POLITECNICO DI TORINO

Master degree

in Physics of Complex Systems

Decision-Making Under Risk: the Role of Correlation in Fostering the Low-Carbon Transition

SUPERVISORS
Prof. Stefano Battiston
UZH Zürich - DBF, Department of Banking and Finance

Prof. Alfredo Braunstein
Politecnico di Torino - DISAT

CANDIDATE
Carola Botto

CO-SUPERVISORS
Prof. Marcello Restelli
Politecnico di Milano - DEIB

PhD student Alan Roncoroni
UZH Zürich - DBF, Department of Banking and Finance

Academic year 2019/2020

Cara nonna Cia,
questo è per te che hai sempre vissuto come se non ci fosse un traguardo, spero mi sia concesso di imparare a vivere una vita così senza fine.

Abstract

Building a successful portfolio is a challenging task requiring thoughtful choices of financial assets and contracts that best meet investors' risk profiles. Due to their characteristics, securities like zero-coupon bonds represent an attractive asset class for a wide range of investors. Nonetheless, securities are not riskless since they are characterized by a certain probability of default. Moreover, building larger profitable portfolios, relationships between firms, and correlations between different contracts can not be neglected. In our work, we leveraged on these three aspects (default probability, correlation, and risk attitudes) to help investors in making wiser decisions within the complexities of the real economy and sustainable finance perspectives. After the Paris agreement, signed in December 2015, the community of financial supervisors and central banks agrees that climate changes must be considered while assessing the soundness of both corporate and governmental bonds investments. Financial risks could result from the sign of a new agreement or the adoption of a different climate policy aiming at encouraging sustainable business strategies. We discovered that correlation can play a crucial role in fostering an endogenous transition to a low-carbon economy, changing the status quo of portfolio management. The impact of a transition on business and portfolio strategy is then assessed by a scenario-based approach, exploiting different risk metrics and portfolio selection models (e.g. Markowitz portfolio theory, expected utility maximization, stochastic dominance ordering). In the last part of this work, we highlighted advantages and limitations of the mixture model approach in evaluating the repercussions of shocks stemming from the disordered introduction of new climate policy, both on bonds' expected values and their default probability distributions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Decision Theory refers to the study of a rational agent that chooses among a set of options. In decisions under risk, the outcome of a choice depends on which state of the world turns out to be the actual one. The choice is said to be risky since the decision-maker, when choosing, knows only the probability of the possible outcomes. As an example of a decision-maker, an investor building a portfolio chooses the financial contracts that best meet her risk profile. Generally, a certain level of correlation between the contracts is present, positive or negative depending on the existing relationship between the firms under analysis. After the Paris agreement, signed in December 2015, the community of financial supervisors and central banks agrees that climate changes must be considered while assessing the soundness of both corporate and governmental bonds investments. Although the low-carbon transition is a hot topic on the international agenda, to date, the status quo of investment portfolios is still largely focused on conventional assets, mainly related to carbon-intense activities.In the financial context, *brown* is referred to contracts mainly related to carbon-intensive businesses such as high industrial and mining sectors. Conversely, *green* is addressed to contracts that foster low-carbon activities. A low-carbon transition concerns a shift towards an economy based on a low-carbon production and consumption system, reducing CO_2 emissions into the atmosphere; in addition, it requires massive investments in low-carbon technologies. If on one hand, the transition could be promoted by the adoption of a climate policy, for instance, issued by the EU, on the other, such a policy could be a potential source of financial risk. Adding capital into green activities, boost people to develop new climate-related projects, therefore improving the technology and reducing production costs. Thus, other investors, more skeptical at the beginning, are persuaded to invest too, de-facto promoting the transition. To this end, it is crucial to assess the implications of the transition on the current economy as well as to comprehend the correlation structure between green and brown contracts.

The purpose of this thesis is to investigate through a probabilistic approach the role of the correlation in portfolio management within the context of the low-carbon transition.

In **Chapter 2**, we start making a review of the mathematical tools necessary to appreciate the result of our study. Then, in **Chapter 3**, we present the state-of-the-art in Decision Theory; the most widespread economic theories and portfolio selection models letting the investor achieve an optimal portfolio are therefore shown and exploited. In particular, we leverage on Markowitz Portfolio Theory and Expected Utility Maximization theory. We consider a portfolio of zero-coupon, defaultable corporate bonds with maturity T, modeled as binary variables with default probability q_i. To study the optimal portfolio allocation we start by considering the situation of two investments and one investor and we analyze how the portfolio is affected by three parameters: correlation, default probability, and investor's risk attitude.

During the study, the Copula function is employed to calculate the joint probability distribution between bonds. However, applying the Gaussian Copula, the correlation coefficient between the Gaussian test variables changes with respect to the Bernoulli ones of our interest. An in-depth investigation of this implicit link is delineated in **Chapter 4**. In addition, in the same chapter, we outline also the correlation coefficient upper and lower bounds. Starting with the next chapter, the number of bonds addressed becomes N. In **Chapter 5** we contribute to analytically respond to the investors' concern about choosing the most efficient portfolio among two, simply by knowing the probability of win/fail of individual bonds. The last step is to include the correlation between bonds. In **Chapter 6** we proceed by exploiting the Mixture Model: the correlation comes again into play, but it is driven by an exogenous macroeconomic variable. We thus outline some hypotheses on how the correlation changes according to the exogenous parameter. Besides, we discuss also the asymptotic limit of a portfolio with $n \to \infty$. Then, we assess how a shock (focusing in particular on a Climate policy shock) can impact on the portfolio risk. Finally, we point out the limitations of this model concerning our objectives and develop some further insights.

Chapter 2

Theoretical Background

In this chapter, we consider a portfolio of defaultable corporate bonds with maturity T. For sake of simplicity, we focus on zero-coupon bond, modelled as binary variables with default probability q_i. We consider both the case of independent bonds as well as the case of correlated bonds. Then, we consider a situation with multiple future scenarios, where each scenario is characterized by different default probabilities q_i of individual bonds. Finally, we describe the evolution of portfolio value and portfolio payoff. Relying on known mathematical tools such as convolution, generating functions, and copulas we analyse under which conditions such tools are suitable to study our portfolio of bonds.

2.1 Convolution and Generating Functions

Convolution is useful since it is a tool, through which the probability density function (pdf) of a sum of random variables can be computed, starting from the single pdf. For instance, let us a portfolio of k corporate bonds, where the payoff of each bond is distributed according to a specific pdf. The probability density function of the entire portfolio's payoffs changes under the following different situations:

A. Every payoff index is independent and identically distributed;

B. Every payoff is independent of the other, but they are distributed according to different pdf;

C. The payoffs are correlated among themselves.

For cases A and B the convolution is crucial to find the probability density function of the portfolio. We start analyzing case B, since knowing this, the other one (case A) could be easily derived. Every bond payoff is considered as a mutually independent random variable X_i. By consequence, the value of the total portfolio Y will be the linear combination of these k random variables:

$$Y = \sum_{i=1}^{k} b_i X_i \tag{2.1}$$

each one weighted according to a specific weight b_i, corresponding to the percentage of the total wealth invested in the i-*th* index. If the variables are normally distributed with a mean μ_i and a variance σ_i^2, the probability distribution of Y remains a Gaussian with $E[Y] = \sum_{i=1}^{k} b_i \mu_i$ and $var[Y] = \sum_{i=1}^{k} b_i^2 \sigma_i^2$. Precisely thanks to their mathematical tractability and to the possibility of obtaining analytical results, usually in literature, financial variables are assumed to follow a Gaussian distribution. With Gaussian variables, for instance, the convolution product is obtained for free, since no calculation are needed. In all the other non-Gaussian cases, we need to compute explicitly the convolution product to get the distribution of Y. Sometimes, solving the convolution integral is not easy, therefore, to get the result, one can exploit the generating or the characteristic functions, depending on if the single index payoff has a discrete or continuous distribution, respectively. The generating function $G_X(z)$ of a discrete random variable X is defined as the expected value of z^X:

$$G_X(z) = E[z^X] = \sum_{n=0}^{\infty} p_X(n) z^n \tag{2.2}$$

Therefore, for the random variable Y defined in Equ. 2.1 with $b_i = 1 \quad \forall i$, the generating function becomes

$$G_Y(z) = G_{X_1+X_2+\ldots+X_k}(z) = E[z^Y] = E[z^{Y_i=\sum_{i=1}^{k} X_i}] = \prod_{i=1}^{k} E[z^{X_i}] = \prod_{i=1}^{k} G_{X_i}(z) \tag{2.3}$$

where X_i are independent. If the X_i are also identically distributed the expression simplifies even more: $G_{X_i}(z) = G_X(z) \quad \forall i$, and hence $G_Y(z) = [G_X(z)]^n$. On the contrary, if X is a continuous random variable, one can choose to adopt either the characteristic function

$$\phi_X(t) = E[e^{itX}] = \int_{-\infty}^{+\infty} dx p_X(x) e^{-itx} \tag{2.4}$$

or the moment generation function

$$M_X(t) = E[e^{-tX}] = \int_{-\infty}^{+\infty} dx p_X(x) e^{-tx} \tag{2.5}$$

Usually, characteristic function is preferred because, corresponding to the Fourier transform, it is always calculable, while the moment generation one, which can be seen as the Laplace transform, might lead to a non-convergent integral. The idea behind the use of these transforms is to make the calculations easier. If independent, convolving variables in Fourier or Laplace space is straightforward since the convolution product corresponds to a standard multiplication. However, after having obtained the generating or characteristic function of a sum of random variables, one has to perform the reverse transformation to get the joint probability distribution, which corresponds to the n-*th* term of the series. Sometimes it is easier cause convolving variable of the same type return still a variable of that type. However, when the variables are not i.i.d or does not follow a standard distribution we can get some help from complex analysis. Starting from the definition of generating function of Equ. 2.2 one can observe that this is no more than a Taylor's expansion. Multiplying both sides by $\frac{1}{z^{n+1}}$ we obtain

$$\frac{1}{z^{n+1}}G_X(z) = \sum_{n=0}^{\infty} p_X(n)\frac{z^n}{z^{n+1}} = \sum_{n=0}^{\infty} p_X(n)\frac{1}{z}. \tag{2.6}$$

The right-hand side is now a Laurent's series. Hence, we know that the coefficient of the $1/z$ term i.e. a_{-1} corresponds to the n-*th* term of the original Taylor's series:

$$a_{-1} = p_X(n) = \frac{1}{2\pi i}\oint \frac{G_X(z)}{z^{n+1}}dz \tag{2.7}$$

then, performing the complex integration we have back the n-th term, and hence the probability distribution of the sum.

2.1.1 Convolution of Bernoulli-like bonds

In the following, we try to figure out some results for the case of a portfolio of corporate bonds with maturity T. Following the notation of Battiston and Monasterolo (2019) the bond value is defined as a binary variable, which can default with probability q_i

$$\text{Bond Value} = v_i(T) = \begin{cases} R_i = 1 - LGD_i, & \text{w.p.} \quad q_i \\ 1, & \text{w.p.} \quad 1-q_i \end{cases}$$

$$X_i(T) = \begin{cases} R_i = 1 - LGD_i, & \text{w.p.} \quad q_i \\ x_i, & \text{w.p.} \quad p_i = 1-q_i \end{cases}$$

where R_i is the Recovery Rate, i.e. the percentage of notional recovered upon default, which corresponds to $1 - LGD$, that is Loss Given Default, i.e. the percentage of losses. Consequently, one define the bond loss as 1-the bond value

$$\text{Bond Loss} = l_i(T) = \begin{cases} LGD_i, & \text{w.p.} \quad q_i \\ 0, & \text{w.p.} \quad 1 - q_i \end{cases}$$

Then, we sum every gain and every loss of the k bonds in the portfolio by weighting them with w_i, i.e the amount (as a percentage of the entire wealth) of i-th bond purchased, obtaining:

$$\text{Portfolio Value} = Z(T) = \sum_{i=1}^{k} v_i w_i \tag{2.8}$$

$$\text{Portfolio Loss} = Y(T) = \sum_{i=1}^{k} l_i w_i \tag{2.9}$$

To tackle the issue of writing the full probability distribution of the portfolio (ptf) we start by computing the generating function of the portfolio loss. We choose the generating function since we are dealing with discrete random variable, while we take the ptf loss instead of the ptf value since it would return a formula similar to the generating function of a binomial

$$\begin{aligned} G_Y(\eta) = E[\eta^Y] &= \prod_i \left[q_i \eta^{LGD_i w_i} + 1 - q_i \right] \\ &= \prod_i \frac{1}{w_i} \left[q_i \eta^{u_i} + 1 - q_i \right] \end{aligned} \tag{2.10}$$

where defining $u_i = LGD_i \cdot w_i$ we get

$$p(u_i) du_i = p(l_i) dl_i \implies p(u_i) = \frac{p(l_i)}{w_i} = \frac{q_i}{w_i}$$

As said before, the trickiest part of this approach is getting back the probability density function, which corresponds with the n-th term of the series. Although the variables we are dealing with in the present dissertation are not so different from Bernoulli's rv, they can't be considered as such. Hence, to obtain some analytical results it is suggested to approach the problem via complex analysis solution and then plot the result. Otherwise, and this is what we did to check it, the sum can be computed numerically and then, plotting the histogram, the shape of the distribution is shown. Considering m (number of simulations) large, one can find that the normalized sum of n (number of bonds) independent variables tends to a Gaussian distribution. Getting this result, even if the single variables are not Gaussian, makes sense because it is a direct consequence of the Central Limit Theorem.

2.2 Correlated Random Variables

A quite different story happens when variables are correlated with each other (case C). Indeed, as said above, for two independent random variables, the probability density function of their sum is the convolution of the density functions of single addends. While, if we consider the case of two correlated random variables, the situation is a bit trickier. Let consider, for instance, two correlated random variables X and Y, where their sum is $Z = X + Y$. If one intends to compute the probability density function of $z = g(x, y)$, the ordinary strategy is to integrate the joint density function $p_{XY}(x, y)$ throughout the domain where $g(x, y) < z$. In this way, the cumulative distribution $P_Z(z)$ is obtained. Then, differentiating with respect to Z, the density function $p_Z(z)$ is obtained.

$$P_Z(z) = \int_D \int p_{XY}(x, y)dxdy = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} p_{XY}(x, y)dy \right] dx \tag{2.11}$$

$$p_Z(z) = \frac{dPZ(z)}{dZ} = \int_{-\infty}^{\infty} p_{XY}(x, z - x)dx \tag{2.12}$$

When X and Y are independent, the joint density function separates into a product of the two marginal density functions $p_X(x)$ and $p_Y(y)$, and the procedure we are about to describe, using the inverse relation $y = z - x$, leads directly to the convolution. But now we are considering X and Y to be correlated, so the joint cannot factorize and this no longer leads to the convolution. In general, the mean is unaffected by the correlation, whereas the variance is made larger or smaller according to whether the correlation is positive or negative, respectively

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y \tag{2.13}$$

Again, in the special case of Gaussian distributed random variables, given the mean and the variance we can easily find the distribution of the sum, even in presence of correlation. For the other cases, in order to find the joint probability, a new tool becomes effective: the copula function. Thanks to the copula, we can get the joint probability and therefore, deriving it with respect to Z, we will have the pdf of the entire portfolio.

2.3 Copula function

"*The "era of i.i.d." is over: and when dependence is taken seriously, copulas naturally come into play. It remains for the statistical community at large to recognize this fact. And when every statistics text contains a section or a chapter on copulas, the subject will have come of age.*" - Schweizer (2007)

In the following, a bit of theory about copula has been summarized, whereas a exhaustive discussion of the topic can be found looking at Sklar (1996), Schweizer (1991) and Nelsen (2006).
As one can grasp, the word “Copula” reminds the grammatical expression linking subject and predicate. Sklar (1959) himself argued that it would be a suitable name also to define a function connecting the multidimensional distribution of $d \geq 2$ variables to their one-dimensional marginals, and from there this name was officially adopted in Probability Theory and Statistics.

Definition 1 (d-dimensional Copula). For $d \geq 2$, a d-dimensional Copula (in short, d-copula) is a d-variate distribution function from $\mathbb{I} = [0,1]^d$ to $[0,1]$ whose univariate marginals are uniformly distributed on $\mathbb{I}$. It satisfies the following properties:

1. $C(1, 1, a_m, 1, \cdots, 1) = a_m, \quad \forall a_m \text{ in } [0,1], \text{ with } m \leq d$
2. $C(a_1, a_2, \cdots, a_d) = 0, \qquad \text{if } a_m = 0 \text{ for any } m \leq d$
3. C is d-increasing

Equivalently, a d-copula may be seen as a d-dimensional cumulative probability distribution function whose domain is $\mathbb{I}^d$ and whose one-dimensional margins are uniform on $\mathbb{I}$. In Sklar (1959), the author present what can be seen as the most important theorem about copula functions.

Theorem 1 (Sklar's Theorem)**.** *If H is a d-dimensional probability distribution function with one-dimensional margins $F_1, \cdots, F_d$, then there exists a d-dimensional copula C such that, $\forall x_i$ in* $\mathbf{R}$

$$H(x_1, \cdots, x_n) = C(F_1(x_1), \cdots, F_d(X_d)) \tag{2.14}$$

If H is continuous, then C is unique; while for what concerns discrete random variables, C is uniquely determined on the Cartesian product $(RanF_1)X(RanF_2)X(RanF_d)$. The theorem can be applied also in reverse: if C is an d-dimensional copula and $F_1, \cdots, F_d$ are one-dimensional distribution functions, then the function H defined in by a d-dimensional probability distribution function with univariate marginal $F_1, \cdots, F_d$. For example, given two (d=2) random variables X, Y with joint pdf $p_{XY}(x, y)$ and distribution $H(x, y) = \int^x dx' \int^y dy' p_{XY}(x', y')$, the Copula is thus defined by the identity

$$H(x, y) = C(F_X(x), F_Y(y)) \tag{2.15}$$

with marginals $F_X(x)$ and $F_Y(y)$ respectively. The idea is that the mapping

$$(X, Y) \rightarrow (U = F_X(x), V = F_Y(y)) \tag{2.16}$$

transforms the original random variable (rv) into another rv having uniform marginal. Consequently, letting F^{-1} denotes the inverse function of F, defined by

$$F^{-1}(u) = \sup\{x : F(x) < u\} \tag{2.17}$$

it is straightforward that:

$$C(u,v) = H(F_X^{-1}(u), F_Y^{-1}(v)) \tag{2.18}$$

Therefore, finding the multivariate joint distribution is reduced to the study of the copula $C(u,v)$, which contains information on statistical dependence of U and V. The main advantage of copula function is that it enables to specify the marginal univariate distributions separately from the specification of the dependence structure among the variables, which is provided by the type of copula chosen. It is important to underline the distinction between the concept of dependence embedded in copula function and the concept of linear correlation, i.e. the standard Pearson's correlation coefficient ρ. Although the linear correlation represents the standard tool used in risk management to measure the comovement of markets, it may turn out to be a flawed instrument to capture any non-linear relationships among the variables. As regards the dependence structure proper of the copula, it is scale-invariant, which imply that any measure of dependence right based on copula is non parametric and does not depend on the type of distribution of the original variables. Some examples of non parametric measures are: Spearman's rho, Kenndall's tau, and Blomqvist's beta. For a complete discussion have a look at Schmid and Schmidt (2007). These parametric measures are called rank correlations and they measure to what extent large and small values of one rv are associate with large and small values of another one. In the following two of those are reported to evidence that they do not depend on the original marginal probability distributions, but they are directly linked to the copula.

$$\rho_S = 12\int_0^1\int_0^1 C(u,v)dudv - 3 \tag{2.19}$$

$$\tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1 \tag{2.20}$$

Since the copula is bounded both from above and below

$$Max(F_X(x) + F_Y(y) - 1, 0) \leq H(x,y) \leq Min(F_X(x), F_Y(y)) \tag{2.21}$$

it may be proved that substituting in the equations 2.19 and 2.20 the maximum and minimum copulas gives values of -1 and 1 respectively. Differently from the linear correlation measure, where this is not verified, here we obtain 1 for both Spearman's and Kendall's if the two variables are perfectly dependent, while -1 corresponds to a perfect negative dependence.

2.3.1 Gaussian Copula

There are many parametric copula families available, which are characterized by specific parameters that control the strength and the structure of dependence. One of the most popular parametric copula is the Gaussian, which is frequently used to model the trend of credit portfolios and also to study its risk. To have a clearer view of this concept let us consider again the example of a bivariate copula, but taking here the Gaussian one. We take two random variables X and Y of which we know the probability density function (pdf) and the cumulative distribution function (cdf); the aim is to find the joint $F_{XY}(x, y)$. The idea is to map both X and Y into two standard normal variables T_1 and T_2 with a correlation ρ_{12} and a cdf $F_{T_1 T_2}(t_1, t_2)$ which is a bivariate Gaussian distribution which can be written knowing means, variances and correlation. Then, thanks to the normalization property of pdf, the area under the curve stays the same even the variables change, leading to the following equality

$$F_X(x) = F_{T_1}(t_1) = \mathcal{N}(t_1) \tag{2.22}$$

where $\mathcal{N}$ stands for a standard normal distribution. At this point, we can invert the previous relation obtaining

$$t_1 = \mathcal{N}^{-1}(F_X(x)) = \mathcal{N}^{-1}(u) \tag{2.23}$$

$$t_2 = \mathcal{N}^{-1}(F_Y(y)) = \mathcal{N}^{-1}(v) \tag{2.24}$$

Finally, since we know the cdf of T_1 and T_2, to get the joint probability distribution of X and Y, we need just to substitute t_1 and t_2 obtaining

$$\begin{aligned}
F_{XY}(x, y) &= C(F_X(x), F_Y(y)) && (2.25)\\
&= C^{Gauss}(u, v) && (2.26)\\
&:= \mathcal{N}(t_1, t_2) && (2.27)\\
&= \mathcal{N}(\mathcal{N}^{-1}(u), \mathcal{N}^{-1}(v)) && (2.28)\\
&= \mathcal{N}(\mathcal{N}^{-1}(F_X(x)), \mathcal{N}^{-1}(F_Y(y))) && (2.29)
\end{aligned}$$

Usually, since the securities are normally distributed, one can go also further, considering the so-called *1-factor* Gaussian copula. This is very useful, in particular, in case we deal with N random variables. The core of this approach is performing a change of variable from the rv modeling a bond in the portfolio X_i to an expression that allows reducing the number of variables needed to estimate the correlations among the assets: $X_i = a_i M + \sqrt{1 - a_i^2 Z_i}$. Indeed, assuming that M and Z_i are normally distributed, and the loading factor a_i belongs to $[0, 1]$, the correlation among two different assets X_i and X_j is totally encoded in

$$Corr[X_i X_j] = \frac{\sigma_{X_i X_j}}{\sigma_{X_i}\sigma_{X_j}} = \frac{E[X_i X_j] - E[X_i]E[X_j]}{1} = E[X_i X_j] = a_i a_j E[M^2] = a_i a_j \tag{2.30}$$

Finally, to estimate the correlation among all the assets, one need to know just N values, which are the a_i of every asset, instead of from $\frac{N(N-1)}{2}$, i.e. the number of entries of the initial symmetric correlation matrix.

As one can expected, the Gaussian copula is not the only example of copula family existing in literature, although it is the most used. There exist in fact many other parametrizations, such as Gumbel, Student-t, Clayton and so on. They do not differ so much in the grade of association between variable they provide, but more in which portion of the distribution the correlation is stronger. For example, to model crisis periods in the financial market, the correlation between big losses, i.e. in the right tail of the distribution, needs to be remarked. In this case, a Student-t copula is preferable to a Gaussian one, since it better catches this peculiarity. An interesting discussion about some drawbacks of Gaussian copula can be found in Donnelly and Embrechts (2010), where the authors analyse the mathematical models used by the finance industry in the aftermath of the 2007-2008 financial crisis.

2.3.1.1 Numerical Simulation

During the simulation performed in this work, the dependence between corporate bonds (which are modelled as binary variables) is expressed via a Gaussian Copula. The steps below describe the effective construction of what is known as a bivariate Gaussian copula.

1. Generate pairs of values from a bivariate normal distribution. The statistical dependence between these two variables is determined by the Pearson's ρ correlation parameter.

2. Apply the normal cumulative distribution function Φ to the original standard normal random variables N_i with $i = 1, 2$. It results in a random variable $U = \Phi_1(N_1)$ that is uniform on the interval [0,1]. Notice that by computing the linear correlation between the marginals we obtain the rank correlation coefficient of Spearman among the original normals

$$\begin{aligned} \rho^{(\text{Normal})}_{Spearman} &= \frac{Cov(\Phi_1(N_1), \Phi_2(N_2)}{\sqrt{Var(\Phi_1(N_1))Var(\Phi_2(N_2))}} \\ &= \frac{Cov(U, V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \rho^{(\text{Uniform})}_{Pearson} \end{aligned} \tag{2.31}$$

3. Perform with respect to every variable U and V the inverse transformation of Equ. 2.17, by exploiting the desired function, i.e. those whose joint distribution we intended to find.

In general, because the transformation works on each component separately, the two resulting random variables need not even have the same marginal distributions. Following these steps we create dependent random variables with arbitrary marginal distributions. In our case we transform the uniform random variables got after step 2 in Bernoulli exploiting a threshold. In this way, maintaining the dependence structure of the Gaussian copula, we address the problem of joining the discrete distributions of two correlated random variable.

Chapter 3

Decision Theories under Comparison - The Two Horses Race

In this section we present and exploit the main economic theories and portfolio selection models and rules such as Markowitz Portfolio Theory, Expected Utility Maximization, Stochastic dominance ordering. To study the optimal portfolio allocation we consider the situation of two investments and one investor and we work on three parameters: correlation, default probability, and investor's risk attitude. Our contributions result in analyzing how an optimal portfolio is affected by those three parameters. In particular, the results obtained are interpreted under a sustainable finance perspective, in order to figure out the most suitable conditions to promote the low-carbon transition.

3.1 Markowitz Portfolio Theory

The standard problem of portfolio optimization aims to subdivide the initial budget into the activities available for the investment. The optimization is made respect to the vector of weights that define how the portfolio is constituted. A pioneer in this research was Markowitz, who modeled it as a two-objectives optimization problem, i.e. looking for the portfolio that simultaneously maximizes the return and minimizes the risk. However, since the two criteria conflict with each other, such a portfolio does not exist. The idea is therefore to fix one of the two objective to a specific value, which is strictly related to the degree of risk one is willing to take, while optimizing the other.

$$\min_{\vec{w}} \frac{1}{2}\vec{w}^T \mathbf{C}\vec{w} \quad \text{with} \quad \vec{w}^T\vec{\mu} = \eta; \quad \vec{w}^T\mathbf{1} = 1 \tag{3.1}$$

$$\max_{\vec{w}} \vec{w}^T\vec{\mu} \quad \text{with} \quad \vec{w}^T\mathbf{C}\vec{w} = \xi; \quad \vec{w}^T\mathbf{1} = 1 \tag{3.2}$$

where $\vec{w}$ is the vector of weights, $\vec{\mu}$ of means values, $\mathbf{C}$ the covariance matrix; moreover, η and ξ are the values for mean and variance respectively that the investor is willing to accept and thus they depend on the risk aversion of the investor. These trade-off portfolios are usually called efficient portfolios and their mapping to the expectation/variance space is the efficient frontier. Thanks to its mathematical tractability, Markowitz's portfolio theory, also known as the mean-variance analysis, is one of the most popular theories of asset allocation. Some of the assumptions at the basis of this theory are extracted directly from Markowitz (1952) and may be summed into:

1. The risk of a portfolio is based on the volatility of returns;
2. The markets are efficient, a short sale is not allowed;
3. The information on returns and risk, known by the investors, are fair and correct;
4. Only one single period of investment is considered;
5. The investors are *rational*, meaning that they want to maximize their utility, which is concave and increasing;
6. The investors are risk-averse and try to minimize the risk and maximize the return;
7. Investors base decisions on expected returns and variance (or standard deviation) of these returns from the mean;

As regard of point 2, the volatility coincides with the standard deviation. It involves possible deviations of the obtained returns from their expected value, hence a high standard deviation is intuitively named as high risk. As a simple and intuitive risk measure, it is widely accepted among professional investors as well as academics. Although, it too has its drawbacks. The main one is that it consider both the good and the bad deviations from the mean as risky. Thus, the only way to increase the return on the portfolio is to increase the level of risk that the investor is willing to accept. An investor can reduce the risk if his portfolio consists of combinations of instruments that are not perfectly positively correlated. Diversifying, in this case, brings the same expected return for the same portfolio, but with reduced risk. Consider for instance to have one investor and two assets, the variance of the resulting portfolio results into:

$$var[Z] = var[w_A A + w_B B] = w_A^2 var[A] + w_B^2 var[B] + 2\rho w_A w_B \sqrt{var[A]}\sqrt{var[B]} \tag{3.3}$$

where $Z = w_A A + w_B B$ is the 2-assets portfolio payoff and A and B are the variables modeling the singe assets.

If they correlate 0 (i.e. they are unrelated) the variance of the portfolio is the sum of all the individual asset variances, weighted by the square of the amount invested in each of them. If they correlate 1 (i.e. they are positively correlated), the standard deviation of the portfolio yield is the sum of the standard deviations of the fraction-weighted assets returns held in the portfolio. The definition of an *efficient portfolio*, contained in Markowitz (1959) says that: a portfolio is efficient if it is not possible to obtain a greater expected return without increasing the standard deviation or, in other words, reducing the standard deviation without giving up the return. Within the set of portfolios defined as efficient, several sub-sets can be defined according to the level of risk expressed by the investor. To sum up, through Markowitz's optimization, one can find all the feasible portfolios offering the best compromise between risk and return. From a computational point of view, the most common approach is one that seeks to minimize the variance with respect to the weights, while the mean is constrained to have a certain value between the minimum and the maximum return, which are the two known points of the efficient frontier. The variance is indeed a quadratic function of the weights and, for this reason, the problem can be solved in the field of quadratic programming. Moreover, if the objective to be optimized has a convex function, as happen for the variance, the efficiency of the algorithm increases. However, even if with higher computational cost, one can do also the opposite, maximizing the return with the volatility set at a certain value. In this case, one can exploit linear programming algorithms with quadratic constraints.

