International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2024, 14(2), 77-88.

Simulating Credit Loss Distributions: Empirical Versus the Vasicek Model

Natasa Milonas[1], Gary van Vuuren[2]*

[1]Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria, South Africa, [2]Centre for Business Mathematics and Informatics, North-West University, Potchefstroom, 2351, South Africa. *Email: vvgary@hotmail.com

Received: 10 November 2023 **Accepted:** 28 February 2024 **DOI:** https://doi.org/10.32479/ijefi.15698

ABSTRACT

Because credit losses can be substantial, managing credit risk is a focus area of risk measurement and management. It is important for financial institutions to select credit risk models that accurately forecast losses. The Basel Committee on Banking Supervision (BCBS) chose the closed-form single risk factor Vasicek model for regulatory capital calculations. In this article, its forecast accuracy is compared with empirical loss distributions using simulated probabilities of default and losses given default. The effect of altering probabilities of default on asset correlations was analysed and how this affects credit portfolio loss distributions. The robustness of the Vasicek model against five different portfolios with unique compositions was explored: results highlight two key findings. Firstly, the Vasicek model is a good approximation of credit losses for a portfolio that does not contain dominating loans (it is, after all, based on the assumption of large-scale homogeneity). Secondly, the Vasicek model is a good approximation for expected loss (ELs) but lacks accuracy when determining extreme unexpected losses (ULs). Finally, credit capital requirements as a function of two variables are presented which reveals novels ways of viewing these values.

Keywords: Credit risk, Vasicek Distribution, ASRF Model
JEL Classifications: C3, C5, G1, M4

1. INTRODUCTION

Financial institutions are exposed to many different types of risks and without adequate risk management practices these risks could have catastrophic financial implications. The three main risks that a financial institution is exposed to are market risk, credit risk and operational risk (Gregory, 2012). Operational risk is the risk an entity faces as a result of daily activities such as loss due to errors by systems and individuals as well as damages to the entity. Market risk arises from losses due to poor investments, day-to-day price or interest rate movements, or by shocks to the market that alter the price of a financial asset. While the possibility of a party failing to make a contractual payment and resulting in the debt issuer not being paid back in full, is known as credit risk (Gregory, 2012).

Although risk comes in different forms, there is not always a precise separation between them. When operational risk occurs, depending on the severity of the losses incurred because of that risk, the entity may also experience market risk. If the stock is listed, the operational risk may affect the entity's stock price, decreasing the value of their financial assets and thus resulting in the inability of the entity to pay their debt.

The variables which determine credit risk such as probability of default (PD), exposure at default (EAD) and loss given default (LGD) were explored as well as how credit rating agencies determine these variables for each obligor. Further, what an asymptotic single risk factor is and how correlation between obligors and the systematic risk factor are affected by default probabilities are unpacked. The focus is specifically on the distribution of credit losses within a portfolio and how the single risk factor Vasicek model is able to accurately approximate the empirical distribution. Vasicek (2002) provides a closed-form approximation for the distribution of credit losses by making use of average input variables. The Vasicek framework is analysed to determine whether EL and UL are accurately approximated through the simulation of various portfolios. The robustness of the Vasicek framework is explored by investigating homogeneous portfolios, mixed portfolios and how varying probabilities of default affect the shape of the distribution. In addition, some of the shortcomings of the Vasicek framework were investigated and how the loss distribution is affected by the inclusion of a few large loans within the portfolio. Finally, the effects of varying risk factors have on the capital requirements determined by the Vasicek model are explored and presented.

The remainder of this article proceeds as follows: Section 2 presents a contemporary literature review, and Section 3 sets out the data and methodology. Section 4 presents the results and Section 5 concludes.

2. LITERATURE REVIEW

Credit risk arises due to losses as a result of counterparties failing to fulfill contractual obligation payments (Kliěstik and Cu´g, 2015). Although credit risk is traditionally defined as the risk of losses due to defaults on loans, credit risk may arise from multiple financial activities such as investments, trading on capital markets, equity securities, outstanding invoices, etc. (Kliěstik and Cu´g, 2015). The challenge with credit risk modeling is that defaults are random and unexpected. Default events may result in large losses for financial institutions. Thus, these ELs need to be quantified so that financial institutions can adequately capitalise themselves. Credit risk is managed in multiple different ways by financial institutions. Before the existence of any models, risk experts would use subjective decisions or rating agencies to determine if a potential borrowers request would be accepted (Adamko et al., 2014).

The first mathematical models to model credit risk were introduced by Beaver (1966) and Altman (1968) which predicted an entities probability of failure by taking their financial reports into account. Beaver (1966) used a ratio to determine the probability that a financial institution would fail to pay their contractual obligations. Altman (1968) proposed a Z-score which is a linear combination of five joint weighted business ratios to determine the likelihood of an entity facing bankruptcy in the next 2 years. $Z < 1.8$ suggests a high probability of bankruptcy, while a $Z > 3.0$ would guarantee that the entity would not default. Altman et al. (1977) improved the Z-score model and renamed it the ZETA model which was 70% accurate compared with the Z-score's 36% (Harjans, 2018).

Structural models consider business failures to be endogenous events and are affected by capital structures. These models assume that credit risk events are a result of a change in the value of the company. These models assume that default occurs once the value of the company falls below a given threshold (Kliěstik and Cu´g, 2015). At that point in time, once the threshold is breached, the business is assumed to no longer have sufficient assets to cover all obligations. Merton (1974) introduced the first structural model which was based on the Black and Scholes (1973) option pricing framework. Merton (1974) derived a formulation for risky bonds with a flat risk-free term structure which could be used either evaluate the PD of a company or the credit spread. Vasicek (1984) expanded on Merton's approach by finding a closed-form credit loss distribution for a short-term loan. Initially Vasicek's model was based on a single risk factor but has since been expanded to include multiple risk factors by Pykhtin (2004).

