Global sensitivity analysis of credit risk portfolios

D. Baur, J. Cariboni & F. Campolongo
European Commission, Joint Research Centre, Italy

Abstract

This paper proposes the use of global sensitivity analysis to evaluate latent factor credit risk models. Our claim is that this type of sensitivity analysis is superior to a local approach in providing the risk modeler with a broader picture of the risk contributions of the key elements to a credit risk model. The main finding is that default probabilities and the correlation of the latent variables are considerably more important than the multivariate distribution and hence the copula of the latent variables.

Keywords: credit risk model, latent factor model, uncertainty analysis, global sensitivity analysis.

1 Introduction

Modelling credit risk portfolios is one of the most challenging tasks in finance of current days. In this work we model credit portfolio losses following a latent variable approach. In a latent variable credit risk model, default of an obligor occurs if a latent variable falls below a certain threshold. Dependence between individual default events is caused by dependence between latent variable.

These credit risk models can result in very different loss distributions depending on several factors. Objective of our study is to assess the riskiness of credit risk portfolios and to study the sensitivity of commonly used risk measures with respect to three key input factors of the model.

Previous studies of the sensitivity of latent factor credit risk models to selected model inputs are available in the literature (Frey et al. [1], and Kiesel and Kleinow [2]). However in these works the sensitivity analysis has been performed by following a local approach, i.e. by evaluating the impact of changing one input factor at a time, while our aim is to evaluate the sensitivity of
a latent factor model more thoroughly by performing a global analysis. Results show that the global approach provides the risk manager with more information than the commonly applied local sensitivity analysis and is therefore the proper approach to use in this kind of problem settings.

The following section, 2, describes the credit risk model under investigation. Section 3 introduces the concept of global sensitivity analysis. Section 4 presents the sensitivity analysis experiment with its results and section 5 concludes.

2 The model

This section describes a credit portfolio model based on the latent variables approach. The dependence of $m$ individual obligors is modelled via the dependence of $m$ underlying latent variables $X_1, X_2, \ldots, X_m$. The underlying variables $X_1, X_2, \ldots, X_m$ are assumed to be driven by one common factor $Z$ and idiosyncratic shocks $\epsilon_j$ according to the equation:

$$X_j = \sqrt{a_j}Z + \sqrt{(1-a_j)}\epsilon_j \quad j = 1, \ldots, m$$  \hspace{1cm} (1)

where $a_j$, varying in $(0,1)$, is the factor loading, representing the exposure of obligor $j$ to factor $Z$. The random variables $Z$ and $\epsilon_j$ are assumed to be independently and identically distributed with mean zero and variance one. The terms $\sqrt{a_j}$ and $\sqrt{(1-a_j)}$ are used to ensure a constant variance of $X_j$ independent of the factor loading.

For $j = 1, \ldots, m$ let the random variable $Y_j$ be the default indicator for obligor $j$ taking values in $\{0,1\}$. We interpret 1 as default and 0 as non-default. Let $D_j$ be the cut off point for default of obligor $j$ so that

$$Y_j = 1 \iff X_j \leq D_j$$

The probability of default $\pi_j$ is given by

$$\pi_j = P(Y_j = 1) = P(X_j \leq D_j)$$

while the joint default probability for obligors $i$ and $j$ can then be written as

$$\pi_{ij} = P(Y_i = 1, Y_j = 1) = P(X_i \leq D_i, X_j \leq D_j).$$
In order to focus on the number of joint defaults we assume all losses to be equal to one. The number of joint defaults is obtained by Monte Carlo simulation according to the following steps:

(i) for each obligor the distribution $X_j$ is simulated according to equation (1);

(ii) from the distributions obtained for the $X_j$’s and the input default probabilities the cutoff points of the obligors $D_j$, $j = 1,...,m$, are derived;

(iii) from the obtained multivariate distribution $\mathbf{X}$ a number $I$ of draws is randomly selected and compared with the cutoff vector $\mathbf{D}$ so that the distribution of joint defaults is obtained.

