**.
This new probability measure $\mathbb{Q}$, is different from the **Statistical probability** $\mathbb{P} = \{p_u, p_d\}$ defined at the beginning. $\mathbb{Q}$ is said to be a risk-neutral measure, while $\mathbb{P}$ is said to be a physical, real-world, measure.
Where does the name **Risk-neutral** come from?
Once one defines the risk-neutral probabilities q_u and q_d, the previous formula becomes:

$$\Pi_0(g) = \frac{1}{1+r_f}[q_u \cdot g(u) + qd \cdot g(d)] \tag{3.18}$$

which is nothing else than an *expectation* under the new probability measure $\mathbb{Q}$:

$$\Pi_0(g) = \frac{1}{1+r_f}\mathbb{E}^{\mathbb{Q}}[g(Z)] \tag{3.19}$$

The name risk-neutral comes exactly form this fact: under the assumption of this new probability measure $\mathbb{Q}$ the price of the derivative can be seen as a discounted expected value of the payoff, which is independent of investors' attitude towards risk.
This feature will be resumed later on, in the Black-Scholes-Merton model price valuation.

Going back to the general case of a multi-period binomial model, given a derivative with payoff $g(Z_1, ..., Z_N$, its arbitrage-free risk-neutral price is:

$$\Pi_0(g) = \frac{1}{(1+r_f)^N}\mathbb{E}^{\mathbb{Q}}[g(Z_1, ..., Z_N)] \tag{3.20}$$

We can consider two types of payoff function now:

- A **path-dependent payoff** which depends on the precise succession of "up" and "down" moves.
- A **path-independent payoff** which only depends on the total number of "up" and "down" moves, and not on their distribution.

Under the risk-neutral probability measure $\mathbb{Q}$, the distribution of the total number of "up" and "down" moves, without taking into account their succession, is given by the **binomial distribution**.
The probability of having n "up" and $(N-n)$ "down" can be written as:

$$\mathbb{Q}(\{n \text{ "up" and } (N-n) \text{ "down"}\} = \frac{N!}{n!(N-n)!}q_u^n q_d^{N-n} \tag{3.21}$$

Consequently, giving a *path-independent* derivative with payoff $g(Z_1, ..., Z_N) = g(s\prod_{i=1}^N Z_i)$, its arbitrage-free risk-neutral valuation price $\Pi_0(g)$ is given by:

$$\Pi_0(g) = \frac{1}{(1+r_f)^N}\mathbb{E}^{\mathbb{Q}}[g(s\prod_{i=1}^N Z_i)] \tag{3.22}$$

$$= \frac{1}{(1+r_f)^N} \cdot \sum_{n=0}^N g(s \cdot u^n \cdot d^{N-n})\frac{N!}{n!(N-n)!}q_u^n q_d^{N-n} \tag{3.23}$$

where:

- $\sum_{n=0}^{N}$ indicates that we are summing over all possible payoff scenarios that depend on the number of "up" and "down".
- $\frac{N!}{n!(N-n)!}$ is the binomial coefficient, usually written as $\binom{N}{n}$. It indicates the number of paths that leads to the *same* payoff.
- $q_u^n q_d^{N-n}$ indicates the risk neutral probability of having n "up" moves and $N-n$ "down" moves.
- $g(s \cdot u^n \cdot d^{N-n})$ indicates the payoff associated to each path.

As we have seen in this section, the binomial model is a *discrete time* model, but what happens if we consider a *continuous time* evaluation? The Black-Scholes-Merton model represents a suitable answer to this question.

3.2 The Model

The Black–Scholes model is a continuous time model that assumes the financial market to consist of at least one risky asset, usually called the stock, and one risk-free asset, usually called the bank account.

- **Bank Account**: it is a risk-free asset with a continuously compounded risk-free rate $r > 0$. Its value comes from the solution of the following ordinary differential equation:

  $$dB_t = B_t r dt \tag{3.24}$$

  with $B_0 = 1$. The solution of such ordinary differential equation is:

  $$B_t = e^{rt} \quad \text{for all} \quad t \in (0, T) \tag{3.25}$$

  Notice that here the absence of risk is reflected by the absence of Brownian motion, which makes the value of B_t a deterministic function of time.
  Note. *r and r_f defined before are the same thing, we changed notation in order to have a sloppier one.*

- **Stock**: it is a risky asset, whose value comes from the solution of the following stochastic differential equation of a geometric Brownian motion:

  $$dS_t = S_t \mu dt + S_t \sigma dW_t \tag{3.26}$$

  with $S_0 = s > 0$. In the latter, μ is the **drift** and σ is the **volatility**.
  Notice that the asset is risky because its evolution is random (due to the Brownian motion). Knowing that $(W_t)_{t\in(0,T)}$ is a Markov process, also the the price $(S_t)_{t\in(0,T)}$ of the risky asset is a Markov process.
  Remark. *A Markov process is a stochastic process in which the transition probability to a specific state of the system depends only on the immediately preceding state, and not on how we get into this preceding state.*
  Exploiting Itˆo's formula, one can find the solution of such stochastic differential equation:

  $$S_t = s e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad \text{for all} \quad t \in (0, T) \tag{3.27}$$

  As one can easily see, the risky asset price S_t has a log-normal distribution. All the properties of S_t have been described in Chapter 2.

Finally we list here the market assumptions on which the Black-Scholes model is based:

- No arbitrage opportunity: there is no way to make a risk-free profit.
- It is always possible to borrow and lend any amount, even fractional, of cash at the risk-free rate.
- It is always possible to buy and sell any amount, even fractional, of the stock, this includes short selling.
- The above transactions do not incur any fees or costs.

3.3 From Hedging to Pricing

As already said the Black-Scholes-Merton model is a continuous-time model, and this gives us the possibility of continuously rebalancing our portfolio over a certain time period $[0, T]$.

Definition 17. ***(Dynamic portfolio)*** *A* ***Dynamic portfolio*** *in continuous time is a couple $\theta = (\theta^B, \theta^S)$, where the two components $\theta^B = (\theta_t^B)_{t\in[0,T]}$ and $\theta^S = (\theta_t^S)_{t\in[0,T]}$ are continuous-time stochastic processes. At every time $t \in [0, T]$, the value of the portfolio θ is given by:*

$$V_t(\theta) = \theta_t^B B_t + \theta_t^S S_t \tag{3.28}$$

Here, θ_t^B can be interpreted as the number of risk-free asset units held at time t, while θ_t^S can be interpreted as the number of risky asset units held at time t.
For a multi-period binomial model we have seen that at generic time t, the self-financing condition can be written as:

$$\theta_{t-1}^B B_t + \theta_{t-1}^S S_t = \theta_t^B B_t + \theta_t^S S_t \tag{3.29}$$

The value of a portfolio can be in general written as:

$$\begin{aligned}
V_t(\theta) &= V_0(\theta) + \sum_{k=1}^{t} \Delta V_k(\theta) &(3.30)\\
&= V_0(\theta) + \sum_{k=1}^{t} (V_k(\theta) - V_{k-1}(\theta)) &(3.31)\\
&= V_0(\theta) + \sum_{k=1}^{t} (\theta_k^B B_k + \theta_k^S S_k - \theta_{k-1}^B B_{k-1} - \theta_{k-1}^S S_{k-1}) &(3.32)\\
&= V_0(\theta) + \sum_{k=1}^{t} (\theta_{k-1}^B B_k + \theta_{k-1}^S S_k - \theta_{k-1}^B B_{k-1} - \theta_{k-1}^S S_{k-1}) &(3.33)\\
&= V_0(\theta) + \sum_{k=1}^{t} \theta_{k-1}^B \Delta B_k + \sum_{k=1}^{t} \theta_{k-1}^S \Delta S_k \quad \forall t = 1...N &(3.34)
\end{aligned}$$

where we used the self-financing condition.
Suppose that we have N intervals $[k, k-1]$. If we let $N \to \infty$ we can replace summations with integrals, obtaining:

$$V_t(\theta) = V_0(\theta) + \int_0^t \theta_u^B dB_u + \int_0^t \theta_u^S dS_u \tag{3.35}$$

The last term $\int_0^t \theta_u^S dS_u$ must be considered as a stochastic integral because of the presence of the Brownian motion inside the definition of S_t.
The latter can be written in its infinitesimal form as a stochastic differential equation:

$$dV_t(\theta) = \theta_t^B dB_t + \theta_t^S dS_t \tag{3.36}$$

inserting the dynamics:

$$\begin{cases} dB_t = B_t r dt \\ dS_t = S_t \mu dt + S_t \sigma dW_t \end{cases} \tag{3.37}$$

we finally obtain:

$$dV_t(\theta) = (\theta_t^B B_t r + \theta_t^S S_t \mu)dt + \theta_t^S S_t \sigma dW_t \tag{3.38}$$

$$= \underbrace{(\theta_t^B B_t r + \theta_t^S S_t r)}_{V_t(\theta) r} dt + \theta_t^S S_t(\mu - r)dt + \theta_t^S S_t \sigma dW_t \tag{3.39}$$

$$= V_t(\theta) r dt + \theta_t^S S_t \sigma (\frac{\mu - r}{\sigma} dt + dW_t) \tag{3.40}$$

which is the self-financing condition for a continuously rebalanced portfolio.

Definition 18. ***(Payoff)*** *The* ***payoff*** *of a path-independent European derivative with maturity T written on the risky asset is defined as $G(S_t)$, where $G : \mathbb{R}_+. \to \mathbb{R}$*

Let us now focus on the initial question: *Does there exist a self-financing dynamic portfolio such that $V_T(\theta) = G(S_T)$ with probability 1?.*
Recalling that $(S_t)_{t\in[0,T]}$ is a Markov process, the price of an European derivative with payoff $G(S_T)$ is given by a function of t and the current price S_t of the underlying. Let us denote $\Pi_t(G)$ the price of the derivative at time t for $t \in [0, T]$, it is reasonable to assume that:

$$\Pi_t(G) = F(t, S_t) \quad \text{for all} \quad t \in [0, T] \tag{3.41}$$

for some function $F : [0, T] \times \mathbb{R}_+ \to \mathbb{R}$.

We are ready to determine a self-financing portfolio $(\theta_t)_{t\in(0,T)}$ composed by $\theta_t = (\theta_t^B, \theta_t^S)$, such that $V_T(\theta) = G(S_T)$ with probability 1.
Knowing that for self-financing dynamic portfolio holds:

$$dV_t(\theta) = (\theta_t^B B_t r + \theta_t^S S_t \mu)dt + \theta_t^S S_t \sigma dW_t \tag{3.42}$$

and knowing that if we apply Itô's formula to a certain $F(t, S_t)$ representing the price of the derivative, function of time t and underlying price S_t, we get:

$$dF(t, S_t) = (F_t(t, S_t) + F_x(t, S_t)S_t\mu + \frac{1}{2}F_{xx}(t, S_t)S_t^2\sigma^2)dt + F_x(t, S_t)S_t\sigma dW_t \tag{3.43}$$

where $F_t(t, S_t) = \frac{\partial F(t,S_t)}{\partial t}$, $F_x(t, S_t) = \frac{\partial F(t,S_t)}{\partial S_t}$, $F_{xx}(t, S_t) = \frac{\partial^2 F(t,S_t)}{\partial S_t^2}$.

Remark. (Law of One Price): *Let X and Y be two random variables representing the payoffs at some maturity T of two portfolios and let us denote by Π_t^X and Π_t^Y their prices at time $t \in [0, T]$. Suppose that $\mathbb{P}(X = Y) = 1$, if the no-arbitrage principle holds, then:*

$$\Pi_t^X = \Pi_t^Y \tag{3.44}$$

Hence, if two portfolios yield the same payoff at maturity, then their market values should always coincide.
This means that if $V_T(\theta) = G(S_T)$ at the maturity T, than the hedging portfolio and the derivative should have the same value also at each generic time t.

We are led to conclude that $dV_t(\theta) = dF(t, S_t)$.
Comparing (3.42) with (3.43) we are able to determine θ_t^B and θ_t^S:

$$\theta_t^S = F_x(t, S_t); \qquad \theta_t^B = \frac{F_t(t, S_t) + \frac{1}{2}F_{xx}(t, S_t)S_t^2\sigma^2}{B_t r} \tag{3.45}$$

Again, from the law of one price we must have that $V_t(\theta) = F(t, S_t)$:

$$V_t(\theta) = \theta_t^B B_t + \theta_t^S S_t = \frac{F_t(t, S_t) + \frac{1}{2}F_{xx}(t, S_t)S_t^2\sigma^2}{B_t r} B_t + F_x(t, S_t)S_t = F(t, S_t) \tag{3.46}$$

which, with some basic algebra, leads to the famous **Black-Scholes partial differential equation (PDE)**:

$$\begin{cases} F_t(t, S_t) + \frac{1}{2}F_{xx}(t, S_t)S_t^2\sigma^2 + F_x(t, S_t)S_t r - rF(t, S_t) = 0 \\ F(T, S_T) = G(S_T). \end{cases} \tag{3.47}$$

By solving this equation we are able to determine the pricing function $F(t, S_t)$ of the derivative we are dealing with.

3.4 Pricing of Derivatives

For some payoff functions $G(S_T)$ the Black-Scholes PDE can be explicitly solved. When an explicit solution cannot be found, one can apply numerical schemes in order to obtain an approximate solution for $F(t, S_t)$.
We recall here the **Black-Scholes PDE**:

$$\begin{cases} F_t(t, S_t) + \frac{1}{2}F_{xx}(t, S_t)S_t^2\sigma^2 + F_x(t, S_t)S_t r - rF(t, S_t) = 0 \\ F(T, S_T) = G(S_T) \end{cases} \tag{3.48}$$

One can notice that, by construction, the drift parameter μ does not appear in the equation, and this is crucial.
The fact that the Black-Scholes-Merton PDE is independent of μ means that does not involve risk preferences. The drift parameter represent the expected return on the stock, the higher the level of investors' risk aversion is, the higher μ will be for any given stock.
Since μ is not relevant, one can substitutes it with the risk-free rate r in the stochastic differential equation of the risky asset, but if the risky asset has drift r, then the market participants are risk-free. If the market participants are risk-free, we are living in a *Risk-neutral World*, then prices are given by discounted expectations of future payoffs: **Risk-neutral Valuation**.

Definition 19. ***(Feynman-Kac formula)***: *Let the function $F : [0, T] \times \mathbb{R}_+ \to \mathbb{R}$ be the solution of the Black-Scholes partial differential equation:*

$$\begin{cases} F_t(t, S_t) + \frac{1}{2}F_{xx}(t, S_t)S_t^2\sigma^2 + F_x(t, S_t)S_t r - rF(t, S_t) = 0 \\ F(T, S_T) = G(S_T) \end{cases} \tag{3.49}$$

Then, the function $F(t, S_t)$ satisfies:

$$F(t, S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[G(S_T)|S_t] \tag{3.50}$$

where $\mathbb{Q}$ is a probability measure under which $(S_t)_{t\in[0,T]}$ satisfies the following stochastic differential equation:

$$dS_t = S_t r dt + S_t \sigma dW_t^{\mathbb{Q}} \quad \text{with} \quad S_0 = s \tag{3.51}$$

where we replaced μ with r. $(W_t^{\mathbb{Q}})_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}$.

Once again we can appreciate that with the assumption of a new probability measure $\mathbb{Q}$ under which μ is replaced with r, the price of the derivative can be seen as a discounted expected value of the payoff, which is independent of investors' attitude towards risk.
Let us now focus on the **risk-neutral probability** $\mathbb{Q}$.
Until now we have seen that the stochastic differential equation for S_t can appear in two different ways:

$$dS_t = S_t \mu dt + S_t \sigma dW_t \quad \text{under } \mathbb{P} \tag{3.52}$$

$$dS_t = S_t r dt + S_t \sigma dW_t^{\mathbb{Q}} \quad \text{under } \mathbb{Q} \tag{3.53}$$

where $W_t = (W_t)_{t\in[0,T]}$ is the Brownian motion under the probability $\mathbb{P}$, while $W_t^{\mathbb{Q}} = (W_t^{\mathbb{Q}})_{t\in[0,T]}$ is the Brownian motion under the risk-neutral probability $\mathbb{Q}$.
Clearly, the two quantities must have the same value, then:

$$S_t \mu dt + S_t \sigma dW_t = S_t r dt + S_t \sigma dW_t^{\mathbb{Q}} \tag{3.54}$$

From this equation we can determine the relationship between the two different Brownian motions:

$$dW_t^{\mathbb{Q}} = \frac{\mu - r}{\sigma} dt + dW_t \tag{3.55}$$

where the factor $\frac{\mu-r}{\sigma}$ is called **market price of risk**.
Observe that in our model we are not assuming that agents are risk-neutral, the formula only says that the price of derivatives is calculated **as if** we live in a risk neutral world. Agents are allowed to have any attitude to risk, but prices remain **preference free**, namely independent of risk attitudes.
Note. *For the change of probability measure we exploited the* ***Girsanov theorem****.*

3.5 Black-Scholes pricing formula for Call & Put options

In this subsection we present a particular example of pricing formula with its complete derivation. This will be useful later on, when we will present the pricing formula for *Credit Default Swaps*.
Let us now consider a **Call option** with maturity T and strike price K. For the sake of completeness, we recall the definition of a Call option:

Definition 20. ***(Call option)*** *A Call option with maturity T and strike price K is a financial derivative instrument which gives the right to buy one unit of the underlying at maturity T at a pre-fixed price K. The corresponding payoff is:*

$$G(S_T) = max[S_T - K; 0] = (S_T - K)^+ \tag{3.56}$$

Now, we fix an arbitrary time $t \in [0, T]$, by risk-neutral valuation we have that:

$$\begin{aligned} \Pi_t^{\text{Call}} &= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+ | S_t] && (3.57) \\ &= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[(S_T - K)\mathbb{I}_{(S_T > K)} | S_t] && (3.58) \\ &= e^{-r(T-t)} (\underbrace{\mathbb{E}^{\mathbb{Q}}[S_T \mathbb{I}_{(S_T>K)} | S_t]}_{\mathbf{B}} - K \underbrace{\mathbb{E}^{\mathbb{Q}}[\mathbb{I}_{(S_T>K)} | S_t]}_{\mathbf{A}}) && (3.59) \end{aligned}$$

Let us evaluate the two pieces separately:

- **A.** From the definition of expected value we get:

$$\mathbb{E}^{\mathbb{Q}}[\mathbb{I}_{(S_T>K)}|S_t] = 1 \times \mathbb{Q}(S_T > K|S_t) + 0 \times \mathbb{Q}(S_T \leq K|S_t) \tag{3.60}$$
$$= \mathbb{Q}(S_T > K|S_t) \tag{3.61}$$

Recalling that the Geometric Brownian Motion $dS_t = S_t r dt + S_t \sigma dW_t^{\mathbb{Q}}$ has solution:

$$S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t^{\mathbb{Q}}} \tag{3.62}$$

then:

$$S_T = S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})} \tag{3.63}$$

Replacing this inside the conditional probability we get:

$$\mathbb{Q}(S_T > K|S_t) = \mathbb{Q}(S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})} > K|S_t) \tag{3.64}$$
$$= \mathbb{Q}(\log S_t + (r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}) > \log K|S_t) \tag{3.65}$$
$$= \mathbb{Q}(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}} > \frac{1}{\sigma}[\log \frac{K}{S_t} - (r - \frac{\sigma^2}{2})(T-t)]|S_t) \tag{3.66}$$
$$= \mathbb{Q}\left(\frac{W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}}{\sqrt{T-t}} > \frac{\log \frac{K}{S_t} - (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}|S_t\right) \tag{3.67}$$

By recalling the properties of Brownian motion, we can state that the left side of the inequality follows a Gaussian distribution, namely $\frac{W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}}}{\sqrt{T-t}} \sim N(0,1)$. Renaming it as $Z := \frac{W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}}}{\sqrt{T-t}}$, and defining a new variable $d_2 := \frac{-\log \frac{K}{S_t} + (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$ in order to have a more compact notation, we get:

$$\mathbb{Q}(S_T > K|S_t) = \mathbb{Q}(Z > -d_2) = \mathbb{Q}(Z < d_2) \tag{3.68}$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{-\frac{Z^2}{2}} dZ =: \Phi(d_2) \tag{3.69}$$

$\Phi(d_2)$, which is nothing else than the probability that the random variable Z takes value below d_2, represents the risk-neutral probability of exercising the Call option.