3.2 Horses Race with Mean-Variance

In the simulation performed, we deal with the case of one investor and two projects on which she decides to allocate her wealth. For simplicity of discussion, the optimal investment problem is modeled as a hazardous bet on horse racing. Every horse has a probability of success p_i and consequently a certain return if the bet is successful (for example, the bettor wins twice as much as invested) and a recovery rate $\in [0, 1)$ if the project fails. For the purpose of this exercise, we allow the correlation coefficient ρ to take any values in the range $[-1, 1]$. This means that, if they are positively correlated, they cooperate to win, while, if they are negatively correlated, when one wins the other has a higher chance of losing. The goal is to understand how the bettor's optimal portfolio allocation changes as the external parameters (default/failing probability, correlation, risk attitude of the investor) change too. It may happen, for instance,that after a change in correlation, the bettor decides to drastically change her investment strategy, moving from a diversified investment to mono-horse one. In the next we display the optimization formula, here in the "version" aiming at maximizing the expected return but at fixed variance. Even if it is characterized by a high computational cost, through this version every term of the expression below is more easily interpretable. Indeed, with the Lagrangian multiplier λ one can settle the "importance" given to the variance, with respect to the mean. A higher and positive λ lead to a shrinkage of the difference between $E[Z]$ and $\frac{1}{2}\lambda var[Z]$, reflecting the behaviour of a risk-averse investor.

$$\max_{\vec{w}} \left\{ E[Z] - \frac{1}{2}\lambda var[Z] : \vec{w}^T \cdot \vec{\mathbf{1}} = 1 \right\} = \max_{w_i \forall i} \left\{ w_i \mu_i - \frac{1}{2}\lambda w_i \sigma_{ij}^2 w_j : \sum_{i=1} w_i = 1 \right\} \tag{3.4}$$

where Z is the portfolio payoff, $\vec{\mu}$ is the vector of the expected value of the horses racing payoffs and σ_{ij}^2 is the ij element of the covariance matrix.

3.2.1 Calibration of the setting: issues and insights

Before going into details we briefly mention some issues we faced in performing the simulation and some insights gathered from the calibration of the setting. First, since our horses are correlated, we exploit a Gaussian copula to calculate the joint probability distribution between the Bernoulli's variables. In particular, as anticipated at the end of the previous chapter, the procedure starts by generating two gaussian vectors with correlation coefficient ρ, that is the correlation coefficient only among the gaussians. Then, taking the marginals of each of the two gaussians variable (i.e. of each column of the matrix) we obtain a nx2 matrix **U** of uniformly distributed variables which are still correlated, although the correlation coefficient is changes. The link between the two correlation coefficient see Equ. 4.12. Finally, to create two vector of drowns made by the realization of the two correlated Bernoulli variables A and B, we exploit a threshold. Since in **U** the numbers are uniformly distributed the probability to draw a number bigger than p_A, which is the probability that A wins, corresponds to p_B, the winning probability of horse B. We make some considerations about how the correlation is transferred from Gaussian to Bernoulli in the following Chapter.

Second,we wondered how could we quantify the risk aversion of an investor, in other words, which values assign to lambda. We discovered that there exist numerous qualitative psychological test to help the investor determines her psychological profile. For instance, the *PASS test* by W.G. Droms, the *Baillard, Biehl and Kaiser test* (which distinguish investors on a scale from "confident" to "anxious" and "careful" to "impetuous") that of *Barnewal test* which distinguishes between passive and active investors, or the *Bonpian test*, with eight types of investor. Leveraged on their suggestion we assigned to λ a number from 1 (lowest risk aversion) to 5 (highest risk aversion) to an investor.

Third and last, concerning the investor attitude towards risk, we handled the parameter λ in order to test if we could describe also a risk lover behaviour. As expected (since part of the Markowitz's hypotheses), the model fails. Indeed we discovered that the risk lover does not provide significant results because increasing the points of the probability, a gap is formed, which is an indication of an issue. This is because there is a point beyond which risk lover is so fond of risk that she prefers a lower expected return in order to have more variance. Such a choice cannot be defined as efficient precisely because a lower return should be accepted in exchange for a lower risk, but not for the contrary. This risk lover can be defined colloquially blinded by risk. Another problem related to the risk lover that led us to exclude it from this first analysis is the fact that its behaviour is influenced by the initial guess; since there are two equivalent maximum to the extremes of the domain, the algorithm chooses which of the two to prefer according to the initial point.

3.2.2 Case study: Comparisons between Three Risk Averters

In the following we illustrates the result of the mean-variance analysis applied to a portfolio of two bond (or horses) having different probability of success p_i with $i \in \{A, B\}$, but same payoff. The payoff of B in case of winning is computed to be inversely proportional to the probability of winning, i.e $x_{Bwins} = 1/p_{Bwins} \approx 1.11$. As concerns of the recovery rate, it is fixed at zero, meaning that if the horse loses, the investor loses all the money invested in it. Then, since we want the two payoffs equal, $x_{Awins} = x_{Bwins} \approx 1.11$. Hereafter, the objective function of the optimization procedure is displayed:

$$\begin{aligned} f(w) &= E[Z] - \frac{1}{2}\lambda var[Z] \\ &= wE[A] + (1-w)E[B] - \frac{1}{2}\lambda\left(w^2 var[A] + (1-w)^2 var[B] + 2\rho w(1-w)\sqrt{var[A]var[B]}\right) \end{aligned} \tag{3.5}$$

Here p_A, i.e. the probability of horse A wins is varied from 0.8 to 0.95, while p_B is fixed to 0.9. In every subplot of Figure 3.1 a different investor behaviour is shown.

- $\lambda = 0$ corresponds to a risk neutral investor;
- $\lambda = 2$ refers to an investor that is risk averse but not so conservative;
- $\lambda = 6$ is a risk averter very conservative.

The behaviour for a risk neutral is easily explicable. Having $\lambda = 0$ the variance doesn't enter the optimization, hence the bettor chooses only based on the expected value. Needing only to maximize the expected value, she would funds her money in the horse having the highest mean. In our example the horses are modeled as Bernoulli variables. This means that, given the same payoff, the expected value of A would be higher than B, as soon as the probability of A overcomes the one of B. Looking at Figure 3.1, as long as $p_A < 0.9$, the investor will put everything on B, then for $p_A = 0.9$ he will equidistribute his wealth, according to the initial guess given to the optimization algorithm. For $p_A > 0.9$, as expected, the investor will move his wealth to A. In general, B remains advantageous as long as A is less likely than B. For what concern the two risk averters, the two plots (top-right and bottom-left) are quite similar even if the severity of the riskiness of the investor changes.

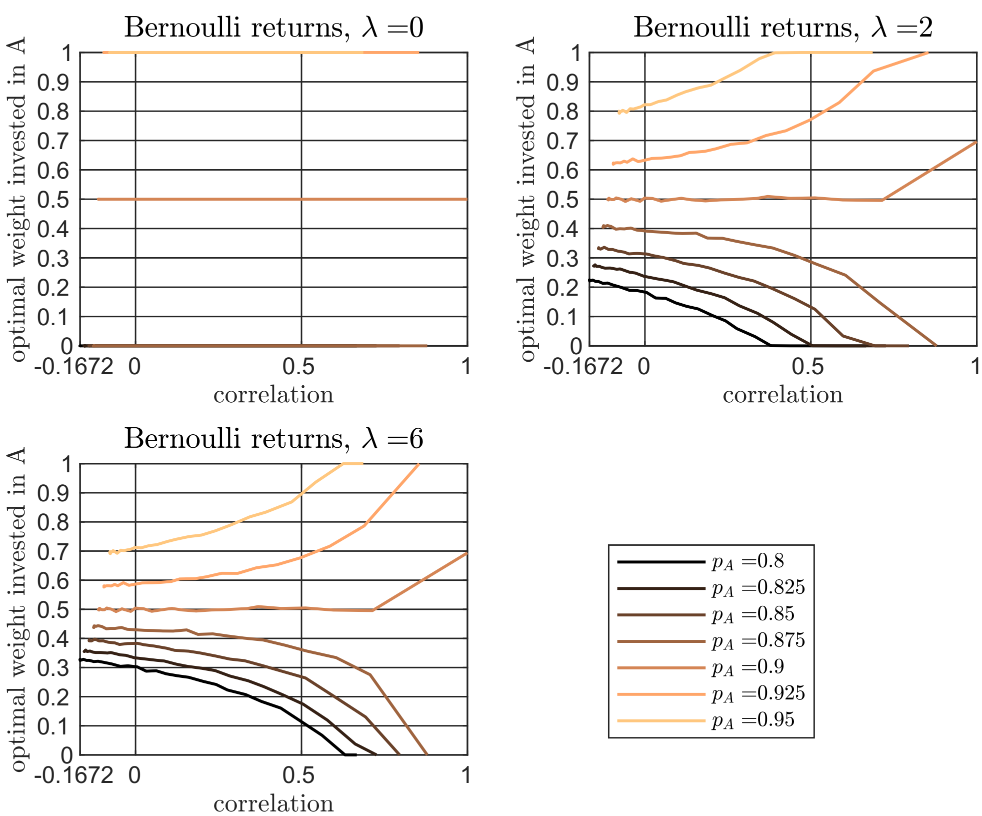

Figure 3.1: **Optimal weight** found by mean-variance method as the correlation coefficient varies, **with λ fixed**. Every curve in each subplot corresponds to a specific probability value p_A, while $p_B = 0.9$.

Looking at how the brownish curves changes varying the correlation, one can see that they are approximately constant as long as the value of the correlation coefficient does not exceed a certain value, which is higher if the probabilities of winning are similar. From this point on, the curves more or less slowly go to 1 or 0 depending on whether p_A is greater or less than p_B, respectively. Besides, if we have a look at the graph with $\lambda = 6$ it can be seen that the investor will maintain a diversified portfolio for more time. This behaviour is consistent with his very conservative attitude which led him to diversify. It is interesting for our analysis highlighting that when the correlation diminishes the investor modifies her portfolio towards a more diversified one. This behavior is emphasized for a more conservative risk averter. Indeed, observing the bottom right subplot in Figure 3.2, one can see that, even the investment A has a bigger probability of success than B, decreasing the correlation the most conservative investor is propelled to diversify his wealth. If we think about these results in light of a sustainable finance perspective, assuming for a moment that A and B are no longer the horses participating in a race but two bonds, an attitude of this kind may support a low-carbon transition. From the current composition of investment portfolios emerges a predominance of carbon-intense corporate bonds, i.e. bonds that are somehow related to intensive CO_2 production. Therefore, it is reasonable to think that conventional bonds most of the time are preferable. Now imagining that the correlation between low and high carbon intense bonds is > 0.5 and that brown bonds have a higher probability of success. Figures 3.1 and 3.2 show that if there exists an external event able to change the correlation between investments, decreasing it, at that point we could see a shift in capital towards more sustainable investments. Hence, we can think about a decrease in correlation as a positive fact in support of the sustainable investments. A stronger connection with market data and conditions is illustrated in the following subsection.

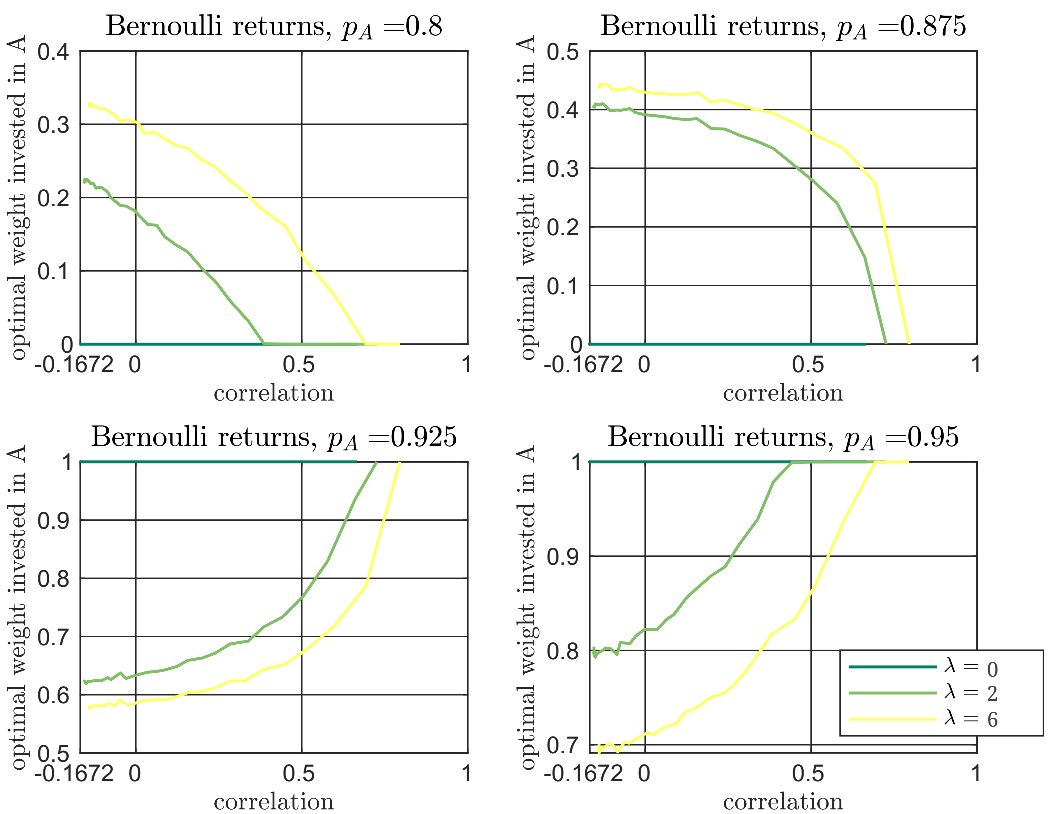

Figure 3.2: **Optimal weight** found by mean-variance method as the correlation coefficient varies, whereas $\mathbf{p_A}$ **fixed**. Three risk trends are shown in each subplot: $\lambda = 0$ (Risk neutral), $\lambda = 2$ (Risk averse, less conservative), $\lambda = 6$ (Risk averse, very conservative).

3.2.3 Market Equilibrium Condition

We consider now to be at market equilibrium meaning that

- $x_{A_{wins}} = 1/p_{A_{wins}}$
- $x_{B_{wins}} = 1/p_{B_{wins}}$

It follows that the expected values of A and B are equal, whatever the probability of winning: $E[A] = E[B] = 1$. To analyse this case, we consider $p_{B_{wins}} = 0.9$ fixed, as the previous analysis, whereas can moves in a wider range $p_{A_{wins}} \in (0, 1]$. By assuming again to deal with market data instead of a betting, the values chosen for $p_{A_{wins}}$ and $p_{B_{wins}}$ acquire new meaning. B is chosen in a way that could represent a common sovereign bond, hence 10% of return with probability 0.9. A, instead, embodies the behaviour of bonds with low probability of default. Indeed, a low winning probability may model the sales conditions of very unstable bonds, such as the Argentinian Tango bonds. Generally speaking, unstable bonds are usually linked to states that risk to bankrupt, which, especially during economic crisis such as the current devastating one caused by the Covid-19 pandemic, navigate in more and more stormy waters. Similarly, a pretty low $p_{A_{wins}}$ could represent a corporate bond issued by a company that may attempt the so called “gambling for resurrection” to save itself from bankruptcy. With only two variables (A and B) the solution can be calculated analytically. The expected value of the portfolio return is

$$\begin{aligned} &E[Z] = wE[A] + (1 - w)E[B] \\ &\text{but having} \quad E[A] = E[B] \quad \text{it will be} \\ &E[Z] = E[B] = 1 \text{ , always.} \end{aligned} \tag{3.6}$$

Therefore, the mean-variance optimization (see Equ. 3.5) results into the simple minimization of the variance without any constraints on the mean. Moreover, as one can see from the formula above, the optimal weight $w*$ assigned to A is unaffected by λ. This means that the $w*$ will be the same for any risk-averse investor, regardless of her more ($\lambda > 5$) or less conservative ($\lambda < 5$) nature. For what concern the risk neutral investor, i.e. with $\lambda = 0$, the expected utility results in $E[f(w)] = E[Z(w)] = E[B] = 1$. It follows that here any portfolio can be said “optimal”, since they are all equivalent. That means that, in the numerical simulation, the initial guess chosen, will also represent the optimal investment. These behaviours are shown in Fig. 3.3. Let’s now look again the picture from right to left, that is as the correlation decreases. As hint before, when the correlation is very large and positive, it is not convenient to diversify and therefore the investor will invest in the asset with less variance, since the optimization method chosen (mean-variance) at market equilibrium, as mentioned earlier, depends only on the variance. Decreasing the correlation, the tendency is to invest

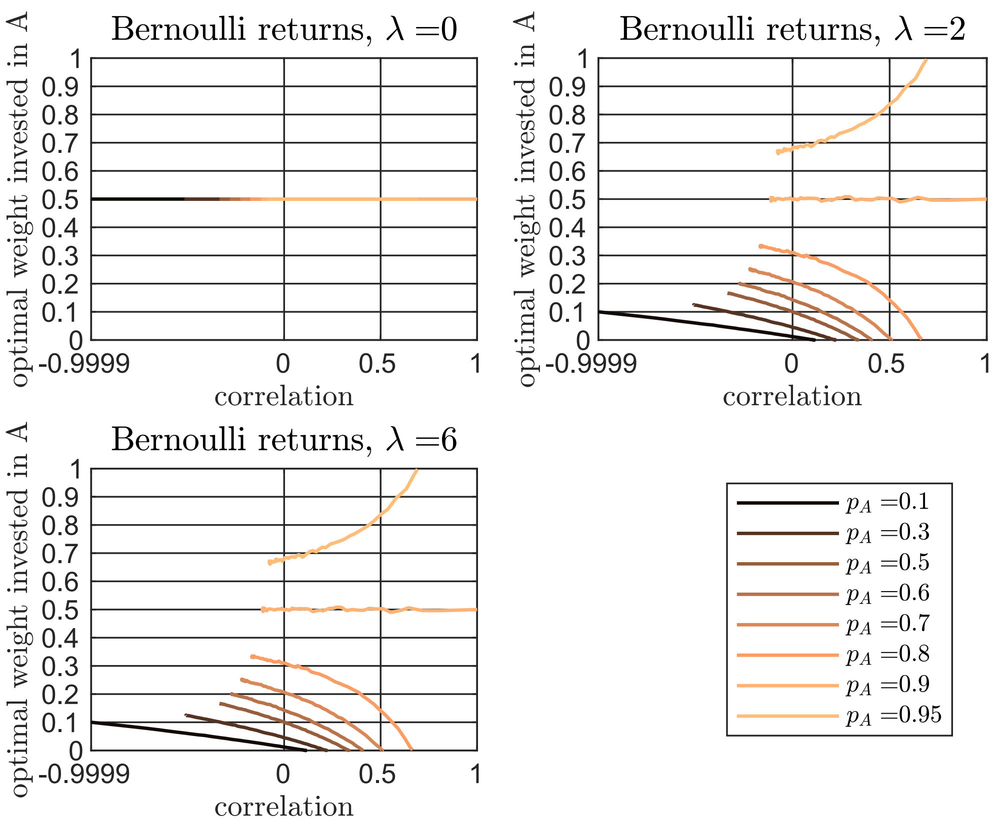

Figure 3.3: **Optimal weight** found by mean-variance method at **market equilibrium** as the correlation coefficient varies, whereas λ **fixed**. Here we are considering $\mathbf{p_B = 0.9}$

.

more fairly in the two investments. As one could see in section 4.1.2 the upper and lower bounds of the linear correlation coefficient change according to the probabilities of the two variables. Therefore, some values are not attainable and this is the reason why the curves, representing the probability of A to win, have different existence domains. The presence of different domains is evident also from Figure 3.4. For instance, the value $\rho \approx 1$ can be attained only when the two horses behave the same, i.e. $p_A = p_B = 0.9$. Every curve represents indeed a value of Pearson's ρ. One can see that going from a lighter brown to a darker brown curve, at fixed value of probability of A to win, the correlation decreases. This lead to an increase of the amount set up on A. Besides, as one can expect, growing with the probability of A to win, at fixed ρ ,the amount in A grows too.

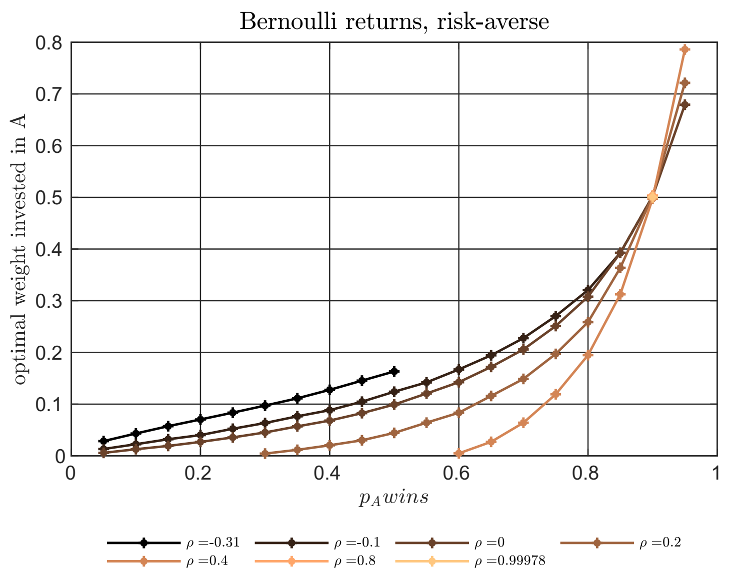

Figure 3.4: **Optimal weight** found by mean-variance method at **market equilibrium** as the probability of A wins, whereas λ **fixed**. Here we are considering $\mathbf{p_B = 0.9}$

.

3.3 Expected Utility Maximization

The investor, defined *rational* on the basis of a certain number of axioms contained in Von Neumann and Morgenstern (1947), is faced with several options, each characterized by a certain level of risk, among which she must make a choice. To do so, the decision-maker uses a utility function that weights each possible result. In the end, her choice will fall on the decision that leads her to reach the maximum possible utility, i.e. her weighting function. For what concern the marginal utility, it corresponds to the slope of the utility function at that wealth level. Chosen a level of utility U(x*), and given the expected mean μ and the standard deviation σ, one can define the indifference curve in (μ, σ) plane, that is the locus of points (μ, σ) with constant expected utility, i.e. equal to U(x*). The utility function satisfies the basic hypotheses presented by Von Neumann and Morgenstern (1947), i.e. it is determined up to linear transformations, it is not decreasing, and the objective is to maximize its expected value. The utility is a function of the wealth, but in our case, of the expected portfolio payoff. The probability distribution of the portfolio can be of any type, discrete or continuous, with finite or infinite support. One way to estimate the investor's degree of risk aversion is thought the certainty equivalent, i.e. the amount of money an investor would accept to avoid the risky investment. To calculate the certainty equivalent, CE, one has to simply take the inverse image (or preimage) of the expected utility. In the case of risk aversion, this CE results to be less than the expected value of the portfolio payoff. Another fundamental concept to describe risk aversion is the risk premium. It is defined as the difference between the expected value and the certainty equivalent of the same investment. The risk premium is the maximum amount that the individual would pay for the investment to be risk-free. The larger and more positive this difference is, the greater the concavity of the utility and the more the investor will prefer to get a lower but certain return, rather than take some risk. Another way to measure her risk attitude is through Arrow-Pratt's risk aversion coefficient. It is also how we are going to account risk in this section. There exist two form of the risk coefficient. One is the Absolute Risk Aversion coefficient (ARA),

$$ARA = \lambda = -\frac{U''(Z)}{U'(Z)} \tag{3.7}$$

With Equ. 3.7 one can easily find whether and how risk aversion varies as wealth increases, by analyzing if the absolute risk aversion coefficient is an increasing, constant or decreasing function of wealth. The other coefficient is the Relative Risk Aversion one

$$RRA = -Z\frac{U''(Z)}{U'(Z)} \tag{3.8}$$

where Z is the portfolio payoff.

3.3.1 Classes of Utility Functions

Generally, among the classes of utility functions available in the literature (see for instamce Brunnermeier (2015)), the most used is the one that present HARA (Hyberbolic Absolute Risk Aversion). This class includes utility functions called CARA, CRRA, and quadratic utility functions. The choice falls on this class because of its mathematical tractability. A utility function exhibits HARA if its absolute risk aversion is a hyperbolic function, as

$$\lambda = -\frac{U''_{HARA}(w)}{U'_{HARA}(w)} = \frac{r}{1 + \alpha\mathbb{E}[W]} \tag{3.9}$$

depending on the values of r and α, the utility function changes. By varying only these two coefficients it is possible to describe all possible psychological profiles of the investor. Setting r positive, (=1 for simplicity), null or negative, we are considering an investor respectively risk averse, risk neutral and risk lover. Then, once r is chosen, by varying α you change the concavity of the utility function, increasing or decreasing the risk. The next two paragraphs formalizes some definition regarding the different risk attitudes.

3.3.1.0.1 Risk Aversion A risk averter investor can be defined throughout different ways. Hereafter two of them are presented. For a more detailed discussion see pag. 89 of Levy (2015) to more details.

1. A risk averter will not play a fair game;

2. A risk averter will be ready to pay a positive risk premium to insure their wealth.

3.3.1.0.2 Risk Seeking utility function The utility function, whose expected value is maximized, is convex for a risk-seeker. The effect of the concavity takes the decision-maker to choose, between two options with the same expected return, the one more spread, i.e. with higher variance. This concept rely on the mean-preserving spread, which is recalled at section 3.4.0.1).

3.3.1.1 CRRA or CARA?

As shown in Equ. 3.9, in the most general version of the HARA utility function two free parameters have to taken into account. To simplify, we investigate the different 1-parameter subclasses to choose the best ones for our dissertation. CARA (= Constant Absolute Risk Aversion) is a convenient assumption, in particular if the return are normally distributed. However, most agree that absolute risk aversion is not constant, but decreases with the level of wealth, i.e. it presents DARA (= Decreasing Absolute Risk Aversion). This means that richer people are willing to take a greater risk. It makes sense if you think that a wealthy person, even when faced with a large loss, will not feel the same effect as a less wealthy person. For this reason, CARA functions are rejected in favor of DARA ones. A family of function belonging to the DARA class, is the one having Constant Relative Risk Aversion (CRRA). According to Chiarella et al. (2008), fundamentalist traders use constant relative risk aversion utility function (CRRA) to evaluate their propensity to the risk. For these reasons, we turn our choice to this function, exploiting here its power version

$$U(Z) = Z^{\eta}, \tag{3.10}$$

where

$$\text{Risk Attitude} = \begin{cases} Aversion, & \text{if} \quad 0 < \eta < 1 \\ Neutral, & \text{if} \quad \eta = 1 \\ Seeking, & \text{if} \quad \eta > 1 \end{cases}$$

Below we show either that the absolute risk aversion coefficient is decreasing or that the relative one is constant. Then, we compute the first and second derivatives of a power utility.

$$\begin{aligned} U'(Z) &= \eta Z^{\eta-1} \\ U''(Z) &= \eta(\eta-1)Z^{\eta-2} \end{aligned} \tag{3.11}$$

Keeping in mind that we defined λ as the absolute risk aversion coefficient, we will have:

$$ARA = \lambda = -\frac{U''(Z)}{U'(Z)} = -\frac{\eta(\eta-1)Z^{\eta-2}}{\eta Z^{\eta-1}} = \frac{1-\eta}{Z} \tag{3.12}$$

which, as expected, is a decreasing function of Z, i.e. the portfolio payoff. By consequence, the relative risk aversion coefficient will be

$$RRA = -Z\frac{U''(Z)}{U'(Z)} = -Z\frac{\eta(\eta-1)Z^{\eta-2}}{\eta Z^{\eta-1}} = 1-\eta \tag{3.13}$$

3.3.2 Direct Maximization of a CRRA Utility Function

We decide now to move forward, setting up an original experiment where the efficient portfolio is found by direct optimization of the utility function. As mentioned at the beginning of this section, the efficient portfolio is found by deriving directly the maximum of the expected power utility function, defined in Equ. 3.10. In order to analyze the problem from each of its perspectives, we tune again (as was done for the mean variance) the 3 parameters available to us: the probability of the bond A to receive the premium, the linear correlation coefficient between the two bonds; investor's risk attitude. Then, we combine them obtaining 6 different experiments. The set up of each experiment is the following

- The market is at equilibrium;
- The risk attitude is defined by the parameter η proper of the power utility. Remember that
  - Risk Aversion: $0 < \eta < 1$
  - Risk Neutral: $\eta = 1$
  - Risk Seeking: $\eta > 1$
- $p_B = 0.9$;
- p_A takes each value in the range $[0.05, 1]$, with an increasing step of 0.05;

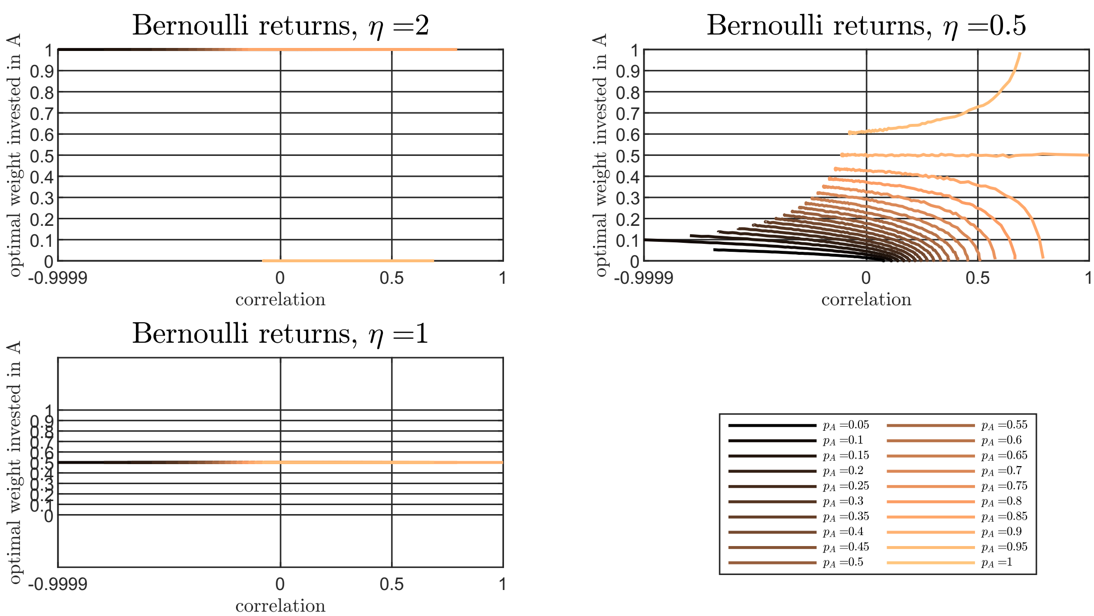

Figure 3.5: **Optimal weight** found by expected utility maximization at **market equilibrium** as the probability of A to win and the correlation are varying, whereas η is **fixed**. Here we are considering $\mathbf{p_B = 0.9}$.

