University of Pisa
Sant’Anna School of Advanced Studies
Department of Economics and Management
Master of Science in Economics
Master’s Thesis

Distance to Default
as a Predictor of Financial Distress

Supervisor:
Prof. Giulio Bottazzi

Assistant Supervisor:
Dott. Federico Tamagni

Candidate
Roberta Pullano

March 7, 2014
Abstract

Probit models, which adopt both accounting and market information, have become one of the most suitable statistical method in forecasting firm’s default.

Comparing Merton’s Structural model with the Naïve alternative approach suggested by Bharath and Schumway (2008), my work shows that the Naïve predictor, in spite of its simple formulation, has a strong forecasting power. The latter retains Merton’s functional form but avoids the most criticised aspect of Merton’s model: the simultaneous calculation of unobservable variables. Using a dataset of publicly traded Italian firms between 1990 and 2011, a probit estimation shows that Merton’s distance to default is not a sufficient statistic to predict a firm’s financial distress. Data analysis underlines the similarity in distribution of the main components of Merton’s and Naïve models and the strong correlation between them. However, Merton’s functional form and market-based approach remain relevant in credit-risk models.
## Contents

**Introduction** 1  

1 Literature Review 3  
  1.1 Non Structural Models .................................. 6  
  1.2 Structural Models ..................................... 7  

2 Distance to Default 11  
  2.1 Merton’s model ........................................ 11  
     2.1.1 Assumptions ................................... 11  
     2.1.2 Model setting .................................. 13  
     2.1.3 Distance to Default: a definition .............. 15  
  2.2 Merton’s non-linear system ............................ 17  
     2.2.1 Merton’s system with rescaled variables ...... 18  
  2.3 The Naive predictor ................................... 22  

3 Data description and The Default-risk Zone 26  
  3.1 Data .................................................... 26  
  3.2 Default-risk Selection Criteria ........................ 27  
  3.3 Merton and Naive: Summary statistic ................. 31  

4 Forecasting Default-risk: Probit Models 34  
  4.1 Probit Models .......................................... 34  
     4.1.1 The Marginal Effect .............................. 35  
     4.1.2 Goodness of fit .................................. 36  
  4.2 Results ................................................ 36  
     4.2.1 Probit models per year ........................... 37  
     4.2.2 Pooled probit models ............................. 39  

Conclusions 1
## CONTENTS

### A Stochastic calculus

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1</td>
<td>Probability space and filtration</td>
<td>2</td>
</tr>
<tr>
<td>A.2</td>
<td>Stochastic processes</td>
<td>4</td>
</tr>
<tr>
<td>A.3</td>
<td>Brownian motion and Ito’s Lemma</td>
<td>5</td>
</tr>
<tr>
<td>A.3.1</td>
<td>Geometric Brownian motion</td>
<td>8</td>
</tr>
<tr>
<td>A.3.2</td>
<td>Ito’s Lemma</td>
<td>9</td>
</tr>
<tr>
<td>A.3.3</td>
<td>The lognormal distribution</td>
<td>10</td>
</tr>
</tbody>
</table>

### B The Black-Scholes-Merton scenario

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1</td>
<td>The market model</td>
<td>12</td>
</tr>
<tr>
<td>B.2</td>
<td>Portfolio strategies</td>
<td>13</td>
</tr>
<tr>
<td>B.3</td>
<td>Contingent Claims and Arbitrage</td>
<td>15</td>
</tr>
<tr>
<td>B.3.1</td>
<td>European options</td>
<td>15</td>
</tr>
<tr>
<td>B.3.2</td>
<td>The Put-Call parity</td>
<td>16</td>
</tr>
<tr>
<td>B.4</td>
<td>Solution of the Black-Scholes PDE for a Call Option</td>
<td>17</td>
</tr>
</tbody>
</table>
Introduction

Is it possible to predict a company’s default in a reasonable period of time before it occurs? Are we able to do it with a certain knowledge and to prevent it? Are credit-risk models able to quantify the amount of a firm’s accumulated risk? Is there a univocal and accepted measure to calculate corporate’s default risk? Answers to these questions are not trivial and many researchers and practitioners are still looking for them. The great certainty is that there is a wide number of external and internal factors which could lead a firm to default.

Default occurs when a debtor is unable to meet the legal obligation of debt repayment. Models that forecast corporate default have always raised interest among academics and practitioners. The assessment of credit risk concerns all credit institutions. Recently banks have devoted more resources than usual in it. In particular, one of the purposes of Basel II accord was to create standards and regulations on the amount of capital that financial institutions must put aside to reduce the risks associated with their investing and lending practices. In particular the accord regulates the amount of capital a bank need to invest in riskless security to cover its credit risk, that is the probability that its counterparty will default. Since riskless investments do in general generate lower returns, banks are interested in estimator able to measure their credit exposure. In this work my attention is focused on one of these estimators, a measure called Distance to default. It can be seen as a quantitative indicator of how much a company is far from, or near to, default.

Merton was the first to introduce the distance to default. His structural model for credit-risk is based on market information, in order to produce a distance to default. The latter is then mapped into a probability of default. The importance of Merton’s model can be found in its market-based approach and in its theoretical basis. In spite of its strengths, it has been criticised for its unrealistic assumptions and its analytical limits. The burdensome calculation of unobservable variables through a non-linear system
required by Merton’s estimator may not be crucial to obtain the same result. This is the conclusion of Bharath and Schumway (2008), who created a similar default predictor which maintains Merton’s model functional form but bypasses the non-linear system solving.

In this work, I compare Merton’s model and the Naïve alternative proposed by Bharath and Schumway (2008). My purpose is to investigate if Merton’s alone is a sufficient statistic in order to predict financial distress and to compare the real contribution of Merton’s and Naïve models using a probit analysis.

This work is organised as follows: Chapter 1 reviews the positive and negative aspects of the most famous default risk models. Chapter 2 focuses on Merton’s model, the Naïve alternative and the resolution of Merton’s non-linear system. Chapter 3 describes the dataset, selection criteria used to define a default-risky zone for firms at risk and includes a summary statistic of the most important parameters of Merton’s and Naïve models. The first part of Chapter 4 contains a description of probit regressions. They are useful to identify and characterize factors which may contribute to forecast a company’s default. The last part contains empirical results, followed by Conclusions. This work is provided with two theoretical Appendices concerning stochastic calculus and the Black-Scholes-Merton scenario.
Chapter 1

Literature Review

An investor, in order to measure the default risk associated to a certain firm, can follow different methods. The first one is related to rating agencies. They evaluate the capability of a company to repay its debt and classify its credit worthiness according to a rating. The most famous and recognised rating agencies are Moody, Standard & Poor and Fitch. An other possible path is to look at non-structural models. They calculate credit risk by referring to accounting information. The third possibility is related to structural models, that use option-pricing theory to calculate credit risk. Merton’s model is the most known model of this kind. It extracts information from the market in order to produce a probability of firm’s default.

There is a large literature who compares and analyses both structural and non-structural approaches.

Accounting data-based bankruptcy prediction models filter the relevant information from publicly available accounts to assess bankruptcy risk. In a way, traditional accounting models exploit structured fundamental analysis using published financial statements. They are typically estimated by searching for the linear combination of ratios that best differentiates between samples of non-failed and failed firms through discriminant or logit models.

Despite the widespread use of the accounting-based bankruptcy prediction models in the literature, they are often criticised for their lack of theoretical grounding. Hillegeist et al. (2004) argue that accounting data is by nature historical and prepared on a “going concern” assumption, hence their use in predicting future, especially one that involves violating the “going concern” assumption itself is fundamentally flawed. Similarly, Agarwal and Taffler (2008) acknowledge that accounting numbers are subject to reporting
standards (such as conservatism and historical cost accounting) that might hinder a true representation of the economic value of assets. Moreover, accounting numbers can be manipulated by the management.

In addition, there are methodological issues associated with the development of accounting-based bankruptcy prediction models. For instance, Zmijewski (1984) argues that such models are biased as they typically oversample failed firms during model development. Mensah (1984) argues that as ratios change over time, a regular re-estimation of the models is necessary to maintain their utility. However, Begley et al. (1996) and Hillegeist et al. (2004) find that simply updating the model coefficients does not improve the performance, hence such models have to be redeveloped periodically.

Contingent claims-based bankruptcy prediction structural models overcome many of the fundamental shortcomings of accounting-based models. First, in efficient markets, prices reflect both historical financial information (i.e., accounting data) as well as the individual and market-wide outlook of a business. Second, market prices are less likely to be influenced by accounting policies. Third, while accounting-based models typically lack theoretical underpinnings, contingent claims-based models have impeccable theoretical grounding as they draw on the Black and Scholes (1973) and Merton (1974) option pricing framework. In these models, equity is viewed as a call option on the firm’s assets, and the probability of going bankrupt is simply the probability that the call option is worthless at maturity (i.e., market value of total assets is less than the face value of total liabilities).

However, implementation of the contingent claims framework for bankruptcy prediction is far from straightforward. First, Saunders and Allen (2002) argue that such models are unable to differentiate between the different durations of debt since they assume a zero-coupon bond for all liabilities. Second, Avramov et al. (2010) argue that distressed firms are prone to suffer from market microstructure problems such as thin trading or limitations to short-selling which might result in prices deviating from fair values for extended periods. Perhaps more importantly though, some key variables required for these models (e.g., asset volatility, expected asset returns, and market value of assets) are unobservable and need to be approximated introducing potentially large errors.

The competing arguments in accounting and contingent claims-based bankruptcy prediction frameworks enforce a trend in literature that argues for combining the two information sources. Sloan (1996) finds that market prices do not accurately reflect the information from company accounts, hence, accounting data can be used to complement market data. Pope (2010) argues
for combining the accounting and finance disciplines. In line with these arguments, latest hazard models dismantle the strict separation of accounting and market data while incorporating the informational benefit of both.

The evidence in the existing literature that compares the performance of contingent claims and accounting-based approaches shows that the theoretical superiority does not necessarily imply a higher explanatory power. Hillegeist et al. (2004) compare the Ohlson (1980) o-score and Altman (1968) z-score with the contingent claims-based measure using information content tests. While they claim their contingent claims based model carries more information about future bankruptcy, they also find that the two measures carry different information as neither measure subsumes the other. Further, they do not provide tests of classification accuracy of the different models. Reisz and Perl (2007) model equity as a down-and-out call option and compare it with the vanilla call option framework as well as the Altman (1968) z-score using the ROC curve. They show that for a one-year prediction horizon their down-and-out framework outperforms the Merton (1974) framework but, in contrast to Hillegeist et al. (2004), underperforms the z-score model. Likewise, Agarwal and Taffler (2008a) find that there is no significant difference in the explanatory power between z-score and the Merton (1974) based approach. Similar to Hillegeist et al. (2004) they demonstrate that the two measures carry different information about bankruptcy risk. In addition, using the approach of Bichlinger and Leippold (2006) they show that z-score is a better pricing tool for lenders.

The second strand of literature tests the performance of hazard models against accounting-based models. Shumway (2001) compares a hazard model to the accounting-based alternatives and finds that the majority of previously used accounting variables from Altman (1968) and Zmijewski (1984) have little power in forecasting bankruptcy. He further finds that a combination of accounting and market-driven variables such as past stock returns and idiosyncratic risk increases forecasting accuracy significantly. Charalambakis et al. (2009) find similar results for the UK market. Chava and Jarrow (2004) extend the study of Shumway (2001) and provide further confirmatory evidence that the accounting variables previously used in the literature add little predictive power when market variables are also included. Finally, Bharath and Schumway (2008) show that the strength of structural models lies in their functional form, not in the criticised calculation of unobservable variables. They compare a Naive alternative which retains this functional form, with Merton’s model. The results show that the latter has less forecasting power than their naive predictor.
1.1 Non Structural Models

The first scoring model for firms was developed by Altman (1968). Known as Z-score, it uses five financial ratios to attribute a credit score to companies. These ratios, obtained from a discriminant analysis model, are weighted differently. The five ratios are:

1. working capital/total assets;
2. retained earnings/total assets;
3. earnings before taxes and interest/total assets;
4. market value of equity/book value of total liabilities;
5. sales/total assets.

An extension of this approach has been the use of linear or non-linear regression models to directly estimate the probabilities of default. These models allow several ratios and assorted financial data to be considered simultaneously and provide descriptive statistics for the estimated parameters. Furthermore, they can explicitly model non-linearities between the financial variables and the score and, finally, directly compute the probability of default. Logit and probit models are often used. Typically, the greatest variations in the probabilities of default come from ratios capturing firms profitability, level of indebtedness, liquidity and solvency. These models can be estimated on cross-section or panel data.

Several banks use this method for privately owned and publicly traded firms, either by buying a model or its extension, such as Moodys RiskCalc, or by programming their own estimation method. A problem they frequently encounter involves building an adequate database. Very often, credit files are not computerized or contain no historical data.

