COLLANA DEL
DIPARTIMENTO DI ECONOMIA

EVALUATION OF CREDIT RISK UNDER CORRELATED DEFAULTS: THE CROSS-ENTROPY SIMULATION APPROACH

Loretta Mastroeni - Giuseppe D'Acquisto - Maurizio Naldi

ISSN 2279-6916 Working papers
(Dipartimento di Economia Università degli studi Roma Tre) (online)

Working Paper n° 193, 2014

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. l del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

REDAZIONE:
Dipartimento di Economia
Università degli Studi Roma Tre
Via Silvio D'Amico, 77 - 00145 Roma
Tel. 0039-06-57335655 fax 0039-06-57335771
E-mail: dip_eco@uniroma3.it
http://dipeco.uniroma3.it

DIPARTIMENTO DI ECONOMIA

EVALUATION OF CREDIT RISK UNDER CORRELATED DEFAULTS: THE CROSS-ENTROPY SIMULATION APPROACH

Loretta Mastroeni - Giuseppe D'Acquisto - Maurizio Naldi

Evaluation of Credit Risk under Correlated Defaults: the Cross-Entropy Simulation Approach

Loretta Mastroeni[a], Giuseppe D'Acquisto[b], Maurizio Naldi[b]

[a] *Università di Roma Tre, Dipartimento di Economia, Via Silvio D'Amico 77, 00145 Rome, Italy*

[b] *Università di Roma "Tor Vergata", Dipartimento di Ingegneria Civile e Ingegneria Informatica, Via del Politecnico 1, Rome, Italy*

Abstract

Credit risk, associated to borrowers defaulting on their debts, is an ever growing source of concern for lenders. The presence of correlation among defaults may be described by the t-copula model. However, the typically large number of variables involved calls for a simulation approach. A simulation method, based on the use of the Cross-Entropy (CE) technique, is here proposed as an alternative to non-adaptive Importance Sampling (IS) techniques so far presented in the literature, the main advantage of CE being that it allows to deal easily with a wider range of probability models than ad hoc IS. The method is validated through a comparison of its results with the crude MonteCarlo and the Exponential Twist approaches. The proposed Cross-Entropy technique is shown to provide accurate results even when the sample size is several orders of magnitude smaller than the inverse of the probability to be estimated.

Keywords: Credit risk, Cross-Entropy, Copula models
JEL codes: C150, G320

1. Introduction

The evaluation of credit risk has acquired a growing importance in financial activities. Lending money exposes the lender to the risk of losing it if the borrower does not comply. The possibility of the borrower defaulting on its debt is always present, but the number of defaults grows dramatically during periods of economic crisis. For a lender possessing a portfolio of credit securities, diversification may reduce the overall risk, but not eliminate it altogether. In addition, managing a large portfolio implies by itself a large exposure (and potential loss), since the number of obligors is large. For the well being of the financial company acting as a lender it is of paramount importance that the credit risk is evaluated and controlled.

Though defaults hit the single borrower, they tend to be correlated, since many of them are caused by common factors exogenous to the defaulting company. Examples of common risk factors are the economic cycle, or a recession hitting a specific country or industry sector, as surveyed in (Lucas, 1995). Of particular interest to lenders is the possibility of many obligors defaulting concurrently, an event that may be spurred by the common risk factors and goes under the name of extremal dependence. The cases of major interest are those where a very large loss (i.e., above a threshold hereafter referred to as the critical loss) is experienced by the portfolio owner. This may be the maximum tolerable loss or a proxy for the maximum loss. An example is the Value-at-Risk, which is the maximum loss that may happen with a given (yet significant) probability and is often used as a measure of risk (Alexander, 2009), having been applied in a number of contexts (Mastroeni and Naldi, 2011b) (Mastroeni and Naldi, 2011c). In this case losses may be large, and the lender needs to evaluate the probability associated to such loss events in order to take the appropriate hedging measures. Empirical financial data support extremal dependence as a significant defaulting mechanism, as shown in (Mashal and Zeevi, 2002). In other cases we may be interested in examining the critical states of the portfolio (i.e., the combination of states of the obligors that lead to an overall loss exceeding the critical one). and their relevance, rather than evaluate the probability of the set of critical states. In addition, we may want to identify all the critical states, rather than just the sample obtained through simulation (Naldi et al., 2013). For example, we may want to identify the obligors most present in such extreme cases. Such obligors may act as signals that the portfolio is at risk of generating losses over the critical threshold. Monitoring those critical obligors would allow to build an early warning system for the loss of the whole portfolio.

The dependence of the portfolio loss on a large number of individual and common risk factors has been the subject of modelling research in the latest years. A survey of credit risk models is reported in (Lando, 2009). An established framework is the latent variable approach, where the probability of default for an obligor is evaluated as the probability that an associated continuous latent variable exceeds a threshold. The normal copula model has been the first one based on this approach and assumed that the latent variable followed a multivariate normal distribution as described in (Crouhy et al., 2000). However, the alternative t-copula model has later been proposed in (Mashal and Zeevi, 2002), since it better describes the extremal dependence phenomenon. In that model a multivariate t-Student distribution is used instead of the normal one.

Unfortunately, there is no analytical solution for the t-copula model. Given the large number of variables involved, the task of estimating the probability of large losses is tackled by simulation. Since we are typically interested in the events leading to large losses, whose probability is hopefully quite small, the use of crude MonteCarlo (MC) simulation is not possible: the required size of the simulation sample would be too large due to the need to reduce the variance of the estimate down to acceptable values. Instead, Importance Sampling (IS) has been applied. In IS the probability density of the variables of interest is artificially biased towards larger values so to increase the probability of threshold exceedance; the effect of the bias is then eliminated through the use of the likelihood ratio as a weighting factor for the occurrences of the rare event of interest (see (Rubinstein, 1981) for an introduction to Importance Sampling). In (Bassamboo et al., 2008) two variants of IS have been proposed and compared for the t-copula case, namely the exponential twisting and the hazard rate twisting. After an extensive set of simulation tests the first has resulted to perform better.

However, both the approaches presented in (Bassamboo et al., 2008) require the use of a biasing procedure applicable just strictly under the working hypotheses for the probability distributions of the variables included in the model. In addition, the best performing algorithm requires to resort to acceptance-rejection procedures for the generation of pseudo random numbers (a quite slower generation technique with respect, e.g., to inversion algorithms, as found in (Rubinstein, 1981)) and to compute offline the likelihood ratio values needed in the weighting step of IS. The computational effort to accomplish the simulation task may therefore be very large.

We aim at reducing the computational effort to compute the probability of large losses in the presence of a t-copula through the use of an alternative simulation technique. In this paper we propose the use of Cross-Entropy (CE). Cross-Entropy, introduced in (Rubinstein and Kroese, 2004), may be classified as an adaptive Importance Sampling method, in which biasing is accomplished through an iterative procedure: at each step it relies on the simulation results obtained in the previous step, so to progressively approach the optimal bias. The method has been applied successfully in a number of fields (see the large reference list in (Rubinstein and Kroese, 2004) for papers published before 2002 or, e.g., the more recent in (D'Acquisto and Naldi, 2005)). Its characteristics are here exploited to remove the aforementioned shortcomings associated to the Exponential Twist IS approaches. Namely, we apply the CE approach to the t-copula model (following that same approach used for the normal copula in (Naldi and D'Acquisto, 2008) and (Mastroeni and Naldi, 2011a)). We propose a smoothed version of the classical CE method, where smoothing is applied both in the selection of the biasing variables and in the updating of the biasing parameters. We validate our smoothed CE approach by comparing the results with those obtained with the crude MonteCarlo technique and with non-adaptive Importance Sampling (namely, the Exponential Twist of (Bassamboo et al., 2008)). Finally, we examine the impact of several factors appearing in the model on the overall probability that the losses exceed a given threshold. We find that the outcome of CE simulation is quite close to those obtained by alternative approaches and that a sample size of 10^5 is satisfactory in most cases to obtain an accurate estimate of the probability of large losses. A summarized version of some results has been presented in (D'Acquisto et al., 2012).

The paper is organized as follows. After a brief description of the t-copula model, provided in Section 2, a full description of the CE simulation algorithm is reported in Sections 3 through 5. In Section 6 we report the simulation results and compare them with the alternative approaches.