As we are interested in extreme events, we concentrate on the upper tail of the distribution and focus on its 95 and the 99.5 quantiles. In particular we study the uncertainty in these variables (uncertainty analysis) and then their sensitivity to the following model input factors:

1. the distribution of the factor $Z$, which determines the (marginal) distributions of the latent variables $X_j$ and their dependence structure;
2. the degree of correlation among the latent variables, represented by the factor loadings;
3. the probabilities of default, representing the rating of obligors in the portfolio.

Note that, although the copula is not an explicit element of this model, changing the distribution of $X_j$ or changing the copula are two alternative ways to act on the multivariate distribution.

### 3 On uncertainty and sensitivity analysis

Each mathematical model incorporates several input factors which characterize the real process being modelled. These inputs are usually subject to many sources of uncertainty that consequently produces an uncertainty in the model output. Thus, an understanding of the sensitivity of the model outputs to the uncertainty in the input values is necessary in order to develop confidence in the model and its predictions.

Uncertainty analysis quantifies the uncertainty that arises in the model output due to the uncertainty in the model input factors. Results of uncertainty analysis can be used to establish whether or not, given all the sources of uncertainty, the modeller can place confidence in the model outcomes.

Uncertainty analysis is performed via a Monte Carlo approach. Each of the input factors involved in the analysis is assigned a probability density function (pdf), reflecting its uncertainty, i.e. our imprecise knowledge of the input value.
Then, a $k$-dimensional sample of size $N$, $k$ being the number of input factors, is generated from the pre-selected pdf's through an appropriate sampling design. Finally, the model is evaluated at each sample point to obtain $N$ realizations of the output used to estimate its empirical pdf. The estimation of the output pdf allows to quantifying the uncertainty associated with the model output which is due to the uncertainty in the inputs. Note that results of the uncertainty analysis strongly depend on the pdf’s assigned to the inputs. Furthermore, for an effective analysis it is crucial that the entire range of variation of each factor is explored.

Sensitivity analysis (SA) is the study of how the uncertainty in the output can be apportioned, qualitatively or quantitatively, to its different sources. In general sensitivity analysis is used to order by importance the input factors according to the percentage of the output variation (uncertainty) that they are accounting for. Results of sensitivity analysis can be used to improving reliability of these outcomes. To this end further research is addressed to improve the estimates of the crucial inputs and to reduce the degree of uncertainty in the estimates of their values.

A large number of sensitivity analysis methodologies are available in the literature. The choice of the method to adopt to perform a sensitivity experiment on a model depends on a number of aspects: the properties of the model under study (linearity, additivity, monotonicity), the number of input factors involved in the analysis, the computational time needed to evaluate the model, and, last but not least, the objective of the analysis.

Very often in the literature, SA is conceived as a local measure of the effect of a given input on the output, estimated by computing a derivative. This type of approach, called local sensitivity analysis, is based on sampling designs that vary one input factor at a time, while holding all the others fixed at determined values. Although the local sensitivity approach has some advantages, it has to be underlined that this approach is practicable only when the input factors are allowed small variations around a base value, or when the input-output relationship is assumed to be linear (see Saltelli et al. [3], [4]). In contrast in problem settings such as risk analysis, decision support, environmental appraisal, where the degree of variation of the input factors is material and/or the model is non-linear, the linear sensitivities alone are not likely to provide a reliable measure of sensitivity. In these cases, the use of a global approach allowing all input factors to vary simultaneously and that permits material variations, is mandatory.

A review of global sensitivity analysis methodologies and a more detailed description of their features can be found in Saltelli et al. [3], [4].

