MODELLING CREDIT RISK: THE LOSS DISTRIBUTION OF A LOAN PORTFOLIO

Guillermo Magnou, FRM¹

May 2018

Abstract

The aim of this work is to present a methodology that allows in a simple way to compute the regulatory capital for credit risk. The Vasicek model is a popular one-factor model that derives the limiting form of the portfolio loss. This model will allow calculating different risk measures such as, for example, the expected loss (EL), the value at risk (VaR) and the Expected Shortfall (ES). For this study, three different portfolios were proposed: the first was a homogeneous portfolio that had the same weighting among all loans, then a portfolio with unequal weights was considered and finally a mixed portfolio with different weights and different probabilities of default was used. Monte Carlo simulation with 100,000 scenarios served as our benchmark. It was observed that the Vasicek model correctly estimates the results of the homogeneous portfolio. On the other hand, when the portfolio is not homogeneous (portfolio unequal weights and mixed) the Vasicek model correctly estimates the mean (Expected Losses) but underestimates the Value at Risk and the Expected Shortfall. This is because the approximation of the Vasicek model is good on average but not at the extremes.

Keywords: Portfolio Loan, Credit Risk, Loss Distribution, Vasicek Model, Risk Measures, Expected Loss, Value at Risk, Expected Shortfall.

1. INTRODUCTION

Credit risk concerns the risk of loss arising from an obligor inability to honor its obligations. Among other sources of risk, it is the most important one that financial institutions have to deal due to the large exposures concentrated in the portfolios. Because of this, financial institutions must quantify credit risk at portfolio level. For this purpose several risk measures based on the portfolio loss distributions will be presented. The expected and unexpected loss, are defined as the expectation and standard deviation, respectively, of the portfolio loss variable. Further risk measures are the Value-at-Risk (VaR) and the Expected Shortfall (ES) which will be presented in the next section. All of these risk measures have a lot of advantages as, for example, the aggregation from a single position to the whole portfolio. Moreover, diversification effects and netting can be reflected and the loss distributions are comparable across portfolios.

From a regulator’s perspective a clear understanding of the techniques commonly used would enhance supervisory oversight of financial institutions. The motivation to develop credit risk models stemmed from the need to develop quantitative estimates of the amount of economic capital needed to support a financial institution’s risk taking activities. Minimum capital requirements have been coordinated internationally since the Basel Accord of 1998 published by the Basel Committee on Banking Supervision housed at the Bank for International Settlements (BIS).

When estimating the amount of economic capital needed to support their credit risk activities,

¹ Master Mathematical Engineering, University of the Republic Uruguay. Chef Risk Officer (CRO) in Microfinanzas del Uruguay S.A. Mail: guillermo.magnou@microfin.com.uy.
financial institutions employ an analytical framework that relates the overall required economic capital for credit risk to their portfolio’s probability density function (PDF) of credit losses, also known as loss distribution of a credit portfolio. Figure 1 shows this relationship.

The Basel Committee established three models to determine the capital requirement for credit risk: i) Standard Approach (SA), ii) Foundational Internal-Rating Based Approach (FIRD) and iii) Advanced Internal-Rating Based Approach (AIRB). Figure 2 indicates the different ways to determine the capital required for credit risk and the parameters according to the models.

Today, Uruguayan regulation only recognizes the Standardized Approach (SA), for this reason, it will be proposed to introduce more advanced methodologies. These methodologies are based on the Vasicek model that will be discussed later.

The paper is structured as follows: first, it will present some risk measures definitions. After it will present the Vasicek Models, then it will present the result of the study and end with the conclusion.
2. RISK MEASURES

Financial risk is the prospect for financial loss due to unforeseen changes in underlying risk factors (these factors are those that provide uncertainty in financial results). Financial risks can be classified in different ways, such as market risk (or the risk of loss arising from unexpected changes in market prices or market rates), credit risk (or the risk of loss arising from the failure of a counterparty to make a promised payment), liquidity risk, operational risk (or the risk of loss arising from the failures of internal systems or the people who operate in them) and others (as legal risk, reputational risk) [3].

This section discusses statistical summaries of the loss distribution that quantify the portfolio risk. These summaries will be called as risk measures. First, the risk factor, loss distribution, expected (\( EL \)) and unexpected loss (\( UL \)) will be described. Then the so-called coherence axioms, which are properties that are considered desirable for risk measures, will be presented. Subsequently, two widely used financial risk measures will be discussed: the Value at Risk (\( VaR \)) and the Expected Shortfall (\( ES \)).

