

DEPARTMENT OF ACCOUNTING AND FINANCE

Don't Fall from the Saddle: the importance of higher moments of credit loss distributions

**Jan Annaert, Crispiniano Garcia Joao Batista,
Jeroen Lamoot & Gleb Lanine**

UNIVERSITY OF ANTWERP
Faculty of Applied Economics



Stadscampus
Prinsstraat 13, B.213
BE-2000 Antwerpen
Tel. +32 (0)3 220 40 32
Fax +32 (0)3 220 47 99
<http://www.ua.ac.be/tew>

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF ACCOUNTING AND FINANCE

**Don't Fall from the Saddle:
the importance of higher moments
of credit loss distributions**

**Jan Annaert, Crispiniano Garcia Joao Batista,
Jeroen Lamoot & Gleb Lanine**

RESEARCH PAPER 2007-010
JUNE 2007

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.213
phone: (32) 3 220 40 32
fax: (32) 3 220 47 99
e-mail: joeri.nys@ua.ac.be

The papers can be also found at our website:
www.ua.ac.be/tew
(research > working papers)

D/2007/1169/010

Don't Fall from the Saddle: the importance of higher moments of credit loss distributions¹

Jan Annaert
University of Antwerp

Crispiniano Garcia Joao Batista
Dexia Bank

Jeroen Lamoot²
Banking Finance and Insurance Commission (CBFA)

Gleb Lanine
Dexia Bank

February 2007

Abstract

The original Panjer recursion of the CreditRisk+ model is said to be unstable and therefore to yield inaccurate results of the tail distribution of credit portfolios. A much-hailed solution for the flaws of the Panjer recursion is the saddlepoint approximation method. In this paper we show that the saddlepoint approximation is an accurate and robust tool only for credit portfolios with low skewness and kurtosis of the loss distribution. However, often credit portfolios are heterogeneous with large skewness and kurtosis. We show that for such portfolios the commonly applied saddlepoint approximations (the Lugannani-Rice and the Barndorff-Nielsen formulas) are not reliable. We explain it by the dependence of the high-order standardized cumulants and the relative error on the saddlepoint. The more the cumulants and the relative error fluctuate for variations in the value of the saddlepoint (from 0 to the upper bound) the less accurate the saddlepoint approximation is. Hence, the saddlepoint approximations are not a universal substitute to the Panjer recursion algorithm. We also provide users of CR+ with a set of diagnostics to identify beforehand when the saddlepoint approximations are prone to failure.

JEL.

Keywords: CreditRisk+, saddlepoint approximations, Lugannani-Rice formula, Barndorff-Nielsen formula, credit VaR.

¹ We would like to thank Michael Gordy as well as participants at the First Conference of Advanced Mathematical Methods for Finance for helpful comments and suggestions. The remaining errors are the sole responsibility of the authors. The views and findings of this paper are entirely those of the authors. They do not necessarily reflect the views of the Banking, Finance and Insurance Commission (CBFA) or Dexia Bank.

² Corresponding author: Jeroen.Lamoot@cbfa.be

1. Introduction

Analytical models such as CreditRisk+ (CR+) of Credit Suisse First Boston (CSFB) have some advantages over Monte Carlo based credit risk models. By offering a full analytic description of the portfolio loss of a credit portfolio the CR+ model allows a quick calculation of the loss distribution, thereby avoiding time-consuming Monte Carlo simulations. Additionally, the model also requires a relatively limited dataset (Kurth, et al., 2002). However, Wilde (2000), Gordy (2002), and Haaf, Reiß and Schoenmakers (2003) draw attention to the numerical fragility and therefore inaccuracy of the Panjer recursion algorithm of CR+. Its numerical instability arises from an accumulation of numerical round-off errors due to the summation of numbers of similar magnitude but opposite sign. The larger the number of independent risk factors, the larger the number of obligors and the smaller the standardized loss unit, the longer the polynomials in the recurrence equation and, hence, the larger the possibility for round-off errors to accumulate. As an alternative to the Panjer recursion algorithm, Gordy (2002) introduces the Lugannani-Rice (LR) saddlepoint approximation (SPA) for fast and accurate computation of tail percentiles of the loss distribution in CR+. He finds that the SPA is extremely fast, accurate and robust for large portfolios with complex risk factor structures, exactly in the situations for which the Panjer recursion algorithm fails. Nevertheless, the SPA is less accurate in situations where the Panjer recursion is fast and reliable such as small portfolios with only one risk factor. Feuerverger and Wong (2000) test the accuracy of the Lugannani-Rice SPA introduced by Gordy (2002) as well as the alternative Barndorff-Nielsen (BN) SPA (Jensen, 1995) for large complex portfolios whose payoff functions contain linear (direct holdings in the underlying risk factor assets) and nonlinear (due to derivative securities) terms. They conclude that both SPAs are fast and accurate for any portfolio for which the risk factors are normally distributed with correctly specified covariance and for which a delta-gamma approximation of the nonlinear terms is appropriate.

Although the standard normal SPA is said to be extremely accurate in the tails and even exact for normal, gamma and inverse Gaussian distributions (Daniels, 1980), some notes of caution are expressed in the literature on their reliability for other distributions. Booth and Wood (1995), Beran and Ocker (2003), Studer (2001), and Merino and Nyfeler (2003) warn that the standard normal SPA should be applied cautiously as it may produce large errors for certain distributions.

Our purpose is, first, to show that caution is warranted when applying the LR and BN³ formulas in the CR+ framework on certain credit portfolios. Second, we provide users of CR+ with a set of diagnostics to identify beforehand when the LR and the BN SPA can yield highly inaccurate distribution estimates. Third, we also show the sensitivity of the accuracy of the LR and BN estimates for concentrations to single exposures. Initially, we test the accuracy and the robustness of the LR and the BN formulas on a set of credit portfolios compiled from a large real-life portfolio. In contrast to Gordy (2002), we find that there exists a class of credit portfolios for which both SPAs perform poorly, even in a simple one risk factor setting. Both formulas yield accurate approximations of the credit loss distributions for homogenous portfolios⁴. However, in practice credit portfolios are often heterogeneous and are thereby characterized by a loss distribution with large skewness and kurtosis. We show that applying the LR and BN formulas to such credit portfolios yields unreliable approximations in the tails of the loss distribution. Moreover, the LR formula fails in the sense that sometimes negative probabilities are generated. The large relative errors of the SPA are explained by the dependency of the third $\zeta_3(\theta)$ and the fourth $\zeta_4(\theta)$ standardized cumulants and the relative error on the saddlepoint θ . We find that the larger $\zeta_3(\theta)$, $\zeta_4(\theta)$ and hence the relative error are at $\theta = 0$, the less accurate the SPA is. To provide the users of CR+ with some guidance for the application of the LR and BN formulas, we assess the accuracy and robustness of both formulas on numerous simulated portfolios of different credit quality, ranging from absolutely homogeneous to extremely heterogeneous. In general, the skewness and the kurtosis of high credit quality portfolios are very sensitive to concentrations on single exposures. For these portfolios we thus find that the accuracy of the LR and BN formulas deteriorates fast as the concentration on the exposure increases. Moreover, the level of concentration does not have to be extremely high before both formulas' accurateness deters strongly. To conclude, we warn against applying both SPAs on credit portfolios without first checking the third and the fourth moments of the loss distribution. At the same time we corroborate Gordy's (2002) results that under a simple one risk factor specification of CR+, the Panjer recursion algorithm gives accurate results for both homo- and heterogeneous portfolios, regardless of the third and the fourth moments of the credit loss distributions.

³ The LR and BN formulas are two favoured SPA by practitioners.

⁴ A homogenous portfolio is understood to be a portfolio in which the maximum exposure to a single obligor is less than 1% of total book value.

The plan of the paper is as follows. The next section gives a short explanation of the SPA technique. The third section gives a discussion of the literature that describes certain conditions to be fulfilled to obtain accurate SPA estimates. In the fourth section we test the accuracy of the Panjer recursion algorithm, the LR and the BN formulas on five real-life credit portfolios with loss distributions exhibiting varying degrees of skewness and kurtosis. In the fifth section we show for a multitude of simulated portfolios the sensitivity of the LR and BN estimates to the skewness and kurtosis of the loss distributions (resulting from concentrations on single exposures). The last section concludes.

2. Saddlepoint approximations

Martin et al. (2001) and Gordy (2002 and 2004) were the first to employ the technique of SPA for the calculation of the loss distribution of credit portfolios. Underlying the SPA technique is the Edgeworth expansion. This expansion gives an approximation of the density of a centered random variable for which the closed-form solution is not known. In general, Edgeworth expansions are found to work well at the center of a distribution but not at the tails of a distribution. The trick of the SPA is to use the Edgeworth expansion precisely where it works well; in the center of a distribution. The density of the original distribution $f(x)$ is estimated by ‘tilting’ $f(x)$ to a new distribution $f(\theta; x)$ which is centered around x . This new tilted distribution is then approximated by the Edgeworth expansion and finally the mapping is inverted to obtain the approximation of $f(x)$.⁵ SPAs are of interest in many different applications as they are found to provide accurate approximations to tail probabilities. However, a prerequisite to employ the SPA technique is that the cumulant generating function (cgf) of the distribution must have a tractable form (Jensen (1995), Studer (2001), Gordy (2002)).

