

DOTTORATO DI RICERCA IN METODI MATEMATICI PER L'ECONOMIA,  
L'AZIENDA, LA FINANZA E LE ASSICURAZIONI

# An analysis of credit risk financial indicators

DOTTORANDA  
Sara Cecchetti

DIRETTORE DI TESI  
Prof.ssa Giovanna Nappo

COORDINATORE DEL DOTTORATO  
Prof. Fausto Gozzi

XXII CICLO

# Contents

<b>1</b>	<b>A Dynamic Default Dependence Model</b>	<b>7</b>
1.1	Introduction . . . . .	7
1.2	Credit Derivatives . . . . .	9
1.2.1	Credit Default Swap . . . . .	9
1.2.2	Portfolio Credit Derivatives and CDOs . . . . .	9
1.3	Mathematical Preliminaries . . . . .	11
1.3.1	Lévy subordinator . . . . .	11
1.3.2	Time normalization condition . . . . .	13
1.4	The Multivariate Default Model . . . . .	13
1.4.1	Heterogeneous case with $r$ different homogeneous classes . .	19
1.5	Applications . . . . .	29
1.5.1	Pricing CDO tranches . . . . .	29
1.5.2	The Lévy Subordinator . . . . .	31
1.5.3	The calibration to iTraxx quotes . . . . .	33
1.6	Distress Dependence and Systemic Risk . . . . .	35
1.6.1	The Stability Index . . . . .	36
1.6.2	The Distress Dependence Matrix . . . . .	36
1.6.3	The Probability of Cascade Effects . . . . .	37
1.7	Conclusions . . . . .	38
A	Appendix . . . . .	39
<b>2</b>	<b>On the Relationship between the Risk of Default and the Yield-to-Maturity of Bonds</b>	<b>41</b>
2.1	Introduction . . . . .	41
2.2	The mathematical model . . . . .	43

2.3	Numerical examples . . . . .	45
2.4	Conclusion . . . . .	46
A	Appendix (proofs) . . . . .	47
	References . . . . .	53

# Introduction

The recent financial crisis has highlighted the importance of one of the main components of the financial risk: the credit risk. Of particular interest in finance has become the modeling of credit risk for measuring portfolio risk and for pricing securities exposed to credit risk, as defaultable bonds and credit derivatives.

In this thesis we focus on modeling the aggregated risk of a portfolio, and on studying the relationship between default risk and yield rates for defaultable bonds. In particular we analyze financial indicators of the default risk of a portfolio of credit risky assets and of a bond.

This first chapter aims at contributing to the literature on the pricing of portfolios of credit derivatives (such as CDOs or basket CDS) where the goal is to compute the joint probability of default of a portfolio of risky assets. In this regard, the risk of default of each asset in the portfolio depends on mostly two sources of randomness: an individual risk factor and a common market factor. The latter represents the uncertainty affecting all assets simultaneously. Our scope is to model the aggregate portfolio risk. Such main theme is fundamental in finance both for the valuation of many credit derivatives and for extracting information from market prices that can be relevant from a macro-prudential point of view (such as estimating joint probabilities of default or probabilities of default conditional on other assets being in default). The recent international financial crisis has just highlighted the missing of correct models for valuating credit derivatives as CDO. From the theoretical point of view we develop a dynamic multivariate default model. A recent paper of Mai and Scherer (2008) uses a stochastic time change to introduce dependence in a portfolio of credit risky assets. In that paper the default times are modeled as random variables with possibly different marginal distribution. By restricting the time change to suitable Lévy subordinators the authors can separate the dependence structure and the marginal default probabilities. Using a so-called time normalization they compute the survival copula of all default times. In order to compute the portfolio loss distribution and to apply their model to the pricing of CDO tranches, an homogeneous portfolio is assumed, in which all default times share the same marginal distribution. Our model develops the ideas of Mai and Scherer (2008) by introducing the case of different

marginal distributions for the different assets in the portfolio; we aim at introducing heterogeneity in the model by allowing for an heterogeneous portfolio, as in the implied copula model of Hull and White (2009). In particular we define and model a cumulative dynamic hazard process as a Lévy subordinator, which allows for jumps and induces positive probabilities of joint defaults, and we model the dependence structure by the implied survival copula, which is related to the choice of the subordinator. In our model we allow the main asset classes in the portfolio to have different cumulative default probabilities and corresponding different cumulative hazard processes. We find an analytical closed formula for the distribution of the portfolio-loss process under this heterogeneous assumption, and we prove an approximation formula for the loss distribution that is useful for empirical applications. Once we have specified a suitable Lévy subordinator, our model can be calibrated to portfolio-CDS spreads and CDO tranche spreads, properly choosing the model parameters that determine the dependence structure. Once we have developed the model for the heterogeneous case we may use it to determine the possible testable implications that our theoretical model for the dependence of the joint defaults has for the characteristics of the CDS prices. From an empirical point of view we calibrate the parameters of our model to the tranches of the iTraxx Europe, which is a basket of 125 CDS on the European firms. Once we have estimated the multivariate default distribution of the companies included in the iTraxx we can follow Segoviano and Goodhart (2010) and analyze the distress dependence in the portfolio computing indicators of systemic risk by estimating a set of stability measures that incorporate changes in distress dependence that are consistent with the economic cycle. Examples of these stability measures are the Stability Index (that reflects the expected number of firms becoming distressed given that at least one firm has become distressed), the Distress Dependence Matrix (in which we estimate the set of pairwise conditional probabilities of distress providing some insights into inter-linkages and likelihood of contagion between the firms), and the Probability of Cascade Effects (that characterizes the likelihood that one or more institutions becomes distressed given that a specific firm becomes distressed). These stability measures can be used to verify which firms are more systemically relevant for the index as a whole.

In the second chapter we study the relationship between the risk of default and the yield-to-maturity of a bond. Credit Default Swap (CDS) spreads and bond spreads (i.e. the spreads between the bond yield rate and the risk free rate) have become commonly used as default risk indicators for risk analysis (see for example Bank of England, 2009, and Fitch, 2010a, 2010b). However, both CDS and bond spreads depend not only on variables directly linked to the risk of default but also on the specific structure of the contract. In particular, in this chapter we show that bond spreads can be misleading if used to infer the

default probability of the issuers, and consequently the yield-to-maturity must be cautiously interpreted as an indicator of the bond default risk.<sup>1</sup> In fact, the yield rate, for given default probabilities and recovery rates, can considerably vary as a function of the residual life and the coupon value of the bond. In particular, when there is default risk, bonds with high coupons are more likely to have high yield rates too (and viceversa). The intuition behind this result is that greater coupons are associated with less than proportional price increases, because there is a probability, linked to the default likelihood, that the coupons are not actually payed-off. Bond prices which are relatively low with respect to the nominal payment flows (on which the yield is computed) determine nominal yields which are relatively higher. This implies that bonds with higher default risk can have lower yields (and viceversa), just as a consequence of their coupon structures. Also the slope of the yield curve must be cautiously interpreted as in general it does not convey enough information to establish if the default risk is higher in one period than in another period. We show that a downward sloping yield curve does not necessarily imply that the default probability on shorter maturities is higher than on longer maturities. This result arises from the fact that also the yield curve slope is linked to the coupon rate: taking fixed the other variables (in particular the default probability and the recovery rate) bonds with low coupons determine decreasing yield curves, while bonds with high coupons imply increasing yield curves. The intuition behind this result is that higher coupons determine losses relatively higher for bonds with longer maturity in case of default, and consequently higher yields for these bonds. On the contrary, when the coupons are low, the nominal losses in case of default are similar both for short term and long term bonds;<sup>2</sup> it follows that the prices of longer maturity bonds, for which the losses are likely in far away time horizons, are relatively higher, and the corresponding yields are lower. Most of the literature available on the valuation of fixed income securities (see for example Fabozzi, 2003, 2007), is focused on the interest rate risk (i.e. the bond's price sensitivity to the change in interest rates) and the concept of duration is used to describe the relationship between the bond maturity and coupon rate, and the bond price sensitivity (a longer maturity and a lower coupon rate are linked to a greater price sensitivity to interest rate changes); in this context a higher bond yield is considered as a premium for the higher interest rate risk. In this chapter we study similar financial indicators that could be used in presence of credit (or default) risk to properly evaluate the relationship

---

<sup>1</sup>In particular, without loss of generality, we assume that the risk-free rate is constant in the considered time horizon, so that the yield level provides information analogous to the spread between the yield rate and the risk-free rate. Anyway the inferred considerations are still valid, on a quality level, even in case the risk-free rate curve is not flat.

<sup>2</sup>For the zero-coupon bonds, for example, in case of default the loss is always only given by the amount lost on the bond final pay-off.

between the defaultable bond yield and its default probability.

# Chapter 1

## A Dynamic Default Dependence Model

### 1.1 Introduction

This paper aims at contributing to the literature on the pricing of portfolios of credit derivatives (such as CDOs or basket CDS) where the goal is to compute the joint probability of default of a portfolio of risky assets. In this regard, the risk of default of each asset in the portfolio depends on mostly two sources of randomness: an individual risk factor and a common market factor. The latter represents the uncertainty affecting all assets simultaneously. Our scope is to model the aggregate portfolio risk. Such main theme is fundamental in finance both for the valuation of many credit derivatives and for extracting information from market prices that can be relevant from a macro-prudential point of view (such as estimating joint probabilities of default or probabilities of default conditional on other assets being in default). The recent international financial crisis has just highlighted the missing of correct models for valuating credit derivatives as CDO.

From the theoretical point of view we want to develop a dynamic multivariate default model. A recent paper of Mai and Scherer (2008) uses a stochastic time change to introduce dependence in a portfolio of credit risky assets. In that paper the default times are modelled as random variables with possibly different marginal distribution. By restricting the time change to suitable Lévy subordinators the authors can separate the dependence structure and the marginal default probabilities. Using a so-called time normalization they compute the survival copula of all default times. In order to compute the portfolio loss distribution and to apply their model to the pricing of CDO tranches, an homogeneous portfolio is assumed, in which all default times share the same marginal distribution. Our model develops the ideas of Mai and Scherer (2008) by introducing the case of



different marginal distributions for the different assets in the portfolio; we aim at introducing heterogeneity in the model by allowing for an heterogeneous portfolio, as in the implied copula model of Hull and White (2009). In particular we define and model a cumulative dynamic hazard process as a Lévy subordinator, which allows for jumps and induces positive probabilities of joint defaults, and we model the dependence structure by the implied survival copula, which is related to the choice of the subordinator. In our model we allow the main asset classes in the portfolio to have different cumulative default probabilities and corresponding different cumulative hazard processes. We find an analytical closed formula for the distribution of the portfolio-loss process under this heterogeneous assumption, and we prove an approximation formula for the loss distribution that is useful for empirical applications.

Once we have specified a suitable Lévy subordinator, our model can be calibrated to portfolio-CDS spreads and CDO tranche spreads, properly choosing the model parameters that determine the dependence structure. Once we have developed the model for the heterogeneous case we may use it to determine the possible testable implications that our theoretical model for the dependence of the joint defaults has for the characteristics of the CDS prices.

From an empirical point of view we calibrate the parameters of our model to the tranches of the iTraxx Europe, which is a basket of 125 CDS on the European firms. Once we have estimated the multivariate default distribution of the companies included in the iTraxx we can follow Segoviano and Goodhart (2010) and analyze the distress dependence in the portfolio computing indicators of systemic risk by estimating a set of stability measures that incorporate changes in distress dependence that are consistent with the economic cycle. Examples of these stability measures are 1) the Stability Index, that reflects the expected number of firms becoming distressed given that at least one firm has become distressed; 2) the Distress Dependence Matrix, in which we estimate the set of pairwise conditional probabilities of distress providing some insights into inter-linkages and likelihood of contagion between the firms; 3) the Probability of Cascade Effects that characterizes the likelihood that one or more institutions becomes distressed given that a specific firm becomes distressed. These stability measures can be used to verify which firms are more systemically relevant for the index as a whole.

## 1.2 Credit Derivatives

A credit derivative is a derivative security that is primarily used to transfer, hedge or manage credit risk. The credit derivative payoff is conditioned on the occurrence of a credit event. The credit event is defined with respect to a reference credit (or several reference credits), and the reference credit asset(s) issued by the reference credit. If the credit event has occurred, the default payment has to be made by one of the counterparties. A credit event is a precisely defined default event, such as bankruptcy, failure to pay, obligation default, repudiation/moratorium, ratings downgrade below a given threshold, changes in the credit spread. A default payment is the payment which has to be made if a credit event has happened.

### 1.2.1 Credit Default Swap

In a single-name credit default swap (CDS) the default seller B agrees to pay the default payment to the default buyer A if a default has happened. The default payment is structured to replace the loss that a typical lender would incur upon a credit event of the reference entity. If there is no default of the reference security until the maturity of the default swap, counterparty B pays nothing. On the other side the default buyer A pays a fee for the default protection. In the most common version the fee is paid at regular intervals until default or maturity. If a default occurs between two fee payment dates, A still has to pay the fraction of the next fee payment that has accrued until the time of default.

### 1.2.2 Portfolio Credit Derivatives and CDOs

When dealing with a portfolio, we must consider the risk of a clustering of defaults and of joint defaults. Basket and portfolio credit derivatives are instruments used to manage risks of this type.

Collateralised Debt Obligations (CDOs) are financial products to securitise portfolios of defaultable assets: loans, bonds or credit default swaps. The assets are sold to a specific company and investors are offered the opportunity to invest in notes issued by this company. These obligations are collateralized by the underlying debt portfolio. The different notes are structured in order to offer risk/return profiles that are specifically targeted to the risk appetite and investment restrictions of different investors groups. A simple CDO has the following components:

- The underlying portfolio is composed of defaultable assets issued by issuers  $C_i$  with notional amounts  $K_i$ ,  $i = 1, \dots, I$ . The total notional is  $K = \sum_{i=1}^I K_i$ .

- The portfolio is transferred into a special created company, the special purpose vehicle (SPV).
- The SPV issues notes:
  - an equity (or first-loss) tranche with notional  $K_E$ ;
  - several mezzanine tranches with notional  $K_{M_1}$ ,  $K_{M_2}$ ,  $K_{M_3}$ , etc.;
  - a senior tranche with notional  $K_S$ .
- If during the existence of the CDO one of the bonds in the portfolio defaults, the recovery payments are reinvested in default-free securities.
- At maturity of the CDO, the portfolio is liquidated and the proceeds are distributed to the tranches, according to their seniority ranking.

The key point of the CDO is the final redistribution of the portfolio value according to the seniority of the notes. First, the senior tranche is served. If the senior tranche can be fully repaid, the most senior mezzanine tranche is repaid. If this tranche can also be fully repaid, then the next tranches are paid off in the order of their seniority, until finally the equity tranche is paid whatever is left of the portfolio's value. The payoffs are function of the losses.

1. The first losses hit the equity tranche alone. Until the cumulative loss amount has reached the equity's notional  $K_E$ , the other tranches are protected by the equity tranche.
2. Cumulative losses exceeding  $K_E$  affect the first mezzanine tranche, until its notional is used up.
3. After this, the subsequent mezzanine tranches are hit in the order of their seniority.
4. Only when all other tranches have absorbed their share of the losses will the senior tranche suffer any losses.

In the standard CDOs the underlying portfolio can consist of bonds (collateralized bond obligation, or CBO) or loans (collateralized loan obligation, or CLO).

We have instead a synthetic CDO when credit default swaps are used instead of bonds or loans in the underlying portfolio.

Basically, once a CDO is constructed by partitioning the credit portfolio in tranches with different seniority, a tranche represents a certain loss piece of the overall portfolio which is defined by its lower and upper attachment points. The protection seller receives periodic premium payments depending on the remaining

nominal and the spread of this tranche, while the protection buyer is compensated for losses affecting this tranche. Pricing of a tranche corresponds to assessing the spread such that the expected discounted payment streams of this tranche agree.