The main benefit of non-structural models is their precision in estimating probabilities of default. Furthermore, they are easy to use for financial institutions equipped with strong database management systems. Beaulieu (2003) demonstrates how data from a Canadian bank can yield very precise probabilities of default. On the other hand, these models are not flexible, since they require information from financial statements. Thus, it proves very difficult to update probabilities of default over the course of a year. Some institutions may demand financial statements on a quarterly basis, but these are rarely audited. Another criticism is the absence in accounting data of anticipations regarding the future. They reflect the past well, but
1.2 Structural Models

tell us nothing of the future. Market data are more relevant to forecasting probabilities of default.

1.2 Structural Models

The pioneers of structural models were Black and Scholes (1973) and Merton (1974). They are also known as “firm-value” models because they have a common concept: a company defaults on its debt if the value of the asset of the company falls below a certain \textit{default point}.

The option pricing theory governs such models because the central idea is that the market value of the firm is modeled as an option on the firm’s asset with strike price equal to the face value of the firm’s debt.

As a result, structural models improved credit risk analysis and particularly the valuation of default-prone risky bonds. They show how and how much default risk is related to the capital structure and asset volatility of the firm. For this reason the structural approach is a useful tool in the analysis of counterparty risk for banks and in the risk analysis of portfolios of securities.

The first and most important structural model is \textit{Merton’s model}, which treats the equity of a firm as an european call option on its assets, with maturity fixed in time. At the horizon date, stockholders exercise the option if the value of the assets exceeds that of the debt, and then reimburse the debt and share the surplus. Otherwise, the firm is in default and stockholders do not exercise their option. Their loss is then equal to the initial investment. Thus, the probability of default is the probability that the option is not exercised. To evaluate this probability, another measure is needed: the distance to default, which is equal to the gap between the mean asset value and the value of the debt, normalized by the standard deviation of the asset value. The shorter this distance, the greater the probability of default.

To improve the basic Merton model, several extensions have been suggested in the literature, including Moodys KMV. The paper by Crosbie and Bohn (2003) refers to this model. It uses a proprietary model and it estimates a new credit risk measure denoted as Expected Default Frequency (EDF), which represents the probability of default of a given firm. Through its large historical database, Moody created an empirical distribution of distances to default and the EDF is obtained using a software called Credit Monitor, which maps actual default rates to obtain default probabilities.

Later on, it was improved by Vasicek and Kealhofer to obtain a model called VK which is a generalization of Merton’s KMV. It includes more classes and
1.2 Structural Models

maturities of debt in the capital structure of the firm. The most relevant extension of Merton’s model is from Brockman and Turtle (2003) and Avellaneda and Zhu (2001), who propose using barrier options, which were introduced into the literature by Brennan and Schwartz (1978), Leland (1994), and Briys and de Varenne (1997). They use the down-and-out option on the assets: lenders hold a portfolio of risk-free debt and a short put option combined with a long down-and-out call option on the firm’s assets. The last part gives them the right (but not the obligation) to place the company into bankruptcy when they anticipate that its financial health can only deteriorate. Thus, rather than stockholders who wait for the debt to mature before exercising a standard European call, this option makes it possible to place the firm into bankruptcy as soon as the value of its assets reaches the barrier at any time until the debt’s maturity. This theory is in response to the main criticism levelled at Merton’s model: it does not account for the possibility that the firm may default before the debt matures.

Brockman and Turtle (2003) demonstrate that Merton’s standard call option model is a special case of the barrier option model, and test their model on U.S. data. They empirically verify that the barriers are statistically different from zero, thus rejecting the standard European call option for all years, capital structures, and industries studied. Finally, they show that their model, with a barrier option, dominates Altman’s (1968) Z-score non-structural model. It is important to emphasize that they may not use the most advanced version of the non-structural model. Duan, Gauthier, and Simonato (2004) demonstrate that estimating the parameters of the Brockman and Turtle (2003) model by maximum likelihood yields results that resemble those from the iterative estimation method used in this literature when the theoretical model is Merton’s, or when the capital structure is fixed. The appeal of the maximum-likelihood method is that it allows for statistical inference or, more specifically, calculating descriptive statistics for the estimated parameters, such as the value of the firm. Another important aspect in the contribution by Duan, Gauthier, and Simonato (2004) is that the correspondence between the two estimation methods is not necessarily perfect when we insert an additional parameter into the structural model to account for the capital structure, as when Brockman and Turtle (2003) estimate three parameters (the value of the firm, its standard deviation, and a parameter for the capital structure owing to the barrier option), instead of two. In this particular instance, the maximum-likelihood method dominates, since it yields unbiased estimates of the parameters.

Structural models have positive and negative aspects, and their usefulness depends on the purpose of the specific analysis. They make use of market
information to refer at a firm’s value and this is possibly the best source of information as it reflects the sum of many investors’ forecasts. Assuming the lack of any insider information, it is difficult that any individual’s, or committee’s, forecast is better. However, market expectations could be caught by surprise, generating large and sudden changes in credit valuation.

Moreover, structural models are difficult to calibrates and for this reason are not suitable for the frequent marking to market of credit contingent securities. They are also computationally burdensome.

Structural approach has been criticized for overestimating the probabilities of default (Duan and Fulop 2005). The presence of trading noise on the exchange introduces randomness into the correlation between unobservable asset values and stock prices, thus annulling the one-to-one relationship between these two values. This relationship is very important, however, in applications of the maximum-likelihood method to unobservable data on assets. Duan and Fulop (2005) demonstrate that the presence of trading noise can affect the standard deviation in the Merton model. On the basis of their sample of securities, they find an average increase of 7.64 per cent in the standard deviation, with a maximum of 25 per cent, which has an effect on the projected probabilities of default.

To avoid most of these problems, many practitioners and academics developed an hybrid model, which incorporate both structural and non structural characteristics. Tudela and Young (2003) present an application of the hybrid model. They estimate various models on data from non-financial English firms for the period 1990-2001. For their estimates of probabilities of default in the structural model, the researchers use data on firms that did, and did not, default. Thus, they first verify whether the two firm types are assigned different predicted probabilities of default, i.e. Type I and II errors\(^1\). They suggest that the estimated probabilities of default are successful in discriminating between failing and non-failing firms. The results of the error tests are satisfactory. Classifying defaults as those firms with an estimated probability of default greater than, or equal to, 10 \%, the Type I error is relatively modest at 9.2\% (with a Type II error of 15\%).

Second, they compare their hybrid model with other non-structural models to see whether the added probability of default variable is significant for explaining probabilities of default. Their implementation of the Merton’s approach outperforms a reduced-form model based solely on company’s account data. But their hybrid model, obtained by combining company’s

\(^1\)Type I errors are defined as the percentage of actual failures classified as non-failures, Type II errors are the percentage of non-failures classified as failures
account information and the Merton’s approach, outperforms their implementation of the Merton’s model, if only marginally.
Third, they measure the performance of their model with power curve and accuracy ratio type instruments and they confirm that the hybrid model outperforms other models. They also perform dynamic analysis and find that the probabilities of default rise as the date of default nears.
The conclusion of their work is that, over a one-year interval, the mean of the probability of default for the non-defaulting firms is 5.44 per cent, while that percentage rises to 47.33 per cent for those that did default. Tudela and Young (2003) further confirm that the PD variable is significant in their probit model, increasing the estimated likelihood.
Also Moody and FitchSolutions developed proprietary hybrid models; in their quantitative research Special Report published in 2007, Fitch’s analysts describe barrier options with a down-and-out call option.
Chapter 2

Distance to Default

2.1 Merton’s model

In 1974, Robert Merton defined a model for assessing credit risk; in particular, in the line with the option pricing theory recently developed by Black and Scholes, he related credit risk to the capital structure of the firm, and therefore, to observable variables. Merton, by modeling a firm’s market value, market value volatility and liability structure uses contingent claim analysis to construct the probability of default for each firm in the sample at any given point in time.

He defines a firm as defaulted when the value of its assets is less than the promised debt repayment at maturity. To do this, he needs an estimate of the market value of a firm and its volatility. Market value can be recovered by the sum of firm’s equity and debt. He assumes that equity’s market value is a call option\(^1\) on the value of the firm’s asset while market value of debt can be derived from the Put-Call parity.\(^2\)

2.1.1 Assumptions

To develop his model, Merton makes some assumptions related to the market structure, and the firm’s capital structure. I will review them in turn.

- **Perfect Market.**
  
The following assumptions are formulated for the sake of convenience.

---

\(^1\)For a more specific description of Contingent Claims theory see Appendix B

\(^2\)see Appendix B
1. There are no transaction costs, taxes, or problems of indivisibilities of assets;
2. Each investor can buy and sell as much of an asset as he wants at the market price, i.e. the market is liquid.
3. The risk free interest rate for borrowing and lending is the same and it is constant through the time horizon;
4. Short-sales of all assets are allowed.

- **Firm’s capital structure.**
  Suppose that the company has two classes of claims, both functions of the firm’s asset value \( V(t) \) and time; hence the company structure is characterized as follows:

\[
V(t) = E(V, t) + D(V, t),
\]

where \( E(V, t) \) represents equity’s market value, and \( D(V, t) \) the market value of debt.

In detail, we have the following simplifying hypotheses:

1. **Equity, seen as the residual claim.**
   According to the firm’s capital structure, if at \( T \) the bondholders will receive their debt in full, the equity holders receive the remaining part of the asset. This is the reason why Merton gives to the equity its option nature: it is seen as a Call option on the underlying firm’s asset, with strike price \( F \) equal to the face value of the firm’s debt. At maturity \( T \) its value is

\[
E(V(T), T) = \max(V(T) - F, 0). \tag{2.1}
\]

2. **A single, homogeneous class of debt.**
   The firm promises to pay the bond whose face value is \( F \) to the bondholders at maturity \( T \). In the event this payment is not met, the bondholders take over the company and the shareholders receive nothing.

   By referring to the Put-Call parity and to the capital structure, the debt value at maturity can be written as:

\[
D(V(T), T) = F - \max(F - V(T), 0) = \min(V(T), F). \tag{2.2}
\]

\(^3\)The definition of an european option is explained in Appendix B.
2.1 Merton’s model

The company cannot issue any new senior claims and cannot pay cash dividends prior to the maturity date of debt.

3. Firm’s assets dynamics.
The firm’s assets are tradable assets and follow a Geometric Brownian motion:

\[ dV = V(\mu_V dt + \sigma_V dW), \quad (2.3) \]

with \( \mu_V \) instantaneous expected rate of return, \( \sigma_V \) the volatility of the firm and \( dW \) the standard Wiener process (or Brownian Motion). Consequently \( V(t) \) is lognormal distributed with value at time \( t \):

\[ V(t) = V(0)e^{(\mu_V - \frac{1}{2}\sigma_V^2)t + \sigma_V \sqrt{t} W_t}, \]

where \( \mu_V = r - \frac{1}{2}\sigma_V^2 \) and \( V(0) \) is the value of the assets at an initial date \( t = 0 \).

2.1.2 Model setting

According to assumption 2.1, the market value of firm’s equity is modeled as a call option on the firm’s asset \( V \) with maturity \( T \) and strike price equal to the face value of its debt, \( F \). By using the Black and Scholes pricing formula we can write the equity’s value at time \( t = 0 \) as:

\[ E(0, V(0)) = V(0)N(d_1(0)) - Fe^{-\mu_V T}N(d_2(0)), \quad (2.4) \]

where

\[ d_1(0) = \ln\left(\frac{V(0)}{F}\right) + (\mu_V + \frac{1}{2}\sigma_V^2)T \]

\[ \sigma_V \sqrt{T}, \quad (2.5) \]

\[ d_2(0) = d_1(0) - \sigma_V \sqrt{T}, \quad (2.6) \]

and \( N \) represents the cumulative Normal distribution function. The pricing formula holds for every successive time \( t \in [0, T] \):

\[ E(t, V(t)) = V(t)N(d_1(t)) - Fe^{-\mu_V T}N(d_2(t)) \quad (2.7) \]

---

4. The definition of the Geometric Brownian Motion is illustrated in Appendix A
5. The Black and Scholes pricing formula has been derived in Appendix B.4
2.1 Merton’s model

with

\[ d_1(t) = \frac{\ln\left(\frac{V(t)}{F}\right) + (\mu_V + \frac{1}{2}\sigma_V^2)\tau}{\sigma_V \sqrt{\tau}} \]  

(2.8)

\[ d_2(t) = d_1(t) - \sigma_V \sqrt{\tau} \]  

(2.9)

where \( \tau = T - t \) is the length of time.