Let f_n be the density of $\bar{X} = (X_1 + \dots + X_n) / n$, where the X_i are i.i.d. random variables. The low-order standard normal SPA of f_n has the following form

$$f_n(x) = \exp\left(n(K(\theta) - \theta x)\right) \cdot \sqrt{n/2\pi\sigma^2(\theta)} \cdot \left(1 + O\left(n^{-1}\right)\right), \quad (1)$$

⁵ In this paper only final formulas for SPA which are useful for the further discussion are presented. An extensive discussion of SPA can be found in the textbook by Jensen (1995).

where n denotes the number of random variables, $\sigma^2(\theta)$ the variance of the tilted distribution, $K(\theta)$ the cgf and O the relative error term. The saddlepoint θ is chosen in such a way that the first cumulant (i.e. the mean) equals x , the point for which we want to know the density. The high-order standard normal SPA formula is as follows:

$$f_n(x) = \exp(n(K(\theta) - \theta x)) \cdot \sqrt{n/2\pi\sigma^2(\theta)} \cdot \left(1 + \frac{1}{n} \left(\frac{\zeta_4(\theta)}{8} - \frac{5\zeta_3(\theta)^2}{24} \right) + O(n^{-2}) \right), \quad (2)$$

where ζ_i are the i th standardized cumulants defined as $\zeta_i(\theta) = \kappa_i(\theta) / \kappa_2^{i/2}(\theta)$, with κ_i the i th cumulant. Tail probabilities are given by the following formula

$$\Pr(\bar{X} \geq x) = \frac{\exp(n(K(\theta) - \theta x))}{\sqrt{n\theta\sigma(\theta)}} \times \left(B_0(\lambda) + \frac{\zeta_3(\theta)}{6\sqrt{n}} B_3(\lambda) + \frac{1}{n} \left(\frac{\zeta_4(\theta)}{8} B_4(\lambda) - \frac{\zeta_3(\theta)^2}{72} B_6(\lambda) \right) + O(B_0(\lambda)n^{-3/2}) \right), \quad (3)$$

where $B_i(\lambda)$ are the Esscher functions (Jensen, 1995).

The SPA (3) for tail probabilities is cumbersome and does not have a simple relation to other well-known quantiles in statistics. Practitioners favour more simple formulas expressed in interpretable quantiles, such as the LR approximation (Jensen, 1995):

$$\Pr(\bar{X} \geq x) = 1 - G(x) \approx 1 - \Phi(\omega) + \phi(\omega) \left(\frac{1}{\nu} - \frac{1}{\omega} \right), \quad (4)$$

where $G(x)$ is the cumulative density function (cdf) of random variable \bar{X} , Φ and ϕ denote resp. the cdf and density of the standard normal distribution, $\omega = \text{sign}(\theta) \sqrt{2n[\theta x - K(\theta)]}$ and $\nu = \theta \sqrt{nK''(\theta)}$. In the last expression $K''(\theta)$ is the second derivative of K w.r.t. θ .

An alternative to the LR formula is the BN formula (Jensen, 1995):

$$\Pr(\bar{X} \geq x) = 1 - G(x) \approx 1 - \Phi \left(\omega + \frac{1}{\omega} \log \frac{\nu}{\omega} \right). \quad (5)$$

For the empirical analysis we focus on the LR and the BN formulas as both formulas are the favoured SPA by practitioners for the calculation of the loss distribution of portfolios.

3. Conditions for accurate SPA

The authors Wood et al (1993), Booth and Wood (1995), Studer (2001) and Beran and Ocker (2003) study the reliability of standard normal SPAs and find that for specific distributions the SPAs can give highly inaccurate results. Typically, these distributions have large third and fourth standardized cumulants. Studer (2001) concludes that the standard normal SPA can give very good results only for well-behaved distributions, i.e. distributions whose fourth cumulant is not too large. For distributions with large higher cumulants, the Edgeworth expansion is found to yield imprecise probability estimates in the center of its distribution (exactly where it should work well) (Jensen, 1995). Therefore, the large higher standardized cumulants result in a large relative error for the probability estimate of the standard SPA. Moreover, a less documented but elementary condition for the SPA technique to give accurate estimates is for the relative error $(A(\theta) = \frac{1}{n} \left(\frac{\zeta_4(\theta)}{8} - \frac{5\zeta_3(\theta)^2}{24} \right) + \dots)$ to be independent of the saddlepoint (θ) (or $A(\theta) = A$) (Barndorff-Nielsen and Cox (1979), Daniels (1954, 1980), and Jensen (1995)).

Daniels (1980) proves that only for three distributions – the normal, the gamma and the inverse-normal distribution – the independence condition is fulfilled and, hence, the standard normal SPA is exact, i.e. SPA estimates give for all n exactly the densities of the distributions. For the normal distribution the higher standardized cumulants (and thus the relative error) equal zero, hence the relative error and the standardized cumulants are independent of the saddlepoint. The relative error for the gamma distribution remains constant (but not zero) for all possible values of θ .⁶ A special case is the inverse Gaussian distribution, Jensen (1995) shows that although the standardized cumulants $\zeta_i(\theta)$ depend on the saddlepoint θ and can even increase without bound, the SPA is exact, because the relative error is identical to zero for varying θ .

⁶ Strictly speaking, the fact that the relative error is not zero implies that the *renormalized* SPA, i.e. the SPA multiplied by a constant such that it integrates to one, is exact.

The condition of independence of the relative error on the saddlepoint does not hold for the one risk factor CR+ specification.⁷ Limiting us to the leading term of the asymptotic expansion for the relative error

$$\left(\zeta_4(\theta)/8 - 5\zeta_3(\theta)^2/24\right), \quad (6)$$

we show in the appendix A that the error term (6) depends on the saddlepoint θ .⁸ Moreover, the error term (6) has the limit $-\sigma^2/12$, where σ^2 is the variance of the risk factor and $\hat{\theta}$ is the upper bound of the valid range of saddlepoints, calculated according to the inequality derived by Gordy (2002). In the appendix we also show that for the CR+ specification the higher order standardized cumulants $\zeta_3(\theta)$, $\zeta_4(\theta)$ depend on θ and have the following limits:

$$\lim_{\theta \rightarrow \hat{\theta}} \zeta_3(\theta) = 2\sigma \quad \lim_{\theta \rightarrow \hat{\theta}} \zeta_4(\theta) = 6\sigma^2. \quad (7)$$

In practice we are more interested in distributions for which the SPA formulas will give accurate instead of exact estimates. Accuracy is the best thing to exactness. Daniels (1954) shows that for several density specifications the standard normal SPA yields accurate estimates, as those specifications approximate the normal or the gamma distributions for the saddlepoint $\theta \rightarrow \hat{\theta}$. More recent results from Daniels (1980), Booth and Wood (1995) and Jensen (1995) suggest that examining the behaviour of the third and the fourth standardised cumulants and the leading error term in the asymptotic expansion (6) across the valid range of θ may provide a useful diagnostic tool to ascertain the accuracy of the SPA.

4. The importance of higher moments for the saddlepoint approximations

In this section we illustrate the importance of the third and the fourth moments on the accuracy and the robustness of the VaR-estimates given by the Panjer recursion, the LR and BN SPAs for the CR+ model, by applying the methods on a set of portfolios randomly sampled from a real-life credit portfolio.

⁷ The conclusions derived here are also valid for the multifactor CR+, though the limits for the multifactor model differ from those derived for the one factor model. However, we do not present the limits of the standardized cumulants and correction terms for the multifactor CR+ as the proof is messy and without additional insight.

⁸ As the relative error is an asymptotic expansion specification (6) is not sufficient to prove the independence as the remaining part of the asymptotic expansion may eliminate any dependency. However, we restrict ourselves to the specification in (6) as in empirical applications this term will dominate.

The dataset

The large real-life credit portfolio consists of 3,000 exposures for which we know the exposure size and the unconditional probability of default. Each of the five sampled credit portfolios consists of 1,000 obligors.

Table 1. Summary statistics.

This table provides summary statistics for the five real-life credit portfolios randomly sampled from a large real-life credit portfolio. Each of the sampled portfolios consists of 1,000 obligors. We raise the skewness and kurtosis of loss distribution of the portfolios by increasing the maximum exposure to a single borrower and increasing the difference between the first and second largest exposures. In the first part of the table we report the statistics of the exposure size distributions. Summary statistics of the credit loss distributions are reported in the second part of the table. The summary statistics of the credit loss distributions were calculated for one systematic risk factor CR+ specification with LGD=0.3 and $\sigma^2=1$.