The subplots in Figure 3.5 trace the behaviour of the ones in the previous sections. Since the utility of a risk neutral investor (η=0)is equal to the payoff itself, maximizing the expected utility coincides with performing the expectation value of the portfolio payoff at market equilibrium, which return one as illustrates in Equ. 3.6. Therefore, for her any portfolio can be said "optimal", since they are all equivalent. Here the optimal weight turn to be 0.5, which corresponds exactly to the initial guess of the algorithm. Regarding the risk lover, she always make a sharp choice: all in A or all in B. In particular here, differently from Fig. 3.3 she is more prompt to invest in the bond with highest payoff, although the tiny probability of winning. The behaviour of the risk averter is fully investigate in Fig.

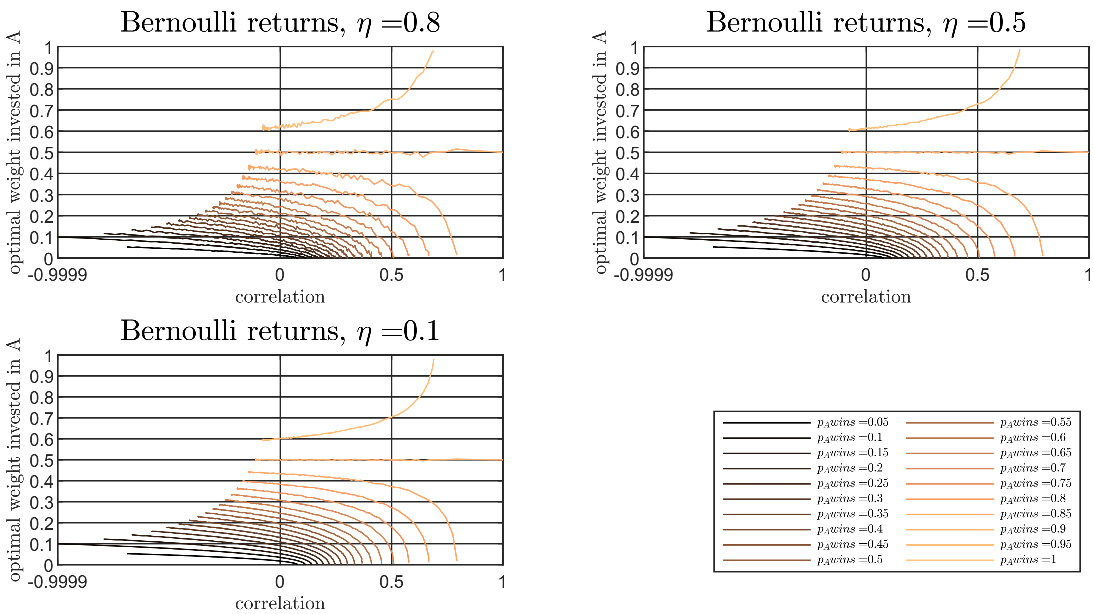

Figure 3.6: **Optimal weight** found by expected utility maximization at **market equilibrium** as the probability of A to win and the correlation are varying, whereas η is **fixed**. Here we are considering $\mathbf{p_B = 0.9}$ and only **risk averse investors** (from the less conservative, having $\eta = 0.8$ to the most one $\eta = 0.1$)

3.6. It can be seen that the line describing a less conservative attitude ($\eta = 0.8$) is less clear than the one of a more conservative decision-maker ($\eta = 0.1$). The overall fashion follows Fig. 3.3 and reinforces the fact that this may be the crucial point dominating the interplay between correlation and financial contracts. Indeed, decreasing the correlation, the tendency is to invest more fairly in the two investments, leading to a strong alteration of the status quo of portfolio management. Besides, as happens in the previous sections, some values of the correlation are not attainable and this is the reason why the curves, representing the probability of A to win, have different existence domains. Looking at the Fig. 3.7 stands out that the change in the allocation arises in a tiny range of correlation. This can suggest that it may exist a "threshold" value for the correlation among the Bernoulli's variables below which, a fixed p_A the portfolio totally changes its aspect, favouring a diversified one. As always, the most interesting cases (which in the end are also those most likely to happen in reality) are those related to a risk averse investor. The three subplots relative to this attitude show a trend equal to that of Figure 3.8 and consequently also the considerations are similar.

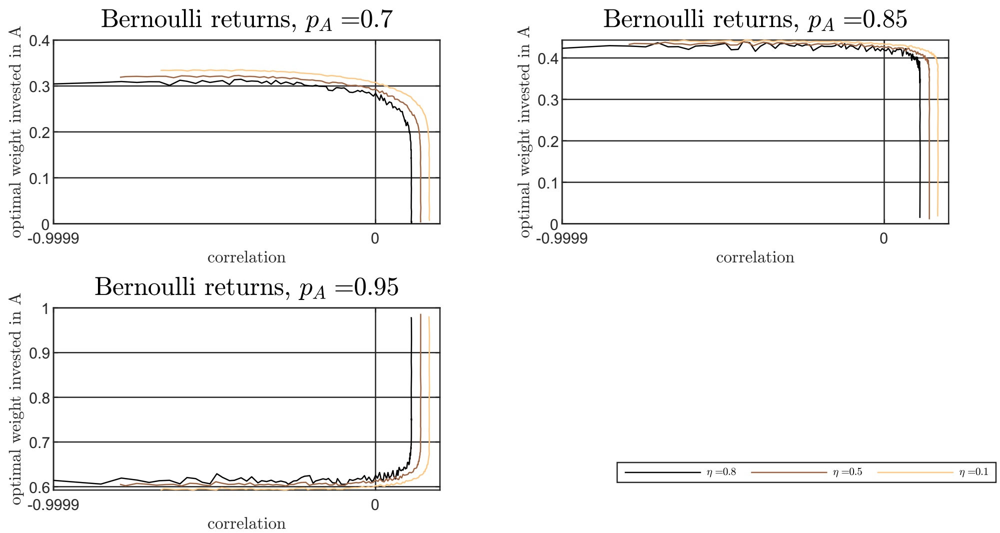

Figure 3.7: **Optimal weight** found by expected utility maximization at **market equilibrium** as the risk attitude, related to η, and the correlation are varying, whereas $\mathbf{p_A}$**wins** is **fixed**. Here we are considering $\mathbf{p_B = 0.9}$.

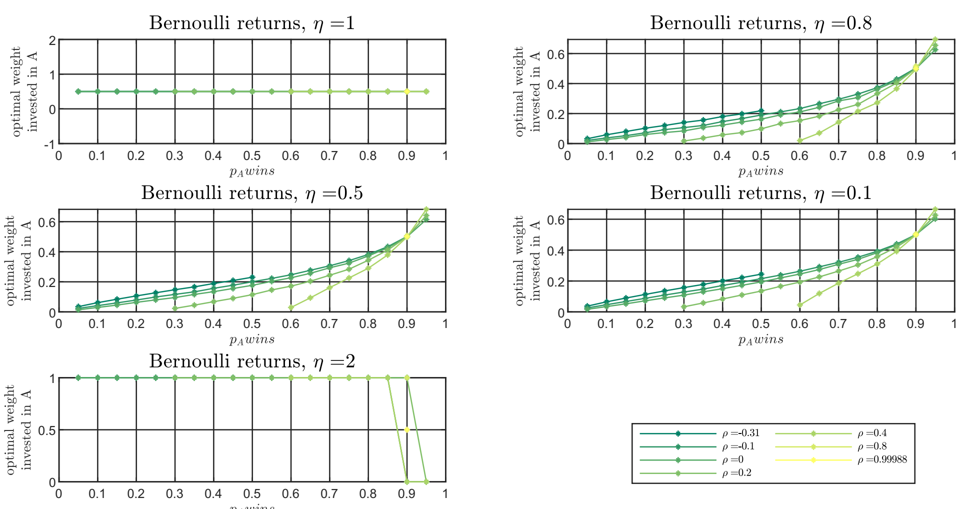

Figure 3.8: **Optimal weight** found by expected utility maximization at **market equilibrium** as the probability of A to win and the correlation are varying, whereas η is **fixed**. Here we are considering $\mathbf{p_B = 0.9}$ **and 5 risk attitudes**.

The three figures (one for each risk attitude) coming next, display the optimal allocation of a portfolio made up by two bonds (supposing one green and the other brown). The probability of A to win and risk attitudes are varying, whereas linear correlation ρ is fixed. $p_B = 0.9$ as usual. One can notice that either the correlation is positive or negative increasing p_A the amount invested in A increases too.

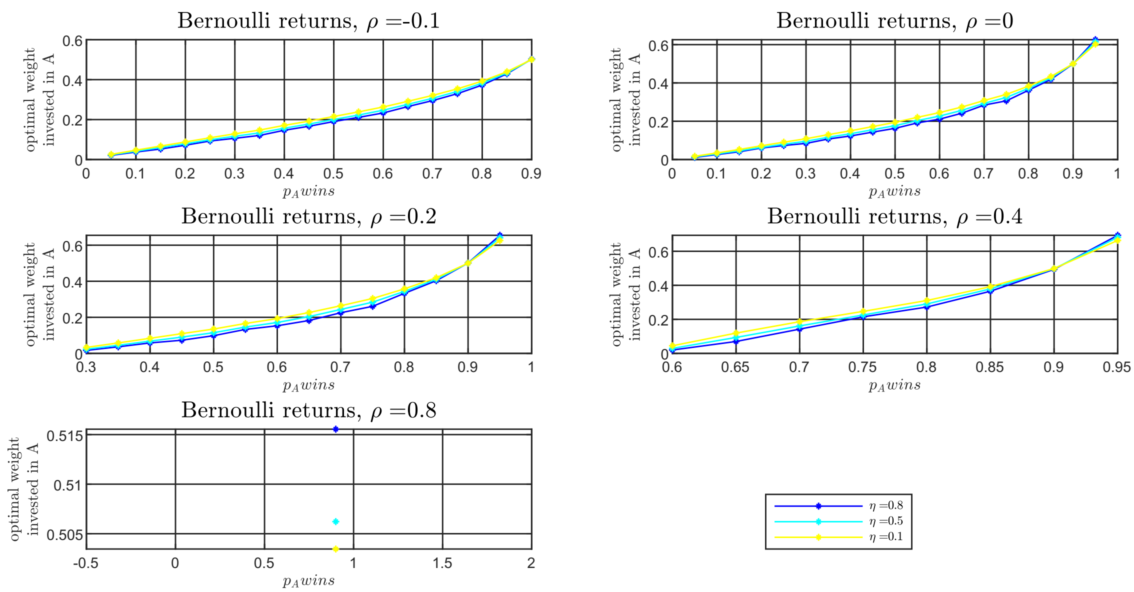

Figure 3.9: **Optimal weight** found by expected utility maximization considering only **risk averse investors** (from the less conservative, having $\eta = 0.8$ to the most one, with $\eta = 0.1$)

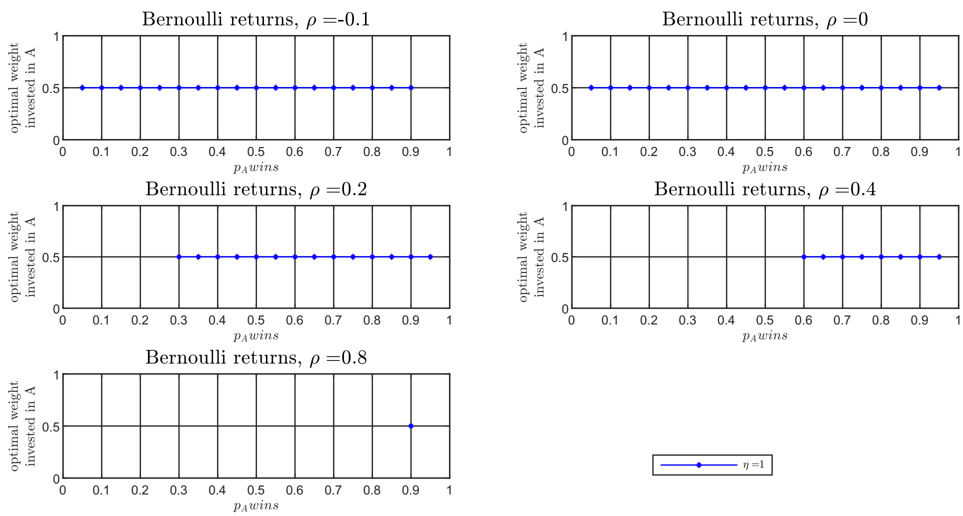

Figure 3.10: **Optimal weight** found by expected utility maximization considering only a **risk neutral investor**, $\eta = 1$.

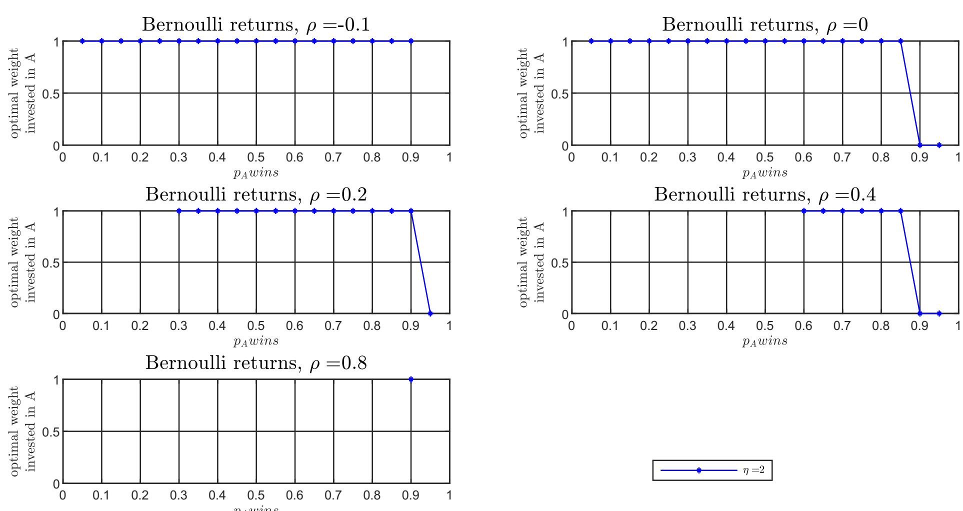

Figure 3.11: **Optimal weight** found by expected utility maximization considering only a **risk seeking investor**, $\eta > 1$.

With the next two pictures a few last insights are gained. Observing the Figure 3.12 you can see that while changing risk attitude, i.e. moving moving towards the right on the x-axis, you pass from a mono-asset portfolio to a diversified one. Moreover, at fixed value of ρ the weight assigned to A increases with its probability of success, while it is practically constant changing only the predisposition of the averter towards the risk.

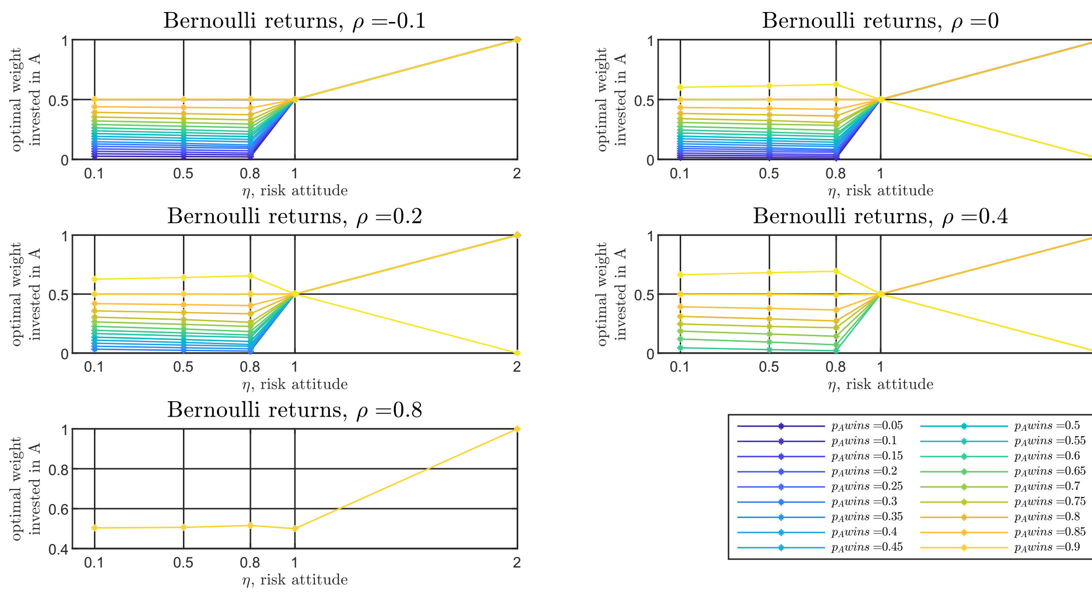

Figure 3.12: **Optimal weight** found by expected utility maximization at **market equilibrium** as the risk attitude (embodied in η) and the correlation coefficient ρ are varying, whereas ρ is **fixed**.

Figure 3.13 illustrates that a larger and more positive correlation corresponds to a less balanced distribution of funds between A and B. So if $p_A < p_B$ the optimal weight of $\rho = 0.4$ is less than $\rho = 0.2$. This makes sense because the more the two horses/bonds are positively correlated the more it seems to have only one bond rather than two. Hence, since their behavior in response to external events will be the same, the greater weight will go on the safer, though less profitable, investment. The precise value of the weight will depend on the balance between higher gain and lower risk performed by the Von-Neumann optimization.

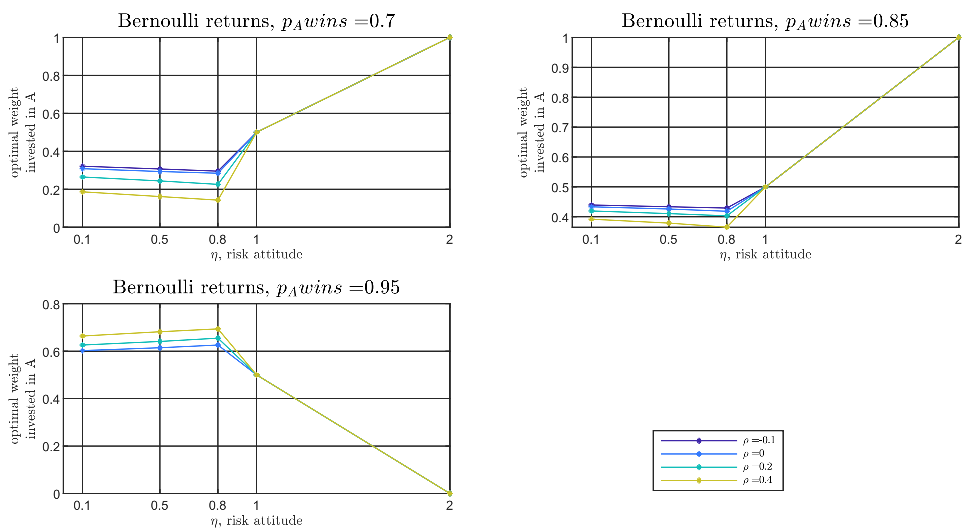

Figure 3.13: **Optimal weight** found by mean-variance method at **market equilibrium** as the risk attitude (embodied in η) and the probability of A to win are varying, whereas $\mathbf{p_A}$**wins** is **fixed**.

3.3.3 Markowitz encounters Expected Utility

In this subsection we are going to investigate how to link the two optimization methods presented above. You will discover that you don't always directly maximize the function that describes the utility, but you can proceed with an approximation. Thus, we start expanding the utility function around the point Z^*, i.e the wealth wished. The only hypothesis you need to perform the expansion is that $U(Z)$ is smooth in Z, where Z is the level of wealth of the investor, which in our case coincides with the portfolio payoff.

$$\begin{aligned} U(Z) &\approx U(Z^*) + U'(Z)\big|_{Z=Z^*} \cdot (Z - Z^*) + \frac{1}{2}U''(Z)\big|_{Z=Z^*} \cdot (Z - Z^*)^2 + o(Z^2) \\ &= \underbrace{U(Z^*) - U'(Z^*) \cdot Z^*}_{costant} + U'(Z^*) \cdot Z + \frac{1}{2}U''(Z^*) \cdot (Z - Z^*)^2 + o(Z^2) \\ &= \text{cost.} + U'(Z^*) \cdot \left[Z + \frac{1}{2}U''(Z^*) \cdot (Z - Z^*)^2 \right] + o(Z^2) \end{aligned} \tag{3.14}$$

From Equ.3.7 we get $U''(Z) = -\lambda(Z) \cdot U'(Z)$ and thus follows

$$U(Z) \approx U'(Z^*) \cdot \left[Z - \frac{1}{2}\lambda(Z^*) \cdot (Z - Z^*)^2 \right] + \text{cost.} \tag{3.15}$$

In our exercise the point around which the Taylor's expansion is performed is the expected value of the portfolio payoff, i.e. $E\big[Z\big]$. Now, taking the expectation of the Formula 3.15 and, neglecting the constants, we get

$$\begin{aligned} E\big[U(Z)\big] &\propto E\big[Z\big] - \frac{1}{2}\lambda \cdot E\Big[(Z - E\big[Z\big])^2\Big] \\ &= E\big[Z\big] - \frac{1}{2}\lambda \cdot var\big[Z\big] \end{aligned} \tag{3.16}$$

The latter (Equ. 3.16) illustrates the same result of the Equation 3.5 in section 3.2.2. This implies that as long as the utility function can be said to be smooth in its argument, one can always approximate the expected utility optimization with the mean-variance one. For what concerns finding a perfect match between the two methods of portfolio optimization, the accordance between expected utility theory and mean-variance principle appears only under specific assumptions regarding the utility function of the investor and/or the probability distributions of the target variable, here the portfolio payoff. The two theories perfectly match either in case of normally distributed return and negative exponential expected utility or when the utility is quadratic, no matter about the portfolio return distribution. For the latter, the expected value of a quadratic utility will be a linear function of the expected portfolio payoff and the expected value of the payoff squared (which is equal to the portfolio variance plus the expected value squared). In this case, the expected utility has not the standard form usually used in the mean-variance optimization. Despite so, it still remain a function of the variance, therefore, for any given expected return an investor with quadratic utility wishes to minimize variance and hence chooses a mean-variance efficient portfolio. For the former, the reason lies in the fact that the Gaussian distribution has only the first two moments different from zero. Therefore, by doing a Taylor's expansion of the exponential utility, it will be naturally truncated at the second order. As reported by Markowitz (2014) himself, there is some confusion in the literature about this. It is often thought that mean-variance analysis is only valid in the two cases listed above. But this is not the case. Having Gaussian return distributions or quadratic utilities is only a sufficient condition for the use of mean-variance, but not necessary. This means that even if we are not in those two cases, the mean-variance is still valid. Indeed, Markowitz (1959) observes that if the utility function can be approximated by its Taylor's expansion, truncated at second order (which turns out to be a quadratic function), then, taking its expected value, we get a function that depends only on the mean value and the variance. In Markowitz (2014) the author tries to answer the question "Why not just maximize the expected utility?" (i.e avoiding any kind of approximation?). The answer lies in the fact that the mean-variance analysis is a practical way to approximately maximize the EU since:

1. It has a lower computational cost;

2. It does not impose to specify further parameters of the probability distributions of returns, apart from the average value and the variance (in case the utility function is known and the Taylor expansion is used to approximate it);

3. It does not require the investor to specify the functional form of his utility. Choosing a mean-variance efficient portfolio the decision-maker will select a portfolio with a maximum or the nearly maximum expected utility (Levy and Markowitz (1979)).

In Kroll et al. (1984) the author presents either the classes in which the approximation works well or some improvements to the theory. It says that the approximation works in case of a log utility (the one proposed by Bernoulli (1954)) or a power (Cramer). Besides, in the same article a study about a higher-order moment utility expansion done by Errington is presented.

3.4 Stochastic Dominance as a Criterion for Efficiency

This section illustrates how stochastic dominance can be exploited to find an efficient portfolio. We start by reporting the concept of dominance from Hanoch and Levy (1969). Then, several condition (both necessary and sufficient) for efficiency are listed.

We take two random variables X and Y, with cumulative distribution F and G, respectively. One can say that X dominates Y (or alternatively that F dominates G) if the expected utility of X is greater than the expected utility of Y, for every utility, as long as it is increasing. The key point linking dominance and efficiency is that "every sufficient condition for dominance is an efficient criterion" (Hanoch and Levy (1969)). This implies that any investor who is willing to maximize her portfolio, will choose a dominant investment. For the mean-variance analysis, the variable having larger mean and smaller variance will be always chosen by risk averse investors. However, Markowitz's criterion is neither sufficient nor necessary for dominance, while we are maximizing the Expected Utility (EU). The only two counterexamples, showing that the mean-var criterion is a sufficient condition (even if not necessary) for dominance, are the ones mentioned in the previous section and presented in Hanoch and Levy (1969):

- Normal returns;
- Quadratic utility;

It is suggested that stochastic dominance rules do not replace the mean-variance rule but offer an different approach, complimenting rather than substituting it. In the following we report some theorems defining the efficiency criterion for dominance, under different risk attitudes.

Theorem 2 (First order Stochastic Dominance)**.** *Let F and G two cumulative distributions and U(x) a non decreasing function. A necessary and sufficient condition for FSD of F over G is:*
$F(x) \leq G(x)$, *for every* x *and* $F(x_0) < G(x_0)$ *for some* x_0.

Hence, if the two cumulative distributions intersect, the above statement will not be met. From the condition of F always less than G follows $E_F X > E_G Y$, which is a necessary condition for dominance for any variance and any risk aversion level. The only constraint here is to have a non-decreasing utility function. In Levy (2015) "an *optimal* decision rule is defined as a decision rule, which is necessary and sufficient for dominance". The FSD rule is the optimal rule for U, since it is a sufficient and necessary condition for dominance. Thus, considering the subset of increasing utility function an optimal rule is defined as: $E_F U(X) > E_G U(X) \Longleftrightarrow FDG$. For the second stochastic ordering, the applicability is restricted to the utility functions which are concave, i.e. those that refer to a risk averter.

Theorem 3 (Second order Stochastic Dominance)**.** *Let F and G two cumulative distributions, which differ at least in one point. A necessary and sufficient condition for SSD, for every U(x) non decreasing and concave is:*
$\int_{-\infty}^{x} [G(t) - F(t)]dt \geq 0$

Proof. Both theorems are proved in Hanoch and Levy (1969). □

It means that the two functions F and G can intersect more than once; the important thing is that the area subtended by F up to x is always smaller than the area subtended by G. Like before, $E_F X > E_G Y$ is a necessary condition for dominance of X over Y, and, again, there are no restrictions for variance. The variance of X does not to smaller than the variance of Y, for any concave function. A particular case of the Theorem 3 is when there is only one point of intersection. In this case, in fact, a necessary and sufficient criterion to have stochastic dominance of F over G, is having $F < G$ below the intersection point, given $E_F X \geq E_G Y$. The only case in which the variance is a necessary condition for dominance, even though not sufficient, is when the two random distributions have equal means, i.e. $E_F(X) = E_G(X)$. In this case to say that "F dominates G at second order" one needs: $\sigma_F < \sigma_G$. Moreover, there exists also higher order stochastic dominance and special rule for all risk seeker investors. The rule defining second order stochastic dominance, only in the case of risk seeking, is given below.

Theorem 4 (Risk seeking Stochastic Dominance). *Let F and G be two investments whose density functions are f(x) and g(x), respectively. Then F dominates G by $\overline{SSD}$, i.e. at second order and for all risk seekers, if and only if*

$\int_b^x [G(t) - F(t)]dt \geq 0$, for all $x \in [a, b]$ and there is at least one x_0 for which there is a strict inequality.

Proof. See pages 217-218 of Levy (2015) □

One might erroneously think that if F dominates G by SSD then G dominates F by $\overline{\text{SSD}}$. This is incorrect, and the proof can be found in Levy (2015). To summarise, the theorems presented so far ensure that whatever the investor is (risk-averse or risk-loving), if the sufficient conditions are met, she will choose the dominant investment.