Eq. 2.1 assumes that firm’s equity follows a Geometric Brownian Motion

\[ dE = E(\mu_E dt + \sigma_E dW) \]  

(2.10)

with \( \mu_E \) the instantaneous expected rate of return per unit time and \( \sigma_Y \) the instantaneous variance of return. By Ito’s lemma, we can write the dynamics of \( E(V(t),t) \) as

\[ dE = E_V dV + E_t dt + \frac{1}{2} E_{VV} (dV)^2. \]

Since we know that, since \( V \) is a Geometric Brownian Motion it is \( (dV)^2 = \sigma_V^2 V^2 dt \) and then, substituting \( dV \) from equation 2.3 it becomes

\[ dE = \left( \frac{1}{2} \sigma^2 V^2 E_{VV} + E_t + \mu_V V E_V \right) dt + \sigma_V V E_V dW. \]

Comparing diffusion terms in 2.10 we have the following system

\[
\begin{cases}
\mu_E F = \frac{1}{2} \sigma^2 V E_{VV} + E_t + \mu_V V E_V \\
\sigma_E E = \sigma_V V E_V \\
dW_E = dW
\end{cases}
\]  

(2.11)

From pricing theory can be shown that the delta of a call coincides with \( N(d_1) \). Hence, the second equation of the system can be written as

\[ \sigma_E = \frac{V}{E} N(d_1) \sigma_V. \]  

(2.12)

This model linking equity and asset volatility holds only instantaneously. Actually the market leverage does not behave exactly in that way. For example, if the market leverage is fast-growing then 2.12 will tend to underestimate asset volatility and thus default probability will be minimized as the firm’s credit risk declines. Conversely, if the market leverage is decreasing

\textsuperscript{6}Ito’s Lemma statement and an intuition of the proof are analyzed in deep in Appendix A
2.1 Merton’s model

rapidly then 2.12 will overestimate asset volatility. As a consequence the probability of default will be overstated as the firm’s credit risk improves. Asset volatilities calculated on the basis of the above system generate default probabilities with very little discriminatory power. This is one of the main drawbacks of the model.

2.1.3 Distance to Default: a definition

The concept of distance to default is expressed as the distance between firm’s asset and the default point, relative to firm’s asset volatility $\sigma_V$. The default point is constant, fixed at the face value of debt. Formally we can define the default probability under risk neutral measure as

$$P(t) = Pr\{V(T) < F\} = Pr\{V(t)e^{(r-\frac{\sigma_V^2}{2})\tau + \sigma_V W_\tau} < F\}.$$  

Finally we have

$$Pr\{z < \frac{\ln(F/V(t)) - (r - \frac{\sigma_V^2}{2})\tau}{\sigma_V \sqrt{\tau}}\} = Pr\{z < -\frac{\ln(V(t)) + (r - \frac{\sigma_V^2}{2})\tau}{\sigma_V \sqrt{\tau}}\} = N(-d_2(t)) = N(-DD).$$

So, we can call Distance-to-default the risk measure directly related to the creditworthiness of the firm:

$$DD = \frac{\ln(F/V(t)) + (r - \frac{\sigma_V^2}{2})\tau}{\sigma_V \sqrt{\tau}}.$$  \hspace{1cm} (2.13)

Therefore, Merton’s default probability can be calculated as

$$\pi_{merton} = N(-DD).$$  \hspace{1cm} (2.14)

With this default probability most papers refers to the Expected Default Frequency, that corresponds to

$$EDF = N(-DD).$$

Notice that the smaller is DD, the higher the probability of default. Summarizing, the firm’s default over some horizon, from $t = 0$ until $H$, is described by six variables:

1. The current asset value $V(t)$;
2. The distribution of asset value at maturity $H$;
3. The volatility of the asset value at $H$;

4. The level of the default point;

5. The expected rate of growth of the asset over the horizon;

6. The length of the horizon, $H$.

Figure 2.1: Distance to default. Source: Crosbie and Bohn, *Modeling Default Risk*.

An important note to discuss here is the role of the Normal distribution. The model defines the default point as a constant, but in practice it is a random variable. As a firm is near default, it tends to adjust its liabilities. For example, it is common that in commercial firms there is an increase of debt, the opposite happens for financial institutions as they are approaching default. The difference reflects the level of liquidity of the corporates’ asset, and then their capacity to modify their leverage in that situation. For this reason, the uncertainty related to stochastic movements of liabilities can be captured by an empirical distribution, rather than a simple Normal one. Moody’s KMV uses its large historical database to estimate an empirical distribution, and it calculates default probabilities based on that distribution.

The opinion of Hans Ne Byström(2007) is different: he argues that the assumption that the default barrier is equal to the book value of debt is not
necessary wrong; it is the book value of debt that has to be paid back, not its market value.

## 2.2 Merton’s non-linear system

The fundamental problem of Merton’s approach has to be identified with the valuation of the firm’s asset and its volatility since they are not directly observable in the market. For this reason, Merton focuses on a system of two non-linear equations in two variables: 2.7 and 2.12, through which firm’s asset and volatility can be derived.

\[
\begin{align*}
E(t) &= V(t)N(d_1(t)) - e^{-r\delta t}FN(d_2(t)) \\
\sigma_E(t) &= \frac{V(t)}{F}N(d_1(t))\sigma_V(t),
\end{align*}
\]

with

\[
\begin{align*}
d_1(t) &= \frac{\ln\left(\frac{V(t)}{F}\right) + (r + 0.5\sigma_V(t)^2)\delta t}{\sigma_V(t)\sqrt{\delta t}}, \\
d_2(t) &= \frac{\ln\left(\frac{V(t)}{F}\right) + (r - 0.5\sigma_V(t)^2)\delta t}{\sigma_V(t)\sqrt{\delta t}}, \\
\delta t &= T - t.
\end{align*}
\]

The first equation allows us to relate the market value of equity with the firm’s asset value, while the second one governs the relationship between equity volatility and firm’s volatility. The solution of the previous non linear equations represents the most difficult aspect of Merton’s model. This section is aimed at describing some procedures used to solve it and at presenting an alternative calculation.

Research papers adopt different methods to extract numerical solutions. Many researches and practitioners have pointed out that in general the real behavior of assets and volatility is not compatible with the above system. Crosbie and Bohn (2003) explain that “In practice the market leverage moves around far too much for the second equation of 2.15 to provide reasonable results.” For this reason they and Vassalou and Xing (2004) implemented a more complex iterative procedure. Barath and Schumway (2008) follow their line of thought and adopt the same algorithm. It proceeds as follows:

1. They start with an initial value of \( \sigma_V = \sigma_E\left(\frac{E}{E+F}\right) \), which is used in the second equation of 2.15 to infer the market value of firm’s assets every day from the previous year;
2. Then, they calculate the implied log return on assets each day in order to generate new estimates of $\sigma_V$ and $\mu$;

3. They iterate on $\sigma_V$ in this manner until it converges, with tolerance fixed at $10^{-3}$.

Fitch Solutions (2007) criticizes this iterative procedure. As discussed in Ericsson and Reneby (2005), the procedure assumes constant equity volatility, which is inconsistent with a structural model. Moreover this approach underestimates volatility for those firms with rapidly falling asset value due to changes in leverage. For this reason Fitch’s Analysts use a double loop convergence approach. More specifically:

1. They fix an initial value of $\sigma_V = \sigma_F$ and adopt Euler’s method to solve for the unobservable asset value given in the first equation of 2.15;

2. Once they obtain the asset value, the corresponding volatility is calculated for the next loop, and the process is repeated until the calculated volatility difference between iterations converges to below a fixed tolerance.

In a Dissertation of Oxford by Lu (2008), another iterative approach is presented:

1. The initial volatility of the firm’s asset is replaced by the volatility of equity; substituting this new value in the first function of 2.15 he derives the corresponding value of the firm’s asset.

2. the value of the firm’s asset calculated in step 1 is used in the second function in equation 2.15 to get the corresponding volatility of equity.

3. If equity volatility calculated in step 2 is equal to the real volatility of equity, the program stops. Otherwise, we need to readjust the volatility of the firm’s asset, and iterate the step 1 and 2 until the condition in step 3 is reached.

### 2.2.1 Merton’s system with rescaled variables

My approach is totally different from the procedures described in the previous section. I focus on the non-linear system structure and by studying the nature of the unknown variables I work on a riparamaterisation. First of all my aim is to calculate annual default probabilities and for this
reason I would like to extract from Merton’s system only firm’s annual asset and volatility values for each year from 1990 to 2011. Hence, it can be written in a simplified form by fixing the maturity of debt $T = 1$ year:

\[
\begin{cases}
    E = VN(d_1) - e^{-r}FN(d_2) \\
    \sigma_E = \frac{V}{F}N(d_1)\sigma_V
\end{cases}
\]  

(2.16)

with

\[
\begin{align*}
    d_1 &= \frac{\ln\left(\frac{V}{F}\right) + r + 0.5\sigma_V^2}{\sigma_V} \\
    d_2 &= \frac{\ln\left(\frac{V}{F}\right) + r - 0.5\sigma_V^2}{\sigma_V}
\end{align*}
\]

The unknown variables, $V$ and $\sigma_V$, have totally different order of magnitude. To get an idea, the first one has a magnitude of about $10^9$ while the second is around $10^{-1}$. My approach is to obtain a new version of the system and get an easier formulation. Then the numerical solutions are calculated directly in matlab by using the fsolve command. Let’s denote with $\lambda'$ and $\eta'$ the following ratios:

\[
\begin{align*}
    \lambda' &= \frac{V}{F} \\
    \eta' &= \frac{\sigma_V}{\sigma_E}
\end{align*}
\]

and then the system becomes:

\[
\begin{cases}
    \lambda'N(d_1) - e^{-r}N(d_2) - \frac{E}{D} = 0 \\
    \lambda'\eta'N(d_1) - \frac{E}{D} = 0
\end{cases}
\]  

(2.17)

with

\[
\begin{align*}
    d_1 &= \frac{\ln(\lambda') + r + 0.5\eta'^2\sigma_E^2}{\eta'\sigma_E} \\
    d_2 &= \frac{\ln(\lambda') + r - 0.5\eta'^2\sigma_E^2}{\eta'\sigma_E'}
\end{align*}
\]

Now the unknowns of the system are $\lambda'$ and $\eta'$. Thanks to this transformation the problem related to the strong divergence between the original
unknown variables is solved. The ratio $\eta'$ can be very small because the market value of firm’s equity is more volatile than its assets. For this reason, in order to avoid potential negative solutions caused by the optimization algorithm of fsolve, I impose a logarithmic transformation. Thus, let consider $\lambda = \ln(\lambda')$ and $\eta = \ln(\eta')$. After this change of variables the first equation can be written as:

$$e^\lambda N(d_1) - e^{-r} N(d_2) - \frac{E}{D} = 0$$  \hspace{1cm} (2.18)

![3D plot](image)

Figure 2.2: 3D plot of $e^\lambda N(d_1) - e^{-r} N(d_2) - \frac{E}{D}$.

The second equation becomes:

$$e^{\lambda + \eta} N(d_1) - \frac{E}{D} = 0$$  \hspace{1cm} (2.19)

Equations 2.18 and 2.19 constitute the new version of Merton’s non-linear system:

$$\begin{cases}
  e^\lambda N(d_1) - e^{-r} N(d_2) - \frac{E}{D} = 0 \\
  e^{\lambda + \eta} N(d_1) - \frac{E}{D} = 0
\end{cases}$$  \hspace{1cm} (2.20)
2.2 Merton’s non-linear system

Figure 2.3: 3D plot of $e^{\lambda + \eta}N(d_1) - \frac{E}{D}$.

with

\[
d_1 = \frac{\lambda + r + 0.5e^{2\eta}\sigma_E^2}{e^{\eta}\sigma_E}
\]

\[
d_2 = \frac{\lambda + r - 0.5e^{2\eta}\sigma_E^2}{e^{\eta}\sigma_E}
\]

Next Figure reports the graphical representation of 2.18 and 2.19 in function of the two unknown variables $\lambda$ and $\eta$. The first is depicted in blue while the second one in gold. The graph is created for a specific firm in a precise year; the parameters correspond to the company A2A in 2000.

Finally, the system’s solution can be obtained easily by using the command `fsolve` in MATLAB, with initial point:

\[
P_0 = [\lambda_0, \eta_0] = [\ln(1 + \frac{E}{D}), 0].
\]

Matlab commands are reported in the next page. It includes the definition of the two functions employed for the non-linear system and a short algorithm for its implementation.
2.3 The Naïve predictor

Sreedhar T. Barath and Tyler Shumway (2008) proposed a simplified version of Merton’s model. They highlight that the model’s strong point can be found in its functional form, rather than in its forecasting power. Their purpose is to show that the default probability enhanced by Merton is not a sufficient statistic to predict bankruptcy, i.e. is possible to construct a reduced-form model with better forecasting properties.

This Naïve alternative preserves Merton’s functional form, which cannot be completely replaced by a linear combination of variables used to calculate the probability.