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5
Exposure size distribution					
Maximal exposure*	0.94%	1.50%	1.73%	4.35%	4.29%
Minimal exposure*	0.00%	0.00%	0.00%	0.00%	0.00%
Skewness	3.59	5.07	5.86	12.46	12.23
Kurtosis	19.34	37.02	50.30	248.28	198.98
Credit loss distribution					
Expected Loss*	0.33%	0.04%	0.03%	0.03%	0.03%
St Deviation*	0.35%	0.06%	0.05%	0.05%	0.06%
Skewness**	2.03	3.14	4.37	4.71	5.68
Kurtosis**	9.11	18.15	36.34	60.66	81.27

* expressed in percentage to book value

** skewness and kurtosis were calculated without preliminary dividing the portfolio into exposure bands and therefore differ from the results reported in Appendix D.

To introduce heterogeneity in the portfolios we increase the size of the maximum exposure to a single borrower from 0.94% of total book value in portfolio 1 to 4.35% and 4.29% in portfolios 4 and 5 respectively.⁹ At the same time we gradually increase the difference between the first and the second largest exposures. By doing so, we are able to increase the skewness and the kurtosis of the exposure size and the loss distributions. All portfolios, except portfolio 1, have almost

⁹ Concentrations on a single exposure as in portfolios 4 and 5 may not appear at the level of a financial institution as they are prohibited by regulatory requirements (eg. EU large exposure rule). This level of concentration may arise, though, in certain subportfolios of the institution.

identical values of expected loss and standard deviation in relative terms, but differ thus significantly in their skewness and kurtosis. Portfolios 3 and 4 are built in such a way that they differ remarkably only in kurtosis. As shown in appendix B due to the heterogeneity, portfolios 4 and 5 exhibit loss distributions with spikes far in the tails.¹⁰ Summary statistics of exposure size and credit loss distributions are reported in table 1.

The set-up

VaR estimates for different levels of confidence are computed via the original Panjer recursion algorithm and the LR¹¹ and the BN formulas. We restrict our analysis to a simple specification of the CR+ model, assuming exposure to only one systematic risk factor¹² with variance $\sigma^2 = 1$ and loss-given-default (LGD) equal to 0.3. For a further specification of the employed methodology to estimate the loss distribution via the Panjer recursion and the LR in the CR+ framework we refer to respectively CSFP (1997) and Gordy (2002). The Panjer, LR and BN estimates are compared with the loss distribution estimates given by the Monte Carlo simulation of the CR+ model with one systematic risk factor. This procedure has the advantage that we can compute a confidence interval around the VaR numbers.

In the one factor specification the probability of default for obligor i conditional on the risk factor x is given by

$$p_i(x) = \bar{p}_i x, \tag{9}$$

where \bar{p}_i is the unconditional default probability of obligor i (for example, given by rating agencies or bank experts).¹³ The risk factor x is assumed to be an independent gamma variable with mean one and variance one. CR+ assumes that defaults follow a Poisson distribution with

¹⁰ The distributions are estimated using a Monte Carlo simulation, cfr. infra.

¹¹ In order to avoid root-solving for θ , a strategy based on interpolation is applied (Gordy, 2002). This strategy is efficient when it is required to calculate VaRs for several probabilities. At first, the upper bound $\hat{\theta}$ is computed and then a fine grid of 1,000 values in the open interval $(0, \hat{\theta})$ is formed. At each point in the grid we calculate the pairs of losses and corresponding probabilities and then interpolate to find the loss corresponding to a concrete target solvency probability.

¹² The idiosyncratic risk is not included in the model. It is the simplest and at the same time the most prudent specification of CR+, because it assumes the highest correlation between the default probabilities. Inclusion of the idiosyncratic risk does not change the conclusions, but instead raises questions about the weights attached to the risk factors.

¹³ The generalized K -factor specification of CR+ is given in Gordy (2002).

intensities $p_i(x)$. We therefore perform the Monte Carlo simulations in four steps: (1) simulate the realization of the risk factor x from the gamma distribution $\Gamma(1, 1)$; (2) compute the probability of default for each obligor from equation (9); (3) simulate default events for each obligor from the Poisson distribution and calculate the total portfolio loss; (4) repeat this procedure $N=100,000$ times.

The error of a VaR estimate is calculated as the difference between the estimate given by the Panjer recursion, the LR or BN SPA and the Monte Carlo estimate. To assess the significance of the error of the VaR estimate we derive confidence intervals as described in Pritsker (1997). To construct a 95% confidence interval for the p^{th} percentile of a distribution generated by the Monte Carlo simulation, we solve for L and H such that $\Pr[V_L \leq VaR \leq V_H] = p$, where V_i , for $i = 1, \dots, N$, are the portfolio values generated by the simulation and sorted in ascending order. V_L and V_H are the bounds for the confidence interval. These bounds are chosen such that:

$$\begin{aligned} \sum_{i=L}^{H-1} \binom{N}{i} p^i (1-p)^{N-i} &\geq 0.95 \\ \sum_{i=L+1}^{H-1} \binom{N}{i} p^i (1-p)^{N-i} &\leq 0.95. \end{aligned} \tag{10}$$

Moreover, H and L are chosen such that the confidence interval is as close to symmetric as possible.

The accuracy of the VaR estimates

In table 2 we report Monte Carlo, Panjer, LR and BN VaRs for all five portfolios and the probability that the Monte Carlo loss will exceed these estimates. We show the results only for VaR_α at 90%, 95%, 97.5%, 99% and 99.5% levels, which are the most relevant quantiles in practice. To ease the comparison of the portfolios, we express VaR as a percentage of total book value. We also test whether deviations from the Monte Carlo simulations are statistically significant.

We conclude that under the simple one factor specification the Panjer recursion algorithm indeed performs better than the LR and the BN formulas, although its accuracy decreases for higher skewness and kurtosis of the loss distribution. This decrease in accuracy of the Panjer recursion is

consistent with Gordy (2002) who finds that the potential for errors increases as the skewness of the loss distribution is increased¹⁴. Indeed, the Panjer recursion algorithm gives correct approximations in the tail of Portfolio 1, but for portfolios 4 and 5 the approximations are correct only starting from $\alpha = 97.5\%$.

The accuracy of the LR and BN formulas, however, deteriorates significantly when the skewness and the kurtosis of the loss distributions increase. For example, in portfolio 1, with minimum skewness and kurtosis, the difference between the VARs given by the LR or BN formulas and the Monte Carlo VaR are statistically insignificant. Starting from portfolio 3 both SPAs produce incorrect estimates with statistically significant deviations over the whole tail of the loss distribution. There is no systematic pattern in the behavior of the LR and the BN formulas. In some cases the VaR's are significantly overestimated (portfolios 2 and 3), in others significantly underestimated (portfolios 4 and 5 for $\alpha = 90\%$). Hence, the users of the SPAs are not always on the safe side, especially for heterogeneous portfolios highly inaccurate results can be obtained.

As a robustness check for these results we assess the extent of the deviations between the estimated distribution of the Panjer recursion, LR and BN and the distribution calculated via the Monte Carlo simulation. This goodness-of-fit of the distributions is assessed by the Kolmogorov-Smirnov test. The results are summarized in the last column of table 2. For each portfolio we compare the tail distribution produced by the Panjer algorithm and both SPAs against the Monte Carlo simulation. In not a single case we can reject the null hypothesis about the equality of distributions produced by the Panjer algorithm and Monte Carlo simulations. On the contrary, we reject the null hypothesis of equality of the distributions produced by the LR or BN formula and the Monte Carlo simulation for portfolios 4 and 5.

¹⁴ Gordy (2002) shows that the larger the number of sectors, the larger the portfolio and the smaller the standardized loss unit, the longer the polynomials in the recurrence equation, so the greater the opportunity for round-off errors to accumulate and for the Panjer recursion algorithm to fail. However, for small and medium-sized portfolios (less than 5,000 obligors) and a simple one factor CR+ specification, the original Panjer recursion algorithm is stable and reliable.

Table 2. Risk estimates and distribution characteristics of the credit portfolios.

In this table we report Monte Carlo, Panjer, LR and BN VaRs for the five portfolios and the probability that Monte Carlo loss will exceed these estimates. We show results only for VaR_{α} at 90%, 95%, 97.5%, 99% and 99.5% levels. VaR is expressed as a percentage of the total book value. The goodness-of-fit is estimated by the Kolmogorov-Smirnov test.