- **B.** Starting from the solution of Geometric Brownian Motion we get:

$$\mathbb{E}^{\mathbb{Q}}[S_T \mathbb{I}_{(S_T>K)}|S_t] = \mathbb{E}^{\mathbb{Q}}[S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})} \cdot \mathbb{I}_{[S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})} > K]}|S_t] \tag{3.70}$$
$$= S_t e^{r(T-t)} \mathbb{E}^{\mathbb{Q}}[e^{-\frac{\sigma^2}{2}(T-t)+\sigma(W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}})} \cdot \mathbb{I}_{\left[\frac{W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}}}{\sqrt{T-t}} > -d_2\right]}|S_t] \tag{3.71}$$

As before, we can introduce a random variable $Z := \frac{W_T^{\mathbb{Q}}-W_t^{\mathbb{Q}}}{\sqrt{T-t}}$ which follows a gaussian distribution. Furthermore, we can remove the conditioning with respect to S_t because the Brownian motion $W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}$ is independent of S_t, hence:

$$\mathbb{E}^{\mathbb{Q}}[S_T \mathbb{I}_{(S_T>K)}|S_t] = S_t e^{r(T-t)} \mathbb{E}^{\mathbb{Q}}[e^{-\frac{\sigma^2}{2}(T-t)+\sigma\sqrt{T-t}Z} \cdot \mathbb{I}_{[Z>-d_2]}] \tag{3.72}$$

$$= S_t e^{r(T-t)} \int_{-\infty}^{+\infty} e^{-\frac{\sigma^2}{2}(T-t)+\sigma\sqrt{T-t}Z} \cdot \mathbb{I}_{[Z>-d_2]} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} dZ \tag{3.73}$$

$$= S_t e^{r(T-t)} \int_{-d_2}^{+\infty} e^{-\frac{\sigma^2}{2}(T-t)+\sigma\sqrt{T-t}Z-\frac{Z^2}{2}} \frac{1}{\sqrt{2\pi}} dZ \tag{3.74}$$

$$= S_t e^{r(T-t)} \int_{-d_2}^{+\infty} e^{-\frac{1}{2}(Z-\sigma\sqrt{T-t})^2} \frac{1}{\sqrt{2\pi}} dZ \tag{3.75}$$

Note. *In the above derivation we have exploited the fact that, by definition, the expectation value of a normally distributed random variable X is:*

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f(x^2) dx \quad \textit{with probability density function} \quad f(x^2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{3.76}$$

Making a change of variable in the integral: $y := Z - \sigma\sqrt{T-t}$, we get:

$$\mathbb{E}^{\mathbb{Q}}[S_T \mathbb{I}_{(S_T>K)}|S_t] = S_t e^{r(T-t)} \int_{-d_2-\sigma\sqrt{T-t}}^{+\infty} \underbrace{e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} dy}_{\text{density function of a normal distribution} \sim N(0,1)} \tag{3.77}$$

Notice that the integral is nothing else than the probability that a normally distributed random variable takes values above $-d_2 - \sigma\sqrt{T-t}$.
By symmetry arguments, this is equal to the probability that a normally distributed random variable takes values below $d_2 + \sigma\sqrt{T-t}$:

$$\int_{-d_2-\sigma\sqrt{T-t}}^{+\infty} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} dy = \int_{-\infty}^{d_2+\sigma\sqrt{T-t}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} dy =: \Phi(d_2 + \sigma\sqrt{T-t}) \tag{3.78}$$

Finally, we obtain:

$$\mathbb{E}^{\mathbb{Q}}[S_T \mathbb{I}_{(S_T>K)}|S_t] = S_t e^{r(T-t)} \Phi(\underbrace{d_2 + \sigma\sqrt{T-t}}_{d_1}) \tag{3.79}$$

which represents the value at time t of one unit of the underlying in case of exercise.

Summing the two contributions **A** and **B**, we end up with the pricing formula for a Call option:

$$\Pi_t^{Call} = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2) \tag{3.80}$$

Recalling the **Put-Call parity** relation:

$$\Pi_t^{Call}(K) - \Pi_t^{Put}(K) = S_t - K P_t(T) \quad \text{for all} \quad t \in [0,T] \tag{3.81}$$

one is able to find the pricing formula for a Put option:

$$\Pi_t^{Put}(K) = \Pi_t^{Call}(K) - S_t + K e^{-r(T-t)} \tag{3.82}$$

$$= S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2) - S_t + K e^{-r(T-t)} \tag{3.83}$$

$$= K e^{-r(T-t)} (1 - \Phi(d_2)) - S_t (1 - \Phi(d_1)) \tag{3.84}$$

$$= K e^{-r(T-t)} \Phi(-d_2) - S_t \Phi(-d_1) \tag{3.85}$$

where in the last step we exploited the symmetry properties of a normal distribution.
Similarly as before, $\Phi(-d_2)$ represents the risk-neutral probability of exercising the Put option, while $S_t \Phi(-d_1)$ represents the value at time t of one unit of the underlying in case of exercise.

3.6 Pro & Cons of the Black-Scholes Model

Pro

- **Provides a Framework**: The Black-Scholes model provides a theoretical framework for pricing options. This allows investors and traders to determine the fair price of an option using a structured and defined methodology that has been tested.
- **Allows for Risk Management**: By knowing the theoretical value of an option, investors can use the Black-Scholes model to manage their risk exposure to different assets. The Black-Scholes model is therefore useful to investors not only in evaluating potential returns but understanding portfolio weakness and deficient investment areas.
- **Allows for Portfolio Optimization**: The Black-Scholes model can be used to optimize portfolios by providing a measure of the expected returns and risks associated with different options. This allows investors to make smarter choices better aligned with their risk tolerance and pursuit of profit.
- **Streamlines Pricing**: On a similar note, the Black-Scholes model is widely accepted and used by practitioners in the financial industry. This allows for greater consistency and comparability across different markets and jurisdictions.

Cons

- **Limits Usefulness**: As stated previously, the Black-Scholes model is only used to price European options and does not take into account that American options could be exercised before the expiration date.
- **Lacks Cash-flow Flexibility**: The model assumes dividends and risk-free rates are constant, but this may not be true in reality. Therefore, the Black-Scholes model may lack the ability to truly reflect the accurate future cash-flow of an investment due to model rigidity.
- **Assumes Constant Volatility**: The model also assumes volatility remains constant over the option's life. In reality, this is often not the case because volatility fluctuates with the level of supply and demand.
- **Misleads Other Assumptions**: The Black-Scholes model also leverages other assumptions. These assumptions include that there are no transaction costs or taxes, the risk-free interest rate is constant for all maturities, short selling of securities with use of proceeds is permitted, and there are no risk-less arbitrage opportunities. Each of these assumptions can lead to prices that deviate from actual results.

3.7 Black Scholes Pricing Formula from Fokker-Planck Equation

The Fokker-Planck equation is an equation that describes the time evolution of the probability density of a Brownian particle.
Let us consider a single Brownian particle in one dimensional space. We define the *probability density function* for the particle to have a certain position x at a certain time t as $f(x,t)$. The probability to find the Brownian particle in the interval $(x, x+dx)$ at time t is $f(x,t)dx$. Let us now consider the following stochastic differential equation:

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \tag{3.86}$$

The solution of such a stochastic differential equation provides a description of the possible trajectories or paths that the Brownian particle can follow over time, each with a certain associated probability density.
The probability density associated with each of these possible trajectories over time is given by the solution of the **Fokker-Planck equation**:

$$\frac{\partial}{\partial t}f(x,t) = -\frac{\partial}{\partial x}\big(\mu(x,t)f(x,t)\big) + \frac{\partial^2}{\partial x^2}\big(D(x,t)f(x,t)\big) \tag{3.87}$$

where as we saw for Einstein's theory $D = \frac{\sigma^2}{2}$. The solution of the Fokker-Planck equation would provide a function that describes the probability of finding the Brownian particle in a specific state x at time t.
Let us now try to derive the Black-Scholes pricing formula for Call options starting from the Fokker-Planck equation instead of the Black-Scholes partial differential equation. Before, we solved Black-Scholes PDE, where we derived with respect to t and S_t, through Feynman-Kac formula in order to find the Call option pricing formula. Now we will find another partial differential equation, this time involving partial derivatives with respect to K and T, that if solved will lead us to the same Call option pricing formula.
As we already mention many times, stock price's trend is often modelled by the Geometric Brownian Motion, namely:

$$dS_t = S_t r dt + S_t \sigma dW_t \tag{3.88}$$

Its associated Fokker-Planck equation becomes:

$$\frac{\partial}{\partial t}f(S_t,t) = -\frac{\partial}{\partial S_t}\big(rS_t f(S_t,t)\big) + \frac{\partial^2}{\partial S_t^2}\left(\frac{\sigma^2}{2}S_t^2 f(S_t,t)\right) \tag{3.89}$$

Let us assume the Call option price at time t is determined, as always, by the present value of the risk-neutral expected payoff:

$$F(t,S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[max(S_T-K;0)|S_t] \tag{3.90}$$

Also consider $f(S_t,t,S_T,T)$ to be the transition probability density for the risk-neutral random walk where S_t is the asset price at time t. Being K the strike price and T the maturity, the price of the Call option at time t is given by:

$$F(t,S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[max(S_T-K;0)|S_t] \tag{3.91}$$

$$= e^{-r(T-t)}\int_0^{+\infty} max(S_T-K;0)f(S_t,t,S_T,T)dS_T \tag{3.92}$$

$$= e^{-r(T-t)}\int_K^{+\infty} (S_T-K)f(S_t,t,S_T,T)dS_T \tag{3.93}$$

Differentiating with respect to K we get:

$$\frac{\partial}{\partial K}F(t,S_t) = -e^{-r(T-t)}\int_K^{+\infty} f(S_t,t,S_T,T)dS_T \tag{3.94}$$

If we derive another time with respect to K, we obtain the following expression for the probability density function:

$$f(S_t,t,S_T,T) = e^{r(T-t)}\frac{\partial^2}{\partial K^2}F(t,S_t)\Big|_{S_T=K} \tag{3.95}$$

If we now derive the pricing formula $F(t, S_t)$ with respect to the maturity T, we get:

$$\frac{\partial}{\partial T}F(t, S_t) = -rF(t, S_t) + e^{-r(T-t)} \int_K^{+\infty} (S_T - K)\frac{\partial f(S_t, t, S_T, T)}{\partial T} dS_T \tag{3.96}$$

Inserting the Fokker-Planck equation:

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial S_T}(rS_T f) + \frac{\partial^2}{\partial S_T^2}\left(\frac{\sigma^2}{2}S_T^2 f\right) \tag{3.97}$$

where we have omitted the dependence on t and S_t in order to have a sloppier notation, we obtain:

$$\frac{\partial}{\partial T}F(t, S_t) = -rF(t, S_t) + e^{-r(T-t)} \int_K^{+\infty} (S_T - K)\left(-\frac{\partial}{\partial S_T}(rS_T f) + \frac{\partial^2}{\partial S_T^2}\left(\frac{\sigma^2}{2}S_T^2 f\right)\right) dS_T \tag{3.98}$$

Integrating by parts twice we get:

$$\frac{\partial}{\partial T}F(t, S_t) = -rV + \frac{1}{2}e^{-r(T-t)}\sigma^2 K^2 f + re^{-r(T-t)} \int_K^{+\infty} S_T f dS_T \tag{3.99}$$

After some algebra one can obtain:

$$\frac{\partial}{\partial T}F(t, S_t) = \frac{1}{2}\sigma^2 K^2 \frac{\partial^2}{\partial K^2}F(t, S_t) - rK\frac{\partial}{\partial K}F(t, S_t) \tag{3.100}$$

Once solved, this partial differential equation results in the same Call option pricing formula we obtained after solving the Black-Scholes partial differential equation through Feynman-Kac formula. This method is rarely used to find Call option prices, but is widely used to find σ, the stock volatility, knowing the pricing formula as a function of K and T.

Chapter 4

Greek Letters

References: [9] [1] [6]

A financial institution that sells an option to a client in the over-the-counter markets is faced with the problem of managing its risk. If the option is the same as one traded on an exchange, the institution can easily manage the risk buying the same option it has sold. But in case the option does not have a counterpart in the standardized products traded in the exchanges, manage its exposure is more difficult.
Some alternative solutions to this problem are called "Greek Letters". Each Greek Letter measure a different dimension to the risk in an option position and the aim of a trader is to manage the "Greeks" in such a way that the risk is acceptable.
Greek Letters are partial derivatives representing the sensitivity of the price of a derivative instrument (such as options) to changes in one or more underlying parameters on which the value of an instrument or a portfolio is dependent. The Greeks in the Black-Scholes-Merton model are relatively easy to calculate and they are particularly useful for hedging. The most common Greek letters are **delta**, **gamma**, **vega**, **theta**, and **rho**.

Delta Let us consider a derivative with maturity T with the risky asset as underlying, and let us denote by $F(t, S_t)$ its price in the Black-Scholes model, for $t \in (0, T)$.
Delta (Δ) is defined as the rate of change of the option price with respect to the price of the underlying asset. It represents the slope of the curve that relates the option price to the underlying asset price:

$$\Delta_t = \frac{\partial}{\partial S_t} F(t, S_t) \tag{4.1}$$

In the Black-Scholes model, for all $t \in (0, T)$, it holds that:

- For a *Call* option:

$$\Delta_t^{Call} = \frac{\partial}{\partial S_t} F^{Call}(t, S_t) \tag{4.2}$$

$$= \frac{\partial}{\partial S_t} (S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2)) \tag{4.3}$$

$$= \Phi(d_1) \tag{4.4}$$

- For a *Put* option:

$$\Delta_t^{Put} = \frac{\partial}{\partial S_t} F^{Put}(t, S_t) \tag{4.5}$$
$$= \frac{\partial}{\partial S_t}(F^{Call}(t, S_t) - S_t + Ke^{-r(T-t)}) \tag{4.6}$$
$$= \underbrace{\frac{\partial}{\partial S_t} F^{Call}(t, S_t)}_{=\Delta_t^{Call}} - \underbrace{\frac{\partial}{\partial S_t} S_t}_{=1} + \underbrace{\frac{\partial}{\partial S_t} Ke^{-r(T-t)}}_{=0} \tag{4.7}$$
$$= \Delta_t^{Call} - 1 \tag{4.8}$$
$$= \Phi(d_1) - 1 \tag{4.9}$$

by using the Put-Call parity relation.

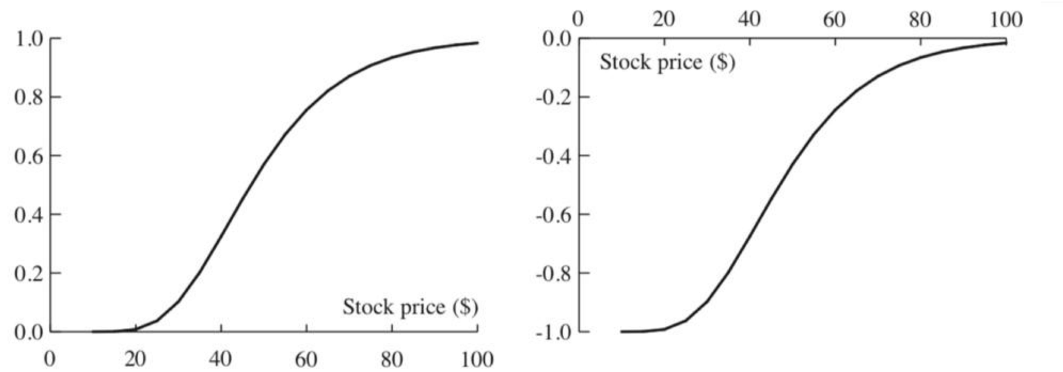


Figure 4.1: Delta of a Call and Put options with $K = 50$, $r = 0$, $\sigma = 0.25$ $T = 2$. [6]

Gamma Let us consider a derivative with maturity T with the risky asset as underlying, and let us denote by $F(t, S_t)$ its price in the Black-Scholes model, for $t \in (0, T)$.
Gamma (Γ) is defined as the rate of change of the portfolio's Delta with respect to the price of the underlying asset. It is the second partial derivative of the portfolio with respect to asset price:

$$\Gamma_t = \frac{\partial^2}{\partial S_t^2} F(t, S_t) = \frac{\partial}{\partial S_t} \Delta_t \tag{4.10}$$

In the Black-Scholes model, for all $t \in (0, T)$, it holds that:

$$\Gamma_t^{Call} = \Gamma_t^{Put} = \frac{\partial^2}{\partial S_t^2} F^{Call}(t, S_t) \tag{4.11}$$

and considering a Call option we have:

$$\Gamma_t^{Call} = \frac{\partial}{\partial S_t} \Phi(d_1) = \Phi'(d_1) \frac{\partial}{\partial S_t} d_1 \tag{4.12}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}} \frac{\partial}{\partial S_t} d_1 = \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}} \frac{1}{S_t} \frac{1}{\sigma\sqrt{T-t}} \tag{4.13}$$

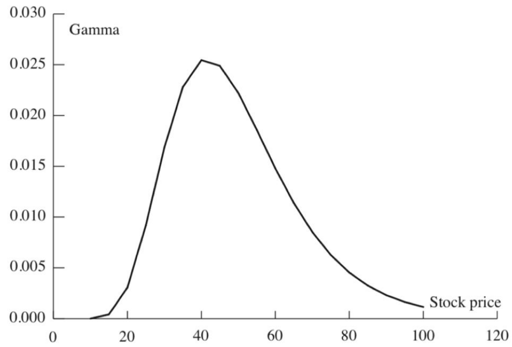


Figure 4.2: Gamma of a Call/Put option with $K = 50$, $r = 0$, $\sigma = 0.25$ $T = 2$. [6]

Theta Let us consider a derivative with maturity T with the risky asset as underlying, and let us denote by $F(t, S_t)$ its price in the Black-Scholes model, for $t \in (0, T)$.
Theta (Θ) is defined as the rate of change of the value of the portfolio with respect to the passage of time with all else remaining the same. Theta is sometimes referred to as the *time decay* of the portfolio:

$$\Theta_t = \frac{\partial}{\partial t} F(t, S_t) \tag{4.14}$$

Usually, when theta is quoted, time is measured in days, so that Theta is the change in the portfolio value when one day passes with all else remaining the same.
In the Black-Scholes model, for all $t \in (0, T)$, it holds that:

$$\Theta_t^{Call} = -\frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}} \frac{S_0 \sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(d_2) \tag{4.15}$$

$$\Theta_t^{Put} = -\frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}} \frac{S_0 \sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(-d_2) \tag{4.16}$$

Theta is usually negative for an option. This because, as time passes with all else remaining the same, the option tends to become less valuable. When the stock price is very low, Theta is close to zero.

Vega Let us consider a derivative with maturity T with the risky asset as underlying, and let us denote by $F(t, S_t)$ its price in the Black-Scholes model, for $t \in (0, T)$.
Vega (V) is defined as the rate of change of the value of the portfolio with respect to the volatility of the underlying asset:

$$V = \frac{\partial}{\partial \sigma} F(t, S_t) \tag{4.17}$$

If Vega is highly positive or highly negative, the portfolio's value is very sensitive to small changes in volatility. If it's close to zero, then volatility changes have relatively little impact on the value of the portfolio.
For a Call/Put option, Vega assumes the same form, which is:

$$V = S_0 \sqrt{T-t} \frac{1}{\sqrt{2\pi}} e^{\frac{-d_1^2}{2}} \tag{4.18}$$

Vega is typically expressed as the amount of money per underlying share that the option's value will gain or lose as volatility rises or falls by 1 percentage point. All options (both calls and puts) will gain value with rising volatility.

Rho Let us consider a derivative with maturity T with the risky asset as underlying, and let us denote by $F(t, S_t)$ its price in the Black-Scholes model, for $t \in (0, T)$.
Rho (ρ) is defined as the rate of change of the value of the portfolio with respect to the interest rate, it measures the sensitivity of the portfolio to a change in the interest rate when all else remains the same:

$$\rho = \frac{\partial}{\partial r} F(t, S_t) \tag{4.19}$$

In the Black-Scholes model, for all $t \in (0, T)$, it holds that:

$$\rho_t^{Call} = K(T-t)e^{-r(T-t)}\Phi(d_2) \tag{4.20}$$

$$\rho_t^{Put} = -K(T-t)e^{-r(T-t)}\Phi(-d_2) \tag{4.21}$$

Delta and Gamma Hedging Let us consider a portfolio with value $V(t, S_t)$, which depends only on one type of underlying, and suppose that we want to immunize it against changes in the underlying price. If the portfolio is already delta neutral then its value is not sensible to changes in underlying price, in formulas:

$$\Delta_V = \frac{\partial V}{\partial S_t} = 0 \tag{4.22}$$

but what if $\Delta_V \neq 0$?
The main idea is to add a certain number of units of an option or underlying itself to the portfolio in order to make it delta neutral. We denote the pricing function of the chosen derivative as $F(t, S_t)$, and the number of units of that derivative as x_F. At this point the value of the portfolio is given by:

$$P(t, S_t) = V(t, S_t) + x_F \cdot F(t, S_t) \tag{4.23}$$

In order to make this portfolio delta neutral we have to choose x_F in such a way $\frac{\partial P}{\partial S_t} = 0$, this is given by the following equation:

$$\frac{\partial V}{\partial S_t} + x_F \cdot \frac{\partial F}{\partial S_t} = 0 \tag{4.24}$$

which gives the solution:

$$x_F = -\frac{\Delta_V}{\Delta_F} \tag{4.25}$$

We provide here an example to make things more clear. Assume that we are a financial institution and we have sold a particular derivative with price F_{t,S_t}, we want to hedge it buying a certain amount of the underlying asset in order to have a **delta** equal to zero, we get the following equation:

$$\frac{\partial}{\partial S_t}[-F(t, S_t) + x_s] \cdot S_t = 0 \tag{4.26}$$

with solution:

$$x_s = \Delta_F = \frac{\partial F(t, S_t)}{\partial S_t} \tag{4.27}$$

which gives us the number of units of the underlying that we need to buy in order to manage the risk incoming from exercising this option. It is crucial to see that a delta hedge works well only for a short time interval. As time goes by, the value of S_t will change and so does the Δ. To have a well-hedged portfolio we should rebalance our portfolio as soon as delta changes.
Clearly, if we rebalance it often we will have a good hedging but we could face high transaction costs. For this reason we should take into account another greek letter, **gamma**.