The model focuses on two main objectives: it has to perform as Merton model, and has to be simple. In fact it does not require to solve the non-linear system represented in Section 2.2. The Naïve alternative is based on the strong assumption that the market value of firm’s debt coincides with its face value $F$:

\[ D_{\text{naive}} = F \]

And that the riskiness of debt is correlated to equity risk in this measure:

\[ \sigma_{D_{\text{naive}}} = 0.05 + 0.25\sigma_E. \]
% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
% Workbook: C:\Users\birobe\Documents\MATLAB\MertonSystemDataset.xlsx
% Worksheet: Sheet1
% To extend the code to different selected data or a different spreadsheet,
% generate a function instead of a script.
% Auto-generated by MATLAB on 2014/01/24 17:43:33
% Import the data
% [~, ~, raw] = xlsread('C:\Users\birobe\Documents\MATLAB\MertonSystemDataset.xlsx', 'Sheet1');
raw = raw(2:1483,:);
% Replace non-numeric cells with 0.0
R = cellfun(@(x) ~isnumeric(x) || isnan(x), raw);
raw(R) = {0.0};
% Create output variable
A = cell2mat(raw);
% Clear temporary variables
clearvars raw R;

%Merton non linear system resolution with fsolve.
%Variables:
%V= asset value; sigmaV= sigma asset; E=equity; sigma=sigma equity;
%D=total current liabilities+0.5*long term debt; T=maturity=1 year; r= risk-free interest rate.
%EonD=E/D
%Unknowns: sol=[l(1), l(2)]:
% l(1)=lambda=log(V/D);
% l(2)=eta=log(sigmaV/sigma);
% y=[V,sigmaV]=[exp(l(1))*D, exp(l(2))*sigma]
D=A(:,4)+0.5*A(:,6);
EonD=A(:,3)./D;
sigma=A(:,10);
sol=zeros(length(A),2);
y=zeros(length(A),2);
options=optimset('MaxFunEvals',20000, 'MaxIter',10000, 'TolFun',10^-10, 'TolX',10^-10);
for i=1:length(A)
sol(i,:)=fsolve(@(l)f1(l,EonD(i),1,A(i,10),A(i,12)),[log(1+EonD(i)),1], options);
end
% Vector of Non-linear functions:
function y=f1(l,EonD,T,sigma,r)
y=[exp(l(1))*(1-erf((l(1)+(r+0.5*(sigma^2)*exp(2*l(2))))/(sigma*exp(l(2))*sqrt(T)*sqrt(2))))/2-exp(-r*T)*(1-erf(-(l(1)+(r-0.5*(sigma^2)*exp(2*l(2))))/((sigma*exp(l(2))*sqrt(T)*sqrt(2))))))]/2-EonD;
end

Figure 2.4: Matlab program for the solution of Merton’s system with rescaled variables.
The equation is justified by the fact that firms close to default have very risky debt. The first element represents the term structure volatility while the second element is the percentage of equity risk related to debt risk. Instead of solving the non-linear system, Sreedhar T. Barath and Tyler Shumway approximate asset volatility by adding the percentage of equity risk and debt risk contained in the firm’s asset:

\[
\sigma_{V_{\text{naive}}} = \frac{E}{E + D_{\text{naive}}} \sigma_E + \frac{D_{\text{naive}}}{E + D_{\text{naive}}} \sigma_{D_{\text{naive}}}
\]

\[
\sigma_{V_{\text{naive}}} = \frac{E}{E + F} \sigma_E + \frac{F}{E + F} (0.05 + 0.25 \sigma_E)
\]

(2.21)

and

\[
\text{naive } V = E + F
\]

(2.22)

Finally, they set the expected return on the firm’s asset equal to the firm’s stock return of the previous year, in order to capture the same information of the iterative procedure used to solve Merton’s non linear system:

\[
\mu_{\text{naive}} = \mu_{it-1}
\]

To calculate the naïve Distance to default and the Expected default frequency, they use the same structure of Merton’s model:

\[
DD_{\text{naive}} = \ln\left(\frac{E+F}{F}\right) + \frac{(\mu_{it-1} - 0.5 \sigma_{V_{\text{naive}}}^2)T}{\sigma_{V_{\text{naive}}} \sqrt{T}}
\]

\[
\pi_{\text{naive}} = N(-DD_{\text{naive}}).
\]

The choices made by Sreedhar T. Barath and Tyler Shumway are totally arbitrarily, but they proved that the easy Naïve predictor has better forecasting power than Merton’s one.

For what concerns the choice of the face value of debt, Barath and Shumway fix \( F \) as total current liabilities plus 0.5 of long-term debt. But many researchers have made different choices. For example, Huang and Huang (2002) and others have pointed out that firms often continue to operate with negative net worth. They also notice that the implied default costs must be extremely high to explain the relatively low recovery rates on corporate bonds. It is more reasonable to specify a default point that is some fraction \( \beta \leq 1 \) of debt principal.

Hans Ne Byström (2007) did his studies on both non-financial firms and US bank holding companies. He used total debt to have a proxy of the its
2.3 The Naïve predictor

face value. This choice has been justified by saying that the exclusion of long-term debt has two different reasons. First, when firms service their long-term debt, the interest payments are part of their short-term debt. Second, the larger the long-term debt, the harder it is for a firm to roll-over its short-term debt. This increases the likelihood of default.

However, in Merton-Naïve comparative study the exact specification of debt level is important but not critical. My choice is to follow Barath and Schumway (2008), Crosbie and Bohn (2008), Dionne et al. (2006) Fitch Solution (2007) and to fix the face value of debt equal to total current liabilities plus one half of long-term debt.
Chapter 3

Data description and The Default-risk Zone

The aim of this Chapter is to introduce two topics. The first one includes a description of the data source and the creation of the dataset. The second one is related to the identification and selection of potential defaulters contained in my sample and a summary statistic.

3.1 Data

My analysis encompasses time series of publicly traded active firms in the Italian market from 2000 to 2011\(^1\) without financial institutions. For the data collection I use Thomson Reuters datastream. I extract annual values from datastream’s ACTIVE series, while market closing prices are extracted quarterly. Below are listed all the selected items.

- Market value;
- Market closing price;
- Current liabilities total;
- Long term debt;
- Total assets;

\(^1\)Initially the dataset contained time series from 1990 but I considered only this interval of time because companies had no data available before 2000.
3.2 Default-risk Selection Criteria

- Net income;
- Earnings before interests and taxes (EBIT);
- Interest coverage ratio;
- Assets to equity ratio;
- Debt to Equity ratio;

The scale factor used by datastream is variable; market value is in millions of euro while financial statement data are expressed in thousands of euro. The initial dataset contains 189 publicly traded active firms. After cleaning data, the final sample is characterised by 173 Italian companies. This is the sample used for the statistical analysis reported in Section 3.2.4 and the probit estimation described in Chapter 4.

3.2 Default-risk Selection Criteria

Although Thomson Reuters datastream is very comprehensive, it does not offer any default indicator for Italian firms. For this reason, before the analysis starts, I need to define a set of companies with high risk profile which I consider as potential defaulters. Hence, I focus on financial criteria to assess whether a firm belongs to a “Default-risk Zone” or not. My strategy consists in two steps. The first step is to filter the companies contained in my sample on the basis of their interest coverage ratio. The interest coverage ratio belongs to the category of Coverage Ratios and focuses on the income statement. It measures the ability of a firm to cover its debt payments, i.e. how many times a company’s EBIT\(^2\) could cover its interest payments. Formally, it is computed as the ratio between EBIT and Interest payments:

\[
IC = \frac{EBIT}{IP}.
\]

A higher interest coverage ratio indicates stronger solvency, offering a greater assurance that the firm can service its debt from operating earnings. Together with other financial indicators, the interest coverage is used to assess the firm’s level of credit risk through a synthetic rating\(^3\).

\(^2\)EBIT stands for earnings before interests and taxes
\(^3\)A synthetic rating consists in playing the role of a ratings agency and assign a rating to a firm based upon its financial ratios.
Table 3.1 is the result of the estimation of a sample of US companies’ credit risk. It reports the correspondence between interest coverage ratios and ratings. In the first case, an interest coverage greater than 12.50 indicates that the firm’s EBIT is 12.5 times greater than its interest payments. As a consequence, the corresponding company is assigned to a high rating class because of its wide capacity to repay its interest payments. The same procedure is applied to every firm for which an interest coverage ratio is identified and a specific rating assigned. The ratings AAA and AA mean that the company has a high credit quality; A and BBB mean medium credit quality. Firms which possess credit ratings below the designations BB, B, CCC are considered to be low credit quality, and their bonds are commonly referred to as “junk bonds”.

Hence, in order to find risky companies, I restrict the sample by considering those firms with interest coverage ratio below the value of 3.

In the second step I focus on a particular category of financial ratios. Extensive academic research has examined the importance of ratios not only in predicting stock returns but also credit failure. Altman (1968), Ohlson (1980), Hopwood et al. (1994) and many others have found that financial statement ratios are effective in selecting investments and in predicting financial distress. Ratio analysis is useful to derive and communicate value of companies,
3.2 Default-risk Selection Criteria

but it is just an “indicator” of some aspects of a company’s performance. Financial ratios provide insights of a company’s financial flexibility or ability to obtain the cash required to grow and meet its obligations, even if unexpected circumstances develop. Common ratio categories include activity, liquidity, solvency, profitability, and valuation. Each category measures a different aspect of the company’s business, but all of them are useful in evaluating a company’s overall ability to generate cash flows from operating its business and the associated risks.

I limit my attention only to **solvency ratios**. Solvency refers to a company’s ability to fulfil its long-term debt obligations. Assessment of a firm’s ability to pay for them (i.e. to make interests on principal payments) generally includes an in-depth analysis of the components of its financial structure. Solvency ratios provide information regarding the relative amount of debt in the company’s capital structure and the adequacy of earnings and cash flow to cover interest expenses and other fixed charges.

My strategy is to consider the set of active firms with interest coverage ratio greater than 3, i.e. those companies with a high rating class. In this sample, I look at the distributions of the following solvency ratios:

1. **Debt-to-book equity Ratio**. This ratio belongs to the category of Debt Ratios; it focuses on the balance sheet and measure the amount of debt capital relative to equity capital.

2. **Financial leverage Ratio**. It is the ratio between a firm’s total assets and book equity. The higher the financial leverage ratio, the more leveraged the firm is in the sense of using debt to finance assets.

3. **Debt-to-equity Ratio**. Its meaning is the same as Debt-to-book equity ratio, but here the equity corresponds to its market value.

I refer to the first two measures as an indicator of a firm’s balance-sheet state and to the third one as a firm’s market situation.

The last step is to take in each corresponding distribution the value to the 75th quantile which represents a benchmark for my classification.

Hence, in a specific year, a firm with interest coverage ratio less than 3 is defined at risk if its three ratios are greater than the corresponding benchmark values. This procedure is applied on the companies’ sample for every year.

---

4More other parameters should be considered such as current ratio, working capital over total assets, retained earnings-to-assets ratio. But in Datastream most of that values were not available for all the Italian companies.
3.2 Default-risk Selection Criteria

Table 3.2 reports the values relative to the 75th quantile of the three distributions from 2000 to 2011.

**Table 3.2: Benchmark values**

<table>
<thead>
<tr>
<th>Year</th>
<th>TD/BE</th>
<th>TA/BE</th>
<th>TD/E</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>1.28</td>
<td>3.90</td>
<td>0.89</td>
</tr>
<tr>
<td>2001</td>
<td>1.17</td>
<td>3.53</td>
<td>0.70</td>
</tr>
<tr>
<td>2002</td>
<td>1.09</td>
<td>3.31</td>
<td>0.66</td>
</tr>
<tr>
<td>2003</td>
<td>1.31</td>
<td>3.99</td>
<td>0.84</td>
</tr>
<tr>
<td>2004</td>
<td>1.19</td>
<td>3.73</td>
<td>0.87</td>
</tr>
<tr>
<td>2005</td>
<td>1.17</td>
<td>3.42</td>
<td>0.70</td>
</tr>
<tr>
<td>2006</td>
<td>1.38</td>
<td>4.06</td>
<td>0.73</td>
</tr>
<tr>
<td>2007</td>
<td>1.27</td>
<td>3.69</td>
<td>0.61</td>
</tr>
<tr>
<td>2008</td>
<td>1.24</td>
<td>3.96</td>
<td>0.72</td>
</tr>
<tr>
<td>2009</td>
<td>1.03</td>
<td>3.44</td>
<td>0.96</td>
</tr>
<tr>
<td>2010</td>
<td>1.07</td>
<td>3.53</td>
<td>0.90</td>
</tr>
<tr>
<td>2011</td>
<td>1.31</td>
<td>3.77</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Table 3.2: This Table reports values of the 75th quantile of the three ratios’ distributions of active firms with interest coverage ratio greater than 3. $TD/BE$ is Total debt-to-book equity ratio, $TA/BE$ is total assets-to-book equity ratio and $TD/E$ is total debt-to market value of equity ratio.