Sufficient conditions	
FSD	$F(x) \leq G(x)$ $minF(x) \geq maxG(x)$
SSD	$\int_a^x [G(t) - F(t)]dt \geq 0$ FSD
$\overline{\textbf{SSD}}$	$\int_b^x [G(t) - F(t)]dt \geq 0$

Table 3.1: **Sufficient Conditions** for Dominance

3.4.0.1 Which Investment is the Riskiest?

In this last paragraph we deal with a common question in quantitative risk management: "How can one state that one investment is riskier than an other?". In Rothschild and Stiglitz (1970) the authors made a list of answers to this question. They are all based on the assumption that X and Y have the same expectation value. According to the authors all the following definition are equivalent in stating that Y is riskier than X.

1. if Y is equal to X plus noise;
2. if the investor is risk averse and $E[U(X)] \geq E[U(Y)]$;
3. if Y has heavier tails than X.
   This is called the "fat tail criterion". It is based on the definition of the mean preserving spread (MPS), which is a step function with the aim to shift the density from the center to the tails, increasing its risk. If the difference between the pdf (or pmf) of Y and X is a MPS, than Y the riskiest variable.

All these three approaches result in the same answer to the question of who has a higher risk, and also define a partial order between the random variables. The fourth approach is related to variance. In fact, oftentimes in literature, it is found that Y is riskier than X if the variance of X is less than that of Y. This is true, even if not for all the risk averter, if X and Y have the same mean value. However, the order given to the variables based on the criterion of the variance is a total order and not partial as before. This means that if $E(X) = E(Y)$ the variables can always be ordered according to their variance. Nevertheless, it is important to repeat that even if X has a smaller variance than Y, this does not mean that its expected utility is greater than Y. This may happen for some investors, but not for all. As reported before, in fact, in the case of equal mean value, the condition on the variance is necessary, but not sufficient to the dominance and therefore will not define an optimal rule.

Chapter 4

Correlation Bounds

The purpose of this chapter is to characterize the minimum and maximum correlation coefficients between two random variables, given their marginal distribution. In particular, we focus on the case of discrete random variables. Applying the Copula function, in fact, the correlation coefficient between the Gaussian test variables is different with respect to the Bernoulli ones of our interest. The relevance of the following analysis emerges precisely in investigating and understanding this implicit link. Leveraging on the studies of Hoeffding and Fréchèt, we contribute to characterize analytically the case of two Bernoulli random variables. For the said case, either the bound for the linear correlation coefficient (Pearson) or the rank one (Spearman) are outlined. The analytical findings are then checked numerically and compared to some results in the literature.

4.1 Upper and Lower Bounds for a Two Variable Distribution

Based on Whitt (1976), we consider $H = (F, G)$ to be the set of all the cumulative distribution functions on $\mathbb{R}^2$, having F and G as marginal cumulative distribution function (cdf) and finite variance. It is straightforward H^* and H_*, that are the distribution functions having maximum and minimum correlation, respectively belong to the set H.

Theorem 5 (Hoeffding)**.** *In* Π *there exists for all* $(x, y) \in \mathbb{R}^2$

$H^*(x,y)=min[F(x),G(y)]$ *and* $H_*(x,y)=max[0, F(x)+G(y)-1]$.

Theorem 6. *For any marginal cumulative distributions F and G with positive variances, the pair of random variables*

$[F^{-1}(U), G^{-1}(U)]$ *has cdf* H^* *whereas the pair* $[F^{-1}(U), G^{-1}(1-U)]$ *has cdf* H_*

Proof. Both theorems are proved in two different ways at pages 1282-1283 of Whitt (1976). $\square$

Thanks to these results, one can compute the upper and lower bounds of the correlation raising between two random variables we call X_1 and X_2. Indeed, given X_1 and X_2 two random variables with finite variance and cdf F_1 and F_2 respectively, we can find the correlation bounds simply by using the marginal distributions, and thus avoiding dealing with the original variables. Hereafter, U is a uniform random variable and F_i^{-1} is the inverse of the cdf $F_i, \forall i = 1, 2$. The upper and lower bounds are thus defined

$$\overline{\rho} := \max_{X_1, X_2} corr(X_1, X_2) = corr(Y_1, Y_2), \tag{4.1}$$

where $Y_1 = F_1^{-1}(U)$, $Y_2 = F_2^{-1}(U)$,

$$\underline{\rho} := \min_{X_1, X_2} corr(X_1, X_2) = corr(Y_1, \tilde{Y_2}), \tag{4.2}$$

where $\tilde{Y_2} = F_2^{-1}(1 - U)$.

In the following, relying on the discussion of Leonov and Qaqish (2017), we provide some analytical findings for Pearson's correlation among bivariate discrete distributions. In particular we start with the most general case and then we focus on the Bernoulli-Bernoulli one.

4.1.1 Discrete Bivariate Case

Let consider X_1 which takes values in $A = \{a_1, \dots, a_m\}$ e X_2 which, on the other hand, takes values in $B = \{b_1, \dots, b_n\}$. A and B are sorted. Now, given the marginal $\alpha = \mathbf{P}(X_1 = a_i)$ and $\beta = \mathbf{P}(X_2 = b_j)$, we calculate the joint that maximizes the $corr(X_1, X_2)$. We create two sets of values S_1 and S_2. $S_1 = \{0\} \cup \{F_{X_1}(a_i), i = 1, \dots, m\}$ defines a partition of the interval [0,1] in m-1 subsets $[0, f_1), (f_1, f_2], \cdots (f_{m-1}, 1]$, where $f_1 = \alpha_1$ and $f_i = f_{i-1} + \alpha_i$. On each of these sub-intervals, the function $F_{X_1}^{-1}$ is constant and takes the value a_i in the i-th interval. Similarly, $S_2 = \{0\} \cup \{F_{X_2}(b_j), j = 1, \dots, n\}$, splits the interval [0, 1] in n-1 subsets $[0, g_1], (g_1, g_2], \cdots (g_{j-1}, 1]$, where $g_1 = \beta_1$, and $g_j = g_{n-1} + \beta_j$. Again, across any of these sub-intervals, the function $F_{X_2}^{-1}$ is constant and assumes the value b_j in the j-th interval.

Then, considering $S = S_1 \cup S_2$, the range [0,1] is divided into K intervals, with $max(m, n) \leq K \leq m + n - 1$, which represents the different K combinations of the events' realizations X_1 and X_2. Each interval I_k thus corresponds to the event $\{X_1 = a_i, X_2 = b_j\}$ whose probability is simply the length of the interval itself. At this point, to get $\overline{\rho}$ you only need to calculate $E[X_1, X_2]$, which will be:

$$E[X_1, X_2] = \int_0^1 F_1^{-1}(u) F_2^{-1}(u) du \tag{4.3}$$

4.1.2 Bernoulli Distribution

We now focus on the case in which $X_1 = Bern(p_1)$ e $X_2 = Bern(p_2)$. We wish to derive the relationship existing between the marginals and the correlation, in order to calculate, given certain p_1 and p_2, the maximum range in which ρ can take value. Referring to Equ.4.1 and 4.2 we define

$$Y_i = Bern^{-1}(u) = \mathbf{1}(U > 1 - p_i) = \mathbf{1}(U > q_i) \quad \forall i \in \{1, 2\} \tag{4.4}$$

and

$$\tilde{Y_2} = Bern^{-1}(1 - u) = \mathbf{1}(1 - U > 1 - p_2) = \mathbf{1}(U < p_2) \tag{4.5}$$

As the linear correlation coefficient between two variables, for example considering Y_1 and Y_2, is defined as:

$$\rho = \frac{E[Y_1 Y_2] - E[Y_1]E[Y_2]}{\sqrt{var[Y_1]var[Y_2]}} \tag{4.6}$$

the only element to know is $E[Y_1 Y_2]$. Indeed, the first and the second moment of the distribution are known, being Y_i a Bernoulli. Then, exploiting Equ. 4.3 we obtain

$$\begin{aligned} E[Y_1 Y_2] = \int_{\max_i q_i}^{1} du &= 1 - \max_{i=1,2} q_i \\ &= 1 + \min_{i=1,2}(-q_i) \\ &= \min_{i=1,2}(1 - q_i) \\ &= \min_{i=1,2} p_i \end{aligned} \tag{4.7}$$

Therefore from Eq. 4.1 we compute the maximum correlation:

$$\begin{aligned} \overline{\rho} &= \frac{\min p_i - p_1 p_2}{\sqrt{p_1(1 - p_1) \cdot p_2(1 - p_2)}} \\ &= \sqrt{\frac{(\min p_i)^2 (1 - \max p_i)^2}{p_1(1 - p_1) \cdot p_2(1 - p_2)}} \\ &= \sqrt{\frac{(\min p_i)^2 (1 - \max p_i)^2}{\min p_i \cdot (1 - \min p_i) \cdot \max p_i \cdot (1 - \max p_i)}} \\ &= \sqrt{\frac{\min p_i}{\max p_i} \cdot \frac{1 - \max p_i}{1 - \min p_i}} \end{aligned} \tag{4.8}$$

While, through Eq. 4.2, we get the minimum ρ. In this case we need to consider two different situations depending on whether the intersection between Y_1 and $\tilde{Y_2}$ is empty or not.

1. If $p_2 < q_1$, then the overlap region is zero, as well as $E[Y_1\tilde{Y_2}] = 0$. So:

$$\begin{aligned} \underline{\rho} &= \frac{0 - p_1 p_2}{\sqrt{p_1(1-p_1) \cdot p_2(1-p_2)}} \\ &= -\sqrt{\frac{(p_1 p_2)^2}{p_(1-p_1) \cdot p_2(1-p_2)}} \\ &= -\sqrt{\frac{p_1}{1-p_1} \cdot \frac{p_2}{1-p_2}} \end{aligned} \tag{4.9}$$

2. If $p_2 > q_1$, then the region will be $p_2 - q_1$ long and therefore $E[Y_1\tilde{Y_2}] = p_2 - q_1$. Then:

$$\begin{aligned} \underline{\rho} &= \frac{p_2 - (1-p_1) - p_1 p_2}{\sqrt{p_1(1-p_1) \cdot p_2(1-p_2)}} \\ &= -\sqrt{\frac{(p_2(1-p_1) - 1(1-p_1)}{p_(1-p_1) \cdot p_2(1-p_2)}} \\ &= -\sqrt{\frac{1-p_1}{p_1} \cdot \frac{1-p_2}{p_2}} \end{aligned} \tag{4.10}$$

So finally, we discover that in the case of two Bernoulli random variables, the Pearson coefficient varies between -1 and +1 if and only if the two variable X_1 and X_2 are of the same *type*. Being of the same type, as explained in Haught (2016), means that one is function of the other: $X_1 = aX_2 + b$ with $a > 0, b \in \mathbb{R}$. Adapting this concept to the case of two binary variables results in the need to have same probability, but not necessary same domain. Notice that, $\underline{\rho} =$ -1 is attained if and only if X_1 and -X_2 are of the same type; whereas $\overline{\rho} = 1$ is obtained when X_1 and X_2 are of the same type.

4.1.2.1 Exercise: Numerical Checking of the bounds

With the following exercise, we pursue to check by a numerical simulation the validity of the results already achieved. To achieve our purpose, we consider again a horse racing between two horses A and B. The upper and lower bounds for the correlation coefficient between the Bernoulli are computed numerically while the probabilities $p_1 = p_{A_{wins}}$ and $p_2 = p_{B_{wins}}$ change. As one can see in Figure 4.1 the curves, concerning either $\overline{\rho}$ or $\underline{\rho}$, are symmetrical to the point $(0.5, 0)$. We notice that, when the variables A and B have equal distributions, i.e. equal probability of winning, the maximum correlation coefficient is 1. This result is coherent with the definition of "same type" given some lines above. Similarly, every time the probability of one to wins coincides with the probability of the other to loose, we reach $\underline{\rho} = -1$. In all the other intermediate cases the upper and lower bounds are calculated by Equ.4.8 and Equ.4.9 or 4.10, respectively. In the next we focus for a moment to the case of horses

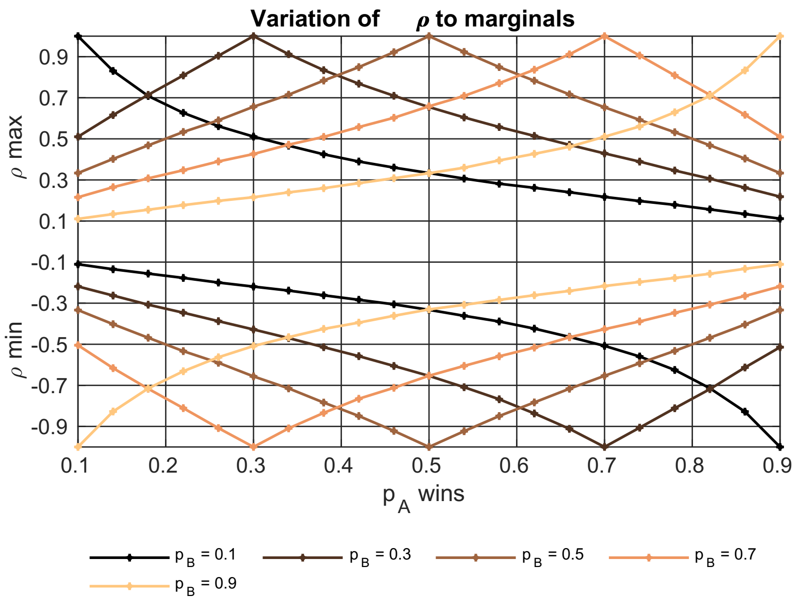

Figure 4.1: **Bernoulli upper and lower bounds**. Every point in the graph corresponds to a pair $(p_{A_{wins}}, p_{B_{wins}})$. Every line shows how $\overline{\rho}$ ($\underline{\rho}$) evolves in the upper (lower) plane, varying $p_{A_{wins}}$, but a $p_{B_{wins}}$ fixed.

with the same chance of winning, $p = 0.9$ and the same payoff $\frac{1}{p}$. The theory says that the maximum possible correlation in this case (calculating through Equ. 4.8) turns out to be 1. This makes sense because the two variables are of the same type and they may be perfectly positively related. To obtain the minimum coefficient we use the Equ. 4.10 since, being $p_1 + p_2 = 0.9 + 0.9 > 1$, we are in case 2, and we get $\underline{\rho} = -0.11$. This value can be justified because the "winning" event is much more likely than the opposite for both horses. Indeed, to see $\underline{\rho} = -1$ here, we should have a situation where every time the A horse wins, B looses and vice versa. Here, however, but the probability of obtaining simultaneously a winning of A and a losing of B is very low, since the "loss" event has a probability of happening only 0.1.

To sum up, this easy exercise ensured us that there exists a match between how we employed the copula function to simulate the behaviour of two correlated Bernoulli variables.

4.2 From $\rho_{Gaussian}$ To $\rho_{Bernoulli}$

As anticipated in the setting of the Horse-racing experiment, we make now some considerations about how the correlation is transferred from Gaussian to Bernoulli. Hence, starting by looking at the relationship between the two linear correlation coefficients, we pursue to understand how the use of the Gaussian copula affects the final result. We first investigate the Formula 4.6, noticing that $\rho_{Gaussian}$ enters only in the calculus of expected value of the product of variables. We then define $Y_1 = Bern^{-1}[\Phi(Z_1)] = \mathbf{1}[\Phi(Z_1) < p_1]$, where U of Equ. 4.6 corresponds intentionally with $\Psi(Z_1)$, i.e. the Gaussian marginal distribution of Z_1. The same is repeated for Y_2. To compute $E[Y_1 Y_2]$ one has to solve the integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{1}[\Phi(z_1) < p_1]\mathbf{1}[\Phi(z_2) < p_2]\phi(z_1, z_2, \rho_G)dz_1 dz_2 \tag{4.11}$$

where $\phi(z_1, z_2, \rho_G)$ is the bivariate gaussian distribution with correlation ρ_G. Being a function of ρ_G, the Equ. 4.11 provides the relation between the two coefficients and therefore the shape on the curve in Figure 4.2. In this figure the two linear correlation coefficients, $\rho_{Gaussian}$ and $\rho_{Bernoulli}$, are compared. The yellow curve in the picture derives from the integration, while the purple one is found by exploiting the Copula function. In the simulation we compute $\rho_{Bernoulli}$ while $\rho_{Gaussian}$ is varied between [-1,+1]. As expected (and discussed in the previous section), $\rho_{Bernoulli}$ is limited above and below by the $\overline{\rho}$ and $\underline{\rho}$ calculated above. The picture show the case of two identical binary variables with probability to succeed $p = 0.9$, however the increasing trend is maintained the same even for other values of p. In the displayed case, the relationship between Bernoulli and Gaussian coefficient appears exponential-like. It means that fitting the curve we can find the functional form that better approximate describes the curve. Although, we decided not to pursue this way since we discovered that the functional form of $\rho_{Bernoulli}$ changes according to the probability the two random variable.

4.3 Spearman's Rho Bounds for Discrete Distributions

In the case of the rank correlation coefficient Spearman's rho, the limits between X_1 and X_2 can be deduced thanks to the relation existing between $\rho_{Pearson}$ and $\rho_{Spearman}$. As shown by Equ. 3.1, the Spearman correlation coefficient arising between two random variables is equal to the linear correlation existing between the marginal $F_1(X_1)$ and $F_2(X_2)$ of the same RVs. Additionally, if the two distributions are continuous, the limits of Spearman's correlation are always +1 and -1. This is because the marginal cumulative function of any continuous random variable is a uniform variable distributed over the range [0,1]. So, whatever the distribution of X, the linear correlation coefficient

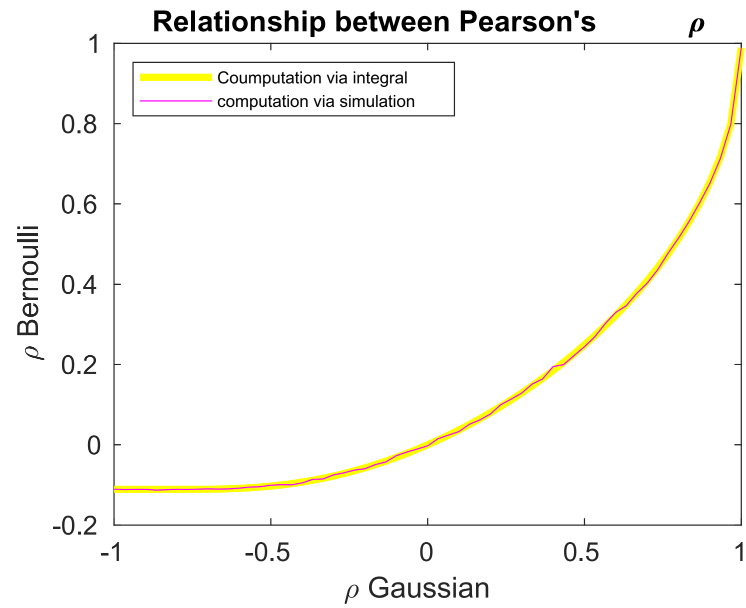

Figure 4.2: **Relationship** existing between the two linear correlation coefficients: ρ **Gaussian and** ρ **Bernoulli**. Here ρ_G is taken to assume 61 different values in the range [-1,+1].

between two marginal functions corresponds to the linear correlation coefficient between two uniforms U_1 and U_2. Since both U_1 and U_2 and U_1 and $-U_2$ can be called of the same type, the range between which ρ_S can move coincides with the maximum, always -1, 1. This is one of the advantages that explain why Spearman's coefficient is preferable to Pearson's. Moreover, it depends only on the family of the chosen copulation function and it is independent under monotonic transformations. However, for two discrete distributions, this is no longer true because the marginal cumulative distribution is discontinuous. If the two variables under examination have the same cdf, the upper limit is 1 taking $X_1 = X_2$. Otherwise, the general considerations presented in subsection 4.1.1 can be used to compute $E[F_1(X_1)F_2(X_2)]$ and, thus, the upper and lower bounds for ρ_S. In the following, we are going to show those bounds for two Bernoulli random variables. Knowing the cdf for the Bernoulli, we compute the probabilities of the following three events:

1. $F(X_i) = 0 \rightarrow \mathbf{P}(F(X_i) = 0) = 0$;

2. $F(X_i) = q_i \rightarrow \mathbf{P}(F(X_i) = q_i) = q_i$;

3. $F(X_i) = 1 \rightarrow \mathbf{P}(F(X_i) = 1) = 1 - q_i = p_i$;

The result is still a binary variable which, with probability p_i takes value 1, while with probability q_i takes value q_i. To compute the maximum joint cdf we follow the procedure presented in subsection 4.1.1. The result is shown in figure 4.3. The three intervals correspond to: $I_1 = \min_i q_i, I_2 =$

$|q_1 - q_2|, I_3 = \min_i p_i$. Therefore applying the method of finding the overlapped region we get

$$E[F_1(X_1), F_2(X_2)] = q_1 q_2 \cdot I_1 + \max q_i \cdot I_2 + 1 \cdot I_3 \tag{4.12}$$

Then, using the general definition for the linear correlation coefficient (e.g. Equ.4.6), the upper bound for ρ_S is found. Analogously, the computation of the lower bound is straightforward.

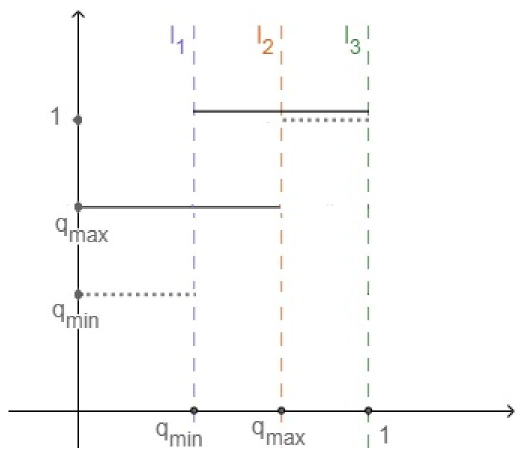

Figure 4.3: **Graphical way to compute** $E[X_1 X_2]$. Partitioning of the [0, 1] interval used to implement Equ. 4.3 in subsection 4.1.1. The pmf of X_1 is $\alpha = (q_{min}, p_{max})$ whereas of X_2 is $\beta = (q_{max}, 1 - p_{min})$. Each of the three subintervals $I_k, k = 1, 2, 3$ corresponds to unique (i, j) pair and event $\{X_1 = a_i, X_2 = b_j\}$, and together for the maximum joint pmf.

Chapter 5

From Two to N Bonds - Zero Correlation

In this chapter we move to a portfolio of n bonds, which are independent. We consider two investors owning each a portfolio of n independent and identically distributed bonds, modeled with n Bernoulli variables. As the portfolio grows (passing from 2 of the previous sections to n) the investors' concern is a fortiori finding the dominant portfolio, in order to suffer the least loss possible. We contribute to analytically respond to the investors' concern about choosing the dominant portfolio, simply by knowing the probability of win/fail of individual bonds. The default probability is p_1 in the case of portfolio 1, while it turns out to be p_2 if they belong to portfolio 2. Since we are dealing with a sum of n i.i.d. Bernoulli random variables (as zero correlation is present), the distribution of the portfolio is a Binomial. The analytical solutions proposed are two: first through a formal discussion about first-order stochastic dominance, then leveraging on the coupling method.

5.1 First Order Stochastic Dominance

Being $B(n, p_2)$ and $B(n, p_1)$ the Binomial distribution modeling the two portfolio, we aim at proving that $B(n, p_2)$ dominates $B(n, p_1)$ iff $p_2 > p_1$. We begin by recalling the definition

Definition 2 (First order Stochastic Dominance). Let F and G two cumulative distributions and U(x) a non decreasing utility function. A necessary and sufficient condition for which F first-order stochastically dominates G is: $F(x) \leq G(x)$, for every x and $F(x_0) < G(x_0)$ for some x_0.

In the following we call G the cumulative function of X ($X \sim G_X = B(n, p_1)$) and F the cumulative function of Y ($Y \sim F_Y = B(n, p_2)$). Then, to demonstrate that the sufficient condition for dominance is met if and only if $p_2 > p_1$, we study the monotony of the binomial cumulative function with respect

to p. If the distribution function is decreasing or not increasing with p, we will have the dominance of F over G.
In fact, if we look at the Figure 5.1, one can see that the green curve, let's call it F, dominates the blue curve, G because “it is underneath”, verifying the condition of the Definition 2. The dominant distribution, as expected, has higher probability: $p_{green} > p_{blue}$. Let us now consider the cumulative

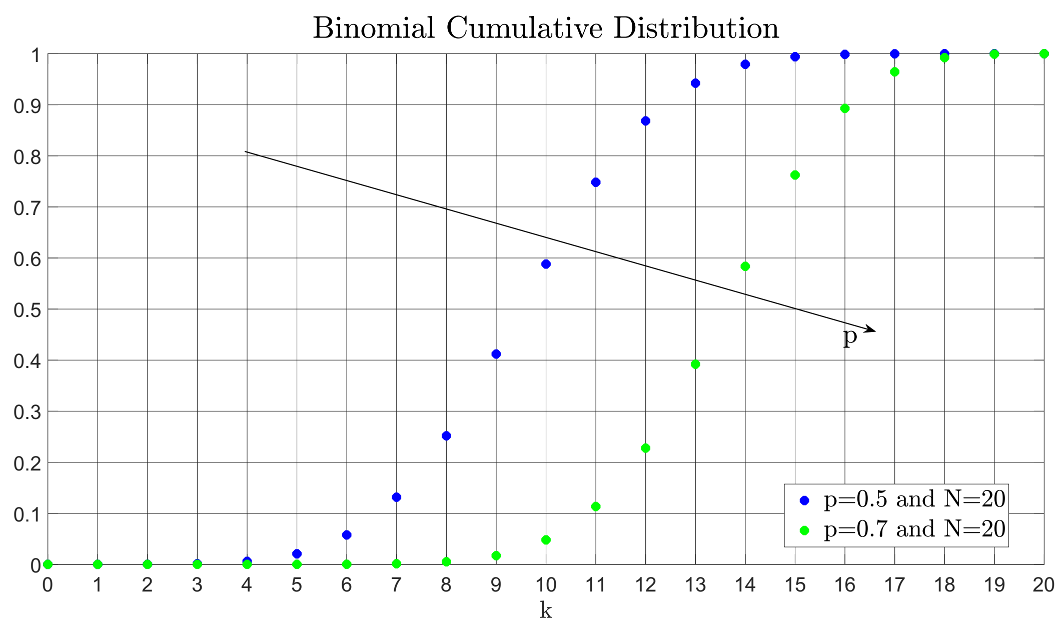

Figure 5.1: **Binomial cumulative distributions** varying p

function of a binomial variable:

$$F(k;n,p) = \sum_{i=n-k}^{n} \frac{n!}{i!(n-i)!}(1-p)^i \cdot p^{n-i} \tag{5.1}$$

By calculating the derivative with respect to p you get its point-wise variation with respect to p:

$$\begin{aligned}
\frac{\partial F(n,k,p)}{\partial p} &= \sum_{i=n-k}^{n} \frac{n!}{i!(n-i)!}(1-p)^{i-1} \cdot p^{n-i-1} \cdot (-ip + (n-i) - np + ip) \\
&= \sum_{t=0}^{k} \frac{n!}{(n-t)!t!}(1-p)^{n-t-1} \cdot p^{t-1} \cdot (t-np) \qquad [t = n-i] \\
&= \sum_{t=0}^{\lfloor np \rfloor} \frac{n!}{(n-t)!t!}(1-p)^{n-t-1} \cdot p^{t-1} \cdot (t-np) \\
&+ \sum_{t=\lceil np \rceil}^{k} \frac{n!}{(n-t)!t!}(1-p)^{n-t-1} \cdot p^{t-1} \cdot (t-np) \\
&= A + B.
\end{aligned} \tag{5.2}$$

As long as $k \leq np$, where np corresponds to the mean value of the distribution, the sum is negative. When $k > np$ instead, we can break the derivative in two terms, one ranging from 0 to $\lfloor np \rfloor$ (let's call it A), the other ranging from $\lceil np \rceil$ to k (B). The limit case occurs when $k = n$. A is a sum, as mentioned above, of terms that are all negative, while B is a sum of positive terms. Therefore, the derivative does not change sign (becoming positive) until B is lower in absolute value than A. To prove that the monotony remains constant, we take the limit case. In fact, if even in the limit case the derivative is not positive, then, *a fortiori*, it will not be positive for any values $np < k < n$.

$$\begin{aligned}
A &= -\frac{n!(1-p)^{n-\lfloor np \rfloor} p^{\lfloor np \rfloor}(\lfloor np \rfloor + 1)}{(1-p)(n-(\lfloor np \rfloor + 1))!(\lfloor np \rfloor + 1)!} \\
&= -\frac{n!(1-p)^{n-\lfloor np \rfloor - 1} p^{\lfloor np \rfloor}(\lfloor np \rfloor + 1)}{(n-(\lfloor np \rfloor + 1))!(\lfloor np \rfloor + 1)!} \\
&= -\frac{\Gamma(n+1)(1-p)^{n-\lfloor np \rfloor - 1} p^{\lfloor np \rfloor}}{\Gamma(n-\lfloor np \rfloor) \cdot \Gamma(\lfloor np \rfloor + 1)} \\
&= -\frac{\Gamma(n+1)(1-p)^{n-\lceil np \rceil} p^{\lceil np \rceil - 1}}{\Gamma(n-\lceil np \rceil + 1) \cdot \Gamma(\lceil np \rceil)} \qquad (\lfloor np \rfloor + 1 = \lceil np \rceil) \\
B &= \frac{n!(1-p)^{n-\lceil np \rceil} p^{\lceil np \rceil - 1} \lceil np \rceil}{(n-\lceil np \rceil)! \lceil np \rceil !} \\
&= \frac{\Gamma(n+1)(1-p)^{n-\lceil np \rceil} p^{\lceil np \rceil - 1}}{\Gamma(n-\lceil np \rceil + 1) \cdot \Gamma(\lceil np \rceil)}
\end{aligned}$$

An easier way to further check the result, showing that A and B cancel out in the limit case of $k = n$, is presented in the following. For $k = n$, Equ.5.1 coincides with the normalization condition proper of every pdf. Indeed, the probability density function is non negative everywhere, and its integral over the entire space is equal to 1. Therefore, Equ.5.1 results to be a constant, precisely equal to 1, that derived with respect to p, as in Equ.5.2, gives 0, as aimed to show.