2. Defaults in financial markets

The default of the obligor is the riskiest event a lender may face, since it leads to the loss of the money lent. For a lender possessing a portfolio of securities, defaults may occur concurrently for a number of its obligors (because of the inter-relationships existing on financial markets), increasing the extent of the loss. In order to evaluate the overall risk, we need a model to describe how defaults take place and their inter-dependencies. A model widely employed in the literature for this purpose is the t-copula model. In this section, we describe the t-copula model for the credit risk of a portfolio of securities under extremal dependence. Throughout the rest of this paper, we will adopt that model.

We consider a lender owning a portfolio of loans, bonds, and financial instruments, who has lent money to n obligors. We indicate by $\mathcal{I} = \{1, 2, \ldots, n\}$ the set of obligors. Each obligor is bound to give back the money under the issuing conditions but may occasionally default, i.e. fail to comply. In that case the lender loses its money. At any time, the set $\mathcal{I}$ of obligors is then partitioned in the two subsets $\mathcal{I}_\mathrm{d}$ and $\mathcal{I}_\mathrm{c} = \mathcal{I} \setminus \mathcal{I}_\mathrm{d}$, comprising respectively the defaulting obligors and the compliant ones (so that $\max|\mathcal{I}_\mathrm{d}| = \max|\mathcal{I}_\mathrm{c}| = n$). We describe first the behaviour of the individual obligor and then the joint effects of the behaviour of all obligors on the portfolio.

We indicate by p_i the probability that the i-th obligor defaults and associate that probability to the binary variable Y_i, which takes the value 1 if the i-th obligor defaults and 0 otherwise, so that $p_i = P[Y_i = 1]$. This variable is a direct indicator of the default event. In turn, we associate each default indicator to a latent (continuous) variable X_i by the relationship

$$Y_i = \begin{cases} 1 & \text{if } X_i \geq x_i \\ 0 & \text{if } X_i < x_i \end{cases}, \tag{1}$$

i.e. $Y_i = \mathbb{I}\{X_i \geq x_i\}$ where $\mathbb{I}\{\cdot\}$ is the indicator function and x_i is a suitable threshold. The intrinsically discrete phenomenon of defaults (represented by the binary variable Y_i) is now described by a continuous variable (X_i).

In a portfolio of n securities, the loss actually incurred by the lender is the sum of the amounts lost on each security. If the obligor i has borrowed the amount a_i, the overall loss for the lender is

$$l = \sum_{i=1}^{n} a_i Y_i = \sum_{i=1}^{n} a_i \mathbb{I}\{X_i \geq x_i\} = \sum_{i \in \mathcal{I}_\mathrm{d}} a_i. \tag{2}$$

Since defaults occur randomly, the composition of the set of defaulting obligors is itself random. For any composition of $\mathcal{I}_\mathrm{d}$, the loss is given by the random variable L, for which the following probability holds

$$\mathbb{P}[L = l] = \mathbb{P}[Y_1 = \mathbb{I}\{1 \in \mathcal{I}_\mathrm{d}\}, \ldots, Y_i = \mathbb{I}\{i \in \mathcal{I}_\mathrm{d}\}, \ldots, Y_n = \mathbb{I}\{n \in \mathcal{I}_\mathrm{d}\}]. \tag{3}$$

If defaults occur independently of each other, the probability of joint defaults is simply the product of the individual state probabilities

$$\mathbb{P}[L = l] = \prod_{i=1}^{n} \mathbb{P}[Y_i = \mathbb{I}\{i \in \mathcal{I}_\mathrm{d}\}] = \prod_{i \in \mathcal{I}_\mathrm{d}} p_i \prod_{i \in \mathcal{I}_\mathrm{c}} (1 - p_i). \tag{4}$$

However, it has since long been observed that the default occurrences among different obligors are not independent (see (Frey and McNeil, 2003) and (Dietsch and Petey, 2004)). Some obligors may operate in the same industrial or service sector, or in the same geographical area, and be subject to the same negative contingencies. Even if that's not the case, they all share the same overall economic cycle. Each latent variable must therefore incorporate such common dependence, accounting for a number of risk factors. In this paper, we adopt the t-copula model. If we label by η_i the individual risk factor pertaining specifically to the obligor i, we must include in each latent variable also a term accounting for the common dependence, which we indicate by Z (the common risk factor). In addition to the individual risk factor and the common risk factor, the t-copula model considers a common shock factor W. The common risk factor appears as an additive term, while the common shock factor models larger common defaulting causes. The relation between the latent variable and the three risk factors is

$$X_i = \frac{\rho Z + \sqrt{1-\rho^2}\eta_i}{Q}, \tag{5}$$

where ρ is the weight of the common risk factor. When the common shock factor assumes small values, all the latent variables are likely to be large, meaning that a large number of simultaneous defaults occur.

If we make the following assumptions on the probability distributions of the variables at hand (in addition to considering them independent of each other)

$$Z \sim N(0,1), \tag{6}$$

$$\eta_i \sim N(0,1), \tag{7}$$

$$Q = \sqrt{\frac{R}{\beta}}, \tag{8}$$

$$R \sim \chi^2(\beta), \tag{9}$$

the latent variable X_i follows a Student t-distribution, hence the t-copula name. The probability density function of the shock factor Q is

$$f_Q(x) = 2\left(\frac{\beta}{2}\right)^{\beta/2} \frac{x^{\beta-1}}{\Gamma(\beta/2)} e^{-\beta x^2/2}. \tag{10}$$

The t-copula model has been first proposed by Mashal and Zeevi ((Mashal and Zeevi, 2002)), as an advance over the previously proposed multivariate normal model (incorporated in the Creditmetrics model (Gupta et al., 1997)). The multivariate normal model considers the presence of just the additive common risk factor (in fact it can be derived by the model described by eqn. 5 by simply omitting the common shock factor W), underestimating the chance of many simultaneous defaults in the portfolio. As noted in (Crouhy et al., 2000), while it was legitimate to assume normality of the portfolio changes due to market risk, that is no longer the case for credit returns which are by nature highly skewed and fat-tailed. In the case of credit risk there is limited upside to be expected from any improvement in credit quality, while there is a substantial downside risk due to downgrading and default. As claimed in (Mashal and Zeevi, 2002), the t-copula has two major advantages over the multivariate normal model. While in the normal model dependence is captured only via correlation, the t-dependence structure retains the use of correlation while introducing a single additional parameter (β, representing the degrees-of-freedom), that allows for extreme co-movements. In particular, increasing β decreases the tendency of underlying to exhibit extreme co-movements. The second advantage is that the t-copula supports extreme co-movements regardless of the marginal behavior of the individual assets.

3. A Cross-Entropy formulation of the rare event estimation problem

Defaults cause losses. However large the may be, their probability is of paramount importance to assess the relevance of the associated risk. In this section, we describe a simulation method to estimate that probability, based on the Cross-Entropy approach. We draw largely on the original formulation of the approach due to (Rubinstein and Kroese, 2004).

Our problem is to evaluate the probability that the overall loss exceeds a given threshold

$$\gamma = \mathbb{P}(L > l). \tag{11}$$

Excepting the case of independence, the evaluation is not amenable to an analytic solution, given the large number of random variables involved. We must resort to simulation. Since the events leading to large losses are those of greatest interest, the threshold l is typically quite large. In addition, the events leading to large losses are typically rare if the portfolio is not built under an extremely risky profile. We are then left with the problem of defining a simulation method for rare events.

Since the loss L actually depends on a number of random variables, we can signal that dependence by aggregating them all in the random vector $\mathbf{S} = \{Z, \eta_1, \ldots, \eta_n, Q\}$. Since all the random variables introduced in Section 2 follow parametric models (defined either by their first two moments or by their scale and shape parameters), we can also build the vector $\mathbf{u}$ comprising all those parameters. The probability density function (pdf) for the random vector $\mathbf{S}$ is then $f(\mathbf{s}; \mathbf{u})$.

An alternative expression for the probability of losses being larger than the given threshold is therefore

$$\gamma = \mathbb{P}_{\mathbf{u}}[L > l] = \mathbb{E}_{\mathbf{u}}\left[\mathbb{I}\{L(\mathbf{S}) > l\}\right], \tag{12}$$

where the symbol $\mathbb{E}$ stands for the expected value, and the subscript $\mathbf{u}$ indicates that the operations involved (computation of the probability or extraction of the expected value) are performed adopting the set $\mathbf{u}$ of parameter values for the pdf of $\mathbf{S}$.