Portfolio 1	90	95	97.5	99	99.5	KS value
MC VaR [‡]	0.782	1.022	1.271	1.589	1.836	
Panjer VaR [‡]	0.783	1.027	1.271	1.593	1.837	0.026
Pr[MC loss > Panjer VaR] (%)	9.959	4.917	2.495	0.989	0.499	
LR VaR [‡]	0.779	1.022	1.264	1.586	1.828	0.026
Pr[MC loss > LR VaR] (%)	10.077	5.007	2.549	1.013	0.509	
BN VaR [‡]	0.782	1.025	1.268	1.590	1.833	0.026
Pr[MC loss > BN VaR] (%)	9.989	4.954	2.523	0.999	0.505	
Portfolio 2						
MC VaR [‡]	0.115	0.167	0.224	0.311	0.374	
Panjer VaR [‡]	0.112	0.166	0.223	0.308	0.374	0.051
Pr[MC loss > Panjer VaR] (%)	10.360***	5.100	2.527	1.038	0.500	
LR VaR [‡]	0.126	0.185	0.245	0.324	0.385	0.179
Pr[MC loss > LR VaR] (%)	8.498***	4.020***	1.999***	0.858***	0.460	
BN VaR [‡]	0.128	0.187	0.247	0.326	0.387	0.205
Pr[MC loss > BN VaR] (%)	8.244***	3.949***	1.958***	0.845***	0.454**	
Portfolio 3						
MC VaR [‡]	0.075	0.113	0.152	0.210	0.258	
Panjer VaR [‡]	0.072	0.112	0.152	0.211	0.264	0.051
Pr[MC loss > Panjer VaR] (%)	10.499***	5.088	2.514	0.987	0.462	
LR VaR [‡]	0.080	0.141	0.200	0.278	0.337	0.231
Pr[MC loss > LR VaR] (%)	9.065***	2.957***	1.173***	0.393***	0.224***	
BN VaR [‡]	0.086	0.144	0.202	0.280	0.339	0.256
Pr[MC loss > BN VaR] (%)	8.062***	2.800***	1.128***	0.389***	0.221***	
Portfolio 4						
MC VaR [‡]	0.087	0.127	0.173	0.257	0.305	
Panjer VaR [‡]	0.084	0.125	0.173	0.255	0.306	0.051
Pr[MC loss > Panjer VaR] (%)	10.479***	5.163**	2.501	1.023	0.485	
LR VaR [‡]	0.049	0.054	0.062	0.409	0.514	0.667***
Pr[MC loss > LR VaR] (%)	20.973***	18.952***	15.913***	0.114***	0.031***	
BN VaR [‡]	0.055	0.165	0.277	0.417	0.520	0.308***
Pr[MC loss > BN VaR] (%)	18.426***	2.760***	0.778***	0.104***	0.030***	
Portfolio 5						
MC VaR [‡]	0.089	0.129	0.181	0.262	0.320	
Panjer VaR [‡]	0.086	0.127	0.179	0.259	0.317	0.077
Pr[MC loss > Panjer VaR] (%)	10.630***	5.176**	2.555	1.045	0.509	
LR VaR [‡]	0.044	0.051	0.062	0.087	0.556	0.769***
Pr[MC loss > LR VaR] (%)	24.908***	21.507***	16.967***	10.425***	0.062***	
BN VaR [‡]	0.050	0.185	0.303	0.451	0.562	0.282*
Pr[MC loss > BN VaR] (%)	21.798***	2.387***	0.599***	0.184***	0.059***	

Asterisks indicate the level of significance as follows: (*) 0.10, (**) 0.05, (***) 0.01.

We use the following abbreviations: **MC** – Monte Carlo simulations; **LR** – the Lugannani-Rice formula; **BN** – the Barndorff-Nielsen formula; **KS** – Kolmogorov-Smirnov test.

[‡] - reported as a percentage of the total book value.

Moreover, for the portfolios characterized by loss distributions with large skewness and kurtosis the LR formula proves to be fragile.¹⁵ The effect of the instability of the LR formula can be well shown via the loss exceedance curve, which gives for each loss value the probability of exceeding this loss. For a loss distribution to be a reliable approximation of the true loss distribution at least two properties must hold. First, the loss exceedance curve must be a non-increasing function of the loss value. It may be flat on some intervals implying that the probability of the loss occurring in those intervals is zero, but an increasing function would imply negative probabilities for some losses. Second, the loss exceedance curve can never have negative values. Appendix C shows the loss exceedance curves for the first and the fifth credit portfolios.¹⁶ The loss exceedance curve of the first portfolio obeys the two properties. However, the loss distribution of portfolio 5 can never be a close approximation of the true loss distribution. The loss exceedance curve of the fifth portfolio is not continuously decreasing, signaling that the credit loss distribution given by the LR formula is totally unreliable.

The accuracy condition of SPA

As we discussed in section 3, to assess the accuracy of the LR and BN estimates we examine the sensitivity of the third and the fourth standardized cumulants and the error term for variations of θ (across the range 0 to $\hat{\theta}$).

In appendix D we plot the values of the third and fourth standardized cumulants and the relative error versus θ for the one systematic risk factor CR+ model with $\tau=1$ ¹⁷. Under these assumptions the limits of $\zeta_3(\theta)$ and $\zeta_4(\theta)$ as given in (7) are 2 and 6 respectively. The limit of the relative error (6) is -0.083. At $\theta=0$ the third standardized cumulant of portfolio 1 is 2.026, the fourth is 6.117

¹⁵ The fragility of the LR approximation has also been reported by Booth and Wood (1995).

¹⁶ For portfolios 2 and 3 we obtain similar results as for portfolio 1 whereas the results for portfolio 4 confirm those of portfolio 5. We refrain from reporting them to save space.

¹⁷ In order to check whether these conclusions hold in a multifactor model, we increased the number of risk factors to 2 and 5. We came to exactly the same conclusions as in the one factor model. In order to save space, we do not report the figures of the third and fourth standardized cumulants and the relative error for multifactor CR+. These figures are available on request.

and the relative error is -0.09.¹⁸ They are initially very close to the limits, and converge fast as $\theta \rightarrow \hat{\theta}$. In fact, the cumulants remain almost stable as $\theta \rightarrow \hat{\theta}$, therefore it is not surprising that the LR and the BN formulas give accurate estimates for these portfolios. In contrast to portfolio 1, the third and the fourth standardized cumulants of portfolios 4 and 5 are initially far from the limits¹⁹ and vary considerably with θ , causing huge variation in the relative error. Therefore they converge to the limits very slowly explaining the inaccuracy of the LR and BN estimates for these portfolios.

To summarize, the LR and BN formulas yield accurate estimates for the credit loss distributions of which the variation of the third and fourth standardized cumulants ($\zeta_3(\theta)$, $\zeta_4(\theta)$) remains limited across the valid range of θ . In these cases the standardized cumulants or the relative error of the loss distributions are seen to be relatively small and to converge relatively fast towards their limits. However, if for a loss distribution the standardized cumulants and the error term vary strongly with the saddlepoint (and deviate significantly from their limits) then the slower they converge to their limits and the less accurate the SPA is.

Therefore, a general rule of thumb to assess the potential accuracy of the employed SPA formula is to, first, compute the third and the fourth standardized cumulants at $\theta = 0$. If the standardized cumulants and the correction term diverge significantly from their limits, then the formula is likely to produce large relative errors.

Several solutions have been brought up to increase the accuracy of the standard normal SPA. First, Studer (2001) points at higher-order SPA (2) as a partial solution for the large relative error given by the standard low-order SPA. Although this solution leads to a decrease in the relative error it does not solve the inaccuracy problem as often this error remains quite large. Including fifth and sixth standardized cumulants into (2) probably would add little in terms of accuracy. Instead, the formula will be more complicated and the practical intuition behind the formula will

¹⁸ Before calculating the standardized cumulants all exposures are grouped in 100 exposure bands. Therefore the standardized cumulants at $\theta = 0$ differ from the skewness and kurtosis reported in table 1, where the results are obtained without preliminary dividing portfolios into exposure bands.

¹⁹ The third standardized cumulants of portfolios 4 and 5 are 3.9 and 4.9 respectively. The fourth standardized cumulants are 38.2 and 59.4 respectively.

be lost. Second, to capture the bimodal shape of the loss distribution for their credit portfolio composed of a few large exposures Beran and Ocker (2003) introduce a ‘recursive’ SPA. This approximation consists of applying the standard normal SPA on a sample of the credit portfolio for which the extreme exposures are excluded and afterwards adjusting the preliminary estimated distribution for the extreme exposures. However, although, the recursive approximation considerably improves the accuracy of the distribution estimates, the error often remains important. These two proposed solutions still use the normal distribution as the underlying distribution for the applied SPA formulas. The third solution, proposed by Wood et al (1993), and Booth and Wood (1995) is to replace the normal distribution in the LR and the BN formulas by another distribution. For the approximation of tail probabilities of the first passage time of a random walk with drift process they propose the Inverse Gaussian as an alternative underlying distribution. The modified (inverse Gaussian based) LR and BN SPAs prove to give highly accurate estimation results in the tails when the distribution is characterized by large higher moments. However, the modified (inverse Gaussian based) SPA can never totally replace the normal based SPA as the distribution estimates show to be less accurate when the higher standardized cumulants are low. Hence, one ‘standard-to-fit-all’ SPA has not been developed yet.

5. Concentration and the accuracy of the SPAs

In the previous section we showed the potential inaccuracy of the LR and BN distribution estimates for the CR+ framework and gave the explanation for the inaccuracies. In this section we assess the sensitivity of the accuracy of the LR and BN formulas to different levels of skewness and kurtosis of the loss distributions, derived by altering the level of concentration to a single exposure.