5.2 Coupling Method

Another way to prove the claim is through coupling. Coupling is a probabilistic technique with a wide range of applications. In probability theory, coupling is a proof technique that allows one to compare two unrelated variables by "forcing" them to be related in some way. The idea behind the coupling method is that, to compare two unrelated probability measures, it is sometimes more useful to construct a "ad-hoc" joint probability space with same marginal of the original space.
Let's consider now our two investors with binomial probability distribution $B(n, p_1)$ and $B(n, p_2)$. The binomial distribution describes the number of hits in a Bernoulli process, i.e. the random variable $S_n = X_1 + X_2 + \cdots + X_n$ which sums n independent Bernoulli random variables $Ber(p)$. We can represent the investors as binomial since we are assuming they own a portfolio with n bond, which can default with probability $1 - p_1$ and $1 - p_2$, for *Investor 1* and *2* respectively. Moreover, every bond is a Bernoulli variable and it contributes in equal weight to forming the portfolio, since the same amount of money is invested in each bond. Hence, if both investors have the same number of bonds in the portfolio, the investor with the highest p, i.e. the probability that a bond does not default (in our case the highest is p_2) will succeed more often than the other. More specifically, this means that, for any fixed k (that is the number of bonds which have not default), the probability that *Investor 1* produces at least k successes should be less than the probability that *Investor 2* produces at least k successes. Coupling simplifies this problem.

- *Investor 1 trial*:

$$X_i = \begin{cases} 1, & p_1 \\ 0, & 1 - p_1. \end{cases}$$

- *Investor 2 trial*:

$$Y_i = \begin{cases} 1, & p_2 \\ 0, & 1 - p_2. \end{cases}$$

where $0 < p_1 \leq p_2 < 1$ without loss of generality. In the new probability space defined by the coupling, the sequence of X_i remains unchanged, so $X_i = X'_i$, while for the second investor, I define a new binary sequence Y'_i such that:

- if $X_i = 1 \implies Y'_i = 1, \ w.p. \ 1$;

- if $X_i = 0 \implies Y_i' = 1,\ w.p.\ \frac{p_2 - p_1}{1 - p_1}$.

One of the property of the coupling is that the marginal probability distribution of the new Y_i' is equal to the original Y_i one. This is checked in the following, where we will neglect the subscript i.

$$\begin{aligned} P(Y' = 1) &= \sum_x P(X = x, Y' = 1) \\ &= P(Y' = 1|X = 1) \cdot P(X = 1) + P(Y' = 1|X = 0) \cdot P(X = 0) \\ &= 1 \cdot p_1 + \frac{p_2 - p_1}{1 - p_1} \cdot (1 - p_1) \\ &= p_2. \end{aligned} \tag{5.3}$$

As a result, using the chain rule for probability, the joint and the marginal probability become:

	$X = 1$	$X = 0$	
$Y' = 1$	$1 \cdot p_1$	$\frac{p_2 - p_1}{1 - p_1} \cdot (1 - p_1)$	p_2
$Y' = 0$	0	$1 - p_2$	$1 - p_2$
	p_1	$1 - p_1$	

Now, since we are dealing with Bernoulli rv, we can list all the combinations of X and Y', where $Y' \geq X$. It can be seen that the sum of the probabilities of all those events returns 1.

$$P(Y_i' \geq X_i) = 1 \qquad i.e. \quad Y_i' \geq X_i \quad a.s. \tag{5.4}$$

Indeed, the only combination with $X > Y'$ is when $X = 1$ and $Y' = 1$, but it has probability 0, as shown in Table 5.2. This makes sense also intuitively because for every bet where *Investor 1* gets 1, *Investor 2* always wins, but even when *Investor 1* loses, *Investor 2* will still have a chance of winning which depends on the difference between p_2 and p_1.

In the new probability space defined by the coupling, X and Y' are coupled $\forall i$, so they can be compared on every game. From Equ.5.4 we have:

$$P(Y_i' - X_i \geq 0) = P(\alpha_i \geq 0) = 1 \tag{5.5}$$

Then we define: $S_n' := \sum_{i=0}^n Y_i'$, $S_n := \sum_{i=0}^n Y_i$ and $T_n := \sum_{i=0}^n X_i = \sum_{i=0}^n X_i'$. Since there exists a coupling between X_i and Y_i' for all i, as a result, considering both X_i and Y_i' to be i.i.d., there is also a coupling among their sums. In the new probability space defined by the S_n' and T_n variables, the marginal probability distributions of S_n and S_n' coincide. To show it we exploit the generating function tool

$$
\begin{aligned}
G_{S'_n}(z) &= E[z^{S'_n}] = E[z^{\sum_i^n Y'_i}] && (5.6)\\
&= \prod_{i=1}^{n} E[z^{Y'_i}] = \prod_{i=1}^{n} G_{Y'_i}(z)\\
&\overset{\text{i.i.d}}{=} \Big(E\big[z^{Y'}\big]\Big)^n = (1 - p_2 + z \cdot p_2)^n\\
&= \sum_{k=0}^{n} \binom{n}{k} p_2^k (1-p_2)^{n-k} z^k = \sum_{k=0}^{n} f(k; n, p_2) z^k\\
G_{S_n}(z) &= E[z^{S_n}] = E[z^{\sum_i^n Y_i}] && (5.7)\\
&= \prod_{i=1}^{n} E[z^{Y_i}] = \prod_{i=1}^{n} G_{Y_i}(z)\\
&\overset{\text{i.i.d}}{=} \Big(E\big[z^{Y}\big]\Big)^n = (1 - p_2 + z \cdot p_2)^n
\end{aligned}
$$

After showing the relationship between S'_n and T_n and the fact that the marginal ones remained unchanged, now, we wish to show if such a coupling has the property:

$$
P\big(S'_n \geq T_n\big) = 1 \tag{5.8}
$$

Equ. 5.8 is equivalent to:

$$
P\Bigg(\sum_{i=0}^{n} Y'_i \geq \sum_{i=0}^{n} X_i\Bigg) = 1 \tag{5.9}
$$

$$
P\Bigg(\sum_{i=0}^{n} (Y'_i - X_i) \geq 0\Bigg) = 1 \tag{5.10}
$$

$$
P\Bigg(\sum_{i=0}^{n} \alpha_i \geq 0\Bigg) = 1 \tag{5.11}
$$

With the last equality we are seeking the probability of the event "sum of positive elements greater or equal than zero". Intuitively the probability of that event is one because of Equ. 5.5, which says that the difference between Y' and X is always positive. $\alpha_i \geq 0 \quad \forall i \longrightarrow \sum_{i=0}^{n} \alpha_i \geq 0$. In the following, it is shown in a more formal way. Indeed, one can see that

$$
P\Bigg(\sum_{i} \alpha_i < 0\Bigg) \leq P\big(\text{at least one } \alpha_i < 0\big) = 1 - P\big(\text{all } \alpha_i \geq 0\big) \tag{5.12}
$$

Since the α_i are assumed to be independent random variables, seeking the probability that all $\alpha_i \geq 0$ means:

$$P(\text{all } \alpha_i \geq 0) = P(\alpha_1 \geq 0, \cdots, \alpha_n \geq 0) = \prod_{i=1}^{n} P(\alpha_i \geq 0) = 1^n = 1 \tag{5.13}$$

$$\implies P\left(\sum_i \alpha_i < 0\right) \leq 0 \tag{5.14}$$

$$\implies P\left(\sum_i \alpha_i < 0\right) = 0, \quad \text{since the probability is never negative} \tag{5.15}$$

$$\implies P\left(\sum_i \alpha_i \geq 0\right) = 1 - P\left(\sum_i \alpha_i < 0\right) = 1 \tag{5.16}$$

To conclude the demonstration we report the Theorem 4.23 on page 163 of Roch (2015). The author demonstrates that given two real random variables, one can state that Y stochastically dominates X if and only if there is a coupling (X', Y') of X and Y such that $Y' > X'$ almost surely. In our case, the coupling is defined by (S_n^*, S_n). The relationship between coupling and stochastic dominance is widely described in Whitt (2014). Thus, given Equ. 5.8, for every $k \leq n$, we have the domination of S_n^* over S_n, that is:

$$\begin{gathered} P(S_n^* \leq k) \leq P(S_n \leq k) \\ P(Y_1' + \cdots + Y_n' \leq k) \leq P(X_1 + \cdots + X_n \leq k) \\ F(k; n, p_2) \leq G(k; n, p_1) \end{gathered} \tag{5.17}$$

The final inequality is exactly what we were seeking: given $p_2 > p_1$ the distribution of the variables having probability p_2 dominates the other one.

Chapter 6

N Bonds and Correlation - A Mixture Model Approach

We now turn the discussion to a portfolio of n bonds, exhibiting a correlation. Since we pursue to manage the risk related to the low-carbon transition, in this chapter, the portfolio of investment is identified by its loss distribution. One option to model the dependence between the variables of the said portfolio is through the Copula function, following the procedure in section 2.3.1. However, since we are interested in obtaining some analytic results, we proceed by exploiting another method: the Mixture Model. With this method, the correlation between bonds is introduced through an exogenous factor Ψ, which outlines multiple future scenarios, each one characterized by different default probabilities $p_i(\psi)$ of individual bonds. Our contributions result, firstly, in addressing how the correlation changes according to the exogenous parameter. Then, leveraging on the Mixture Model approach, the impact of a shock (focusing in particular on a Climate policy shock) on the portfolio risk is assessed. Additionally, we examine the asymptotic limit of a portfolio with $n \to \infty$ bonds, providing some personal reflections on how it can be joined to the correlation. Finally, we point out both the limitations and the advantages of this model concerning our objectives.

6.1 Introduction to Mixture Model

Mixture Model is a model to manage credit risk, i.e. the risk that the value of a portfolio changes due to unexpected changes in the credit quality of issuers or trading partners (McNeil et al. (2005)). In literature, Mixture Models are argued as the most useful way of analyzing and comparing one-period portfolio securities. Surely, one of the biggest concerns in managing the credit risk of a portfolio (as one may guess by looking at the previous chapters) is dealing with bonds dependence. Default dependence between two or more firms is supported by valuable economic reasons, which can be clustered into two groups:

1. Firms are affected by common macroeconomic factors. Their financial health fluctuates together with systemic variables, such as changes in economic growth.

2. Firms have some direct economic links between them, such as a strong borrower–lender relationship.

Typically in credit risk management is assumed that direct business relations are less significant in explaining default dependence, because of the huge size of the standard portfolio loans (McNeil et al. (2005)). On the other hand, dependence due to common factors is of crucial prominence and will be a recurring theme in our analysis. The Mixture Model is precisely based on the assumption of default conditional stochastic independence, knowing the realizations of the common underlying stochastic factors. Thus, given a realization of the factors Ψ, the defaults of individual firms $Y_i = 1$ are assumed to be independent. The correlation between different obligors Y_i arises only because the random variables, we exploit to model them, are a function of a common stochastic variable Ψ. Known that, they no longer share anything. Since of particular interest for us, the default correlation, i.e. the linear correlation of the default indicators, is recalled hereafter:

$$\rho(Y_i, Y_j) = \frac{E(Y_iY_j) - E(Y_i)E(Y_j)}{\sqrt{var(Y_i)var(Y_j)}} \tag{6.1}$$

Dealing with binary variables we have

$$E[Y_i] = E[Y_i^2] = P(Y_i = 1), \tag{6.2}$$

$$var(Y_i) = P(Y_i = 1) - P(Y_i = 1)^2 \tag{6.3}$$

$$E[Y_iY_j] = P(Y_i = 1, P_j = 1) \tag{6.4}$$

so the expression in Equ. 6.1 evolves into:

$$\rho(Y_i, Y_j) = \frac{P(Y_i = 1, Y_j = 1) - P(Y_i = 1)P(Y_j = 1)}{\sqrt{P(Y_i = 1)P(Y_i = 1)^2 \cdot P(Y_j = 1)(1 - P(Y_j = 1)^2)}} \tag{6.5}$$

6.2 Bernoulli Mixture Model

Definition 3. Given some $r < n$ and a r-dimensional random vector $\mathbf{\Psi} = (\mathbf{\Psi_1}, \cdots, \mathbf{\Psi}_r)$, the random vector $\mathbf{Y} = (\mathbf{Y}, \cdots, \mathbf{Y}_n)$ follows a <u>Bernoulli mixture model</u> with factor vector $\mathbf{\Psi}$ if there are functions

$p_i : \mathbb{R}^r \to [0,1]$, with $1 \leq \ i \leq n$ such that conditional on $\mathbf{\Psi}$ the components of $\mathbf{Y}$ are independent Bernoulli rvs satisfying

$$P(Y_i = 1|\mathbf{\Psi} = \vec{\psi}) = p_i(\vec{\psi}) \tag{6.6}$$

Therefore we have:

$$P(Y_1 = y_1, \ldots, Y_n = y_n|\mathbf{\Psi} = \vec{\psi}) = \prod_{i=1}^{n} p_i(\vec{\psi})^{y_i}(1 - p_i(\vec{\psi}))^{1-y_i} \tag{6.7}$$

while the unconditional joint default probability is obtained by integrating over the factor vector $\mathbf{\Psi}$. For simplicity in the treatment, from here on, we will focus only on the *one factor* case, i.e. having $r = 1$. It means that we suppose to have only one external macroeconomic variable responsible for the dependence among bonds. One useful formula which will frequently return throughout the entire section is the Total Probability Rule. By applying it to our case, we find that the default probability of company i is such that

$$P(Y_i = 1) = \int_{Dom(\Psi)} P(Y_i = 1|\Psi = \psi) \cdot P(\Psi) d\psi \tag{6.8}$$

here in the case of an external factor continuously distributed. From Equ. 6.8 follows that:

$$P(Y_i = 1) = E[P(Y_i = 1|\Psi)] = E[p(\Psi)] \tag{6.9}$$

6.2.1 Exchangeable Bernoulli Mixture Model

A farther simplification occurs when the variable Y_i modeling the bonds are considered exchangeable; this means that the conditional probability of default $p_i(\psi)$ is the same for every bond. Having a homogeneous group of bonds suggests that we are no more interested in knowing which firm defaults, but the number of firms defaulting in the portfolio. To count the number of default obligors at maturity time T we define a new random variable:

$$N = \sum_{i=1}^{n} Y_i \tag{6.10}$$

Following McNeil et al. (2005) we introduce a simpler notation to indicate the default probabilities which is worth for all the Exchangeable Mixture Models and not only for the Bernoulli's one.

- Probability that bond $i, \forall i$ defaults:

$$\pi := P(Y_i = 1) \tag{6.11}$$

- Joint probability that k bonds default together:

$$\pi_k := P(Y_1 = 1, \cdots, Y_k = 1) \quad k < n \tag{6.12}$$

Given these ingredients, since we are working with exchangeable variables, Equ. 6.5 becomes:

$$\rho_Y = \rho(Y_i, Y_j) = \frac{\pi_2 - \pi^2}{\pi - \pi^2} \tag{6.13}$$

whereas the conditional probability that the bond i defaults after the realization of the exogenous factor ψ (Equ. 6.6) changes into: $P(Y_i = 1|\Psi = \psi) = p_i(\psi) = p(\psi)$ where we neglected the subscript i since we are considering exchangeable rvs. To further simplify, the conditional probability of default, which is a function of the exogenous rv Ψ, is defined as a new rv, called Q.

$$Q := p(\Psi) \text{ with df } G(q) \tag{6.14}$$

Then we have:

$$\sum_{\forall k\text{-combination}} P(Y_1 = 1, Y_2 = 1, \ldots, Y_k = 1|\Psi = \psi) = P(N = k|Q = q) = \binom{n}{k} q^k \cdot (1-q)^{n-k} \tag{6.15}$$

meaning that, conditional on $Q = q$, the number of defaultable bonds N is a rv distributed according to a binomial with parameters n and q. Regarding instead the probability density function of the number of defaults in the portfolio, it can be found by integration over all the possible realizations of the external factor as follows

$$P(N = k) = \binom{n}{k} \int_0^1 q^k \cdot (1-q)^{n-k} dG(q) \tag{6.16}$$

With the change of variable displayed in Formula 6.14 and thanks to Equ. 6.9, the single and joint probability of default (defined in Equ. 6.11 and Equ. 6.12 respectively) simplify:

$$\pi = P(Y_i = 1) = E[Q] \qquad \forall i \in \{1, \ldots, n\} \tag{6.17}$$

$$\pi_k = P(Y_{i_1} = 1, \ldots, Y_{i_k} = 1) = E[Q^k] \qquad \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}, 1 < k \leq n \tag{6.18}$$

6.2.2 Case study: Beta mixing distribution

Below, the case of a One factor Exchangeable Bernoulli Mixture Model is further investigated. Thus, we consider a collection of bonds which belong to the same category, for instance, all labelled as *green*. Then, given its handiness, we choose a Beta mixing variable Q, i.e with $G(q) \sim Beta(\alpha, \beta)$. Indeed, leveraging on its recursion properties, as shown in McNeil et al. (2005), one can get interesting results. Both the single and joint default probabilities are easily computable

$$\pi = E[Q] = \frac{\alpha}{\alpha + \beta} \tag{6.19}$$

$$\pi_2 = E[Q^2] = \frac{\alpha(\alpha + 1)}{(\alpha + \beta + 1)(\alpha + \beta)} = \pi \frac{\alpha + 1}{\alpha + \beta + 1} \tag{6.20}$$

as well as the default correlation

$$\rho = \frac{var[Q]}{E[Q](1 - E[Q])} = \frac{1}{\alpha + \beta + 1} \tag{6.21}$$

Now that we have all the elements to proceed we aim to inquire the connection between the parameters proper of the mixing distribution and the correlation, providing an original interpretation. We start by showing how the Beta probability distribution function changes while we tune its parameters α and β. In Fig. 6.1 we just change α whereas β is considered fixed at the value of 1.5. One can notice that by increasing α the biggest portion of the probability density moves more and more to the right, i.e. towards 1. This means that high values of α lead to high conditional default probability q. On the

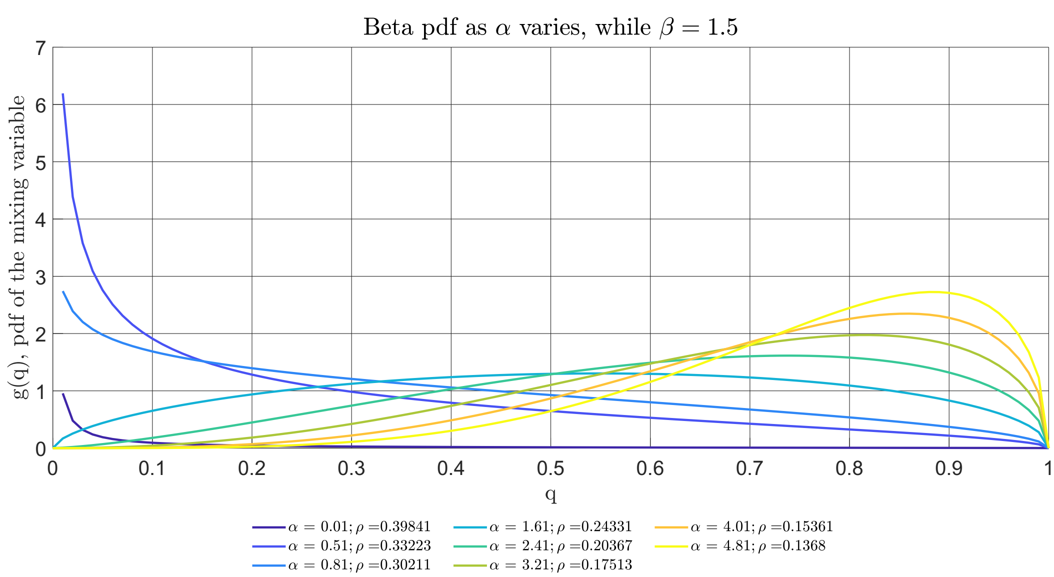

Figure 6.1: **Probability density function of the mixing variable Q**. Every curve refers to a different value for α while β remain fixed and equal to 2

other hand, by fixing α and varying β the density would be shifted toward zero. Similarly, a mixing distribution with a high β corresponds to a situation where the conditional probability of default is very low. One can question the reason why diminishing α and β we get a higher correlation. One answer can be found by looking at Equ. 6.19 and 6.20: in the limit of α and β tend together to zero, it follows that $\pi_2 \to \pi$, because both α and β are negligible with respect to 1 in the parenthesis both in the numerator and in the denominator. This means that it does not matter how many bonds fail. Being in the limit perfectly co-monotonic, a portfolio formed by only one bond on which the investor put all her wealth behaves exactly as a diversified portfolio of n bonds. This implies that if one bond defaults all the bonds will default too and this happens with probability E[Q]. Whether a large correlation is attainable when the Beta seems like a Bernoulli with mean 0.5 leads us to relate the correlation to the variance of the mixing distribution. It appears that the higher is the variance of Q, the largest the correlation. Indeed, a Bernoulli with mean 0.5 is the case having the largest variance.

Figure 6.2 shows how the correlation and the unconditioned default probability change as α and β varies. Looking at the full line displaying the behavior of the correlation, one can notice that,

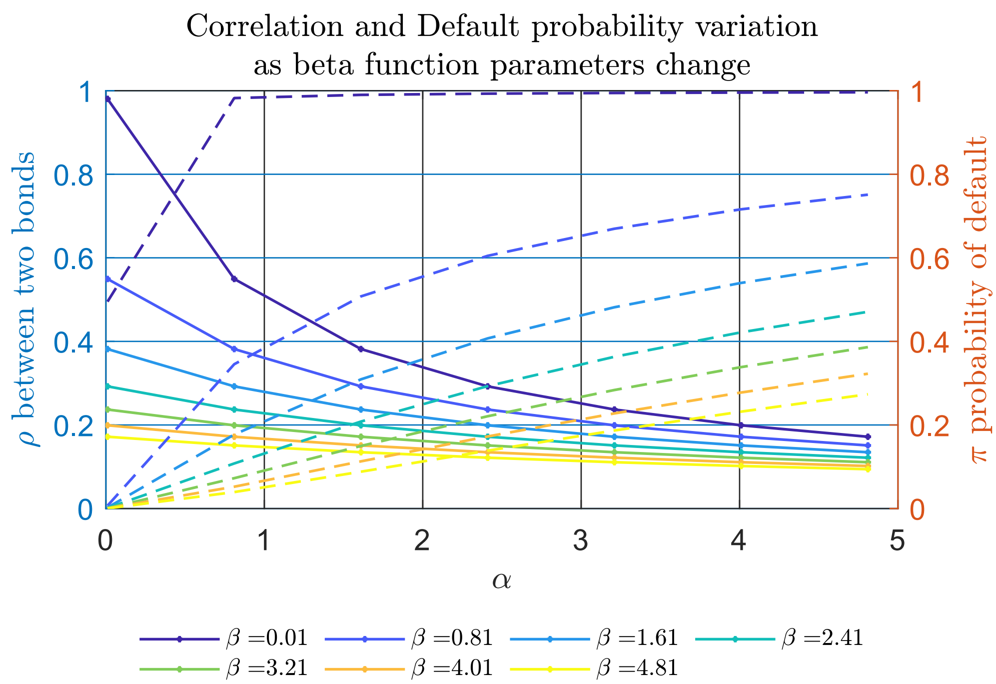

Figure 6.2: Dependence between **correlation, single probability of default and the parameters of the mixing distribution**

inverting α and β the graph for the correlation would be the same since α and β are symmetrical in Equ. 6.21. Besides, since the distributions with both $\alpha \approx \beta << 1$, as said, are the ones having the highest variance, the correlation takes there the highest value and then decreases as α or β increase. The lower bound for ρ is indeed met when both α and β tend to ∞. A Beta distribution with both its parameters very large gathers the density around a single point which is the mean value of Q, which leaves no more space to uncertainty. Therefore $\pi = P(Y_i = 1) = P(Y_i = 1|Q = q) = q$. Being only one scenario attainable the default probability conditional to that scenario coincides exactly with the unconditional one. The external factor does not provide us with more information about default and does not result in a shock on the default probabilities. Since we guess that the uncertainty carries the correlation, the fact of not having here a correlation is therefore justified. Given that the probability of an individual default equals to q, once I know π the bonds in the portfolio are just i.i.d. random variables. For what concerns the dashed curve, i.e. the ones related to the probability of default, we see that, as expected, the probability of default increases dramatically from zero to a non-zero value increasing α. This is expected behavior since the α parameter for the Beta distribution drives the probability density towards high value for Q, which corresponds to bad states for the firm values. As further evidence of what has been said above, we see that when both parameters go to zero with the same rate, the probability of default is exactly 1/2, equally balanced between the worst-case scenario (with $q = 1$) and the most desirable one(with $q = 0$). On the other hand, one can see that taking α fixed, the correlation grows always of the same amount while β becomes bigger. Except for $\beta = 0.01$, which makes sense since corresponds to a probability density concentrated around 1, whichever value α takes.

6.2.2.1 Handling the Loss distribution

As mentioned at the beginning of this section, with the Beta distribution one can analytically find the portfolio losses distribution

$$P(N = k) = \binom{n}{k} \frac{\mathcal{B}(\alpha + k, \beta + n - k)}{\mathcal{B}(\alpha, \beta)} \tag{6.22}$$

where $\mathcal{B}(\alpha, \beta)$ denotes the Beta function. Looking at the graphs of the probability density function of the losses (Figure 6.3 and 6.4) we see that the curves follow perfectly the ones in Figure 6.1 (i.e. related to the probability distribution of Q). Given a set of Beta opportunities, the shape of the unconditional probability density function of losses is determined by how the probabilities of default of a single firm conditional to a scenario q distribute, while the Beta distribution is changing. Therefore, selecting α and β for the variable describing the factor, we are choosing the distribution of the portfolio loss.

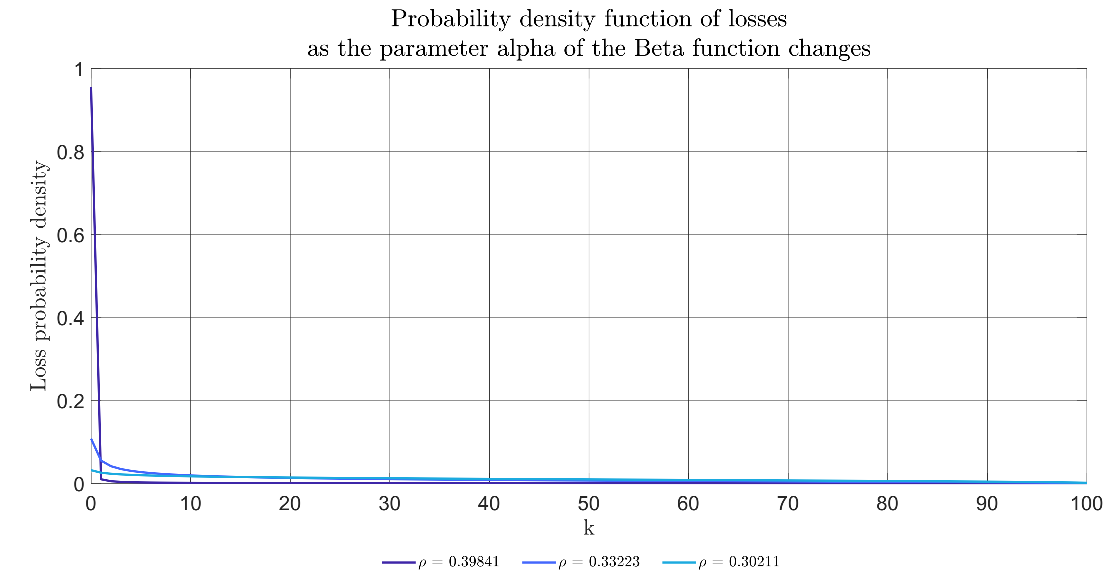

Figure 6.3: **Beta-Binomial probability density** function, with $\beta = 1.5$ and $\alpha < 1$

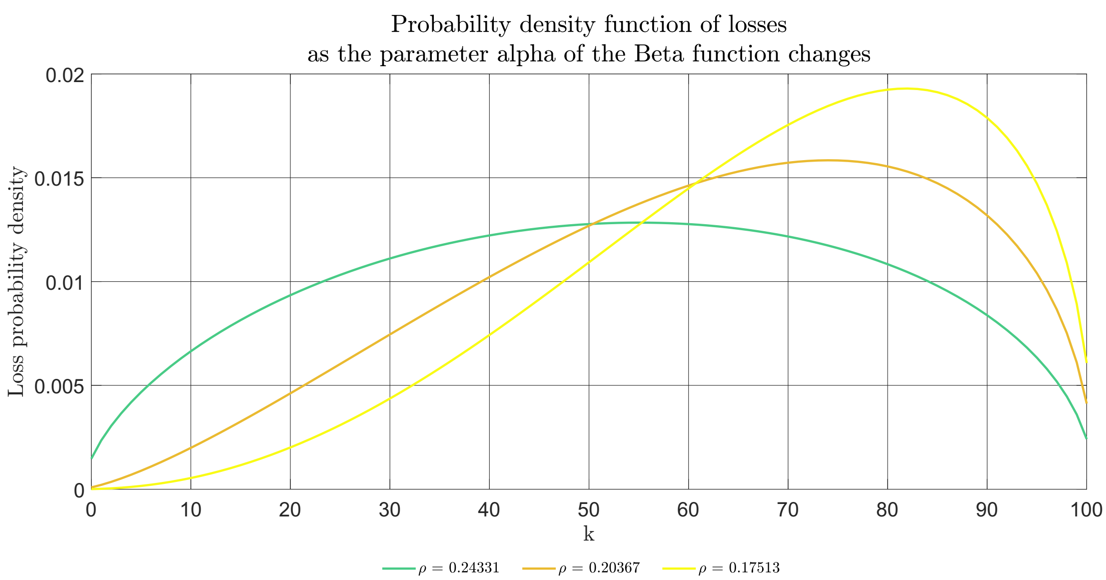

Figure 6.4: **Beta-Binomial probability density** function, with $\beta = 1.5$ and $\alpha > 1$

Additionally, inspecting the cumulative distribution functions behaviour, one may see a relationship of first order stochastic dominance (FSD) among the Beta distributions at fixed β (here equal to 1.5), as displayed in Figure 6.5. In particular, the function with the biggest α, the yellow one, dominates the others. Moreover, a relationship of FSD can be defined also between the distribution having $\alpha = 1.5$ and fixed whereas β varies (see Figure 6.6). Notice that, here, the function which *dominates* the other corresponds to the best portfolio, which is foreseen to suffer fewer losses. Even it may seem in contradiction with the discussion at Chapter 3, it is not since here the random variable $Y_i = 1$ indicates default while in Chapter 3 the analysis is made by considering the portfolio payoff. Under this perspective, here, one should prefer the dominated distribution rather than the dominant one. The fact that the *dominate* function is the less risky is supported also by computing the c-quantile of the different loss distribution by varying the scenario. Notice that here we use the notation of c-quantile rather than the α-quantile to avoid blending in with the α parameter of the Beta distribution. We consider then $c = 10\%$, meaning that with the 10% of probability, the portfolio can suffer a total number of defaults bigger than the preimage of $1 - c$, represented in the figures by a gray dashed line.