In this work we investigate the sensitivity of the portfolio model described in section 2 through a global sensitivity approach, based on Monte Carlo simulation. In our exercise this approach is made necessary as we cannot rely on the assumptions needed to apply a local analysis, i.e. we have no reasons to believe a priori that the model is linear. Furthermore we aim to explore wider ranges of variation for the input factors with respect to those examined in previous works by Frey et al. [1] and Kiesel and Kleinow [2].
Our goal is to determine what input factors in the credit risk models are more responsible for variation in the model outcome and therefore need a better determination. In other words, we are facing the following problem, labelled as Factor Prioritisation, "what factor, once fixed to its true albeit unknown value, would give the greatest reduction in the variance of the output?".

The following global sensitivity measures are estimated for each input factor $i$: $S_i$, which is a measure of the main effect, and $S_{Ti}$, which is a measure of the total effect, i.e. a measure that takes into account also the effect of factor $i$ due to its interactions with other factors (two or more factors are said to interact when their effect on the output cannot be expressed as the sum of their single effects). $S_i$ tells us what percentage of the total output variance is due to the $i$th factor, while $S_{Ti}$ is used to determine what factors are irrelevant and can be fixed at any given value within their range without significantly reduce the output variance ($S_{Ti}=0$ if a sufficient condition to state that a factor is irrelevant). If a model is linear the sum of the main effects is sufficient to explain the total output variance, i.e. $\sum_i S_i = 1$. When a model is nonlinear, which is interactions among inputs are presents, than $\sum_i S_i < 1$, as the portion of output variance due to interactions effects is not captured by the $S_i$. A full description of these measures, including technicalities on how to compute them can be found in Saltelli et al. [3], [4].

It can be demonstrated that the $S_i$ indices are the proper measure to rank factor in order of importance when the problem is that of Factor Prioritisation also in the presence of interactions.

4 The sensitivity analysis experiment

In the work of Frey et al. [1] and Kiesel and Kleinow [2] sensitivity analysis experiments were performed on latent variable credit risk models by following a local approach, i.e. by evaluating the effect of changing one input factor at a time. The main finding of their work can be summarised in the deduction that the copula, determining the distribution of $X_j$, considerably contributes to determining the number of joint defaults for a given quantile of the loss distribution.

Here we aim to perform a sensitivity experiment following a global approach, which allows for material variation of the input factors values and which does not require any preliminary assumptions on the model such as linearity.

Our experiment starts with the choice of the output variables of interest, as different choices can lead to very different results. In this work the output variables of interest are the higher quantiles of the distribution of joint defaults: the 95 and the 99.5 quantiles.
As second step we choose the input factors and assign the correspondent pdf. These are:

(i) The dependence structure among the latent factors $X_j$, defined through the choice of the distribution for the common factor $Z$. This is modelled by a trigger factor which may assume three possible values corresponding to the following distributions for $Z$: a Gaussian distribution, a t-distribution with 10 or 4 degrees of freedom.

(ii) The degree of correlation among the obligors, represented by a $m$-dimensional vector of factor loadings $a_j$. A trigger factor is defined to choose among 5 possible determinations of the vector of factor loadings. These 5 determinations are generated a priori to represent 5 different correlation structure among the obligors:

- very low correlation, the $a_j$ are sampled from uniform distributions on [0-0,15];
- low correlation, the $a_j$ are sampled from uniform distributions on [0,15-0,30];
- medium-low correlation, the $a_j$ are sampled from uniform distributions on [0,30-0,45];
- medium correlation, the $a_j$ are sampled from uniform distributions on [0,45-0,60];
- medium-high correlation, the $a_j$ are sampled from uniform distributions on [0,60-0,75].

(iii) The rating portfolio composition, represented by a $m$-dimensional vector of default probabilities $\pi_j$. A trigger factor is used to sample among nine possible determinations of the vector of default probabilities. These nine determinations are randomly generated a priori under the assumption that the $\pi_j$ are uniformly distributed within a fixed range. The lower and upper bounds of the range depend on the obligors rating:

- high rated obligors (AAA class), the $\pi_j$ are sampled from uniform distributions on [0 -0,05];
- medium rated obligors (BBB class), the $\pi_j$ are sampled from uniform distributions on [0,05 -0,10];
- low rated obligors (CCC class), the $\pi_j$ are sampled from uniform distributions on [0,10 -0,15].