Following structural models, reduced-form models were introduced which are based on the assumption that default is exogenous and use credit spreads as an input to determine probabilities of default (Kliěstik and Cu´g, 2015). These models can be split into two categories: Intensity and credit migration models (Adamko et al., 2014). Intensity models model the randomness of default as a time of jump in a one jump random process (Adamko et al., 2014). Credit migration models transition between credit ratings with the use of a Markov process. Reduced models do not require capital structure information and PD is modelled as a random Poisson process. Jarrow and Turnbull (1995) published the first reduced-form credit model which incorporated credit rating information.

Hybrid models are a combination of structural models and reduced models. After the 1990's, banks and consultants developed credit models whereby potential losses were determined using a predetermined confidence level (Kliěstik and Cu´g, 2015). These types of models were motivated by Basel II and the growing importance of risk management. These models are specifically known as Value-at-Risk (VaR) models.

2.1. Credit Risk

Jorion (2011) defines credit risk as the risk of an economic loss from the failure of a counterparty to fulfill its contractual obligations. Credit risk is split into pre-settlement risk and settlement risk. Pre-settlement risk is the risk over the whole life of the obligation which arises from the counterparty's failure to perform on an obligation such as making a contractual payment resulting in default (Gregory, 2012). Whereas settlement risk is short-term and is the risk carried from the time a payment is made till the time it is received. Settlement risk is caused by different time zones, exchange rates and more. Credit risk traditionally refers only to pre-settlement risk and includes default on loans, bond or derivative transactions.

2.1.1. Risk measures

The distribution of credit risk losses is influenced by the following variables (Jorion, 2011):

- PD: Is the probability of a counterparty defaulting over a specific period.
- Credit Exposure (CE) is the market value of the loan that the counterparty is exposed to, which is also known as EAD.
- Recovery rate (RR) is the percentage of the EAD that the entity can recover. This allows the entity to use the counterparty's assets to decrease the outstanding credit.
- LGD is the percentage of the EAD that an entity loses from the obligor defaulting. This proportion is not recovered from assets and yields the final amount the entity loses due to credit risk. Thus, $LGD = 1\text{-}RR$.

Jorion (2011) defines default as a discrete state that a counterparty may be in, which takes the value one if the counterparty is in default and the value zero if the counterparty is not in default. This process follows a Bernoulli distribution with a PD equivalent to the corresponding PD.

2.1.2. Portfolios and diversification

In modern banking a single loan is far riskier than a group (or portfolio) of loans. This is because a single loan has a single PD and when held alone holds considerable credit risk for a bank as no obligor has a PD of 0. This means that if an obligor has taken a large loan out they will either default or not, resulting in a large financial loss to the bank if default occurs. On the other hand, a portfolio is comprised of multiple loans and allows the bank to diversify their risk. Diversification ensures that a single loan does not account for all of the capital held by an entity in case of default since the risk of default is diversified across all loans (Jorion, 2011). Since different obligors have different PDs, this ensures that only in catastrophic circumstances will all obligors default causing a bank to go bankrupt.

2.1.3. Distributions

The distribution of credit loss for a portfolio of *N* loans from different *i* obligors is (Jorion, 2011),

$$CL = \sum_{i=1}^{N} D_i \cdot EAD_i \cdot LGD_i \tag{1}$$

Where

$$B_i\left(k, PD_i\right) = \begin{cases} 1\, with\ probability\ p_i \\ 0\, with\ probability\, 1 - p_i \end{cases} \tag{2}$$

Where B_i is a Bernoulli distribution. EAD_i, LGD_i are the EAD and LGD for obligor respectively.

Since all variables in (1) may be random and assuming they are independent the expected credit loss may be calculated using,

$$E\left[CL\right] = \sum_{i=1}^{N} E\left[D_i\right] \cdot E\left[EAD_i\right] \cdot E\left[LGD_i\right] \tag{3}$$

Where $E[D_i]=p_i$

ECL or EL is the average loss that a bank may incur this risk is accounted for in the interest rate that the bank would offer a counterparty, which is known as the risk premia (Kiliç, 2007). Banks fund general provisions or loan loss reserves to absorb expected credit losses. Losses are not predictable and since ELs are an average, a second provision must be introduced for when there are spikes in losses.

ULs are the losses that occur over and above the EL due to the volatility in PD and LGD (BCBS, 2005). Since the timing of ULs is not known, risk prima may absorb a portion of the ULs. However, the market will not permit risk prima prices to cover the full ULs and thus a capital buffer is required, this is known as a bank's capital requirement. To determine the amount of capital needed to act as a loss-absorbing function, a confidence interval must be selected. Since a bank cannot protect themselves 100% from losses, a given percentile of loss must be chosen. According to BCBS (2005), a confidence level of 99.9% is sufficient since losses that exceed the ULs are predicted to occur only every 1000 years because the PDs used are annual PDs. The high confidence level is chosen to counteract any under estimations by the PD, LDG, EAD and model uncertainties.

UL is the measure of the amount of deviation from the ELs which is a given percentile of loss less the EL (Jorion, 2011):

$$UL = V\left[CL\right]_{0.999} - E\left[CL\right]. \tag{4}$$

The losses above the given percentile are considered catastrophic losses. Since holding capital for such losses would not be optimal, insurance may be taken by the bank to protect against catastrophic losses. However, the event is extremely unlikely, and it would be extremely expensive to insure against such an event (BCBS, 2005).