Gamma is defined as the second derivative of the option price with respect to that underlying asset price, as well as the first derivative of the delta with respect to the underlying asset price: $\Gamma = \frac{\partial \Delta}{\partial S_t} = \frac{\partial^2 F(t,S_t)}{\partial S_t^2}$. If gamma is high we will have to rebalance often, while for low gamma we can keep our initial delta hedge for a longer period of time. In the end, we would like to form a portfolio both **gamma** and **delta neutral**.
In order to do this we need two different derivatives in our hedge, with pricing function $F(t, S_t)$ and $G(t, S_t)$. The value of the portfolio will be:

$$P(t, S_t) = V(t, S_t) + x_F \cdot F(t, S_t) + x_G \cdot G(t, S_t) \tag{4.28}$$

In order to have a portfolio gamma and delta neutral we need to solve the following system of equations:

$$\begin{cases} \Delta_V + x_F \cdot \Delta_F + x_G \cdot \Delta_G = 0 \\ \Gamma_V + x_F \cdot \Gamma_F + x_G \cdot \Gamma_G = 0 \end{cases} \tag{4.29}$$

Sometimes it is useful to exploit the fact that the stock itself has zero gamma, so we could choose a portfolio consisting in $V(t, S_t)$ and $F(t, S_t)$ gamma neutral (notice that in general this portfolio won't be delta neutral) and subsequently add the underlying stock in order to make it delta neutral, in this case the gamma neutrality will be maintained by the fact that the stock is gamma neutral. Formally we will have the following equation:

$$P(t, S_t) = V(t, S_t) + x_F \cdot F(t, S_t) + x_s \cdot S_t \tag{4.30}$$

and the following system has to be solved:

$$\begin{cases} \Delta_V + x_F \cdot \Delta_F + x_s = 0 \\ \Gamma_V + x_F \cdot \Gamma_F = 0 \end{cases} \tag{4.31}$$

The solutions are:

$$x_F = -\frac{\Gamma_V}{\Gamma_F} \tag{4.32}$$

$$x_s = \frac{\Gamma_V \Delta_F}{\Gamma_F} - \Delta_V \tag{4.33}$$

In an ideal world, traders working for a financial institution would be able to re-balance their portfolios very frequently in order to maintain all Greeks equal to zero. In practice this is not possible. When managing a large portfolio dependent on a single underlying asset, traders usually make Delta zero, or close, at least once a day by trading the underlying asset. Unfortunately, a zero Gamma and a zero Vega are less easy to achieve because it's difficult to find options or other nonlinear derivatives that can be traded in the volume required at competitive prices.
Maintaining Delta neutrality for a small number of options on an asset by trading daily is usually not economically feasible because of trading cost. But when a derivatives dealer maintains Delta neutrality for a large portfolio of options on an asset, the trading costs per option hedged are likely to be much more reasonable.

Chapter 5

Credit Risk

References: [9] [1] [8] [16]

In the following chapter we will consider an important risk for financial institutions: Credit Risk. Credit risk arises from the possibility that counterparties may default in repaying their debt. This chapter will discuss the different approaches to estimating the probability that a certain company will default and will explain the differences between risk-neutral and real world probabilities of default.

5.1 Estimating Default Probabilities from Historical Data

In our financial system, an important role is covered by Rating Agencies, such as Moody's, Standard & Poor's, and Fitch. They are institutions that assess the financial strength of companies and government entities, especially their ability to meet principal and interest payments on their debts. The rating assigned to a given debt shows an agency's level of confidence that the borrower will honor its debt obligations as agreed. Each agency uses unique letter-based scores to indicate if a debt has a low or high default risk and the financial stability of its issuer.
In the following table we provide the cumulative default rates (in %), over a 20-year period, of bonds that had a particular rating, which is emitted by Moody's in this specific case, at the beginning of the period.

Term (years):	*1*	*2*	*3*	*4*	*5*	*7*	*10*	*15*	*20*
Aaa	0.000	0.013	0.013	0.037	0.106	0.247	0.503	0.935	1.104
Aa	0.022	0.069	0.139	0.256	0.383	0.621	0.922	1.756	3.135
A	0.063	0.203	0.414	0.625	0.870	1.441	2.480	4.255	6.841
Baa	0.177	0.495	0.894	1.369	1.877	2.927	4.740	8.628	12.483
Ba	1.112	3.083	5.424	7.934	10.189	14.117	19.708	29.172	36.321
B	4.051	9.608	15.216	20.134	24.613	32.747	41.947	52.217	58.084
Caa–C	16.448	27.867	36.908	44.128	50.366	58.302	69.483	79.178	81.248

Figure 5.1: Average cumulative default rates from Moody's. [9]

The probability of a bond defaulting during a particular year can be calculated from the table, for example for a "Baa" rated bond the default probability during the second year is $0.495\% - 0.177\% = 0.318\%$.

As one can easily see, the default probability is an increasing function of time. This because the bond issuer is initially considered to be creditworthy, but as the time goes on the possibility that its financial health will decline grows. Viceversa, for bonds with low credit rating the default probability is a decreasing function of time. The reason behind this fact is that for this type of bonds the earlier years may be the most critical, the longer the issuer survives, the greater is the chance that its financial health improves.

Let us now define the cumulative default rate at a certain time t as $C(t)$, we can see that the *unconditional default probability* $P_{default}(\Delta t)$ during a certain time period $\Delta t = t_f - t_i$ is given by:

$$P_{default}(\Delta t) = C(t_f) - C(t_i). \tag{5.1}$$

For example the unconditional default probability for a "Caa" rating bond during the third year is given by $C(3) - C(2) = 36.908\% - 27.867\% = 9.041\%$.
Following the same reasoning, we can define the cumulative probability that a bond will survive by the end of a certain period of time as:

$$P_{survive}(t) = 100\% - C(t). \tag{5.2}$$

For example the cumulative probability of surviving at the end of the second year for a "Caa" rating bond is given by $100\% - 27.867\% = 72.133\%$.
Finally the probability that a bond will default during a certain time period **conditional** on no earlier default can be defined as:

$$P^{cond.}_{default}(\Delta t) = \frac{P_{default}(\Delta t)}{P_{survive}(t_i)} = \frac{C(t_f) - C(t_i)}{100\% - C(t_i)} \tag{5.3}$$

Let us consider a shorter period of time Δt. We define the **hazard rate** $\lambda(t)$ at time t as a time dependent coefficient such that $\lambda(t)\Delta t$ is the default probability between time t and $t + \Delta t$ conditional on no earlier default.
Remembering that $P_{survive}(t) = 100\% - C(t)$ is the cumulative probability that a bond will survive until time t, then the conditional default probability between time t and $t + \Delta t$ is:

$$P^{cond.}_{default}(\Delta t) = \frac{C(t+\Delta t) - C(t)}{100\% - C(t)} = \frac{P_{survive}(t) - P_{survive}(t+\Delta t)}{P_{survive}(t)} = \lambda(t)\Delta t \tag{5.4}$$

From which:

$$P_{survive}(t+\Delta t) - P_{survive}(t) = -\lambda(t)\Delta t P_{survive}(t) \tag{5.5}$$

Taking the limit $\Delta t \to 0$ we have:

$$\frac{dP_{survive}(t)}{dt} = -\lambda(t)P_{survive}(t) \tag{5.6}$$

Integrating and defining $Q(t)$ as the probability of default by time t, we finally have:

$$Q(t) = 1 - e^{-\int_0^t \lambda(\tau)d\tau} = 1 - e^{-\lambda^a(t)t} \tag{5.7}$$

where $\lambda^a(t)$ is the *average hazard rate* between time 0 and t.

5.2 Estimating Default Probabilities from Bond Yield Spreads

Another approach to estimate the default probabilities is to look at the bond yield spreads.
A bond yield spread is defined as the excess of the promised yield on the bond over the risk-free rate. The usual assumption is that the excess yield is a compensation for the possibility of default.

Remark. A **coupon bond** with maturity T, coupon rate c and payment frequency m is a security which pays one unit of money in T years and c/m amount of money every $1/m$ year. We call *maturity* the termination date of the contract and c is a proportion of the *principal amount*, which by convention is set equal to 1.
Let us call $P_0(T, c, m)$ the value today (at time $t = 0$) of a coupon bond, which depends on the maturity, the coupon rate, and the payment frequency. It is the sum of:

- **A series of coupon payments**. The amount of each coupon is constant and equal to c/m. In total there are mT coupon payments (m coupon every year), each taking place at time $t = k/m$ and discounted with the corresponding *zero rate* $R_0(k/m)$. $R_0(T)$ is the interest rate applied to an investment that starts today and ends in T years from now, without any intermediate payment.
- **A cash flow of the principal amount**, occurring at maturity T

In formula:

$$P_0(T, c, m) = \frac{c}{m} \sum_{k=1}^{mT} e^{-\frac{k}{m} \cdot R_0(\frac{k}{m})} + e^{-T \cdot R_0(T)} \tag{5.8}$$

As one could easily guess, $R_0(\frac{k}{m}) \neq R_0(T)$, because they are evaluated at different times. We may ask if there exists a synthetic indicator which measures the return on investing in a bond. One possibility is the **bond yield**.

Definition 21. ***(Bond Yield)*** *the bond yield is the* ***single*** *discount rate that, if applied to* ***all*** *cash flows of the bond , gives a bond value equal to its market price. Mathematically, the bond yield y associated to a coupon bond with price $P_0(T, c, m)$ is the solution to the equation:*

$$\frac{c}{m} \sum_{k=1}^{mT} e^{-\frac{k}{m} \cdot y} + e^{-T \cdot y} = P_0(T, c, m) \tag{5.9}$$

Suppose now that the bond yield spread for a T-year bond is $s(T)$ per year. This means the average loss rate on the bond between time 0 and time T should be approximately $s(T)$ per year. Another expression for the average loss rate is $\lambda^a(T)(1 - R)$, where R is the estimated *recovery rate* (the recovery rate is defined as the bond's market value a few days after a default, as a percent of its face value).
This means that it is approximately true that:

$$\lambda^a(T)(1 - R) = s(T) \tag{5.10}$$

or

$$\lambda^a(T) = \frac{s(T)}{1 - R}. \tag{5.11}$$

This approximation mimics the behaviour of $s(T)$ very well.

The method we have just presented for calculating default probabilities depends on the choice of a risk-free rate because bond yields spreads are calculated starting from its value.
Another extremely useful credit spread estimate is provided by **Credit Default Swap (CDS) spread** that does not depend on the risk-free rate chosen.
CDS spreads will be discussed in detail in the next chapter and will be crucial for the evaluation of a carbon risk factor, which is the main core of this essay.

5.3 Comparison of Default Probability Estimates

The default probabilities estimated from historical data are usually much less than those derived from bond yield spreads. Why do we see such big differences between those hazard rates?

Rating	*Historical hazard rate*	*Hazard rate from bonds*	*Ratio*	*Difference*
Aaa	0.04	0.60	17.0	0.56
Aa	0.09	0.73	8.2	0.64
A	0.21	1.15	5.5	0.94
Baa	0.42	2.13	5.0	1.71
Ba	2.27	4.67	2.1	2.50
B	5.67	8.02	1.4	2.35
Caa and lower	12.50	18.39	1.5	5.89

Figure 5.2: Average hazard rate per year (in %). [9]

Hazard rates implied from bond yield spreads are dependent on the choice of a risk-free rate, so they are risk-neutral estimates. They can be used to calculate expected cash flows in a risk-neutral world whenever there is a credit risk. The value of cash-flows is obtained exploiting a risk-neutral valuation by discounting the expected cash-flows at a risk-free rate. Treasury rates are usually used as risk-free rates.
Hazard rates implied from historical data are, instead, real-world default probabilities because they are calculated starting from real data from the past.
The different results shown in the table above arises directly from the difference between real world and risk-neutral default probabilities.
One possible reason advanced to explain this results could be that bond traders may be allowing for depression scenarios much worse than anything seen during the period covered by historical data. Anyway, the most important reason is that bonds prices depend on the financial health of a certain period: if the prices of bonds decrease (during a crisis for example), they become riskier. To bear the risk usually investors demand for a higher yields, resulting in an increased bond yield spreads and so an increased hazard rates.
At this stage it is natural to ask whether we should use real-world or risk-neutral default probabilities in the analysis of credit risk. The answer depends on the purpose of the analysis.
For valuing credit derivatives or estimating the impact of default risk on the pricing of instruments, risk-neutral default probabilities should be used. This because the analysis calculates the present value of future cash-flows and this should be independent from risk preferences (as shown for the Black-Scholes-Merton model) and this involves a risk-neutral valuation.
When carrying out analyses to calculate the potential future losses from defaults, real-world default probabilities should be used.

5.4 Merton Model for Default Probabilities

Until now we have estimated a company's real-world probability of default through credit ratings. Unfortunately, credit ratings are revised relatively infrequently, and this led some analysts to argue that equity prices can provide a more efficient way for estimating default probabilities.
The firm issues two classes of securities: asset and zero-coupon bond. The equity receives no dividends. The bonds represent the firm's debt obligation maturing at time T with principal value D. If at time T the firm's asset value exceeds the promised payment D, the lenders are paid the promised amount and the shareholders receive the residual asset value. If the asset value is lower than the promised payment, the firm defaults, the lenders receive a payment equal to the asset value, and the shareholders get nothing. The model assumes that a company will default if its assets' value at the maturity falls below the debt. Let us suppose, for simplicity, that the firm has only one asset. We define:

- V_0 as the value of the company's asset today.
- V_T as the value of the company's asset at maturity T.
- σ_V as the volatility of the company's asset (constant).
- D as the debt repayment due at maturity T.

When the debt matures on date T, if there is enough value in the firm to meet this payment, which means that $V_T > D$, debtholders will receive the full face value D, while equityholders receive the balance $V_T - D$. However, if the value of the firm's assets on date T is insufficient to meet the debtholders claims, which means $V_T < D$, the debtholders receive the total assets value, and the equityholders receive nothing. Thus, the amount D_T received by the debtholders on maturity T can be expressed with:

$$D_T = \begin{cases} D & \text{if } V_T \geq D \\ V_T & \text{otherwise} \end{cases} \tag{5.12}$$

The payoff received by the debtholders at the maturity T may also be expressed with the expression:

$$D_T = D - \max\{D - V_T; 0\} \tag{5.13}$$

where:

- The first term D represents the payoff from investing in a risk-free zero coupon bond maturing at time T with a face value D.
- The second term $\max\{D - V_T; 0\}$, is the payoff from a short position in a put option on the firm's assets with a strike price D and maturity date T.

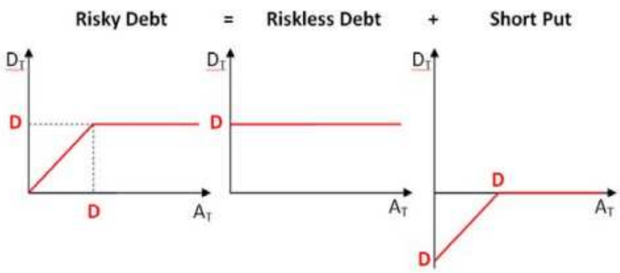


Figure 5.3: Decomposition of debt value at the time T. [8]

The decomposition illustrated above defines a procedure to valuate the **present value of risky debt**, consisting of two steps:

1. Identifying present value D of the risk-free debt.
2. Subtracting the present value of the Put option.

The first step of the procedure is straightforward, typically the formula of continuous compounding of interest is used, while the second step is clearly valuing the put option, for which we use the Black-Scholes option pricing model.

$$\Pi_0^{Put} = De^{-rT}\Phi(-d_2) - V_0\Phi(-d_1) \tag{5.14}$$

where

$$d_1 = \frac{\ln\frac{V_0}{D} + (r + \frac{\sigma_V^2}{2})T}{\sigma_V\sqrt{T}} \tag{5.15}$$

$$d_2 = d_1 - \sigma_V\sqrt{T}. \tag{5.16}$$

The value of the Put option determines the price differential between today's risky and risk-free value of the debt, so the market value of debt D_0 can be identified as:

$$D_0 = De^{-r \cdot T} - \Pi_0^{Put}. \tag{5.17}$$

The significant problem appearing while attempting a practical implementation of the Merton's model of debt valuation is that both the firm's asset value V_0 and its volatility σ_V are usually unobservable, but it is possible to use prices of traded securities issued by the firm to identify these quantities implicitly.
Let us define the **Equity** as the amount of money that would be returned to a company's shareholders if all of the assets were liquidated and all of the company's debt was paid off in the case of liquidation. Suppose that the firm is publicly traded with observable equity prices, we define:

- E_0 as the value of the company's equity today.
- E_T as the value of the company's equity at maturity T.
- σ_E as the volatility of the company's equity.

A company's equity, which is its assets' value minus its liabilities, can be considered as a call option on the value of its assets, with debt acting as a strike price:

$$E_T = (V_T - D)^+ = \begin{cases} V_T - D & \text{if } V_T \geq D \\ 0 & \text{otherwise} \end{cases} \tag{5.18}$$

If $V_T < D$ at the maturity T, it is reasonable that the company declares default. Then the value of the equity is zero.
If $V_T > D$ at the maturity T, the company should make the debt repayment, and the value of the equity is $V_T - D$.
Using the Black-Scholes-Merton pricing formula for call options, the value today of the equity is:

$$E_0 = V_0\Phi(d_1) - De^{-rT}\Phi(d_2) \tag{5.19}$$

where

$$d_1 = \frac{\ln\frac{V_0}{D} + (r + \frac{\sigma_V^2}{2})T}{\sigma_V\sqrt{T}} \tag{5.20}$$

$$d_2 = d_1 - \sigma_V \sqrt{T}. \tag{5.21}$$

From the Black-Scholes-Merton pricing formula we got that $\Phi(d_2)$ is the risk-neutral probability of exercising the call option at the maturity. In our case this is translated into the probability that, at time T, $V_T > D$, which means no default. By the parity property of the Gaussian distribution, the probability of not exercising the call option at the maturity ($V_T < D$), which is a synonymous of default, will be $\Phi(-d_2)$.
To calculate this, we require V_0 and σ_V. As we already said, none of them is directly observable, however if the company is traded we can observe E_0.
Then, from the Itô's formula, one can derive the following relation:

$$\sigma_E E_0 = \Phi(d_1)\sigma_V V_0 \tag{5.22}$$

where σ_E can be estimated from historical data, and $\Phi(d_1)$ is the probability that the random variable Z takes values below d_1.
Finally solving the system of two equations with two variables:

$$\begin{cases} E_0 = V_0\Phi(d_1) - De^{-rT}\Phi(d_2) \\ \sigma_E E_0 = \Phi(d_1)\sigma_V V_0 \end{cases} \tag{5.23}$$

one can get the values V_0 and σ_V for the calculation of the risk-neutral default probability $\Phi(-d_2)$.
The Merton's model provides a good ranking of default probabilities. Then, we can assume $\Phi(-d_2)$, which is a risk-neutral default probability calculated from an option pricing model, as the best instrument for providing a real-world default probability.

Let us define D_0 as the market price of the debt at time zero. The value of the asset is equal to total value of the two sources of financing: equity and debt, so that the present value of the debt can be expressed as:

$$D_0 = V_0 - E_0 \tag{5.24}$$

Substituting what we have found, we obtain:

$$D_0 = V_0 - V_0\Phi(d_1) + De^{-rT}\Phi(d_2) \tag{5.25}$$

$$= V_0(1 - \Phi(d_1)) + De^{-rT}\Phi(d_2) \tag{5.26}$$

$$= V_0\Phi(-d_1) + De^{-rT}\Phi(d_2) \tag{5.27}$$

We could ask ourselves which could be the yield to the maturity for the risky debt. By definition we have:

$$D_0 = De^{-y\cdot T} \tag{5.28}$$

where we have used D instead of D_T because the present value of the risky debt must be calculated starting from the hypothetical bond debt D, not the real one D_T which will be effectively repaid.
Comparing the two expressions we can get the yield for the risky debt:

$$De^{-y\cdot T} = V_0\Phi(-d_1) + De^{-rT}\Phi(d_2) \tag{5.29}$$

$$y = -\frac{1}{T}\ln\left(\frac{V_0}{D}\Phi(-d_1) + e^{-rT}\Phi(d_2)\right) \tag{5.30}$$

Finally, the **credit spread**, defined as the difference between the yield on the risky debt and the risk–free rate, can be simply calculated as:

$$s = y - r. \tag{5.31}$$

Chapter 6

Credit Derivatives

References: [9] [8] [1]

In this chapter we introduce an important development in derivatives market: Credit Derivatives. Credit derivatives are contracts where the payoff depends on the creditworthiness of one or more companies or countries. Here we will explain how credit derivatives work and how they are valued. Credit derivatives allow companies to trade credit risk. Before the introduction of these financial instruments, once a banks or other financial institutions had assumed a credit risk, they could just wait. Now they can manage their portfolios protecting themselves entering into credit derivative contracts. Banks have historically been the biggest buyers of credit protection and insurance companies have been the biggest sellers.
Credit derivatives can be categorized into two groups:

- **Single-name credit derivatives**: The most popular one is a Credit Default Swap. There are two sides to the contract: buyers and sellers of protection on a company/country; the payoff depends on the creditworthiness of the company/country and it is emitted in case the underlying defaults.
- **Multi-name credit derivatives**: The most popular is the CDO (Collateralized Debt Obligation) which is a portfolio of debt instruments with a complex structure where the cash-flows are channelled to different categories of investors.

This chapter will explain how credit default swaps work and how they are valued, finally we will introduce collateralized debt obligations.