It is possible that a firm which in 2000 lied in the Default-risk Zone, in 2001 may not satisfy the criteria and, as a consequence, it can be pulled out. This dynamic strategy allows companies to enter or exit from the Default-risk Zone. Table 3.3 shows the number of companies defined at risk. In 2009 the firms are 39 and it is the highest number in the sample. 2010 and 2011 follow with, respectively, 30 and 22 firms. On the contrary, in 2000 potential defaulters are only 4, followed by 6 firms in 2006.
Firms in the Default-risk Zone.

<table>
<thead>
<tr>
<th>Year</th>
<th>No. of firms at risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>4</td>
</tr>
<tr>
<td>2001</td>
<td>7</td>
</tr>
<tr>
<td>2002</td>
<td>20</td>
</tr>
<tr>
<td>2003</td>
<td>15</td>
</tr>
<tr>
<td>2004</td>
<td>17</td>
</tr>
<tr>
<td>2005</td>
<td>16</td>
</tr>
<tr>
<td>2006</td>
<td>6</td>
</tr>
<tr>
<td>2007</td>
<td>15</td>
</tr>
<tr>
<td>2008</td>
<td>19</td>
</tr>
<tr>
<td>2009</td>
<td>38</td>
</tr>
<tr>
<td>2010</td>
<td>30</td>
</tr>
<tr>
<td>2011</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 3.3: Number of firms with high risk profile for each year from 2000 to 2011.

3.3 Merton and Naive: Summary statistic

The inputs of Merton’s model include the market value of equity $E$, the face value of debt $F$, the risk free rate $r$ $^5$, the time period $T$ (which is equal to 1 year and for this reason is omitted in the formula), and the annualized standard deviation of stock returns $\sigma_E$. The last one is obtained from historical returns. We get market values of asset and its volatility for each firm by solving the Merton’s non-linear system with rescaled variables described in Section 2.2.1. Hence, Merton distance to default is calculated as:

$$DD_{merton} = \frac{\ln \left( \frac{F}{V} \right) + \mu - 0.5\sigma_E^2}{\sigma_V}$$ \hfill (3.2)

Naive probability is less complicated than Merton’s and needs a smaller set of parameters: market value of equity and its volatility, $F$, $r$ and returns on assets of the previous year $\mu_{t-1}$:

$$DD_{naive} = \frac{\ln \left( \frac{E+F}{F} \right) + \mu_{t-1} - 0.5\sigma_{V_{naive}}^2}{\sigma_{V_{naive}}}$$ \hfill (3.3)

$^5$The risk free rate is obtained from the Treasury Department, available at www.dt.tesoro.it/it/debito_pubblico/dati_statistici/principali_tassi_di_interesse
In the first place, a comparison between Merton’s model and the Naïve alternative can be made by looking at descriptive statistic, reported in Table 3.4. An important remark should be made about the strong similarity between Naïve and Merton’s predictors. The distribution of $\pi_{merton}$ is highly similar to the Naïve alternative $\pi_{naive}$, as it is between $DD_{naive}$ and $DD_{merton}$. The same occurs for the distributions of $\sigma_{V_{merton}}$ and $\sigma_{V_{merton}}$. This is confirmed also by looking at the strong correlation between them. We can notice that the two volatilities, $\sigma_{naive}$ and $\sigma_{merton}$, have the highest correlation (0.9694), followed by the correlation between $DD_{merton}$ and $DD_{naive}$ (0.8551) and between $\pi_{naive}$ and $\pi_{merton}$ (0.6450). Thus the two default measures move together.
Table 3.4: Summary statistic

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$</td>
<td>2120</td>
<td>7980</td>
<td>1.16</td>
<td>100000</td>
</tr>
<tr>
<td>$F$</td>
<td>2040</td>
<td>6920</td>
<td>2.621</td>
<td>65000</td>
</tr>
<tr>
<td>$\sigma_E(%)$</td>
<td>15.62</td>
<td>13.82</td>
<td>1.35</td>
<td>308.25</td>
</tr>
<tr>
<td>$\mu(%)$</td>
<td>0.27</td>
<td>3.18</td>
<td>-37.62</td>
<td>23.75</td>
</tr>
<tr>
<td>$\mu_{t-1}(%)$</td>
<td>0.43</td>
<td>3.88</td>
<td>-30.78</td>
<td>94.37</td>
</tr>
<tr>
<td>$r(%)$</td>
<td>4.52</td>
<td>0.59</td>
<td>3.52</td>
<td>5.72</td>
</tr>
<tr>
<td>$\sigma_{V_{naive}}(%)$</td>
<td>12.24</td>
<td>10.33</td>
<td>1.9</td>
<td>287.79</td>
</tr>
<tr>
<td>$V_{naive}$</td>
<td>4170</td>
<td>13800</td>
<td>15.5</td>
<td>142000</td>
</tr>
<tr>
<td>$\pi_{naive}(%)$</td>
<td>1.32</td>
<td>5.92</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>$\sigma_{V_{merton}}(%)$</td>
<td>8.07</td>
<td>10.4</td>
<td>0.07</td>
<td>301.5</td>
</tr>
<tr>
<td>$V_{merton}$</td>
<td>4070</td>
<td>13600</td>
<td>14.8</td>
<td>141000</td>
</tr>
<tr>
<td>$DD_{naive}$</td>
<td>9.099</td>
<td>9.272</td>
<td>-11.744</td>
<td>139.816</td>
</tr>
<tr>
<td>$DD_{merton}$</td>
<td>14.0599</td>
<td>13.984</td>
<td>-85.261</td>
<td>170.3994</td>
</tr>
<tr>
<td>$\pi_{merton}(%)$</td>
<td>1.76</td>
<td>12.09</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

$corr(DD_{naive}, DD_{merton})=0.8551$
$corr(\pi_{naive}, \pi_{merton})=0.6450$
$corr(\sigma_{naive}, \sigma_{merton})=0.9694$

Table 3.4: This Table provides summary statistic for the variables used in Merton and Naive models. $E$ is the market value of equity in millions of euro and is taken from datastream item MV. $F$ is the face value of debt in millions of euro (computed as total current liabilities plus 0.5 long-term debt). $r$ is the risk-free interest and $\mu$ is a proxy of the firms’ returns calculated as net income over total assets. $\mu_{t-1}$ is firm’s return of the previous year. $V_{merton}$ is the market value of firm assets in millions of euro and $\sigma_V$ is the asset volatility measured in percentage per annum. These last two variables are generated as the result of solving the new version of Merton’s non-linear system described in Section 2.2.1. $V_{naive}$ and $\sigma_{naive}$ are derived from equations 2.21 and 2.22. $\pi_{merton}$ and $\pi_{naive}$ are the expected default frequencies in percentage calculated by Merton and Naive models. The two distances to default are calculated from 3.2 and 3.3. $corr(DD_{naive}, DD_{merton})$ is the correlation between the two factors. The same calculation is made for $\sigma_{naive}$ and $\sigma_{merton}$, $\pi_{naive}$ and $\pi_{merton}$. 
Chapter 4

Forecasting Default-risk: Probit Models

A common approach in empirical credit risk literature is to use standard discrete choice models, such as logit or probit models. They are suitable methods for the selection of default main determinants. In particular, probit models are appropriate to validate Merton’s model since the distance to default is mapped into a default probability through a Standard Normal distribution.

4.1 Probit Models

In this Section I evaluate the performance of probit models in order to assess naïve predictive power and to show that Merton’s model is an insufficient statistic to predict a firm’s financial distress.

In particular, I try to identify the real impact of Merton and naïve risk measures represented by their distances to default ($DD_{merton}$ and $DD_{naive}$). Default risk is seen as a dichotomous variable $y$: if a firm belongs to the Default-risk Zone in the evaluated year then $y$ takes value 1, otherwise 0. Thus, the probability of observing the default event is modeled as a binary outcome $y$ and then estimated conditional upon a set of explanatory variables.

In probit models the default probability depends on the covariates $x_i$ through a standard regression for each firm $i$: $y_i = x_i \beta + u_i$.

The latter is then mapped into an estimated probability by a function $f()$:

$$y_i = f(x_{i,1}\beta_1 + x_{i,2}\beta_2 + .. + x_{i,n}\beta_n + u_i)$$ (4.1)
There are infinite functions, and some of these were developed as alternative models for estimation. In probit regressions the functional form $f(\cdot)$ is expressed by the Cumulative Normal Distribution:

$$f(x\beta) = \int_{-\infty}^{x\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = N(x\beta)$$

Figure 4.1: CDF of a Standard Normal variable

Then a probit model can be written as:

$$y_i = N(x_i\beta + u_i)$$  \hspace{1cm} (4.2)

where $N()$ is the cumulative distribution function of a standard normal variable.

The focus of the analysis is to evaluate the probability that $y_i$ is equal to 1. Formally, this statement can be translated in the following probit regression:

$$Pr(y_i = 1|x_{i,1}, x_{i,2},..., x_{i,n}) = N(x_i\beta)$$  \hspace{1cm} (4.3)

Some variations of Eq 4.3 are employed in Section 4.2.

### 4.1.1 The Marginal Effect

In regressions, $\beta_k$ measures the effect of the exogeneous variable $k$ on the average value of $y$. The average value of a dichotomous variable is equal to
the probability that it assumes the value of one. Hence, in the probit model we can write:

\[
\frac{dPr(y = 1)}{dx_k} = \frac{1}{2\pi} e^{-\frac{x_k^2}{2}} \beta_k = N(x\beta)\beta_k
\]

In the previous formula, \( \beta_k \) appears as a multiplicative factor and determines the sign of the effect, since the other factor is necessarily positive. But the effect of \( x_k \) on \( Pr(y = 1) \) is attenuated by a nonlinear function of \( x\beta \) (i.e. all \( k \) of the \( x_k \)'s). Thus, we see that the effect of a change in \( x_k \) on the probability of the response \( y = 1 \) is clearly related to, though not completely determined by, \( \beta_k \).

The sign of \( \beta_k \) determines the direction of the effect, and this effect tends to be larger when \( \beta_k \) gets larger. So, qualitatively, the interpretation of \( \beta_k \) is the same as in the linear regression model. But, since its magnitude varies with the values of the independent variables, the description of that effect is not so simple. In Section 4.2 the Tables of results 4.1 and 4.2 report marginal effects of the covariates in each probit regression.

### 4.1.2 Goodness of fit

One of the measures employed in the literature to analyse the goodness of fit of a probabilistic model is the Brier Score (Brier 1950), adopted to assess the relative explanatory power of alternative models of distress prediction. In our case, \( i \) represents the firm. Thus, for each company \( i \), the Brier score is computed as:

\[
\frac{1}{N} \sum_{i=1}^{N} (y_i - p_i)^2
\]

where \( N \) is the total number of companies, \( p_i \) is the estimated probability of default of firm \( i \) based on coefficient estimates of the probit regressions, and \( y_i \) is the actual realization of \( y \) for firm \( i \): default \( (y_i = 1) \) or non-default \( (y_i = 0) \).

The lower the Brier score is, the higher is the performance of the model.

### 4.2 Results

In this section the probit estimation's results are presented and commented.

The probit analysis is based on data in an interval between 2001 and 2011 for the sample of 173 active companies. In each year the number of firms at
risk can vary from a maximum of 39 in 2009 to a minimum of 6 in 2001. Several variations of Equation 4.3 are explored in the following, including different sets of regressors. Two groups of probit regressions are presented in this section. The first group includes regressions which are computed for each year observed. These probit models are estimated using only $DD_{naive}$ and $DD_{merton}$ as covariates. The second group consists of pooled probit models. They are tested for all the observations contained in the sample and are aimed to verify the validity of the first group’s results. In these regressions, Naïve and Merton are performed with other covariates which are used to calculate them.

### 4.2.1 Probit models per year

The first group consists of probit specifications made year by year. They are described in Model I and Model II:

- **Model I.** This model includes only the $DD_{naive}$ covariate and it may be written as:

  $$Pr(y_T|x_t) = N(\beta_{0,t} + \beta_{1,t}DD_{naive})$$

  for each year $T$.

- **Model II.** A second regression contains only the Merton’s distance to default $DD_{merton}$:

  $$Pr(y_T|x_t) = N(\beta_{0,t} + \beta_{1,t}DD_{merton})$$

  for each year $T$.

Table 4.1 reports marginal effects of $DD_{naive}$ and $DD_{merton}$ and the relative models’ performance indicated by the Brier score. The results show that, in Model I, $DD_{naive}$ is always significant with a negative marginal effect. Its sign indicates the direction of the regressor and the negativity is compatible with the theoretic meaning of distance to default: the greater the distance, the lower is the probability of default. This is a strong signal that the Naïve predictor’s contribution is relevant in forecasting default risk. On the contrary, Model II shows that $DD_{merton}$ is not significant in more than 50% of the years observed. Moreover, its Brier score is higher than in Model I except for year 2002, where the performances of the two models are the same. This values indicate that the probit regressions with $DD_{naive}$ perform better than those in which Merton’s distance
4.2 Results

to default is tested.