2.1.4. Credit loss distribution

Over the years due to variations in losses, being expected and unexpected, banks have been able to model the distribution of losses. BCBS (2005) states that although a portfolio may be consistent over several years, the losses that are actually experienced by a bank vary and are dependent on the severity of events. Figure 1 shows the distribution of credit losses.

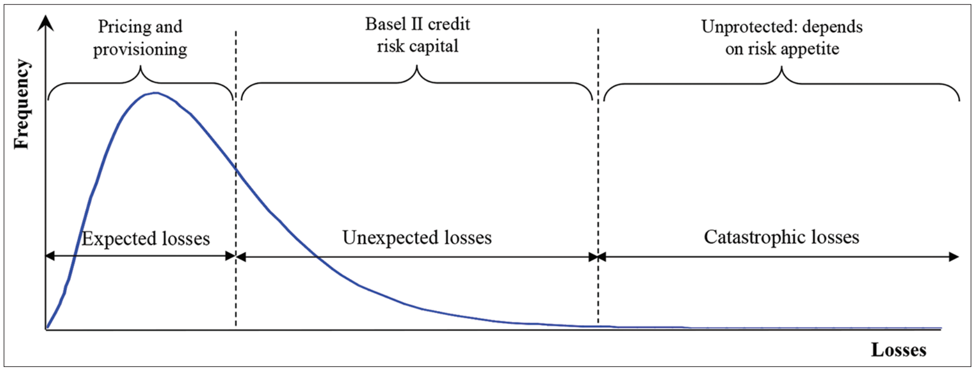
We see that for a given magnitude of losses, the probability of a certain number of losses for a single year is modeled as a probability density function (PDF). The area underneath the curve in Figure 1 is equivalent to 100% of the EAD. The shaded area would indicate a catastrophic/worst case scenario where there are almost 100% losses, and the bank may not have adequate capital to cover these losses. The PDF is skewed to the right, insinuating that there are more smaller losses than larger losses. The highest point of the graph is the mode pf the PDF and is the most frequent potential loss, while the average loss is represented by the vertical dotted line. The losses before and including the vertical dotted line in Figure 1 is the EL, while everything after the dotted line and before the shaded area is the UL. For a confidence interval of 99.9%, the EL plus the UL is known as Credit Value-at-Risk (VaR) (BCBS, 2005).

Banks have an incentive to minimise their capital requirements to free up economic resources for other investments, but they must still be able to meet their debt obligations or suffer insolvency. Thus, the balance between risk and reward of holding capital must be determined by an efficient model.

2.1.5. Effect of maturity adjustments

Banks hold credit portfolios with varying maturities. Long-term credit is riskier than short-term credit because of the possibility of a credit downgrade in that period or a loss of market value of loans. Capital requirements therefore increase with maturity. Maturity effects are negatively correlated with PDs: The lower the PD the higher the maturity effect on the loan and vice versa (BCBS, 2005). It is more likely that borrowers with lower PDs downgrade over time than if these borrower already had higher PDs. The maturity adjustment is thus a function of both maturity and time.

Figure 1: Credit loss distribution (BCBS, 2005)



The derivation of the BCBS maturity adjustment is achieved by applying a mark-to-market credit risk model like the KMV Portfolio Manager. The regression model that is implemented to determine the maturity adjustments allows for the following:

1. Adjustments are linear and increase as maturity (M) increases,
2. The adjustment slope with regards to maturity will decrease as the PD increases,
3. When the maturity is set to 1 year, the function yields the value 1 and the resulting capital requirement is consistent with an Asymptotic Single Risk Factor model.

The capital requirement is simply multiplied by the maturity adjustment:

$$\frac{1+(M-2.5)\cdot f}{1-1.5\cdot f} \tag{5}$$

where $f=(0.11852-0.05478\cdot\log(PD))^2$.

Since only the capital requirement is adjusted, the distribution does not change other than being scaled by the maturity adjustment factor. According to BCBS (2011), capital requirements for credit exposures have a minimum maturity requirement of 1 year.

2.2. Credit Ratings and LGD Scorecards

When determining the PD for a company, credit ratings are used. A credit rating is an *"evaluation of creditworthiness"* issued by a credit rating agency (Jorion, 2011). Moody's Investor Service (2022) defines credit ratings as *"Opinions of relative credit risk of fixed-income obligations with an original maturity of 1 year or more. These ratings address the possibility that a financial obligation will not be honoured as promised. Such ratings reflect both the likelihood of default and any financial loss suffered in the event of default."*

Each rating is based on a set of input rating variables that when coupled with a methodology produce a credit rating for a corporate entity. Credit ratings are depicted by a specific letter or series of letters. These letters carry their own PD (S&P Global, 2022; Fitch, 2022; Moody's Investor Service, 2022).

In addition, credit rating agencies are also able to determine the LGD of an entity. LGD scorecards use qualitative and quantitative factors to estimate the LGD at the exposure level for a corporate entity (S&P Global, 2021).

2.3. ASRF Framework

To understand an Asymptotic Single Risk Factor (ASRF) model, the following definitions are required:

- Portfolio invariance occurs when the capital required for a specific loan is only dependent on the risk of that loan and is independent of the portfolio it is added to.
- Idiosyncratic risk is risk that is unique to a single obligor such as losing their job, being in an accident, etc. while systematic or system-wide risk affects all borrowers to a degree such as the unemployment rate, inflation, GDP, etc.

BCBS (2005) states that a strong influence on the structure of a portfolio model is portfolio invariance. Gordy (2003) showed that only ASRF models are portfolio invariant. By the law of large numbers, ASRF models are derived from ordinary portfolio models. When given a fine-grained portfolio, there are a large amount of relatively small loans and large individual exposures are limited to a small portion of the portfolio (BCBS, 2005). Idiosyncratic risks associated with each obligor are diversified away and only systematic risk affecting multiple obligors has an overall effect on the portfolio losses (Gordy, 2003). ASRF models employ only one risk factor which represents the state of the economy. All obligors are linked to each other by this single factor (Lee et al., 2009).