## 1.3 Mathematical Preliminaries

### 1.3.1 Lévy subordinator

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. A one-dimensional Lévy process on this probability space is a càdlàg stochastic process  $\Lambda = \{\Lambda_t\}_{t \geq 0}$  starting at  $\Lambda_0 = 0$ , which is stochastically continuous and has independent and stationary increments. A *Lévy subordinator* is a particular Lévy process in which almost all paths are non-decreasing. A Lévy subordinator is completely characterized by two characteristics: a drift  $\mu \geq 0$  and a positive measure  $\nu$  on  $(0, \infty)$ . Basically  $\Lambda$  is a process that grows linearly with a constant drift and is affected by random upward jumps. The process drift is  $\mu \geq 0$ , while the expected number of jumps greater than or equal to  $x$  within a unit of time is given by the Lévy measure  $\nu$  of the interval  $[x, \infty)$ . In particular, if  $\Lambda_t$  is a Lévy subordinator with related Lévy measure  $\nu$ , then there exists  $\mu \geq 0$  such that

$$\Psi(\lambda) = \mu\lambda + \int_0^\infty (e^{\lambda t} - 1) \nu(dt), \quad \lambda \leq 0, \quad t \geq 0.$$

Remark that the function  $\Psi$  is with negative values,  $\Psi(0) = 0$  and, unless  $\Lambda_t \equiv 0$ , is strictly increasing. The function  $\Psi$  is the *Laplace exponent* of  $\Lambda$  that completely determines the process via its Laplace transform

$$\mathbb{E}(e^{\lambda \Lambda_t}) = e^{t\Psi(\lambda)}.$$

Three fundamental examples of Lévy subordinators, that we will use in this paper, are the following Lévy processes:

- the *Inverse Gaussian Subordinator*;
- the *Gamma Subordinator*;
- the *Compound Poisson Subordinator*.

#### The Inverse Gaussian Subordinator

The *Inverse Gaussian (IG) subordinator*  $\Lambda^{IG} = \{\Lambda_t^{IG}\}_{t \geq 0}$  belongs to the class of infinite activity subordinators, meaning that processes of this class jump infinitely often within a unit interval of time. The IG Lévy measure as well as the density of

the underlying infinitely divisible distribution are well known. In particular, given an IG subordinator with parameters  $\eta, \beta > 0$ , we have that  $\Lambda_t^{IG}$  follows an Inverse Gaussian  $IG(\beta t, \eta)$ -distribution with density

$$f_{IG}(x) = \frac{\beta t}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{\eta\beta t} e^{-\frac{1}{2}\left(\frac{\beta^2 t^2}{x} + \eta^2 x\right)} \mathbf{1}_{\{x>0\}}.$$

The corresponding Lévy measure is given by

$$\nu_{IG}(dx) = \frac{1}{\sqrt{2\pi}} \beta x^{-\frac{3}{2}} e^{-\frac{1}{2}\eta^2 x} \mathbf{1}_{\{x>0\}} dx.$$

### The Gamma Subordinator

The *Gamma* ( $\Gamma$ ) *subordinator*  $\Lambda^\Gamma = \{\Lambda_t^\Gamma\}_{t \geq 0}$  is another Lévy process that belongs to the class of infinite activity subordinators. Given a Gamma subordinator with parameters  $\eta, \beta > 0$ , we have that  $\Lambda_y^\Gamma$  follows a Gamma  $\Gamma(\beta t, \eta)$ -distribution with density

$$f_\Gamma(x) = \frac{\eta^{\beta t}}{\Gamma(\beta t)} x^{\beta t - 1} e^{-x\eta} \mathbf{1}_{\{x>0\}},$$

where

$$\Gamma(y) = \int_0^\infty t^{y-1} e^{-t} dt.$$

The corresponding Lévy measure is given by

$$\nu_\Gamma(dx) = \beta e^{-\eta x} \frac{1}{x} \mathbf{1}_{\{x>0\}} dx.$$

### The Compound Poisson Subordinator

The *Compound Poisson subordinator* is of the form

$$\Lambda_t = \mu t + \sum_{k=1}^{N_t} J_k,$$

where  $\{J_k\}_{k \in \mathbb{N}}$  are i.i.d. random variables with a cumulative distribution  $D$  with support on the positive axis, and  $N = \{N_t\}_{t \geq 0}$  is a Poisson process with intensity  $\beta$  which is independent of  $\{J_k\}_{k \in \mathbb{N}}$ . The Lévy measure corresponding to this Lévy subordinator has the form  $\nu_P(dy) = \beta dD(y)$ . This subordinator has upward jumps of magnitude  $J_k$  and the expected number of jumps within a unit time interval is  $\beta$ . For instance we can assume that  $D$  is the exponential distribution with parameter  $\eta > 0$ , so that we get a subordinator depending on the only two parameters  $\eta$  and  $\beta$ .

### 1.3.2 Time normalization condition

In the model of Mai and Scherer [3] the following *time normalization condition* is needed to be able to separate the marginal distributions and the dependence structure, given by a copula that, under such condition, doesn't depend on the marginal distributions.

**Definition 1** (Time normalization (TN)). *Let  $F$  be a cumulative distribution function with  $F(0) = 0$ . Let  $\Lambda = \{\Lambda_t\}_{t \geq 0}$  be a stochastic process which is almost surely non-decreasing and such that  $\Lambda_0 = 0$ . We say that  $\Lambda$  satisfies (TN) for the distribution  $F$  if  $\mathbb{E}[F(\Lambda_t)] = F(t)$ , for each  $t \geq 0$ .*

If in particular we consider the cumulative distribution function of an exponential random variable with parameter  $(-\lambda) > 0$ ,  $F(t) = (1 - e^{\lambda t})\mathbf{1}_{\{t > 0\}}$ , and a Lévy subordinator  $\Lambda$  with characteristics  $(\mu, \nu)$ , we have that

$$\Lambda \text{ satisfies (TN) for } F \Leftrightarrow \Psi(\lambda) = \lambda$$

.

## 1.4 The Multivariate Default Model

The model developed in this paper is an extension of the Mai and Scherer model [3].

Consider  $n$  defaultable firms with default times

$$\tau_1, \tau_2, \dots, \tau_n.$$

In [3] these default times are supposed to be characterized by individual factors, given by their marginal distribution functions  $G_i(t)$  (with  $G_i(0) = 0$ ,  $G_i(t) < 1$  for each  $t \geq 0$ ,  $i = 1, \dots, n$ ), and linked by a common factor, a Lévy subordinator  $\Lambda_t$  with Laplace exponent  $\Psi(\lambda)$ . In particular for each firm is considered, by a time transformation, the related cumulative hazard function  $h_i(t) = -\log(1 - G_i(t))$ ; this function  $h_i(t) : [0, \infty) \rightarrow \mathbb{R}$  is non negative, increasing, with  $h_i(0) = 0$  and such that  $\lim_{t \rightarrow \infty} h_i(t) = \infty$ . The firms survival functions are so defined as  $\overline{G}_i(t) := e^{-h_i(t)}$ ,  $t \geq 0$ .

In order to construct the default times such that they have the pre-specified marginal distributions and the dependence structure given by the subordinator, as threshold factors are considered  $n$  exponential times  $E_i$ , i.i.d. exponential random variables with parameter 1, also independent by the Lévy subordinator  $(\Lambda_t)_{t \geq 0}$  satisfying (TN) for the unit exponential law, and so with Laplace exponent  $\Psi$  satisfying  $\Psi(-1) = -1$ . In [3] the  $i$ -th default time is defined by

$$\tau_i = \inf\{t > 0 : \Lambda_{h_i(t)} > E_i\},$$

and so can be considered as first-jump time of a Poisson process with the stochastic clock  $\{\Lambda_{h_i(t)}\}_{t \geq 0}$ .

In our paper we consider the existence of other individual factors that, together with the individual hazard function and the common subordinator, define the default times as

$$\tau_i = \inf\{t > 0 : a_i \Lambda_{h_i(t)} + b_i h_i(t) > E_i\}.$$

In other words, for each  $i = 1, \dots, n$ , there exists a Lévy subordinator

$$\Lambda_i(t) = a_i \Lambda_t + b_i t$$

and a function  $h_i$  such that

$$\tau_i = \inf\{t > 0 : \Lambda_i(h_i(t)) > E_i\}.$$

Working with the survival distributions, let us first compute the marginal distributions and in a second step the joint distribution.

### The marginal distributions

For each default time the marginal survival distribution is given by

$$\begin{aligned} \bar{F}_i(t) &:= \mathbb{P}(\tau_i > t) = \mathbb{E}(\mathbb{P}(\tau_i > t | \mathcal{F}_\infty^\Lambda)) = \mathbb{E}(e^{-\Lambda_i(h_i(t))}) = \mathbb{E}(e^{-a_i \Lambda_{h_i(t)} - b_i h_i(t)}) \\ &= e^{\Psi(-a_i) h_i(t) - b_i h_i(t)} = e^{-(b_i - \Psi(-a_i)) h_i(t)} = (\bar{G}_i(t))^{b_i - \Psi(-a_i)} \end{aligned}$$

Let us observe that assuming the parameters constraint  $b_i - \Psi(-a_i) = 1$ , we would obtain that the default times marginal distributions are in fact  $G_i(t)$ .

We can also obtain the inverse function

$$\begin{aligned} \bar{F}_i^{-1}(u) = t(u) \quad \Leftrightarrow \quad e^{-(b_i - \Psi(-a_i)) h_i(t(u))} = u \quad \Leftrightarrow \quad -(b_i - \Psi(-a_i)) h_i(t(u)) = \log u \\ \Updownarrow \\ h_i(t(u)) = -\frac{\log u}{b_i - \Psi(-a_i)} \end{aligned} \quad (*)$$

that means, as  $h_i$  is strictly increasing and continuous, and so invertible,

$$\bar{F}_i^{-1}(u) = h_i^{-1}\left(-\frac{\log u}{b_i - \Psi(-a_i)}\right).$$

Notice that (\*) will be fundamental to compute the survival joint copula, as we'll see later.

## The joint distribution

For the joint survival distribution, conditioning to the sigma-algebra  $\mathcal{F}_\infty^\Lambda$ , we get

$$\begin{aligned}\bar{F}_{\tau_1, \dots, \tau_n}(t_1, \dots, t_n) &:= \mathbb{P}(\tau_i > t_i, i = 1, \dots, n) = \mathbb{E}(\mathbb{P}(\tau_i > t_i, i = 1, \dots, n | \mathcal{F}_\infty^\Lambda)) \\ &= \mathbb{E}\left(\prod_{i=1}^n e^{-\Lambda_i(h_i(t_i))}\right) = \mathbb{E}\left(\prod_{i=1}^n e^{-a_i \Lambda_{h_i(t_i)} - b_i h_i(t_i)}\right) \\ &= \mathbb{E}\left(\prod_{i=1}^n e^{-a_i \Lambda_{h_i(t_i)}}\right) e^{-\sum_{i=1}^n b_i h_i(t_i)}.\end{aligned}$$

To compute the expected value

$$\mathbb{E}\left(\prod_{i=1}^n e^{-a_i \Lambda_{h_i(t_i)}}\right) = \mathbb{E}\left(e^{-\sum_{i=1}^n a_i \Lambda_{h_i(t_i)}}\right)$$

we introduce the permutation  $\sigma_i(\mathbf{t}) = \sigma_i(t_1, \dots, t_n)$  such that

$$h_{(i)}(\mathbf{t}) := h_{\sigma_i(\mathbf{t})}(t_{\sigma_i(\mathbf{t})})$$

is a reordering of  $h_i(t_i)$ , that means that

$$h_{(i-1)}(\mathbf{t}) \leq h_{(i)}(\mathbf{t}), \quad i = 1, \dots, n,$$

where we assume by convention  $h_{(0)}(\mathbf{t}) = 0$ .

Let us also introduce the following notation:

$$\theta_j(\mathbf{t}) = \sum_{i=j}^n a_{\sigma_i(\mathbf{t})}, \quad j = 1, \dots, n$$

so that

$$\theta_{j+1}(\mathbf{t}) = \theta_j(\mathbf{t}) - a_{\sigma_j(\mathbf{t})}, \quad j = 1, \dots, n$$

where we assume by convention that  $\theta_{n+1}(\mathbf{t}) = 0$ .



We have

$$\begin{aligned}
\sum_{i=1}^n a_i \Lambda_{h_i(t_i)} &= \sum_{i=1}^n a_{\sigma_i(\mathbf{t})} \Lambda_{h_{\sigma_i(\mathbf{t})}(t_{\sigma_i(\mathbf{t})})} = \sum_{i=1}^n a_{\sigma_i(\mathbf{t})} \Lambda_{h_{(i)}(\mathbf{t})} \\
&= \sum_{i=1}^n a_{\sigma_i(\mathbf{t})} \sum_{j=1}^i (\Lambda_{h_{(j)}(\mathbf{t})} - \Lambda_{h_{(j-1)}(\mathbf{t})}) \\
&= \sum_{j=1}^n (\Lambda_{h_{(j)}(\mathbf{t})} - \Lambda_{h_{(j-1)}(\mathbf{t})}) \sum_{i=j}^n a_{\sigma_i(\mathbf{t})} \\
&= \sum_{j=1}^n (\Lambda_{h_{(j)}(\mathbf{t})} - \Lambda_{h_{(j-1)}(\mathbf{t})}) \theta_j(\mathbf{t}).
\end{aligned}$$

Now, as a subordinator is a process with independent increments, we can compute

$$\begin{aligned}
\mathbb{E}\left(\prod_{i=1}^n e^{-a_i \Lambda_{h_i(t_i)}}\right) &= \mathbb{E}\left(e^{-\sum_{i=1}^n a_i \Lambda_{h_i(t_i)}}\right) = \mathbb{E}\left(e^{-\sum_{j=1}^n (\Lambda_{h_{(j)}(\mathbf{t})} - \Lambda_{h_{(j-1)}(\mathbf{t})}) \theta_j(\mathbf{t})}\right) \\
&= \prod_{j=1}^n \mathbb{E}\left(e^{-\theta_j(\mathbf{t}) (\Lambda_{h_{(j)}(\mathbf{t})} - \Lambda_{h_{(j-1)}(\mathbf{t})})}\right) = \prod_{j=1}^n e^{\Psi(-\theta_j(\mathbf{t})) (h_{(j)}(\mathbf{t}) - h_{(j-1)}(\mathbf{t}))} \\
&= e^{\sum_{j=1}^n h_{(j)}(\mathbf{t}) (\Psi(-\theta_j(\mathbf{t})) - \Psi(-\theta_{j+1}(\mathbf{t})))}
\end{aligned}$$

where in the last equation we have considered that  $h_{(0)}(\mathbf{t}) = 0$ ,  $\theta_{n+1}(\mathbf{t}) = 0$  and so  $\Psi(-\theta_{n+1}(\mathbf{t})) = 0$ . Finally, considering also the relationship between  $\theta_{j+1}(\mathbf{t})$  and  $\theta_j(\mathbf{t})$ , we can compute

$$\begin{aligned}
\overline{F}_{\tau_1, \dots, \tau_n}(t_1, \dots, t_n) &:= \mathbb{P}(\tau_i > t_i, i = 1, \dots, n) = \mathbb{E}\left(\prod_{i=1}^n e^{-a_i \Lambda_{h_i(t_i)}}\right) e^{-\sum_{i=1}^n b_i h_i(t_i)} \\
&= e^{-\sum_{j=1}^n h_{(j)}(\mathbf{t}) (\Psi(-\theta_{j+1}(\mathbf{t})) - \Psi(-\theta_j(\mathbf{t})))} e^{-\sum_{i=1}^n b_i h_i(t_i)} \\
&= e^{-\sum_{j=1}^n h_{(j)}(\mathbf{t}) (\Psi(-\theta_j(\mathbf{t}) + a_{\sigma_j(\mathbf{t})}) - \Psi(-\theta_j(\mathbf{t})))} e^{-\sum_{i=1}^n b_{\sigma_i(\mathbf{t})} h_{\sigma_i(\mathbf{t})}(t_{\sigma_i(\mathbf{t})})} \\
&= e^{-\sum_{j=1}^n (\Psi(-\theta_j(\mathbf{t}) + a_{\sigma_j(\mathbf{t})}) - \Psi(-\theta_j(\mathbf{t})) + b_{\sigma_j(\mathbf{t})}) h_{\sigma_j(\mathbf{t})}(t_{\sigma_j(\mathbf{t})})}
\end{aligned}$$