Then uncertainty and sensitivity analysis are performed in a Monte Carlo fashion on a portfolio of 1000 obligors ($m=1000$). An input sample of size $N=$
16.384 is generated through an appropriate sampling strategy. Then the model is evaluated at each input sample point to produce $N$ output realizations that are analysed to produce uncertainty and sensitivity analysis results.

Results of the uncertainty analysis are in Table 1, which shows the main descriptive statistics of the output empirical pdf.

Table 1: Basic statistics of the outputs’ empirical pdf.

<table>
<thead>
<tr>
<th></th>
<th>95 quantile</th>
<th>99.5 quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>303</td>
<td>569</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>172</td>
<td>243</td>
</tr>
<tr>
<td>Minimum</td>
<td>52</td>
<td>73</td>
</tr>
<tr>
<td>Maximum</td>
<td>795</td>
<td>999</td>
</tr>
</tbody>
</table>

The obtained statistics point out that the average number of joint defaults is rather different for the two selected quantiles. This underlies the need for an analyst to set the objective at the beginning of the exercise. Shifting the interest from one quantile to another may in fact lead to different results.

The standard deviation values indicate that results are rather volatile, especially those referring to the higher quantile. This stresses the need of a global sensitivity analysis to assess the relative contribution of the various input factors to such a considerable variability.

Results of sensitivity analysis are in Table 2. The two sensitivity measures, $S_i$ and $S_{ij}$, are shown for each of the selected input factors for the two quantiles.

Table 2: The sensitivity measures for the 3 input factors and the 2 outputs.

<table>
<thead>
<tr>
<th></th>
<th>95 quantile</th>
<th>99.5 quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree of correlation.</td>
<td>0.237</td>
<td>0.318</td>
</tr>
<tr>
<td>Dependence structure</td>
<td>0.021</td>
<td>0.035</td>
</tr>
<tr>
<td>Rating portfolio composition</td>
<td>0.672</td>
<td>0.774</td>
</tr>
</tbody>
</table>

The obtained indices show the relative importance of the three factors at different quantiles and allows for the conclusion that more than the 80% of the total variance can be explained by the degree of correlation among obligors and the portfolio rating composition in both cases (the sum of their first order indices is greater than 0.8). The relative importance of the portfolio rating composition and the degree of correlation among the obligors depends on the quantile: at the 95 quantile the composition of the portfolio is accounting for most of the variability in the output and therefore needs to be accurately determined; at the 99.5 quantile it is more relevant to focus on the determination of the degree of correlation among obligors.
The effects of the dependence structure is almost negligible for the 95 quantile ($S_{Ti} = 0.035$) and rather low with respect to the others also at the 99.5 quantile. The latter result implies that the choice of the distribution of the common factor $Z$, which determines the dependence structure among the obligors, is not dramatically affecting the quantiles of interest.

5 Conclusions

This paper has introduced the concept of global sensitivity analysis to evaluate credit risk models. One of our main findings is that in order to draw reliable conclusions it is essential to establish a priori the objective function of the analysis, as focusing on different quantiles of the distribution of the number of joint defaults may lead to different conclusions.

In our framework we have shown that when the focus is on the 99.5 quantile of the distribution of number of joint defaults, modelling the degree of correlation is more effective in reducing the output uncertainty than focusing on the portfolio composition. In contrast, when the interest is on the 95 quantile, more careful must be placed in the credit portfolio construction, since the portfolio composition explains more than the 65% of the unconditional variance. Moreover results makes evident that the dependence structure of the latent variables (i.e. their multivariate distribution) is much less influent than other factors when the aim is that of reducing the output variance.

References