Risk Factor, Loss Distribution, Expected and Unexpected Loss

Consider a credit portfolio consisting of \( N \) obligors and let \( V_t \) denote its current value. The portfolio value is assumed to be observable at time \( t \). The portfolio loss over the time interval from \( t \) to \( t + 1 \) is written as

\[
L_{t+1} = -(V_{t+1} - V_t)
\]

Because \( V_{t+1} \) is unknown, \( L_{t+1} \) is random from the perspective of time \( t \). The distribution of \( L_{t+1} \) will be referred to as the loss distribution. The portfolio value \( V_t \) will be modeled by a function of time and a set of \( d \) underlying risk factor. It will be assume that each obligor \( i \) is characterized by three risk factors:

- Probability of default (\( PD \)) is the average percentage of obligor that will default over a one-year period.
- Exposure at default (\( EAD \)) gives an estimate of the amount outstanding if the borrower defaults.
- Loss given default (\( LGD \)) represents the proportion of the exposure (\( EAD \)) that will not be recovered after default.

The most common measure of credit risk is expected loss (\( EL \)), which is the average loss in value of the credit portfolio over a given time period or the lifetime of the credit instrument. Depending on the nature of credit instruments, two alternative ways to estimate the expected loss are by looking at default events only and by looking at the loss in value due to the changes in credit quality or credit eating [13]. Typically, the expected loss of a loan portfolio (e.g. credit card, home and auto loans, personal lending) can be measured as

\[
EL = \sum_{i=1}^{N} PD_i \times LGD_i \times EAD_i
\]

Unlike Expected Loss, the Unexpected Loss (\( UL \)) is not an aggregate of individual loss but rather depends on loss correlations between all loans in the portfolio [2]. The deviation of losses from the EL is usually measured by the standard deviation of the loss variable. The portfolio standard deviation of credit losses can be decomposed into the contribution from each of the individual credit facilities:
\[ UL = \sum_{i=1}^{N} \sigma_i \rho_i \]

where \( \sigma_i \) denotes the stand-alone standard deviation of credit losses, and \( \rho_i \) denotes the correlation between credit losses on the overall portfolio.

**Coherent Measures of Risk**
Artzner et al. (1999) [1] argue that an appropriate measure of risk should satisfy a set of properties termed as the axioms of coherence. Let financial risk be represented by a set \( M \) interpreted as portfolio losses, i.e. \( L \in M \). Risk measures are real-valued functions \( \rho: M \to \mathbb{R} \).

The amount \( \rho(L) \) represents the capital required to cover a position facing a loss \( L \). The risk measure \( \rho \) is coherent if it satisfies the following four axioms:

1. **Monotonicity:** \( L_1 \leq L_2 \Rightarrow \rho(L_1) \leq \rho(L_2) \).
2. **Positive homogeneity:** \( \rho(\lambda L) = \lambda \rho(L), \forall \lambda > 0 \).
3. **Translation invariance:** \( \rho(L + l) = \rho(L) + l, \forall l \in R \).
4. **Subadditivity:** \( \rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2) \).

Monotonicity states that positions that lead to higher loss in every state of the world require more risk capital. Positive homogeneity implies that the capital required to cover a position is proportional to the size of that position. Translation invariance states that if a deterministic amount \( l \) is added to the position, the capital needed to cover \( L \) is changed by precisely that amount. Subadditivity reflects the intuitive property that risk should be reduced or at least not increased by diversification, i.e. the amount of capital needed to cover two combined portfolios should not be greater than the capital needed to cover the portfolios evaluated separately.

**Value at Risk**
Value-at-Risk is defined as the sufficient capital to cover, in most instances, losses from a portfolio over a holding period of a fixed number of days [8]. Assume a random variable \( X \) with continuous distribution function \( F \) models losses on a certain financial portfolio over a certain time horizon. \( VaR_\alpha \) can then be defined as the \( \alpha \)-th quantile of the distribution \( F \):

\[ VaR_\alpha = F^{-1}(1 - \alpha) \]

where \( F^{-1} \) is defined as the inverse of the distribution function \( F \). \( VaR_\alpha \) is the risk measure chosen in the Basel II Accord for the evaluation of capital requirements. For this paper we compute a 0.1% \( VaR \) over a one-year holding period.

However, by definition \( VaR_\alpha \) gives no information about the size of the losses that occur with probability smaller than 1-\( \alpha \), i.e. the measure does not tell how bad it gets if things go wrong [18]. Given these problems with \( VaR_\alpha \), will be present an alternative measure which satisfies this.