6.1 Credit Default Swaps

The most popular credit derivative is a **Credit Default Swap**. It is a contract that provides insurance against the risk of a default by a certain company. The company is called *reference entity* and the default of it is called *credit event*.
The buyer of insurance obtain the right to sell bonds issued by the company for their face value when a credit event occurs (i.e., it has defaulted). The total face value of the bonds that can be sold is know as the credit default swap's *notional principal*.

Remark. *A bond (or obligation) is a credit title issued by a country, a financial institution, or a private company in order to accumulate liquidity.*
Whoever emits a bond is committed to returning the face value of the title at the maturity of the

bond, plus and additional amount of money, which is a fraction of the face value, pre-established by an interest rate. The quality of a bond is measured by rating agencies who emit a specific rate, indicator of its degree of solvency.

The buyer of a CDS makes periodic payments to the seller until the end of the life of the CDS or until a credit event occurs. These payments are typically made every 3 months and they are measured in *basis points* (1 basis points = 0.01% of the principal amount).
Let us make an example to illustrate how the deal is structured.
Example. Suppose that two parties enter into a 5-year credit default swap with notional principal of $100 million at 90 basis points per year for protection against default by the reference entity, with payments being made every 3 months.
Two scenarios can present:

- The reference entity does not default: in this case, the buyer pays 22.5 basis points each quarter of year until the maturity of the CDS. The amount payed every 3 months is the 0.225% of $100 million, so $225, 000, for a total of $0.9 million per year. At the maturity (5-years), the buyer does not receive any payoff and has payed $4.5 million.
- The reference entity defaults: there is a credit event.
  Suppose that the credit event occurs 2 months into the fourth year, the buyer has the right to sell bonds issued by the reference entity for $100 million. The final payment to close the 3 months cycle is required, but no further payments will be required. At the end, the buyer receives $100 million and has payed $3.825 million.

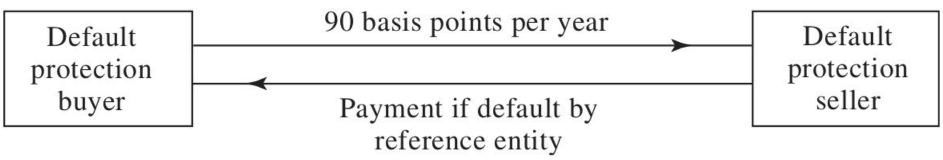


Figure 6.1: How a Credit Default Swap works. [9]

The total amount payed per year, as a percentage of the notional principal, to buy protection (90 basis points in our example) is called **CDS *spread***. This indicator will be crucial for the last part of this elaborate.

6.2 Hedging with Credit Default Swaps

Credit default swaps are often used to manage the risk of default that arises from holding debt. A bank, for example, may hedge its risk that a borrower may default on a loan by entering into a CDS contract as the buyer of protection. If the loan goes into default, proceeds from the CDS contract cancel out the losses on the underlying debt.
Suppose that an investor buys a 5-year corporate bond with yield 7% per year. For the sake of completeness we remind that a bond yield is the single discount rate that, if applied to all cash-flows of the bonds, gives a bond value equal to its market price. If we call the market price as $P_0(T, c, m)$, the bond yield is the solution of the following equation:

$$\frac{c}{m}\sum_{k=1}^{mT} e^{-\frac{k}{m}\cdot y} + e^{-T\cdot y} = P_0(T, c, m) \tag{6.1}$$

In order to manage the default risk, the investor can enter into a 5-year CDS with the same face value to buy protection against the issuer of the bond.
Suppose that the CDS has a CDS spread of 2% (i.e 200 basis points). If the bond issuer does not default, the investor earns 5% of the bond's face value per year, this because the bond yields of 7% is compensated by a CDS spread of 2%. If the bond does default, the investor earns 5% up to the time of the default, and then is allowed to sell the bond for its initial face value.
Doing this we simply turned our financial investment into a risk-free investment. In fact it is easy to see that buying a risky bond with yield 7% and managing its risk using a credit default swap with CDS spread 2% is exactly the same that investing into a riskless bond with risk-free rate at 5%. This shows that the spread of the yield on an n-year bond issued by a company over the risk-free rate should be equal to the company's n-year CDS spread. If it is not, arbitrage opportunities appear

$$\text{CDS spread} = y - r_f. \tag{6.2}$$

We finally define the **CDS *bond basis*** as the difference between the CDS spread and the bond yield spread:

$$\text{CDS bond basis} = \text{CDS spread} - \text{Bond yield spread} \tag{6.3}$$

In absence of arbitrage, the CDS bond basis should be as close as possible to zero.

6.3 Valuation of Credit Default Swaps

The CDS spread for a particular reference entity can be calculated from default probability estimates.
Suppose that the average hazard rate λ^a of the reference entity is 2% per year.
Recalling that the probability of default by time t is given by

$$Q(t) = 1 - e^{-\lambda^a(t)t} \tag{6.4}$$

then the probability of survival by time t can be written as

$$S(t) = 1 - Q(t) = e^{-\lambda^a(t)t}. \tag{6.5}$$

According to the fact that the probability of default during a year is the probability of survival to the beginning of the year ($S(t_{i-1})$) minus the probability of survival to the end of the year ($S(t_i)$), we can write:

$$Q(t_i) = S(t_{i-1}) - S(t_i) \tag{6.6}$$

where $Q(t_i)$ is the probability of default during the year t_i.
Remark. Before we defined the probability of default by the time t as $Q(t) = 1 - e^{-\lambda^a(t)t}$, now we are defining something different, $Q(t_i)$ gives us the probability of default during a single year i, not until the year i. This can be written as:

$$Q(t_i) = S(t_{i-1}) - S(t_i) = e^{-\lambda^a(t_{i-1})t_{i-1}} - e^{-\lambda^a(t_i)t_i} \tag{6.7}$$

Let us take a range of 5-years, knowing that the probability of surviving at time t is given by $e^{-0.02t}$ (average hazard rate of 2%), and again the probability of default during a year can be taken as the difference, we can define the following table:

Year	*Probability of surviving to year end*	*Probability of default during year*
1	0.9802	0.0198
2	0.9608	0.0194
3	0.9418	0.0190
4	0.9231	0.0186
5	0.9048	0.0183

Figure 6.2: Default probabilities and survival probabilities. [9]

Let us take a 5-years Credit Default Swap. We will make the following assumptions:

- Defaults always happen halfway through a year.
- Payments on the credit default swap are made once a year, at the end of the year.
- The risk-free interest rate is 5% per year with continuous compounding.
- The recovery rate is at 40%.

For a notional principal of \$1, indicating the CDS spread as s, and recalling that the probability of survival by the end of the year t is $S(t) = e^{-\lambda^a(t)t}$, the *Expected Payment* by the end of the year t is:

$$E_{payment}(t) = S(t) \cdot s \tag{6.8}$$

Therefore, its *Present Value* is nothing else than the Expected Payment discounted by a *Discount Factor* which in our case is:

$$D = Ne^{-r_f t} \tag{6.9}$$

where $r_f = 5\%$ is the risk-free rate, while $N = \$1$ is the notional principal.
We collected all the data in the following table.

Time (years)	*Probability of survival*	*Expected payment*	*Discount factor*	*PV of expected payment*
1	0.9802	0.9802*s*	0.9512	0.9324*s*
2	0.9608	0.9608*s*	0.9048	0.8694*s*
3	0.9418	0.9418*s*	0.8607	0.8106*s*
4	0.9231	0.9231*s*	0.8187	0.7558*s*
5	0.9048	0.9048*s*	0.7788	0.7047*s*
Total				4.0728*s*

Figure 6.3: Present value of expected payments. [9]

The total present value of expected payments $PV_{payments}$ is simply the sum of all the present value of the last column.
As already mentioned above, we set the recovery rate, which is the value of the bonds after the default, at 40%. This means that in case of default the owner of a CDS must be repaid for the 60% of the initial value of the notional principal. Furthermore, assuming that defaults always occur halfway through a year, the default probability during a certain year, which is given by the difference between two consecutive survival probabilities, is also the probability that a payoff occurs halfway through a year.

The owner's *Expected Payoff* at this time can be calculated as:

$$E_{payoff} = Q(t_i) \cdot (1 - R) \cdot N \tag{6.10}$$

where:

- $Q(t_i) = S(t_{i-1}) - S(t_i)$ is the default probability through a year.
- $(1 - R)$ is the percentage repayment of the notional principal.
- N is the notional principal (in our case equal to 1).

As before, the *Present Value* of the Expected Payoff can be thought as the Expected payoff multiplied by a Discount Factor $Ne^{-r_f t}$, but in this case t is a semi-integer number, because again we are assuming that defaults happen halfway through a certain year.
The following table resumes all the calculations:

Time (years)	*Probability of default*	*Recovery rate*	*Expected payoff ($)*	*Discount factor*	*PV of expected payoff ($)*
0.5	0.0198	0.4	0.0119	0.9753	0.0116
1.5	0.0194	0.4	0.0116	0.9277	0.0108
2.5	0.0190	0.4	0.0114	0.8825	0.0101
3.5	0.0186	0.4	0.0112	0.8395	0.0094
4.5	0.0183	0.4	0.0110	0.7985	0.0088
Total					0.0506

Figure 6.4: Present value of expected payoff. [9]

The total present value of expected payoff PV_{payoff} is simply the sum of all the present value of the last column.
As a final step, we consider also the accrual payment made in the event of a default. Considering that we have assumed that defaults only occur halfway through year, the accrual payment will always be $0.5s$, where s again is the CDS spread. Multiplying it by the probability of default halfway during a certain year ($Q(t_i)$), we obtain the *Expected Accrual Payment*:

$$E_{accrual} = 0.5s \cdot Q(t_i) \tag{6.11}$$

For its present value we just have to multiply it by a discounted factor in which t is semi-integer.
The following table resumes all the calculations:

Time (years)	*Probability of default*	*Expected accrual payment*	*Discount factor*	*PV of expected accrual payment*
0.5	0.0198	0.0099*s*	0.9753	0.0097*s*
1.5	0.0194	0.0097*s*	0.9277	0.0090*s*
2.5	0.0190	0.0095*s*	0.8825	0.0084*s*
3.5	0.0186	0.0093*s*	0.8395	0.0078*s*
4.5	0.0183	0.0091*s*	0.7985	0.0073*s*
Total				0.0422*s*

Figure 6.5: Present value of expected accrual payments. [9]

The total present value of expected accrual payment $PV_{accrual}$ is simply the sum of all the present value of the last column.
The **total** expected payments PV_{total} is given by the sum of the two payment contributions:

$$PV_{total} = PV_{payments} + PV_{accrual} \tag{6.12}$$

By equating the present value of the total expected payments (which is proportional to s as can be seen from the tables) and the present value of the expected payoff one can get the value of s which is the **CDS spread** for the 5-years CDS on a certain reference entity.

$$s = \frac{PV_{payoff}}{PV_{total}} \tag{6.13}$$

In this specific case we obtained $s = 0.0123$, that is to say 123 **basis points**.
Note that the previous calculation has been made assuming that defaults can happen only in the middle of the year, this gives us a good approximation, but for a better result more default times must be taken in consideration.

The default probabilities used to value a CDS should be risk-neutral default probabilities, not real-world probabilities, in this way we don not take into account risk-preferences in our valuation formula. Risk-neutral probabilities can be estimated as presented in Chapter 6.
Suppose we do not know the default probabilities, but we know the CDS spread for a 5-years CDS on a certain reference entity. We could reverse the process presented here, in order to obtain the implied average hazard rate per year.

6.4 CDS Forwards and Options

Once the CDS market was established, it was natural for derivatives dealers to trade forward and options on CDS spreads.

1. A **Forward Credit Default Swap** is the *obligation* to buy or sell a particular credit default swap on a reference entity at a certain future time T. The forward contract ceases to exist if the reference entity defaults before the maturity of the contract.

2. A **Credit Default Swap Option** gives the *right* to buy or sell a particular credit default swap on a reference entity at a certain future time T. The option contract ceases to exist if the reference entity defaults before the maturity of the contract.

For example, a trader could negotiate the right to buy a 5-years protection on a company starting in 1 year for a predetermined value of CDS spread (in basis points) entering into a **Call** option. If the 5-year CDS spread for the company in 1 year turns out to be more than the predetermined value, the option will be exercised, otherwise it will not.
Similarly, a trader could negotiate the right to sell a 5-years protection on a company starting in 1 year for a predetermined value of CDS spread (in basis points) entering into a **Put** option. If the 5-year CDS spread for the company in 1 year turns out to be less than the predetermined value, the option will be exercised, otherwise it will not.

Chapter 7

Carbon Default Swap

References: [2] [10] [14] [4] [8]

In this last chapter we present a study case in which, using credit default swap spreads, we construct a carbon risk factor and we show that carbon risk affects firms' credit spread.
With carbon risk we refer to the negative impact of de-carbonization on corporations and equity portfolios. It results to be larger for European firms than North American ones and varies across industries, suggesting that the market recognizes where and which sectors are more favorable to a low-carbon economy transition.
As a consequence of this, from a credit risk point of view, lenders will demand more credit protection for those borrowers perceived to be more exposed to carbon risk.

7.1 Low-Carbon Transition

The transformation required to companies to achieve net-zero targets could generate sizable costs for those sector which are unprepared and deeply carbon-dependent. This could generate significant losses in firms' cash-flows, leading to an economic situation which could undermine their ability to repay debts and, in the most dramatic cases, which could lead to a high probability of default.
There are already many evidences that transition risk influences credit risk. With transition risk we mean the risk resulting from changes in climate policy, technology, and market sentiment during the adjustment to a low-carbon economy, and it could be codified looking to the firms' current carbon emission data.
It is easy to appreciate that the speed and the efficiency of this economic transition varies across sectors, countries, and time, and so does its associated credit risk.
In particular we find that the exposure to carbon risk:

- is more prominent in Europe than in North America;
- varies substantially across industries' sectors;
- is stronger during times in which the attention to climate change is particularly heightened;
- is more peculiar for shorter time horizons, especially in Europe, confirming that lenders expect faster carbon improvements in Europe because of a stricter net-zero policy. Faster changes mean larger costs, and higher firms' credit risk.

In the following chapters we will concentrate on the carbon component of transition risk, briefly called carbon risk, being it the most relevant one by far. Understanding the full impact of the tran-

sition will require the measurements of entire carbon profiles as well as firms' emissions reduction plans. We will provide all these data for different countries and sectors.
We will finally construct a market-implied, high-frequency, forward-looking proxy for carbon risk exposure, showing how the carbon risk affects firms' credit spread, namely the difference between the interest rate on the debt and the risk-free rate.
Although there is no comprehensive theoretical framework linking the low-carbon policy transition to credit dynamics, markets are already recognizing that carbon policies are causing relevant changes in global economies, and this manifests in increased default risk or lower asset values for firms that are pore exposed to transition risks.
Credit risk does not depend only on transition's costs, but also on firms' goals. A huge number of the world's biggest companies have committed to reach a low-carbon economy by setting emissions intensity targets or time limits for reaching net-zero emissions. Clearly, the failure to fulfill those self-imposed commitments may lead to reputational risks and subsequently become a credit risk. Equally, unambitious emissions reduction strategies might become a transition risk.
To conclude, we state that markets recognize that firms' low-carbon transition may occur at different times and different speeds, and this affects their valuation between the lenders.

7.2 The Strategy

The goal of this final part of the thesis is to develop a theoretical argument able to provide us a connection between carbon risk and credit risk. In order to achieve this goal we will use the Merton model which will illustrate us the effects of carbon transition costs on the credit spread: higher carbon costs imply higher default probabilities and higher credit spreads.
To construct a market-implied, high-frequencies and forward-looking carbon risk factor we will use the information contained in the daily Credit Default Swaps (CDS) spreads.
Let us remind briefly what a CDS, and a CDS spread is:
A CDS is a contract that provides insurance against the risk of a default by a certain company. The buyer of a CDS makes periodic payments to the seller until the end of the life of the CDS or until a credit event occurs. These payments are measured in basis points (1 basis points = 0.01% of the principal amount).
*The total amount payed, as a percent of the notional principal, to buy protection is called **CDS spread**. CDS spreads are usually calculated as the difference in value between the yield y on an n-year bond issued by a company and the risk-free rate:*

$$\textit{CDS spread} = y - r_f \tag{7.1}$$

We could ask ourselves why should we use CDS spreads to create a carbon risk factor. Credit Default Swaps offer several advantages over other commonly used credit risk indicators like corporate bonds or ratings. We list the most important here below:

- Credit Default Swaps respond more quickly to changes in financial market conditions than any other debt and credit product. The reason behind this is that CDS are traded on standardized terms: pre-specified maturity, default event, and debt seniority.
- Credit Default Swaps are more liquid than corporate bonds
- Credit Default Swaps are contracts with high tenors, even up to 30 years, as a consequence they allow lenders to have long forward-looking considerations.

The carbon risk factor will be constructed as the difference between the daily median CDS spreads of high-emission-intensity firms, considered as polluting, and low-emission-intensity firms, consid-

ered as clean (emission intensity is a commonly used measure for firms carbon dependence). This difference is extremely useful because it works as an indicator of how the market perceives the differential exposure of polluting and clean firms to carbon risk.
When policy events (e.g. amendments or regulations) are emitted, lenders to more exposed firms may demand for more protection. At the same time lenders to less exposed firms may demand for less protection, and these opposite requests will widen the distance between the price of default protection, the CDS spread, for polluting and clean companies. This kind of difference is usually called CDS wedge. Conversely, a loosening of regulation will narrow the CDS wedge. Because of its construction, the carbon risk factor mimics a lending portfolio in which default protection is bought for a polluting firm and sold for a clean firm.

In the following analysis we will propose three hypothesis to study how, when, and where carbon risk affect more firms' creditworthiness. In order to do so, using daily CDS data, which reflects firms' exposure to carbon risk, we will investigate how firms' CDS spread returns change in response to the carbon risk factor. We will observe this for a large number of European and North American firms for a period from 2013 to 2019.
Subsequently, we will be able to show that under *ordinary conditions* carbon risk is a determinant of credit risk, and since the carbon risk factor reflects the collective expectation of carbon risk, an increase in the carbon risk is accompanied by more credit protection demanded by lenders.
To examine the effect of carbon risk under *extraordinary conditions*, namely when firms experience large shifts in their credit spreads, we will use quantile regressions. As we expected, we will find that the effect of carbon risk is amplified at the tails of the credit spread distribution.

In addition, further analyses to test for geographical and sectoral dependencies have been conducted.
For what concerns geographical dependencies, in Europe, where climate policies are more stringent, there is a very strong relationship between carbon risk exposure and the increase cost of default protection. The same effect is consistently weaker in North America, where climate policies are less stringent and more ambiguous. For example, if we consider a 5-years CDS, a one standard deviation increase in the carbon risk factor for an European firm leads to a rise on 15 basis points in CDS spread.
For what concerns the sectoral level, we will find out that, obviously, carbon-intensive sectors are more affected than less carbon-intensive industries. Limiting our attention to Europe, if we again consider a 5-years CDS, a one standard deviation increase in the carbon risk factor leads to a rise of 84 basis points in CDS spread for an energy firm; while the same increase in the carbon risk factor leads to an increase of 5 basis points in CDS spread for a healthcare firm. This suggest that market recognise which sectors are better positioned for a transition to a low-carbon economy.
Another important observation we will make is that the effect of carbon risk on CDS is stronger during times of heightened public attention to climate change. In fact, lenders appear to be more sensitive to carbon risk when market-wide concern about climate change risk is elevated.

Finally we will provide an analyses of the temporal dimension of carbon risk, extending our understanding of when carbon risk affects firms' creditworthiness. In particular, using information deriving from the CDS spread curve, we will show that a shift in the expected temporal materialization of carbon risk, known also as *carbon risk slope*, positively affects the steepness of the CDS curve slope. Also in this case we have geographical differences. In Europe, the effect on CDS spread is particularly salient for shorter time horizons, suggesting that the market perceives carbon risk to be a short/medium term risk. In North America we will not see the same effect, this is due to the weak and ambiguous current climate change policies.

To conclude, this last chapter will study the amplifying effect of a climate related transition on credit risk: changes induced by a transition to a net-zero economy will cause adjustments in firms' valuations which may contribute to the deterioration of firms' creditworthiness, translating to higher credit risk.
Resuming what we are going to analyze, we will show that:

- Firms with an emission-intensive business model have higher transition costs than low-carbon firms, which lead us to conclude that carbon risk is concentrated in specific sectors, like construction materials, fossil fuels, and utilities.
- Carbon risk can vary substantially across regions. This crucially depends on the local regulations and climate change policy.
- Carbon risk is also sector-specific. After all, decarbonizing the economy will involve large-scale structural changes. There will be sectors for which this transformation will be rapid and will not involve high transition costs, while other sectors that will have to entirely transform their technological basis, or alternatively disappear.
- Carbon risk perception continually changes as climate attentions evolve.

Using news and concern indexes related to carbon risk, we will empirically test in which situations lenders demand more credit protection, advancing our understanding of the effect of market awareness and carbon risk.

7.3 From Carbon Risk to Credit Risk

The risks related to the transition to a low-carbon economy arise from uncertainties regarding the characteristics and nature of the transformation, like the time frame of the transformation and the speed of greenhouse gas emissions reductions.
Since the transition path cannot be observed, but can only be inferred, it is not clear which proxies are the most appropriate.
To date, firms' exposure to carbon risk is almost always codified using firms' emission data. In particular, finance literature has focused on government policies and regulations to limit carbon emissions, approaching the pricing of carbon risk by focusing on how financial assets reflects market concerns about the above mentioned policies.
Measuring financial effects of carbon policies is intricate. Carbon policies can affect firms in multiple ways, both directly and indirectly. Here we list some of them:

- Carbon emissions are tied to fossil fuels, so carbon abatement regulations often translate into higher energy costs for firms.
- High energy prices translate into higher operating costs, so lower cash-flows.
- Firms may increase their investment in research an development to reducing operating costs in the future, but this will provide lower cash-flows in the present.
- Firms may incur greater costs through product modifications in response to changes in carbon restrictions and consumer preferences.