After drawing a random sample $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_N$ of size N of the random vector $\mathbf{S}$, we could get a crude MonteCarlo estimation of the probability of interest by the sample average

$$\hat{\gamma}_{MC} = \frac{1}{N} \sum_i \mathbb{I}\{L\left(\mathbf{S}_i\right) > l\}. \tag{13}$$

However, the problem with MonteCarlo estimation is the large variance associated with the low value of the probability to be estimated. Actually, the normalized standard error of the MonteCarlo estimator has the well known expression (with the final approximation valid for small values of γ)

$$\frac{\sigma_{(\hat{\gamma}_{MC})}}{\mathbb{E}\left(\hat{\gamma}_{MC}\right)} = \sqrt{\frac{1-\gamma}{N\gamma}} \simeq \sqrt{\frac{1}{N\gamma}}. \tag{14}$$

For any required accuracy, the sample size should grow as the square of the inverse of the probability to be estimated: achieving a normalized standard error of 10% for a probability of 10^{-6} would require a sample size of one hundred million, with the resulting large simulation time.

This problem can be overcome by resorting to the Importance Sampling (IS) simulation method where the probability associated to the events of interest (those such that $\mathbb{P}[L > l]$) is artificially increased through the use of a biased pdf $g(\mathbf{s})$ (for a wider description of IS see (Srinivasan, 2002)). Since biasing introduces a distortion on the probability as observed after drawing the random sample of $\mathbf{S}$, the bias has then to be recovered by using the IS estimator

$$\hat{\gamma}_{IS} = \frac{1}{N} \sum_i \mathbb{I}\{L(\mathbf{S}_i) > l\} \frac{f(\mathbf{s};\mathbf{u})}{g(\mathbf{s})}. \tag{15}$$

The use of IS is expected to lead to a more efficient use of the random sample, i.e., to a lower error for the same sample size. The extent of the improvement depends on the proper choice of this biased pdf, which has to be such to reduce the variance of the associated estimator. An ideal zero-variance IS estimator would be attained when the biased pdf is

$$g^*(\mathbf{s}) = \frac{\mathbb{I}\{L(\mathbf{S}_i) > l\} f(\mathbf{s};\mathbf{u})}{\gamma} \tag{16}$$

which unfortunately depends on the same quantity to be estimated and is therefore not feasible.

However, this ideal estimator can be approached by looking for the best biasing pdf within the family $f(\mathbf{s};\mathbf{v})$, where $\mathbf{v}$ is the so-called tilting parameter vector, such that the distance between this newly defined pdf and the optimal one is minimized. A suitable measure of distance is the Kullback-Leibler distance, a.k.a. as Cross-Entropy, defined as the expected value of the logarithm of the ratio of the two pdfs (the biasing one and the optimal $g^*(\mathbf{s})$) computed under the probability measure provided by the pdf to be approached

$$\mathcal{D}(g,f) = \mathbb{E}_g\left[\ln \frac{g^*(\mathbf{s})}{f(\mathbf{s};\mathbf{v})}\right] = \int g^*(\mathbf{s}) \ln g^*(\mathbf{s}) ds - \int g^*(\mathbf{s}) \ln f(\mathbf{s};\mathbf{v}) d\mathbf{s}. \tag{17}$$

Since just the latter term depends on the tilting parameter to be optimized, minimizing the Kullback-Leibler distance is equivalent to choose $\mathbf{v}$ to solve the following maximization problem, where we have used Equation (16) for the ideal optimal pdf:

$$\max_{\mathbf{v}} \int \frac{\mathbb{I}\{L(\mathbf{S}_i) > l\} f(\mathbf{s};\mathbf{u})}{\gamma} \ln f(\mathbf{s};\mathbf{v}) d\mathbf{s}. \tag{18}$$

which in turn is equivalent to the program

$$\max_{\mathbf{v}} \quad \mathbb{E}_{\mathbf{u}}\left[\mathbb{I}\{L(\mathbf{S}_i) > l\} \ln f(\mathbf{s};\mathbf{v})\right]. \tag{19}$$

By the repeated application of the Importance Sampling approach, using again the pdf family $f(\mathbf{s};\mathbf{w})$ with a reference tilting parameter $\mathbf{w}$, the maximization program can finally be written as

$$\max_{\mathbf{v}} \quad \mathbb{E}_{\mathbf{w}}\left[\mathbb{I}\{L\left(\mathbf{S}_i\right) > l\}\frac{f(\mathbf{s};\mathbf{u})}{f(\mathbf{s};\mathbf{w})}\ln f(\mathbf{s};\mathbf{v})\right], \tag{20}$$

whose solution is

$$\mathbf{v}^* = \underset{\mathbf{v}}{\operatorname{argmax}} \quad \mathbb{E}_{\mathbf{w}}\left[\mathbb{I}\{L\left(\mathbf{S}_i\right) > l\}\frac{f(\mathbf{s};\mathbf{u})}{f(\mathbf{s};\mathbf{w})}\ln f(\mathbf{s};\mathbf{v})\right]. \tag{21}$$

The optimal tilting parameter vector can be estimated by solving the corresponding stochastic program, which uses a simulated sample $\mathbf{S}_1,\ldots,\mathbf{S}_N$ extracted from $f(\cdot;\mathbf{w})$, and replacing the expected value with the sample average

$$\max_{\mathbf{v}}\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\{L\left(\mathbf{S}_i\right) > l\}\frac{f(\mathbf{S}_i;\mathbf{u})}{f(\mathbf{S}_i;\mathbf{w})}\ln f(\mathbf{S}_i;\mathbf{v}). \tag{22}$$

The j-th component of the tilting vector can be obtained by zeroing the derivative with respect to the optimizing parameter, obtaining the equation

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\{L\left(\mathbf{S}_i\right) > l\}\frac{f(\mathbf{S}_i;\mathbf{u})}{f(\mathbf{S}_i;\mathbf{w})}\frac{\partial}{\partial v_j}\ln f(\mathbf{S}_i;\mathbf{v}) = 0, \tag{23}$$

which can be solved analytically if the distributions of the random variables of interest belong to a natural exponential family.

The solution provides us with the best approximation of the ideal biasing density (16), which can then be employed in the standard IS simulation method with the estimator provided by Equation (15).

4. Selection of biasing variables

In Section 3, we have provided the general formulation of our estimation problem through the Cross-Entropy approach. However, that formulation leaves the problem open to determine which variables to bias. For a portfolio comprising many securities, that number can be very large. In this section, we highlight the problems related to the presence of a large number of biasing parameters and propose a solution based on the selection of restricted basket of parameters.

In the description of the problem given in Section 3, the potential biasing parameters are all the parameters involved in the probability models of the variables at hand described in Section 2, i.e., the expected values and the variances of all the individual risk factors, the expected value and the variances of the common risk factor, and the parameters of the shock factor. As to the latter component, though Equations (8) and (9) contemplate just a single parameter (the number β of degrees of freedom of the Chi-square distribution), we could define the variable R embedded in the shock factor as following a Gamma distribution, to have a wider choice of biasing parameters, as done in (Bassamboo et al., 2008). The variable R would be modelled through a Gamma distribution with shape parameter $\beta/2$ and scale parameter $2/\zeta$. In fact, when opting for the Gamma model, we can bias both the scale and the shape parameter rather than just β. When $\zeta = 1$, we get back the chi-square distributed R. However, we must mention that some difficulties may arise when generating the pseudo-random number we need to simulate the Gamma distributed variables. If we limit ourselves to the cases where the shape parameter is an integer number, the Gamma distribution reduces to the Erlang distribution. In that case, the Erlang-distributed pseudo-random numbers can be obtained as the sum of a number (equal to the value of the shape parameter, i.e., half the number of degrees of freedom of the chi-square distributed variable R) of pseudo-random numbers following an exponential distribution with expected value equal to half the scale parameter of the Gamma distribution (see, for example, chapter 3.6.2 of (Rubinstein, 1981)). However, if that is not the case (the shape parameter is not an integer number), the distribution function of the pseudo-random numbers to be generated does not possess a closed form. The generation of those pseudo-random numbers requires then an acceptance-rejection generation method, quite slower with respect to those based on the inversion of the cumulative distribution function. Enclosing the shock factor among the biasing variables may therefore lead to a significant slowing of the simulation procedure.

We can aggregate all the parameters defining the model in the vector $\mathbf{u} = \left\{\mu_Z, \sigma_Z^2, \mu_{\eta_1}, \sigma_{\eta_1}^2 \ldots, \mu_{\eta_n}, \sigma_{\eta_n}^2, \beta, \zeta\right\}$. With this enhanced formulation, the number of potential biasing parameters rises to $2n+4$. Notwithstanding the linear relationship, this number can be in the order of hundreds if the portfolio is well diversified.