Remark that, for the monotonicity property of the Laplace exponent, we have

$$\Psi(-\theta_j(\mathbf{t}) + a_{\sigma_j(\mathbf{t})}) - \Psi(-\theta_j(\mathbf{t})) + b_{\sigma_j(\mathbf{t})} > 0.$$

Alternately, introducing the permutation  $\sigma_j^{-1}(\mathbf{t})$ , i.e. the inverse permutation of  $\sigma_j(\mathbf{t})$ , we can write

$$\begin{aligned}
\overline{F}_{\tau_1, \dots, \tau_n}(t_1, \dots, t_n) &= e^{-\sum_{j=1}^n (\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+a_j) - \Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+b_j))} h_j(t_j) \\
&= \prod_{j=1}^n \overline{G}_j(t_j)^{-\left(\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+a_j) - \Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+b_j)\right)} \\
&= \prod_{j=1}^n \overline{G}_j(t_j)^{-\left(\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+a_j) - \Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+b_j) + \Psi(-a_j) + b_j - \Psi(-a_j)\right)} \\
&= \prod_{j=1}^n (\overline{G}_j(t_j)^{b_j - \Psi(-a_j)})^{-\left(\frac{\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+a_j) - (\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})) - \Psi(-a_j))}{b_j - \Psi(-a_j)} + 1\right)} \\
&= \prod_{j=1}^n (\overline{F}_{\tau_j}(t_j))^{-\left(\frac{\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})+a_j) - (\Psi(-\theta_{\sigma_j^{-1}(\mathbf{t})}(\mathbf{t})) - \Psi(-a_j))}{b_j - \Psi(-a_j)} + 1\right)}.
\end{aligned}$$

## The survival copula

The survival copula is defined as

$$\begin{aligned}
\widehat{C}_{\tau_1, \dots, \tau_n}(u_1, \dots, u_n) &= \overline{F}_{\tau_1, \dots, \tau_n}(\overline{F}_{\tau_1}^{-1}(u_1), \dots, \overline{F}_{\tau_n}^{-1}(u_n)) \\
&= \overline{F}_{\tau_1, \dots, \tau_n}(t_1(u_1), \dots, t_n(u_n)) = \overline{F}_{\tau_1, \dots, \tau_n}(\mathbf{t}(\mathbf{u}))
\end{aligned}$$

where  $\mathbf{t}(\mathbf{u})$  is the vector with components  $t_i(u_i)$ , by which we mean  $\overline{F}_{\tau_i}^{-1}(u_i)$ . So we have

$$\widehat{C}_{\tau_1, \dots, \tau_n}(u_1, \dots, u_n) = e^{-\sum_{j=1}^n (\Psi(-\theta_j(\mathbf{t}(\mathbf{u})) + a_{\sigma_j(\mathbf{t}(\mathbf{u}))}) - \Psi(-\theta_j(\mathbf{t}(\mathbf{u})) + b_{\sigma_j(\mathbf{t}(\mathbf{u}))}))} h_{\sigma_j(\mathbf{t}(\mathbf{u}))}(t_{\sigma_j(\mathbf{t}(\mathbf{u}))}(u_{\sigma_j(\mathbf{t}(\mathbf{u}))}))$$

Let us introduce the following notation:

$$\widehat{\sigma}_j(\mathbf{u}) := \sigma_j(\mathbf{t}(\mathbf{u}))$$

and the following characterization, directly in terms of the vector  $\mathbf{u}$ :

$$s_i(\mathbf{u}) := -\frac{1}{b_i - \Psi(-a_i)} \log u_i = \log(u_i)^{-\frac{1}{b_i - \Psi(-a_i)}}.$$

Thank's to the relationship (\*), to arrange in order of increasing magnitude  $h_i(t_i(u_i))$  we can put in an increasing order  $-\frac{\log u_i}{b_i - \Psi(-a_i)}$ ; we consider the permutation  $\widehat{\sigma}_j(\mathbf{u})$  as the permutation (not necessarily unique) such that

$$-\frac{\log u_{\widehat{\sigma}_{j-1}(\mathbf{u})}}{b_{\widehat{\sigma}_{j-1}(\mathbf{u})} - \Psi(-a_{\widehat{\sigma}_{j-1}(\mathbf{u})})} \leq -\frac{\log u_{\widehat{\sigma}_j(\mathbf{u})}}{b_{\widehat{\sigma}_j(\mathbf{u})} - \Psi(-a_{\widehat{\sigma}_j(\mathbf{u})})} \quad j = 2, \dots, n.$$

It follows that, considering the following reordering for  $s_i(\mathbf{u})$

$$s_{(1)}(\mathbf{u}) \leq s_{(2)}(\mathbf{u}) \leq \dots \leq s_{(n)}(\mathbf{u}),$$

the permutation  $\widehat{\sigma}_j(\mathbf{u})$  is defined by

$$-\frac{\log u_{\widehat{\sigma}_j(\mathbf{u})}}{b_{\widehat{\sigma}_j(\mathbf{u})} - \Psi(-a_{\widehat{\sigma}_j(\mathbf{u})})} = s_{(j)}(\mathbf{u}).$$

Let us also introduce the following notation:

$$\widehat{\theta}_j(\mathbf{u}) := \theta_j(\mathbf{t}(\mathbf{u})) = \sum_{i=j}^n a_{\sigma_i(\mathbf{t}(\mathbf{u}))} = \sum_{i=j}^n a_{\widehat{\sigma}_i(\mathbf{u})},$$

For the copula computation we have the following lemma:

**Lemma 1.** *The survival copula of the vector  $\tau_1, \dots, \tau_n$  is*

$$\begin{aligned} \widehat{C}_{\tau_1, \dots, \tau_n}(u_1, \dots, u_n) &= e^{\sum_{j=1}^n \left( \Psi(-\widehat{\theta}_j(\mathbf{u}) + a_{\widehat{\sigma}_j(\mathbf{u})}) - \Psi(-\widehat{\theta}_j(\mathbf{u}) + b_{\widehat{\sigma}_j(\mathbf{u})}) \right) \frac{\log u_{\widehat{\sigma}_j(\mathbf{u})}}{b_{\widehat{\sigma}_j(\mathbf{u})} - \Psi(-a_{\widehat{\sigma}_j(\mathbf{u})})}} \\ &= \prod_{j=1}^n u_{\widehat{\sigma}_j(\mathbf{u})}^{\frac{\Psi(-\widehat{\theta}_j(\mathbf{u}) + a_{\widehat{\sigma}_j(\mathbf{u})}) - \Psi(-\widehat{\theta}_j(\mathbf{u}) + b_{\widehat{\sigma}_j(\mathbf{u})})}{b_{\widehat{\sigma}_j(\mathbf{u})} - \Psi(-a_{\widehat{\sigma}_j(\mathbf{u})})}} \\ &= \prod_{j=1}^n \left( \left( u_{\widehat{\sigma}_j(\mathbf{u})} \right)^{\frac{1}{b_{\widehat{\sigma}_j(\mathbf{u})} - \Psi(-a_{\widehat{\sigma}_j(\mathbf{u})})}} \right)^{\Psi(-\widehat{\theta}_j(\mathbf{u}) + a_{\widehat{\sigma}_j(\mathbf{u})}) - \Psi(-\widehat{\theta}_j(\mathbf{u}) + b_{\widehat{\sigma}_j(\mathbf{u})})}. \end{aligned}$$

In our model we'll assume the parameters constraint  $b_i - \Psi(-a_i) = 1$ , by which the marginal survival distributions are exactly  $\overline{G}_i$ . We will calibrate the model as in Mai and Scherer [3] in two steps: we'll first calibrate the marginal distributions, and then the copula.

**Cor 1.** *In case  $b_i = 0$ ,  $a_i = 1$  and  $\Psi(-1) = -1$  we find the Mai and Scherer copula given by*

$$\widehat{C}_{\tau_1, \dots, \tau_n}(u_1, \dots, u_n) = \prod_{i=1}^n u_{(i)}^{\Psi(-(i-1)) - \Psi(-i)}.$$

**Proof of Cor 1.** *In fact, with those parameters,*

$$s_i(\mathbf{u}) := -\frac{\log u_i}{b_i - \Psi(-a_i)} = -\log u_i, \quad i = 1, \dots, n,$$

and so the permutation  $\widehat{\sigma}_j(\mathbf{u})$  is linked to the permutation  $\sigma_i$  of the reordering of  $u_i$ , as follows

$$u_{\widehat{\sigma}_j(\mathbf{u})} = u_{(n-j+1)} = u_{\sigma_{n-j+1}}.$$

Moreover  $\widehat{\theta}_j(\mathbf{u}) = n - j + 1$  and so the copula becomes

$$\prod_{j=1}^n u_{(n-j+1)}^{\Psi(-(n-j)) - \Psi(-(n-j+1))} = \prod_{i=1}^n u_{(i)}^{\Psi(-(i-1)) - \Psi(-i)}.$$

### 1.4.1 Heterogeneous case with $r$ different homogeneous classes

Let us in general suppose that according to their rating classes our  $n$  firms can be divided into  $r$  different classes with the related  $r$  parameters  $a_i$ ,  $b_i$  and hazard function  $h_i$ . In this case we can consider the first  $m_1$  firms of type 1, other  $m_2$  firms of type 2, and so on up to the remaining  $m_r$  firms of class  $r$ , with  $n = m_1 + m_2 + \dots + m_r$ . In other words, denoting  $M_1 = m_1$ ,  $M_2 = m_1 + m_2$ ,  $M_\ell = M_{\ell-1} + m_\ell$ ,  $\ell \leq r$ , we assume

$$\left\{ \begin{array}{l} a_i = a^{(1)} \text{ and } b_i = b^{(1)}, \text{ for } i = 1, \dots, m_1 = M_1 \\ a_i = a^{(2)} \text{ and } b_i = b^{(2)}, \text{ for } i = M_1 + 1, \dots, m_1 + m_2 = M_2 \\ \dots \\ a_i = a^{(r)} \text{ and } b_i = b^{(r)}, \text{ for } i = M_{r-1} + 1, \dots, M_r (= n) \end{array} \right. \quad (*m)$$

For the reason previously explained, we assume  $b_i - \Psi(-a_i) = 1$  for each  $i$ .

Other two reasonable assumptions about the  $r$  classes are the following: the related hazard functions  $h^{(i)}(t)$  are such that, for each  $t$ ,

$$h^{(1)}(t) \leq h^{(2)}(t) \leq \dots \leq h^{(r)}(t) \quad (*h)$$

and

$$a^{(1)} \leq a^{(2)} \leq \dots \leq a^{(r)} \quad (*a).$$

So we have

$$a^{(1)} \Lambda_{h^{(1)}(t)} \leq a^{(2)} \Lambda_{h^{(2)}(t)} \leq \dots \leq a^{(r)} \Lambda_{h^{(r)}(t)}.$$

With our further assumption

$$b^{(l)} - \Psi(-a^{(l)}) = 1 \quad \Leftrightarrow \quad b^{(l)} = 1 + \Psi(-a^{(l)}), \quad l = 1, \dots, r \quad (*p)$$

and considering that  $\Psi$  is strictly increasing, we also get

$$b^{(1)} \geq b^{(2)} \geq \dots \geq b^{(r)}. \quad (*b)$$

For each firm in a class  $(l)$ ,  $l = 1, \dots, r$ , we can denote the common subordinator as

$$\Lambda^{(l)}(t) := a^{(l)} \Lambda_{h^{(l)}(t)} + b^{(l)} h^{(l)}(t).$$

Let us denote the classes  $\mathcal{M}_\ell := \{M_{l-1} + 1, \dots, M_l\}$  and remark that our assumptions imply that, if  $i \in \mathcal{M}_\ell$  and  $j \in \mathcal{M}_{\ell+1}$ , we have the survival distributions stochastic ordering

$$\mathbb{P}(\tau_i > t) = \overline{G}_i(t) \geq \mathbb{P}(\tau_j > t) = \overline{G}_j(t).$$

In other words the default risk classes  $\mathcal{M}_\ell$  are ordered such that the first class is the less risky, while the last class is the more risky.

As we'll see these assumptions will simplify the computation of the portfolio loss distribution, and in particular allow us to easily compute the survival probability involving all the variables  $\tau_i$

$$\mathbb{P}(\tau_1 > t, \tau_2 > t, \dots, \tau_n > t).$$

In fact we have the following proposition:

**Prop. 1.** *Under the above assumptions  $(*a)$ ,  $(*b)$ ,  $(*h)$  and  $(*p)$  on the parameters constraints, we get*

$$\begin{aligned} & \mathbb{P}(\tau_i > t, \forall i \in I_j, \forall j = 1, \dots, r) \\ &= e^{-\sum_{j=1}^r \left( \Psi(-\sum_{\ell=j+1}^r k_\ell a^{(\ell)}) - \Psi(-\sum_{\ell=j}^r k_\ell a^{(\ell)}) + k_j (\Psi(-a^{(j)}) + 1) \right) h^{(j)}(t)} \end{aligned} \quad (1.1)$$

where  $k_j = |I_j|$ .

Let us remark that this formula depends on the parameters  $a^{(j)}$  but not on  $b^{(j)}$ .

**Proof of Prop 1.** *We have*

$$\begin{aligned} \sum_{i=1}^n a_i \Lambda_{h_i(t_i)} &= \sum_{\ell=1}^r m_\ell a^{(\ell)} \Lambda_{h^{(\ell)}(t)} \\ &= \sum_{\ell=1}^r m_\ell a^{(\ell)} \sum_{j=1}^i (\Lambda_{h^{(j)}(t)} - \Lambda_{h^{(j-1)}(t)}) \\ &= \sum_{j=1}^r (\Lambda_{h^{(j)}(t)} - \Lambda_{h^{(j-1)}(t)}) \sum_{i=j}^r m_\ell a^{(\ell)} \\ &= \sum_{j=1}^n (\Lambda_{h^{(j)}(t)} - \Lambda_{h^{(j-1)}(t)}) \theta_j^r, \end{aligned}$$

where

$$\theta_j^r := \sum_{\ell=j}^r m_\ell a^{(\ell)}.$$

We can thus compute

$$\begin{aligned} \mathbb{P}(\tau_1 > t, \tau_2 > t, \dots, \tau_n > t) &= \mathbb{E}\left(\prod_{i=1}^n e^{-a_i \Lambda_{h_i}(t)}\right) e^{-\sum_{i=1}^n b_i h_i(t)} \\ &= e^{-\sum_{j=1}^r h^{(j)}(t) \left(\Psi(-\theta_j^r + m_j a^{(j)}) - \Psi(-\theta_j^r)\right)} e^{-\sum_{i=1}^r b^{(j)} h^{(j)}(t)} \\ &= e^{-\sum_{j=1}^r \left(\Psi(-\theta_j^r + m_j a^{(j)}) - \Psi(-\theta_j^r) + m_j b^{(j)}\right) h^{(j)}(t)} \\ &= e^{-\sum_{j=1}^r \left(\Psi(-\sum_{\ell=j+1}^r m_\ell a^{(\ell)}) - \Psi(-\sum_{\ell=j}^r m_\ell a^{(\ell)}) + m_j b^{(j)}\right) h^{(j)}(t)} \end{aligned} \quad (1.2)$$

$$\begin{aligned} &\mathbb{P}(\tau_i > t, \forall i \in I_j, \forall j = 1, \dots, r) \\ &= e^{-\sum_{j=1}^r \left(\Psi(-\sum_{\ell=j+1}^r k_\ell a^{(\ell)}) - \Psi(-\sum_{\ell=j}^r k_\ell a^{(\ell)}) + k_j b^{(j)}\right) h^{(j)}(t)} \\ &= e^{-\sum_{j=1}^r \left(\Psi(-\sum_{\ell=j+1}^r k_\ell a^{(\ell)}) - \Psi(-\sum_{\ell=j}^r k_\ell a^{(\ell)}) + k_j (\Psi(-a^{(j)}) + 1)\right) h^{(j)}(t)} \end{aligned} \quad (1.3)$$

where  $k_j = |I_j|$ .