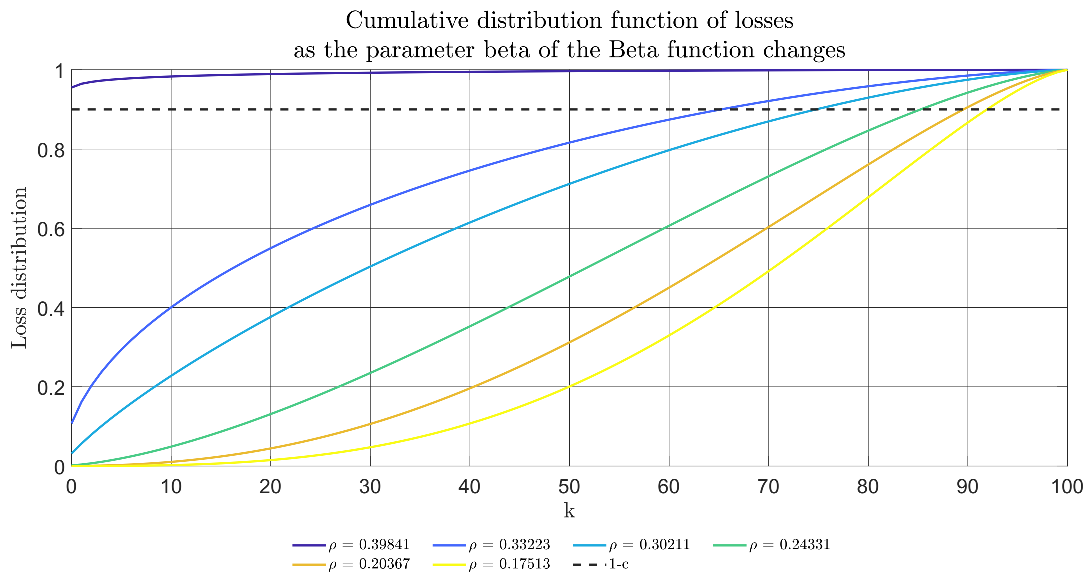

Figure 6.5: **Beta-Binomial Cumulative** distribution function, with $\beta = 1.5$ **fixed** and α variable

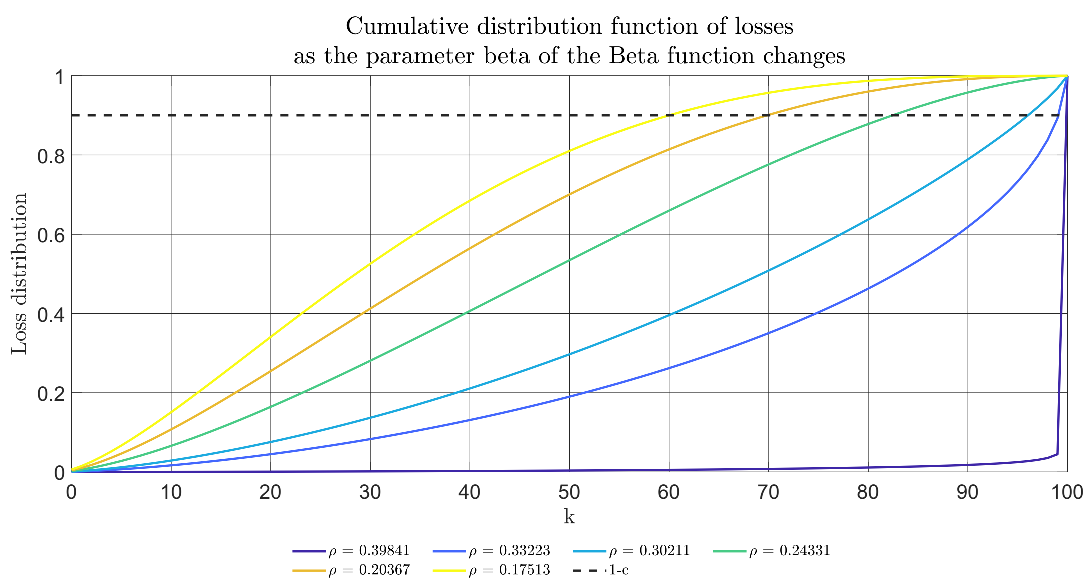

Figure 6.6: **Beta-Binomial Cumulative** distribution function, with $\alpha = 1.5$ **fixed** and β variable

Concerning the correlation, as one can guess looking at Figure 6.7 that it does not provide an univocal rule to determine the shape of the loss distribution. Indeed, the two charts compare two portfolio with an expected number of losses totally different. The right graph is a super-safe portfolio, while the left one is dramatically bound to fail. This is caused by the fact that Equ. 6.21 is symmetric and thus it is as not considering α and β as two separate parameters but as only one. The correlation is not influenced from the outcomes of the exogenous factor itself but from the effect it provokes on the default probability. Analyzing deeper Figure 6.7 we see that for $\alpha > \beta$ the probability density is peaked

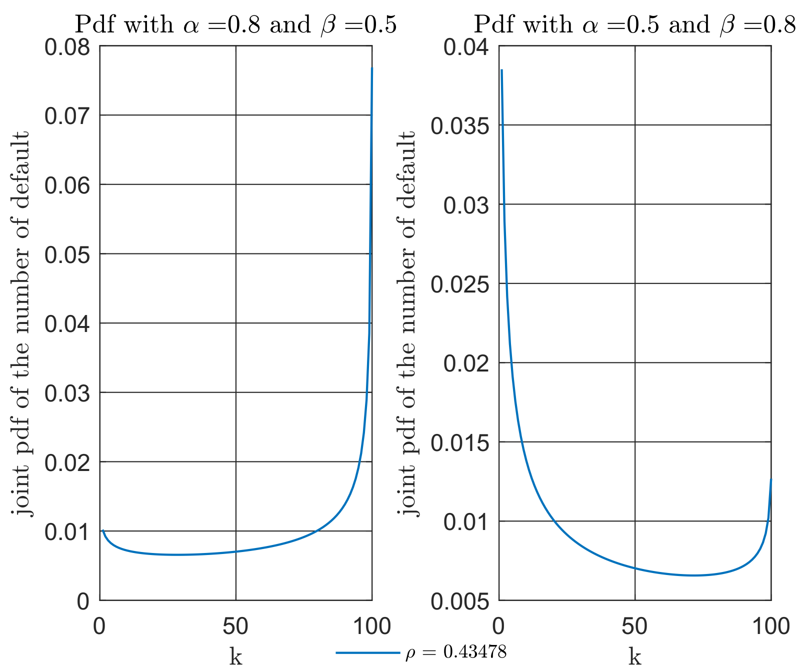

Figure 6.7: **Symmetry of the correlation**: probability density function of the portfolio having the same $\rho = 0.43$, but with α and β reversed.

on the value $k \simeq n$, which means that it is more probable that the bonds default all together than to default individually or in small groups. This makes sense since if we plot the mixing distribution $g(q)$ with those values for α and β it is clear that for $\alpha, \beta < 1$ and $\alpha > \beta$ the default probability of a single variable after the realization of the external factor has a higher peak near $q \simeq 1$ than near 0. Quite the opposite occurs in the specular case: peak on $q \simeq 0$ and hence a higher probability that the joint of the portfolio is unbalanced towards $k \simeq 0$. However, the values reached by the peaks in both plots are very small. This is because the probability of the portfolio is nothing more than a beta-binomial distribution spread on a domain $[0, n]$, with n number of bonds in the portfolio, which is n times the original one.

Like many other mixing distributions, the Beta belongs to a two-parameter family. So, to specify entirely the model, we can either fix the default probability π and the default correlation ρ or equivalently the first two moments of the mixing distribution π and π_2 (see Equ. 6.11 and 6.12). Setting the correlation ρ and the probability of a single bond defaults π allows us to choose values that have a clear and well-defined meaning, compared to setting α and β, which might see more like two empty parameters. In this last paragraph, we wish to point out the effect of the external variable on the total

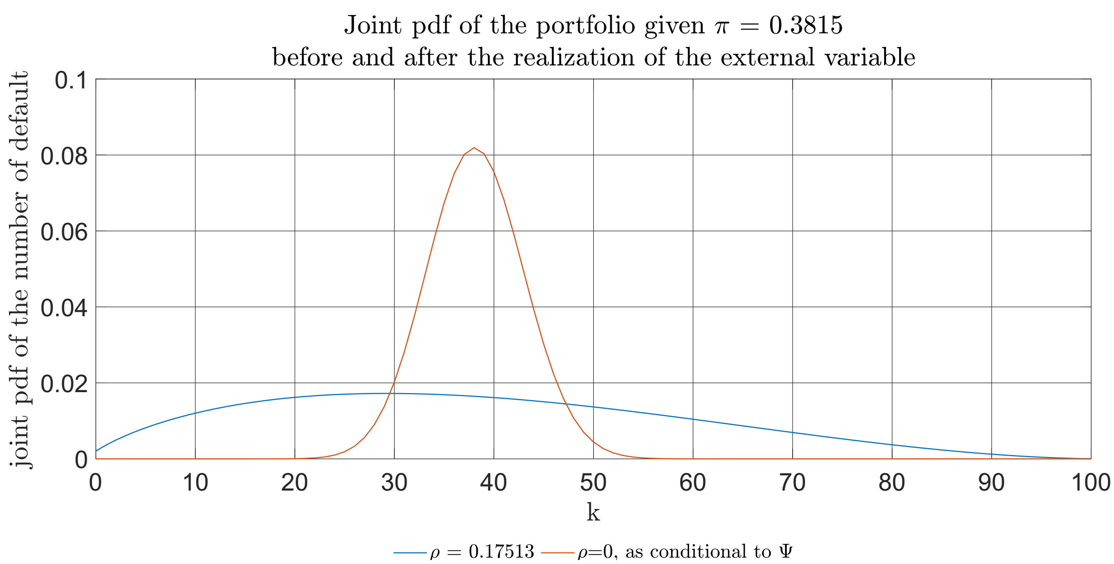

Figure 6.8: **Comparison between Beta-Binomial($\rho \neq 0$) and Binomial($\rho = 0$)**

loss distribution. Figure 6.8 shows two curves: the blue curve corresponds to a Beta-binomial portfolio loss distribution in case of a single probability of default $\pi = 0.3815$ and a correlation of $\rho = 17513$; on the other hand, the red one represents a Binomial density function with n bonds and average in $n \cdot \pi$. The red curve draws the distribution of bonds once the external factor is known, which corresponds also to the loss distribution with $\rho = 0$. The red curve, being much steeper than the blue one, tells that the portfolio will suffer losses ranging from 20 to 60. It is evident that, although the correlation is very small, the shape of the unconditional distribution (the blue one) flatten, and the portfolio has a non zero probability to incur any losses in the interval varying from 0 to 90. The fact that the range becomes broader increases the uncertainty on the effective loss of the portfolio, and lowering the peak of 75%. Not knowing the distribution of the outcome therefore leads to a correlation that leaves much more uncertainty about the performance of the portfolio at maturity.

6.2.3 Numerical Evaluation of the Correlation caused by the Common Factor Ψ

The Mixture Model is based on the assumption that correlation across obligors in credit events is driven by a common dependence on a set of systematic risk factors. In this section, we examine this statement via numerical simulation. We select two bonds from a portfolio of exchangeable bonds. We consider then that, depending on whether or not an external event takes place, the distribution of returns has a particular shape. SETTING OF THE PROBLEM:

- Two homogeneous bonds: let consider for instance two *green* bonds.
- The Exogenous factor is a binary random variable. We imagine that this factor correspond to a Climate Policy or Agreement similar to the Paris Agreement; for this reason we will call it PA. Thus, the old Ψ can take two values and we call it PA. If $PA = 0$ the agreement is not signed, whereas $PA = 1$ means that the agreement is signed. Depending on the value taken on by PA, the scenario changes. The probability density function which models the bond's return changes according to PA too.
- We consider two Scenarios, called HIGH or LOW depending on whether the mean of the distribution of returns is high (near 1), meaning agreement signed or low (near 0), agreement unsigned.
- The Probability that a bond defaults Conditional to the external event is obtained throughout a *threshold* on the distribution of returns.
- The Probability density function of the Returns conditional to the external event is supposed to be a Beta with parameters $\alpha(PA)$ and $\beta(PA)$. According to which scenario is realized, the Beta distribution is more flatten to the right or to the left. Thus, the Beta models the return, in a sense that given the scenario we know precisely if the variables are extracted from a low or high mean distribution.
- The Beta distribution of returns may be envisioned as representing the health state of a firm.
- The Paris Agreement is signed with probability $p_{realized}$, which can be seen as the bookmakers' quotation on the probability with which the event may or may not happen.
- We call X_{Mixed} and Y_{Mixed} the two random variable unconditional to the external factor, while X_{LOW}/Y_{LOW} and X_{HIGH}/Y_{HIGH} are the variables, knowing which scenario is realized. The "Mixed" variables, i.e. the ones which are not conditional, are created mixing HIGH and LOW with a proportion imposed by $p_{realized}$.

- To get from the distribution of the returns the Number of bonds that Default one can use the aforementioned threshold. This means that all the values of the returns below than a certain fixed threshold are considered as default, i.e. $Y = 1$, whereas the ones above take $Y = 0$.

The numerical simulation provide us a proof for our intuition: the correlation between $X_{Mixture}$ and $Y_{Mixture}$ arises because every vectors' element depends on the outcome of PA, which given its uncertainty, contains some risk. Playing with the free parameters of the simulation, first of all, one can see that the correlation changes while $p_{realized}$ is tuned. In particular, it reaches the maximum when $p_{realized} = \frac{1}{2}$. This condition can be compared to the one related to the Beta distribution having maximum variance, and therefore maximum correlation. Having two equiprobable scenarios means that the uncertainty and unknown on which scenario dominates increase, which leads to a larger correlation between the outcomes. Conversely, if the bookmakers know, with probability almost 1, if a generic Climate Policy will be signed or not, there is no more uncertainty about which scenario is fullfilled, and thus the impact on the bonds' returns is known too. Secondly, the correlation is even more influenced by the "severity" of the scenario, i.e. how bad is the scenario labeled as LOW and good the scenario *HIGH*. The more the forecasts on the distribution of returns conditional to a scenario (i.e. the beta distributions) are different, the higher the correlation between the bonds. Evidence of this can be seen comparing Figure 6.10 with 6.12. Another way through which the value of the correlation may be explained is by looking at the definition of Pearson's linear correlation coefficient. One can see that $(X_{Mixture} - \mu_X)(Y_{Mixture} - \mu_Y)$ is positive if and only if both $X_{Mixture}$ and $Y_{Mixture}$ are on the same side of the respective mean values. Figure 6.9 corresponds to a correlation of 0.75 between the two green bonds. This number makes sense since most of the points stay in the top-right corner or the bottom-left one, and represent values of returns that are respectively jointly above or below the average. Thus, $\rho > 0$ whether $X_{mixture}$ and $Y_{mixture}$ take simultaneously values that are larger or smaller than the respective mean. On the other hand, $\rho < 0$ if when $X_{mixture}$ is higher than its mean, $Y_{mixture}$ is lower. Moreover, the higher is the trend to have the same behavior (both above and below or one above e the other below) bigger is the absolute value of the coefficient. Concerning the case of $\rho = 0$, as we can see from Figure 6.11 the points seem randomly scattered over the entire plane.

6.2.3.0.1 Testing with Negative Correlation In this paragraph we briefly display the case of a couple of non Exchangeable bonds. The purpose is to test the simulation even for a non homogeneous situation. This means that we are dealing with one *brown* and one *green* bond. Hopefully, the introduction of a Climate Policy should influence positively the green bond and more negatively the brown one. In the simulation, we model the carbon-intense bond by the variable $X_{Mixture}$. Differently from before, for the brown bond, to an unsigned agreement is associated with the high scenario (thus a high probability to have a higher return), while the probability conditional to the sign of the Climate agreement is defined by the low scenario. The last two figures (Figure 6.13 and 6.14) represent precisely two bonds negatively correlated.

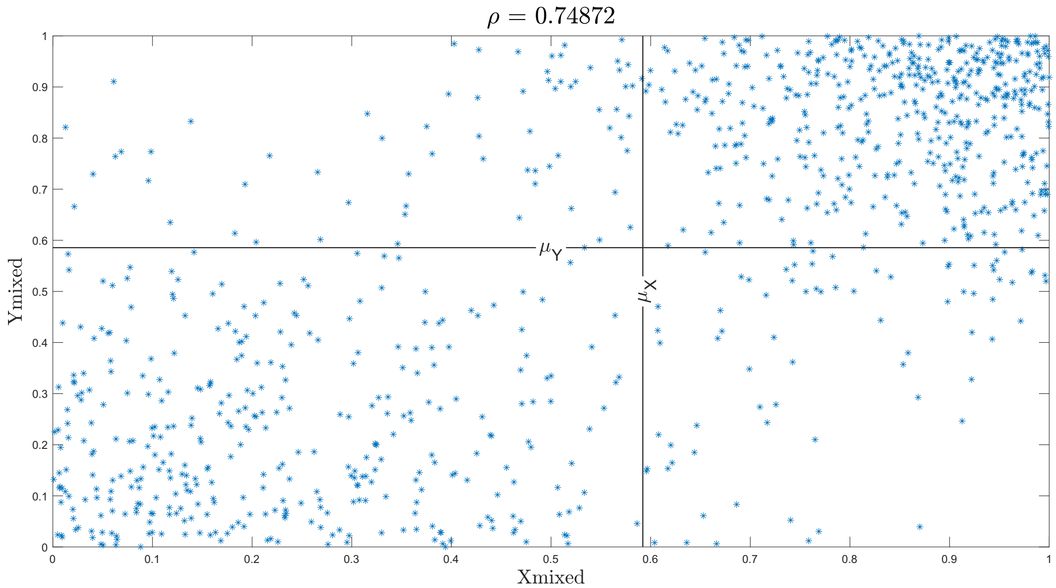

Figure 6.9: **Visual** representation of **Default correlation** between two **homogeneous** bonds, **highly-positive correlated**.

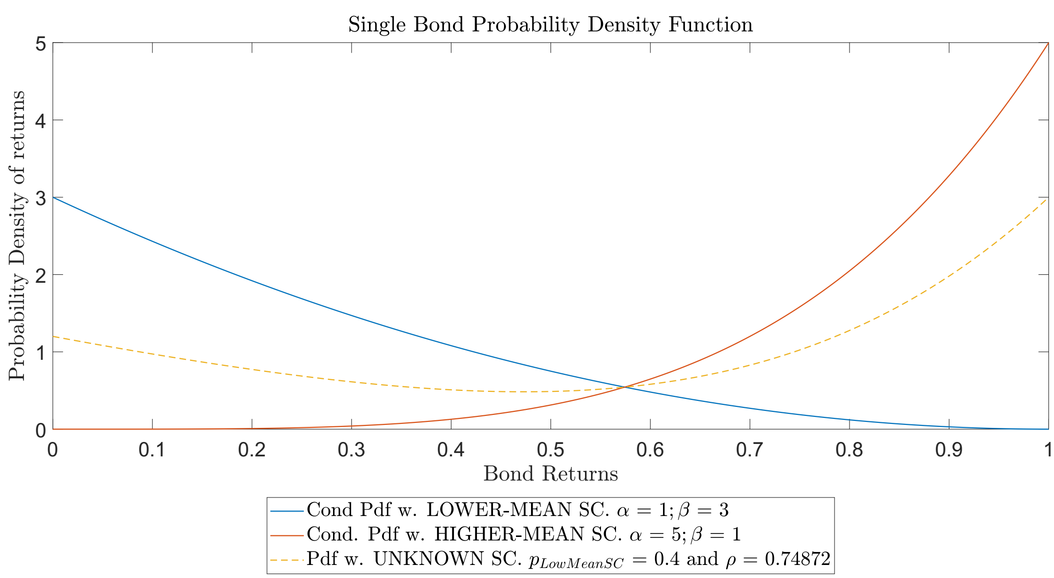

Figure 6.10: **Comparison between single default probabilities**: (Red curve) conditional to the high scenario, (Blue curve) conditional to the low one, (Dashed curve) with the scenario unknown. $\rho \approx 0.75$**, the scenarios do not overlap**.

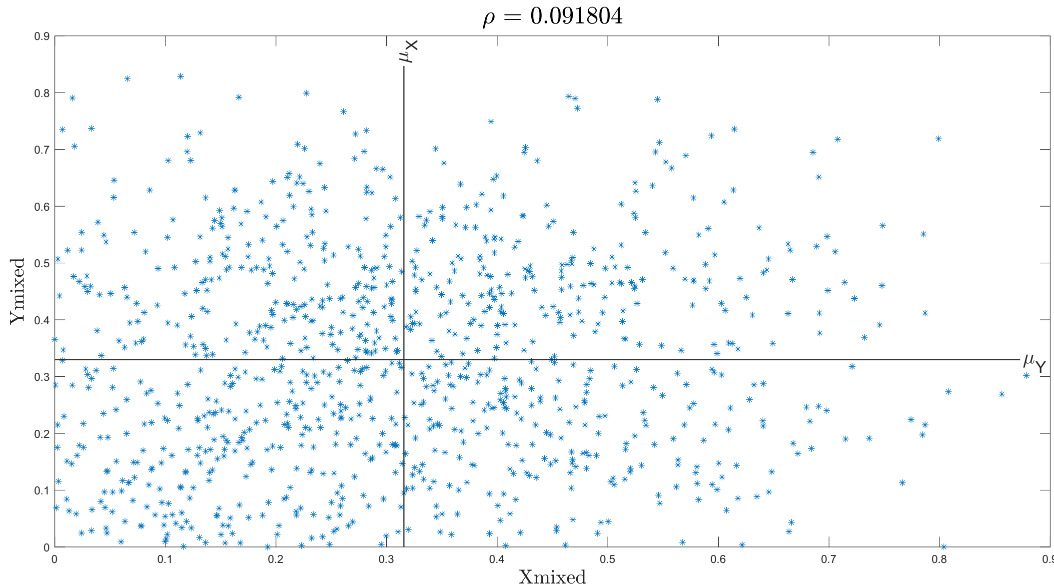

Figure 6.11: **Visual** representation of **Default correlation** between two **homogeneous** bonds, **weakly correlated**.

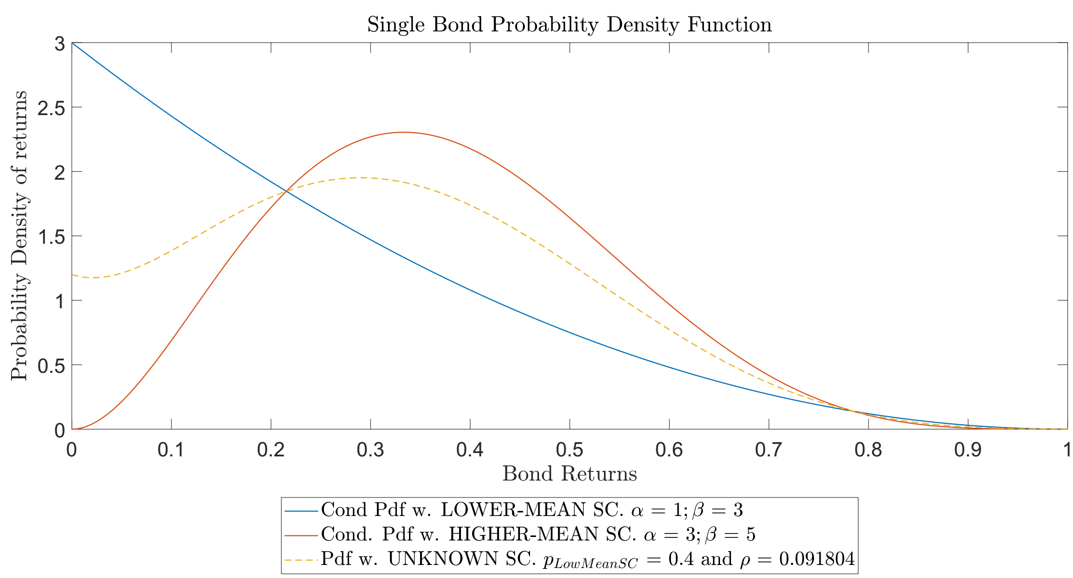

Figure 6.12: **Comparison between single default probabilities**: (Red curve) conditional to the high scenario, (Blue curve) conditional to the low one, (Dashed curve) with the scenario unknown. $\rho \approx 0.09$**, the scenarios overlap**.

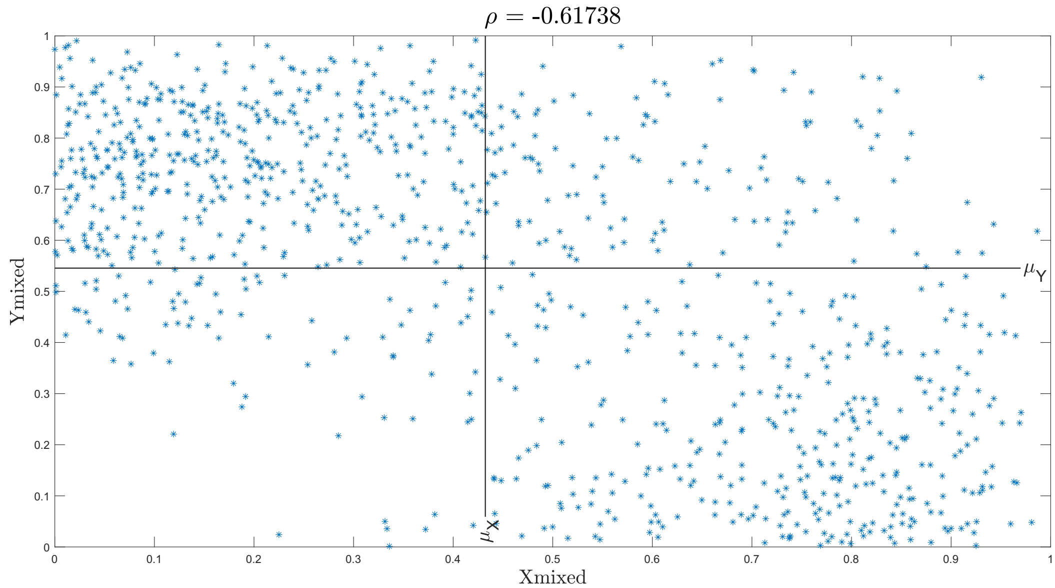

Figure 6.13: **Visual** representation of **Default correlation** between one ***green*** and one ***brown*** bond, hence **negatively correlated**.

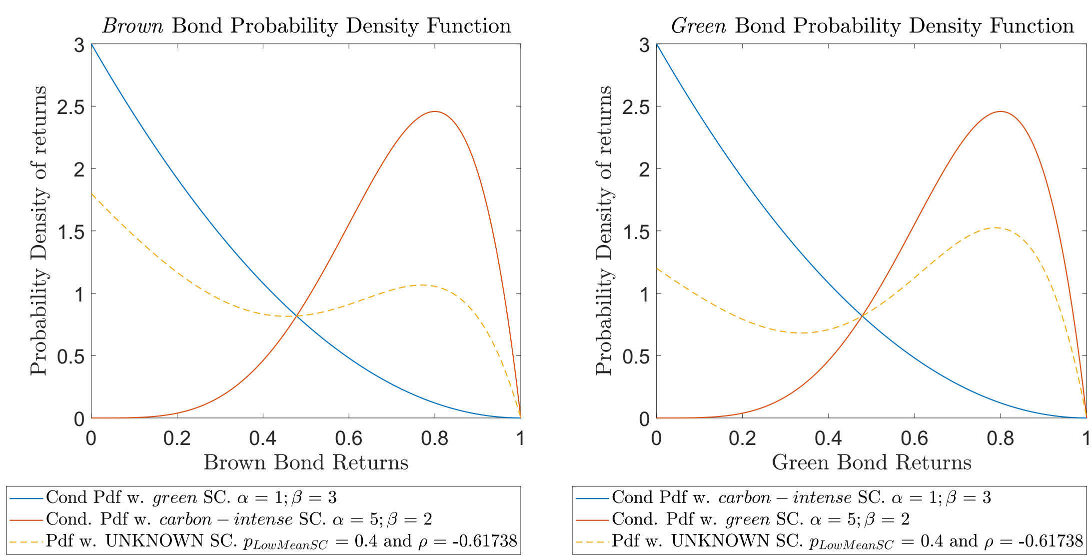

Figure 6.14: **Comparison between single default probabilities**: (Red curve) conditional to the high scenario, (Blue curve) conditional to the low one, (Dashed curve) with the scenario unknown. $\rho \approx -0.62$**, the scenarios do not overlap**.