Without questions, these costs could significantly affect firms' cash-flows and financial health, undermining their capacity to generate profits to repay their debt, possibly leading them to higher probabilities of default.

Firms' transition to a low-carbon model can occur at different times and different speeds, depending on firms' size, sector, geographical location, and other characteristics. Basically, firms can face many different challenges, and carbon risk will affect them differently depending on how and where they do their business.
To illustrate how carbon risk is connected to credit risk, we use Merton's model for credit risk, let us briefly remind how this model acts.

Merton proposed a model where a company's value, which is its assets' value minus its liabilities, can be considered as a call option on the value of its assets, with debt acting as a strike price. The firm issues two classes of securities: equity and zero-coupon bond. The equity receives no dividends. The bonds represent the firm's debt obligation maturing at time T with principal value D. If at time T the firm's asset value exceeds the promised payment D, the lenders are paid the promised amount and the shareholders receive the residual asset value. If the asset value is lower than the promised payment, the firm defaults, the lenders receive a payment equal to the asset value, and the shareholders get nothing. The model assumes that a company will default if its assets' value at the maturity falls below the debt. Defining E_T, V_T, and D as the equity value, asset value and debt repayment, respectively, at the maturity, we get:

$$E_T = (V_T - D)^+ = \begin{cases} V_T - D & \text{if } V_T > D \\ 0 & \text{otherwise} \end{cases} \tag{7.2}$$

- *If $V_T < D$ at the maturity T, it is reasonable that the company declares the default. Then the value of the equity is zero.*
- *If $V_T > D$ at the maturity T, the company should make the debt repayment, and the value of the equity is $V_T - D$.*

Using the Black-Scholes-Merton pricing formula we get that the today value of the equity is:

$$E_0 = V_0\Phi(d_1) - De^{-rT}\Phi(d_2) \tag{7.3}$$

where

$$d_1 = \frac{\ln\frac{V_0}{D} + (r + \frac{\sigma_V^2}{2})T}{\sigma_V\sqrt{T}} \tag{7.4}$$

$$d_2 = d_1 - \sigma_V\sqrt{T} \tag{7.5}$$

Defining D_0 as the market price of the debt at time zero. The value of the asset is equal to total value of the two sources of financing: equity and debt, so that the present value of the debt can be expressed as:

$$D_0 = V_0 - E_0 \tag{7.6}$$

$$= V_0\Phi(-d_1) + De^{-rT}\Phi(d_2) \tag{7.7}$$

Note. *For a complete derivation see "**Merton Model for Default Probabilities**" in Chapter 6.*

Integrating carbon costs into Merton's model we get a theoretical foundation for a straightforward translation of carbon risk into credit risk.
In presence of carbon regulations, firms' cash-flows are reduced due to restrictions on carbon emissions. As a consequence, we introduce the **Carbon Tax Rate** and we label it δ: our working

assumption is that each firm, depending on their exposure to carbon risk, pays a certain amount δ per unit of time, where $0 < \delta < r$. Assuming that the value of the firm's asset V_t follows a Geometric Brownian Motion, the dynamics of the firm value are:

$$dV_t = V_t(r - \delta)dt + V_t \sigma dW_t^{\mathbb{Q}} \tag{7.8}$$

where $W^{\mathbb{Q}}$ is a Brownian motion under the risk-neutral probability measure $\mathbb{Q}$.
As we already did, under the risk-neutral probability measure $\mathbb{Q}$ we can use the risk-neutral pricing valuation: today's price of the equity can be seen as the discounted expectation value of Call option's future payoff. In our case:

$$E_0 = e^{-r\cdot T}\mathbb{E}^{\mathbb{Q}}[(V_T - D)^+] \tag{7.9}$$
$$= e^{-r\cdot T}\left[\mathbb{E}^{\mathbb{Q}}[V_T \cdot \mathbb{I}_{V_T>D}] - D\mathbb{E}^{\mathbb{Q}}[V_T \cdot \mathbb{I}_{V_T>D}]\right] \tag{7.10}$$

By evaluating separately the two pieces, exactly as we did in Chapter 3 but substituting r with $(r - \delta)$, we obtain:

$$\mathbb{E}^{\mathbb{Q}}[V_T \cdot \mathbb{I}_{V_T>D}] = V_0 e^{(r-\delta)T}\Phi(d_1) \tag{7.11}$$

$$\mathbb{E}^{\mathbb{Q}}[V_T \cdot \mathbb{I}_{V_T>D}] = \mathbb{Q}[V_T > D] =: \Phi(d_2) \tag{7.12}$$

where now:

$$d_1 = \frac{\ln\frac{V_0}{D} + (r - \delta + \frac{\sigma_V^2}{2})T}{\sigma_V\sqrt{T}} \tag{7.13}$$
$$d_2 = d_1 - \sigma_V\sqrt{T} \tag{7.14}$$

Putting all together we get the relation:

$$E_0 = e^{-r\cdot T}\mathbb{E}^{\mathbb{Q}}[(V_T - D)^+] \tag{7.15}$$
$$= e^{-r\cdot T}\left[\mathbb{E}^{\mathbb{Q}}[V_T \cdot \mathbb{I}_{V_T>D}] - D\mathbb{E}^{\mathbb{Q}}[V_T \cdot \mathbb{I}_{V_T>D}]\right] \tag{7.16}$$
$$= e^{-r\cdot T}\left[V_0 e^{(r-\delta)T}\Phi(d_1) - D\Phi(d_2)\right] \tag{7.17}$$
$$= V_0 e^{-\delta\cdot T}\Phi(d_1) - De^{-r\cdot T}\Phi(d_2) \tag{7.18}$$

Once again is important to remark the fact that:

- $\Phi(d_2) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} e^{-\frac{Z^2}{2}}dZ$ is nothing else than the probability that the random variable Z takes value below d_1, represents the risk-neutral probability of exercising the Call option.
- $V_0 e^{-\delta\cdot T}\Phi(d_1)$ is the value at time $t = 0$ of one unit of the asset in case of exercise, in which we take into account the losses due to the carbon tax rate.

The market price of the debt at time zero D_0 is equal to the sum of equity and debt, so its present value can be expressed as:

$$D_0 = V_0 - E_0 \tag{7.19}$$
$$= V_0 - V_0 e^{-\delta\cdot T}\Phi(d_1) + De^{-r\cdot T}\Phi(d_2) \tag{7.20}$$
$$= V_0(1 - e^{-\delta\cdot T}\Phi(d_1)) + De^{-r\cdot T}\Phi(d_2) \tag{7.21}$$
$$= V_0 e^{-\delta\cdot T}\Phi(-d_1) + De^{-r\cdot T}\Phi(d_2) \tag{7.22}$$

We define the yield on the firm's risky debt as the solution of the following equation:

$$D_0 = De^{-y \cdot T} \tag{7.23}$$

Equating (7.22) with (7.23) we obtain:

$$De^{-y \cdot T} = V_0 e^{-\delta \cdot T} \Phi(-d_1) + De^{-r \cdot T} \Phi(d_2) \tag{7.24}$$

Recalling that D is the face value at the maturity of the zero-coupon bond issued by the firm, defining D^* its value at time $t = 0$, then:

$$D = D^* e^{r \cdot T} \tag{7.25}$$

The **credit spread** at time $t = 0$ is defined as the difference between the yield on the firm's risky debt and the risk-free interest rate, namely $s = y - r$, so by inserting the (8.25) into (8.24) we can obtain the value of s, which is clearly dependent on δ:

$$D^* e^{(-y+r) \cdot T} = V_0 e^{-\delta \cdot T} \Phi(-d_1) + De^{-r \cdot T} \Phi(d_2) \tag{7.26}$$

$$s(\delta) = y - r = -\frac{1}{T} \ln \left(\frac{V_0}{D^*} e^{-\delta \cdot T} + \Phi(d_2) \right) \tag{7.27}$$

We can now express the conditional probability of default as a function of the carbon tax rate δ:

$$PD(\delta) := \mathbb{Q}(V_T < D) = \Phi(-d_2) \tag{7.28}$$

and observe that, when higher carbon-related costs materialize, firms may respond by increasing carbon risk, in fact:

$$\frac{\partial PD(\delta)}{\partial \delta} = \frac{\phi(-d_2)\sqrt{T}}{\sigma} > 0 \tag{7.29}$$

where $\phi(-d_2)$ is a certain function of d_2 after the derivation.

It is not difficult to understand that high-emitting firms (*polluting firms*, P) may incur greater costs compared to low-emitting firms (*clean firms*, C). This is due to the fact that the first ones have to suffer greater technological implementations, ending with the relation $\delta_C \leq \delta_P$.

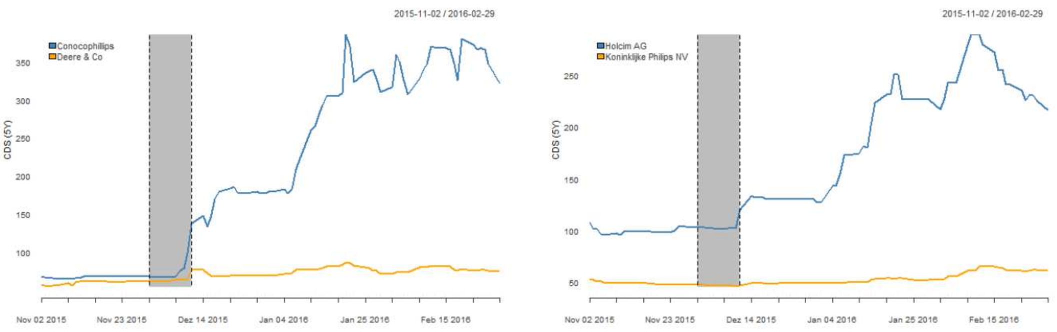
Combining the latter with the fact that default probabilities have a monotonic relation with the carbon tax rate, we find a link between the carbon risk exposure and the credit spread. The higher is the carbon tax rate, which reduces the firm's cash flow, the higher is the default probability, and consequently higher is the demand of protection by the lenders, increasing the credit spread.

7.4 Measuring Carbon Risk

Examining how the market perceives the firms' exposures to carbon risk requires a measurement of firms' carbon profiles. The latter is commonly proxied by firms' current emissions and emission intensity, although it should be supplemented by firm-specific future emissions reduction targets and strategies.
Once determined a theoretical relationship between carbon risk and credit spreads, we will analyze the changes in the credit spreads, which reflect the evolution of market's perception, to measure the carbon risk. In order to do this, we will utilize the information contained in the spreads of Credit Default Swaps contracts.
As we already mentioned, CDS contracts have three main advantages:

- CDS are traded on standardized terms (maturity, payoff, spread);
- CDS spreads respond quickly to changes in credit and market conditions;
- CDS are contracts with maturity up to 30-years, and this allows lenders to take into account forward-looking considerations on carbon risk within different time horizons.

Example. *In order to visualize what we have just said, we provide data on two polluting firms (ConocoPhillips and Holcim AG) and two clean firms (Deere & Company and Philips NV) in North America and Europe. For what concerns North American firms, ConocoPhillips is a multinational corporation ranked as the 21st World's Top 100 Polluters, while Deere & Company, the world's largest agricultural equipment manufacturer, has demonstrated leading practice in controlling and reducing their emissions in recent years. For what concerns European firms, Holcim AG is a global manufacturer of construction materials, considered as an emissions-intensive firm, while Philips NV is a healthcare company that boasts emissions reductions through an increased use of renewable energy.*



(a) North America (b) Europe

Figure 7.1: *Evolution of the 5Y-CDS spreads of ConocoPhillips (blue) and Deere & Co (orange) on the left diagram, and Holcim AG (blue) and Koninklijke Philips NV (orange) on the right diagram. The time period spans from 02 November 2015 to 29 February 2016. The gray-shaded area indicates the time period of COP21 (30th Nov 2015 – 12th Dec 2015) [2].*

In Figure 7.1, the CDS spreads for two pairs of companies before and after the Conference of the Parties in Paris in 2015 (culminated with the Paris Agreement) have been plotted. It illustrates that the difference in CDS spreads is approximately constant until the policy event. Post Paris Agreement, the spreads diverge. This could be interpreted as the result of lenders expecting higher carbon impacts for high-emitting firms. They demand more protection, namely more CDSs, over the more exposed firms, ultimately paying higher spreads.

In most studies, firms' exposure to carbon risk is codified using their emission intensity data and argues that high-emitting firms may incur greater costs from changes in policy and product changes in response to changes in consumer preferences. Clearly, the size of these costs and the consequent size of carbon risks are proportional to the size of firms' emissions, and their rate of growth.

Following this argument, we use the information contained in the CDS spreads to construct a proxy that reproduces firms' carbon risk.
Our approach to construct a carbon risk factor is based on how firms' exposure to carbon risk changes, and it can change because of:

- changes in lenders' expectations about the carbon exposure of different firms;
- changes in lenders' perception of carbon risk for a specific firm over time.

Specifically, we partition the universe of firms into different groups according to their emission intensity profile, then we subdivide them into quintiles.

Definition 22. *A quintile is a statistical value of a data set that represents* 20% *of a given population, so the first quintile represents the lowest fifth of the data (*1% *to* 20%*); the second quintile represents the second fifth (*21% *to* 40%*) and so on.*

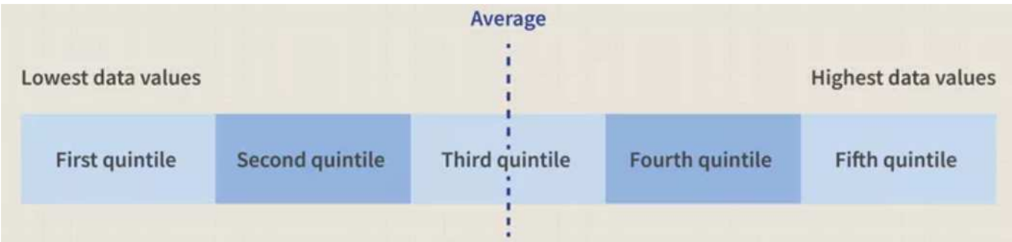


Figure 7.2: *Quintile representation.*

We then define firms below the first quintile as **clean** and we collect their CDS spreads in the set:

$$\gamma_t^m = \{c_{1,t}^m; c_{2,t}^m; ...; c_{n,t}^m\} \tag{7.30}$$

Analogously, we define firms above the last quintile as **polluting** and we collect their CDS spreads in the set:

$$\pi_t^m = \{p_{1,t}^m; p_{2,t}^m; ...; p_{n,t}^m\} \tag{7.31}$$

where:

- m is the tenor (maturity) of CDS contract, in our case we selected $m = \{1, 3, 5, 10, 30\}$.
- t is the in which we are looking at the data.
- $\{1, 2, ...n\}$ is the number of firms we are taking into account: equal for both the set because we divided our sample into quintiles and this implies the same number of events for each quintile.

We obtain the **median cost of default protection** of clean and polluting firms by calculating the median m-years CDS spread level for each tenor m at every time t

$$C_t^m = Med(\gamma_t^m) \tag{7.32}$$

$$P_t^m = Med(\pi_t^m) \tag{7.33}$$

Finally, we introduce the **Carbon Risk Factor (CR)** for a given tenor m at time t as the difference between the median CDS spreads of polluting and clean firms:

$$CR_t^m = P_t^m - C_t^m. \tag{7.34}$$

This difference represents the differential credit risk exposure of polluting versus clean firms, essentially it mimics the dynamics of a portfolio in which default protection is bought for a representative polluting company and sold for a representative clean company.

Note. *We are calling it "representative" because we are considering sets of different firms' CDS spreads and we are calculating its median, this is similar to have a single firm with a single CDS spread equal to the median value.*

When policy events occur, the demand for protection of more (less) exposed firms increases (decreases), widening the CDS wedge, which is the distance between price of default protection for polluting and clean companies. Conversely, if the market expects a loosening of regulation, the is a narrowing of the wedge.

These changes in perceived exposure to carbon risk are represented by the behavior of CR. We are allowed to consider CR as a reliable proxy for lenders' perception of carbon risk exposure.

To illustrate the relevance of CR, in Figure 7.3 and Figure 7.4 we display the evolution of CR over time, for tenors of 1, 5, and 30 years for the universe of CDS of 34 (35) clean (polluted) firms listed in Europe and 82 (73) clean (polluting) firms listed in North America.

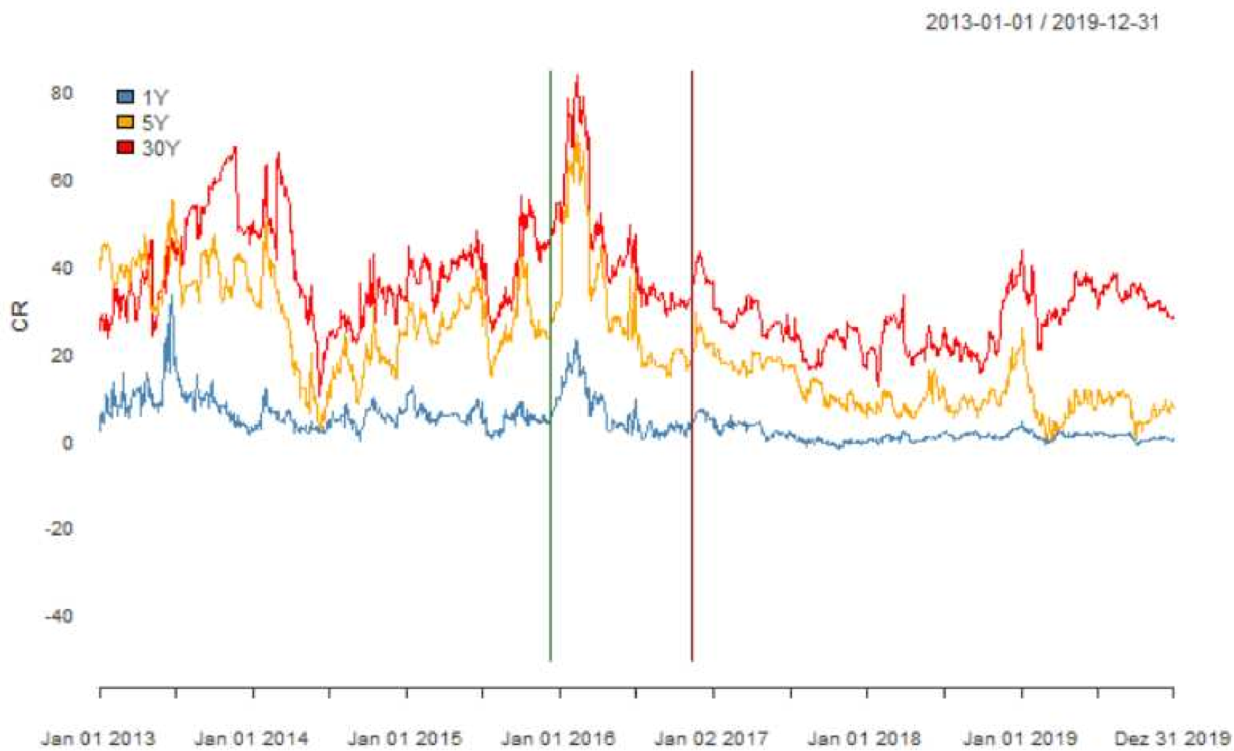


Figure 7.3: Evolution of CR over time for tenors of 1 (blue), 5 (orange) and 30 (red) years for Europe. The vertical lines refer to the Paris Agreement (dark green) and Trump election (brown), respectively [2].

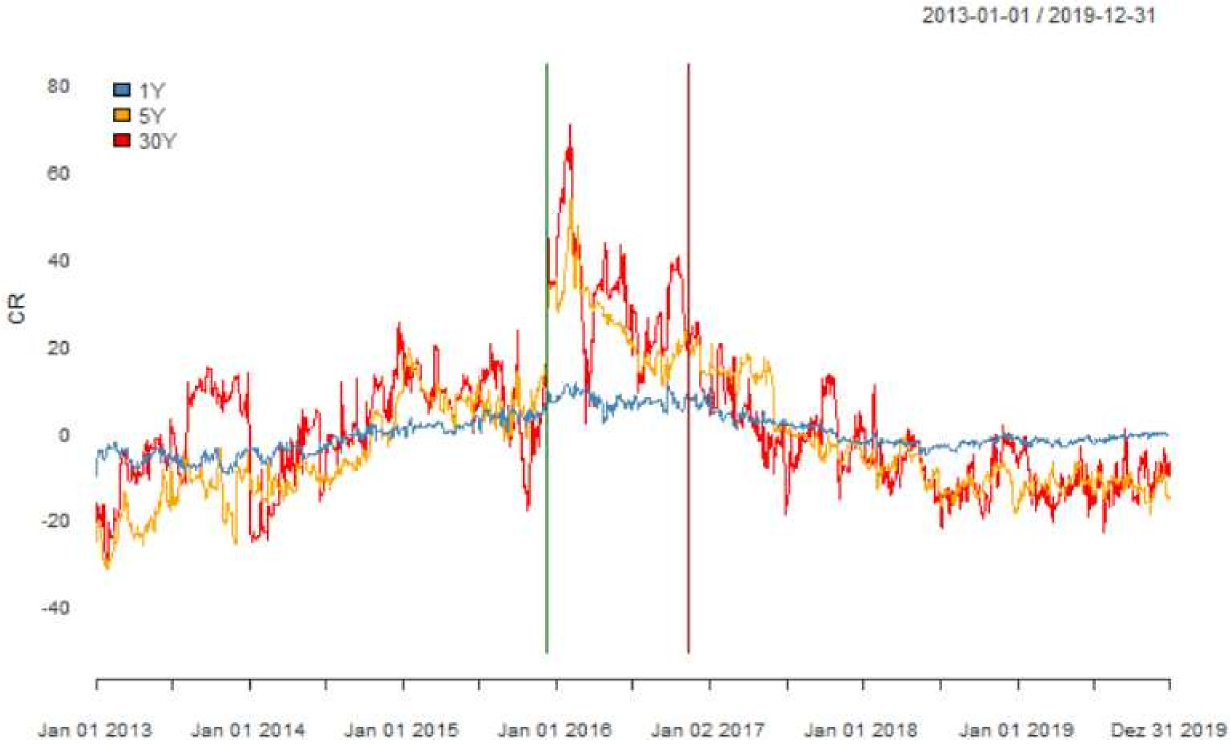


Figure 7.4: Evolution of CR over time for tenors of 1 (blue), 5 (orange) and 30 (red) years for North America. The vertical lines refer to the Paris Agreement (dark green) and Trump election (brown), respectively [2].