## The Correlation Coefficient

An interesting computation involves the default correlation of firms  $i$  and  $j$  up to time  $t$ .

Let us define the stochastic processes  $A^i = \{A_t^i\}_{t \geq 0}$  for  $i = 1, \dots, n$  by

$$A_t^i := \mathbf{1}_{\{E_i < \Lambda_i(h_i(t))\}}$$

so that the  $i$ -th default time can be defined by

$$\tau_i = \inf\{t > 0 : E_i < \Lambda_i(h_i(t))\} = \inf\{t > 0 : A_t^i = 1\}.$$

**Prop. 2.** Consider two firms  $i$  and  $j$  in the rating classes  $\mathcal{M}_m$  and  $\mathcal{M}_n$  respectively, with  $m < n$ ; the covariance coefficient  $\text{Cov}[A_t^i, A_t^j]$  is given by

$$\text{Cov}[A_t^i, A_t^j] = \overline{G}^{(m)}(t) \overline{G}^{(n)}(t) \left( \max(\overline{G}^{(m)}(t), \overline{G}^{(n)}(t)) \left( \Psi(-a^{(m)}) + \Psi(-a^{(n)}) - \Psi(-(a^{(m)} + a^{(n)})) \right) - 1 \right).$$

while the correlation coefficient  $\text{Corr}[A_t^i, A_t^j]$  is given by

$$\text{Corr}[A_t^i, A_t^j] = \frac{\sqrt{\overline{G}^{(m)}(t)} \sqrt{\overline{G}^{(n)}(t)} \left( \max(\overline{G}^{(m)}(t), \overline{G}^{(n)}(t)) \left( \Psi(-a^{(m)}) + \Psi(-a^{(n)}) - \Psi(-(a^{(m)} + a^{(n)})) \right) - 1 \right)}{\sqrt{1 - \overline{G}^{(m)}(t)} \sqrt{1 - \overline{G}^{(n)}(t)}}.$$

If instead  $i$  and  $j$  ( $i \neq j$ ) are in the same rating class  $\mathcal{M}_m$  we have

$$\text{Cov}[A_t^i, A_t^j] = \overline{G}^{(m)2}(t) (\overline{G}^{(m)}(t)^{(-\Psi(-2a^{(m)})+2\Psi(-a^{(m)}))} - 1),$$

and

$$\text{Corr}[A_t^i, A_t^j] = \frac{\overline{G}^{(m)}(t) (\overline{G}^{(m)}(t)^{(-\Psi(-2a^{(m)})+2\Psi(-a^{(m)}))} - 1)}{1 - \overline{G}^{(m)}(t)}.$$

**Proof of Prop 2.** Let us start by the case  $i \in \mathcal{M}_m$ ,  $j \in \mathcal{M}_n$ , with  $m < n$ . In this case we have

$$\begin{aligned} a_i &= a^{(m)} \leq a_j = a^{(n)}, \\ b_j &= b^{(m)} \geq b_j = b^{(n)}, \\ h_i &= h^{(m)} \leq h_j = h^{(n)}, \\ \overline{G}_i(t) &= \overline{G}^{(m)}(t) = e^{-h^{(m)}(t)}, \\ \overline{G}_j(t) &= \overline{G}^{(n)}(t) = e^{-h^{(n)}(t)}. \end{aligned}$$

$$\begin{aligned} \text{Cov}[A_t^i, A_t^j] &= \text{Cov}[1 - A_t^i, 1 - A_t^j] \\ &= \mathbb{P}(\tau_i > t, \tau_j > t) - \mathbb{P}(\tau_i > t)\mathbb{P}(\tau_j > t) \\ &= \mathbb{E}[e^{-(a_i\Lambda_{h_i}(t)+a_j\Lambda_{h_j}(t))}]e^{-b_i h_i(t)-b_j h_j(t)} - \overline{G}_i(t)\overline{G}_j(t) \\ &= \mathbb{E}[e^{-(a^{(m)}\Lambda_{h^{(m)}}(t)+a^{(n)}\Lambda_{h^{(n)}}(t))}]e^{-b^{(m)}h^{(m)}(t)-b^{(n)}h^{(n)}(t)} - \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \\ &= \mathbb{E}[e^{-(a^{(m)}+a^{(n)})\Lambda_{h^{(m)}}(t)-a^{(n)}(\Lambda_{h^{(n)}}(t)-\Lambda_{h^{(m)}}(t))}]e^{-b^{(m)}h^{(m)}(t)-b^{(n)}h^{(n)}(t)} - \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \\ &= \mathbb{E}[e^{-(a^{(m)}+a^{(n)})\Lambda_{h^{(m)}}(t)}]\mathbb{E}[e^{-a^{(n)}(\Lambda_{h^{(n)}}(t)-\Lambda_{h^{(m)}}(t))}]e^{-b^{(m)}h^{(m)}(t)-b^{(n)}h^{(n)}(t)} - \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \\ &= e^{h^{(m)}(t)\Psi(-(a^{(m)}+a^{(n)}))}e^{(h^{(n)}(t)-h^{(m)}(t))\Psi(-a^{(n)})}e^{-b^{(m)}h^{(m)}(t)-b^{(n)}h^{(n)}(t)} - \overline{G}^{(m)}(t)\overline{G}^{(n)}(t). \end{aligned}$$

Now, considering that

$$\begin{aligned} &e^{h^{(m)}(t)\Psi(-(a^{(m)}+a^{(n)}))}e^{(h^{(n)}(t)-h^{(m)}(t))\Psi(-a^{(n)})}e^{-b^{(m)}h^{(m)}(t)-b^{(n)}h^{(n)}(t)} \\ &= e^{h^{(n)}(t)(\Psi(-a^{(n)})-b^{(n)})}e^{h^{(m)}(t)(-b^{(m)}-\Psi(-a^{(n)}))+\Psi(-(a^{(m)}+a^{(n)}))} \\ &= e^{h^{(n)}(t)(\Psi(-a^{(n)})-b^{(n)})}e^{h^{(m)}(t)(-b^{(m)}+\Psi(-a^{(m)})-\Psi(-a^{(m)})-\Psi(-a^{(n)})+\Psi(-(a^{(m)}+a^{(n)}))} \end{aligned}$$

and taking into account condition (\*p) and that

$$\overline{G}^{(m)}(t) = e^{-h^{(m)}(t)} \quad \overline{G}^{(n)}(t) = e^{-h^{(n)}(t)},$$

we can write this formula as

$$= \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) e^{h^{(m)}(t)\left(-\Psi(-a^{(m)})-\Psi(-a^{(n)})+\Psi(-(a^{(m)}+a^{(n)}))\right)}.$$

So we have

$$Cov[A_t^i, A_t^j] = \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \left( e^{h^{(m)}(t)\left(-\Psi(-a^{(m)})-\Psi(-a^{(n)})+\Psi(-(a^{(m)}+a^{(n)}))\right)} - 1 \right)$$

that is equivalent to

$$Cov[A_t^i, A_t^j] = \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \left( e^{-h^{(m)}(t)\left(\Psi(-a^{(m)})+\Psi(-a^{(n)})-\Psi(-(a^{(m)}+a^{(n)}))\right)} - 1 \right)$$

and taking into account that  $\overline{G}^{(m)}(t) = \max(\overline{G}^{(m)}(t), \overline{G}^{(n)}(t))$  we finally have

$$Cov[A_t^i, A_t^j] = \overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \left( \max(\overline{G}^{(m)}(t), \overline{G}^{(n)}(t))^{\left(\Psi(-a^{(m)})+\Psi(-a^{(n)})-\Psi(-(a^{(m)}+a^{(n)}))\right)} - 1 \right).$$

To compute the correlation coefficient  $Corr[A_t^i, A_t^j]$  we need to compute the variance  $Var(A_t^i)$ . We have

$$\begin{aligned} Var(A_t^i) &= Var(1 - A_t^i) \\ &= \mathbb{P}(\tau_i > t)(1 - \mathbb{P}(\tau_i > t)) \\ &= \overline{G}_i(t)(1 - \overline{G}_i(t)). \end{aligned}$$

Of course if the firm  $i$  is in the rating class  $\mathcal{M}_m$ , we can write the previous formula as

$$Var(A_t^i) = \overline{G}^{(m)}(t)(1 - \overline{G}^{(m)}(t)).$$

We can thus compute the correlation coefficient between firm  $i$  and firm  $j$  as

$$\begin{aligned} Corr[A_t^i, A_t^j] &= \frac{Cov[A_t^i, A_t^j]}{\sqrt{Var[A_t^i]}\sqrt{Var[A_t^j]}} \\ &= \frac{\overline{G}^{(m)}(t)\overline{G}^{(n)}(t) \left( \max(\overline{G}^{(m)}(t), \overline{G}^{(n)}(t))^{\left(\Psi(-a^{(m)})+\Psi(-a^{(n)})-\Psi(-(a^{(m)}+a^{(n)}))\right)} - 1 \right)}{\sqrt{\overline{G}^{(m)}(t)(1 - \overline{G}^{(m)}(t))}\sqrt{\overline{G}^{(n)}(t)(1 - \overline{G}^{(n)}(t))}} \\ &= \frac{\sqrt{\overline{G}^{(m)}(t)}\sqrt{\overline{G}^{(n)}(t)} \left( \max(\overline{G}^{(m)}(t), \overline{G}^{(n)}(t))^{\left(\Psi(-a^{(m)})+\Psi(-a^{(n)})-\Psi(-(a^{(m)}+a^{(n)}))\right)} - 1 \right)}{\sqrt{1 - \overline{G}^{(m)}(t)}\sqrt{1 - \overline{G}^{(n)}(t)}}. \end{aligned}$$



If instead  $i, j \in \mathcal{M}_m$ , in the previous formula we have

$$\begin{aligned} a^{(m)} &= a^{(n)}, \\ G^{(m)}(t) &= G^{(n)}(t), \end{aligned}$$

and so we have

$$\text{Corr}[A_t^i, A_t^j] = \frac{\overline{G}^{(m)}(t) (\overline{G}^{(m)}(t)^{(-\Psi(-2a^{(m)}) + 2\Psi(-a^{(m)}))} - 1)}{1 - \overline{G}^{(m)}(t)}.$$

### The Loss Distribution

Let us assume a homogeneous portfolio in which each firm has the same weight. The zero-recovery loss process  $L^n = \{L^n(t)\}_{t \geq 0}$  is defined as

$$L^n(t) := \frac{1}{n} \sum_{i=1}^n A_t^i.$$

Thus  $L^n(t)$  gives the fraction of defaulted names in the portfolio up to time  $t$ .

To compute the portfolio loss distribution we want to compute, for  $k \in \{0, \dots, n\}$ ,

$$\mathbb{P}(nL^n(t) = k).$$

We have the following proposition:

**Prop. 3.** *The distribution of  $L^n(t)$  is given by*

$$\begin{aligned} &\mathbb{P}(nL^n(t) = k) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = n - k}} \dots \sum_{v=0}^k (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \dots \prod_{j=1}^r \frac{m_j!}{k_j! v_j! (m_j - k_j - v_j)!} (\overline{G}^{(j)}(t))^{\Phi^{(j)}}, \end{aligned}$$

where

$$\Phi^{(j)} = \Psi\left(-\sum_{\ell=j+1}^r (k_\ell + v_\ell) a^{(\ell)}\right) - \Psi\left(-\sum_{\ell=j}^r k_\ell a^{(\ell)}\right) + k_j (\Psi(-a^{(j)}) + 1).$$

**Proof of Prop 3.** *To compute the portfolio loss distribution we use formulas (1.2) and (1.3). We have*

$$\mathbb{P}(nL^n(t) = n - k) = \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = k}} \dots \sum \mathbb{P}(\forall j = 1, \dots, r \exists I_j : |I_j| = k_j, \tau_i > t, \forall i \in I_j, ; \tau_{i'} \leq t, \forall i' \in \mathcal{M}_j \setminus I_j)$$

where  $\mathcal{M}_j = \{M_{j-1} + 1, \dots, M_j\}$

$$= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = k}} \dots \sum_{j=1}^r \binom{m_j}{k_j} \mathbb{P}(\tau_i > t, \tau_{i'} \leq t, M_{j-1} < i \leq M_{j-1} + k_j, M_{j-1} + k_j < i' \leq M_j, j = 1, \dots, r)$$

Considering that

$$\begin{aligned} & \mathbb{P}(\tau_i > t, \tau_{i'} \leq t, M_{j-1} < i \leq M_{j-1} + k_j, M_{j-1} + k_j < i' \leq M_j, j = 1, \dots, r) \\ &= \mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) \end{aligned}$$

where

$$A = \{\tau_i > t, M_{j-1} < i \leq M_{j-1} + k_j, j = 1, \dots, r\}$$

and

$$B^c = \{\tau_{i'} \leq t, M_{j-1} + k_j < i' \leq M_j, j = 1, \dots, r\}$$

so that

$$B = \bigcup_{j=1, \dots, r} \bigcup_{M_{j-1} + k_j < i' \leq M_j} \{\tau_{i'} > t\}$$

we have

$$\mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}\left(\bigcup_{j=1, \dots, r} \bigcup_{M_{j-1} + k_j < i' \leq M_j} A \cap \{\tau_{i'} > t\}\right)$$

and, using the inclusion/exclusion formula, we get

$$= \mathbb{P}(A) - \sum_{v=1}^{n-k} \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \dots \sum_{j=1}^r (-1)^{v+1} \binom{m_j - k_j}{v_j} \mathbb{P}(A, \tau_{i'} > t, M_{j-1} + k_j < i' \leq M_{j-1} + k_j + v_j, j = 1, \dots, r)$$

(where by  $\{A, \tau_{i'} > t, \text{etc.}\}$  we mean  $A \cap \{\tau_{i'} > t, \text{etc.}\}$ )

$$\sum_{v=0}^{n-k} \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \dots \sum_{j=1}^r (-1)^v \binom{m_j - k_j}{v_j} \mathbb{P}(A, \tau_{i'} > t, M_{j-1} + k_j < i' \leq M_{j-1} + k_j + v_j, j = 1, \dots, r)$$

having used that

$$\mathbb{P}(A) = \sum_{\substack{v_1 + \dots + v_r = 0 \\ 0 \leq v_j \leq m_j - k_j}} \dots \sum_{j=1}^r (-1)^0 \binom{m_j - k_j}{v_j} \mathbb{P}(A, \tau_{i'} > t, M_{j-1} + k_j < i' \leq M_{j-1} + k_j + v_j, j = 1, \dots, r)$$

$$\mathbb{P}(A) = \sum_{v_j=0, j=1, \dots, r} \cdots \sum_{j=1}^r (-1)^0 \prod_{j=1}^r \binom{m_j - k_j}{0} \mathbb{P}(A, \tau_{i'} > t, M_{j-1} + k_j < i' \leq M_{j-1} + k_j + 0, j = 1, \dots, r)$$

where the condition  $\tau_{i'} > t, M_{j-1} + k_j < i' \leq M_{j-1} + k_j + 0$  is pointless as not involving any index  $i'$ .