To assess this relationship we construct numerous simulated portfolios ranging from absolutely homogeneous to extremely heterogeneous (i.e. highly concentrated on a single exposure). All portfolios include 10,000 obligors. Each portfolio consists of obligors with equal credit quality (AAA, AA, A, BBB, BB, B and CCC). This may be a rather restrictive assumption and may seem out of line with real-life portfolios, however, the assumption is made to better illustrate how the accuracy of the SPA behaves for portfolios of different credit quality. The initial portfolio is a

fully homogeneous one; all exposures are equal to 1,000. Starting from this initial portfolio we gradually increase the size of only one exposure and measure the impact of the concentration on the skewness and the kurtosis of the credit loss distribution.

The relationship between the higher cumulants and the concentration on the single exposure (for a one systematic risk factor CR+ with $\tau = 1$) is depicted in appendix E. The concentration is increased from 0.02% until 10%: higher concentrations are not shown as it is clear on both figures that skewness and kurtosis reach their limits at a 10% concentration. These extreme concentrations, though, are solely retained for illustrative purposes, the following discussion does not depend on these extreme values. The figures of appendix E show that high credit quality portfolios, consisting of only AAA, AA and A obligors, are very sensitive to concentrations on single exposures. Skewness and kurtosis increase sharply and deviate much from the theoretical limits (7) already at a 1% concentration on a single exposure. On the contrary, skewness and kurtosis of lower credit quality portfolios (BBB, BB, B, CCC and lower) are less sensitive to concentrations on single exposures.

The intuitive explanation for this result is the following: the loss distribution for high credit quality portfolios is concentrated around zero and has a long tail. The number of defaults is relatively low and the losses are, given the low default probability, situated far in the tail. Concentrations on single exposures strongly ‘fatten’ and provoke humps or spikes in the tail of the loss distribution for these portfolios, as the concentration translates in a severe loss far in the tail of the distribution. Hence, the concentration strongly increases the level of skewness and kurtosis. In contrast, the mass of the loss distribution for low credit quality portfolios is less concentrated at zero as defaults are more likely. Concentrations also affect the tail of the loss distribution for low credit quality portfolios, however, the impact on the skewness and kurtosis is not as strong as for the high credit quality portfolios (the losses are situated less far in the tail of the distribution).

To assess the impact of the concentration to single exposures (and thereby the skewness and kurtosis) on the accuracy of the LR and BN formulas in the CR+ framework, we compare the analytical VaR, given by the LR and BN formulas, with the VaR computed by the Monte Carlo simulations with 100,000 runs for each portfolio (as we did in the previous section). To ease the

comparison, instead of using absolute values, we compute the ratio of analytical VaR to Monte Carlo VaR. We assess the accuracy by the Mean Absolute Error (MAE)²⁰ in the upper 1% tail of the credit loss distributions:

$$MAE = \frac{\sum_i \left| 1 - \frac{AnalyticalVaR}{MonteCarloVaR} \right|}{Nobs}, \quad (11)$$

where i ranges from 99% till 99.75% with step 0.01%, and $Nobs$ is number of quantiles used in the nominator.

Table 3 reports the results of the accuracy of the LR and BN formulas for the simulated portfolios with different credit quality and increasing concentration on a single exposure²¹. As documented in the previous section the general finding is that as the concentration on a single obligor and thus the skewness and kurtosis increases, both the LR and the BN formulas produce less accurate estimates in the tail of the loss distribution. However, the important finding reported in table 3 is that for the high credit quality portfolios the accuracy already disappears at relatively low levels of concentration. For instance for the AAA portfolio, the accuracy of the LR and BN estimates is certainly lost when the concentration exceeds the 0.2% of the total portfolio. At this level of concentration the average absolute difference (over the 99% to 99.75% probability range) between the LR or BN VaR estimates and the MC estimates amounts to 25%. In other words, the average economic capital (computed via the LR and BN estimates) that would be derived over the 99% to 99.75% range of confidence intervals would be over- and / or underestimated by 25%. Also for the AA and A portfolios the accuracy diminishes rapidly at relatively low levels of concentration. A same level of over- and/or underestimation of the average VaR's (across the range of confidence intervals) is obtained at concentration levels of respectively 0.4% and 0.7%. In contrast the LR and BN formulas keep their accuracy for the low credit quality portfolios (BB and lower) even at extremely high levels of concentration. Note that no threshold level is given in table 3 at which the accuracy becomes unacceptable, the users of the SPAs should decide for themselves which level of accuracy has to be attained.²²

²⁰ The MAE is reported as this measure more conveniently expresses the inaccuracies of the LR and BN estimates as deviations in terms of VaR or economic capital.

²¹ A table which shows the unreliability of the LR, more specifically the break-points for the different portfolios is available on request.

²² Relying on the inferential statistics is of little help here, because often statistical tests show that MSE is significantly different from 0, even for CCC portfolios, while these deviations are not economically significant.

Table 3. Concentration and the accuracy of the LR and the BN formulas.

This table reports the relationship between the concentrations on single exposures and the accuracy of the LR and BN formulas. Several simulated portfolios are employed. The initial portfolio is an absolutely homogeneous portfolios in which all borrowers are homogeneous in credit quality (AAA, AA, A, BBB, BB) and loan size distribution (1,000). From this initial portfolio the size of an exposure is gradually increased. At each step we measure the impact of the concentration on the skewness, the kurtosis and the accuracy of SPAs. $MAE_{99\%}$ reports the mean absolute errors in the upper 1% tail of the credit loss distributions. As the benchmark we use estimates derived from the Monte Carlo simulations with 100,000 runs.

	Max Exp (%)	Skewness	Kurtosis	MAE_{99%} LR	MAE_{99%} BN
AAA	0.1	2.2	6.9	0.041	0.042
	0.2	2.4	11.3	0.243	0.256
	0.3	3.1	28.2	0.604	0.625
	0.4	4.2	67.6	0.779	1.107
	0.5	6.0	138.0	1.118	1.487
AA	0.1	2.0	6.1	0.021	0.023
	0.2	2.1	6.6	0.014	0.015
	0.3	2.2	8.4	0.091	0.097
	0.4	2.5	12.7	0.237	0.247
	0.5	2.9	20.6	0.398	0.411
A	0.1	2.0	6.0	0.010	0.011
	0.3	2.1	6.5	0.011	0.013
	0.5	2.3	9.0	0.101	0.106
	0.7	2.7	15.6	0.254	0.262
	1.0	3.9	38.4	0.567	0.580
BBB	1.0	2.1	6.5	0.010	0.008
	2.0	2.6	11.3	0.048	0.052
	3.0	3.6	23.5	0.165	0.170
	4.1	4.9	40.4	0.291	0.297
	5.0	6.1	59.9	0.373	0.381
BB	1.0	2.0	6.0	0.007	0.009
	3.0	2.2	7.2	0.022	0.020
	5.0	2.7	11.0	0.015	0.018
B	14	2.2	7.1	0.003	0.007
CCC	15	2	6.1	0.005	0.001

Although, the setup of the accuracy assessment of the LR and BN formulas to concentrations may be rather restrictive with respect to real-life portfolios (e.g. the simulated portfolios only consist of exposures of the same credit quality and the portfolios are only concentrated to one single exposure) and therefore may overestimate the sensitivity of the accuracy of the LR and BN estimates to concentrations. However, the results clearly show that concentrations to single exposures of good credit quality can have a significant impact on the skewness and kurtosis of the loss distribution and, hence, importantly affect the accuracy of the LR and BN formulas. Moreover, note that in general financial institutions take into account the credit quality of the counterparty to determine their maximum exposures (or concentration limits). Hence, in general institutions are most significantly exposed or concentrated to high credit quality debt (portfolios).

6. Conclusions

CR+ is an influential method for credit risk modeling that provides an analytical solution to derive the credit loss distribution. The model attracted a lot of attention from practitioners and scholars due to its analytical tractability, which eliminates the need for time-consuming Monte Carlo simulations. However, the original Panjer recursion of the CR+ model is found to be unstable and therefore often yields inaccurate results for the tail of the loss distribution of the credit portfolios. To solve this instability issue alternative solutions are proposed of which the technique of the SPA is shown to be extremely fast, more robust in practical applications and accurate for large portfolios, regardless of the complexity of the risk factor structure. In the statistical literature, though, certain authors argued that for some distributions the commonly employed standard normal SPAs are not always reliable and that sometimes the relative errors of the estimates are quite large.

In this paper we find that the accuracy of the LR and BN estimates (two popular SPA formulas among practitioners) in the CR+ framework strongly depends on the sensitivity of the higher order standardized cumulants and, thus, the relative error ($A(\theta)$) to variations of the saddlepoint (θ) across its valid range. For five credit portfolios compiled from a large-real life portfolio and for a multitude of simulated portfolios, we show that both formulas give highly inaccurate and unreliable estimates when the third and fourth standardized cumulants at $\theta = 0$ deviate

significantly from their theoretical limits. For distributions of portfolios with relatively low third and fourth standardized cumulants we confirm that the LR and BN formulas give highly accurate estimates. Hence, the third and fourth moments are key diagnostics for a preliminary assessment of the accuracy of the standard normal SPA estimates.