The ASRF framework calculates the total expected and ULs by using the conditional EL for an exposure given an appropriately conservative value of the single systematic risk factor. The conditional EL is determined by transforming the average PDs into conditional PDs by making use of a supervisory mapping function (BCBS, 2004). Conditional PDs reflect the default rates associated with the specified systematic risk factor. The same systematic risk factor value is used on all the PDs of the loans in the portfolio. It is important to note that LGDs are not adjusted by the systematic risk factor to yield a conditional LGD. Banks are required to use LGDs that reflect economic-downturn conditions where losses are greater than under normal business conditions. According to BCBS (2004), the total economic resources under the ASRF model necessary to cover EL and UL for a given portfolio are equal to the total conditional EL.

2.4. Asset Correlation

Asset correlation determines the relationship between one borrower's asset and another borrower's asset. However, in the ASRF framework the asset correlation is the relationship that is exhibited by the asset value of the borrower and the general state of the economy or the single systematic risk factor. The asset correlation determines how the credit quality of an obligor changes with respect to the state of the economy (Lee et al., 2009).

2.4.1. Basel II asset correlation

The asset value of a borrower is driven by a factor model (Lee et al., 2009):

$$r_i = \sqrt{\rho_i}\lambda + \sqrt{1-\rho_i}\eta_i, \tag{6}$$

Where r_i is the asset return, λ is the single systematic risk factor, ρ_i is the R^2 representing the portion of systematic risk and η represents the idiosyncratic risk factor of obligor i.

Two obligors are correlated with each other since they are both exposed to the same single systematic risk factor. Their degree of dependence on the single systematic risk factor may vary. The correlation between obligor i and obligor j is (Lee et al., 2009):

$$corr\left(r_i, r_j\right) = \sqrt{\rho_i}\sqrt{\rho_j} \tag{7}$$

According to the Advanced-Internal Ratings Based Approach (A-IRB) of BCBS (2001), the asset correlation parameter ρ is a decreasing function of PD:

$$\rho = a \times \frac{1-e^{-c\times PD}}{1-e^{-c}} + b \times \left[1 - \frac{1-e^{-c\times PD}}{1-e^{-c}}\right] \tag{8}$$

Parameters *a, b* and *c* are dependent on the type of borrower. For corporate borrowers and low asset correlation commercial real estate (CRE) $a = 0.12$, $b = 0.24$ and $c = 50$.

BCBS (2001) assumes that average asset correlation decreases with an increase in PD. This means that the higher the PD of an obligor, the lower the effect of the single systematic risk factor on the obligor shown in Figure 2.

The relationship between asset correlation and default probabilities is a decreasing function for all borrower types. Asset correlation reaches a plateau for each type of borrower as PD increases. Corporate borrowers have the highest asset correlation from all borrower types. For the purpose of this research the size of the firm as a factor will be disregarded since Lee et al. (2009) show that when firm size is accounted for, the negative relationship between asset correlation and PD no longer holds.

2.5. Vasicek Model

The Gaussian Asymptotic Single Risk Factor (ASRF) Model for portfolio credit losses was developed by Vasicek (1987) to produce an estimate for a credit loss distribution for a given portfolio where the correlation between obligors is driven by a single risk factor. The Vasicek single factor model for portfolio credit losses has been generalised to include stochastic EADs and LGDs (Kupiec, 2008). The model is able to adapt to accommodate any distribution and correlation assumptions for LGD and EAD to produce a closed-form distribution (Vasicek, 2002).

Vasicek (1987) followed the approach of Merton (1974) and assumed that a loan defaults only if the value of the borrower assets (A) at the loan maturity (T), assumed to be 1 year, falls below the contractual value (C) of its obligations. Let A_i be the ith obligor's assets, then the asset price dynamics are given by (Vasicek, 2002):

$$dA_i = \beta_i A_i dt + \sigma_i A_i dW_i \tag{9}$$

where β_i and σ_i are the drift and volatility respectively and W_i is a Wiener process with

$$E\left(dW_i\right)^2 = dt \tag{10}$$

And

$$D_i = \begin{cases} 1\, if\ obligor\ is\ in\ default\ with\ probability\ p \\ 0\, if\ obligor\ is\ not\ in\ default\ with\ probability\ 1-p \end{cases} \tag{11}$$

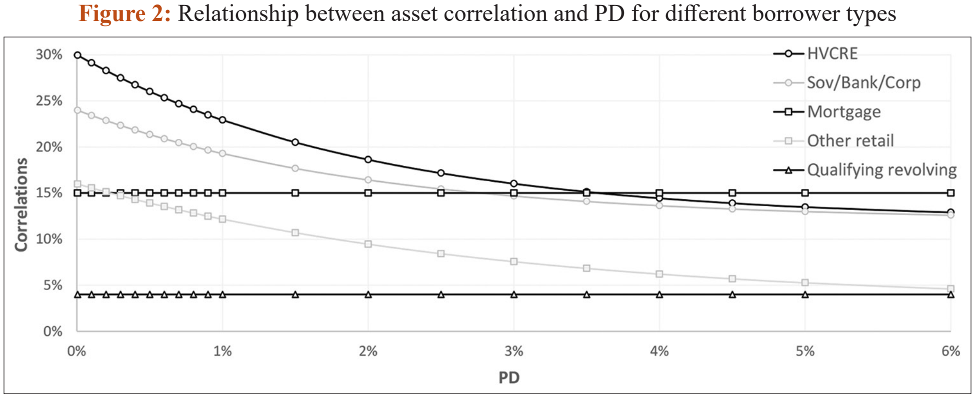
The asset value at maturity, T, is then given by (Vasicek, 2002):

$$\log\left[A_i\left(T\right)\right] = \log\left[A_i\right] + \beta_i T - \frac{1}{2}\sigma_i^2 T + \sigma_i\sqrt{T}W_i \tag{12}$$

so,

$$A_i\left(T\right) = e^{A(0)+\beta_i T - \frac{1}{2}\sigma_i^2 T + \sigma_i\sqrt{T}W_i} \tag{13}$$