6.3 Asymptotic result for Large Portfolios Loss distribution

We now present some of the most remarkable asymptotic results for large portfolios in One factor Bernoulli Mixture Model. The derivation of these results is available in the paper of Frey and McNeil (2003). The authors provide a tool to evaluate approximately the credit loss distribution in a large portfolio of bonds or of other securities. In particular, the crucial point is that the tail of the loss distribution is basically driven by the tail of the mixing distribution. Leveraging on this, we aim at finding its implication on the optimal portfolio allocation. In particular, we first pursue the link between correlation and loss distribution, and then we try to underpin the evidence that the stronger the correlation, the lesser the benefit of diversification, which we studied, although considering only two bonds, in Chapter 3 of this work.

Referring to Frey and McNeil (2003), we start by defining the asymptotic conditional loss function, i.e. the loss that on average the portfolio can incur in the n (number of bonds in the portfolio) large limit. To do so, we need first to define the loss of a single firm L_i as the product between the positive deterministic exposure e_i, which refers to the amount of money that the investor has invested in a particular security, the corresponding default indicator Y_i and the percentages of losses related to the company i, given that default occurs δ_i. In this framework the loss of portfolio of size n is given by $L^{(n)} = \sum_{i=1}^{n} L_i$, where $L_i = e_i \delta_i Y_i$ is the of a single firm. This sequence is assumed satisfying a set of assumptions,

$$[\text{A1}] \quad l_i : \mathbb{R} \to [0,1] | L_i \text{are independent if conditional on } \psi \text{and with mean } l_i(\psi) = E(L_i|\Psi = \psi) \tag{6.23}$$

$$[\text{A2}] \quad \exists C < \infty | \sum_{i=1}^{m} \frac{{e_i}^2}{i} < C \quad \forall m \tag{6.24}$$

Definition 4. [A3] <u>The Asymptotic Conditional Loss Function</u> is a function $\bar{l} : Dom(\Psi) \to \mathbb{R}^+$ such that

$$\lim_{n\to\infty} \frac{1}{n} E[L^{(n)}|\Psi = \psi] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E[L_i|\Psi = \psi] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} l_i(\psi) = \bar{l}(\psi) \quad \forall \psi \in \mathbb{R} \tag{6.25}$$

where we denotes the function $l_i(\psi) : \mathbb{R} \to [0,1]$ as $l_i(\psi) = E(L_i|\Psi = \psi)$. Besides, one can notice that the case of LGD (Loss Given Default) of 100% and bond exposure 1 implies having $L_i = Y_i$, from which it follows $l_i(\psi) = P(Y_i = 1|\Psi = \psi) = p(\psi) \quad \forall \psi$.

Besides, we consider a one-dimensional mixing variable Ψ with distribution function $G(\psi)$. Then we assume that the conditional asymptotic loss function $\bar{l}(\psi)$ is strictly increasing and right continuous and that G is strictly increasing at $q_c(L^{(m)})$, i.e. the quantile of the distribution of Ψ. Then:

$$\lim_{n\to\infty}\frac{1}{n}q_c(L^{(n)}) = \bar{l}(q_c(\Psi)) \tag{6.26}$$

To ensure that the assumption on the monotony of $\bar{l}$ makes sense, we need to assume that low (high) values of Ψ correspond to good (bad) states of the world, with conditional default probabilities and losses given default lower (higher) than average. Therefore, from Equ. 6.26 follows that the tail of the portfolio loss distribution in the case of a large portfolio is determined by the tail of the mixing variable. Remember that the quantile function for a continuous random variable is the inverse distribution function. Saying that the c-quantile $= \hat{\psi}$ means that, with a probability of $c\%$, the portfolio may incur in a loss bigger than $\hat{\psi}$. Thus, a bigger c-quantile means a fatter tail and then a deeper loss. One easy example is the case of exchangeable variables with mixing distributions $G_i(q) = P(Q_i < q), \quad i = 1, 2.$ where of G_1 has a heavier tail than G_2; by consequence it is reasonable to expect that the losses are large for the model 1 rather than the 2.

6.3.1 Tie together Correlation and Loss distribution

6.3.1.1 Link between Correlation and Moments of the distribution

For what concerns how the correlation and the loss distribution are linked, we start leveraging Equ. 6.21, defining the correlation in case of an Exchangeable Bernoulli Mixture model and reported hereafter:

$$\rho = \frac{var[Q]}{E[Q](1-E[Q])} = \frac{var[p(\Psi)]}{E[p(\Psi)](1-E[p(\Psi)])} = \frac{\sigma^2}{\mu(1-\mu)} \tag{6.27}$$

The equation envisages a relation between the correlation coefficient, the mean of the distribution of $p(\Psi)$ (which looking at Equ. 6.17 corresponds also to the unconditional probability distribution of a single bond) and the variance of the mixing distribution. We see that any value in [0,1] can be obtained by a suitable choice of the parameters characterizing the mixing distribution. No negative value are allowed since the variables we are considering are exchangeable, thus they belong to the same category as well as they react in the same way to the external factors.

In particular, $\rho = 1$ is obtained when $\pi = \pi_2$, that is when the probability of default of one bond is equal to the joint default probability of two firms taken together, and the distribution of Q has two peaks: one in 0 and the other one in 1. Having a correlation of 1 under such conditions makes sense since, $\pi = \pi_2$ models a situation where, if one firm defaults, then necessarily two firms taken together will also default. The mixing distribution is not concentrated around a single value, but is has two peaks at the extremities of the domain. It contains maximal uncertainty about the outcome of the scenario, given that the variance takes its maximum value. Once the bad scenario is realized, the probability to see Y_i defaulting is 1, whereas no defaults are registered if the good scenario is attained. Regarding the case of $\rho = 0$, it arises when all the density of the distribution of Q is concentrated around a single point that coincides with the mean π. Intuitively, this means that there is one and a scenario available, with probability $\pi = p(Y_i = 1|\Psi = \psi)$. Therefore, since Equ. 6.11 is true only if $p(\psi) = 1$ the outcome of the external macro factor is known and no more uncertainty derives from the outcomes of the scenario since basically only one of them can exist. Similarly, relying on the assumption grounding the model, if $\pi = p(Y_i = 1|\Psi = \psi)$ is true and $Y_i = 1|\Psi = \psi$ are independent random variables, then the independence structure is transferred back also to the Y_i non conditional and the correlation becomes zero.

Now, looking at equation 6.27 one can question how the correlation changes according to a variation of either the variance or the mean. In the first place, the variance is linked to the mean by means of its domain. Indeed, the standard deviation cannot exceed the distance between the mean and the extremities of the set of departure of μ which, corresponding to the expected value of a probability, is bounded between 0 and 1: $\sigma_{MAX} = min\{\mu, (1-\mu)\}$. Therefore, $\sigma^2_{MAX} = min\{\mu^2, (1-\mu)^2\}$ and then:

$$0 \leq \frac{\sigma^2}{\mu(1-\mu)} \leq min\{\mu^2, (1-\mu)^2\} \tag{6.28}$$

$$0 \leq \rho \leq min\left\{\frac{\mu}{1-\mu}, \frac{1-\mu}{\mu}\right\} \tag{6.29}$$

It follows, as expected, that the maximum of the correlation may be reached only when $\mu = \frac{1}{2}$.
For a generic distribution of Ψ and a generic functional form of p, specifying the mean is not sufficient to fully fix the variance. Considering the case of the mean μ fixed to a specific value, what we know is that the variance is bounded by σ_{MAX}. Then, one can consider the correlation as proportional to the variance. A large correlation can derive from a high variance, meaning that the probability of one default conditional to a particular realization (let's call it ψ_1) of exogenous factor Ψ has the same likelihood as the probability conditioned to all the other possible outcomes ψ_i with $i \neq 1$. The correlation is driven by the uncertainty in the outcomes of Q. This can be somehow related to the concept of mean preserving spread and other risk measure treated in Chapter 3. The portfolio is riskier if the variance is high, at equal average value.
Analysing the correlation by using the variance as a floating parameter is found to be mathematically easier to handle than the one with the mean. For this reason, in the next, we consider the mean only to set an upper bound to the variance. One can see from the Figure 6.15 and 6.16 that, as expected, the variance and the standard deviation take the maximum value exactly where also the correlation reaches its maximum value. In those figures, we have on the X and Y axis the two parameters of the

Beta distribution α and β respectively. The height corresponds to the default correlation coefficient between two firms, whereas the color bar represent the standard deviation or variance. The smallest ρ is reached when $\alpha = \beta = 5$, i.e. the maximum value. We agree with this result since increasing α and β the function squeezes around the mean (which is $1/2$ every time $\alpha = \beta$), meaning that the exogenous factor can take just one value, and the correlation stretches more and more to zero. On the other hand the maximum value of ρ is attained when both α and β tend to zero. As expected that point corresponds also to the one having maximum standard deviation as well as variance, and therefore the biggest uncertainty. Looking at the Figure 6.16, the colour scheme is more pronounced, (as is expected given the square root) allowing to highlight some points in the larger blue zone. In particular the points $(\alpha, \beta) = (5, 0)$, and $(\alpha, \beta) = (0, 5)$ i.e. those at the extremes of the domain having an average value very close to the extremities clearly have the smallest variance.

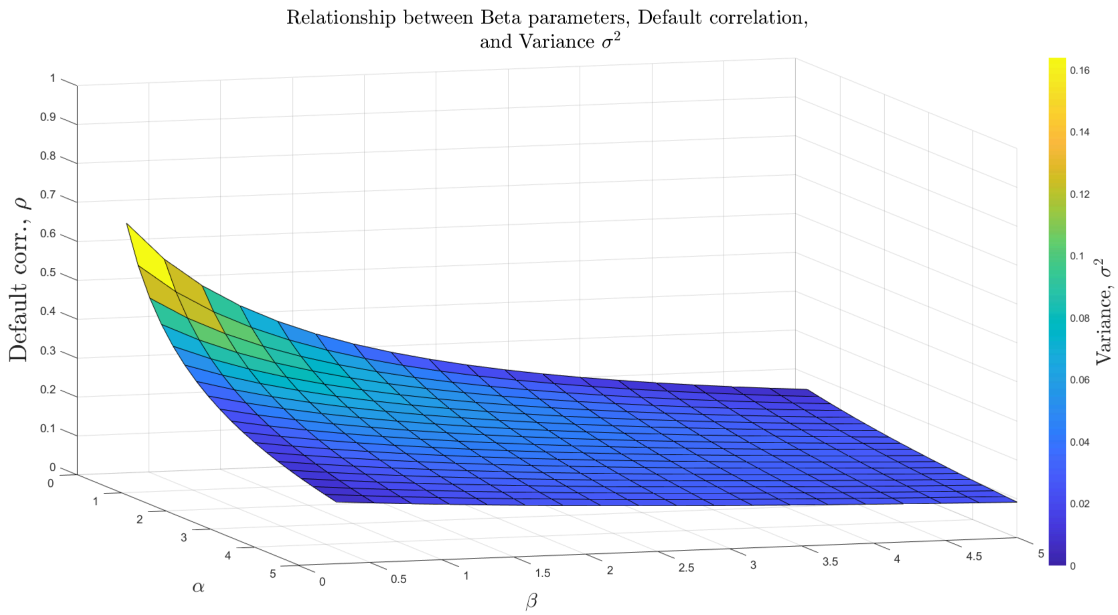

Figure 6.15: **3D** representation of the relationship between **Beta parameters, Default correlation and Variance**

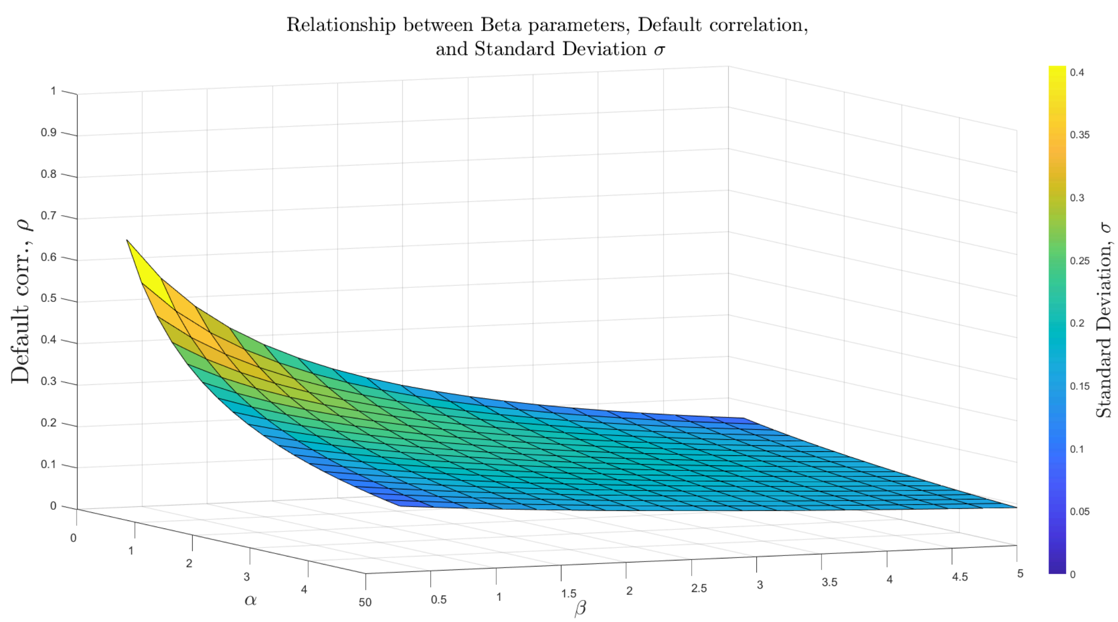

Figure 6.16: **3D** representation of the relationship between **Beta parameters, Default correlation and Standard deviation**

6.3.1.2 Link between Moments and c-quantile

The next step is understanding how the mean and the variance are linked to the c-quantile of the distribution. There exist different methods in literature which compare empirically estimates for mean and standard deviation, starting from the knowledge of the quantiles. We report in the following the Extended Pearson-Turkey Method. This method is a three points approximation for continous random variables having a two-parameter distribution. The Pearson curves are a convenient example of a two-parameter family. Since a wide variety of distributions are close to a two-dimensional manifold, those curves are suitable to approximate all these diverse-appearing distributions (Pearson and Tukey (1965)). To provide the estimate of mean and variance three quantiles (equal to 0.05, 0.50 and 0.95) have to be specify. Here the subscript correspond to $a = 1-c$ since the quantile are computed starting from the left tail.

$$\mu = 0.630q_{0.5} + 0.185(q_{0.95} + q_{0.05}) \tag{6.30}$$

$$\sigma^2 = 0.630(q_{0.5} - \mu)^2 + 0.185[(q_{0.95} - \mu)^2 + (q_{0.05} - \mu)^2] \tag{6.31}$$

where the quantile $q_{0.95}$ having $a = 0.95$ corresponds to the c-quantile representing the fact that with a probability of 0.05, a loss larger than q_c itself can incur (indeed, $a = 0.95 \implies c = 0.05$). What appears clear is that, supposing to know $q_{0.5}$ and $q_{0.05}$ and to take the mean fixed, then the relation between the variance and the c-quantile with $c = 0.05$ has a parabolic trend: $\sigma^2 = cost. + q_c^2 - 2\mu q_c$. Since we are looking for a quantile which is likely above μ, one can notice that the quantile is proportional to the variance, which confirm our previous intuition.

6.3.2 Diversification

Equation 6.26 envisages the behaviour of the loss distribution of the entire portfolio, with losses following a one factor Bernoulli Mixture model. Indeed, the quantile of the loss distribution

$$q_c(L^{(n)}) \approx n \cdot \bar{l}(q_c(\Psi)) \tag{6.32}$$

increases linearly in the size of the portfolio n. This means that when the portfolio becomes asymptotically large the diversification has no further effects. Usually, increasing the number of securities in a portfolio diminishes the risk and then the losses. When n is large one cannot improve the situation: the losses are proportional to the size times a function which is constant once the scenario (i.e. the distribution of Ψ or Q) is known. One can achieve the same result by considering Equ. 6.25 as an equation describing the loss in a $1/n$-portfolio. Given the linearity of the expected value, we can take the $1/n$ inside and ri-define the exposures e_i, hidden in L_i as $e_i' = \frac{e_i}{n}$. Now, if we consider the correlation fixed and positive, by increasing the number of bonds, we are reducing the individual investment. In other words, what we are doing is to further diversify the portfolio. However, the portfolio loss is not subject to any kind of risk-reduction: even with positive correlation, in the limit of large n having one bond more does not improve the portfolio reducing its c-quantile. A different interpretation of the Equ. 6.32 heads to find a quantitative connection between the correlation and a very diversified portfolio (i.e having n large). Imagine to fix n, but to increase the correlation. We notice that, as shown in the previous section, increasing ρ the c-quantile enlarges and, by consequence, the tail of the loss distribution swells. Considering the same level of confidence c, but increasing ρ, implies, thus, a larger quantile, meaning that the portfolio suffers a greater loss than with a lower correlation. Hence, when the correlation grows up at fixed n, having a diversified portfolio loses any added-value.

6.4 Effect of a Policy shock on a portfolio: a Mixture Model interpretation

Finally, we simulate, throughout this simple exercise, how a Climate policy shock can impact the portfolio risk, in particular, analyzing the consequences on the default probability of the single bonds. The setting of the exercise follows the one at section 6.2.3; the only difference here is that we consider a Partially-Exchangeable Model formed by two categories: Green and Brown. For sake of simplicity we take one bond for each group. Then SETTING OF THE PROBLEM:

- Two different bonds: one *green* and one *brown*.
- The Exogenous factor is a binary random variable, modeling a Climate Policy.
- We consider two Scenarios, called *greenish* and *brownish*. The effect of this two scenarios on bonds' returns is different depending on the "color" of the bond.
- With this choice of scenario, the correlation comes to be negative, which in general is not a worrying result as there might be possible that the correlation may drop even under zero, in particular when we deal with securities of different nature.
- The Probability that a bond defaults Conditional to the external event is obtained throughout a *threshold* on the distribution of returns. Below the threshold, it is considered default.
- The Probability density function of the Returns conditional to the external event is supposed to be a Beta with parameters $\alpha(\Psi)$ and $\beta(\Psi)$. According to which scenario is realized, the Beta distribution is more flatten to the right or to the left. The greenish scenario is imagined to favor more the green than the brown; for this reason for the green bond the greenish scenario is positive and results in favorable states, On the other hand, for the brown the greenish scenario increases the probability of the bond to default. The situation is reversed is case the brownish scenario is realized.
- The outcome of the Policy is not green with the $p_{brownish}\%$ of probability.
- The Yellow color should represent the current situation of the individual bonds, since it is the unconditional behaviour.

In the following, exploiting this simple setting, we test the sensibility of the parameters under analysis: α and β of the two scenarios, the *threshold* establishes under which return value (or health state of the firm) there is the default. Different values of the parameters are assigned in each of the five cases, to investigate which and how the other elements of the model are affected, i.e. the correlation among the return, the correlation among the Bernoulli, and the probability of default. The equations underlying this simulation follow the ones presented at the beginning of this chapter, but generalized to the Partially Exchangeable case. Although it may seem quite self-explanatory, with the subscript G we refer to the green bond, while B represents the brown one. The exogenous factor causing the shock is a binary variable: $\Psi = G$ or $\Psi = B$, with probability $p_{greenish}$, $p_{brownish}$, respectively. As said, the Beta distributions, in this section, are exploited to find the effective Bernoulli probability of default. $Y_G = 1|\Psi = G$ means the event: default of green bond conditional to the realization of the greenish scenario.

$$P(Y_G = 1|\Psi = G) = p_G(\Psi = G) \tag{6.33}$$

then applying the threshold *threshold* and counting the number of points fall below one get the Conditional Default probability $p_G^{<t}$. The same procedure is repeated for all the other combinations

$$P(Y_G = 1|\Psi = B) = p_G^{<t}(\Psi = B) \tag{6.34}$$

$$P(Y_B = 1|\Psi = G) = p_B^{<t}(\Psi = G) \tag{6.35}$$

$$P(Y_B = 1|\Psi = B) = p_B^{<t}(\Psi = B) \tag{6.36}$$

Thus, for instance considering the individual default we have:

$$P(Y_G = 1) = P(Y_G = 1|\Psi = B) \cdot P(\Psi = B) + P(Y_G = 1|\Psi = G) \cdot P(\Psi = G) \tag{6.37}$$

$$P(Y_G = 1) = p_G^{<t}(\Psi = B) \cdot p_{brownish} + p_G^{<t}(\Psi = G) \cdot p_{greenish} \tag{6.38}$$

Then to compute the correlation inter-group, i.e. among one green and one brown, one has to compute:

$$P(Y_B = 1, Y_G = 1) = E[P(Y_B = 1, Y_G = 1|\Psi)] \tag{6.39}$$

$$\begin{aligned} P(Y_B = 1, Y_G = 1) = {} & P(Y_G = 1|\Psi = B)P(Y_B = 1|\Psi = B) \cdot P(\Psi = B)+ \\ & P(Y_G = 1|\Psi = G)P(Y_B = 1|\Psi = G) \cdot P(\Psi = G) \end{aligned} \tag{6.40}$$

$$\begin{aligned} P(Y_B = 1, Y_G = 1) = {} & p_G^{<t}(\Psi = B)p_B^{<t}(\Psi = B) \cdot p_{brownish}+ \\ & p_G^{<t}(\Psi = G)p_B^{<t}(\Psi = G) \cdot p_{greenish} \end{aligned} \tag{6.41}$$

Here again the limitation on the bounds of the correlation a has to be taken into account. In the Figure below the ρ displayed in the figure is correlation between the variable unknown before applying the threshold. However, the correlation among the Bernoulli is shown in the table.

CASE AND RELATIVE FIGURES		α PARAM. FOR BROWNISH SC. AND GREEN BOND	DEFAULT CORRELATION	THRESHOLD	$p_{brownish}$
0	Fig.6.17-6.18	4	-0.39	0.5	0.3
1	Fig.6.19-6.20	2	-0.37	0.3	0.65
2	Fig.6.21-6.22	2	-0.53	0.5	0.65
3	Fig.6.23-6.24	2	-0.58	0.5	0.3
4	Fig.6.25-6.26	2	-0.14	0.1	0.3

Table 6.1: **Mixture Model to assess the effects of a Policy shock on a portfolio** - Setting of the Exercise

Case0 well represents the impact of a shock, but does not adequately capture the reality. Observing real data, one can hardly think that today's brown bonds can default with 0.5 of probability (honey-colored column on the right of Figure 6.18) because of the unknowns about the future scenario. The situation modeled is of a policy shock that 7 times out of 10 is achieved by favoring green investments. This fact translates into a lower (but still high for securities) probability of default for green bonds, precisely 0.2%. Case1 is linked to a more unprofitable situation for green bonds because the probability of a greenish deal has dropped to 0.45 compared to 0.7 of Case0. Nevertheless, it reflects more the actual situation as the probability of default is approximately the same for both brown and green. This decrease is due to a lowering of the threshold from 0.5 to 0.3 and to an increase in the probability of the brownish scenario to 0.65. To better evaluate the single parameters, we change only one at a time, keeping the others fixed. This is what has been done from Case1 to Case4. From Case1 to Case2 the threshold is the only parameter which increases. This implies an increase in the absolute value of the correlation that can be explained by a shift of $E[p^{<t}(\Psi)]$ towards 0.5. So the question about uncertainty comes back. The default probability also increases but is due to a greater width of the default space because of the threshold. From Case2 to Case3, while the change in correlation is considered negligible, favoring the greenish scenario increases the unconditional default probability of brown bond, and diminishing the other one. Last, we inquire about changing the threshold but having $p_{brownish}$ fixed. What stands out moving from Case3 to Case4 is that the value for the default correlation among the default indicators is increased. However, this should not be surprising since the threshold drives the distribution of the exogenous factor. A lower correlation implies the existence of a scenario such that conditional to that, the default is more probable than conditional to other scenarios. To conclude, comparing the results, it emerges that the value of the probability distribution of the scenarios ($p_{brownish}/p_{greenish}$) is what actually impacts the unconditional default probability of a single bond. Besides, the correlation is more affected by how much the threshold is shifted to the right (or

left). Indeed, it can be argued that the threshold in this experiment behaves like the distribution of the mixing variable, in the case of an Exchangeable group at the beginning of this Chapter. However, if before we had a Beta distribution, and therefore a continuous scenario, now we deal with a finite number of scenarios, precisely two.