The negative effects of such a large number of biasing parameters are at least two.

The most obvious one is that the simulation time grows proportionally to that number: the addition of any biasing parameters involves an additional amount of computer code to determine its value. Hence, if we want to keep the simulation time as low as possible (considering also the problem of estimating low probability values that led us to opt for an accelerated simulation technique), the number of biasing parameters has to be as low as possible.

A second problem, which is known to affect Importance Sampling and other MonteCarlo-based simulation technique such as particle filtering, is the so-called degeneracy problem of the likelihood ratio, a.k.a. the curse of dimensionality mentioned in (Arulampalam et al., 2002) and (Chan and Kroese, 2011). When the overall number of parameters is very large, the likelihood ratio is obtained by the product of so many terms that it gives rise to largely variable estimates, and the variables that would be i.i.d. may end up with largely different biases. The difficulty of dealing with this problem has led to the practical suggestion of avoiding IS simulation altogether for high dimensional problems. An alternative solution is to limit the number of variable subject to biasing.

In (Bassamboo et al., 2008), just a subset of $n+2$ among the original basket of $2n+4$ is biased, composed of the expected values of the common risk factor and of the individual risk factor, plus the scale parameter of the Gamma distribution governing the shock factor.

Recently Rubinstein has provided an approach to overcome this limitation by the suitable selection of the parameters to be biased in (Rubinstein and Glynn, 2009). In that screening method the parameters are divided into two classes: the bottleneck ones (to be biased) and the non-bottleneck ones (which, not being subject to biasing, do not contribute to the likelihood ratio). The selection is accomplished by performing a MonteCarlo simulation of the system of interest using a threshold l_0 suitably chosen so that the event of interest is not rare. In the process of simulation the stochastic program leading to the optimal biasing parameters is replaced by the following one, where the random sample is extracted from the distribution $f(\mathbf{s};\mathbf{u})$ (i.e., with no bias) and the likelihood ratio is not present

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}_{\{L(\mathbf{S}_i)>l_0\}}\frac{\partial}{\partial v_j}\ln f(\mathbf{S}_i;\mathbf{v})=0. \tag{24}$$

The solution of equation (24) provides us with a CE-derived tilting vector $\hat{\mathbf{v}}$, among whose elements we want to select the bottleneck ones. For this purpose the relative perturbation of each element is computed

$$\delta_{ki} = \frac{\hat{\mathbf{v}}_{ki} - \mathbf{u}_i}{\mathbf{u}_i}, \tag{25}$$

where the index $k = 1, \dots, d$ shows that this procedure is iterated d times. At each iteration just the elements are retained whose relative perturbation lies below some threshold δ. As can be seen, this approach aims at retaining as biasing parameters those moving faster, thereby promoting reactivity as the main factor in choosing the biasing parameters.

In a large number of simulation runs of the main CE algorithm we have however noted that the dynamics of the tilting parameters during adaptive biasing is largely of an oscillatory nature: the convergence of each biasing parameter towards its final value is not monotonic but rather oscillates around a roughly linear trend. A sample picture is reported in Figure 1, where the bias on the expected value of the common risk factor (i.e. μ_Z) converges to its final value as the number of iterations of the CE simulation procedure grows (the sample size was 10^4). If the oscillation associated to a parameter is very large, that parameter, though falling into the bottleneck class according to the criterion described above, may hamper the convergence of the simulation program.

Here we therefore propose, and apply to the case at hand, a different criterion, aiming to promote steadiness rather than reactivity. In this criterion the parameters subject to biasing are those exhibiting the least oscillations around their trend. We therefore proceed by performing the same modified simulation program devised in the screening method by Rubinstein, applying a regression analysis on each parameter, and retaining as biasing parameters those exhibiting the least residuals.

Starting with the choice taken in (Bassamboo et al., 2008), where just the expected values of the risk factors (individual and common) and the scale parameter of the Gamma-distributed shock parameter are biased, and applying the criterion just described, we further restrict the selection by reneging to bias the shock parameter. The number of biasing parameters is so reduced from $2n + 4$ to $n + 1$.

After selecting the parameters to be biased the straightforward step is to use the biasing equations derived in Section 3 for each parameter. We will deviate from this procedure in two ways, aiming at smoothing the convergence process of the biasing parameters towards their final values.

First, some of the variables of interest in the model may be i.i.d. and playing the same role. That's the case for the individual risk factors in the model defined in Section 3. Though the biasing equations would provide different biasing paths for the parameters associated to these variables, we considered it contradictory. In our simulation procedure, all the individual risk factors are biased in the same way. Given the biasing values obtained through the individual applications of the equations of Section 3, the common biasing is obtained as their arithmetic average.

Secondly, notwithstanding the selection operated by biasing just the steadiest parameters, their dynamics during the simulation runs can still be somewhat erratic. In order to avoid the resulting delay in convergence we apply a smoothing procedure (based on the classical simple exponential smoothing algorithm) to the parameters as computed through the biasing equations. This procedure was already suggested in the reference book (Rubinstein and Kroese, 2004). If we indicate by $\hat{\mathbf{v}}$ the set of parameter values as resulting from the straightforward application of the biasing updating equations, by $\mathbf{v}_{t-1}$ the current value of those parameters and by $\mathbf{v}_t$ the tilting vector of parameter values after the smoothed updating, the relationship to apply is

$$\mathbf{v}_t = \alpha \mathbf{v}_{t-1} + (1 - \alpha)\hat{\mathbf{v}}, \tag{26}$$

where α is a smoothing parameter (the larger it is the heavier the smoothing), which we have set equal to 0.8.

5. The simulation procedure

In Section 3, we have provided the general formulation of the estimation of the probability of large losses under the Cross-Entropy approach. Since this would require the simulation of a large number of random variables, with the resulting curse of dimensionality in the computation of the likelihood ratio, we have devised a procedure to safely restrict the number of variables to be biased in Section 4. In this section, we can now describe the complete simulation procedure.

Though we are not going to bias the shock factor, in this section we will provide the more general expressions, which include the scale parameter of the shock factor among the variables to bias. The vector of the parameters prior to bias is $\mathbf{u} = \left\{\mu_Z = 0, \sigma_Z^2, \mu_{\eta_1} = 0, \sigma_{\eta_1}^2, \ldots, \mu_{\eta_n} = 0, \sigma_{\eta_n}^2, \beta, \zeta = 1\right\}$, so that the only parameter actually to set is β. When we bias, we get the

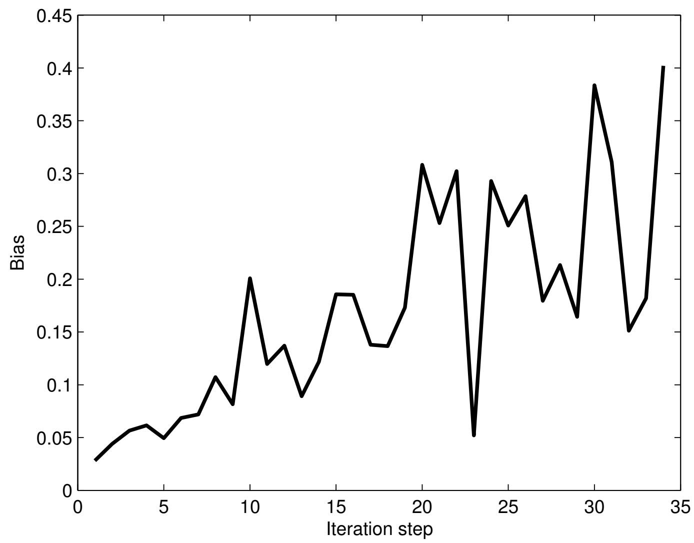

Figure 1: Sample bias dynamics

vector $\mathbf{w} = \{\mu_Z, \sigma_Z^2, \mu_{\eta_1}, \sigma_{\eta_1}^2, \ldots, \mu_{\eta_n}, \sigma_{\eta_n}^2, \beta, \zeta\}$, where the parameter β is still to set, but the parameters subject to bias are gathered in the subset $\{\mu_Z, \mu_{\eta_1}, \ldots, \mu_{\eta_n}, \zeta\}$, which reduces to $\{\mu_Z, \mu_{\eta_1}, \ldots, \mu_{\eta_n}\}$ if we decide not to bias the shock factor.