Figure 2: Relationship between asset correlation and PD for different borrower types



Vasicek (2002) calculates the PD of the *i*th loan as the probability that the value of the borrower's assets at maturity is less than the value of its contractual obligations at maturity:

$$PD_i = P\left[A_i(T) < C_i(T)\right] = P\left[W_i < z_i\right] = N(z_i) \quad (14)$$

Where

$$z_i = \frac{\log(C_i) - \log(A_i) - \beta_i T + \frac{1}{2}\sigma_i^2 T}{\sigma_i \sqrt{T}} \quad (15)$$

represents the default threshold and N is the cumulative normal distribution function. Consider a portfolio consisting of K loans with the following characteristics (Vasicek, 2002):

1. Equal loan amounts in rands (EAD);
2. Equal PD;
3. Correlation between two companies is ρ;
4. Loans have the same maturity T.

We then have that the gross percentage loss of the portfolio before recoveries (L) is (Vasicek, 2002):

$$L = \frac{1}{K}\sum_{i=1}^{K} D_i \quad (16)$$

Where Di is

$$D_i = 1 \, iff \, W_i \leq \phi^{-1}(p_i) \quad (17)$$

By the central limit theorem, the portfolio loss distribution converges to a normal distribution as the portfolio size increases, given that the events of default in the portfolio are independent (Vasicek, 2002). However, since the defaults are not independent, the conditions of the central limit theorem are not satisfied and L is not asymptotically normal. Vasicek (2002) shows that the distribution of losses converge to a limiting form.

Since the variables W_i in (9), are jointly standard normal with equal pair-wise correlation ρ, they can be represented as (Vasicek, 2002):

$$W_i = \sqrt{\rho}\Lambda + \sqrt{1-\rho}\,\eta_i \quad (18)$$

where, $\sqrt{\rho}\Lambda$ represents the systematic risk and $\sqrt{1-\rho}\,\eta_i$ represents the idiosyncratic risk, with the condition that Λ and η_i are mutually independent standard normal variables. Thus, Λ is the single risk factor that affects all obligors and η_i is an obligor specific risk factor (Vasicek, 2002).

When is fixed, the conditional probability of loss on any one loan is (Vasicek, 2002):

$$p(\Lambda) = \mathbb{P}\left[D_i = 1 \mid \Lambda = \lambda\right]$$
$$= \mathbb{P}\left[A_i(T) < C_i(T) \mid \Lambda = \lambda\right]$$
$$= \mathbb{P}\left[W_i < z_i \mid \Lambda = \lambda\right] \quad (19)$$
$$= \mathbb{P}\left[\sqrt{\rho}\Lambda + \sqrt{1-\rho}\,\eta_i < z_i \mid \Lambda = \lambda\right]$$
$$= \mathbb{P}\left[\eta_i < \frac{z_i - \Lambda\sqrt{\rho}}{\sqrt{1-\rho}} \mid \Lambda = \lambda\right]$$
$$= \mathbb{P}\left(\frac{N^{-1}(p) - \Lambda\sqrt{\rho}}{\sqrt{1-\rho}}\right)$$

Where $p(\Lambda)$ is the loan default probability under the given scenario (Λ) which allows one to scenario based testing on the economy by changing and determining the conditional PD based on the single risk factor (Vasicek, 2002). The portfolio loss, L, conditional on Λ converges, by the law of large numbers, to its expectation $p(\Lambda)$ as $K \to \infty$. Therefore, the cumulative distribution function of loan losses on a very large portfolio is (Vasicek, 2002):

$$P\left[L \leq x\right] = N\left(\frac{\sqrt{1-\rho}\,N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right) \quad (20)$$

Portfolio loss is described by a two-parameter distribution with parameters P > 0 and P < 1 (Vasicek, 2002):

$$F(x; p, \rho) = N\left(\frac{\sqrt{1-\rho}\,N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right) \quad (21)$$

which possesses the symmetry property

$$F(x; p, \rho) = 1 - F(1-x; 1-p, \rho) \quad (22)$$

The density of the loss distribution is (Vasicek, 2002):

$$f(x; p, \rho) = \sqrt{\frac{1-\rho}{\rho}}\, e^{\frac{-1}{2\rho}\left(\sqrt{1-\rho}\,N^{-1}(x) - N^{-1}(p)\right)^2 + \frac{1}{2}\left(N^{-1}(x)\right)^2} \quad (23)$$

with the mean of the distribution being $E(L)$ = p and the α-percentile of L is:

$$L_\alpha = F(\alpha; 1-p, 1-\rho) \quad (24)$$

The convergence of the closed-form to the portfolio loss distribution holds for a portfolio that contains a large amount of loans that are a similar size, without being dominated by several large loans (Vasicek, 2002). Thus, the limiting distribution is a good approximation for the portfolio loss under this assumption.

3. DATA AND METHODOLOGY

To analyse the accuracy of the closed form distribution suggested by Vasicek, the empirical distribution is first simulated using 10 000 simulations with 1 000 loans of which all have a maturity of 1 year. This is done with the following variable assumptions:

1. The range of PD is [0.003%; 24.59%] (S&P Global, 2022);
2. The range of LGD is [0.28%;0.59%] (S&P Global, 2021);
3. The range of EAD is [100;1 000].