<table>
<thead>
<tr>
<th>Year</th>
<th>Model I</th>
<th></th>
<th>Model II</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DD\textsubscript{naive}</td>
<td>Brier</td>
<td>DD\textsubscript{merton}</td>
<td>Brier</td>
</tr>
<tr>
<td>2001</td>
<td>$-0.011^{**}$</td>
<td>0.0809</td>
<td>$-0.006$</td>
<td>0.0949</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.026)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>$-0.018^{**}$</td>
<td>0.165</td>
<td>$-0.0083^{*}$</td>
<td>0.1650</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>$-0.0159^{***}$</td>
<td>0.0932</td>
<td>$-0.0161^{**}$</td>
<td>0.1137</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.037)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>$-0.0211^{***}$</td>
<td>0.1081</td>
<td>$-0.0143^{***}$</td>
<td>0.1368</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.039)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>$-0.0150^{***}$</td>
<td>0.1092</td>
<td>$-0.0049^{*}$</td>
<td>0.1255</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>$-0.0021^{**}$</td>
<td>0.0527</td>
<td>$-0.0035$</td>
<td>0.0568</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>$-0.0137^{***}$</td>
<td>0.0950</td>
<td>$-0.0035$</td>
<td>0.1043</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>$-0.0155^{***}$</td>
<td>0.1065</td>
<td>$-0.0079^{**}$</td>
<td>0.1138</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>$-0.0343^{***}$</td>
<td>0.1735</td>
<td>$-0.0088$</td>
<td>0.1942</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.015)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>$-0.0347^{***}$</td>
<td>0.1398</td>
<td>$-0.0024$</td>
<td>0.1546</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>$-0.0191^{***}$</td>
<td>0.1053</td>
<td>$-0.0013$</td>
<td>0.1176</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.009)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Probit analysis computed per year. The second and third column describe Model I. The fourth and fifth represents Model II. \(DD\textsubscript{naive}\) is Naïve distance to default calculated by Eq. 3.3. \(DD\textsubscript{merton}\) is Merton’s distance to default represented by Eq. 3.2. B. score is the Brier score; it is a measure of goodness of fit and is an indicator of the models’ performance. The lower the Brier score, the better is the probit model.
4.2.2 Pooled probit models

The second group of models is described below:

- **Model III.** The probit model has only $DD_{naive}$ as covariate:
  \[
  Pr(y_t|x_t) = N(\beta_{0,t} + \beta_{1,t}DD_{naive})
  \]
  for all $t$ observations.

- **Model IV.** The probit regression contains only $DD_{merton}$. It is employed in order to check if it is significant to predict default:
  \[
  Pr(y_t|x_t) = N(\beta_{0,t} + \beta_{1,t}DD_{merton})
  \]
  for all $t$ observations.

- **Model V.** This model places side by side $DD_{naive}$ and other five covariates which are used to calculate Merton’s and Naïve predictors: the logarithm of firms market value of equity, $ln(E)$, the logarithm of the face value of debt $ln(F)$, a proxy of return on assets at time $t$ and $t-1$ calculated as net income over total assets ($\mu$, $\mu_{t-1}$), and the inverse of equity volatility $1/\sigma_E$:
  \[
  Pr(y_t|x_t) = N(\beta_{0,t} + \beta_{1,t}DD_{naive} + \beta_{2,t}ln(E) + \beta_{3,t}ln(F) + \beta_{4,t}\mu + \beta_{5,t}\mu_{t-1} + \beta_{6,t}\frac{1}{\sigma_E})
  \]
  for all $t$ observations.

- **Model VI.** The same model is computed with $DD_{merton}$ instead of $DD_{naive}$:
  \[
  Pr(y_t|x_t) = N(\beta_{0,t} + \beta_{1,t}DD_{merton} + \beta_{2,t}ln(E) + \beta_{3,t}ln(F) + \beta_{4,t}\mu + \beta_{5,t}\mu_{t-1} + \beta_{6,t}\frac{1}{\sigma_E})
  \]
  for all $t$ observations.
Table 4.2 reports pooled probit models with the corresponding marginal effects and Brier score. Each row contains a different probit regression which corresponds to Model III, Model IV, Model V and Model VI. Model III and IV confirm that both $DD_{naive}$ and $DD_{merton}$ contribute to default. In Model V, Naïve distance to default remains significant also when compared with all the other regressors. Among them only the proxy of asset’s returns is significant in default prediction. Its sign is negative as expected. In Model VI $DD_{merton}$ is placed side by side with all the covariates and its coefficient remains significant along with the coefficient of the asset’s returns. The significance of the latter confirms that Merton distance to default is not a sufficient statistic to forecast financial distress. Moreover, Model III has a Brier score smaller than Model IV. This is an indicator which tells us that Naïve predictor would improve probit performance with respect to Merton’s one. The data provides evidence that the marginal effect of $DD_{naive}$ is always greater than $DD_{merton}$, even if the difference is very small. The robust standard errors confirm the validity of my results.

<table>
<thead>
<tr>
<th>Pooled models: Marginal Effects and Brier score.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$DD_{naive}$</td>
</tr>
<tr>
<td>Model III</td>
</tr>
<tr>
<td>Model IV</td>
</tr>
<tr>
<td>Model V</td>
</tr>
<tr>
<td>Model VI</td>
</tr>
</tbody>
</table>

Table 4.2: This Table reports probit models tested for all the observations in the sample. $DD_{naive}$ and $DD_{merton}$ are Naïve and Merton’s distances to default. $E$ represents the market value of firm’s equity. $F$ is the face value of debt computed as current liabilities plus 0.5 long-term debt. $\mu$ is a proxy of assets’ returns and it is calculated as $\frac{NI}{TA}$, $\mu_{t-1}$ is a proxy of assets’ returns of the previous year. $1/\sigma_E$ is the inverse of firm’s equity volatility. The last column shows the Brier score calculated for each estimation. The coefficients represent each regressor’s marginal effect.
Conclusions

Default, formally defined as a firm’s inability to repay commercial and financial debts, has been the object of extensive analysis within financial economics. The problem of default-risk prediction consists of heterogenous factors that cannot be captured by only one predictor. This work examines the accuracy and the contribution of the Merton’s distance to default in forecasting a firm’s financial distress. It is compared with the Naive alternative proposed by Barath and Schumway (2008), which appears to be very similar in distribution. This similarity is put in evidence by the strong correlation between $\pi_{naive}$ and $\pi_{merton}$ and between the two distances $DD_{naive}$ and $DD_{merton}$.

Looking at probit models, $DD_{merton}$ does not appear to produce a sufficient statistic to predict default. In fact, we can see that other regressors, placed side by side with $DD_{merton}$, remain significant. Moreover, the same characteristics of Merton’s approach can be obtained by constructing a very simple model, which performs surprisingly well in probit models: empirical results show that $DD_{naive}$ is even significant when $DD_{merton}$ is not. However, in the pooled probit estimation, Merton’s distance to default contributes to forecast financial distress.

These results can be summarised in three statements. First, Merton’s model is not a sufficient statistic in forecasting default because of the heterogenous nature of the problem. Second, the analytical limit of Merton’s model can be overtaken through the construction of a default predictor which preserves its functional form and has the same forecasting power. Last, Merton’s formulation remains relevant in default prediction, thanks to its market-based approach and its structural form.
Appendix A

Stochastic calculus

A.1 Probability space and filtration

Probability theory represents our model for future events.

Definition A.1 (Probability space). A probability space is a triplet \((\Omega, \mathcal{F}, P)\) where:

1. \(\Omega\) is the sample space, the set of all possible outcomes;

2. \(\mathcal{F}\) is a \(\sigma\)-algebra on the space \(\Omega\), i.e. a finite collection of subsets of \(\Omega\), \(A_{i=1..n} \subset \Omega\), such that the following properties hold:
   - \(\Omega \in \mathcal{F}\);
   - \(\mathcal{F}\) is closed for unions: \(\forall A_1, \ldots, A_n \in \mathcal{F}, \bigcup_{i=1}^n A_i \in \mathcal{F}\);
   - \(\mathcal{F}\) is closed for the complement: \(\forall A \in \mathcal{F}, A^c \in \mathcal{F}\).

From the properties above, it is easy to show that

(a) \(\emptyset \in \mathcal{F}\), in fact, from properties 1 and 3 we have \(\emptyset = \Omega \setminus \Omega \in \mathcal{F}\);

(b) \(\mathcal{F}\) is closed for intersections: from properties 2 and 3 it follows that

\(\forall A_1, \ldots, A_n \in \mathcal{F}, \bigcap_{i=1}^n A_i \in \mathcal{F} \).

3. \(P\) is the probability, i.e. a normalized, positive measure on the events;
A.1 Probability space and filtration

- It is a function $P : (\Omega, \mathcal{F}) \to [0, 1]$;
- $P(\Omega) = 1$;
- $P$ is countably additive: if $(A_n)_n$ is a sequence of disjoint events, then
  \[ P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) \]

From the properties above, can be shown that $P(A^c) = 1 - P(A)$, and $P(\emptyset) = 0$.

Remark 1. By $A^c$ we denote the complement of the event $A$, $A^c = \Omega \setminus A$.

Hence, with the notion of probability, we relate to the chance of an event occurring.

Definition A.2 (Random variable). A real random variable $X$ on $(\Omega, \mathcal{F})$ is a mapping

\[ X : \Omega \to \mathbb{R} \]

\[ X(w) = \{x \in \mathbb{R} : w \in \Omega\} \]

such that $X$ is $\mathcal{F}$-measurable, i.e.

\[ \{X \leq x\} = \{w \in \Omega : X(w) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R} \]

Example A.1. Let consider the $\sigma$-algebra generated by a random variable $X$:

\[ \sigma(X) := \sigma(\{X \leq x\} : x \in \mathbb{R}) \]

It is the case of a subsigma-algebra on $\Omega$; in particular it is the smallest $\sigma$-algebra on $\Omega$ where $X$ is measurable.

Example A.2. In order to introduce a financial environment, assume that $\Omega$ is a path space, and $\mathcal{F}$ is the information we have at the end of the observations. Suppose that we are interested in finding the information at a specific time $t$. Then we may consider the following sub-$\sigma$-algebra

\[ \sigma(w(t)) = \sigma(\{w(t) \leq x\}, x \in \mathbb{R}^n) \]
A.2 Stochastic processes

Now we can relate to the notion of Filtration. Suppose to have a probability space, we define $\mathcal{F}_i$ a sequence of $\sigma$-algebras, such that

$$\mathcal{F}_i = \{\sigma_t \in (\Omega, \mathcal{F}, P) : \forall t \in [t_0, t_T], \sigma(t_0) \subseteq \sigma(t_1) \subseteq \ldots \subseteq \sigma(t_T)\}$$

When time passes, more information will be progressively revealed; thus, we can think to have a dynamical $\sigma$-algebra, and the $\sigma$-algebra at time $t_i$ contains more informations than at time $t_{i-1}$, for $i = 1, \ldots, T$. Usually, in Finance a time horizon $T$ is fixed. Then we are not interested in what will go on after $T$, and the information at time $T$ verifies $\mathcal{F}_T = \mathcal{F}$.

**Definition A.3** (Stochastic basis). The probability space $(\Omega, \sigma, P)$ with the filtration $\mathcal{F}_i$ defines a stochastic basis, or a filtered probability space:

$$(\Omega, (\mathcal{F}_i)_{t \leq T}, P)$$

A.2 Stochastic processes

Given a probability space $(\Omega, \mathcal{F}, P)$, stochastic processes are a sequence of random variables: $S = (S(t))_t$, such that, if we look at a specific time $t$ it is

$$S(t) : \Omega \rightarrow \mathbb{R}$$

$$\{S(t) \leq x\} \in \mathcal{F} \forall t, \forall x \in \mathbb{R}$$

**Definition A.4** (Adapted processes). Given a filtered space $(\Omega, (\mathcal{F}_i)_{t \leq T}, P)$, we define the process $S(t)_{t \leq T}$ to be adapted to the filtration $(\mathcal{F}_i)_{t \leq T}$ if it is known at time $t$:

$$\{S(t) \leq x\} \in \mathcal{F}_t \forall t$$

Given $t, \omega$ the adapted process is often written as

$$S = S(t, \omega) : \Omega \times [t_0, t_T] \rightarrow \mathbb{R}$$

Suppose to be in a particular state $\omega = \omega^*$, an event of the sample space $\Omega$; then the function defined by

$$S(t, \omega^*)$$

represents a path or a trajectory of the process. We can think of the dynamic of an asset if the event $\omega^*$ occurs.
A.3 Brownian motion and Ito’s Lemma

If we fix a particular time $t = t^*$, we have the function

$$S(t^*, \omega)$$

which it is simply a random variable.