**Expected Shortfall**
Another measure of risk is the expected shortfall (ES) or the tail conditional expectation that estimates the potential size of the loss exceeding \( VaR \) [8]. The expected shortfall is defined as the expected size of a loss that exceeds \( VaR_\alpha \):

\[ ES_\alpha = E(X \mid X > VaR_\alpha) \]

Expected Shortfall, as opposed to Value at Risk, is a coherent risk measure in the sense that
satisfies properties of monotonicity, sub-additivity, homogeneity, and translational invariance [8].

3. VASICEK MODEL

To represent the uncertainty about futures events, it specify a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with sample space \(\Omega\), \(\sigma\)-algebra \(\mathcal{F}\), probability measure \(\mathbb{P}\) and with filtration \((\mathcal{F}_t)_{t \geq 0}\) satisfying the usual conditions. It fix a time horizon \(T > 0\), usually \(T\) equals one year.

Following Merton’s approach (1974)[14], Vasicek [20][21][22] assumes that a loan defaults if the value of the borrower’s assets \((A)\) at the loan maturity \(T\) falls below the contractual value \(B\) of its obligations payable. Let \(A_i\) be the value of the \(i\)-th borrower’s assets, described by the process

\[
dA_i = \mu_i A_i \, dt + \sigma_i A_i \, dX_i
\]

where \(\mu_i\) and \(\sigma_i\) are the drift and volatility of the value, and \(X_i\) is a Brownian motion, i.e. a random walk in continuous time in which the change over any finite time period is normally distributed with mean zero and variance equal to the length of the period, and changes in separate time periods are independent of each other. Solving this stochastic equation one obtains the value of the \(i\)-th firm’s assets at time \(T\) as

\[
A_i(T) = e^{A(0) + \mu_i T - \frac{1}{2} \sigma_i^2 T + \sigma_i \sqrt{T} X_i}
\]

The \(i\)-th firm defaults if \(A_i(T) < B_i\), so the probability of such an event is

\[
p_i = \mathbb{P}[A_i(T) < B_i] = \mathbb{P}[X_i < c_i] = \mathcal{N}(c_i)
\]

where

\[
c_i = \frac{\ln B_i - \ln A_i - \mu_i T + \frac{1}{2} \sigma_i^2 T}{\sigma_i \sqrt{T}}
\]

represents the default threshold and \(\mathcal{N}\) is the cumulative normal distribution function. Consider a homogeneous portfolio consisting of \(N\) loans characterized by:

- Equal dollar amount.
- Equal probability of default \(p\).
- Flat correlation coefficient \(\rho\) between the asset values of any two companies.
- The same term \(T\).

Let \(D_i\) be the default indicator of obligor \(i\) taking the following values.

\[
D_i = \begin{cases} 
1, & \text{if obligor } i \text{ is in default} \\
0, & \text{if obligor } i \text{ is not in default}
\end{cases}
\]

Let \(L\) be the portfolio percentage gross loss (before recoveries),
If the events of default on the loans in the portfolio were independent of each other, the portfolio loss distribution would converge, by the central limit theorem, to a normal distribution as the portfolio size increases. Because the defaults are not independent, the conditions of the central limit theorem are not satisfied and \( L \) is not asymptotically normal. However, the distribution of the portfolio loss does converge to a limiting form.

The variables \( \{X_i\}_{i=1}^N \) are jointly standard normal with equal pair-wise correlation \( \rho \), and can be expressed as:

\[
X_i = \sqrt{\rho}S + \sqrt{1 - \rho}Z_i
\]

Where \( S \) and \( Z_i \) are mutually independent standard normal variables. The firm-value of obligor \( i \) is represented by a common, standard normally distributed factor \( S \) component (usually called systematic factor) and an idiosyncratic standard normal noise component \( Z_i \). The systematic risk can be seen as the macro-economic conditions and affects the credit-worthiness of all obligors simultaneously. The idiosyncratic risk represents conditions inherent to each obligor and this is why they are assumed to be independent of each other.

When the systematic risk is known (or fixed), the conditional probability of loss on any one loan is:

\[
p_i(S) = \mathbb{P}[D_i = 1 \mid S = s] = \mathbb{P}[A_i(T) < B_i \mid S = s] \\
= \mathbb{P}[X_i < c_i \mid S = s] = \mathbb{P}\left[\sqrt{\rho}S + \sqrt{1 - \rho}Z_i < c_i \mid S = s\right] \\
= \mathbb{P}\left[Z_i < \frac{c_i - \sqrt{\rho}S}{\sqrt{1 - \rho}} \mid S = s\right] = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(p_i) - S\sqrt{\rho}}{\sqrt{1 - \rho}}\right)
\]

The quantity \( p_i(S) \) provides the loan default probability under the given scenario (S). This can be interpreted as assuming various scenarios for the economy, determining the probability of a given portfolio loss under each scenario, and then weighting each scenario by its likelihood.