Each loan is determined by randomly selecting a number within the given ranges of each variable assumption. In addition, each simulation is assigned a standard normal random variable which acts as the asymptotic single risk factor. Python makes use of the Mersenne Twister method to generate random numbers (Python, 2022). A seed of two is used to produce comparable and consistent random variables across all portfolios each time the simulation is run. The correlation is determined for each loan using (8), suggested by BCBS (2006).

A critical threshold is determined for each loan by taking the inverse cumulative normal distribution function of the respective PD (Genest and Brie, 2013):

$$T_i = \phi^{-1}\left(p_i\right) \tag{25}$$

Where $\phi(\bullet)$ is the cumulative normal distribution.

Each loan determines a jointly standard normal random variable by applying the effect of the ASRF as in (18) denoted by W_i. The default variable (D_i) is defined as in (2). Since Vasicek's framework is based on (Merton, 1974), obligor *i* defaults if their asset value, W_i in this case, goes below the critical threshold. Thus, (Genest and Brie, 2013)

$$D_i = 1 \, iff \, W_i \le \phi^{-1}\left(p_i\right) \tag{26}$$

If $D_i = 1$, $L = EAD_i \cdot LGD_i$, otherwise losses =0. ELs are determined using (3) and ULs from (4).

PD and LGD are determined by averaging the randomly chosen variables in a single simulation with the EAD equivalent to the total EAD in the respective simulation. Thus, instead of having a range of values, there is a single average value for each variable. The correlation is determined using (8), with the only difference from the empirical being that instead of a range of correlations for each loan, a single correlation for an entire loan book simulation is used. The Vasicek distribution is determined by (23).

The Vasicek EL is:

$$EL_{Vasicek} = PD_{Vasicek} \cdot LGD_{Vasicek} \tag{27}$$

while the UL is:

$$UL_{Vasicek} = LGD_{Vasicek} \cdot \frac{ZPD_{Vasicek}}{\sqrt{1-\rho_{Vasicek}}} + \sqrt{\frac{\rho_{Vasicek}}{1-\rho_{Vasicek}}} N\left(0.999\right) \tag{28}$$

where *N* is the standard normal cumulative density distribution.

To have comparable distributions, the Vasicek distribution is scaled so that the highest point is equivalent to the highest point in the empirical distribution.

4. RESULTS AND ANALYSIS OF EMPIRICAL VERSUS VASICEK RISK MEASURES

The effects of altering a portfolio's composition and how this affects the Vasicek distribution's risk measures were analysed. The first portfolio comprised equal PDs, LGDs and EADs while the second comprised constant PDs with varying EADs and LGDs. The purpose of this portfolio is to investigate the effect three different PDs will have on the correlation and distribution of the credit loss portfolio. The third comprised a range of PDs, LGDs and EADs. The final portfolio was a mixed portfolio with the inclusion of three large loans, i.e., large EADs, which permit the exploration of the robustness of the Vasicek framework (Table 1).

The UL will also be assumed to be the capital requirement. The credit VaR (C-VaR) on all figures represents the 99.9-percentile of losses. The 99.9-percentile was selected as required by BCBS (2006).

4.1. Portfolio 1

We first analyse a homogeneous portfolio which consist of constant PDs, LGDs and EADs. The empirical and Vasicek distributions for this portfolio are depicted in Figure 3a.

In Figure 3a, losses are lognormally distributed, and the Vasicek distribution is a good fit for the empirical distribution. Both EL and UL computed by the Vasicek model are good approximations for the empirical risk measures. A comparison of the different risk measures is given in Table 2.

Portfolio 1 consists of 1 000 loans where each loan has an EAD of 500, which results in a total EAD of 500. From Table 2, it is evident that the EL determined by the Vasicek model underestimates the empirically observed EL by 0.053%. Thus, the empirical EL is approximately 265 more than the Vasicek EL. The 99.9% C-VaR under the Vasicek model is also underestimated by 0.312% compared to the empirically observed 99.9% C-VaR, which translates to a difference of approximately 1 560. The UL computed by the Vasicek model once again underestimates the empirically observed UL by 0.259%, which is expected since the UL is defined as the 99.9% C-VaR less the EL.

When the portfolio is homogeneous, the Vasicek model is a good approximation, but it underestimates the empirical values which may result in a bank underestimating its capital requirements.