We first examine the European case, and we observe that all CR time series (which we will call **CRs**) are non-negative: lenders continually demand more (less) protection for European firms that are perceived to be more (less) exposed to carbon risk. In this case the polluting-minus-clean credit protection portfolio, constructed using CR, would have delivered a positive premium. Furthermore, the graph reflects changes in lenders' demand for default protection in response to policy-relevant events, such as COP21, which called for more ambitious policies and plans to reduce emissions. It is reasonable to argue that policies following this event can increase expected costs for firms that are less prepared for a transition to a low-carbon economy, and benefit firms that are more adequately prepared.
Observing the North American case, we can appreciate that the situation is clearly different. Until the COP21, all CRs were continuously swinging between positive and negative values, denoting an unclear situation in which the carbon behaviour of firms seems to not affect their creditworthiness, probably due to a less rigid carbon regularization. Only after the COP21 the CRs turn positive, indicating a surge in perceived exposure to carbon risk. However, this trend reverts almost immediately after the election of Donald Trump, famous for being a climate change denier, indicating that this event is associated with a decline in carbon risk. As we can see this election impacted principally North American CRs, reflecting the limited effect of US climate policy on European firms. As one can see, the trend of the trajectory for 30-years tenor is very unregular if compared to the one for 1-year tenor. This reflect higher market's concerns about new policies and regulations for the future.

Another valuable information about carbon risk can be extracted by considering the difference between a long and a short tenor CR over a specific time horizon t, we call it $CRslope_t^{mn}$:

$$CRslope_t^{mn} = CR_t^m - CR_t^n \quad \text{with} \quad m > n \tag{7.35}$$

This difference constitutes the slope of the CR factor. Conceptually, starting from a carbon risk exposure for year n, $CRslope_t^{mn}$ provides information describing how the exposure to carbon risk will be perceived over the remaining $m - n$ years.
$CRslope_t^{mn}$ can take take positive and negative values, depending on how the market's perception of carbon risk evolves. A positive $CRslope_t^{mn}$ reflects expectations of a stricter carbon regulatory framework in the next $m - n$ years, while a negative $CRslope_t^{mn}$ reflects expectations of a more permissive carbon regulatory framework.

Figure 7.5: Evolution of CR slope over time for 5-1 years (blue) and 30-5 years (orange) for Europe. The vertical lines refer to the Paris Agreement (dark green) and Trump election (brown), respectively [2].

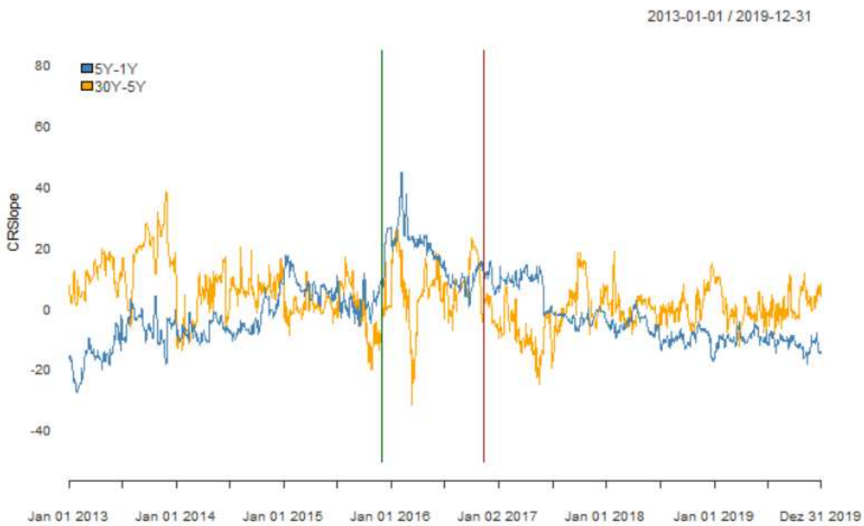


Figure 7.6: Evolution of CR slope over time for 5-1 years (blue) and 30-5 years (orange) for North America. The vertical lines refer to the Paris Agreement (dark green) and Trump election (brown), respectively [2].

The plots in Figure 7.5 and Figure 7.6 again suggest distinct conditions for Europe and North America. The CR slope is always positive in Europe, indicating a perception of continuously growing exposure to carbon risk: the longer the time horizon, the larger the perceived exposure to carbon risk in Europe. Conversely, the perceived future exposure to carbon risk in North America varies continuously and is less clear.

Hypothesis We end up this section listing a theory-motivated series of testable hypotheses:

- **Hypothesis 1** *There is a positive relationship between carbon risk and CDS spread returns*
  In the current section, we argued that CR represents the general perception of carbon risk exposure, such that a higher CR corresponds to a higher perceived carbon risk. We also noticed that a firm with high exposure to carbon risk can see a decline in its valuation, a higher probability of default, and a higher CDS spread.
- **Hypotheses 2** *The effect of carbon risk on CDS spread returns is stronger in Europe than in North America*
  We could appreciate that carbon risk differs across regions due to the different environmental regulations and restrictions on carbon emissions. While Europe has generally been considered a leader in the implementation of strict carbon policies, North American countries result to be more permissive. Consequently, one would expect firms in Europe to face higher costs. Applying this to the Merton model, we are allowed to assume that $\delta_{EU} > \delta_{NA}$, yielding higher expected CDS spreads for firms located and operating in Europe.
- **Hypotheses 3** *The effect of carbon risk on CDS spread returns is stronger during times of heightened attention to climate change*
  Climate policies continually evolve within a rapidly changing social and policy environment. As new information arrives in the market, lenders update their expectations accordingly. Specifically, during times of heightened attention to climate change in the news, lenders will demand more credit protection, thus increasing CDS spreads.

7.5 Data

In the following section we describe our CDS data, we introduce the variables to control for the effects of known determinants of CDS spread return, in order to isolate the impact of carbon risk on CDS spreads, and we report some summary statistics.

7.5.1 Credit Default Swap Spreads

We obtained [2] CDS spread data from Refinitiv for the period January 1, 2013 - December 31, 2020. The dataset covers CDS spreads for tenors of 1, 3, 5, 10, and 30 years for European and North American companies. Each CDS was valuated in US dollars.
We adjusted our data sample excluding:

1. Firms that defaulted during the sample period;
2. Firms that exhibited illiquid CDSs, that are contracts for which no spread movement is recorded for a minimum of 245 consecutive training days;
3. The year 2020 from our sample, in order not to have possible distorting effects due to the COVID-19 pandemic;
4. Financial firms from the sample because of their special business model.

Conversely, we retain firms with large CDS spreads. In total our sample contains 137 European firms and 281 North American firms.

In order to have a more compact range of values, we took into account the log of CDS spread, in particular we calculated the daily CDS spread log returns as:

$$s_{i,t}^{m} = \log\Big(CDS_{i,t}^{m}\Big) - \log\Big(CDS_{i,t-1}^{m}\Big) \tag{7.36}$$

where $CDS_{i,t}^{m}$ is the m-year CDS spread of firm i at day t.
$s_{i,t}^{m}$ quantifies the daily relative change in a firm's CDS spread. The relative change consents a straightforward comparison of credit improvement, or deterioration, across all firms.

7.5.2 Control Variables

To isolate the impact of carbon risk on CDS spreads, we employ a series of control variables that have been identified as determinants of CDS spreads. They can be divided into two categories:

- Firm-specific variables: stock return and stock volatility.
- Market-specific variables: general market conditions, interest rates, and term structure of interest rates,

Remark.(*control variable*) *Experiments attempt to assess the effect of manipulating one or more independent variables on one or more dependent variables. To ensure the measured effect is not influenced by external factors, other variables must be held constant. The constant variables during an experiment are referred to as control variables.*
By controlling these variables, we can isolate the carbon risk effect on firms' default probabilities.

Stock Return: It is calculated as the difference of the log of daily stock prices:

$$r_{i,t} = \log\big(S_{i,t}\big) - \log\big(S_{i,t-1}\big) \tag{7.37}$$

where $S_{i,t}$ denotes the stock price of firm i at time t. Stock return is considered to be one of the main explanatory variables of a firm's probability of default. Higher is the stock return, lesser is the firm default probability. Consequently, we expect a negative relationship between CDS spread and stock return $r_{i,t}$.

Stock Volatility: It is measured as the annualized standard deviation of a firm's returns. The volatility of a firm's assets captures the general business risk of a firm and provides crucial information about the firm's probability of default. A higher stock return volatility means a higher default probability. Hence we expect a positive relationship between CDS spread and changes in stock volatility $\Delta\sigma_{i,t}$.

Median Rated Index (MRI): It is a market condition variable that gives us information about CDS market, in particular the perceived general economic climate. The general assumption is that improvements in market conditions decrease firms' probability of default and automatically lead to lower credit spreads. In order to measure the market climate we look at the changes in the MRI, namely the $\Delta MRI_{i,t}^{m}$. The MRI is defined as the median CDS spread of all firms in the S&P rating categories (Galil et al., 2014). The MRI has a positive relationship with CDS spreads.

Carbon Risk Factor (CR): It is considered as the difference between the median CDS spreads of polluting and clean firms. When policy events occur, the demand for protection of more (less) exposed firms increases (decreases), for this reason we expect a positive relationship between CDS spread and the CR factor.

In Figure 7.7 and Figure 7.8, we present a descriptive statistics for all dependent and independent variables under consideration in both regions. The Δ indicates that we are calculating differences between successive times.

Variable	Mean	Q25	Median	Q75	SD	Min	Max	Skew	Kurt
Europe									
Dependent variables									
$s^1_{i,t}$ (%)	-0.05	-1.02	0.00	0.24	7.31	-555.00	554.96	0.78	1035.52
$s^3_{i,t}$ (%)	-0.06	-1.04	0.00	0.20	3.74	-93.02	123.19	1.55	46.84
$s^5_{i,t}$ (%)	-0.05	-0.65	0.00	0.11	2.20	-85.00	103.68	1.75	81.66
$s^{10}_{i,t}$ (%)	-0.03	-0.44	0.00	0.13	1.62	-67.49	89.16	1.66	144.62
$s^{30}_{i,t}$ (%)	-0.02	-0.42	-0.01	0.19	2.15	-74.53	85.84	0.60	100.22
Independent variables									
$r_{i,t}$ (%)	0.01	-0.79	0.00	0.84	1.64	-44.33	28.98	-0.66	18.88
$\Delta\sigma_{i,t}$ (%)	-0.00	-0.03	-0.00	0.03	0.24	-19.80	15.28	-0.64	960.59
$\Delta\text{MRI}^1_{i,t}$	-0.01	-0.20	0.00	0.15	1.14	-54.69	60.06	1.42	144.78
$\Delta\text{MRI}^3_{i,t}$	-0.03	-0.41	-0.00	0.26	1.86	-113.32	128.25	2.36	404.38
$\Delta\text{MRI}^5_{i,t}$	-0.04	-0.48	-0.01	0.25	2.29	-179.56	174.67	0.93	872.50
$\Delta\text{MRI}^{10}_{i,t}$	-0.04	-0.50	-0.01	0.30	2.52	-226.28	213.96	-2.08	1385.98
$\Delta\text{MRI}^{30}_{i,t}$	-0.04	-0.51	-0.02	0.38	2.96	-235.35	220.58	-1.27	809.32
ΔCR^1_t	-0.00	-0.27	0.00	0.25	1.06	-7.46	13.83	0.88	27.89
ΔCR^3_t	-0.01	-0.50	0.00	0.51	1.32	-9.95	7.58	0.15	10.27
ΔCR^5_t	-0.02	-0.52	0.00	0.49	1.61	-9.75	11.79	0.38	13.21
ΔCR^{10}_t	-0.01	-0.51	0.00	0.52	1.73	-24.38	10.66	-1.85	35.73
ΔCR^{30}_t	0.00	-0.53	0.00	0.54	2.02	-22.06	23.23	-0.55	31.02

Figure 7.7: Mean, 1st quartile, median, 3rd quartile, standard deviation, minimum, maximum, skewness, kurtosis for all independent and dependent variables (no term structure variables) for Europe.[2]

North America									
Dependent variables									
$s^1_{i,t}$ (%)	-0.03	-0.14	0.00	0.10	7.08	-314.63	371.68	0.96	165.07
$s^3_{i,t}$ (%)	-0.03	-0.12	0.00	0.07	3.42	-151.15	149.83	0.40	140.39
$s^5_{i,t}$ (%)	-0.03	-0.12	0.00	0.05	2.40	-84.93	108.81	1.42	95.77
$s^{10}_{i,t}$ (%)	-0.02	-0.11	0.00	0.05	2.58	-164.77	167.00	1.25	252.18
$s^{30}_{i,t}$ (%)	-0.01	-0.13	0.00	0.06	3.16	-218.32	292.52	2.32	499.67
Independent variables									
$r_{i,t}$ (%)	0.03	-0.70	0.01	0.81	1.73	-42.79	43.14	-0.36	26.38
$\Delta\sigma_{i,t}$ (%)	0.00	-0.03	0.00	0.03	0.27	-25.81	24.89	-0.84	1082.45
$\Delta\text{MRI}^1_{i,t}$	-0.01	-0.15	0.00	0.09	0.82	-34.63	38.21	1.45	110.59
$\Delta\text{MRI}^3_{i,t}$	-0.02	-0.25	0.00	0.13	1.50	-88.44	90.83	-0.20	393.00
$\Delta\text{MRI}^5_{i,t}$	-0.03	-0.36	0.00	0.16	2.10	-159.06	170.63	-0.17	947.99
$\Delta\text{MRI}^{10}_{i,t}$	-0.03	-0.47	0.00	0.30	2.56	-178.57	189.77	-0.27	958.66
$\Delta\text{MRI}^{30}_{i,t}$	-0.03	-0.51	-0.01	0.37	2.64	-174.64	197.60	-0.70	859.71
ΔCR^1_t	0.01	-0.21	0.00	0.24	0.70	-3.64	6.80	0.62	12.68
ΔCR^3_t	0.01	-0.35	0.00	0.37	1.18	-9.30	10.53	0.28	19.43
ΔCR^5_t	0.01	-0.49	0.00	0.49	1.58	-10.83	16.18	0.59	17.53
ΔCR^{10}_t	0.01	-0.73	0.00	0.77	2.31	-15.33	16.60	-0.03	12.41
ΔCR^{30}_t	0.01	-0.89	-0.01	0.81	3.21	-20.17	23.51	0.12	12.30

Figure 7.8: Mean, 1st quartile, median, 3rd quartile, standard deviation, minimum, maximum, skewness, kurtosis for all independent and dependent variables (no term structure variables) for North America.[2]

7.6 Methodological Framework: Panel Quantile Regression

Linear regression has been the workhorse for many financial empirical studies, however this framework turned out to give ambiguous results and to not adequately describe the relationship between CDS spread returns and firms' carbon exposure (Collin-Dufresne et al., 2001; Pereira et al., 2018; Kolbel et al., 2022). For this reason we use a quantile regression (QR) approach, which allows us to:

1. provide a more complete description of how carbon risk is linked to the conditional distribution of CDS spread returns;
2. capture the impact of carbon risk beyond known determinants.

Quantile regression extends the classical conditional mean model (linear regression) to a model for different *conditional quantile functions*, allowing us to analyze the effect of different independent variables on the conditional distribution of the dependent variable. This is especially relevant for credit risk, where understanding the effects on the tails of the distribution is essential.

7.6.1 Linear Regression

Linear regression is a statistical tool which estimates the linear relationship between a dependent variable and one or more independent variables. The case of one independent variable is called simple linear regression; for more than one, the process is called multiple linear regression.
In linear regression, the expectation value of the dependent variable is related to the independent variables through some coefficients, called model parameters, that are estimated from the data. Linear regression models are often fitted using the least squares approach, but they may also be fitted in other ways.

Simple Linear Regression Given a single independent variable X, we are interested in its influence on the expectation value of a dependent variable Y. If the relation is linear, we will have:

$$\mathbb{E}[Y|X] = \beta_0 + \beta_1 X + \epsilon \tag{7.38}$$

where β_0, β_1 are the parameters of our model, in particular:

- β_0 is the **intercept** of our model.
- β_1 is the **angular coefficient** of our model.
- ϵ is the error.

The presence of an error term can be justified by three different reasons:

1. The relation could not be perfectly linear.
2. There may be other independent variables acting on Y which we may not know.
3. There might be errors on independent variables estimation.

The most common method used to obtain those coefficients is the **Ordinary Least Squares (ODS)**.
Let us suppose we have N observations on y_i, for every $i \in \{1, ..., N\}$ and let us assume we believe in model (7.42)

$$y_i = a + bx_i + \epsilon_i \tag{7.39}$$

with the assumption that $E(y_i) = a + bx_i$.
Suppose that we know also the *predicted* values of the dependent variables $\hat{y}_i$, which give an estimation of $E(y_i)$, namely:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \tag{7.40}$$

In the least squares approach, we seek the estimators $\hat{\beta}_0$, $\hat{\beta}_1$ that minimize the sum of squares of the deviation $y_i - \hat{y}_i$ of the N observed y_i from their predicted values $\hat{y}_i$:

$$S := \sum_{i=1}^{N} \epsilon_i^2 = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2. \tag{7.41}$$

To find the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize S, we differentiate with respect to $\hat{\beta}_0$ and $\hat{\beta}_1$ and set the result equal to zero:

$$\frac{\partial S}{\partial \hat{\beta}_0} = -2 \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \tag{7.42}$$

$$\frac{\partial S}{\partial \hat{\beta}_1} = -2 \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \tag{7.43}$$

Solving the system, we get the following relations:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \tag{7.44}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{7.45}$$

where we introduced $\bar{x}$ and $\bar{y}$, namely the mean values of the dependent and independent variables.

Multiple Linear Regression We now attempt to predict the expectation value of a dependent variable Y on the basis of an assumed linear relationship with several independent variables X_1, X_2, ..., X_k, namely:

$$\mathbb{E}[Y|\boldsymbol{X}] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \epsilon \tag{7.46}$$

To estimate the β's, we use a sample of N observations, indexed with the index i, on the dependent variable and the associated independent variables.

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + ... + b_k x_{ik} + \epsilon_i \tag{7.47}$$

with the assumption that $E(y_i) = a + bx_i$.
Having N observations, we will end up with N equations for the dependent variable. We can shortly write this system of linear equations in matrix notation as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \dots & x_{Nk} \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix} \tag{7.48}$$

or

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{b} + \boldsymbol{\epsilon} \tag{7.49}$$

Suppose that we know also the *predicted* values of the dependent variables $\hat{y}_i$, which give an estimation of $E(y_i)$, namely:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_k x_{ik} \tag{7.50}$$

As before, using the least squares approach we seek the estimators $\hat{\beta}_0$, $\hat{\beta}_1$, ..., $\hat{\beta}_k$ by minimizing the sum of squares of the deviation $y_i - \hat{y}_i$ of the N observed y_i from their predicted values $\hat{y}_i$:

$$S := \sum_{i=1}^{N} \epsilon_i^2 = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - ... - \hat{\beta}_k x_{ik})^2 \tag{7.51}$$

Remember that $\hat{\beta}_0$, $\hat{\beta}_1$, ..., $\hat{\beta}_k$ give the best estimation for β_0, β_1, ..., β_k parameters in (8.50).
To find the values of $\hat{\beta}_0$, $\hat{\beta}_1$, ..., $\hat{\beta}_k$ that minimize (8.55), we could differentiate S with respect to each $\hat{\beta}_j$, set the results equal to zero, and solve the system. However, the procedure can be carried out in more compact form with matrix notation:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X^T X})^{-1} \mathbf{X^T y} \tag{7.52}$$

where $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k)$.
The latter result can be obtained from an algebra theorem whose proof is not reported in this thesis.