The probability $P$ comes to play if we want to measure the stochastic process, i.e. if we need a distribution of $S$.

Of course we know the Cumulative Distribution Function of $S(t)$, the so called marginal distribution:

$$P(S(t) \leq x), \forall t$$

if we want information about the whole process $S$ we can define the joint distribution

$$P(S(t_1) \leq x_1, S(t_2) \leq x_2, \ldots, S(T) \leq x_T).$$

**Example A.3.** The joint distribution of $S(t_1)$ and $S(t_2)$ is

$$P(S(t_1) \leq x_1, S(t_2) \leq x_2)$$

such that

$$\{S(t_1 \leq x_1) \cap \{S(t_2 \leq x_2) \in \mathcal{F}_{t_2} \subseteq \mathcal{F}_{T} = \mathcal{F}\}$$

Obviously, and we can’t reconstruct the whole process $S$ without extra hypotesis such as independence, or presence of correlation among random variables.

**Definition A.5 (Martingale).** Given a stochastic basis $(\Omega, (\mathcal{F}_t)_{t \leq T}, P)$, an adapted process $M$ is a martingale if

$$E[M(t)|\mathcal{F}_s] = M(s)$$

The meaning of the definition above is that the best prediction of the future value of the process is its current value $M(s)$. In other words, it depends on the information at time $s$.

**A.3 Brownian motion and Ito’s Lemma**

Learning about stochastic processes is the first step to describe and understanding the pricing of options. In particular, we are interested in the so
called Brownian motion or Standard Wiener process.

**Definition A.6** (Brownian motion). Let \((\Omega, (\mathcal{F}_t)_{t\leq T}, P)\) be a filtered space. The process \(W = (W(t))_{t\leq T}\) is a Brownian Motion iff

1. \(W(t)\) is continuous and \(W(0) = 0\)
2. \(W\) is adapted to the filtration
3. for any \(s \leq t\), the increment \(W(t) - W(s)\) is independent of \(\mathcal{F}_s\) and has Normal distribution \(N(0, t - s)\)
4. The paths \(W(t, \omega^*)\) are continuous

**Remark 2.** From the definition, we can say that the increment \(W(t) - W(s)\) has the same joint distribution of any other \(W(t + \tau) - W(s + \tau), \forall s, t, \tau.\) This means that the Brownian motion is and IID process, and the distribution of the joint Gaussian can be written as

\[
N\left(\begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}, \begin{pmatrix}
t_1 & 0 & \cdots & 0 \\
0 & t_2 - t_1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & t_n - t_{n-1}
\end{pmatrix}\right)
\]

Moreover, Marginal distributions are Gaussian for any \(t\): in fact \(W(t) = W(t) - W(0) \sim N(0, t)\)

**Proposition A.3.1.** Brownian Motion \(W(t)\) is a Martingale.

**Proof.** The Brownian Motion can be written as \(W(t) = W(t) - W(s) - W(s), \) with \(s \leq t\) two arbitrary dates. We can apply the definition of Martingale and compute the expectation:

\[
E[W(t)|\mathcal{F}_s] = E[W(t) - W(s) + W(s)|\mathcal{F}_s] = W(s) + E[W(t) - W(s)] = W(s)
\]
given the properties of conditional expectation.

The mean change per unit time for a stochastic process is known as the *drift rate*, and the variance per unit time is known as the *variance rate*. 
Remark 3. The Standard Brownian Motion, $dW$, that has been developed so far, has drift rate of zero and variance rate equal to 1. The drift rate of zero means that the expected value of $W$ at any future time is simply its current value. The variance rate of 1 means that the variance of the variation in $W$ in a time interval $t$ equals $t$.

Definition A.7 (Brownian Motion with drift- Generalized Wiener Process). A Brownian Motion with drift $b$ and volatility $\sigma > 0$ is the process

$$B(t) = bt + \sigma W(t)$$

Remark 4. The above definition is a linear transform of the (Standard) Brownian Motion.

Let analyze the two components on the right-hand side separately; we can observe that:

- The $b dt$ term means that the process $dB$ has an expected drift rate of $b$ per unit time.
  Without the volatility term, the equation becomes $dB = b dt$, which implies $\frac{dx}{dt} = b$. Hence, we get
  $$B = B_0 + bt$$

- The $\sigma dW$ term is related to some variability or noise to the path followed by $B$; The amount of this variability is $\sigma$ plus the variability of the Brownian motion (equal to 1 if the length of the time intervals $dt$ is 1).

Thus, in a small time interval $dt$, the change $dB$ is given by

$$dB = b dt + \sigma \sqrt{dt} N(0, 1)$$

Thus, $dB$ has normal distribution with:

- mean of $b dt$

- standard deviation of $\sigma \sqrt{dt}$
A.3 Brownian motion and Ito’s Lemma

- variance of $\sigma^2 dt$.

Is the Brownian Motion with drift a martingale?

**Proposition A.3.2.** A Brownian Motion with drift is a martingale iff $b = 0$

**Proof.** $E(B(t)|\mathcal{F}_s) = bt + E(\sigma W(t)|\mathcal{F}_s) = bt + \sigma W(s)$ In order to be a martingale, we should have $bt = bs$ but, since it is a deterministic function of $t$, in general it is not verified. The case in which it becomes is $b = 0$.

**Definition A.8** (Ito process). An Ito process can be written as

$$X(t) = b(X,t)t + \sigma(X,t)B(t)$$

The Ito process is a Brownian Motion with drift, in which the drift rate and volatility rate are functions of the value of the underlying variable $X$ and time $t$.

### A.3.1 Geometric Brownian motion

In this section we study the stochastic process used to describe the dynamic of a stock price. It seems reasonable that we can think of a Brownian Motion with drift for stock prices, i.e. we consider a constant expected drift and variance rate. However, the key concept of a stock price and the fundamental element we have to look at, is the expected return which an investor requires. The natural request is that the expected percentage return required by investors on a stock must be independent from the stock’s price. Thus, the assumption of constant expected drift rate is replaced by the assumption that the expected return (i.e. the expected drift divided by the stock price) is constant overtime. If $S(t)$ is the stock price at time $t$, then the expected drift rate is assumed to be $\mu S$, where $\mu$ represents the expected rate of return. Let’s analyze the case in which the volatility rate is equal to zero; then this model implies that

$$dS = \mu S dt$$
from which

\[ \frac{dS}{S} = \mu dt \]

Integrating in the interval \([0,T]\), we have

\[ S(T) = S(0)e^{\mu t} \]

The equations show that, if the variance rate is zero, the stock price grows at a continuous rate of \(\mu\) per unit time.

Obviously, a stock price is also characterized by volatility; in particular we assume that it is proportional to the stock price. Now we can formulate the following

**Definition A.9** (Geometric Brownian Motion). Consider a Brownian Motion with drift \(b\) and volatility \(\sigma\).

The exponential transform process \(S(t) = e^{B(t)} = e^{bt + \sigma W(t)}\)

is called Geometric Brownian Motion.

**Remark 5.** We notice that, for any \(t\), the GBM follows the exponential of a Gaussian variable, so the marginal distributions are lognormals.

### A.3.2 Ito’s Lemma

The price of a stock option and a derivative is a function of the underlying stock’s price and time.

An important result which characterizes the dynamics of any smooth function of the Brownian Motion \(W\) is due to Ito’s Lemma.

**Proposition A.3.3** (Ito’s Lemma). If \(X(t)\) is a stochastic process satisfying \(dX(t) = \mu dt + \sigma dW\) and \(f(t, X(t))\) a deterministic twice continuously differentiable function \((\in C^{1,2}(t, x))\) then \(Y_t = f(t, X(t))\) is also a stochastic process and is given by

\[ dY_t = \sigma \frac{\partial f}{\partial X} dX + \left( \mu \frac{\partial f}{\partial X} + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} \right) dt \]
**Proof.** Here we give just an intuition of the real proof. Consider the deterministic function of two variables \( f(t, x) \). \( f \) varies in response to changes in \((t, x)\) according to the differential

\[
df(t, x) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx
\]

which represents the infinitesimal change \( df(t, x) := f(t + dt, x + dx) - f(t, x) \).

A second order approximation is given by the Taylor expansion

\[
df(t, x) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{1}{2} \left( \frac{\partial^2 f}{\partial x^2} dx^2 + 2 \frac{\partial^2 f}{\partial t \partial x} dt dx + \frac{\partial^2 f}{\partial t^2} dt^2 \right).
\]

Now interpret \( t \) as a time parameter and \( X \) as the stochastic process. The variation can be described by

\[
df(t, X) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X} dX + \frac{1}{2} \left( \frac{\partial^2 f}{\partial X^2} dX^2 + 2 \frac{\partial^2 f}{\partial t \partial X} dt dX + \frac{\partial^2 f}{\partial t^2} dt^2 \right).
\]

But, it can be proved that the term \( dX^2 \) is replaced by \( dt \):

\[
dX^2(t) \sim dt
\]

The intuition of the above approximation is that \( dX(t) = X(t + dt) - X(t) \sim N(0, dt) \) and then we approximate the square increment by its mean:

\[
E(dX^2) = E((dX - E(dX))^2) = Var(dX) = dt
\]

### A.3.3 The lognormal distribution

Now, we can derive the stochastic process followed by the stock price, i.e. the Geometric Brownian Motion. The process has differential \( dS = S(\mu dt + \sigma dW) \), then we can derive through Ito’s formula the differential of the function \( f(S) = \ln S \).

Since

\[
\frac{\partial f}{\partial S} = \frac{1}{S}, \quad \frac{\partial f}{\partial t} = 0, \quad \frac{\partial^2 f}{\partial S^2} = -\frac{1}{S^2}
\]

Then, applying the Ito’s formula, we have

\[
df(S) = \frac{1}{S} dS + \frac{1}{2} \left( -\frac{1}{S^2} \right) dt
\]

by substituting \( dS \) and \( dS^2 \) we obtain

\[
\frac{1}{S}(S(\mu dt + \sigma dW) + \frac{1}{2} \left( -\frac{1}{S^2} \right) (S^2(\mu dt + \sigma dW)^2)]
\]
finally we have
\[ df(S) = (\mu - \frac{1}{2} \sigma^2)dt + \sigma^2 dW \]

This equation indicates that the function \( f(S(t)) = \ln S(t) \) follows a Brownian Motion with constant drift equal to \( \mu - \frac{1}{2} \sigma^2 \) and a constant variance rate \( \sigma^2 \).
This means that the change in \( df \) in \([0,T]\) has distribution
\[
\ln S(0) - \ln S(T) \sim N\left(\mu - \frac{1}{2} \sigma^2 T, \sigma^2 T\right)
\]
\[
\ln S(T) \sim N\left(\ln S(0) + \mu - \frac{1}{2} \sigma^2 T, \sigma^2 T\right)
\]
Thus, the variable \( S(t) \) has a lognormal distribution (because \( \ln S(t) \) is normally distributed).
The solution of the stochastic differential then is
\[ S(t) = S0e^{\mu t - \frac{1}{2} \sigma^2 t + \sigma W(t)} \]
Appendix B

The Black-Scholes-Merton scenario

In this section we present one of the most famous models for pricing financial derivatives. In particular we need the Black-Scholes formula for pricing an *European Option*. Before that, we will construct a particular continuous time market model.

B.1 The market model

**Definition B.1.** The Black-Scholes model consists of two assets with dynamics given by

\[
\begin{align*}
\frac{dB(t)}{B(t)} &= rB(t)dt \\
\frac{dS(t)}{S(t)} &= S(t)(\mu dt + \sigma dW(t))
\end{align*}
\]

where \( r, \mu \) and \( \sigma \) are deterministic constants.

This means that in the market we have

- The *riskless security*, i.e. the Geometric Brownian Motion without the stochastic term \( \sigma dW \)

\[
B(t) = B(0)e^{rt}
\]

\[
B(0) = 1
\]

where \( r \) is the risk free interest rate.

This object is locally deterministic in the sense that, at time \( t \), we have
complete knowledge of the return by simply observing the prevailing short rate $r$ (assumed constant).