We can now evaluate the probability density functions that we need to compute the likelihood ratio. Since the common risk factor, the individual risk factors and the shock factor are all independent, their joint pdf is just the product of their individual pdf's. In the absence of bias, for the i-th sample we have

$$\begin{aligned} f(\mathbf{S}_i; \mathbf{u}) &= f(z_i, \eta_{1i}, \ldots, \eta_{ni}, r_i; \beta, \sigma_z, \sigma_{\eta_1}, \ldots, \sigma_{\eta_n}) \\ &= \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2\sigma_z^2} \frac{1}{2\Gamma(n/2)} \left(\frac{r_i}{2}\right)^{\frac{\beta}{2}-1} e^{-r_i/2} \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\eta_{ji}^2/2\sigma_{\eta_j}^2} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^{n+1} e^{-z_i^2/2\sigma_z^2} \frac{1}{2\Gamma(n/2)} \left(\frac{r_i}{2}\right)^{\frac{\beta}{2}-1} e^{-r_i/2} e^{-\sum_{j=1}^{n} \eta_{ji}^2/2\sigma_{\eta_j}^2} \end{aligned} \tag{27}$$

When we apply bias, the biased pdf is instead

$$f(\mathbf{S}_i;\mathbf{w}) = f(z_i,\eta_{1i},\ldots,\eta_{ni},r_i;\mu_z,\mu_{\eta_1},\ldots,\mu_{\eta_n};\beta,\sigma_z,\sigma_{\eta_1},\ldots,\sigma_{\eta_n},\zeta) = \\ = \left(\frac{1}{\sqrt{2\pi}}\right)^{n+1}\frac{\zeta}{2\Gamma(n/2)}\left(\frac{\zeta r_i}{2}\right)^{\frac{\beta}{2}-1} e^{-(z_i-\mu_z)^2/2\sigma_z^2-\zeta r_i/2-\sum_{j=1}^{n}(\eta_{ji}-\mu_{\eta_j})^2/2\sigma_{\eta_j}^2} \tag{28}$$

Finally, the target pdf, whose parameters we want to estimate, is

$$f(\mathbf{S}_i;\mathbf{v}) = f(z_i,\eta_{1i},\ldots,\eta_{ni},r_i;\mu_z^*,\mu_{\eta_1}^*,\ldots,\mu_{\eta_n}^*,\beta,\sigma_z,\sigma_{\eta_1},\ldots,\sigma_{\eta_n},\zeta^*) = \\ = \left(\frac{1}{\sqrt{2\pi}}\right)^{n+1}\frac{\zeta^*}{2\Gamma(n/2)}\left(\frac{\zeta^* r_i}{2}\right)^{\frac{\beta}{2}-1} e^{-(z_i-\mu_z^*)^2/2\sigma_z^2-\zeta^* r_i/2-\sum_{j=1}^{n}(\eta_{ji}-\mu_{\eta_j}^*)^2/2\sigma_{\eta_j}^2} \tag{29}$$

We can now recall the general expression of the likelihood ratio for the i-th sample

$$\mathcal{L}_i = \frac{f(\mathbf{S}_i;\mathbf{u})}{f(\mathbf{S}_i;\mathbf{w})} = \exp\left[\frac{(z_i-\mu_z)^2-z_i^2}{2\sigma_z^2}+\sum_{j=1}^{n}\frac{(\eta_{ji}-\mu_{\eta_j})^2-\eta_{ji}^2}{2\sigma_{\eta_j}^2}\right]\frac{e^{(\zeta-1)r_i/2}}{\zeta^{\beta/2}}, \tag{30}$$

which, if we don't bias the shock factor, reduces to

$$\mathcal{L}_i = \exp\left[\frac{(z_i-\mu_z)^2-z_i^2}{2\sigma_z^2}+\sum_{j=1}^{n}\frac{(\eta_{ji}-\mu_{\eta_j})^2-\eta_{ji}^2}{2\sigma_{\eta_j}^2}\right]. \tag{31}$$

The biasing equation (23) gets the following particular forms, depending on the parameter we are going to optimize:

$$\sum_{i=1}^{N}\mathbb{I}\{L_i > l\}\mathcal{L}_i\frac{\partial}{\partial\mu_z^*}\ln[f(z_i,\eta_{1i},\ldots,\eta_{ni},r_i;\mu_z^*,\mu_{\eta_1}^*,\mu_{\eta_2}^*,\ldots,\mu_{\eta_n}^*,\beta,\zeta^*)] = 0 \tag{32}$$

$$\sum_{i=1}^{N}\mathbb{I}\{L_i > l\}\mathcal{L}_i\frac{\partial}{\partial\mu_{\eta_j}^*}\ln[f(z_i,\eta_{1i},\ldots,\eta_{ni},r_i;\mu_z^*,\mu_{\eta_1}^*,\mu_{\eta_2}^*,\ldots,\mu_{\eta_n}^*,\beta,\zeta^*)] = 0 \\ j = 1,2,\ldots,n \tag{33}$$

$$\sum_{i=1}^{N} \mathbb{I}\{L_i > l\}\mathcal{L}_i \frac{\partial}{\partial \zeta^*} \ln[f(z_i, \eta_{1i}, \ldots, \eta_{ni}, r_i; \mu_z^*, \mu_{\eta_1}^*, \mu_{\eta_2}^*, \ldots, \mu_{\eta_n}^*, \beta, \zeta^*)] = 0 \tag{34}$$

The solution of Equations (32) through (34) provides us with the following expressions for the biasing parameters

$$\mu_z = \frac{\sum_{i=1}^{N} \mathbb{I}\{L_i > l\}\mathcal{L}_i z_i}{\sum_{i=1}^{N} \mathbb{I}[L_i > l]\mathcal{L}_i}, \tag{35}$$

$$\mu_{\eta_j} = \frac{\sum_{i=1}^{N} \mathbb{I}\{L_i > l\}\mathcal{L}_i \eta_j}{\sum_{i=1}^{N} \mathbb{I}[L > l]\mathcal{L}_i} \qquad j = 1, 2, \ldots, n, \tag{36}$$

$$\zeta^* = \beta \frac{\sum_{i=1}^{N} \mathbb{I}\{L_i > l\}\mathcal{L}_i}{\sum_{i=1}^{N} \mathbb{I}[L_i > l]\mathcal{L}_i r_i}. \tag{37}$$

We recall that in all the previous equations, the expression to use for the likelihood ratio is either that in Equation (30) or that in Equation (31), depending on whether we decide to bias the shock factor or not.

In order to apply correctly Equations (35) through (37), the following condition must hold

$$\sum_{i=1}^{N} \mathbb{I}\{L_i > l\} \neq 0. \tag{38}$$

In fact, if that condition does not hold, the three equations do not provide any useful (i.e. nonzero) term. It is therefore expedient to start the algorithm with a threshold $\hat{l} < l$, to be progressively increased as the bias level grows. Once the actual threshold l is reached the algorithm is run till a stopping criterion is met, which in our case is that the estimated loss probability is within a prescribed relative difference with respect to the one computed at the previous stage (e.g., within 5% or 10%), in order to warrant convergence. In the end, the resulting simulation programme can be synthetically described as Algorithm 1.

6. Simulation results

Throughout Sections 3, 4, and 5, we have described a simulation method based on the Cross-Entropy approach to the computation of losses under the t-copula model. In this section, we apply that method to a number of cases with the two-fold goal: 1) checking the results against those obtained

Algorithm 1 The Cross-Entropy simulation procedure

Require: The number of obligors n
Require: The number of simulation runs N
Require: The loss l
Ensure: The probability of losses larger than l.

Set $\mathbf{v}_0 = \mathbf{u}$, $l_0 = 1$, and the iteration counter $t = 0$
while $l_t < l$ **do**
    Generate a sample $\mathbf{S}_1, \ldots, \mathbf{S}_N$ from the probability density $f(\cdot; \mathbf{v}_{t-1})$
    Obtain the resulting values of the number of defaults $L(\mathbf{S}_1), \ldots, L(\mathbf{S}_N)$, sort them in increasing fashion, and extract the 90% percentile to assign it to the threshold l_t
    Use the same sample to update the tilting vector $\mathbf{v}_t$ by applying equations (35), (36), and (37), which provide the straightforward tilting vector $\hat{\mathbf{v}}$, and then (26)
    Set $t = t + 1$
    Update the threshold l_t
end while
repeat
    Generate a sample $\mathbf{S}_1, \ldots, \mathbf{S}_N$ from the probability density $f(\cdot; \mathbf{v}_{t-1})$
    Estimate the loss probability by eq. (15) with $\mathbf{v} = \mathbf{v}_t$
    Compute the difference Δ with respect to the previous estimate
    if $\Delta <$ Tolerance **then**
        Output the loss probability
        Set the convergence flag as reached
    else
        Use the same sample to update the tilting vector $\mathbf{v}_t$, according to equations (35), (36), and (37)
        Set $t = t + 1$
    end if
until Convergence

through a closed formula or alternative simulation approaches; 2) analysing the impact of several parameters involved in the model on the probability of large losses.