7.6.2 Quantile Regression

Quantile regression is a type of regression deeply used in statistics and econometrics. If linear regression, with the method of least squares, estimates the conditional **mean** (the conditional expectation value) of the dependent variable across values of the independent variables, quantile regression estimates the conditional **median (or other quantiles)** of the dependent variable. Quantile regression is an extension of linear regression used when the conditions of linear regression are not met.
A quantile of order α or α-quantile (with α a real number in the interval $[0, 1]$) is a value q_α that divides the population into two parts, proportional to α and $1 - \alpha$ and characterized by values respectively smaller and larger than q_α. The quantile of order α is the smallest mode q_α for which the cumulative relative frequency, i.e. the sum of the relative frequencies, calculated up to q_α is at least α. The cumulative relative frequencies following that mode will not exceed $1 - \alpha$.
In the case of a *continuous probability density* with cumulative distribution function F, the quantile of order α is defined by $F(q_\alpha) = \alpha$. For this value, the distribution is correctly divided into two parts proportional to α and $1 - \alpha$.
In the case of a *discrete probability density*, the quantile of order α is a value q_α in which the sum of the discrete probabilities is greater than or equal to α, and the sum of the discrete probabilities from that value onwards is greater than or equal to $1 - \alpha$.
Examples:

1. quantile of order $\alpha = 0.1$ is that value of the distribution for which the cumulative relative probability up to and including that value is greater than or equal to 0.1, and the cumulative relative probability from that value, included, onwards is greater than or equal to 0.9.

2. The quantile of order $\alpha = 0.5$ is that value of the distribution for which the cumulative relative probability up to and including that value is greater than or equal to 0.5, and the cumulative relative probability from that value, inclusive, in then it is greater than or equal to 0.5.

In general, we can called tham with different names:

- The **Median** is the quantile of order $\frac{1}{2}$, which divides the population into two equally populated parts.
- The **Quartiles** are the quantiles of order $\frac{1}{4}, \frac{2}{4}, \frac{3}{4}$, which divide the population into four equally populated parts.
- The **Quintiles**, of order $\frac{m}{5}$ with $m \in \{1, 2, 3, 4\}$, are the quantiles that divide the population into 5 equal parts.
- ...
- The **Centiles**, of order $\frac{m}{100}$ with $m \in \{1, ..., 99\}$, are the quantiles that divide the population into 100 equal parts.

Quantile regression expresses the **conditional quantiles** of a dependent variable as a linear function of the independent variables. As for the conditional mean in linear regression, the conditional quantiles can be expressed as the solution of a minimization problem.
For our purposes, we adopt the Quantile regression framework for a panel setup with firm-specific fixed effects.
Let $y_{i,t}$ be the response of firm i at time t (dependent variable) and $\boldsymbol{x_{i,t}}$ the m-dimensional vector containing the independent variables, where $i = 1, ..., N$ and $t = 1, ..., T$. We have the following relation:

$$y_{i,t} = a_i + \boldsymbol{x_{i,t}^T b} + \epsilon_{i,t} \tag{7.53}$$

where a_i are the firm-specific fixed effects parameters and $\epsilon_{i,t}$ the error terms.
Plotting an histogram, we should be able to see the frequency with which each value of the response of a firm $y_{i,t}$ appears, namely its distribution.
Let us suppose that we are interested in a fixed quantile level $\tau \in (0, 1)$ of our distribution, the **conditional quantile** of $y_{i,t}$ given $\boldsymbol{x_{i,t}}$ can be written in the same form of (7.57), but this time the coefficients of the regression depends on the quantile level we have chosen:

$$Q_{y_{i,t}}(\tau|\boldsymbol{x_{i,t}}) = \alpha_{\tau,i} + \boldsymbol{x_{i,t}^T \beta_\tau} + \epsilon_{i,t} \tag{7.54}$$

Numerous estimation techniques have been established, we follow Zhang et al. (2019) and implement a two-stage approach to estimate the parameter vector β_τ. In a first stage, we run a quantile regressions to estimate the fixed effects $\alpha_{\tau,i}$:

$$\left(\tilde{\alpha}_{\tau,i}, \boldsymbol{\tilde{\beta}_\tau}\right) = \underset{\tilde{\alpha}_i \in A_{\tau,i}, \boldsymbol{\tilde{\beta}} \in \Theta_\tau}{\operatorname{argmin}} \frac{1}{T} \sum_{t=1}^{T} \rho_\tau(y_{i,t} - \tilde{\alpha}_i - \boldsymbol{x_{i,t}^T \tilde{\beta}}) \tag{7.55}$$

where $A_{\tau,i} \in \mathbb{R}$, $\Theta_\tau \in \mathbb{R}^m$ and $\rho_\tau = u(\tau - \mathbb{I}_{\{u<0\}})$ denotes the *quantile loss function*. Provided T sufficiently large, it can be demonstrated that $\tilde{\alpha}_{\tau,i}$ gives the best estimation for $\alpha_{\tau,i}$. In a second stage, we then estimate:

$$\boldsymbol{\beta}_\tau = \underset{\boldsymbol{\beta} \in \Theta_\tau}{\operatorname{argmin}} \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} \rho_\tau(y_{i,t} - \alpha_{\tau,i} - \boldsymbol{x_{i,t}^T \beta_\tau}) \tag{7.56}$$

$\alpha_{\tau,i}$ and $\boldsymbol{\beta}_\tau$ obtained with this methodology provide the best estimation for the parameters of the conditional quantile of $y_{i,t}$ given $\boldsymbol{x_{i,t}}$.

7.7 Empirical Results

In this subsection, we examine the relationship between the carbon risk factor and CDS spread returns.

Following the formalism of quantile regression, we set the CDS spread returns as the dependent variable and we include key known determinants of CDS spread returns as the independent variables:

$$Q_{s^m_{i,t}}(\tau|\boldsymbol{x_{i,t}}) = \alpha_{\tau,i} + \beta_{\tau,1} r_{i,t} + \beta_{\tau,2}\Delta\sigma_{i,t} + \beta_{\tau,3}\Delta MRI_{i,t} + \beta_{\tau,4}\Delta CR_t + \epsilon_{i,t} \tag{7.57}$$

where, for the CDS issued by firm i at time t, we took into consideration:

- **firm-specific factors**, namely the stock return $r_{i,t}$ and the volatility $\Delta\sigma_{i,t}$;
- a **common factor** concerning the market condition $\Delta MRI_{i,t}$;
- a proxy for the **carbon risk exposure** ΔCR_t, which stems for all the changes in carbon related sector.

The regression was run for every decile $\tau \in 0.1, ..., 0.9$ to highlight the effect of each independent variable on the entire conditional distribution of CDS spread returns. In this way, we are able to model the relationship between CDS spread returns and the CR factor for firms that overperform, underperform or behave according to the median of the conditional distribution. Basically, the quantile regression allows us to distinctly examine the effect of each independent variable along the entire distribution of credit spread returns and to investigate the impact of carbon risk over these independent variables.

Firstly, we can easily observe a positive relationship between CDS spread returns and the CR factor. This means that an increase in market's perception of carbon risk generates a rise in the CDS spread returns.

Then, starting from the median ($\tau = 0.5$) we can observe that the coefficients or the CR factor are increasing as we move toward the first and the ninth decile. Basically, the more the state of the firms credit deteriorates or improves, the larger the effect of CR. The effect increases almost symmetrically. Going toward the ninth decile of the CDS spread returns distribution we can appreciate a higher value of ΔCR coefficient, this makes sense because if CR is increasing, ΔCR is **positive**, leading to an extreme *positive* CDS spread shock linked to the exposure to carbon risk. Going toward the first decile we still find higher coefficients with respect to the one of the median decile, this because if CR decrease, then ΔCR is **negative**, and this helps firms experiencing a *negative* CDS spread shock.

We tested **Hypothesis 1**: *there is a positive relationship between carbon risk and CDS spread returns.* In Figure 7.9 we reported the various coefficients at any different decile for tenors of 1, 3, 5, 10, and 30 years. The investigation was held in Europe and includes data of 137 European firms from 01/01/2013 to 31/12/2019 in daily frequency.

Estimates $\beta_{\tau,1}, \beta_{\tau,2}, \beta_{\tau,3}, \beta_{\tau,4}$ and their errors, reported in brackets, have been calculated for all the nine deciles and have been scaled by a factor 10^3.

In Figure 7.9 we reported the various coefficients at any different decile for tenors of 1, 3, 5, 10, and 30 years. The investigation was held in Europe and includes data of 137 European firms from 01/01/2013 to 31/12/2019 in daily frequency.

Estimates $\beta_{\tau,1}, \beta_{\tau,2}, \beta_{\tau,3}, \beta_{\tau,4}$ and their errors, reported in brackets, have been calculated for all the nine deciles and have been scaled by a factor 10^3.

	1	2	3	4	5	6	7	8	9
					1Y				
StockReturn	−281.31***	−247.59***	−178.22***	−11211***	−60.86***	−99.17***	−175.62***	−267.20***	−312.87***
	(13.99)	(8.22)	(6.25)	(4.28)	(2.78)	(3.68)	(5.78)	(10.81)	(19.48)
ΔVolatility	−427.47***	−369.58***	−230.60***	−76.49***	26.65**	275.75***	555.11***	828.02***	980.86***
	(46.03)	(34.73)	(35.25)	(18.13)	(9.88)	(20.90)	(25.40)	(23.04)	(34.26)
ΔMRI	1372.18***	1404.51***	1362.87***	1287.57***	123930***	125663***	1348.04***	1462.28***	1550.98***
	(37.32)	(30.38)	(34.78)	(36.30)	(34.67)	(33.54)	(42.29)	(48.26)	(69.59)
ΔCR	472.70***	347.21***	244.04***	174.88***	122.72***	150.17***	231.54***	347.49***	498.42***
	(27.98)	(15.72)	(14.18)	(11.78)	(9.86)	(11.55)	(15.28)	(22.61)	(31.53)
					3Y				
StockReturn	−218.05***	−204.43***	−164.17***	−10989***	−64.93***	−95.65***	−163.75***	−217.69***	−273.69***
	(7.79)	(5.80)	(4.95)	(3.70)	(2.67)	(3.25)	(4.65)	(7.73)	(12.82)
ΔVolatility	−391.99***	−291.47***	−185.26***	−83.30***	26.23*	238.16***	465.15***	639.01***	857.06
	(52.48)	(48.54)	(23.11)	(18.64)	(11.73)	(18.92)	(20.50)	(15.22)	(7.76)
ΔMRI	558.77***	608.43***	606.76***	586.68***	571.57***	580.52***	612.67***	648.76***	679.99***
	(15.30)	(12.89)	(13.59)	(13.48)	(12.22)	(14.24)	(14.11)	(20.86)	(27.32)
ΔCR	283.18***	228.69***	182.44***	132.27***	92.18***	114.66***	168.78***	221.26***	272.84***
	(9.60)	(7.35)	(7.34)	(6.53)	(5.42)	(5.74)	(6.99)	(9.12)	(14.95)
					5Y				
StockReturn	−143.14***	−124.36***	−98.92***	−66.91***	−42.19***	−56.64***	−94.32***	−132.90***	−174.75***
	(4.55)	(3.60)	(2.94)	(2.29)	(1.72)	(1.94)	(2.82)	(4.68)	(9.14)
ΔVolatility	−279.71***	−202.00***	−137.36***	−58.96***	14.33*	136.81***	279.67***	430.51***	564.96
	(11.10)	(25.64)	(19.57)	(12.68)	(6.72)	(10.69)	(11.84)	(8.31)	(4.72)
ΔMRI	303.66***	332.01***	336.54***	334.85***	332.87***	330.28***	346.04***	365.62***	374.22***
	(5.54)	(7.67)	(8.07)	(7.75)	(7.04)	(7.83)	(9.58)	(9.69)	(16.61)
ΔCR	192.34***	168.24***	140.16***	113.92***	94.25***	99.57***	127.03***	152.62***	180.98***
	(6.47)	(4.87)	(4.54)	(4.65)	(4.28)	(4.31)	(4.74)	(5.29)	(8.29)
					10Y				
StockReturn	−110.20***	−88.55***	−70.03***	−50.39***	−33.36***	−42.92***	−67.90***	−94.80***	−131.77***
	(3.06)	(2.38)	(1.95)	(1.65)	(1.25)	(1.41)	(1.82)	(2.94)	(5.87)
ΔVolatility	−214.75***	−162.14***	−111.86***	−52.23***	5.32	87.17***	193.04***	293.51***	401.34
	(10.84)	(8.68)	(13.26)	(9.71)	(3.78)	(7.71)	(5.31)	(4.25)	(6.32)
ΔMRI	219.39***	236.69***	240.25***	236.04***	234.50***	234.73***	243.35***	253.46***	260.76***
	(4.76)	(5.02)	(4.38)	(4.91)	(4.97)	(4.91)	(5.04)	(4.94)	(9.79)
ΔCR	116.78***	95.30***	80.82***	66.27***	54.92***	58.93***	72.89***	90.44***	112.52***
	(3.23)	(3.43)	(3.02)	(2.58)	(2.56)	(2.68)	(2.64)	(3.37)	(5.84)
					30Y				
StockReturn	−102.92***	−83.70***	−67.15***	−48.99***	−36.33***	−44.11***	−66.63***	−92.10***	−126.36***
	(2.58)	(2.26)	(1.86)	(1.53)	(1.32)	(1.39)	(1.95)	(3.31)	(6.64)
ΔVolatility	−220.25***	−158.34***	−102.71***	−48.46***	5.92	88.53***	189.72***	277.96***	395.18
	(6.04)	(12.13)	(13.24)	(10.89)	(4.19)	(6.79)	(9.12)	(7.27)	(6.66)
ΔMRI	254.53***	248.01***	242.34***	240.95***	240.51***	241.22***	248.85***	263.19***	285.87***
	(5.13)	(4.95)	(4.86)	(4.77)	(5.59)	(5.47)	(6.14)	(6.77)	(9.99)
ΔCR	70.96***	63.55***	54.31***	44.89***	36.51***	39.05***	47.27***	59.98***	79.86***
	(2.61)	(2.22)	(1.83)	(1.60)	(1.69)	(1.63)	(1.96)	(2.28)	(4.22)

Figure 7.9: Estimation of the coefficients from the quantile regression model for 1, 3, 5, 10, and 30 years CDS spread returns. The sample includes data of 137 European firms from 01/01/2013 to 31/12/2019 in daily frequency. Estimates and errors (brackets) are reported for all nine deciles. All estimates are scaled by factor 10^3. [2]

7.7.1 Regional Impact of Carbon Risk

Now we test **Hypothesis 2**: *the effect of carbon risk on CDS spread returns is stronger in Europe than in North America.*

A sample of 281 North American firms from 01/01/2013 to 31/12/2019 in daily frequency has been taken into account. Consistent with the prediction, Figure 7.10 shows a substantially weaker relationship between CDS spread returns and the CR factor for North American firms.

For example, considering the 5Y tenor, the coefficient of CR for the median CDS spread return (0.00004) is nearly 2,400 times smaller than its European counterpart (0.0943). Not only are estimates considerably smaller, but also the symmetry in the effect of CR breaks off in the North American sample. In fact, excluding the 10-years tenor, the effect on the ninth decile is at least twice as high as the effect on the first decile, which is occasionally negative, suggesting that in North America, credit risk exposure is particularly relevant when firms' credit spreads deteriorate and not when they improve.

	1	2	3	4	5	6	7	8	9
					1Y				
StockReturn	−31.94***	−17.78***	−4.75***	−0.65***	−0.16***	−1.12***	−7.88***	−28.76***	−58.73***
	(4.25)	(1.43)	(0.46)	(0.07)	(0.02)	(0.10)	(0.74)	(2.84)	(6.76)
ΔVolatility	−139.05***	−60.98***	−8.90***	−0.13	0.15	5.88***	45.23***	162.38***	393.44***
	(20.75)	(6.38)	(1.68)	(0.26)	(0.10)	(0.63)	(3.96)	(14.79)	(23.74)
ΔMRI	147.45***	88.33***	29.52***	7.34***	2.29***	11.26***	49.92***	165.57***	431.97***
	(23.75)	(9.30)	(3.75)	(0.88)	(0.30)	(1.18)	(6.03)	(19.30)	(50.85)
ΔCR	−1.76	1.50**	1.00***	0.21***	0.06***	0.48***	3.63***	15.84***	49.28***
	(1.17)	(0.57)	(0.17)	(0.04)	(0.02)	(0.07)	(0.46)	(1.87)	(6.69)
					3Y				
StockReturn	−48.53***	−25.76***	−14.46***	−8.22***	−3.50***	−8.63***	−16.40***	−30.52***	−58.80***
	(3.41)	(1.50)	(0.83)	(0.54)	(0.22)	(0.49)	(1.00)	(2.33)	(6.17)
ΔVolatility	−194.68***	−79.67***	−27.60***	−5.58***	0.65	31.52***	77.91***	169.53***	366.50***
	(18.62)	(5.88)	(4.07)	(1.10)	(0.71)	(2.82)	(4.97)	(10.32)	(18.30)
ΔMRI	97.06***	70.35***	41.86***	27.24***	13.44***	28.71***	50.78***	94.52***	186.17***
	(7.83)	(4.97)	(3.81)	(2.66)	(1.18)	(2.52)	(4.20)	(7.59)	(15.39)
ΔCR	3.97***	2.30***	1.48***	0.83***	0.32***	0.65***	1.46***	3.59***	10.01***
	(0.66)	(0.38)	(0.24)	(0.14)	(0.07)	(0.14)	(0.25)	(0.53)	(1.72)
					5Y				
StockReturn	−45.91***	−23.04***	−12.94***	−8.68***	−4.45***	−8.42***	−14.57***	−25.85***	−49.55***
	(2.70)	(1.36)	(0.70)	(0.47)	(0.24)	(0.43)	(0.84)	(1.92)	(4.42)
ΔVolatility	−175.67***	−67.39***	−23.18***	−5.19***	0.55	31.39***	71.73***	149.92***	319.64***
	(12.62)	(6.44)	(3.25)	(1.50)	(0.55)	(2.05)	(4.17)	(8.19)	(12.17)
ΔMRI	53.46***	40.40***	25.13***	18.58***	10.06***	18.26***	31.50***	59.79***	112.81***
	(4.00)	(3.45)	(2.25)	(1.55)	(0.80)	(1.35)	(2.15)	(4.72)	(9.64)
ΔCR	−0.19	0.54**	0.23*	0.12	0.04	0.39***	1.11***	2.78***	7.70***
	(0.42)	(0.20)	(0.10)	(0.07)	(0.04)	(0.07)	(0.12)	(0.31)	(0.95)
					10Y				
StockReturn	−39.83***	−19.99***	−11.16***	−6.90***	−3.49***	−6.29***	−10.86***	−19.49***	−40.61***
	(1.60)	(0.96)	(0.51)	(0.31)	(0.18)	(0.28)	(0.56)	(1.59)	(3.47)
ΔVolatility	−144.49***	−59.47***	−20.50***	−3.86***	1.44***	23.62***	54.58***	114.41***	256.02***
	(6.09)	(5.63)	(2.49)	(0.73)	(0.31)	(1.52)	(2.68)	(6.75)	(8.82)
ΔMRI	36.28***	24.94***	15.54***	10.96***	6.56***	10.80***	17.23***	31.42***	56.29***
	(2.71)	(1.68)	(0.90)	(0.71)	(0.39)	(0.63)	(0.96)	(2.91)	(5.21)
ΔCR	3.57***	1.77***	1.02***	0.47***	0.18***	0.39***	0.73***	1.58***	3.81***
	(0.32)	(0.17)	(0.10)	(0.06)	(0.04)	(0.05)	(0.09)	(0.21)	(0.47)
					30Y				
StockReturn	−46.48***	−25.17***	−15.00***	−9.41***	−5.23***	−8.05***	−13.80***	−23.92***	−47.26***
	(1.84)	(0.93)	(0.58)	(0.38)	(0.22)	(0.34)	(0.65)	(1.49)	(3.48)
ΔVolatility	−156.57***	−72.74***	−28.47***	−8.52***	2.64***	28.93***	64.86***	127.40***	267.00***
	(8.94)	(6.13)	(3.63)	(1.72)	(0.74)	(1.94)	(2.65)	(4.43)	(5.87)
ΔMRI	38.51***	26.58***	17.44***	12.30***	8.47***	11.49***	17.31***	28.99***	51.91***
	(2.27)	(1.29)	(0.90)	(0.67)	(0.45)	(0.60)	(0.98)	(1.94)	(3.87)
ΔCR	1.30***	0.36*	0.14	0.00	0.02	0.22***	0.61***	1.32***	2.78***
	(0.32)	(0.15)	(0.10)	(0.06)	(0.04)	(0.06)	(0.10)	(0.20)	(0.48)

Figure 7.10: Estimation of the coefficients from the quantile regression model for 1, 3, 5, 10, and 30 years CDS spread returns. The sample includes data of 281 North American firms from 01/01/2013 to 31/12/2019 in daily frequency. Estimates and errors (brackets) are reported for all nine deciles. All estimates are scaled by factor 10^3. [2]

7.7.2 Sectoral Impact of Carbon Risk

Until now, we ran our quantile regressions being completely general, taking into account an European/North American *average* firm exposed to carbon risk. It is not easy to realize that certain sectors of the economy may have a higher exposure. Empirical data identifies activities directly related to the production of energy and emissions-intensive goods, especially steel and cement as the most exposed categories.