In addition, as the skewness and kurtosis of the loss distribution of high credit quality portfolios are highly sensitive to concentrations on single exposures, the accuracy and reliability of the LR and BN formulas disappear at even relatively low levels of concentration. In this respect, note that financial institutions in general take into account the credit quality of the counterparty to determine their maximum exposures (i.e. concentration limits). This makes that the institutions in general are most significantly exposed or concentrated to the high credit quality debt (portfolios). Therefore, we conclude by warning against applying the LR and BN SPAs without first checking the third and fourth moments of the loss distribution.

At the same time we corroborate the results of Gordy (2002) that for the simple one risk factor CR+ specification, the original Panjer recursion algorithm is accurate and robust, even for highly heterogeneous credit portfolios with large skewness and kurtosis of the loss distribution.

We should stress that the SPA technique is a valid methodology. It overcomes some important restrictions imposed on CR+. For instance, the Panjer recursion algorithm applies only to discrete distributions, which makes it difficult to extend the model to allow for recovery uncertainty. In contrast, the recovery uncertainty is easily incorporated when using the SPA technique. Hence, the method of SPA should not be discarded. As has been briefly discussed in the paper previous statistical research shows that no simple adjustment of the employed SPA formulas will solve the identified inaccuracy issue. Therefore, different SPA formulas may have to be regarded for different loss distributions. Further research though has to identify viable SPA formulas for the different commonly encountered loss distributions.

References

- Barndorff-Nielsen, O. and Cox, D., (1979). Edgeworth and saddlepoint approximations with statistical applications. *Journal of the Royal Statistical Society*; Series B (Methodological), Vol. 41, No. 3, 279-312.
- Beran, J. and Ocker, D., (2003). Small Sample Asymptotics for Credit Risk Portfolios. <http://www.financerisks.com/Credit%20Risk.htm>
- Booth, J.G., Wood A.T.A., (1995). An example in which the Lugannani-Rice saddlepoint formula fails. *Statistics and Probability Letters*, 23, 53-61.
- Credit Suisse Financial Products, (1997). CreditRisk+: A CreditRisk Management Framework, London.
- Daniels, H.E. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, 25(4), 631-650.
- Daniels, H.E. (1980). Exact saddlepoint approximations. *Biometrika*, 67(1), 59-63.
- Feuerverger, A. and Wong, A.C.M. (2000). Computation of Value-at-Risk for Nonlinear Portfolios. *Journal of Risk*, 3(1), 37-55.
- Gordy, M.B., (2002). Saddlepoint Approximation of CreditRisk+. *Journal of Banking and Finance*, 26 (7), 1335-1353.
- Gordy, M.B., (2004) "Saddlepoint approximation," in: Gundlach, V.M., and Lehrbass, F.B. (Eds.), *CreditRisk+ in the Banking Industry*, Heidelberg, Springer
- Haaf, H., Reiß, O. and Schoenmakers, J., (2003). "Numerically stable computation of CreditRisk+". Weierstrass Institute for Applied Analysis and Stochastics, Preprint N. 846, Berlin.
- Jensen, J.L., (1995). Saddlepoint approximations, Oxford: Clarendon Press.
- Kurth, A., Taylor, H., and Wagner, A., (2002). An extended analytical approach to credit risk management. *Economic Notes by Banca Monte dei Paschi di Siena SpA*, 31(2), 237-253.
- Martin, R., Thompson K., and Browne C. (2001). Taking to the Saddle. *Risk*, 14(6).91-94.
- Merino, S. and Nyfeler, M., (2003). Numerical Techniques for Determining Portfolio Credit Risk CreditRisk+ in the Banking Industry, edited by Gundlach M. and Lehrbass F.
- Pritsker, M., (1997). Evaluating Value-at-Risk Methodologies: Accuracy versus Computational Time. *Journal of Financial Services Research*, 12(2-3), 201-242.
- Studer, M., (2001). Stochastic Taylor Expansions and Saddlepoint Approximations for Risk Management. PhD dissertation, ETH Zurich.
- Wilde, T., (2000). CreditRisk+. Credit Derivatives and Credit Linked notes, second edition.
- Wood, A.T.A., Booth, J.G., and Butler, R.W., (1993). Saddlepoint Approximations to the CDF of Some Statistics with Nonnormal Limit Distributions. *Journal of the American Statistical Association*, 88(422), 680-686.

Appendix A. The third and the fourth standardized cumulants in CR+.

In this section we show that the third and the fourth standardized cumulants are dependent on θ and derive analytically their limits. This proof is based on the derivations of cgf for multi-factor CR+ made by Gordy (2002). In order to make the proof simpler we make the following assumptions:

1. The credit portfolio is absolutely homogeneous, i.e. all default exposures are of equal size. In such simple portfolio only one exposure band is possible;
2. The standardized loss unit is equal to the size of the default exposure and thus in the event of default by obligor i , there is a fixed loss $v_i = 1$.
3. All obligors have identical factor loadings to systematic risk factors $w_{ik} = w = 1/K$ for $k = \{0, 1, \dots, K\}$, where K is the number of systematic risk factors. There is no idiosyncratic risk factor.
4. All systematic risk factors have identical variance $1/\tau=1$

In this case $\mu_k = \sum_i w_{ik} \bar{p}_i = \sum_i w \bar{p}_i = \mu$, $\delta_k = \mu_k / (\tau_k + \mu_k) = \delta$,
 $\Omega_k(\theta) = \sum_i w_{ik} \bar{p}_i \exp(v_i \theta) = \sum_i w \bar{p}_i \exp(\theta) = \Omega(\theta)$.

The cgf $K(\theta)$ for CR+ is

$$K(\theta) = (\Omega_0(\theta) - \mu_0) + \sum_{k=1}^K \tau_k \log \left(\frac{\mu_k (1 - \delta_k)}{\mu_k - \delta_k \Omega_k(\theta)} \right) = \psi_0(\theta) + \sum_{k=1}^K \psi_k(\theta).$$

Under the mentioned above assumptions without the idiosyncratic risk factor, the cgf can be written as

$$K(\theta) = \sum_{k=1}^K \tau \log \left(\frac{\mu(1 - \delta)}{\mu - \delta \Omega(\theta)} \right) = \sum_{k=1}^K \psi(\theta) = K\psi(\theta)$$

As in Gordy (2002) we denote D the differential operator, which means that $D^j f(x)$ is the j th derivative of f with respect to x . The j th derivative of $\Omega(\theta)$ is given by the following equation:

$$D^j \Omega(\theta) = \sum_i w \bar{p}_i \exp(\theta) = \Omega(\theta) \quad \forall j \geq 0.$$

The first derivative of $\psi(\theta)$ is:

$$\psi'(\theta) = \tau \left(\frac{\delta D\Omega(\theta)}{\mu - \delta \Omega(\theta)} \right)$$

The expression in the parenthesis is generalised as

$$V_j(\theta) = \frac{\delta \Omega(\theta)}{\mu - \delta \Omega(\theta)} = V_{j+1}(\theta) = V(\theta)$$

Thus the derivatives of $\psi(\theta)$ can be generated recursively by

$$\psi_k'(\theta) = \psi'(\theta) = \tau V_{1,k}(\theta) = \tau V(\theta)$$

$$\psi_k''(\theta) = \psi''(\theta) = \tau [V_{2,k}(\theta) + V_{1,k}(\theta)^2] = \tau [V(\theta) + V(\theta)^2]$$

$$\psi_k'''(\theta) = \psi'''(\theta) = \tau [V_{3,k}(\theta) + 3V_{2,k}(\theta)V_{1,k}(\theta) + 2V_{1,k}(\theta)^3] = \tau [V(\theta) + 3V(\theta)^2 + 2V(\theta)^3]$$

$$\begin{aligned} \psi_k''''(\theta) &= \psi''''(\theta) = \tau [V_{4,k}(\theta) + 4V_{3,k}(\theta)V_{1,k}(\theta) + 3V_{2,k}(\theta)^2 + 12V_{2,k}(\theta)V_{1,k}(\theta)^2 + 6V_{1,k}(\theta)^4] \\ &= \tau [V(\theta) + 7V(\theta)^2 + 12V(\theta)^3 + 6V(\theta)^4] \end{aligned}$$

Derivatives of the cgf is given by

$$D^j K(\theta) = K\psi^j(\theta)$$

Then the third standardized cumulant is

$$\begin{aligned} \zeta_3(\theta) &= \frac{\kappa_3(\theta)}{\kappa_2(\theta)^{3/2}} = \frac{K'''(\theta)}{(K''(\theta))^{3/2}} = \frac{K\psi'''(\theta)}{[K\psi''(\theta)]^{3/2}} = \frac{K\tau [V(\theta) + 3V(\theta)^2 + 2V(\theta)^3]}{[K\tau (V(\theta) + V(\theta)^2)]^{3/2}} = \\ &= \frac{1 + 2V(\theta)}{\sqrt{K\tau (V(\theta) + V(\theta)^2)}} = \frac{1}{\sqrt{K\tau (V(\theta) + V(\theta)^2)}} + \frac{2}{\sqrt{K\tau \left(\frac{1}{V(\theta)} + 1 \right)}} \end{aligned}$$