Conditional on the value of \( S \), the variable \( D_i \) are independent equally distributed variables with a finite variance. The portfolio loss \( L \) conditional on \( S \) converges, by the law of large numbers, to its expectations \( \mathbb{E}(S) \) as \( N \to \infty \). So the cumulative distribution function of loan losses \( L \) on a very large portfolio is in the limit:

\[
\mathbb{P}[L \leq x] = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(p) - \sqrt{\rho}}{\sqrt{1 - \rho}}\right)
\]

This result is given in Vasicek (1991). The portfolio loss is described by two-parameter distribution with the parameters \( 0 < p, \rho < 1 \).

\[
F(x; p, \rho) = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(p) - \sqrt{\rho}}{\sqrt{1 - \rho}}\right)
\]
The distribution possesses the following symmetry property:

\[ F(x; p, \rho) = 1 - F(1 - x; 1 - p, \rho) \]

The loss distribution has the density:

\[
f(x; p, \rho) = \sqrt{\frac{1}{\rho} e^{-\frac{x}{\rho}} \left( N^{-1}(x) - N^{-1}(p) \right)^2 + \frac{1}{2} \left( N^{-1}(x) \right)^2}
\]

The mean of the distribution is \( \mathbb{E}(L) = p \) and the \( \alpha \)-percentile value of \( L \) is given by:

\[ L_{\alpha} = F(\alpha; 1 - p, 1 - \rho) \]

The portfolio loss distribution is highly skewed and leptokurtic.

The convergence of the portfolio loss distribution to the limiting form above actually holds even for portfolios with unequal weights. Let the portfolio weights be \( w_1, w_2, \ldots, w_N \) with \( \sum w_i = 1 \). The portfolio loss:

\[ L = \sum_{i=1}^{N} w_i D_i \]

Conditional on \( S \) converges to its expectation \( p(S) \) whenever \( \sum w_i^2 \to 0 \). In other words, if the portfolio contains a sufficiently large number of loans without it being dominated by a few loans much larger than the rest, the limiting distribution provides a good approximation for the portfolio loss.
4. EMPIRICAL RESULT

In this section an application of the Vasicek model will be presented, considering three portfolios. The first will be a homogeneous portfolio which will have the same weighting among all the loans, then a portfolio with unequal weights will be considered and finally a mixed portfolio with different weights and different probability of default will be used. The portfolios are as following:

- Portfolio 1 - “Homogenous”.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td># of obligors</td>
<td>10,000</td>
</tr>
<tr>
<td>PD</td>
<td>2,00%</td>
</tr>
</tbody>
</table>

- Portfolio 2 - “Unequal weights”.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>1</th>
<th>10</th>
<th>50</th>
<th>100</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td># of obligors</td>
<td>6,750</td>
<td>50</td>
<td>20</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>PD</td>
<td>2,00%</td>
<td>2,00%</td>
<td>2,00%</td>
<td>2,00%</td>
<td>2,00%</td>
</tr>
</tbody>
</table>

- Portfolio 3 - “Mixed”.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>1</th>
<th>10</th>
<th>50</th>
<th>100</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td># of obligors</td>
<td>6,750</td>
<td>50</td>
<td>20</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>PD</td>
<td>2,25%</td>
<td>2,00%</td>
<td>1,75%</td>
<td>1,50%</td>
<td>0,75%</td>
</tr>
</tbody>
</table>

In all cases it will be used LGD = 50%, $\rho = 9\%$ and the total exposure is $\$ 10,000$. The portfolio 2 and 3 are called lower granularity since the largest obligor has an exposure 150 times larger than the smallest obligor. Exposure concentration is not really significant as the weight of the largest obligor is less than 1,5%.

For each of the portfolios, the expected loss ($EL$), the value at risk $VaR_\alpha$ and the expected shortfall $ES_\alpha$ are calculated. Monte Carlo simulation with 100,000 scenarios serves as our benchmark. The standard deviation and the 95% confidence intervals (CI) are reported along with the point estimates.