With our assumptions  $(*h)$ ,  $(*a)$  and  $(*b)$  on the coefficients  $a^{(i)}$ ,  $b^{(i)}$  and the functions  $h^{(i)}(t)$ , we obtain

$$\begin{aligned} \mathbb{P}(nL^n(t) = n - k) &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = k}} \cdots \sum_{j=1}^r \prod_{j=1}^r \binom{m_j}{k_j} \mathbb{P}(\tau_i > t, \tau_{i'} \leq t, M_{j-1} < i \leq M_{j-1} + k_j, M_{j-1} + k_j < i' \leq M_j, j = 1, \dots, r) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = k}} \cdots \sum_{v=0}^{n-k} (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \cdots \sum_{j=1}^r \prod_{j=1}^r \binom{m_j}{k_j} \prod_{j=1}^r \binom{m_j - k_j}{v_j} \mathbb{P}(\tau_i > t, M_{j-1} < i \leq M_{j-1} + k_j + v_j, j = 1, \dots, r) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = k}} \cdots \sum_{v=0}^{n-k} (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \cdots \sum_{j=1}^r \frac{m_j!}{k_j! v_j! (m_j - k_j - v_j)!} \mathbb{P}(\tau_i > t, M_{j-1} < i \leq M_{j-1} + k_j + v_j, j = 1, \dots, r) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = k}} \cdots \sum_{v=0}^{n-k} (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \cdots \sum_{j=1}^r \frac{m_j!}{k_j! v_j! (m_j - k_j - v_j)!} \\ &\quad e^{-\sum_{j'=1}^r \left( \Psi(-\sum_{\ell=j'+1}^r (k_\ell + v_\ell) a^{(\ell)}) - \Psi(-\sum_{\ell=j'}^r k_\ell a^{(\ell)}) + k_{j'} b^{(j')} \right) v^{(j')}(t)} \end{aligned}$$

and in the same way we get

$$\begin{aligned} \mathbb{P}(nL^n(t) = k) &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = n - k}} \cdots \sum_{j=1}^r \prod_{j=1}^r \binom{m_j}{k_j} \mathbb{P}(\tau_i > t, \tau_{i'} \leq t, M_{j-1} < i \leq M_{j-1} + k_j, M_{j-1} + k_j < i' \leq M_j, j = 1, \dots, r) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = n - k}} \cdots \sum_{v=0}^k (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \cdots \sum_{j=1}^r \prod_{j=1}^r \binom{m_j}{k_j} \prod_{j=1}^r \binom{m_j - k_j}{v_j} \mathbb{P}(\tau_i > t, M_{j-1} < i \leq M_{j-1} + k_j + v_j, j = 1, \dots, r) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = n - k}} \cdots \sum_{v=0}^k (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \cdots \sum_{j=1}^r \frac{m_j!}{k_j! v_j! (m_j - k_j - v_j)!} \mathbb{P}(\tau_i > t, M_{j-1} < i \leq M_{j-1} + k_j + v_j, j = 1, \dots, r) \\ &= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = n - k}} \cdots \sum_{v=0}^k (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \cdots \sum_{j=1}^r \frac{m_j!}{k_j! v_j! (m_j - k_j - v_j)!} \\ &\quad e^{-\sum_{j'=1}^r \left( \Psi(-\sum_{\ell=j'+1}^r (k_\ell + v_\ell) a^{(\ell)}) - \Psi(-\sum_{\ell=j'}^r k_\ell a^{(\ell)}) + k_{j'} b^{(j')} \right) v^{(j')}(t)}. \end{aligned}$$

Let us note that the final exponential factor can be written as

$$\begin{aligned}
& e^{-\sum_{j'=1}^r \left( \Psi(-\sum_{\ell=j'+1}^r (k_\ell + v_\ell) a^{(\ell)}) - \Psi(-\sum_{\ell=j'}^r k_\ell a^{(\ell)}) + k_{j'} b^{(j')} \right) v^{(j')}(t)} \\
&= \prod_{j'=1}^r \left( e^{-v^{(j')}(t)} \right)^{\Psi(-\sum_{\ell=j'+1}^r (k_\ell + v_\ell) a^{(\ell)}) - \Psi(-\sum_{\ell=j'}^r k_\ell a^{(\ell)}) + k_{j'} b^{(j')}} \\
&= \prod_{j'=1}^r \left( \overline{G}^{(j')}(t) \right)^{\Psi(-\sum_{\ell=j'+1}^r (k_\ell + v_\ell) a^{(\ell)}) - \Psi(-\sum_{\ell=j'}^r k_\ell a^{(\ell)}) + k_{j'} b^{(j')}}
\end{aligned}$$

so the final result is

$$\begin{aligned}
& \mathbb{P}(nL^n(t) = k) \\
&= \sum_{\substack{0 \leq k_j \leq m_j \\ k_1 + \dots + k_r = n - k}} \dots \sum_{v=0}^k (-1)^v \sum_{\substack{v_1 + \dots + v_r = v \\ 0 \leq v_j \leq m_j - k_j}} \dots \sum_{j=1}^r \prod_{j=1}^r \frac{m_j!}{k_j! v_j! (m_j - k_j - v_j)!} \left( \overline{G}^{(j)}(t) \right)^{\Psi(-\sum_{\ell=j+1}^r (k_\ell + v_\ell) a^{(\ell)}) - \Psi(-\sum_{\ell=j}^r k_\ell a^{(\ell)}) + k_j b^{(j)}}.
\end{aligned}$$

**Theorem 1** (Portfolio-loss distribution approximation). *Let us denote  $L_{m_i}^{(i)}(t)$  the fraction of defaulted names in the rating class  $(i)$  (with respectively  $m_i$  firms) of the portfolio up to time  $t$*

$$L_{m_i}^{(i)}(t) := \frac{1}{m_i} \sum_{j=1}^{m_i} A_t^{M_{i-1}+j}$$

and let us denote

$$L_\infty^{(i)}(t) := 1 - e^{-(a^{(i)} \Lambda_{h^{(i)}}(t) + b^{(i)} h^{(i)}(t))}$$

Under  $\mathbb{P}(\cdot | \mathcal{F}_\infty^\Lambda)$ ,

$$m_i L_{m_i}^{(i)}(t) \sim \text{Bin}(m_i, 1 - e^{-(a^{(i)} \Lambda_{h^{(i)}}(t) + b^{(i)} h^{(i)}(t))}) = \text{Bin}(m_i, 1 - e^{-\Lambda^{(i)}(h^{(i)}(t))}).$$

Moreover for fixed  $t \geq 0$ ,  $L_{m_i}^{(i)}(t)$  tends to the variable  $L_\infty^{(i)}(t)$  in  $L^2$  as  $m_i$  tends to infinity.

Let us now consider the overall portfolio

$$L^n(t) = \frac{1}{n} \sum_{i=1}^r m_i L_{m_i}^{(i)}(t).$$

Let us denote the portfolio loss conditioned average by  $\hat{L}^n(t) := \mathbb{E}[L^n(t) | \mathcal{F}_\infty^\Lambda]$ .

Then we have

$$\hat{L}^n(t) = \frac{1}{n} \sum_{i=1}^r m_i (1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})$$

and

$$L^n(t) - \hat{L}^n(t) \xrightarrow{n \rightarrow \infty} 0,$$

in  $L^2$  for each  $t$ . So we can use  $\hat{L}^n(t)$  as an approximation of  $L^n(t)$  and in particular, for the approximation error, we have the upper bound of  $\sum_{i=1}^r \frac{\sqrt{m_i}}{n}$ .

**Proof of Theorem 1.** In each rating class ( $i$ ) the firms are homogeneous and conditionally independent. So we have that under  $\mathbb{P}(\cdot | \mathcal{F}_\infty^\Lambda)$ ,  $\{A_t^{M_{i-1}+j}\}_{j=1, \dots, m_i}$  are independent and follow a Bernoulli distribution with success probability given by

$$\mathbb{P}(A_t^{M_{i-1}+j} = 1 | \mathcal{F}_\infty^\Lambda) = \mathbb{P}(E_j < \Lambda^{(i)}(h^{(i)}(t)) | \mathcal{F}_\infty^\Lambda) = \mathbb{E}[\mathbb{P}(E_j < \Lambda^{(i)}(h^{(i)}(t)) | \Lambda^{(i)}(h^{(i)}(t))] = 1 - e^{-\Lambda^{(i)}(h^{(i)}(t))}$$

for  $j = M_{i-1} + 1, \dots, M_i$  and for  $i = 1, \dots, r$ .

To show the  $L^2$ -convergence of  $L_{m_i}^{(i)}(t)$  we compute

$$\mathbb{E}[L_{m_i}^{(i)}(t)] = G^{(i)}(t),$$

$$\mathbb{E}[L_{m_i}^{(i)}(t)(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})] = \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2],$$

$$\mathbb{E}[L_{m_i}^{(i)}(t)^2] = \frac{G^{(i)}(t)}{m_i} + \frac{m_i - 1}{m_i} \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2].$$

It thus follows that

$$\begin{aligned} & \mathbb{E}[L_{m_i}^{(i)}(t) - (1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2]^2 \\ &= \mathbb{E}[(L_{m_i}^{(i)}(t))^2] - 2\mathbb{E}[L_{m_i}^{(i)}(t)(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})] + \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2]^2 \\ &= \frac{G^{(i)}(t)}{m_i} + \frac{m_i - 1}{m_i} \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2] - \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2]^2 \\ &= \frac{1}{m_i} G^{(i)}(t) - \frac{1}{m_i} \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2] \xrightarrow{m_i \rightarrow \infty} 0. \end{aligned}$$

Similarly, for the overall portfolio loss approximation, we want to prove the  $L^2$  convergence. We have

$$L^n(t) = \frac{1}{n} \sum_{l=1}^n A_t^l = \frac{1}{n} \sum_{i=1}^r \sum_{j=1}^{m_i} A_t^{M_{i-1}+j}$$

and so

$$\hat{L}^n(t) = \mathbb{E}[L^n(t) | \mathcal{F}_\infty^\Lambda] = \frac{1}{n} \sum_{i=1}^r \sum_{j=1}^{m_i} \mathbb{E}[A_t^{M_{i-1}+j} | \mathcal{F}_\infty^\Lambda] = \frac{1}{n} \sum_{i=1}^r \sum_{j=1}^{m_i} (1 - e^{-\Lambda^{(i)}(h^{(i)}(t))}) = \frac{1}{n} \sum_{i=1}^r m_i (1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})$$

$$\begin{aligned}\mathbb{E}[(L^n(t) - \hat{L}^n(t))^2] &= \mathbb{E}\left[\left(\frac{1}{n} \sum_{i=1}^r \frac{m_i}{m_i} \sum_{j=1}^{m_i} A_t^{M_{i-1}+j} - \frac{1}{n} \sum_{i=1}^r m_i (1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})\right)^2\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^r \frac{m_i}{n} \left(\frac{1}{m_i} \sum_{j=1}^{m_i} A_t^{M_{i-1}+j} - (1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})\right)\right)^2\right].\end{aligned}$$

According to our notation we can write the previous formula as

$$\mathbb{E}\left[\left(\sum_{i=1}^r \frac{m_i}{n} (L_{m_i}^{(i)}(t) - L_\infty^{(i)}(t))\right)^2\right]$$

Thus the thesis follows by using the Minkowsky inequality, as we have

$$\left\| \sum_{i=1}^r \frac{m_i}{n} (L_{m_i}^{(i)} - L_\infty^{(i)}) \right\|_{L^2} \leq \sum_{i=1}^r \left\| \frac{m_i}{n} (L_{m_i}^{(i)} - L_\infty^{(i)}) \right\|_{L^2} \xrightarrow{m_i, n \rightarrow \infty} 0.$$

About the upper bound for the approximation error we have

$$\|L_{m_i}^{(i)} - L_\infty^{(i)}\|_{L^2}^2 = \frac{1}{m_i} G^{(i)}(t) - \frac{1}{m_i} \mathbb{E}[(1 - e^{-\Lambda^{(i)}(h^{(i)}(t))})^2] \leq \frac{1}{m_i}$$

and so

$$\left\| \sum_{i=1}^r \frac{m_i}{n} (L_{m_i}^{(i)} - L_\infty^{(i)}) \right\|_{L^2} \leq \sum_{i=1}^r \frac{m_i}{n} \sqrt{\frac{1}{m_i}} = \sum_{i=1}^r \frac{\sqrt{m_i}}{n}.$$

## 1.5 Applications

As pricing application of our dependence model between default times we consider a portfolio of credit-risky assets, and in particular the iTraxx Europe, that can be considered a synthetic CDO being an equally weighted portfolio of  $n = 125$  CDS contracts on the European firms.

### 1.5.1 Pricing CDO tranches

CDOs are constructed by partitioning the credit portfolio in tranches with different seniority: each tranche represents a certain loss piece of the overall portfolio and is defined via its lower and upper attachment points. In particular in the iTraxx Europe there are  $J = 6$  tranches defined by the following lower and upper attachment points  $l^j$  and  $u^j$ ,  $j = 1, \dots, 6$ :

$$[0\%, 3\%], [3\%, 6\%], [6\%, 9\%], [9\%, 12\%], [12\%, 22\%], [22\%, 100\%].$$

The insurance seller receives periodic premium payments depending on the remaining nominal and the spread of the tranche, while the insurance buyer is compensated for losses affecting his tranche. Let us fix a quarterly payment schedule for five years  $T = \{t_0 = 0 < t_1 < \dots < t_M = 5\}$  ( $M = 20$ ) and assume a constant recovery rate  $R = 40\%$  for all companies. The loss  $L_t^{(j)}$  affecting tranche  $j$  up to time  $t$  is linked to the overall portfolio loss  $(1 - R)L_t^n$  via

$$L_t^{(j)} = \min(\max(0, (1 - R)L_t^n - l^j), u^j - l^j),$$

where  $u^j - l^j$  is a cap to the potential loss equal to the whole tranche. The remaining nominal of the portfolio at time  $t$  is given by  $Nom_t = 1 - L_t^n$ , while the remaining nominal of tranche  $j$  is  $Nom_t^{(j)} = u^j - l^j - L_t^{(j)}$ .

Pricing a tranche corresponds to assessing the fair spread such that the expected discounted payment streams of the tranche agree. Defining the Expected Discounted Default Leg for tranche  $j$  ( $EDDL^{(j)}$ ) as the compensations for defaults that affect tranche  $j$

$$EDDL^{(j)} = \sum_{t_k \in T} e^{-rt_k} (\mathbb{E}[L_{t_k}^{(j)}] - \mathbb{E}[L_{t_{k-1}}^{(j)}]) \quad (1.4)$$

and the Expected Discounted Premium Leg ( $EDPL^{(j)}$ ) as the periodic payments depending on the remaining nominal of tranche  $j$

$$EDPL^{(j)} = \sum_{t_k \in T} \Delta t_k e^{-rt_k} s_T^{(j)} (u^j - l^j - \mathbb{E}[L_{t_k}^{(j)}]) \quad (1.5)$$

the fair spread is thus given by the following ratio:

$$s_T^{(j)} = \frac{\sum_{t_k \in T} e^{-rt_k} (\mathbb{E}[L_{t_k}^{(j)}] - \mathbb{E}[L_{t_{k-1}}^{(j)}])}{\sum_{t_k \in T} \Delta t_k e^{-rt_k} (u^j - l^j - \mathbb{E}[L_{t_k}^{(j)}])}, \quad (1.6)$$

where  $e^{-rt_k}$  are the discount factors and  $\Delta t_k = t_k - t_{k-1}$ .

The CDS spread for each tranche is quoted in basis points and the previous equation is true for each tranche except the first, the so-called equity-tranche, for which market convention is to use a running spread of 500 basis points plus an upfront payment quoted as a percentage of the nominal; for this first tranche we have

$$s_T^{(1)} = (EDDL^{(1)} - 500bp * EDPL^{(1)})/u^1. \quad (1.7)$$

For the pricing of each tranche we need to compute

$$\begin{aligned} \mathbb{E}[L^{(j)}(t)] &= \mathbb{E}[\min(\max(0, (1 - R)L^n(t) - l^j), u^j - l^j)] \\ &= \sum_{k=0}^n \mathbb{P}\left(L^n(t) = \frac{k}{n}\right) * \min\left(\max\left(0, (1 - R)\frac{k}{n} - l^j\right), u^j - l^j\right). \end{aligned}$$

We can compute the exact formula above and the formula obtained with the approximation for the portfolio loss distribution:

$$\mathbb{E}[L^{(j)}(t)] \approx \mathbb{E}\left[\min\left(\text{Max}\left(0, (1-R)\left(\frac{1}{n}\sum_{i=1}^r m_i(1-e^{-\Lambda^{(i)}(h^{(i)}(t))})\right) - l^j\right), u^j - l^j\right)\right].$$

In both cases we need to specify the Lévy subordinator that we use to model the dependence structure.