The fourth standardized cumulant is

$$\begin{aligned} \zeta_4(\theta) &= \frac{\kappa_4(\theta)}{\kappa_2^2(\theta)} = \frac{K''''(\theta)}{(K''(\theta))^2} = \frac{K\tau (V(\theta) + 7V(\theta)^2 + 12V(\theta)^3 + 6V(\theta)^4)}{[K\tau (V(\theta) + V(\theta)^2)]^2} = \\ &= \frac{K\tau [V(\theta) + V(\theta)^2]}{[K\tau (V(\theta) + V(\theta)^2)]^2} + \frac{6K\tau [V(\theta) + V(\theta)^2]^2}{K^2\tau^2 [V(\theta) + V(\theta)^2]^2} = \frac{1}{K\tau} \left[\frac{1}{(V(\theta) + V(\theta)^2)} + 6 \right] \end{aligned}$$

The relative error is

$$\frac{\zeta_4(\theta)}{8} - \frac{5\zeta_3(\theta)^2}{24} = -\frac{1}{12K\tau} \left[\frac{1}{V(\theta) + V(\theta)^2} + 1 \right]$$

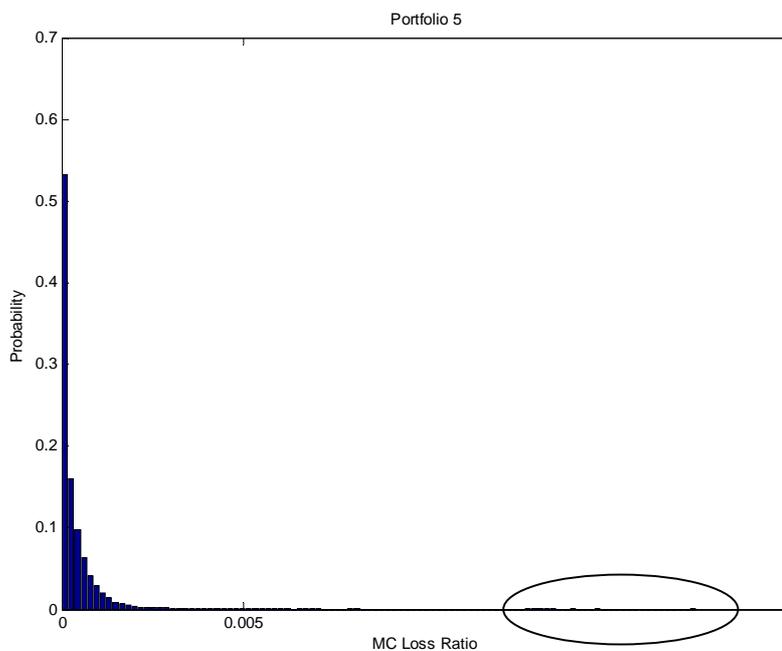
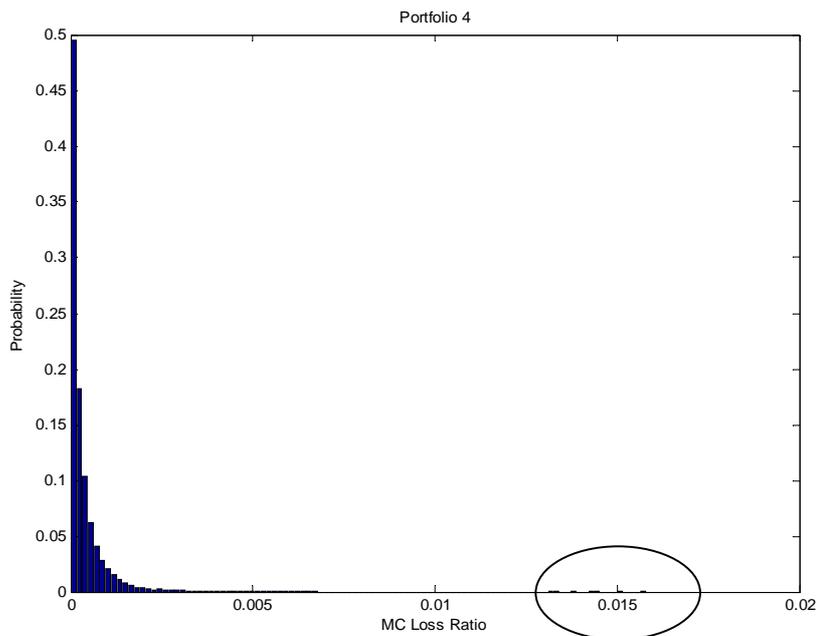
Let $\hat{\theta}$ be the value at which the denominator in $V(\theta)$ is equal to zero. Therefore the denominator in $V(\theta)$ is positive and decreasing for all $\theta < \hat{\theta}$ and equal to zero at $\hat{\theta}$. The numerator in $V(\theta)$ is positive and increasing, so $V(\theta)$ is positive and increasing for all $\theta < \hat{\theta}$. As $\theta < \hat{\theta}$ then

$$V(\theta) \rightarrow \infty \quad \zeta_3(\theta) \rightarrow \frac{2}{\sqrt{K\tau}}, \quad \zeta_4(\theta) \rightarrow \frac{6}{K\tau}, \quad \left(\frac{\zeta_4(\theta)}{8} - \frac{5\zeta_3(\theta)^2}{24} \right) \rightarrow -\frac{1}{12K\tau}$$

In case when $\nu_i > 1$, the dependency of the standardized cumulants and the correction term on θ is more complex, but as we show empirically the standardized cumulants and the correction term have the limits derived above.

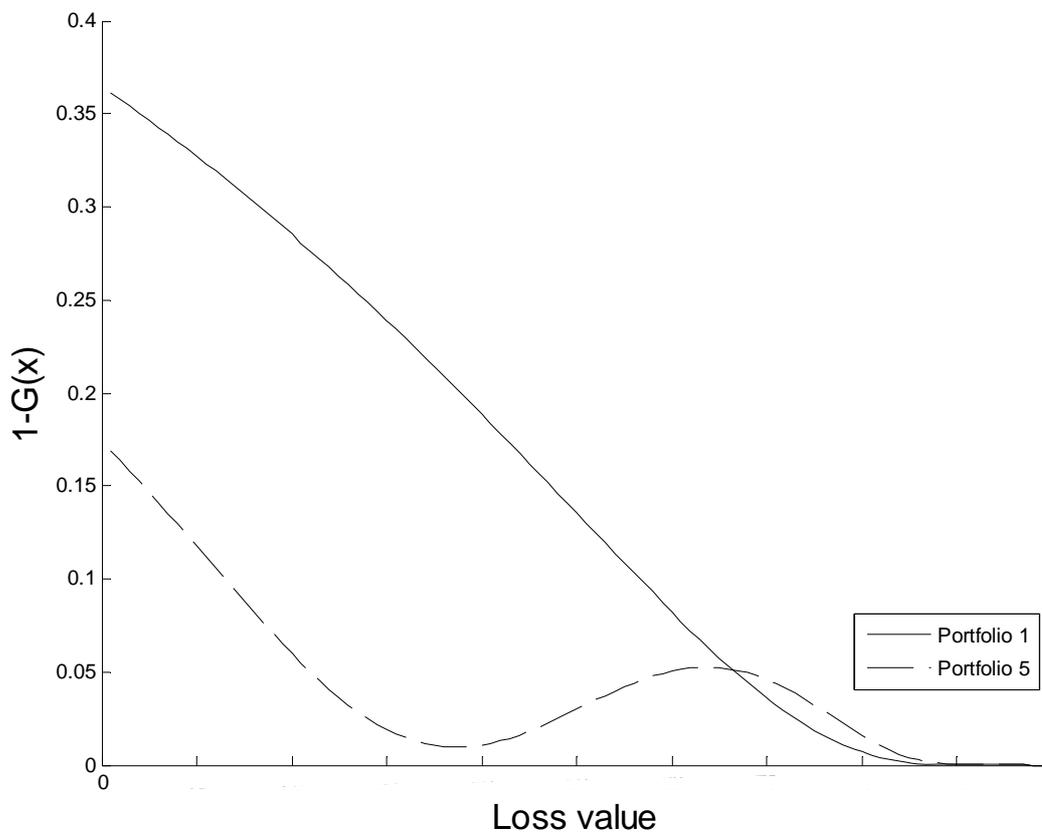
Appendix B: The credit loss distributions for two heterogeneous portfolios

The histograms present simulated loss distributions of credit portfolios 4 and 5, under the assumptions of one systematic risk factor, $LGD=0.3$ and $\sigma^2=1$. In order to ease the comparison between the portfolios, loss is calculated as a percentage to the total book value.