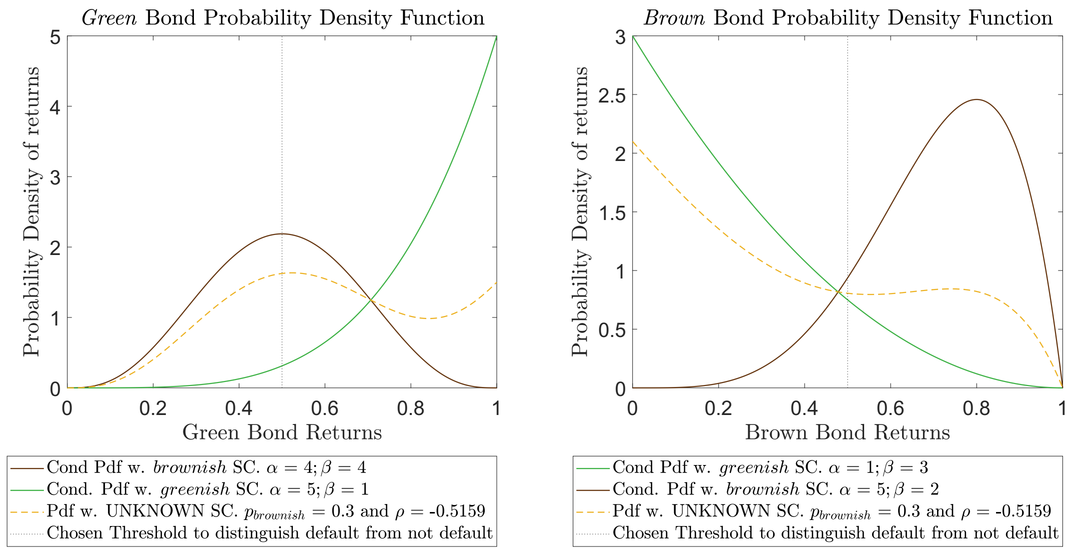

Figure 6.17: **Case 0 - Scenarios**

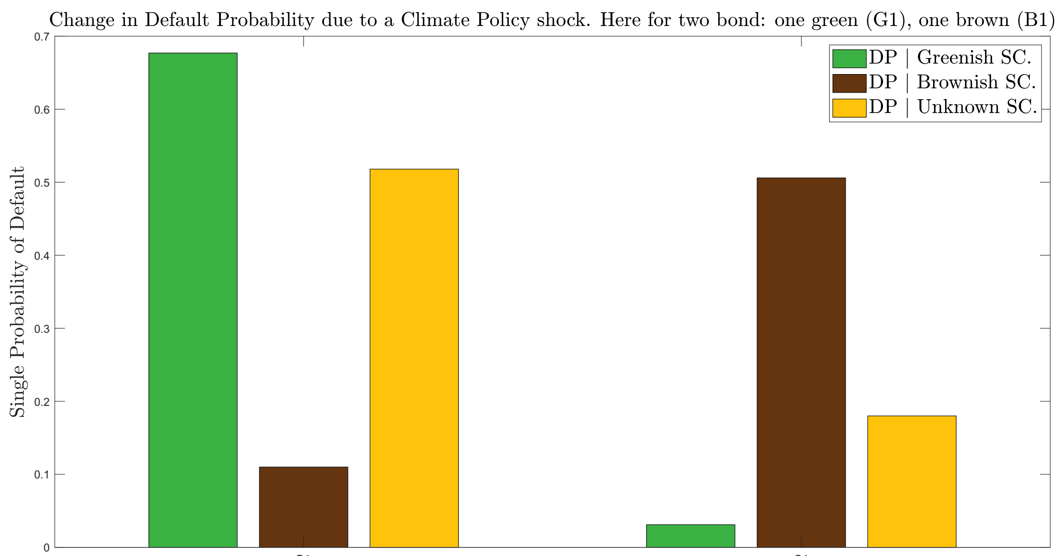

Figure 6.18: **Case 0 - Default** Probability shock

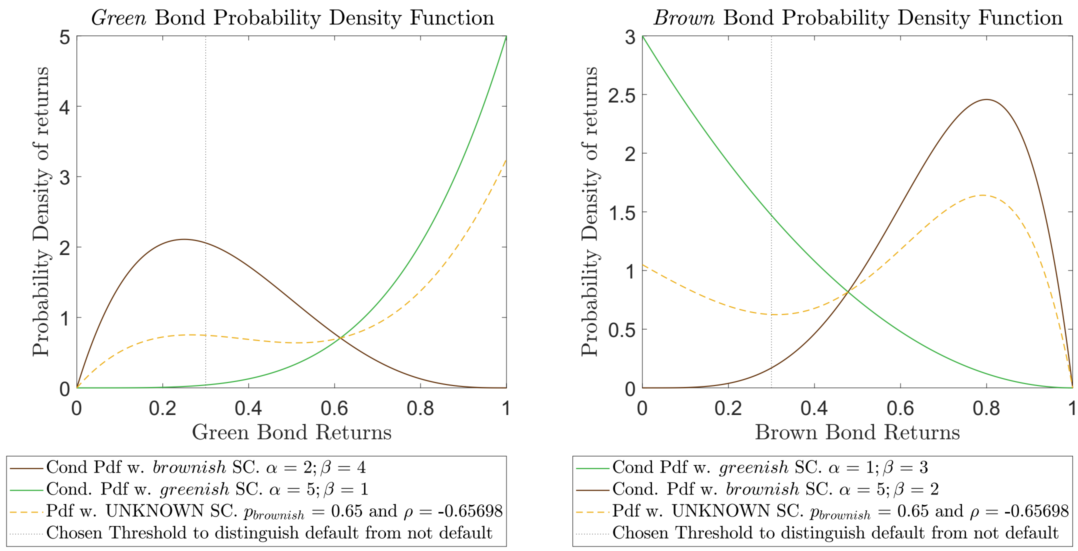

Figure 6.19: **Case 1 - Scenarios**

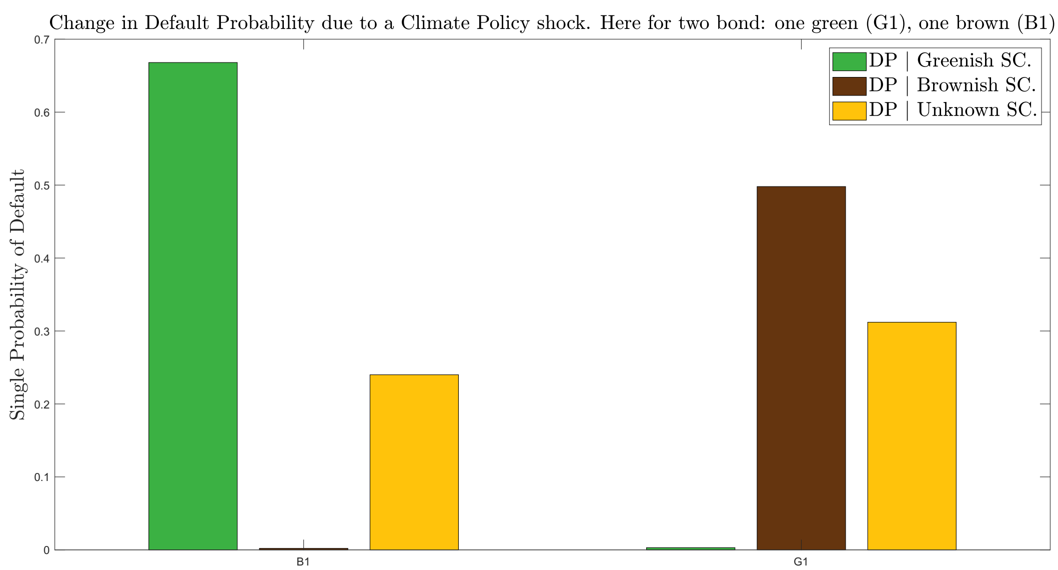

Figure 6.20: **Case 1 - Default** Probability shock

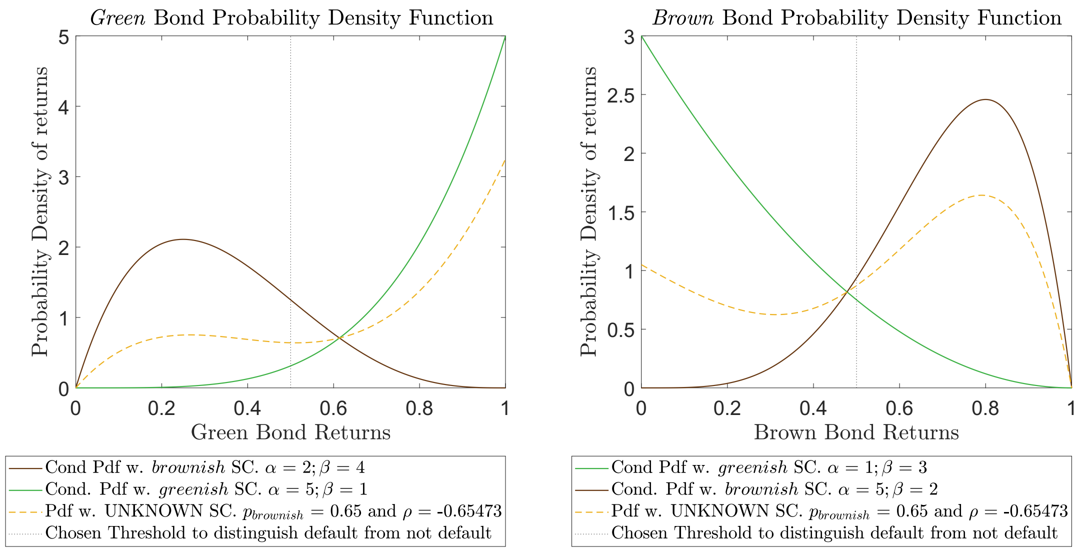

Figure 6.21: **Case 2 - Scenarios**

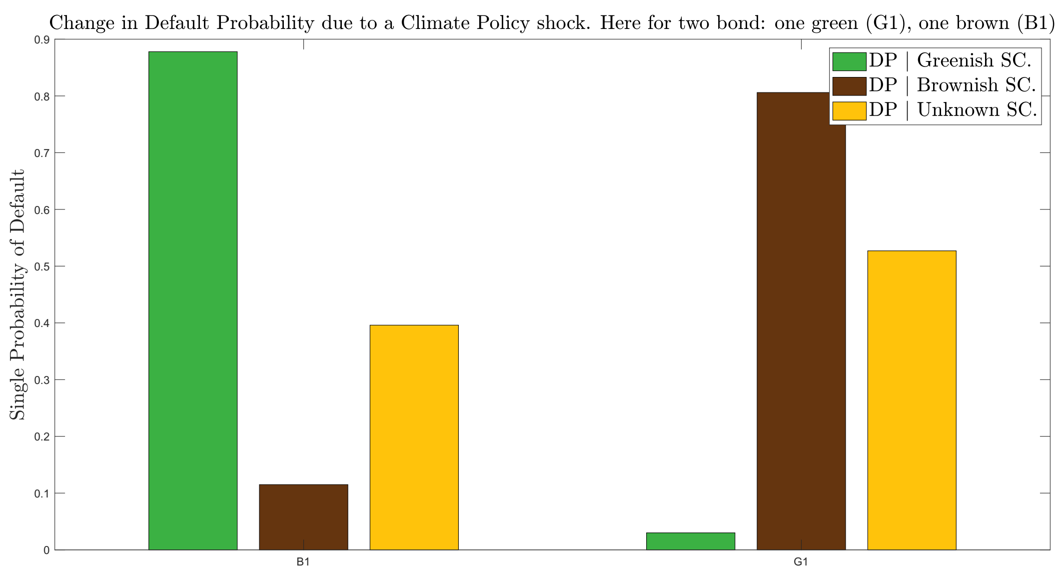

Figure 6.22: **Case 2 - Default** Probability shock

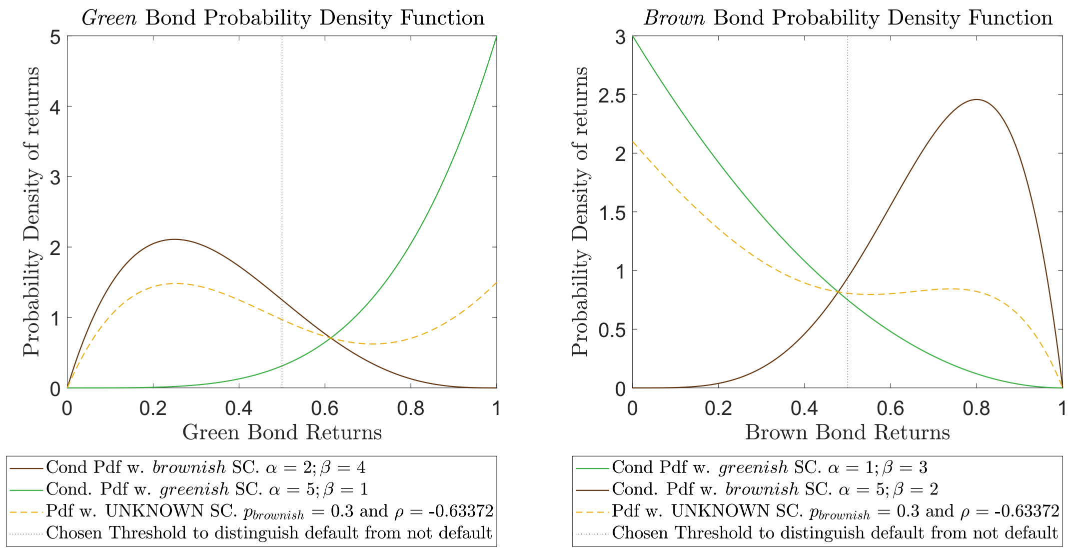

Figure 6.23: **Case 3 - Scenarios**

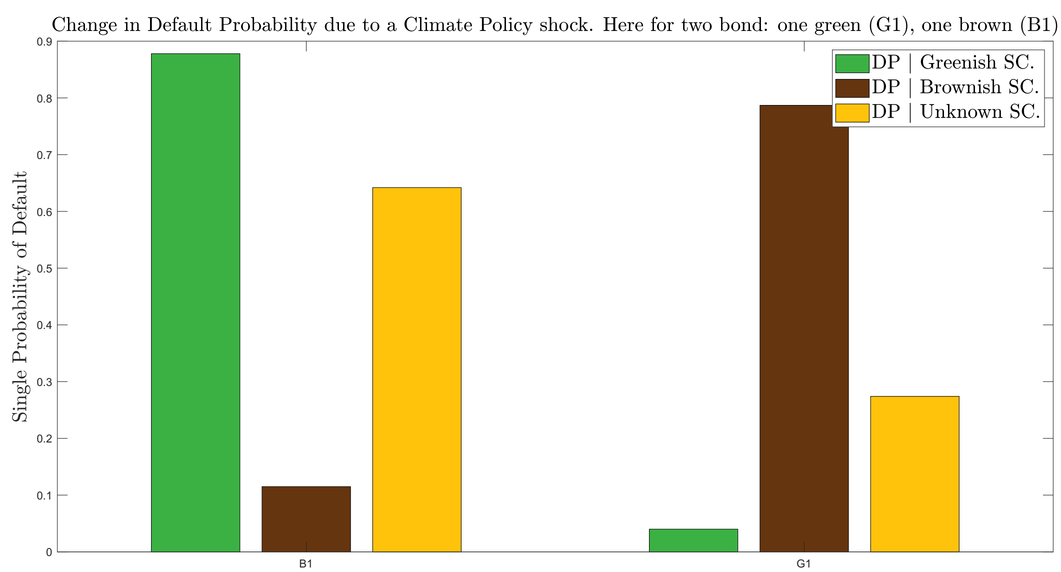

Figure 6.24: **Case 3 - Default** Probability shock

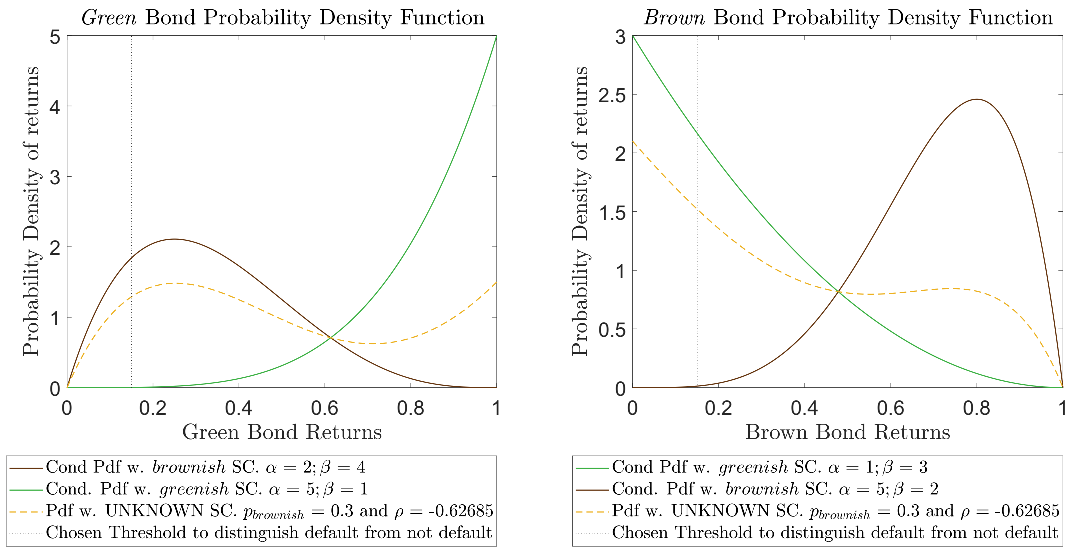

Figure 6.25: **Case 4 - Scenario**

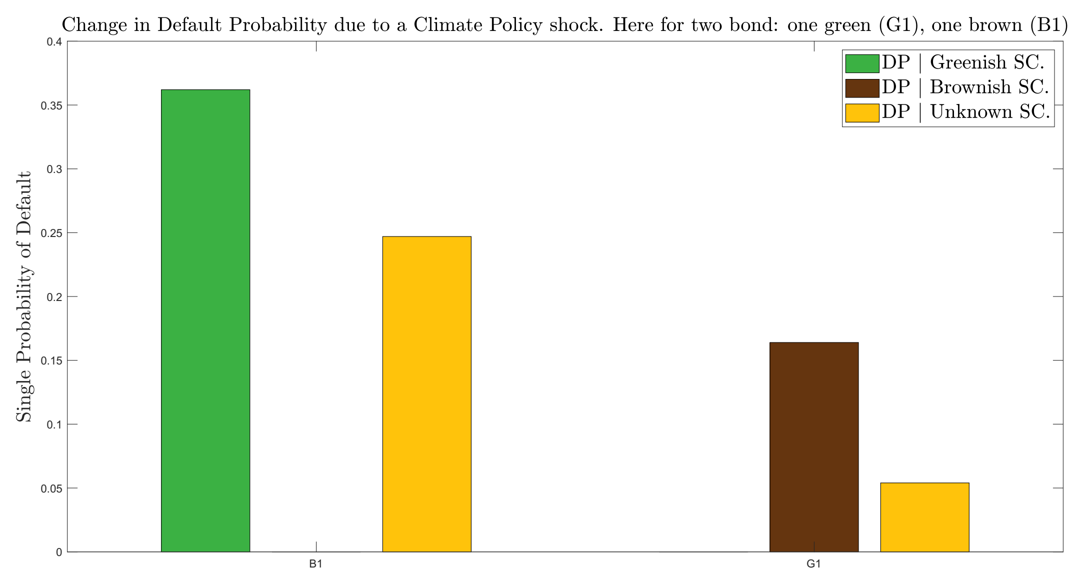

Figure 6.26: **Case 4 - Default Probability shock**

6.5 Model Limitations and Alternatives

In the context of Sustainable Finance, it is crucial to model future events characterized by uncertainty to assess their impact on current events and trends. We wish to address this purpose by choosing a model that allows us to obtain analytical results. One option is the Mixture Model, which has been presented in this section. After having leveraged its useful properties, in the following, we show some of its limitations. The first limitation of Mixture Model approach is a direct consequence of the underlying basic hypothesis. The Mixture Model assumes conditional independence of defaults given common stochastic factors. This means that once the exogenous factor is realized, defaulting bonds are independent: the correlation is zero. Since we aim at studying how the correlation changes depending on the severity of the scenario that appears, (think for instance to the agreement reached in Paris in December 2015) the assumption of conditional independence is tightening for us. In the real market, the correlation does not disappear after a climate-related event has occurred. Hence, with this model, we can only study how the correlation arises from one or more exogenous factors and not how it changes after the realization of those factors. The second limitation we faced is that a closed mathematical form is obtainable only in the simplest case of a homogeneous, i.e. exchangeable, model. Indeed, in the case of an Exchangeable Bernoulli Mixture Model, the default correlation takes values between 0 and 1. This limits our freedom in the portfolio composition. As said, analytically we can only address the situation of n bonds of the same color, which means belonging to the same category. So, we are prevented from analyzing the correlation between two different bonds, let's say one *green* and one *brown*, born from the uncertainty outcome of an external factor. To manage this situation we could exploit a partially-exchangeable model, where the bonds within the same exchangeable group share the risk and behavior with respect to the scenario realized. However, this variation needs a numerical simulation to be performed. Another difficulty arises when one has to grasp the identity of the external factor. The model assumes the existence of a latent, unobserved variable that impacts the distribution of the return of the assets in the portfolio. It may be a macroeconomic variable, an indicator that influences the performance of the economy in general and which, depending on its outcome, may produce a shock over the financial variables under analysis. Hence, being in the Mixture model framework, we assume that the bonds are dependent only because influenced by the same Climate policy or agreement (e.g. the Paris Agreement). However, although it seems a reasonable choice, we have no evidence of that. This leaves open the possibility to make different choices and to interesting research questions, that some recent articles have been started to address. In this regard, we report some results of a study conducted by Broadstock and Cheng (2019). Even though not exhaustive, it gives an overview of what are actually the macroeconomic variables that, belong their analysis, cause most of the correlation between green and brown bonds. The authors first pursue to test if there exists some links among the estimated dynamic correlation patterns and the exogenous factor. In particular, among these external factors they identified:

- changes in financial market volatility;
- economic policy uncertainty;
- daily economic activity;
- oil prices;
- uniquely constructed measures of positive and negative news-based sentiment towards green bonds;

where each one has been represented by several benchmark indices. One of the most relevant results for us regards the fact that from mid-2016 on, the macroeconomic variables OIL and ADS (that is a measure of daily economic activity) start to affect more and more the correlation between green and conventional bond price benchmarks. Moreover, another important point to clarify is what kind of shock one can portray with this model. We may define *shock* in two ways: the common shock and the idiosyncratic shock. We hypothesize that through the Mixture Model we can represent mainly the latter: a shock defined as a deterministic event but how it is implemented and the reactions of the stakeholders are uncertain and of different prominence. Through the Mixture Model, one may not model a shock, understood as an uncertain event that might also not happen. For this reason, trying to model an event such as the Paris Agreement may be done, but only under certain conditions. Besides the incongruities for the correlation arise in the attempt of modeling the event "PA signed or Not", there is also the fact that "sign or not sign" means that the event may also not happen, which is a perfect example of a common shock and as said cannot be tackle. Rather, the shock that can result from an external event be seen as a variable that controls the performance of the economy, which can be associated with the opinion of investors or economic policies and can result in idiosyncratic shocks on the profitability of securities in the portfolio. One has not only to think of an event linked to climate, but also of an international political event that affects the entire financial market, such as the Brexit. The latter is a perfect example of a deterministic shock, that everyone knew it was coming, but whose consequences caught unprepared the actors involved, generating idiosyncratic shocks. One last detail concerning the fact that we have to make assumptions about the distribution of defaults that may appear to come from above. As seen in the previous section, it was difficult to assign appropriate values to α and β being that they seem more like two empty parameters. To tackle this problem, one way can be exploiting a regression structure to study real market data in order to make a reasonable hypothesis on that.

In the following we show two Tables: in one we report the identified limitations and related effects, in the other the advantages of the Mixture Model within our context.

LIMITATIONS	EFFECTS
Assumption of conditional independence.	No analysis on how the correlation changes according to an exogenous event.
Analytical evaluation only in the simplest case of One factor Exchangeable Bernoulli Mixture Model.	Less freedom in the portfolio composition. We can only address the situation of n bonds of the same "colour", that means belonging to the same category.
Existence of a latent, unobserved variable that impacts the distribution of the portfolio.	Difficult to give Ψ an identity. PA cannot be the suitable choice since it is: too less general; difficult making assumption about the probability distribution of default conditional on PA; difficult assuring that, conditional on PA, bond are independent.
Shock, defined as uncertain event, is not modelled. Shock happens, what is uncertain are its implementation and the reactions of the stakeholders.	If Ψ is deterministic, it does not model a *Yes/No* situation

Table 6.2: **Limitations** of Mixture Model method

OPPORTUNITIES
Easy way to approach a correlated portfolio linking the default correlation with an exogenous factor.
Significant results find in the asymptotic limit: portfolio loss is driven by the tail of the mixing variable; the effect of diversification is reduced
Numerically one can represent also an idiosyncratic shock, caused by the certain introduction of a Policy whose effect are uncertain (see Section 6.4). It represents a super simplified version of the reality, characterizing the economy trough two stages. • Time 1: before the factor is realized; • Time 2: after the realization of the factor; green and brown companies suffer a productivity shock that is modulated by the outcome of policy introduced (exogenous factor). With shock in productivity we intend that the single probability of default increase or decrease depending to the result of the factor.
Through statistical inference approach, one can exploit mixture model to fit data, estimating the parameters of the distributions from which the data might come, as well as the probability of coming from each of the mixed distributions. In this way, important but undefined and unobservable variables, such as the already mentioned state of the economy, may be quantified and estimated.

Table 6.3: **Opportunities** of Mixture Model

Chapter 7

Conclusions and Further Development

Managing the correlation is renewed to be a challenging task for the analytical investigation of portfolio risk. In the present work, we investigated, throughout a probabilistic approach, the role of correlation in current portfolio management. Among the methodologies used to model the dependence between the financial contracts, we exploited the Copula function and the Mixture Model. In particular, in Chapter 3 we tested the response of the optimal portfolio to different settings of portfolio parameters (correlation coefficient, default probability, and investor's risk-tolerance). Then, in light of a sustainable finance perspective, we inquired if a specific set of parameters could lead to a low-carbon transition. Indeed, although the low-carbon transition is a hot topic on the international agenda, to date, the status quo of investment portfolios is still largely focused on conventional assets, mainly related to carbon-intense activities. Currently, since the investor with low risk-return profiles seems to dominate the bond market and green contracts are seen as riskier, to invest in green securities they require higher returns otherwise they are reluctant to invest. We modeled the bond as a binary variable, assigning, while evaluating the graphs in Chapter 3, a higher probability of default to the green bonds. The results of the simulations we carried out, shows that decreasing the level of correlation between the bond outcomes leads to a shift in the optimal allocation from a less to a more diversified allocation. Since green investments are often perceived as a novelty, which means higher risk is attached to them, a change in correlation can thus turn to shift capital from brown towards green projects, fostering the transition. The correlation among different securities could change due to several factors; we focused on the impact that external events or policies could have on it. To this end, in Chapter 6, we proceeded by exploiting the Mixture Model: the correlation comes again into play, but it is driven by an exogenous macroeconomic variable. By the use of this method, the correlation between bonds is introduced and controlled through an exogenous parameter, which describes multiple future scenarios, and in a few circumstances, analytic results can be obtained. Among the hypotheses outlined we discovered that the exogenous shock impacts both the portfolio risk, varying the expected losses, and the bond correlation. In particular, the correlation among an exchangeable (homogeneous, e.g only green) group of bonds arises from the shape of the probability distribution of the default conditional to the external factor. It does not have a strong dependence on the outcomes themselves but on the effect these outcomes have on the probability of default. Modeling a heterogeneous group (with both green and brown bonds) requires a numerical simulation since one cannot go further with the analytical computation. In this case, the correlation between green and brown is influenced explicitly also on the probability distribution of the exogenous factor. Besides, we discussed also the asymptotic limit of a portfolio with $n \to \infty$, proposing a way to link the correlation with the portfolio loss throughout the external factor. Then we discussed the limitation of the model we faced and the assumption we made to exploit the model. In conclusion, although little can be done analytically, it is clear the importance of considering the climate risk in terms of financial risk as it would affect the evaluation of the current investment portfolio, implying a different allocation. Future work can involve a statistical analysis of financial data. Although current data is complex and fragmented, processing market data is critical to figure out information on the actual value of the parameters involved. In particular, it would be interesting to assess whether an indirect climate-related external event such as the Covid-19 outbreak has been changing the dependence between green and brown bonds in a way to facilitate the transition or make it harder to achieve. Similarly, a study comparing the different shocks that have occurred in the economy (Paris agreement, Brexit, Covid-19) over the last 10 years could be carried out in order to understand their impact on correlation and probability of default and consequently on investments.

Bibliography

Battiston, S. and Monasterolo, I. (2019). A climate risk assessment of sovereign bonds. *Available at SSRN 3376218*.

Bernoulli, D. (1954). [originally published in 1738; translated in 1954 by Dr L.Sommer] Exposition of a New Theory on the Measurement of Risk. *Econometrica*, 22(1):23–36.

Broadstock, D. C. and Cheng, L. T. (2019). Time-varying relation between black and green bond price benchmarks: Macroeconomic determinants for the first decade. *Finance Research Letters*, 29(C):17–22.

Brunnermeier, M. K. (2015). Pricing models. *FIN501 Asset Pricing I*.

Chiarella, C., Dieci, R., and He, X.-Z. (2008). Heterogeneity, market mechanisms, and asset price dynamics. *SSRN Electronic Journal*.

Donnelly, C. and Embrechts, P. (2010). The devil is in the tails: actuarial mathematics and the subprime mortgage crisis. *Astin Bulletin*, 40(1):1–33.

Frey, R. and McNeil, A. J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6:59–92.

Hanoch, G. and Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, 36(3):335–346.

Haught, M. (2016). An introduction to copula. *IEOR E4602: Quantitative Risk Management*.

Kroll, Y., Levy, H., and Markowitz, H. M. (1984). Mean-variance versus direct utility maximization. *The Journal of Finance*, 39(1):47–61.

Leonov, S. and Qaqish, B. (2017). Correlated endpoints: simulation, modeling, and extreme correlations. *Statistical Papers*, pages 1–26.

Levy, H. (2015). *Stochastic dominance: Investment decision making under uncertainty*. Springer.

Levy, H. and Markowitz, H. M. (1979). Approximating expected utility by a function of mean and variance. *The American Economic Review*, 69(3):308–317.

Markowitz, H. (1952). "portfolio selection". *Journal of Finance*, 7(1):77–91.

Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments*, volume 16. John Wiley New York.

Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, 234(2):346–355.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative risk management: concepts, techniques and tools*. Princeton university press.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York, NY, USA, second edition.

Pearson, E. S. and Tukey, J. W. (1965). Approximate means and standard deviations based on distances between percentage points of frequency curves. *Biometrika*, 52(3/4):533–546.

Roch, S. (2015). Modern discrete probability, chapter 4 - coupling.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. a definition. *Journal of Economic theory*, 2(3):225–243.

Schmid, F. and Schmidt, R. (2007). Multivariate conditional versions of spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis*, 98(6):1123–1140.

Schweizer, B. (1991). Thirty years of copulas. In *Advances in probability distributions with given marginals*, pages 13–50. Springer.

Schweizer, B. (2007). Introduction to copulas. *Journal of Hydrologic Engineering*, 12(4):346–346.

Sklar, A. (1996). *Random variables, distribution functions, and copulas—a personal look backward and forward*, volume Volume 28 of *Lecture Notes–Monograph Series*, pages 1–14. Institute of Mathematical Statistics, Hayward, CA.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231.

Von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton University Press.

Whitt, W. (1976). Bivariate distributions with given marginals. *Ann. Statist.*, 4(6):1280–1289.

Whitt, W. (2014). Stochastic ordering: Theory and overview. *Wiley StatsRef: Statistics Reference Online.*

Acknowledgments

This project closes an important page of my life. Carrying out this Master was a challenge, likely not the best choice I could make for myself. However, despite all the ups and downs, I lived unique experiences on which I will leverage to shape my near future.

In this regard, I would like to mention those who accompanied me during my thesis, first of all, my supervisor at the University of Zurich Prof. Stefano Battiston, and immediately afterward Alan Roncoroni, my spiritual guide, and his assistant. I thank them for the suggestions and reflections, emerged from the numerous (and unfortunately mainly virtual) conversations, and even more for giving me the opportunity to discover the fascinating world of Sustainable Finance.

I would also like to thank for his support and patience, Prof. Alfredo Braunstein, my supervisor at the Politecnico di Torino, and Prof. Marcello Restelli, co-supervisor of Politecnico di Milano.

A special thanks go to my colleagues and friends, who helped me to complete this course of study. I wish to thank my old friends Silvia, Sara, Michela, and Beatrice, that are still there despite the distance and different life choices that have been keeping us apart so many times over the years.

I would like then to thank all my classmates, "quelli forti" as I liked to call them, with whom I felt really free to be myself and never felt judged.

Among them, a special mention goes to Chiara, for friends "Giovanna", my study partner; for having shared some of her determination in facing difficulties with me, and who, among several "bibbiette" and some "SISSA-blackboards printed in color", has also become a special friend. Another pillar that I cannot fail to mention is Valeria, my quasi stepsister. So equal on paper, but so different in real term, thanks to her these years at Politecnico have taken a different tone. Thank you for being so good at understanding me and for your support in the darkest moments.

Then I thank Davide (probably more than a huge thank you he has to be sanctified) for sharing with me the joys and dramas of these two years. For being there when even I was not able, trying to reason with me, and encouraging me never to give up.

Finally, a special thank you goes to Alberto, my mentor (even if he doesn't know it), for every advice in front of a glass of wine, every English lesson, and to always remind me that studies are not everything.

I conclude by expressing my heartfelt thanks to my family, in particular to my mother for being so patient, my sister, for that hour a day she is nice (I am joking Rachele having a sister is one of the most exciting things that could happen to me) and my father, for not having disinherited me yet. I sincerely hope that, despite all my countless faults, I have been able to make you proud of me along these past years, and whatever the future holds for me I will keep doing it.