### 1.5.2 The Lévy Subordinator

For our applications we use the three Lévy subordinators described in the third paragraph:

- the *Inverse Gaussian Subordinator*;
- the *Gamma Subordinator*;
- the *Compound Poisson Subordinator*.

Whatever the choice of the subordinator is, in general the dependence structure is determined by the pair of parameters  $(\eta, \beta)$ , while  $\mu$  is indirectly specified by the Lévy measure of the subordinator via the TN condition for the unit exponential distribution  $\Psi(-1) = -1$ .

In particular, in our paper, having defined the subordinator for each class as  $\Lambda^{(i)}(h^{(i)}(t)) = a^{(i)}\Lambda_{h^{(i)}(t)} + b^{(i)}h^{(i)}(t)$ , we can assume without loss of generality  $\mu = 0$ .

In fact let us denote

$$\hat{b}^{(i)} := b^{(i)} + \mu a^{(i)}$$

and

$$\Psi_0(-a^{(i)}) = \Psi(-a^{(i)}) + \mu a^{(i)}$$

and remember the parameters constraint  $b^{(i)} - \Psi(-a^{(i)}) = 1$ .

We have

$$\Psi_0(-a^{(i)}) = \Psi(-a^{(i)}) + \mu a^{(i)} = b^{(i)} - 1 + \mu a^{(i)} = \hat{b}^{(i)} - 1.$$

At this point we can consider  $\mu = 0$  and  $\hat{b}^{(i)} \geq 0$ . We have  $\Psi_0(-a^{(i)}) \geq -1$ .

Once we have  $(\eta, \beta)$ , we can compute the Laplace exponent  $\Psi_0(-a^{(i)})$  for each subordinator. For the Inverse Gaussian subordinator we have:



$$\begin{aligned}
\Psi_{0,IG}(-a^{(i)}) &= \int_0^\infty (e^{-a^{(i)}s} - 1)\nu_{IG}(ds) \\
&= \frac{1}{\sqrt{2\pi}}\beta \int_0^\infty (e^{-a^{(i)}s} - 1)s^{-\frac{3}{2}}e^{-\frac{1}{2}\eta^2s} ds \\
&= \beta \left( \eta - \sqrt{2a^{(i)} + \eta^2} \right).
\end{aligned}$$

The constraint  $\Psi_0(-a^{(i)}) \geq -1$  is translated into the following constraint for  $\beta$ :

$$0 < \beta \leq \frac{\sqrt{2\pi}}{\int_0^\infty (1 - e^{-a^{(i)}s})s^{-\frac{3}{2}}e^{-\frac{1}{2}\eta^2s} ds} = \frac{1}{-\eta + \sqrt{2a^{(i)} + \eta^2}}$$

for each  $i = 1, \dots, r$ , and so, as  $\max\{a^{(i)}\}_{i=1,\dots,r} = a^{(r)}$ , we want

$$0 < \beta \leq \frac{1}{-\eta + \sqrt{2a^{(r)} + \eta^2}}.$$

For the Gamma subordinator we have:

$$\begin{aligned}
\Psi_{0,\Gamma}(-a^{(i)}) &= \int_0^\infty (e^{-a^{(i)}s} - 1)\nu_\Gamma(ds) \\
&= \beta \int_0^\infty (e^{-a^{(i)}s} - 1)\frac{1}{s}e^{-\eta s} ds \\
&= \beta \ln \left( \frac{\eta}{a^{(i)} + \eta} \right)
\end{aligned}$$

and the constraint  $\Psi_0(-a^{(i)}) \geq -1$  is translated into the following constraint for  $\beta$ :

$$0 < \beta \leq \frac{1}{\int_0^\infty (1 - e^{-a^{(i)}s})\frac{1}{s}e^{-\eta s} ds} = \frac{1}{\ln \left[ \frac{a^{(i)} + \eta}{\eta} \right]}$$

for each  $i = 1, \dots, r$ , and so, as again  $\max\{a^{(i)}\}_{i=1,\dots,r} = a^{(r)}$ , we want

$$0 < \beta \leq \frac{1}{\ln \left[ \frac{a^{(r)} + \eta}{\eta} \right]}.$$

Finally, for the Compound Poisson subordinator, we have

$$\begin{aligned}
\Psi_{0,P}(-a^{(i)}) &= \int_0^\infty (e^{-a^{(i)}s} - 1)\nu_P(ds) \\
&= \beta \int_0^\infty (e^{-a^{(i)}y} - 1)D(y) \\
&= \beta \mathbb{E}[e^{-a^{(i)}J_1} - 1] \\
&= -\frac{a^{(i)}\beta}{a^{(i)} + \eta}.
\end{aligned}$$

and the constraint for  $\beta$  becomes

$$0 < \beta \leq \frac{1}{1 - \mathbb{E}[e^{-a^{(r)}J_1}]} = \frac{a^{(r)} + \eta}{a^{(r)}}.$$

Before describing the calibration to market data, it is worth noting that we can model different grades of default dependence between the firms, from a smaller default dependence corresponding to periods of cyclical up trend (sustained economic growth) to a greater default dependence corresponding to recessions periods, by varying the model parameters  $a^{(i)}$ . In particular we obtain a null correlation when  $a^{(i)} \rightarrow 0$ : in this case the default times become independent. Viceversa the correlation is max when  $a^{(i)} \rightarrow \infty$ .

### 1.5.3 The calibration to iTraxx quotes

According to the iTraxx conventions the payment streams for the calibration are quarterly premium payments with the previously specified attachment points for the tranches. The discount factors required as input are obtained from risk-free par yields. The market quotes to which the model is calibrated comprise the portfolio CDS spreads with maturities three and five years and the spreads for the tranches. In particular we use the CDS spreads to calibrate the marginal distributions, and the tranches spreads to calibrate the subordinator parameters. One week of daily data is used from the seventh series of iTraxx Europe with maturity 5 years ranging from June 20, 2007, to June 26, 2007, and a calibration is run for each of these days.

We divide our basket of 125 firms in two classes according to two different criteria. For the first application we divide the firms according to their spread

level (meaning considering the medium spread in the considered period). For the second application we consider the class of the 25 financial firms and the class of the remaining 100 firms.

The calibration procedure can be split in two different steps, having separated the dependence structure and marginal default probabilities. The first step involves the calibration of the marginal distributions  $G_i$ ,  $\{i = 1, \dots, 125\}$ , for which a piecewise linear intensity is assumed. In particular we consider a CDS with maturity 3 years and a CDS with maturity 5 years. For each firm we have

$$1 - G_i(t) = e^{-h_i(t)} = e^{-\int_0^t \lambda_i(s) ds}$$

where we assume the default intensity  $h_i(t)$  given by

$$h_i(t) = \int_0^t \lambda_i(s) ds = \int_0^t \lambda_i^3 \min\{s, 3\} + \lambda_i^5 (s - 3) \mathbf{1}_{\{s > 3\}} ds,$$

with  $\lambda_i^3$  and  $\lambda_i^5$  being positive intensity parameters which are calibrated to the portfolio-CDS spreads for the 3-year and 5-year contract, respectively. We choose  $\lambda_i^3$  so that the market 3-year CDS spread agrees with our CDS spread computed using the discrete formula

$$modelCDSspread_i^3 = \frac{\frac{1-R}{(1+r_1)} p_1^{i,3} + \frac{1-R}{(1+r_2)^2} p_2^{i,3} + \frac{1-R}{(1+r_3)^3} p_3^{i,3}}{1 + \frac{1-p_1^{i,3}}{1+r_1} + \frac{1-p_1^{i,3}-p_2^{i,3}}{(1+r_2)^2}}$$

where

- $p_t^{i,3}$ ,  $t = 1, \dots, 3$ , are the discrete default probabilities for the event "For the firm  $i$  there will be default in the year  $t$ ", that we compute from our  $G_i(t)$  as

$$p_t^{i,3} = G_i(t) - G_i(t-1)$$

with  $G_i(0) = 0$ ;

- $r_t$ ,  $t = 1, \dots, 3$ , are the risk free interest rates with maturity  $T = t$  used for the discount factors: in practice we directly use the discount factors related to the Germany zero curves, downloaded from Datastream.

For the calibration of  $\lambda_i^5$ , we use instead the 5-year CDS contract and follow the same procedure with

$$modelCDSspread_i^5 = \frac{\frac{1-R}{(1+r_1)} p_1^{i,3} + \frac{1-R}{(1+r_2)^2} p_2^{i,3} + \frac{1-R}{(1+r_3)^3} p_3^{i,3} + \frac{1-R}{(1+r_4)^4} p_4^{i,5} + \frac{1-R}{(1+r_5)^5} p_5^{i,5}}{1 + \frac{1-p_1^{i,3}}{1+r_1} + \frac{1-p_1^{i,3}-p_2^{i,3}}{(1+r_2)^2} + \frac{1-p_1^{i,3}-p_2^{i,3}-p_3^{i,3}}{(1+r_3)^3} + \frac{1-p_1^{i,3}-p_2^{i,3}-p_3^{i,3}-p_4^{i,5}}{(1+r_4)^4}}.$$

The second step involves the calibration of the parameters of the subordinator for the two classes in which the iTraxx underlying firms are divided. The intensity parameters for each firm,  $\lambda_i^3$  and  $\lambda_i^5$ , are thus fixed and from them we compute  $h_i(t)$  for each firm in the iTraxx. We then compute  $h^{(i)}(t)$  for the two considered classes by taking the average value in each class, as  $h^{(i)}(t) = \frac{1}{m_i} \sum_{i=1}^{m_i} h_i(t)$ . The parameters of the subordinator,  $(\eta, \beta)$ , specifying the dependence, as well as the parameters values  $a^{(i)}$ , are calibrated to observed market spreads of the tranches of the CDO. For this, we consider different couples  $(a^{(1)}, a^{(2)})$  and for each couple we define a grid for  $\eta$ . Given  $\eta$ , our subordinator parameters constraint (deriving from the TN condition) defines an interval for  $\beta$ . On this interval,  $\beta$  is chosen so that the tranche equity (i.e. the upfront payment) is perfectly matched. Finally, among the obtained possible parameters,  $(a^{(1)}, a^{(2)}, \eta, \beta)$  are chosen to be the minimizer of the sum of square deviations of market to model spreads over all tranches  $j = 2, \dots, 5$ .<sup>1</sup> For this we solve, using Matlab, the following minimization problem:

$$\min_{(\eta, \beta, a^{(i)})} \sum_{j=2}^5 (\text{marketspread}^j - s_T^j)^2$$

where  $\text{marketspread}^j$  is observed on Bloomberg, and  $s_T^j$  is computed by using 1.6.

## 1.6 Distress Dependence and Systemic Risk

The multivariate default distribution gives us the joint probability of distress, that represents the probability of all the institutions in the system (portfolio) becoming distressed, i.e., the tail risk of the system. It is an empirical fact that the probability that all the banks in the system experience large losses simultaneously, which embeds the distress dependence, increases in times of financial distress, and therefore, in such periods, the financial system's joint probability of distress may experience larger and nonlinear increases than those experienced by the average probabilities of default of the individual institutions. Having estimated the multivariate default distribution of the companies included in the iTraxx, we follow Segoviano and Goodhart (2010) (see [8] and [7]), to analyze the distress dependence in the portfolio by computing a set of indicators of systemic risk. In particular we estimate three stability measures that incorporate changes in distress dependence that are consistent with the economic cycle. The stability measures that we use are:

### 1. *The Stability Index;*

---

<sup>1</sup>We don't consider the super senior tranche [22%, 100%] as this tranche is traded very rarely and we don't have market quotas for it.

2. *The Distress Dependence Matrix*;
3. *The Probability of Cascade Effects*.

Once we have computed these stability measures, we employ them to verify which firms are more systemically relevant for the index as a whole.

### 1.6.1 The Stability Index

The Stability Index (SI) gives a measure of the tail risk of the system, i.e. the common distress of the financial institutions in the system. The SI is based on the conditional expectation of default probability measure developed by Huang (1992) (see [10]), and measures the expected number of other institutions that would fall into distress given that at least a specific institution has become distressed (i.e., were to default). The SI represents a probability measure that conditions on any institution becoming distressed, without indicating the specific bank. In the simplest case of two financial institutions with default times  $\tau_i$  and  $\tau_j$ , let  $\kappa_t$  stand for the number of institutions in default at time  $t$ . Our extreme linkage indicator is the conditional expectation  $\mathbb{E}[\kappa_t | \kappa_t \geq 1]$ . From elementary probability theory we have

$$\begin{aligned} SI = \mathbb{E}[\kappa_t | \kappa_t \geq 1] &= \frac{\mathbb{P}\{\tau_i \leq t, \tau_j > t\} + \mathbb{P}\{\tau_i > t, \tau_j \leq t\} + 2\mathbb{P}\{\tau_i \leq t, \tau_j \leq t\}}{\mathbb{P}\{\tau_i \leq t \text{ or } \tau_j \leq t\}} \\ &= \frac{\mathbb{P}\{\tau_i \leq t\} + \mathbb{P}\{\tau_j \leq t\}}{1 - \mathbb{P}\{\tau_i > t, \tau_j > t\}} \end{aligned}$$

As Huang (1992) shows, this measure can be interpreted as a relative measure of the *system linkage*: when  $SI = 1$  in the limit, the system linkage is weak (meaning asymptotic independence), while as the value of the  $SI$  increases, the system linkage increases (meaning asymptotic dependence).

In our portfolio of 125 firms we divide the group of the 25 financial institutions, that we consider as the entity  $i$ , from the group of the other 100 institutions, that we consider as the entity  $j$ , and we compute the SI with the formula above.

### 1.6.2 The Distress Dependence Matrix

Distress Dependence between two financial institutions is a measure that computes the probability that an institution becomes distressed conditional on another entity becoming distressed. Such measure allows analyzing financial stability. The distress dependence matrix is a matrix based on market data in which are collected pairwise probabilities of financial institutions experiencing distress conditional on other institutions being in distress. It thus accounts for the relationship between

the institutions. Basically the elements of the matrix show the conditional probabilities of distress of the institution in the row, given that the institution in the column falls into distress. For each pair of institutions in the portfolio, we estimate the pairwise conditional probabilities of distress: the probability of distress of institution  $i$  conditional on institution  $j$  becoming distressed is computed as

$$\mathbb{P}(\tau_i \leq t | \tau_j \leq t) = \frac{\mathbb{P}(\tau_i \leq t, \tau_j \leq t)}{\mathbb{P}(\tau_j \leq t)}.$$

Let us denote  $\mathbb{P}(\tau_i \leq t | \tau_j \leq t) := \mathbb{P}(\text{Firm}i | \text{firm}j)$ ; these pairwise conditional probabilities of distress are represented in the following Distress Dependence Matrix:

	Firm1	Firm <i>i</i> ( $i = 2, \dots, 124$ )	Firm125
Firm1	1	$\mathbb{P}(\text{Firm1}   \text{Firm}i)$	$\mathbb{P}(\text{Firm1}   \text{Firm125})$
Firm <i>i</i> ( $i = 2, \dots, 124$ )	$\mathbb{P}(\text{Firm}i   \text{Firm1})$	1	$\mathbb{P}(\text{Firm}i   \text{Firm125})$
Firm125	$\mathbb{P}(\text{Firm125}   \text{Firm1})$	$\mathbb{P}(\text{Firm125}   \text{Firm}i)$	1

This  $125 \times 125$  matrix contains the probability of distress of the institution specified in the row, given that the institution specified in the column becomes distressed. Even if conditional probabilities do not imply causation, this set of pairwise conditional probabilities can provide important insights into interlinkages and the likelihood of contagion between the institutions in the system.