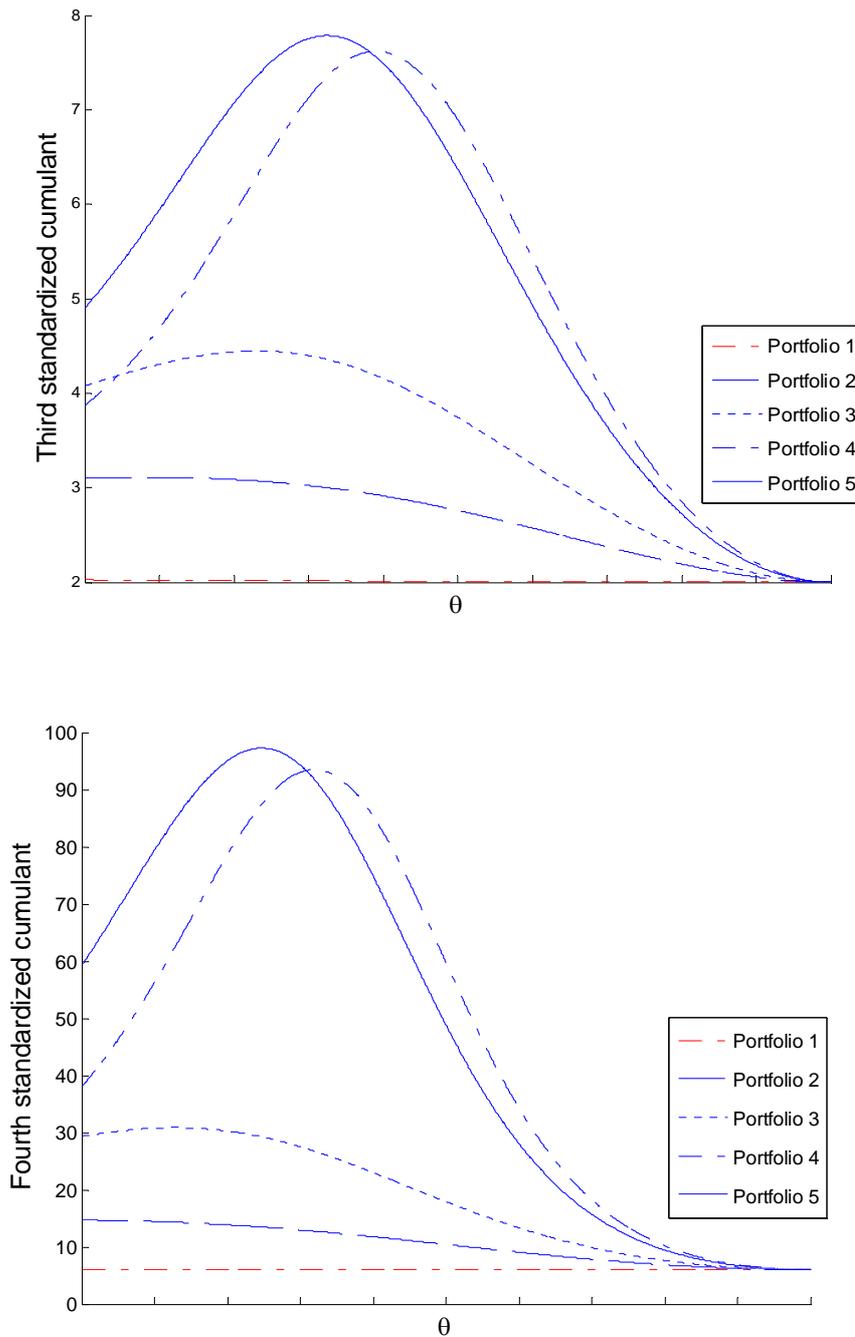
Appendix C. The loss exceedance curves of portfolio 1 and 5.

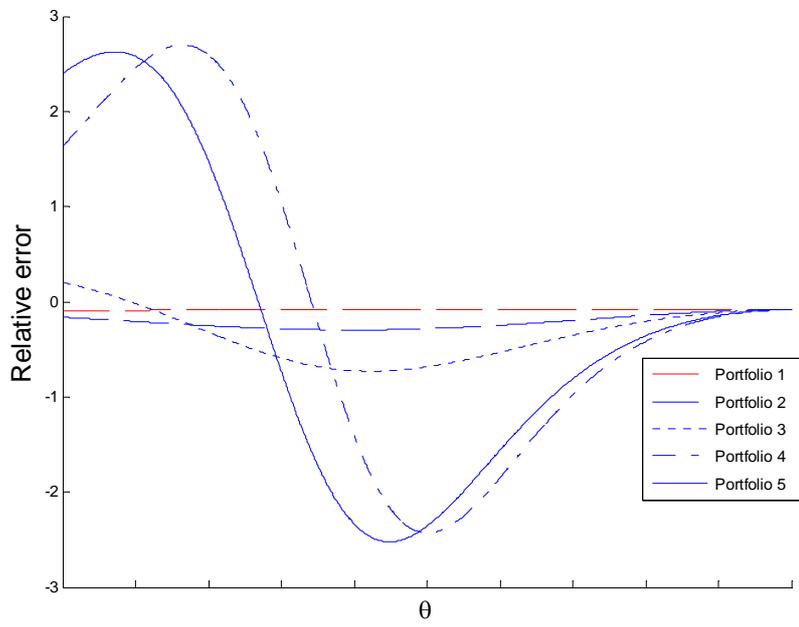
This figure shows the loss exceedance curves produced by the LR formula for the first and the fifth credit portfolios. In order to apply the LR formula a fine grid of 1,000 values in the open interval $(0, \hat{\theta})$ is formed and then at each point in the grid the pairs of losses and corresponding probabilities are computed. In order to ease the comparison between the portfolios, we do not report loss in pecuniary terms. Instead the probabilities are plotted against the points in the grids.



Appendix D. Dependence of $\zeta_3(\theta)$ and $\zeta_4(\theta)$ and the relative error on θ .

In this figure the values of the third, the fourth standardized cumulants and the relative error are plotted against θ for one systematic risk factor CR+ model with LGD=0.3 and $\sigma^2=1$. At first, all exposures were divided in 100 exposure bands and then the standardized cumulants and the relative error were calculated.

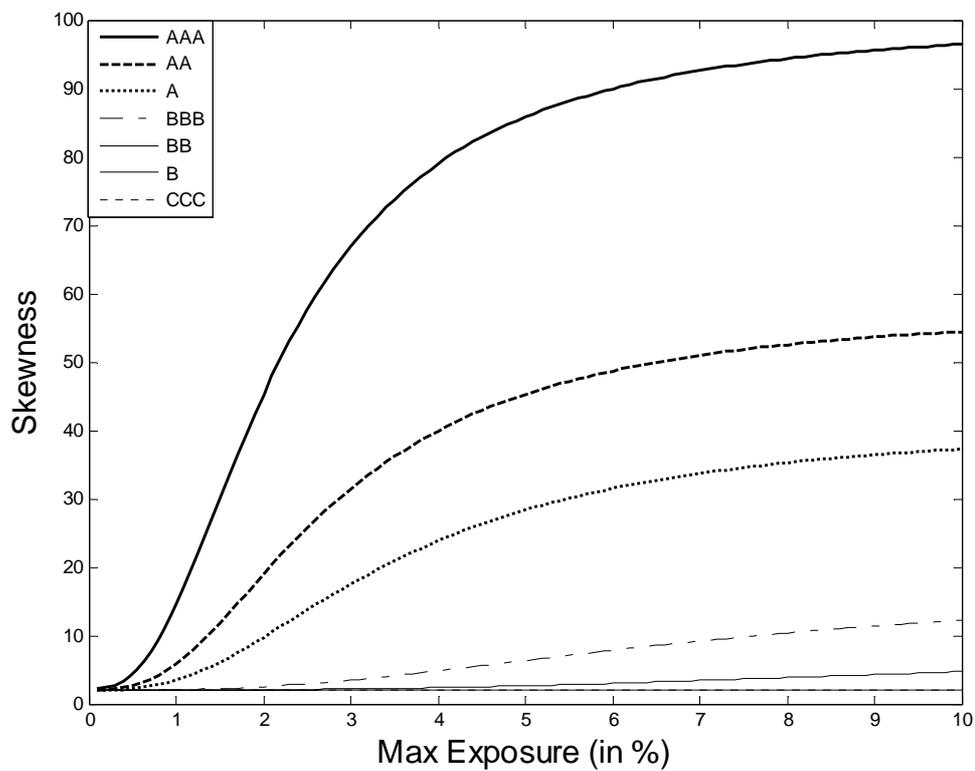




Appendix E. Impact of risk concentration on the skewness and kurtosis of credit loss distribution.

The graphs present the impact of risk concentrations on the skewness and kurtosis of the credit loss distribution under the following assumptions: one systematic risk factor, $\sigma = 1$, LGD=100%. We start from completely homogeneous portfolios in which all borrowers (10,000) are homogeneous in credit quality (AAA, AA, A, BBB, BB, B, CCC) and loan size distribution (1,000). From the initial portfolio we increase gradually the size of only one exposure (expressed relatively to the total portfolio size) and measure its effect on the skewness and kurtosis (at $\theta=0$).

Skewness



Kurtosis