Parameter	Value
n	100
σ_z	1
$\{\sigma_{\eta_i}\}$	1
$\{x_i\}$	1
l	50
$\{a_i\}$	1

Table 1: Parameter values in the absence of the shock factor

6.1. Absence of the shock factor

As recalled, the reason to resort to a simulation approach is the absence of analytical solutions for the probability of large losses. That absence makes it impossible to compare the simulation results against the exact ones to validate the simulation approach. However, we can consider some special cases for which we do own the analytical solution and can accomplish the comparison with the exact result.

We consider the case where the shock factor is absent. The t-copula model reduces to a Gaussian copula, where the set of latent variables are correlated Gaussian variables. The degree of correlation is determined by the weight ρ. If we set $\rho = 0$, we reduce to a set of independent Gaussian variables, for which we can easily compute the probability of any loss through a closed formula. Summing up, we can simulate the Gaussian copula, compare the simulation result with the exact one when $\rho = 0$ and see how the growth of the weight impacts on the overall probability of large losses.

We consider the set of parameter values reported in Table 1.

When $\rho = 0$, the probability that a single latent variable crosses the threshold is (since $\eta_i \sim N(0; 1)$)

$$\mathbb{P}[X_i > x_i] = \mathbb{P}[\eta_i > x_i] = 1 - G(x_i) = 1 - G(1) \simeq 0.1586, \qquad (39)$$

with $G(\cdot)$ being the standard Gaussian cumulative distribution function. Under the assumption of independent defaults, which derives from $\rho = 0$, the number of defaults on the basket of 100 securities follows a binomial distribution. Since $a_i = 1$ the probability of losses larger than 50 equals the probability that there are more than 50 defaults

$$\mathbb{P}[L > 50] = \sum_{i=51}^{100} \binom{100}{i} 0.1586^i (1 - 0.1586)^{100-i} \simeq 1.776 \cdot 10^{-15} \qquad (40)$$

The simulation results are shown in Figure 2. We have used a sample size of 10^5. In Section 3, we have recalled Equation (14) that provides us with the relationship between sample size and accuracy in a crude MonteCarlo simulation. That size would allow to estimate probability values just as low as 10^{-3} with a standard error of 10%. Through the Cross-Entropy approach, we are able to obtain a fairly accurate estimate for much lower values. Even in the lowest range of probability values, corresponding to the case of independent defaults, the simulation provides values which are slightly lower than the exact result.

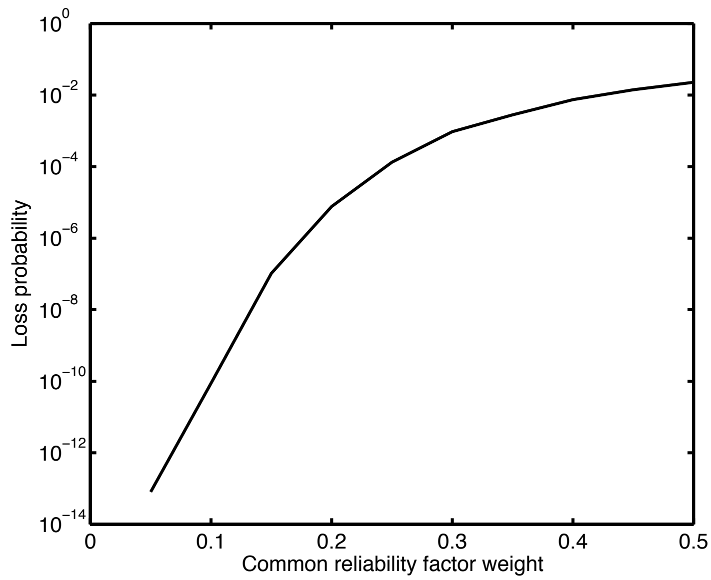

Figure 2: Impact of the common risk factor in the absence of the shock factor

The presence of correlation increases the probability of losses. When $\rho = 0.1$, the probability of large losses increases by 5 orders of magnitude with respect to the absence of correlation. On a singly-log scale, the growth

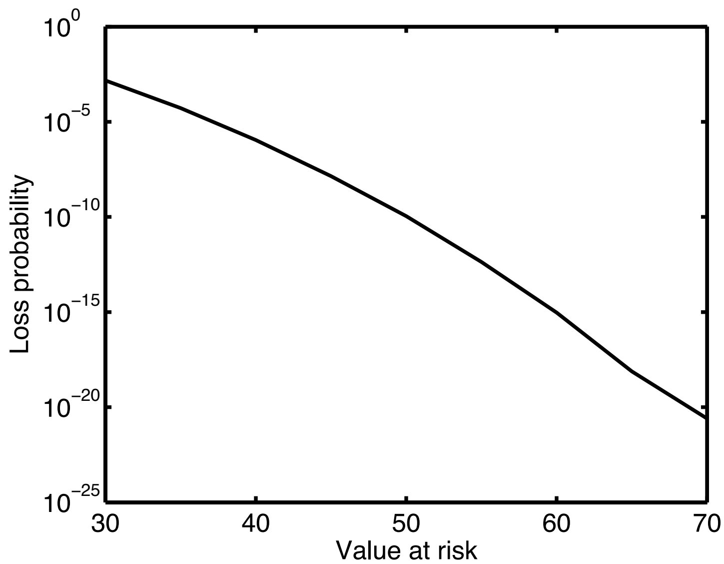

Figure 3: Probability of losses larger than the Value-at-Risk

of the loss probability starts with a linear trend but then bends when $\rho > 0.2$ roughly.

With the same parameter values of Table 1, we can draw the relationship between the Value-at-Risk (VaR) and the probability of losses larger than the VaR. The curve is shown in Figure 3 when $\rho = 0.1$.

6.2. Comparison with alternative simulation approaches

An alternative approach to evaluate the probability of large losses under a t-copula model through simulation has been proposed in (Bassamboo et al., 2008). They applied Importance Sampling, but through a biasing procedure applicable just strictly under the working hypotheses, while the Cross-Entropy approach has a much wider field of applicability, with the only constraint that the biasing variable follow a distribution belonging to the negative exponential family.

The presence of an alternative simulation approach allows us to compare the results obtained with the CE approach. We will consider therefore the same parameter values adopted in (Bassamboo et al., 2008), reported in Table 2.

In our experiment, we have considered a wider range for the individual risk factor: we have them span the range from 1 to 9, rather than the single

Parameter	Value
n	100
σ_z	1
$\{\sigma_{\eta_i}\}$	9
ρ	9
$\{x_i\}$	$0.5\sqrt{n} = 5$
l	25
$\{a_i\}$	1
β	4,8,12,16,20

Table 2: Parameter values adopted in (Bassamboo et al., 2008)

value considered by Juneja. We have set again the number of simulation runs equal to 10^5 and have biased all the risk factors (both the common and the individual ones), but not the shock factor.

The comparison with the results obtained in (Bassamboo et al., 2008) (under the Exponential Twist heading) is shown in Table 3. In the same Table we also report the asymptotic approximations obtained in (Bassamboo et al., 2008), valid when the portfolio is made of a large number of small obligors and we are interested in large losses. Excepting the case of 4 degrees of freedom, the CE approach resulted in values intermediate between those obtained through simulation in (Bassamboo et al., 2008) and the asymptotic values.

Degrees of freedom	Exponential Twist	Cross-Entropy	Asymptotic approx.
4	$4.35 \cdot 10^{-2}$	$3.95 \cdot 10^{-2}$	$5.18 \cdot 10^{-2}$
8	$5.58 \cdot 10^{-3}$	$5.80 \cdot 10^{-3}$	$9.00 \cdot 10^{-3}$
12	$8.96 \cdot 10^{-4}$	$1.10 \cdot 10^{-3}$	$2.15 \cdot 10^{-3}$
16	$1.64 \cdot 10^{-4}$	$2.32 \cdot 10^{-4}$	$6.46 \cdot 10^{-4}$
20	$3.41 \cdot 10^{-5}$	$6.49 \cdot 10^{-5}$	$2.33 \cdot 10^{-4}$

Table 3: Comparison with Exponential Twist results

The use of variances spanning the $[1, 9]$ range rather than the single value considered in (Bassamboo et al., 2008) allows us to evaluate the impact of the dispersion of the individual risk factors. In Figure 4, we see that the variance contributes to increase the probability of losses, but the effect is much more marked as the number β of degrees of freedom grows.