- A *risky asset*. The price process $(S(t))_{t \in \mathbb{R}_+}$ with stochastic differential given by

$$dS(t) = S(t)(\mu dt + \sigma dW)$$

$$S(0) = S_0$$

has the following solution

$$S(t) = S(0)e^{\mu t - \frac{1}{2}\sigma^2 t + \sigma W(t)}$$

In particular, mean and variance of the stock logreturn grow linearly with time.It is shown by the marginals $S(t)$, which satisfy

$$\ln \frac{S(t)}{S_0} \sim N((\mu - \frac{\sigma^2}{2})t, \sigma^2 t)$$

Examining the economic meaning of the drift and volatility, turns out that $\mu$ is exactly the exponential growth of the average stock price:

$$E(S(t)) = e^{\mu t}$$

The volatility is simply the standard deviation of the annual logreturn. These fundamental parameters play also a significant role when we look at the instantaneous behavior of the stock. In fact, by considering the return of $S(t)$ on the interval $[t, t + dt]$, i.e. the ratio $\frac{dS}{S}$, we can write

$$E(\frac{dS}{S} | F_t) = \mu dt$$

In this case, the drift $\mu$ is the *conditional rate of return* of the stock. The same reasoning can be made for the volatility $\sigma^2$, which represents its *conditional rate of variance*.

### B.2 Portfolio strategies

In this context, we can define a *Portfolio strategy* characterized by stock and bond.
B.2 Portfolio strategies

**Definition B.2 (Portfolio strategy).** A portfolio strategy is a couple of adapted processes \((H(t), K(t))\) where:

- \(H(t)\) represents the number of shares of the Stock \(S(t)\)
- \(K(t)\) represents the number of bond shares

Hence, the portfolio value \(V(t)\) can be written as a combination of stock and bond:

\[
V(t) = H(t)S(t) + K(t)B(t)
\]

The most delicate part of such a strategy is to analyze what happens at trading dates. Suppose that trading occurs along specified dates in time \(\{t_0 = 0 < t_1 < t_2 < \ldots < t_n = T\}\), and consider a piecewise constant strategy. In this case, for every interval of time \([t_i, t_{i+1}]\), for \(i = 0..n - 1\), the amount of stock and bond is constant and corresponds to \(H(t_i), K(t_i)\).

At the initial time \(t_0\), we must start with initial wealth \(w_0\), which corresponds to \(V(0) = K(t_0) + H(t_0)S_0\). For times before \(t_1\), the portfolio is

\[
V(t) = K(t_0)B(t) + H(t_0)S(t), \forall t \in [t_0, t_1).
\]

In \(t_1\) the wealth we have corresponds to the value

\[
w_1 = \lim_{t \to t_1} V(t) = V(t_1) = K(t_0)B(t_1) + H(t_0)S(t_1)
\]

since the processes \(B\) and \(S\) are continuous.

To implement the strategy in the second step, we should possess the wealth

\[
w_2 = V(t_1) = K(t_1)B(t_1) + H(t_1)S(t_1)
\]

It can be obtained by external infusion of money if \(w_1 < w_2\), or by consumption if the opposite inequality holds. When \(w_1 = w_2\), a re-allocation of money occurs at date \(t_1\) between stocks and bonds. If the equality holds at any \(t_i\), for \(i = 1..n - 1\), we can write the so called *self-financing equations*

\[
K(t_{i-1})B(t_i) + H(t_{i-1})S(t_i) = K(t_i)B(t_i) + H(t_i)S(t_i)
\]

Remembering that \(K(t_i) = K(t)\) and \(H(t_i) = H(t)\) in \([t_i, t_{i+1})\), and the portfolio value is continuous, we have a global SDE satisfied by \(V\) on \([0, T]\):

\[
dV(t) = K(t)dB(t) + H(t)dS(t)
\]

This means that the variation on the portfolio value are proportional to the change in value of stock and bond; and we can define the following
Definition B.3 (Self-financing strategy). A general portfolio strategy \((H,K)\) (not necessarily piecewise constant) is called self-financing if the equation B.3 holds on \([0,T]\). Any self-financing portfolio is a diffusion.

B.3 Contingent Claims and Arbitrage

In this section we present the fundamental instruments through which the Merton model establishes its main features.

B.3.1 European options

Definition B.4 (European Call Option). An European Call Option with strike price \((or exercise price) K\) and time of maturity \((or exercise date) T\) on the underlying asset \(S\) is a contract defined by the following clauses:

1. The holder of the option has, at time \(T\), the right to buy one share of the underlying stock at the price \(K\) from the underwriter of the option.
2. The holder of the option is in no way obliged to buy the underlying stock.
3. The right to buy the underlying stock at the price \(K\) can only be exercised at the precise time \(T\).

Remark 6. We can observe that the exercise price \(K\) and the time of maturity \(T\) are determined at the time when the option is written, which for us typically will be at \(t = 0\).

Definition B.5 (European Put Option). An European Put Option with strike price \((or exercise price) K\) and time of maturity \((or exercise date) T\) on the underlying asset \(S\) is a contract defined by the following clauses:

1. The holder of the option has, at time \(T\), the right to sell one share of the underlying stock at the price \(K\) from the underwriter of the option.
2. The holder of the option is in no way obliged to sell the underlying stock.
3. The right to sell the underlying stock at the price \(K\) can only be exercised at the precise time \(T\).

The common factor of all these contracts is that they all are completely defined in terms of the underlying asset \(S\), which makes it natural to call them derivative instruments or contingent claims.
**Definition B.6** (Contingent Claim). Consider a financial market with vector price process $S$. A contingent claim with date of maturity (exercise date) $T$, is any stochastic variable

$$X \in \mathcal{F}_T$$

A contingent claim $X$ is called a **simple claim** if it is of the form

$$X = f(S(T))T$$

and the function $f$ is called a contract function.

In other words, a contingent claim is a contract which stipulates that the holder of the contract will obtain $X$ (which can be positive or negative) at the time of maturity $T$.

The condition that $X \in \mathcal{F}_T$ tells us that, at time $T$, it will actually be possible to determine the amount of money to be paid out.

The European call is a simple contingent claim, for which the contract function is given by

$$f(S(T), T) = \max(S(T) - K, 0)$$

Then, we can make the following considerations

- If $S(T) > K$ we have a certain profit by exercising the (call) option in order to buy one share of the underlying stock; this will cost us $K$ and we immediately exercise the option. The net profit of $S(T) - K$.

- If $S(T) < K$ the option has no value; we decide to not exercise it. Thus the profit is equal to zero.

**B.3.2 The Put-Call parity**

So far we have been working with puts and calls separately. To see how their prices must be consistent with each other and to explore common option strategies, let us combine puts and calls.

First, we consider an option strategy referred to **fiduciary call**. It consists of a European call and a risk-free bond $B$, that matures on the option expiration day and has a face value equal to the exercise price of the call. If the price of the underlying is below $B$ at expiration, the call expires worthless and the bond is worth $B$. If the price of the underlying is above $B$ at expiration, the call expires and is worth $S_T - B$. This type of combination is called a fiduciary call because it allows protection against downside losses and is
thus faithful to the notion of preserving capital.

Now we construct a strategy known as protective put, which consists of a European put and the underlying asset. If the price of the underlying is below $B$ at expiration, the put expires and is worth $B - S_T$. If the price of the underlying is above $B$ at expiration, the put expires with no value and the underlying is worth $S_T$. So at expiration, the protective put is worth $B$ or $S_T$.

Thus, the fiduciary call and the protective put end up with the same value. They are, therefore, identical combinations. To avoid arbitrage, their values today must be the same. The value of the fiduciary call is the cost of the call $C_0$, and the cost of the bond, $B/(1 + r)^T$. The value of the protective put is the cost of the put $P_0$, and the cost of the underlying, $S_0$. Thus,

$$C_0 + B/(1 + r)^T = P_0 + S_0$$

This equation is called Put-Call Parity, and is one of the most important results in options.

**Definition B.7** (Hedged portfolio). Given the market model $(S(t), B(t))$, let consider a portfolio $h(t) = (H(t), K(t))$ with value $V_t = H(t)S(t) + K(t)B(t)$. We can say that it is hedged against the contingent claim $C$ (replicates $C$) if, for any $t$ the following condition holds

$$dV_t = dC_t$$

**Definition B.8** (Arbitrage strategy). Given the market model $(S(t), B(t))$, a portfolio strategy $h(t) = (H(t), K(t))$ constitutes and arbitrage opportunity if the following conditions are satisfied:

1. $V(0) = 0$
2. $V(T) \geq 0$, and $P(V(T) > 0) \geq 0$

This means that we start with zero money, and at maturity $T$ we make a profit with positive probability, without spending anything. Indeed, the main assumption we will make is that the market is efficient in the sense that it is free of arbitrage possibilities.

### B.4 Solution of the Black-Scholes PDE for a Call Option

In deriving their formula for the value of an option on the underlying asset, Black and Scholes, in the famous article of the 1973 "The Pricing
of Options and Corporate Liabilities” assume that the market satisfies the following conditions:

1. the short-term interest rate is known and is constant through time.
2. The stock price follows a random walk in continuous time (i.e. a Geometric Brownian Motion) with a variance rate proportional to the square of the stock price; thus the distribution of the stock prices at the end of every finite interval is lognormal.
3. The stock pays no dividends or other distributions.
4. The option is European: it can only be exercised at maturity.
5. There are no transaction costs in buying or selling the stock or the option.
6. There is the possibility to borrow any fraction of the price of the asset to buy it or to hold it, at the short-term interest rate.
7. It is possible the short selling.

Let’s denote with \( C(t) = f(S(t), t) \) the value of the Call Option at time \( t \) on the underlying asset \( S \).

Under these assumptions, it is possible to prove that the price of \( C(t) \) depends only on ”observable” variables.

- the stock price \( S(t) \) of the underlying asset
- the time \( t \)
- the date of maturity \( T \)
- the risk-free rate \( r \)
- the variance \( \sigma \) and other constants present in the contract.

The first step of the PDE method is to construct a self-financing portfolio \( V(t) \) which replicates the (unique) price of a Call for every \( t \in [0, T] \). Let’s apply Ito’s formula to \( C(t) \):

\[
dC = C_t dt + G_S dS + \frac{1}{2}G_{SS}(dS)^2
\]

\[
dC = (G_t + \frac{1}{2}G_{SS}\sigma^2 S^2)dt + G_S dS
\]  \hspace{1cm} (B.4)
If the derivative admits a replicating portfolio, necessarily its dynamics must coincide with B.4:

$$dC(S(t), t) = dV(t) = K(t)dB(t) + H(t)dS(t) = K(t)re^{rt}dt + H(t)dS(t)$$

(B.5)

By equating B.4 and B.5 we have the following relation:

$$(C_t + \frac{1}{2} C_{SS}\sigma^2 S^2)dt + C_SdS = K(t)re^{rt}dt + H(t)dS(t)$$

(B.6)

In order to construct the replicating portfolio, we write two fundamental equations referring to the shares of stock and bond which are needed:

1.

$$H(t) = C_S(t, S(t))$$

(B.7)

The number of shares of the risky stock used for the replication is called the Delta, and it represents the sensitivity of the option to stock changes at any time. Notice that it is simply the derivative of the option price $C$ with respect to the stock, valuated at time $t$.

2.

$$K(t)re^{rt} = C_t + \frac{1}{2} C_{SS}\sigma^2 S^2$$

(B.8)

As $Ke^{rt} = K(t)B(t) = V(t) - H(t)S(t) = C(t, S(t)) - S(t)C_S(t, S(t))$, we can modify the left-hand side of B.8:

$$rC(t, S(t)) - S(t)C_S(t, S(t)) = C_t + \frac{1}{2} C_{SS}\sigma^2 S^2, \forall t \in [0, T]$$

(B.9)

Since $S(t)$ is a lognormal when $t$ is fixed, we can replace it with a positive real parameter $s$ and write the final parabolic valuation PDE, which is the same for all the derivatives:

$$C_t(t, s) + rsC_S(t, s) + \frac{1}{2} \sigma^2 s^2 C_{SS}(t, s) = rC(t, s)$$

(B.10)

The specific valuation formula for the Call Option comes into play when we add the terminal condition: $C(T, S(T)) = max(S(T) - K, 0)$. We face the following Cauchy problem:

$$\begin{cases}
  C_t(t, s) + rsC_S(t, s) + \frac{1}{2} \sigma^2 s^2 C_{SS}(t, s) = rC(t, s) \\
  C(T, S(T)) = max(S(T) - K, 0)
\end{cases}$$

(B.11)
Its unique solution coincides with the famous Black and Scholes pricing formula for a Call option:

\[ C(t, S(t)) = S(t)\phi(d_1(t)) - Ke^{-r\tau} \phi(d_2(t)) \]  

(B.12)

where

\[
\tau = T - t \\
d_1(t) = \frac{\ln(S(t)/K) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}} \\
d_2(t) = d_1(t) - \sigma\sqrt{\tau}
\]

and \( \phi \) is a standard Gaussian CDF.

The above system can be solved through analytical methods, numerical methods, or by using the Feynman-Kac probabilistic representation. We are interested only at the solution, its proof goes beyond the scope of this work.

The price of the corresponding Put option can be reached by direct calculation, or by simply using the Put-Call parity.
Bibliography


[29] John C. Hull, Options Futures, and Other Derivatives,