Table 1 shows the results of the Expected Loss, the Value at Risk to 99.9% and the Expected Shortfall to 99.9% for the three portfolios. In it, it can be seen that the Vasicek model correctly estimates the results of the homogeneous portfolio; this is corroborated given that the estimates of the Vasicek model are within the confidence interval of the simulated model. On the other hand, when the portfolio is not homogeneous (portfolio 2 and 3) the Vasicek model correctly estimates the mean (Expected Losses) but underestimates the Value at Risk and the Expected Shortfall. This is because the approximation of the Vasicek model is good on average but not at the extremes.
<table>
<thead>
<tr>
<th>Portfolio</th>
<th>EL</th>
<th>VaR(_{99.9%})</th>
<th>ES(_{99.9%})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>593,93</td>
<td>688,90</td>
</tr>
<tr>
<td>Benchmark (s.d.)</td>
<td>99,92 (0,25)</td>
<td>585,50 (7,92)</td>
<td>686,28 (15,99)</td>
</tr>
<tr>
<td>95% CI</td>
<td>[99,44;100,41]</td>
<td>[569,98;601,02]</td>
<td>[654,95;717,62]</td>
</tr>
<tr>
<td>Vasicek</td>
<td>100</td>
<td>593,93</td>
<td>688,90</td>
</tr>
<tr>
<td>Benchmark (s.d.)</td>
<td>99,55 (0,27)</td>
<td>624,50 (9,96)</td>
<td>717,78 (12,84)</td>
</tr>
<tr>
<td>95% CI</td>
<td>[99,02;100,07]</td>
<td>[604,97;644,03]</td>
<td>[689,62;739,95]</td>
</tr>
<tr>
<td>Vasicek</td>
<td>100</td>
<td>589,67</td>
<td>683,57</td>
</tr>
<tr>
<td>Benchmark (s.d.)</td>
<td>99,65 (0,26)</td>
<td>619,00 (8,36)</td>
<td>717,18 (13,38)</td>
</tr>
<tr>
<td>95% CI</td>
<td>[99,13;100,17]</td>
<td>[602,62;635,38]</td>
<td>[690,96;743,40]</td>
</tr>
</tbody>
</table>

Table 1: Result of the risk measures for each portfolio.

5. DISCUSSION

The objective of this paper was to present more sophisticated models to determine the regulatory capital for credit risk of those currently used by the Uruguayan regulation (Standardized Approach). For this reason, the Vasicek model was presented to determine the loss distribution of a credit portfolio and its main properties. The Vasicek model proposed in 1991 followed the approach presented by Merton in 1974. Currently, these types of models have been further developed by solving some of the problems found in this document that occur in practice, for example, the problem of homogeneity of the portfolio. However, Vasicek’s model was fundamental from the regulatory point of view to estimate the amount of economic capital needed to support the credit risk activities of financial institutions.

Also it was discussed statistical summaries of the loss distribution that quantify the portfolio risk. The risk factor, loss distribution, expected (\(EL\)) and unexpected loss (\(UL\)) were described. Then the so-called coherence axioms, which are properties that are considered desirable for risk measures, were presented. Subsequently, two widely used financial risk measures were presented: the Value at Risk (\(VaR\)) and the Expected Shortfall (\(ES\)). All of these risk measures have a lot of advantages as, for example, the aggregation from a single position to the whole portfolio. Moreover, diversification effects and netting can be reflected and the loss distributions are comparable across portfolios.

For this study, three different portfolios were proposed: the first was a homogeneous portfolio that had the same weighting among all loans, then a portfolio with unequal weights was considered and finally a mixed portfolio with different weights and different probabilities of default was used. Monte Carlo simulation with 100,000 scenarios served as our benchmark. The Expected Loss, the Value at Risk to 99.9% and the Expected Shortfall to 99.9% were calculated for the three portfolios. It was observed that the Vasicek model correctly estimates the results of the homogeneous portfolio. On the other hand, when the portfolio is not homogeneous (portfolio unequal weights and mixed) the Vasicek model correctly estimates the mean (Expected Losses) but underestimates the Value at Risk and the Expected Shortfall. This is because the approximation of the Vasicek model is good on average but not at the extremes.
Various extensions for non-homogeneous portfolios have been proposed in literature. The granularity adjustment technique was introduced by Gordy (2003) [9]. Wilde (2001) [23] and Martin and Wilde (2002) [12] have derived a general closed-form expression for the granularity adjustment for portfolio VaR. More specific expressions for a one-factor default-mode Merton-type model have been derived by Pykhtin and Dev (2002) [15]. Emmer and Tasche (2003) [4] have developed an analytical formulation for calculating VaR contributions from individual exposures, and Gordy (2004) [10] has derived a granularity adjustment for ES.

Other authors include in the model a stochastic pattern for the loss given default, for example, Frye-Jacobs (2012) [5], Frye (2000) [6], Pykhtin (2003) [17], Tasche (2004) [19], Giese (2005) [7], and Hillebrand (2006) [11]. Finally, a possible extension of this research is use Multi-factor models, like Pykhtin (2004) [16], instead of one-factor models.

REFERENCES