In order to include these findings in our studies, we developed a more specific quantile regression including **sector representative terms** and **interaction terms**. We regrouped firms using a particular 9-sectors classification (Thomson Reuters Business Classification (TRBC) 2020), obtaining:

$$Q_{s^m_{i,t}}(\tau|\boldsymbol{x_{i,t}}) = \alpha_{\tau,i} + \beta_{\tau,1} r_{i,t} + \beta_{\tau,2}\Delta\sigma_{i,t} + \beta_{\tau,3}\Delta MRI_{i,t} + \beta_{\tau,4}\Delta CR_t + \epsilon_{i,t} \tag{7.58}$$

$$+ \sum_{j=5}^{13} \beta_{\tau,j} Sector_{j,i} + \sum_{k=14}^{22} \beta_{\tau,k} Sector_{k,i} \Delta CR^m_t \tag{7.59}$$

where the terms $Sector_{j,i}$ and $Sector_{k,i}$ indicate a specific firm i chosen among the firms belonging to a specific sector j (or k) with $j, k \in \{0, ..., 9\}$.
In Figure 7.11 we report the coefficient estimates of the **interaction terms** for the 5-years sector model of the European and North American samples, respectively. The 9-sectors chosen are:

- **Basic Materials (BM)**: Companies that manufacture chemicals, building materials, and paper products. This sector also includes companies engaged in commodities exploration and processing.
- **Consumer Cyclical (CCGS)**: This sector includes retail stores, auto and auto-parts manufacturers, restaurants, lodging facilities, restaurants, and entertainment companies.
- **Energy**: Companies that produce or refine oil and gas, oilfield-services and equipment companies, and pipeline operators. This sector also includes companies that mine thermal coal and uranium.
- **Healthcare**: This sector includes biotechnology, pharmaceuticals, research services, home healthcare, hospitals, long-term-care facilities, and medical equipment and supplies. Also include pharmaceutical retailers and companies which provide health information services.
- **Industrials**: Companies that manufacture machinery, hand-held tools, and industrial products. This sector also includes aerospace and defense firms as well as companies engaged in transportation services.
- **Real Estate**: This sector includes companies that develop, acquire, manage, and operate real estate properties.
- **Technology**: Companies engaged in the design, development, and support of computer operating systems and applications. This sector also includes companies that make computer equipment, data storage products, networking products, semiconductors, and components.
- **Utilities**: Electric, gas, and water utilities.

Figure 7.11 shows that the coefficients $\beta_{\tau,k}$ on the interaction term between the sector and ΔCR_t is positive and highly significant for Basic Materials, Energy and Utilities. For the remaining sectors, the coefficients are significantly smaller and, in the North American sample, can even be negative or insignificant. These findings support what we already expected: carbon risk impacts firms' valuation differently, and it is concentrated in specific sectors. Therefore, due to higher coefficients, a growing difference in carbon risk exposure could translate into higher credit risk for firms in carbon-intensive sectors. Conversely, businesses in sectors Industrials, Technology, and Healthcare are seen as capable of providing the innovation and technologies necessary to facilitate a low-carbon transformation, being less affected by a growing difference in carbon risk exposure.

	1	2	3	4	5	6	7	8	9
Europe									
BM × ΔCR	263.50***	203.99***	162.23***	136.76***	116.39***	119.03***	150.63***	187.96***	248.33***
	(13.72)	(10.65)	(13.55)	(12.42)	(11.27)	(11.68)	(13.54)	(13.97)	(15.01)
CCGS × ΔCR	−125.98***	−55.06**	−41.43*	−39.23*	−37.43*	−29.45*	−32.98	−51.53**	−97.90***
	(18.72)	(16.90)	(17.07)	(16.05)	(15.01)	(14.74)	(17.02)	(19.89)	(28.60)
Energy × ΔCR	321.07***	379.94***	415.77***	397.75***	405.50***	392.74***	394.60***	414.31***	415.58***
	(25.86)	(23.60)	(31.29)	(31.56)	(38.08)	(37.93)	(39.29)	(38.01)	(32.73)
Healthcare × ΔCR	−53.32	−43.05	−69.39**	−86.88***	−86.04***	−79.89***	−81.81***	−61.35*	−62.44***
	(41.05)	(22.48)	(24.47)	(18.19)	(15.38)	(16.04)	(21.78)	(24.31)	(18.92)
Industrials × ΔCR	−153.48***	−116.20***	−87.11***	−85.56***	−81.18***	−78.38***	−90.75***	−113.46***	−144.85***
	(19.78)	(15.40)	(17.18)	(14.97)	(13.28)	(14.12)	(16.05)	(19.20)	(23.84)
NCGS × ΔCR	−87.68***	−64.77***	−50.52**	−59.43***	−53.54***	−50.87***	−50.57**	−47.28*	−57.45*
	(19.14)	(17.18)	(17.41)	(14.99)	(14.37)	(14.46)	(16.52)	(19.93)	(22.66)
Real Estate × ΔCR	−50.35	−75.95**	−62.13**	−66.15**	−61.22**	−64.54***	−84.64***	−85.31**	−103.82***
	(68.53)	(26.43)	(20.46)	(24.56)	(21.65)	(18.87)	(24.24)	(30.76)	(18.01)
Technology × ΔCR	−118.80***	−75.02***	−38.00*	−31.99	−32.23*	−30.24*	−45.00**	−71.86**	−128.72***
	(27.11)	(18.16)	(17.45)	(16.61)	(15.27)	(14.95)	(17.23)	(22.34)	(32.38)
Utilities × ΔCR	64.30**	106.26***	131.57***	124.31***	118.32***	115.90***	114.41***	94.43**	37.97
	(24.46)	(29.54)	(23.74)	(24.08)	(23.72)	(22.30)	(21.90)	(29.32)	(44.24)
North America									
BM × ΔCR	13.42***	7.32***	2.69***	0.90**	0.52**	1.76***	5.11***	15.82***	47.59***
	(2.21)	(2.10)	(0.66)	(0.33)	(0.19)	(0.35)	(0.87)	(2.56)	(8.73)
CCGS × ΔCR	−26.88***	−10.85***	−4.26***	−1.74***	−0.85**	−1.57***	−4.04***	−9.87***	−29.22**
	(4.79)	(2.94)	(0.88)	(0.47)	(0.27)	(0.47)	(1.03)	(2.91)	(10.44)
Energy × ΔCR	35.99**	10.74***	2.59**	1.32**	0.67*	0.69	1.79	7.54	29.05*
	(12.03)	(2.82)	(0.93)	(0.51)	(0.31)	(0.54)	(1.40)	(5.32)	(12.95)
Healthcare × ΔCR	−31.78***	−12.05***	−4.29***	−1.60***	−0.90**	−2.28***	−5.13***	−12.93***	−35.26***
	(6.18)	(2.40)	(1.00)	(0.44)	(0.29)	(0.49)	(1.00)	(2.79)	(8.86)
Industrials × ΔCR	−5.89	−5.05*	−1.87*	−0.48	−0.46	−1.53***	−3.96***	−10.08***	−24.46*
	(3.37)	(2.43)	(0.78)	(0.41)	(0.23)	(0.41)	(0.94)	(2.74)	(9.91)
NCGS × ΔCR	−20.58***	−9.70***	−3.82***	−1.30**	−0.69**	−2.17***	−4.96***	−12.66***	−38.72***
	(3.42)	(2.49)	(0.84)	(0.41)	(0.24)	(0.42)	(0.98)	(2.84)	(9.60)
Real Estate × ΔCR	−8.94	−5.93*	−2.14**	−0.59	−0.38	−1.66***	−4.68***	−12.97***	−31.79**
	(7.69)	(2.90)	(0.82)	(0.40)	(0.22)	(0.40)	(0.97)	(3.32)	(10.46)
Technology × ΔCR	−13.10***	−8.87***	−3.52***	−1.17***	−0.61**	−1.66***	−4.59***	−11.95***	−36.58***
	(2.76)	(2.31)	(0.73)	(0.35)	(0.21)	(0.37)	(0.91)	(2.72)	(8.83)
Utilities × ΔCR	−5.58*	−3.77	−1.54*	−0.35	−0.27	−1.15**	−3.68***	−11.38***	−32.46***
	(2.76)	(2.21)	(0.72)	(0.36)	(0.21)	(0.38)	(0.92)	(2.62)	(9.00)

Figure 7.11: Estimation of the interaction terms coefficients $\beta_{\tau,k}$ from the quantile regression model for 5 years CDS spread returns. The sample includes data of 137 European firms and 281 North American firms from 01/01/2013 to 31/12/2019 in daily frequency. Estimates and errors (brackets) are reported for all nine deciles. All estimates are scaled by factor 10^3. [2]

7.7.3 Climate Change Attention

We test **Hypothesis 3**: *The effect of carbon risk on CDS spread returns is stronger during times of heightened attention to climate change.*
For this discussion we exploited two indexes, one for Europe and one for North America:

- **Transition Risk Concern (TRC)** index: used for Europe, it scans Reuters News to detect items with a European regional focus that relate to the introduction of new regulations to curb emissions.
- **Media Climate Change Concerns (MCCC)** index: used for North America, it generates an aggregate score based on the number of articles related to climate change in major US newspapers and their tone. Because MCCC index includes news relating to physical climate risk, we use a variant that only incorporates few topics belonging to our area of interest. The adjusted MCCC index thereby provides daily information on the coverage and sentiment of North American carbon-related news and excludes any physical climate component.

We then define a *high-attention day* for Europe (North America) as a day in which the value of the TRC (MCCC) index is above its median.
In order to include climate change attention in our studies, we introduce a more specific quantile regression, including a **climate news term** and an **interaction term**. To do this we construct a new variable $HCNA_t$ (*High Climate News Attention*) wich takes the value 1 if the TRC (MCCC)

is above its median at day t, indicating high attention to climate change for that day in Europe (North America) and 0 otherwise, in formula:

$$HCNA_t = \begin{cases} 1 & \text{if } TRC(MCCC) > Med\{TRC(MCCC)\} \\ 0 & \text{otherwise} \end{cases} \tag{7.60}$$

Finally, the adjusted quantile regression presents as:

$$Q_{s^m_{i,t}}(\tau|\boldsymbol{x_{i,t}}) = \alpha_{\tau,i} + \beta_{\tau,1} r_{i,t} + \beta_{\tau,2}\Delta\sigma_{i,t} + \beta_{\tau,3}\Delta MRI_{i,t} + \beta_{\tau,4}\Delta CR_t + \epsilon_{i,t} \tag{7.61}$$

$$+ \beta_{\tau,5} HCNA_t + \beta_{\tau,6} HCNA_t \Delta CR^m_t \tag{7.62}$$

In Figure 7.12 we show up the coefficients $\beta_{\tau,4}$ and $\beta_{\tau,6}$ obtained for North America. As we

	1	2	3	4	5	6	7	8	9
					1Y				
ΔCR	−62.21***	−5.44***	−0.21**	0.00	0.00	0.01	0.60***	8.74***	72.37***
	(7.96)	(0.89)	(0.07)	(0.00)	(0.00)	(0.00)	(0.15)	(1.77)	(12.82)
ΔCR × HCNA	98.79***	11.63***	1.09***	0.00**	0.00	0.02**	1.02***	10.23***	56.43***
	(10.49)	(1.54)	(0.14)	(0.00)	(0.00)	(0.01)	(0.21)	(1.90)	(14.26)
					3Y				
ΔCR	26.65***	5.01***	1.50***	0.47***	0.13**	0.36***	0.54**	1.69**	11.63***
	(2.82)	(0.71)	(0.21)	(0.10)	(0.04)	(0.09)	(0.18)	(0.62)	(3.18)
ΔCR × HCNA	−16.72***	−0.71	−0.12	0.01	0.02	0.16	1.61***	6.84***	33.13***
	(3.69)	(0.91)	(0.27)	(0.12)	(0.06)	(0.13)	(0.36)	(1.11)	(6.66)
					5Y				
ΔCR	−1.22	−0.23	−0.02	0.00	0.02	0.06	0.17*	0.47	3.68*
	(1.46)	(0.44)	(0.10)	(0.04)	(0.02)	(0.04)	(0.07)	(0.27)	(1.43)
ΔCR × HCNA	5.12**	1.98***	0.41*	0.17*	0.08	0.42***	1.67***	7.30***	32.04***
	(1.89)	(0.57)	(0.16)	(0.08)	(0.05)	(0.10)	(0.23)	(1.02)	(4.90)
					10Y				
ΔCR	8.69***	2.63***	1.01***	0.35***	0.11***	0.20***	0.23**	0.55	−0.00
	(1.05)	(0.28)	(0.10)	(0.05)	(0.03)	(0.05)	(0.09)	(0.30)	(1.20)
ΔCR × HCNA	−4.08***	−0.66	−0.19	−0.03	0.09	0.31***	1.13***	4.67***	23.48***
	(1.23)	(0.37)	(0.15)	(0.08)	(0.05)	(0.08)	(0.16)	(0.57)	(3.30)
					30Y				
ΔCR	7.29***	1.66***	0.58***	0.16*	0.06	0.18**	0.19*	0.44	0.26
	(0.99)	(0.25)	(0.13)	(0.07)	(0.04)	(0.06)	(0.09)	(0.25)	(0.71)
ΔCR × HCNA	−7.53***	−0.72*	−0.16	0.00	0.09	0.22**	0.95***	3.34***	16.76***
	(1.30)	(0.36)	(0.17)	(0.10)	(0.07)	(0.08)	(0.15)	(0.46)	(1.86)

Figure 7.12: Estimation of ΔCR and $HCNA_t\Delta CR^m_t$ from the quantile regression model for 1, 3, 5, 10, and 30 years CDS spread returns. The sample includes data of 281 North American firms from 01/01/2013 to 29/06/2018 in daily frequency. Estimates and errors (brackets) are reported for all nine deciles. All estimates are scaled by factor 10^3. [2]

expected, the table shows that the coefficient of the interaction term $HCNA_t\Delta CR$ is positive and higher for 1-year tenor of CDS, indicating a strengthening effect of carbon risk when attention to climate change is high. This observation is persistent across all deciles and clearly the effects are more pronounced at the extremes of the conditional distribution. Viceversa, for tenors longer than 3 years the coefficients start to decrease, and CDS spread returns seem to be less affected by heightened attention to climate change. This is probably due to the general short-lived impact of news, even though the effect of climate change attention does not completely vanish.

In Figure 7.13 we report the estimation results for Europe. What we found here is a little bit unexpected and less clear. While attention to climate change seems to be relevant for higher tenors (especially for 3-years and 10-years tenors), news about adjustments in European carbon regulations seems to be irrelevant for 1-year tenor CDS spread returns. This contradicts the third hypothesis, however it is consistent with the findings that the effect of carbon risk on CDS spread returns for the 1-year tenor is substantially large. When market-wide concern about climate change risk is elevated, lenders appear to only be more sensitive to carbon risk for longer tenors.

	1	2	3	4	5	6	7	8	9
1Y									
ΔCR	551.67***	392.13***	274.80***	178.65***	124.82***	164.28***	248.73***	361.48***	472.87***
	(27.52)	(26.52)	(20.56)	(16.13)	(13.50)	(15.25)	(19.88)	(29.01)	(49.68)
ΔCR × HCNA	−76.13	−30.65	−39.48	−19.57	−11.91	−20.45	1.46	76.74***	189.62***
	(40.07)	(28.23)	(21.90)	(17.17)	(12.83)	(13.65)	(15.44)	(15.21)	(22.56)
3Y									
ΔCR	344.32***	236.58***	182.21***	122.14***	87.18***	108.04***	173.25***	233.74***	297.92***
	(10.42)	(9.24)	(8.80)	(7.50)	(6.29)	(6.25)	(7.42)	(11.04)	(21.64)
ΔCR × HCNA	−41.83**	17.34	12.06	21.00*	14.42*	29.48***	31.38***	57.78***	90.46***
	(13.94)	(10.67)	(9.78)	(8.52)	(7.16)	(7.52)	(7.99)	(8.70)	(15.34)
5Y									
ΔCR	210.45***	174.22***	140.25***	107.75***	86.13***	88.75***	116.09***	144.75***	172.42***
	(8.31)	(6.18)	(5.00)	(5.26)	(4.93)	(4.65)	(5.24)	(6.86)	(11.12)
ΔCR × HCNA	−33.45***	−18.34**	−9.72*	−1.41	7.68	17.37**	21.06***	28.79***	41.74***
	(7.21)	(5.58)	(4.87)	(5.43)	(5.67)	(5.81)	(6.35)	(7.60)	(11.64)
10Y									
ΔCR	100.48***	68.77***	54.67***	44.37***	35.38***	42.11***	56.40***	74.41***	93.40***
	(4.22)	(2.63)	(3.08)	(2.25)	(2.27)	(2.49)	(2.90)	(3.04)	(5.55)
ΔCR × HCNA	23.68***	26.11***	21.26***	13.90***	10.25**	10.12***	15.00***	21.25***	26.79***
	(4.59)	(3.88)	(3.74)	(3.27)	(3.37)	(2.98)	(3.36)	(3.74)	(5.88)
30Y									
ΔCR	90.60***	75.02***	63.68***	48.09***	42.21***	45.84***	55.08***	64.18***	73.74***
	(4.43)	(3.54)	(2.57)	(2.29)	(2.44)	(2.52)	(2.80)	(2.80)	(6.69)
ΔCR × HCNA	−32.23***	−26.95***	−21.58***	−12.45***	−13.99***	−14.57***	−13.89***	−2.10	17.28***
	(5.07)	(3.88)	(2.87)	(2.66)	(2.59)	(2.60)	(2.50)	(2.43)	(4.35)

Figure 7.13: Estimation of ΔCR and $HCNA_t \Delta CR_t^m$ from the quantile regression model for 1, 3, 5, 10, and 30 years CDS spread returns. The sample includes data of 137 European firms from 01/01/2013 to 31/12/2019 in daily frequency. Estimates and errors (brackets) are reported for all nine deciles. All estimates are scaled by factor 10^3. [2]

7.8 Summary

The process of decarbonization due to net-zero carbon emissions policies surely requires a huge economic transformation, and these changes can generate large costs with consequential risk of default, especially for firms that are unprepared for the transition. In order to understand the impact of these transformations we need measure firms' carbon risk exposure.

Theoretical arguments indicate that firms' exposure to carbon risk might be detected in their credit spreads. Therefore, we exploited the information contained in CDS spreads to construct the CR factor: a market-implied forward-looking proxy for carbon risk exposure. In this last chapter we proposed a method for constructing this proxy for carbon risk exposure and studied how it affects firms' creditworthiness. In particular, we found out a positive relationship between this proxy for exposure to carbon risk and firms' cost of default protection. The observed relationship was significantly stronger in Europe, more stringent in carbon regulations, than in North America.

Using quantile regressions, we showed that the magnitude (coefficients) of the exposure to carbon risk differs considerably along quantiles of the distribution of CDS spread returns.

We found out that exposure to carbon risk also varies across industries. While we observed a high sensitivity to carbon risk in the CDS spreads of carbon-intensive sectors (like Energy, Basic Materials, Utilities), other sectors (like Industrials, Technology, Healthcare) seems to be better suited to a low-carbon transformation.

Further analysis suggested us that the effect of carbon risk on CDS spread returns is stronger during times of heightened attention to climate change news. During these periods, lenders demand more credit protection for those borrowers perceived to be more exposed to carbon risk. These results provide some quantitative assessment of carbon risk economic impact. In particular, they suggest that an improvement in current carbon emissions disclosures and emissions reduction strategies would facilitate better assessment of firms' carbon and credit risk.

Bibliography

[1] Tomas Björk. *Arbitrage theory in continuous time.* Oxford university press, 2009.

[2] Alexander Blasberg, Rüdiger Kiesel, and Luca Taschini. Carbon default swap–disentangling the exposure to carbon risk through cds. *Available at SSRN 3856993*, 2021.

[3] Giorgia Callegaro and Marzia De Donno. Stochastic calculus. *Lecture Notes*, 2015/2016.

[4] John Y Campbell. Some lessons from the yield curve. *Journal of economic perspectives*, 9(3):129–152, 1995.

[5] Abdelmoula Dmouj. Stock price modelling: Theory and practice. *Masters Degree Thesis, Vrije Universiteit*, 2006.

[6] Claudio Fontana. Mathematics for financial risk and derivatives. *Lecture Notes*, 2021/2022.

[7] Erwin Frey and Klaus Kroy. Brownian motion: a paradigm of soft matter and biological physics. *Annalen der Physik*, 517(1-3):20–50, 2005.

[8] John Hull, Izzy Nelken, and Alan White. Merton's model, credit risk, and volatility skews. *Journal of Credit Risk Volume*, 1(1):05, 2004.

[9] John C Hull and Sankarshan Basu. *Options, futures, and other derivatives.* Pearson Education India, 2016.

[10] Roger Koenker. *Quantile regression*, volume 38. Cambridge university press, 2005.

[11] Francesco Mainardi and Paolo Pironi. The fractional langevin equation: Brownian motion revisited. *arXiv preprint arXiv:0806.1010*, 2008.

[12] Albert P Philipse. Brownian motion. *Undergraduate lecture notes in physics*, 2018.

[13] Krishna Reddy and Vaughan Clinton. Simulating stock prices using geometric brownian motion: Evidence from australian companies. *Australasian Accounting, Business and Finance Journal*, 10(3):23–47, 2016.

[14] Alvin C Rencher and G Bruce Schaalje. *Linear models in statistics.* John Wiley & Sons, 2008.

[15] Viktor Stojkoski, Trifce Sandev, Lasko Basnarkov, Ljupco Kocarev, and Ralf Metzler. Generalised geometric brownian motion: Theory and applications to option pricing. *Entropy*, 22(12):1432, 2020.

[16] Tomasz Zieliński. Merton's and kmv models in credit risk management. *Studia Ekonomiczne*, 127:123–135, 2013.