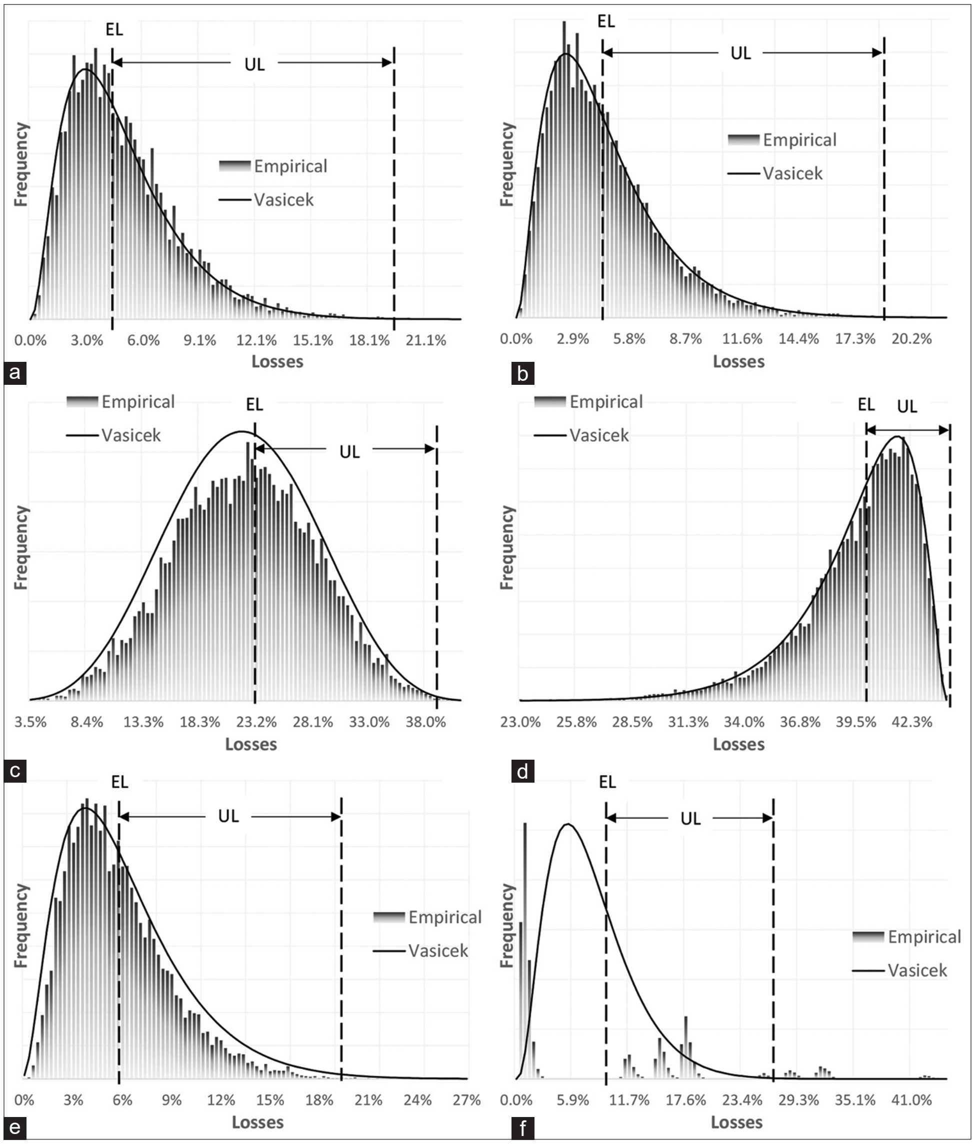
Table 1: Parameterisation for each loan within the various portfolios

Portfolio	Type	PD (%)	LGD (%)	EAD
1	Homogeneous	12	40	500
2.1, 2.2, 2.3	Constant PDs	10, 50, 90	40	500
3	Mixed	(0.03; 24.59)	(28;59)	(100, 1 000)
4	Mixed with 3 large loans	(0.03; 24.59)	(28;59)	(100,1 000) + 3·1 000 000

Table 2: Portfolio 1-risk measures as % total EAD (Empirical vs. Vasicek)

Risk measure	Empirical	Vasicek	Empirical/Vasicek
EL	4.85	4.80	1.011
C-VaR	18.56	18.25	1.017
UL	13.71	13.45	1.019

Figure 3: (a) Portfolio 1-Credit loss distribution, (b) Portfolio 2.1 (PD=10%)-Credit loss distribution, (c) Portfolio 2.2 (PD=50%)-Credit loss distribution, (d) Portfolio 2.3 (PD=90%)-Credit loss distribution, (e) Portfolio 3-Credit loss distribution, (f) Portfolio 4-Credit loss distribution



4.2. Portfolio 2

In portfolio 2 the effect of three different constant PDs on varying LGDs and EADs is investigated. The total EAD for portfolio 2.1, 2.2 and 2.3 is 554 212. This portfolio provides insight on the effect PDs and correlations have on distributions of losses.

4.2.1. Portfolio 2.1

The first scenario sets the PD=10%. Since this PD is within the standard range, a major difference from previous results is not expected (Figure 3). Losses are lognormally distributed in Figure 3b and both the EL from the empirical distribution and the Vasicek model align well. While the 99.9% C-VaR is a close fit, the Vasicek C-VaR is slightly underestimated. There is little to no difference in the shape and results of the distributions when PD=10% compared to portfolio 1. This is attributed to the PD being very similar to the previous PD range (Table 3).

In Table 3, the Vasicek model overestimates the risk measures. Differences between the EL and C-VaR are 0.024% and 0.192% respectively. Even though the LGDs and the EADs are random, by fixing the PD, the risk measures obtained by the Vasicek model are still an accurate approximation.

4.2.2. Portfolio 2.2

The PD is now increased to 50% to analyse the effect it has on the distribution. From Figure 2, once the PD exceeds 15%, the correlation plateaus at 12%. Figure 3c illustrates the loss distribution for a constant PD of 50%.

The portfolio loss distribution is roughly normally distributed, the normality shape due to the shift in ELs. EL as a percentage of EAD shift by approximately 17.5% compared to the EL observed for portfolio 2.1. This is consistent with the assumption that a higher PD will result in greater losses. An increase in the C-VaR as a percentage of EAD of ≈ 20% is observed. Due to the higher PD, the EL as a percentage of total exposure is significantly higher compared to the previous portfolios analysed. Due to the high PD greater losses are observed than when $PD \in [0.003\%; 24.59\%]$.

The correlation is dependent on the PD, which determines how much the single risk factor will affect the borrower defaulting. Since the PDs are high and consistent, they are all affected in the same way by the single risk factor with the only difference being the exposure amounts and the LGD (Table 4).

The Vasicek EL very slightly overestimates the empirically observed EL by 0.039%. In addition, the Vasicek C-VaR overestimates the empirically observed C-VaR by 0.19% and the Vasicek UL overestimates the empirically observed UL by 0.15%.