### 1.6.3 The Probability of Cascade Effects

The Probability of Cascade Effects is an indicator that measures the likelihood that one, two, or more institutions, up to the total number of financial institutions in the system become distressed given that a specific financial institution becomes distressed. In this way it quantifies the potential "cascade" effects in the system given the distress in a specific financial institution, and so this measure can be considered as an indicator that allows to quantify the systemic importance of a specific institution if it becomes distressed. For this systemic risk indicator we divide our portfolio into four groups: the group of the financial institutions with lower spread (denoted FLS), the group of the financial institutions with higher spread (denoted FHS), the group of the non financial institutions with lower spread

(denoted NFLS) and the group of the non financial institutions with higher spread (denoted NFHS). If we consider the higher spread financial institutions group (FHS) in distress, the Probability of Cascade effects can be defined as:

$$\begin{aligned}
PCE = & \mathbb{P}(FLS|FHS) + \mathbb{P}(NFLS|FHS) + \mathbb{P}(NFHS|FHS) \\
& - [\mathbb{P}(FLS \cap NFLS|FHS) + \mathbb{P}(FLS \cap NFHS|FHS) + \mathbb{P}(NFLS \cap NFHS|FHS)] \\
& + \mathbb{P}(FLS \cap NFLS \cap NFHS|FHS). \tag{1.8}
\end{aligned}$$

## 1.7 Conclusions

A multivariate default times model for a portfolio of assets exposed to credit risk was constructed using a conditional independence approach with a stochastic time-change as common factor. The dependence structure was kept separated from the parameters of the marginal default probabilities by choosing a suitable Lévy subordinator as stochastic clock. Thanks to this separation between the univariate marginals and the dependence structure, the implied copula of the default times could be computed explicitly. Under the assumption of an heterogeneous portfolio, a closed formula for the portfolio loss distribution was obtained, and an approximation for large portfolios was presented. The model efficiency was demonstrated by calibrating it to observed portfolio-CDS and CDO spreads, using an appropriate Lévy subordinator. Measures of portfolio systemic risk were computed.

# A Appendix

## Copula Functions

**Definition 2** (Copula function). *A copula is an  $n$ -dimensional distribution function  $C : [0, 1]^n \rightarrow [0, 1]$  of a random vector  $(U_1, \dots, U_n)$ , where the marginal law of  $U_i$  is the uniform distribution on  $[0, 1]$  for all  $i \in \{1, \dots, n\}$ .*

Copula functions are very popular in the study of multivariate distribution functions thanks to their role in imposing a dependence structure on predetermined marginal distributions. Their importance derives from the *Sklar's theorem* that proves that any multivariate distribution function can be characterized by a copula, and that copula functions, together with univariate marginal distribution functions, can be used to construct multivariate distribution functions.

**Theorem 2** (Sklar's theorem). *Let  $H$  be an  $n$ -dimensional distribution function with marginals  $F_1, \dots, F_n$ .*

*Then it exists an  $n$ -copula  $C$  such that, for each  $\mathbf{x} \in \mathbf{R}^n$ ,*

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)).$$

*If the marginals  $F_1, \dots, F_n$  are all continuous, then  $C$  is unique; otherwise  $C$  is univocally determined on  $(\text{Ran}F_1 \times \text{Ran}F_2 \times \dots \times \text{Ran}F_n)$  (where  $\text{Ran}F_i$  denotes the rank of  $F_i$ ). Conversely, if  $C$  is an  $n$ -copula and  $F_1, \dots, F_n$  are distribution functions, then the function  $H$  above defined is an  $n$ -dimensional distribution function with marginals  $F_1, \dots, F_n$ .*

The proof of this theorem can be found in [9].

The main feature of the Sklar's theorem is that for continuous multivariate distribution functions, the univariate marginals and the multivariate dependence structure can be separated, and the dependence structure can be represented by a copula.

Let  $F$  be an univariate distribution function. Let's remember that the generalised inverse of  $F$  is defined as  $F^{-1}(t) = \inf\{x \in \mathbf{R} | F(x) \geq t\}$  for each  $t$  in  $[0, 1]$ , with the usual convention that  $\inf() = -\infty$ .

An important corollary of the Sklar's theorem, that is fundamental in the study of copulas and their applications, is the following:

**Cor 2.** *Let  $H$  be an  $n$ -dimensional distribution function with continuous marginals  $F_1, \dots, F_n$  and copula  $C$ . Then for each  $\mathbf{u} \in [0, 1]^n$ ,*

$$C(u_1, \dots, u_n) = H(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)).$$



## Levy Processes

**Definition 3** (Lévy process). *A Lévy process is any continuous-time stochastic process  $X = \{X_t : t \geq 0\}$  such that*

1.  $X_0 = 0$  almost surely;
2. *It has independent increments: for any  $0 \leq t_1 < t_2 < \dots < t_n < \infty$ ,  $X_{t_2} - X_{t_1}$ ,  $X_{t_3} - X_{t_2}$ , . . . ,  $X_{t_n} - X_{t_{n-1}}$  are independent;*
3. *It has stationary increments: for any  $s < t$ ,  $X_t - X_s$  is equal in distribution to  $X_{t-s}$ ;*
4.  $t \rightarrow X_t$  is almost surely right continuous with left limits.

*The most well-known examples of Lévy processes are the Wiener process and the Poisson process.*

## Chapter 2

# On the Relationship between the Risk of Default and the Yield-to-Maturity of Bonds

### 2.1 Introduction

Credit Default Swap (CDS) spreads and bond spreads (i.e. the spreads between the bond yield rate and the risk free rate) have become commonly used as default risk indicators for risk analysis. For example, in Fitch (2010b), we find an analysis of the CDS spread history and implied annual probability of default for the U.S. broker-dealers over the past years, while in Fitch (2010a) is studied how the directional momentum in CDS spreads affects its performance as indicator of default risk during a stress period. (On the link between CDS spreads and default probability see also Bank of England, 2009). However, both CDS and bond spreads depend not only on variables directly linked to the risk of default but also on the specific structure of the contract. In particular, in this chapter we show that bond spreads can be misleading if used to infer the default probability of the issuers, and consequently the yield-to-maturity must be cautiously interpreted as an indicator of the bond default risk.<sup>1</sup> In fact, the yield rate, for given default probabilities and recovery rates, can considerably vary as a function of the residual life and the coupon value of the bond. In particular, when there is default risk, bonds with high coupons are more likely to have high yield rates too (and viceversa). The intuition behind this result is that greater coupons are associated with

---

<sup>1</sup>In particular, without loss of generality, in this chapter we assume that the risk-free rate is constant in the considered time horizon, so that the yield level provides information analogous to the spread between the yield rate and the risk-free rate. Anyway the inferred considerations are still valid, on a quality level, even in case the risk-free rate curve is not flat.

less than proportional price increases, because there is a probability, linked to the default likelihood, that the coupons are not actually payed-off. Bond prices which are relatively low with respect to the nominal payment flows (on which the yield is computed) determine nominal yields which are relatively higher. This implies that bonds with higher default risk can have lower yields (and viceversa), just as a consequence of their coupon structures.

Also the slope of the yield curve must be cautiously interpreted as in general it does not convey enough information to establish if the default risk is higher in one period than in another period. We show that a downward sloping yield curve does not necessarily imply that the default probability on shorter maturities is higher than on longer maturities. This result arises from the fact that also the yield curve slope is linked to the coupon rate: taking fixed the other variables (in particular the default probability and the recovery rate) bonds with low coupons determine decreasing yield curves, while bonds with high coupons imply increasing yield curves. The intuition behind this result is that higher coupons determine losses relatively higher for bonds with longer maturity in case of default, and consequently higher yields for these bonds. On the contrary, when the coupons are low, the nominal losses in case of default are similar both for short term and long term bonds;<sup>2</sup> it follows that the prices of longer maturity bonds, for which the losses are likely in far away time horizons, are relatively higher, and the corresponding yields are lower.

Most of the literature available on the valuation of fixed income securities (see for example Fabozzi, 2003, 2007), is focused on the interest rate risk (i.e. the bond's price sensitivity to the change in interest rates) and the concept of duration is used to describe the relationship between the bond maturity and coupon rate, and the bond price sensitivity (a longer maturity and a lower coupon rate are linked to a greater price sensitivity to interest rate changes); in this context a higher bond yield is considered as a premium for the higher interest rate risk. In this chapter we study similar financial indicators that could be used in presence of credit (or default) risk to properly evaluate the relationship between the defaultable bond yield and its default probability.

---

<sup>2</sup>For the zero-coupon bonds, for example, in case of default the loss is always only given by the amount lost on the bond final pay-off.

## 2.2 The mathematical model

Let us consider a discrete times model with times  $n = 0, 1, \dots, N$ , in which the probability of being in default at time  $n \geq 1$ , conditioning to not having incurred in default in the previous  $n - 1$  periods, is constant and equal to  $\lambda$ . It thus follows that the unconditioned probability of default has a geometric distribution, i.e. the unconditional probability that the default time is equal to  $n$  is given by  $\mathbb{P}(\tau = n) = \lambda(1 - \lambda)^{n-1}$  for  $n \geq 1$ .

Let us assume a constant risk-free rate  $r$  and consider a defaultable bond with a constant conditional default probability  $\lambda$ , nominal value 100, maturity  $N$ , annual coupons  $c$  and constant recovery rate  $R/100$  in case of default. Under the hypothesis of risk neutral investors, the price at time 0 of the bond with maturity  $N$ ,  $P(N)$ , is given by the expected value of the future payment flows (expected payoff) discounted at the risk-free rate:

$$P(N) = \sum_{n=1}^N c \frac{(1 - \lambda)^n}{(1 + r)^n} + \sum_{n=1}^N R \frac{\lambda(1 - \lambda)^{n-1}}{(1 + r)^n} + 100 \frac{(1 - \lambda)^N}{(1 + r)^N}. \quad (2.1)$$

The first term in the right side of equation (2.1) is the present value of the coupons, which are payed only when there is no default before the time they are due; the second term represents present value of the recovery rate, that is payed when the default happens; the last term is the present value of the bond face value, that is payed at maturity only in case of no previous default.

We define the bond yield-to-maturity as in Hull (2008):

**Definition 4.** *The bond yield-to-maturity is the unique value  $y$  such that*

$$P(N) = \sum_{n=1}^N \frac{c}{(1 + y)^n} + \frac{100}{(1 + y)^N}. \quad (2.2)$$

Another concept that we need when we value fixed income securities is that of par yield, which corresponds to the coupon rate for which the bond is quoted at par:

**Definition 5.** *The par yield is the coupon rate for which the price of the bond is equal to its par value (100), i.e. is the unique value  $c_{par}/100$  such that*

$$P(N) = \sum_{n=1}^N c_{par} \frac{(1 - \lambda)^n}{(1 + r)^n} + \sum_{n=1}^N R \frac{\lambda(1 - \lambda)^{n-1}}{(1 + r)^n} + 100 \frac{(1 - \lambda)^N}{(1 + r)^N} = 100. \quad (2.3)$$

We aim at showing that the yield-to-maturity must be carefully interpreted as a bond default risk indicator because its value depends also on other bond characteristics, such as the maturity and the coupon value. Let us start with the following

lemma that shows that the par yield is unique for all bonds, independently from the maturity:

**Lemma 2.** *There exists one unique coupon value,  $c_{par}$ , for which all bond prices are equal to 100, independently from the bond maturities. Such unique value is given by*

$$c_{par} = \frac{100r + (100 - R)\lambda}{1 - \lambda}. \quad (2.4)$$

Since a bond quoted at par with coupons  $100y$  has yield-to-maturity equal to  $y$ <sup>3</sup>, the yield-to-maturity of a bond with coupons equal to  $c_{par}$  is equal to the par yield. It follows that the term structure for bonds with coupons equal to  $c_{par}$  is flat at the value  $c_{par}/100$ , whatever is the value of the other variables (default rate, recovery rate, risk-free rate). Said in other words, equation (2.4) shows that the values of the default rate, the recovery rate and the risk-free rate determine the value  $c_{par}$  which, however, is independent of the maturity.

From Lemma 2 we have that bonds with different maturity but analogous characteristics in terms of default risk (default probability and recovery rate) have the same yield when the coupons are equal to  $c_{par}$ . Unfortunately this is an exception, as it is inferred in the following proposition:

**Proposition 1.** *The yield of a bond with maturity  $N$  is an increasing function of the bond coupons.*

According to Proposition 1, given two bonds with coupons  $c_1 = c_{par} + \Delta_1$  and  $c_2 = c_{par} + \Delta_1 + \Delta_2$  (with  $\Delta_i > 0$ ,  $i = 1, 2$ ), there exist  $\delta_1$  and  $\delta_2$  (with  $\delta_i > 0$ ,  $i = 1, 2$ ) such that the related yields are given by  $y_1 = c_{par}/100 + \delta_1$  and  $y_2 = c_{par}/100 + \delta_1 + \delta_2$ .<sup>4</sup> An interesting result which arises from the proof of Proposition 1 is that when  $R = 0$  or  $\lambda = 0$  one has  $\delta = 0$  (as equation (A.12) in the proof is actually an identity). This means that when the recovery rate is equal to zero, the bond yield is always equal to the par yield, independently from the other variables value. We get an analogous result when the default probability is equal to zero; in particular, in this last case, the par yield is equal to the risk-free rate.

The following proposition extends the previous results by showing that the term structure for bonds with different residual life depends also on the coupon rate (assumed to be the same for all the bonds).

**Proposition 2.** *The yield term structure for  $N$  bonds with same characteristics except the maturity ( $n = 1, \dots, N$ ) is upward sloping (downward sloping) when the coupon rate is higher (lower) than  $c_{par}$ . The term structure of the bond prices follows analogous trends.*

<sup>3</sup>Cf. the proof of Proposition 1 in the Appendix

<sup>4</sup>An analogous relationship holds when...

## 2.3 Numerical examples

Let us show now a few numerical examples that highlight what we have proved formally. Let us consider the case in which  $r = 3\%$ ,  $R = 80$  e  $\lambda = 1\%$ . In this case, using equation (2.4), we have  $c_{par} = 3.23\%$ . The related bond yields for different maturities and coupon rates are:

N	c=0%	c=1%	c=5%
1	3.21%	3.21%	3.25%
5	3.15%	3.18%	3.27%
10	3.08%	3.13%	3.30%
15	3.00%	3.09%	3.32%

This example highlights two interesting features. The first one is that, for a given coupon rate, the yield is a monotonic function of the bond residual life. In particular, when the coupon rate is lower (higher) than 3.23% the yield curve is downward (upward) sloping. The second relevant feature is that the yield is an increasing function of the coupon rate, for given maturity. The intuition behind this result is that greater coupons are associated with less than proportional price increases, because there is a probability, linked to the default likelihood, that the coupons are not actually payed-off. In this case, bond prices are relatively low with respect to the nominal cash flow (on which the yield is computed) and nominal yields are relatively higher. Let us remark, anyway, that the greater bond risk, and the associated greater nominal yield, are not determined by a bigger default probability or a smaller recovery rate, but are given by other characteristics of the bond (maturity and coupon rate). It follows that it is possible to have different nominal yields for given default probabilities and recovery rates and viceversa.

Another problem that arises when bond yields are used to extract information about the issuer default probability, is that this indicator does not reflect the effective bond duration, that depends on the default rates. Bonds with higher default rates have lower effective maturity, and this can be reflected in relatively higher prices (as the payment flows are discounted for shorter time periods) that determine particularly low yield rates (computed on the bond residual life). Let us for example assume  $r = 3\%$ ,  $R = 80$  and  $\lambda = 10\%$ ; under these assumptions we have  $c_{par} = 5,56\%$  and the following yields:

N	c=0%	c=3%	c=10%
1	5.10%	5.35%	5.89%
5	4.27%	5.01%	6.38%
10	3.41%	4.67%	6.79%
15	2.74%	4.42%	7.06%

This example shows that the yield rate of a zero-coupon bond ( $c = 0\%$ ) with long maturity ( $N = 15$ ) can even be lower than the risk-free rate, in spite of the high default rate. This happens because with such high default rates the probability that the face value of the bond is payed-off at maturity is just about 20%. In the remaining 80% of cases the bond will be repayed before maturity and discounting the payoff for a shorter time period offsets the loss of 20% on the nominal.<sup>5</sup>

Finally, let us consider the case in which  $r = 3\%$  and  $R = 80$ :

	$\lambda = 1\%$	$\lambda = 5\%$	$\lambda = 10\%$
N	$c=15\%$	$c=2\%$	$c=0\%$
1	3.31%	4.12%	5.10%
5	3.43%	3.96%	4.27%
10	3.53%	3.77%	3.41%
15	3.60%	3.61%	2.74%

We can observe that the yield for the first two bonds with maturity 15 years are basically analogous, notwithstanding that the annual default probability of the second bond is 5 times higher than the annual default probability of the first bond. Let us also note that the zero-coupon bonds with maturities 10 and 15 years have yield rates lower than those of the first two bonds for the same maturities, even if the annual default probabilities for the zero-coupon bonds are much higher.