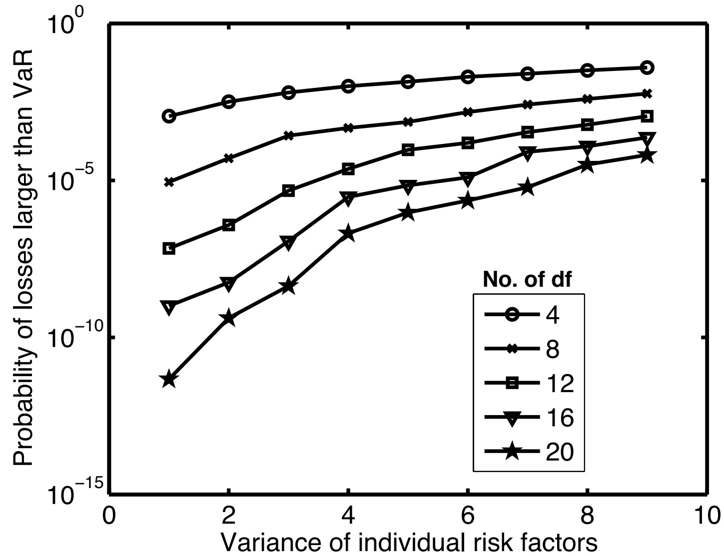

Figure 4: Probability of losses larger than the VaR for non unitary variances

6.3. Impact of default model parameters

After validating the model through comparison with crude MonteCarlo for the case of independent defaults, with both an alternative simulation approach and an asymptotic analysis, we can now employ the Cross-Entropy based simulator to explore the dependence of the probability of large losses on several parameters embedded in the model established in Section 2.

In particular, we consider the following model parameters, which widen the choice made in (Bassamboo et al., 2008):

1. the number of obligors n;
2. the weight ρ of the common risk factor;
3. the number β of degrees of freedom associated to the shock factor;
4. the variances $\sigma_{\eta_i}, i = 1, 2, \ldots, n$ of the individual risk factors.

We look at the way the probability of large losses varies as these parameters change individually. The Cross-Entropy method has been applied with samples of 10^4 elements as well as 10^5 to examine the impact of the sample size

Parameter	A	B	C	D
n	100	100	100	50÷200
$\{\sigma^2_{\eta_i}\}$	9	9	1÷9	9
β	12	4÷20	4÷20	12
$\{x_i\}$	5	5	5	$0.5\sqrt{n}$
l	25	25	25	$0.25n$
$\{a_i\}$	1	1	1	1
ρ	0.1÷0.9	0.25	0.25	0.25

Table 4: Parameter values

on the estimation accuracy. As reference values, against which we compare the CE results, we consider again a crude MonteCarlo (MC) simulation with samples of 10^6 elements. As recalled in Section 3 through Equation (14), that size allows us to estimate with a decent accuracy probability values as low as 10^{-4}. For that reason, in most of the results reported hereafter the range of values of the overall loss probability encompasses not too low values (typically larger than 10^{-4}), though the CE method is capable of estimating rarer events where the MonteCarlo method would require prohibitively large sample sizes.

The impact of the four parameters listed above is analysed respectively for the four scenarios, A through D, defined in Table 4.

We now examine separately the impact of each parameter.

Impact of the weight of the common risk factor. The parameters are set as in Case A of Table 4. The weight of the common risk factor is varied in the range [0.1-0.9]. The limiting values $\rho = 0$ and $\rho = 1$ correspond to two interesting cases, the latter of which can be evaluated quite simply. When $\rho = 0$ we have $X_i = \eta_i/Q$, so that the correlation among the latent variables is due just to the common shock factor Q. Instead, when $\rho = 1$ we have $X_i = Z/Q$, so that all the latent variables are identical, namely equal to a single random variable Ω distributed according to a Student's t-law with β degrees of freedom. In the general case where the thresholds a_i are different of one another, the possible values for the loss are given by the sum of all the elements of the subsets of the set $\{a_1, a_2, \dots, a_n\}$. For each value of l for which we can identify a subset $\mathcal{I}_u \subseteq \mathcal{I}$ such that $l = \sum_{i\in\mathcal{I}_u}$, the probability of losses is

$$\mathbb{P}[L = l] = \sum_{i \in \mathcal{I}_\mathrm{u}} \mathbb{P}[\Omega > x_i] \tag{41}$$

If the thresholds are all equal ($x_i = x$), the loss is either zero or $\sum_{i=1}^{n} a_i$, with probability respectively given by

$$\mathbb{P}[L = 0] = [\Omega < x], \tag{42}$$

$$\mathbb{P}[L = \sum_{i=1}^{n} a_i] = [\Omega > x], \tag{43}$$

Both cases (thresholds equal or different) can be evaluated, e.g., by the use of Table 26.10 in (Abramowitz and Stegun, 1964).

In Figure 5 (where the two Cross-Entropy datasets are labelled respectively 'small sample', corresponding to a sample size of 10^4, and 'large sample', sample size of 10^5, for simplicity), the loss probability has a very small dependence on the common risk factor weight, exhibiting a wide peak before pointing downward to the limiting value as $\rho \to 1$, where the exact loss probability is $1.55 \cdot 10^{-4}$. For the cases reported the standard relative error of the reference MonteCarlo estimate (i.e. its coefficient of variation) is always better than 10%. As to the performance of the Cross-Entropy approach the large sample case follows closely the reference MonteCarlo curve, while the tenfold reduction in the sample size appears to result in deviations from the MonteCarlo curve as large as 25%. It is to be noted that this CE simulation allows to detect the non-monotonic trend of the dependence on ρ, apparent in Figure 5, overlooked in previous IS simulations as those reported in (Bassamboo et al., 2008).

Impact of the number of degrees of freedom As to the effect of the number of degrees of freedom of the variable R in the shock factor, we consider the values $\beta = 4, 8, 12, 16, 20$. The remaining parameters are set as in (Bassamboo et al., 2008) and reported in Table 4 as Case B. The resulting probability $P[L > l]$ is plotted in Figure 6. Here the difference among the reference MonteCarlo and Cross-Entropy is negligible for both sample sizes. Again, the sample size in MonteCarlo simulation is such to have a relative standard error always better than 10%. As expected, the loss probability grows with the number of degrees of freedom, since the expected value of the common shock factor grows with β and therefore shifts back the overall value of the latent variable. The dependence of the loss probability on β is nearly exponential.

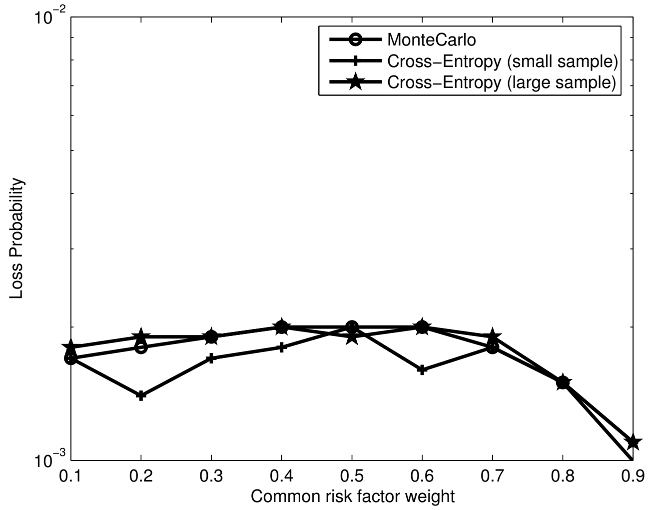

Figure 5: Impact of the common risk factor weight

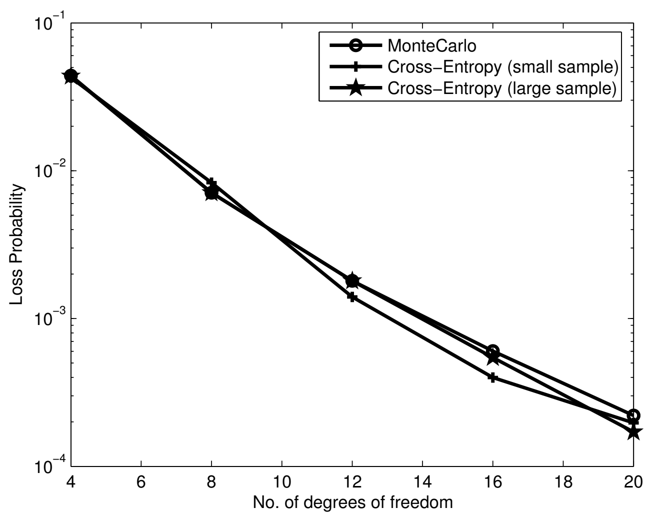

Figure 6: Impact of the number of degrees of freedom

Impact of the variance of the individual risk factors In this case we haven't got at our disposal the results obtained under Exponential Twist in (Bassamboo et al., 2008) and therefore set the other parameters as Case C in Table 4. We varied the variance in the range $1 \div 9$ and the number of degrees of freedom in the range $4 \div 20$. The resulting probability $P[L > l]$ is plotted in Figure 7. A smaller variance leads to rarer defaults, especially when the number of degrees of freedom associated to the shock parameter grows. While the curves appear to be quite regular for the larger sample size, the smaller sample size may be inadequate to provide accurate estimates when we have at the same time small variances of the individual risk factors and many degrees of freedom in the probability model of the shock factor.

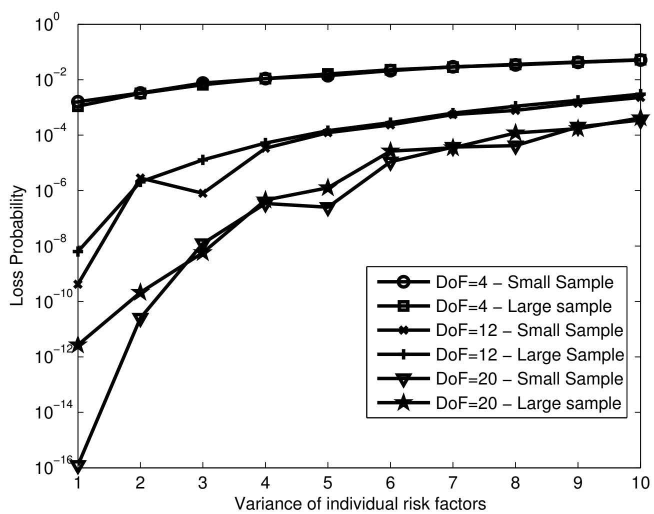

Figure 7: Impact of the variance of individual risk factors

Impact of the number of obligors This is labelled as Case D in Table 4. The number of obligors is set as $n = 50, 100, 150, 200$. The resulting probability $P[L > l]$ is plotted in Figure 8. Here we see that the CE results closely follow the MC ones, with some deviations when the sample size used in CE reduces.

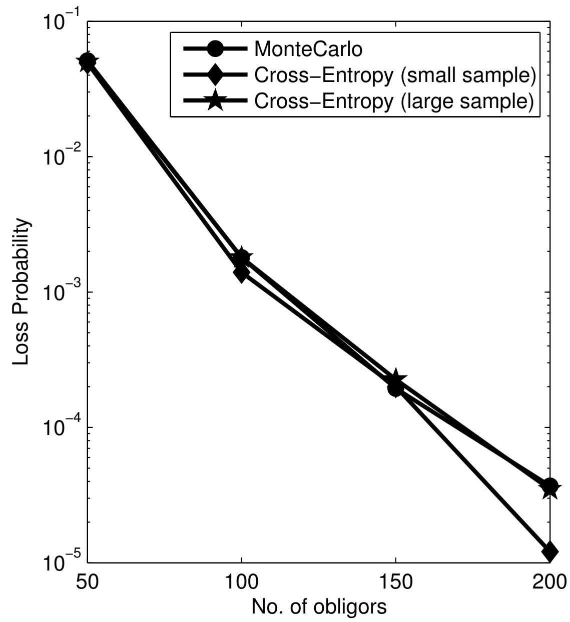

Figure 8: Impact of the number of obligors

7. Conclusions

A simulation method, based on the Cross-Entropy technique, has been proposed to evaluate the probability of large losses for a portfolio of securities subject to correlated default risk, described by a t-copula model. The comparison with crude MonteCarlo and non-adaptive Importance Sampling show that the CE technique allows to reach an adequate accuracy with a limited number of samples (10^5 for most cases). The CE-based method allows to extend the range of models that can be dealt with through Importance Sampling. The dependence of the loss probability on each one of the model parameters (number of obligors, weight of the common risk factor, number of degrees of freedom associated to the shock factor, and variance of the individual risk factors) has been evaluated for a large set of cases.

Acknowledgments

The authors wish to thank the referee Alberto Maria Bersani for his thoughtful review of the paper and his suggestions.

References

Abramowitz, M., Stegun, I., (1964). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York.

Alexander, C., (2009). Market Risk Analysis, Value at Risk Models. volume 4. Wiley.

Arulampalam, M., Maskell, S., Gordon, N., Clapp, T., (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. IEEE Transactions on Signal Processing, 50 (2), 174 –188.

Bassamboo, A., Juneja, S., Zeevi, A.J., (2008). Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation. Operations Research, 56 (3), 593–606.

Chan, J., Kroese, D., (2011). Rare-event probability estimation with conditional Monte Carlo. Annals of Operations Research, 189 (1), 43–61.

Crouhy, M., Galai, D., Mark, R., (2000). A comparative analysis of current credit risk models. Journal of Banking and Finance, 24 (1), 59–117.

D'Acquisto, G., Mastroeni, L., Naldi, M., (2012). Simulation of correlated financial defaults through smoothed cross-entropy, in: 14th International Conference on Computer Modelling and Simulation, UKSim 2012, Cambridge, United Kingdom, pp. 129–134.

D'Acquisto, G., Naldi, M., (2005). Cross entropy-based adaptive optimization of simulation parameters for markovian-driven service systems. Simulation Modelling Practice and Theory, 13 (7), 619–645.

Dietsch, M., Petey, J., (2004). Should SME exposures be treated as retail or corporate exposures? A comparative analysis of default probabilities and asset correlations in French and German SMEs. Journal of Banking and Finance, 28 (4), 773 – 788.

Frey, R., McNeil, A., (2003). Dependent defaults in models of portfolio credit risk. Journal of Risk, 6 (1), 59–92.

Gupta, G., Finger, C., Bhatia, M., (1997). Credit Metrics Technical Document. Technical report, J.P. Morgan & Co., New York.

Lando, D., (2009). Credit risk modeling, in: Mikosch, T., Kreiss, J.P., Davis, R.A., Andersen, T.G. (Eds.), Handbook of Financial Time Series. Springer Berlin Heidelberg, pp. 787–798.

Lucas, D.J., (1995). Default correlation and credit analysis. The Journal of Fixed Income, 4 (4), 76–87.

Mashal, R., Zeevi, A., (2002). Beyond correlation: Extreme co-movements between financial assets. Working paper, Columbia University.

Mastroeni, L., Naldi, M., (2011a). Options and overbooking strategy in the management of wireless spectrum. Telecommunication Systems, 48 (1-2), 31–42.

Mastroeni, L., Naldi, M., (2011b). Compensation policies and risk in service level agreements: A value-at-risk approach under the on-off service model, in: 7th International Workshop on Internet Charging and QoS Technologies, ICQT 2011, Paris, France, pp. 2–13.

Mastroeni, L., Naldi, M., (2011c). Long-range evaluation of risk in the migration to cloud storage, in: 13th IEEE Conference on Commerce and Enterprise Computing, CEC 2011, Luxembourg, pp. 260–266.

Naldi, M., D'Acquisto, G., (2008). A normal copula model for the economic risk analysis of correlated failures in communications networks. J. UCS, 14 (5), 786–799.

Naldi, M., D'Acquisto, G., Mastroeni, L., (2013). Solution space size in credit risk simulation, in: 15th International Conference on Computer Modelling and Simulation UKSim 2013, pp. 101–106.

Rubinstein, R., (1981). Simulation and the Monte Carlo Method. John Wiley and Sons, New York.

Rubinstein, R., Kroese, D., (2004). The Cross-Entropy Method. Springer, New York.

Rubinstein, R.Y., Glynn, P.W., (2009). How to Deal with the Curse of Dimensionality of Likelihood Ratios in Monte Carlo Simulation. Stochastic Models, 25 (4), 547–568.

Srinivasan, R., (2002). Importance Sampling. Springer.