## 2.4 Conclusion

The theoretical results achieved in this paper, and the empirical evidence allow us to infer the following considerations:

- The bond yield rates depend not only on the default probability and the recovery rate (and of course the risk-free rate), but also on the bond structural features as the coupons value and the nominal residual life.
- The bond yields must be cautiously used to deduce information about the issuers default probability.
- The issuers default probability must be explicitly computed using the bond structural features and the related prices.

---

<sup>5</sup>Let us observe that if the risk-free rate were lower, the gain in terms of expected value coming from the anticipated repayment could not be sufficient to compensate the loss. For example, when  $r = 2\%$ , the yield rate for the bond with maturity 15 years is equal to 2,24%, that is higher than the risk-free rate.

## A Appendix (proofs)

### Proof of Lemma 2

For  $N = 1$  one has

$$P_1 = c \frac{1 - \lambda}{1 + r} + R \frac{\lambda}{1 + r} + 100 \frac{1 - \lambda}{1 + r} \quad (\text{A.5})$$

so that the value of the coupon  $c_{par}$  for which  $P_1 = 100$  is equal to

$$c_{par} = \frac{100r + (100 - R)\lambda}{1 - \lambda}. \quad (\text{A.6})$$

From equation (2.1) one has

$$P(N) = P(N - 1) + \left( c \frac{(1 - \lambda)^N}{(1 + r)^N} + R \frac{\lambda(1 - \lambda)^{N-1}}{(1 + r)^N} + 100 \frac{(1 - \lambda)^N}{(1 + r)^N} - 100 \frac{(1 - \lambda)^{N-1}}{(1 + r)^{N-1}} \right). \quad (\text{A.7})$$

The value of  $c_{par}$  from equation (A.6) is such that the second term in equation (A.7) is equal to zero so that, for induction, the price of a bond with arbitrary maturity and coupons equal to  $c_{par}$  is always equal to 100.  $\square$

### Proof of Proposition 1

Let us start the proof showing that coupons higher than  $c_{par}$  imply, with all the other factors being the same, yields higher than the par yield, and viceversa. Let us first show that, for any bond with coupon  $c$ , we have

$$100 = \sum_{n=1}^N \frac{c}{(1 + c/100)^n} + \frac{100}{(1 + c/100)^N}. \quad (\text{A.8})$$



In fact we have

$$\begin{aligned}
\sum_{n=1}^N \frac{c}{(1+c/100)^n} + \frac{100}{(1+c/100)^N} &= c \sum_{n=1}^N \left( \frac{1}{1+c/100} \right)^n + \frac{100}{(1+c/100)^N} \\
&= c \frac{\frac{1}{1+c/100} - \left( \frac{1}{1+c/100} \right)^{N+1}}{1 - \frac{1}{1+c/100}} + \frac{100}{(1+c/100)^N} \\
&= c \frac{\frac{100}{100+c} \left( 1 - \left( \frac{100}{100+c} \right)^N \right)}{c \frac{1}{100+c}} + \frac{100}{\left( \frac{100+c}{100} \right)^N} \\
&= 100 - 100 \left( \frac{100}{100+c} \right)^N + 100 \left( \frac{100}{100+c} \right)^N \\
&= 100.
\end{aligned} \tag{A.9}$$

Let us remark that this is true in particular when  $c = c_{par}$ , in which case we have

$$100 = \sum_{n=1}^N \frac{c_{par}}{(1+c_{par}/100)^n} + \frac{100}{(1+c_{par}/100)^N}. \tag{A.10}$$

Lemma 1 tells us that the yield of any bond with maturity  $N$  is equal to the par yield when the coupons are equal to  $c_{par}$ .

Moreover, from equation (2.1) we have that when a bond has a coupon which differs from  $c_{par}$  by  $\Delta$ , the variation of its yield  $\delta$  is imolcilty defined by

$$100 + \sum_{n=1}^N \Delta \frac{(1-\lambda)^n}{(1+r)^n} = \sum_{n=1}^N \frac{c_{par} + \Delta}{(1+c_{par}/100 + \delta)^n} + \frac{100}{(1+c_{par}/100 + \delta)^N}. \tag{A.11}$$

If by absurd we assume  $\delta = 0$ , equation (A.11) would become, using (A.10),

$$\Delta \sum_{n=1}^N \frac{(1-\lambda)^n}{(1+r)^n} = \Delta \sum_{n=1}^N \frac{1}{(1+c_{par}/100)^n} = \Delta \sum_{n=1}^N \left( \frac{1-\lambda}{1+r - R\lambda/100} \right)^n \tag{A.12}$$

where in the last identity we have used the  $c_{par}$  definition in (A.6). We can easily observe that the last term in equation (A.12) is actually greater (lower) than the first term when  $\Delta > 0$  ( $\Delta < 0$ ). It follows that  $\delta$  must be positive (negative) when  $\Delta$  is positive (negative) in order to satisfy equation (A.11).

Let us now note that the yield increment  $\delta$  is always lower than  $\delta_{sup} = \frac{R\lambda/100}{1-\lambda}$  (i.e., the bond yield has an upper bound, whatever is the coupon value). In fact, we have on one end, using definition 5, when  $c = c_{par} + \Delta$ ,

$$\begin{aligned}
100 + \sum_{n=1}^N \Delta \frac{(1-\lambda)^n}{(1+r)^n} &= \sum_{n=1}^N c_{par} \frac{(1-\lambda)^n}{(1+r)^n} + \sum_{n=1}^N \frac{R\lambda(1-\lambda)^{n-1}}{(1+r)^n} \\
&\quad + 100 \frac{(1-\lambda)^N}{(1+r)^N} + \sum_{n=1}^N \Delta \frac{(1-\lambda)^n}{(1+r)^n} \\
&\geq \sum_{n=1}^N c_{par} \frac{(1-\lambda)^n}{(1+r)^n} + 100 \frac{(1-\lambda)^N}{(1+r)^N} + \sum_{n=1}^N \Delta \frac{(1-\lambda)^n}{(1+r)^n},
\end{aligned} \tag{A.13}$$

where the equation is satisfied only for  $R = 0$  or  $\lambda = 0$ ; on the other end, when  $\delta = \delta_{sup}$ , equation (A.11) becomes

$$\begin{aligned}
100 + \sum_{n=1}^N \Delta \frac{(1-\lambda)^n}{(1+r)^n} &= \sum_{n=1}^N \frac{c_{par} + \Delta}{(1 + c_{par}/100 + \delta_{sup})^n} + \frac{100}{(1 + c_{par}/100 + \delta_{sup})^N} \\
&= \sum_{n=1}^N \frac{c_{par} + \Delta}{(1 + c_{par}/100 + \frac{R\lambda}{100(1-\lambda)})^n} + \frac{100}{(1 + c_{par}/100 + \frac{R\lambda}{100(1-\lambda)})^N} \\
&= \sum_{n=1}^N c_{par} \frac{(1-\lambda)^n}{(1+r)^n} + 100 \frac{(1-\lambda)^N}{(1+r)^N} \sum_{n=1}^N + \Delta \frac{(1-\lambda)^n}{(1+r)^n}.
\end{aligned} \tag{A.14}$$

It follows that, except in the particular cases  $R = 0$  or  $\lambda = 0$ , in which  $\delta = \delta_{sup} = 0$ , we need  $\delta < \delta_{sup}$  in order to satisfy equation (A.11).

To complete the proof let us see what happens to a bond yield when to a first coupon increment  $\Delta_1 > 0$  we add a further increment  $\Delta_2 > 0$ .<sup>6</sup> In this case we must show that the bond yield  $c_{par}/100 + \delta_1$  (related to the bond with coupon  $c_{par} + \Delta_1$ ), increases of the positive value  $\delta_2$  implicitly defined by

$$\begin{aligned}
100 + \sum_{n=1}^N \Delta_1 \frac{(1-\lambda)^n}{(1+r)^n} + \sum_{n=1}^N \Delta_2 \frac{(1-\lambda)^n}{(1+r)^n} \\
= \sum_{n=1}^N \frac{c_{par} + \Delta_1 + \Delta_2}{(1 + c_{par}/100 + \delta_1 + \delta_2)^n} + \frac{100}{(1 + c_{par}/100 + \delta_1 + \delta_2)^N}.
\end{aligned} \tag{A.15}$$

---

<sup>6</sup>The proof for negative variations is analogous.

As previously, if by absurd we assume  $\delta_2 = 0$ , equation (A.15) would become

$$\begin{aligned} \Delta_2 \sum_{n=1}^N \frac{(1-\lambda)^n}{(1+r)^n} &= \Delta_2 \sum_{n=1}^N \frac{1}{(1+c_{par}/100+\delta_1)^n} \\ &= \Delta_2 \sum_{n=1}^N \left( \frac{1-\lambda}{1+r-R\lambda/100+\delta_1(1-\lambda)} \right)^n. \end{aligned} \quad (\text{A.16})$$

As  $\delta_1 < \delta_{sup}$ , we can note that the last term in equation (A.16) is actually greater than the first. It follows that we need  $\delta_2 > 0$  to satisfy equation (A.15).  $\square$

## Proof of Proposition 2

To prove that coupons greater than  $c_{par}$  imply increasing yield curves we must show that, for a bond with maturity  $N$  and coupon  $c_{par} + \Delta$ , with  $\Delta > 0$ , the  $\delta$  implicitly defined by the following equation is positive.<sup>7</sup>

$$\begin{aligned} 100 + \sum_{n=1}^N \Delta \frac{(1-\lambda)^n}{(1+r)^n} &= \\ &= \sum_{n=1}^N \frac{c_{par} + \Delta}{(1+c_{par}/100+\delta_{N-1}+\delta)^n} + \frac{100}{(1+c_{par}/100+\delta_{N-1}+\delta)^N}, \end{aligned} \quad (\text{A.17})$$

where  $\delta_{N-1}$  is the difference with respect to  $c_{par}$  for a bond with maturity  $N-1$  and coupon equal to  $c_{par} + \Delta$ .

Let us start assuming, by absurd, that  $\delta = 0$ . We could write equation (A.17), using (A.6), as

$$\begin{aligned} \Delta \frac{(1-\lambda)^N}{(1+r)^N} &= \frac{c_{par} + \Delta}{(1+c_{par}/100+\delta_{N-1})^N} \\ &+ \frac{100}{(1+c_{par}/100+\delta_{N-1})^N} - \frac{100}{(1+c_{par}/100+\delta_{N-1})^{N-1}} \\ &= (c_{par} + \Delta) \left( \frac{1-\lambda}{1+r-R\lambda/100+\delta_{N-1}(1-\lambda)} \right)^N \\ &- (c_{par} + 100\delta_{N-1}) \left( \frac{1-\lambda}{1+r-R\lambda/100+\delta_{N-1}(1-\lambda)} \right)^N \\ &= (\Delta - 100\delta_{N-1}) \left( \frac{1-\lambda}{1+r-R\lambda/100+\delta_{N-1}(1-\lambda)} \right)^N \end{aligned} \quad (\text{A.18})$$

<sup>7</sup>In case of decreasing yield curves for  $\Delta < 0$  the proof is similar.

Let us now observe that  $\Delta > 100\delta_{N-1}$ , as we have, for any bond with maturity  $N - 1$ , coupon equal to  $c_{par} + \Delta$ , and related yield  $\frac{c_{par}}{100} + \delta_{N-1}$ , the following equation:

$$100 + \Delta \sum_{n=1}^{N-1} \frac{(1-\lambda)^n}{(1+r)^n} = \sum_{n=1}^{N-1} \frac{c_{par} + \Delta}{(1 + c_{par}/100 + \delta_{N-1})^n} + \frac{100}{(1 + c_{par}/100 + \delta_{N-1})^{N-1}}. \quad (\text{A.19})$$

If we had  $\delta_{N-1} = \frac{\Delta}{100}$  this equation wouldn't be satisfied, being  $\Delta \sum_{n=1}^{N-1} \frac{(1-\lambda)^n}{(1+r)^n} > 0$  and having, for each bond with maturity  $N - 1$  and coupon  $c_{par+\Delta}$ , the equation

$$100 = \sum_{n=1}^{N-1} \frac{c_{par} + \Delta}{(1 + \frac{c_{par+\Delta}}{100})^n} + \frac{100}{(1 + \frac{c_{par+\Delta}}{100})^{N-1}}. \quad (\text{A.20})$$

So we need  $\delta_{N-1} < \frac{\Delta}{100}$  in order to satisfy equation (A.19).

Let us also remember that  $\delta_{N-1} < \delta_{sup} = R\lambda/(100(1-\lambda))$ .

It follows that the last term in equation (A.18) is greater than the first term. This is an absurd and it thus follows that we need  $\delta > 0$  to satisfy the equation (A.17).

We can complete the proof showing that coupons higher (lower) than the par yield imply, with the other factors being the same, increasing (respectively decreasing) bond prices curves. In fact, as for  $c > c_{par}$  ( $c < c_{par}$ ) the second term in equation (A.7) is greater (smaller) than zero, we get

$$100 < P_1 < \dots < P_n \dots < P_N, \quad \text{se } c > c_{par} \quad (\text{A.21})$$

$$100 > P_1 > \dots > P_n \dots > P_N, \quad \text{se } c < c_{par}. \quad (\text{A.22})$$

We can thus conclude that when the bond coupons are higher than  $c_{par}$  we get an increasing price structure, and viceversa.  $\square$

# Bibliography

- [1] Schonbucher, P. J. (2007), *Credit derivatives pricing models*. Wiley Finance Series.
- [2] Cherubinu, U., Luciano, E., Vecchiato, W. (2004), *Copula Methods in Finance*. Wiley Finance Series
- [3] Mai, J.F. and Scherer, M. (2008), "A tractable multivariate default model based on a stochastic time-change", *International Journal of Theoretical and Applied Finance*, Vol. 12, No. 2 (2009), 227-249.
- [4] Mai, J.F. and Scherer, M. (2009), "Lévy-frailty copulas", *Journal of Multivariate Analysis*, 100 (2009), 1567-1585.
- [5] Hull, J. and White, A. (2006), "Valuing Credit Derivatives Using an Implied Copula Approach", *Journal of Derivatives*, Fall 2006.
- [6] Hull, J. and White, A. (2009), "An improved implied copula model and its application to the valuation of bespoke CDO tranches", *Working Paper*, June 2009.
- [7] Segoviano, M. A. and Goodhart, C. (2010), "Distress Dependence and Financial Stability", *Central Bank of Chile Working Papers*, April 2010.
- [8] Segoviano, M. A. and Goodhart, C. (2009), "Banking Stability Measures", *IMF Working Papers*, January 2009.
- [9] Nelsen, R. B. (1999), *An Introduction to Copulas*. Springer, New York.
- [10] Huang, X. (1992), "Statistics of Extreme Values", *Tinbergen Institute Research Series, Ph.D. Thesis No. 22*. Erasmus University.
- [11] Bank of England, "Financial Stability Report" (December 2009, issue No. 26).

- [12] Fabozzi, Frank J. (2003), *Fixed Income Portfolio Management*. John Wiley & Sons, Inc.
- [13] Fabozzi, Frank J. (2007), *Fixed Income Analysis*. CFA Institute Investment Series.
- [14] Fitch Ratings Financial Institutions (Special Report, October 12, 2010), "CDS Spreads and Default Risk - Interpreting the Signals".
- [15] Fitch Ratings Financial Institutions (Special Report, November 8, 2010), "CDS Spreads and Default Risk - U.S. Broker-Dealers".
- [16] Hull, John C. (seventh edition), *Options, Futures and other Derivatives*. PEARSON Prentice Hall.