4.2.3. Portfolio 2.3

Finally, PDs are increased to 90%. In this case, the PD is almost guaranteed and thus even higher losses are expected. The distribution for this portfolio is illustrated in Figure 3d.

The distribution of losses in Figure 3d is highly skewed to the right. This is due to the high number of large losses expected when the portfolio is comprised of PDs that are 90%. Figure 3d illustrates that both the EL and the C-VaR under the empirical distribution and the Vasicek model are nearly identical, making the Vasicek distribution a good approximation. A summary of the risk measures for portfolio 2.3 is given in Table 5.

In Table 5, the EL is high compared to portfolios 2.1 and 2.2. The EL as a percentage of EAD increased by approximately 17.6% and 35.19% for portfolio 2.2 and 2.1 respectively. Further, the C-VaR as a percentage of EAD increased by approximately 5.58% and 25.84% for portfolio 2.2 and 2.1 respectively. The ratio of all the risk measures between the empirical and Vasicek are almost equivalent to one. This suggests that as PD are increased, the accuracy of the Vasicek model increases. As PDs increase, the difference between the EL and C-VaR decreases, which results a smaller UL. For high PDs, the EL accounts for a larger portion of the credit losses which results in the UL contributing to a smaller portion of the portfolio losses.

Table 3: Portfolio 2.1 (PD=10%)-Risk measures as % total EAD (Empirical vs. Vasicek)

Risk measure	Empirical	Vasicek	Empirical/Vasicek
EL	4.37	4.40	0.994
C-VaR	17.94	18.14	0.989
UL	13.57	13.74	0.988

Table 4: Portfolio 2.2 (PD=50%)-Risk measures as % total EAD (Empirical vs. Vasicek)

Risk measure	Empirical	Vasicek	Empirical/Vasicek
EL	21.95	21.97	0.998
C-VaR	38.20	38.39	0.995
UL	16.25	16.41	0.991

Table 5: Portfolio 2.3 (PD=90%) Risk measures % total EAD (Empirical vs. Vasicek)

Risk measure	Empirical	Vasicek	Empirical/Vasicek
EL	39.57	39.58	0.999
C-VaR	43.78	43.715	1.002
UL	4.21	4.13	1.020

When PDs increase, the EL increases. Since the asset correlation is linked to PD, the higher the PD the lower the correlation resulting in the single risk factor having less of an effect on the obligor defaulting. Borrower default thresholds decrease as PD increases and correlation decreases. The lower threshold thus increases the probability of the borrower defaulting which results in a negatively skewed distribution when the portfolio comprises borrowers with high PDs. Due to the use of constant PDs, the portfolio is undiversified which increases the risk of greater losses. Thus, PDs play a key role in determining loss severity and directly affect correlations. The relationship between UL and PD can be thought of as follows: UL increases as PD increases until a certain point. Once the PD increases past a threshold, the UL decreases because for increasing PDs, a larger portion of portfolio losses are accounted for by the *EL* resulting in lower ULs.

4.3. Portfolio 3

In portfolio 3, the effect of a mixed portfolio with a total EAD of 554 211 was analysed. This portfolio is the most realistic portfolio since each loan in a portfolio is unique or diversified based on its PD, LGD and EAD. The credit loss distribution for the empirical and Vasicek distribution for a "real world" scenario is illustrated in Figure 3e.

The Vasicek distribution provides a good fit for the empirical distribution. Larger deviations between the empirically observed C-VaR and the Vasicek C-VaR which is expected since this portfolio consists of random PDs, EADs and LGDs. The risk measures for portfolio 3 are summarised in Table 6.

The Vasicek model underestimates the EL by 0.040% compared to the empirically observed EL, which translates to an underestimation of approximately 222. When all variables are random, the Vasicek EL is still a good approximation to the empirically observed EL and is more accurate compared to the homogeneous portfolio (portfolio 1). In contrast, the Vasicek C-VaR is overestimated by 1.74% compared to the empirically observed C-VaR, which results in the Vasicek UL being overestimated by approximately 9 900. Thus, when PD, EAD and LGD are all random, the Vasicek model does not approximate the UL well.

4.4. Portfolio 4

The final portfolio has three dominating loans resulting in a total EAD of 3 552. This portfolio is introduced to illustrate that the convergence of the portfolio loss distribution does not hold when dominating loans are included in the portfolio as stated by Vasicek (1991). In Figure 3f.

The loan losses for the empirical distribution are scattered and concentrated at certain loss levels, which no longer conforms to the usual shape observed in previous portfolios. The Vasicek distribution (Figure 3f) does not fit the empirical distribution when a few dominating loans are included. Vasicek C-VaR and the empirically observed C-VaR differ significantly (Table 7).

The Vasicek EL is underestimated by 0.53% when compared to the empirically observed EL which is equivalent to ≈ 18 600. The biggest drawback of the Vasicek model in a portfolio with a few dominating loans, is the estimation of the C-VaR. A comparison of the Vasicek C-VaR and the empirically observed C-VaR shows that the Vasicek model severely underestimates the C-VaR by 17.83%, nearly 20% of the total portfolio EAD. The ratio of the empirically observed UL is nearly 2 times larger than the Vasicek UL. The Vasicek UL underestimates the empirically observed UL by 17.30%, which translates to approximately 615 000. Financial institutions must ensure that there are no dominating loans within a portfolio when using the Vasicek model to approximate credit losses.

4.5. Impact of Variables on Capital Ratio Determined by the Vasicek Model

In this section the impact of two variables on capital requirements is explored, namely the effect of PD and LGD, PD and maturity as well as LGD and maturity on credit risk capital requirements.

Table 8 shows the ranges for the risk variables as well as the constant value when investigating the effect of the other variables.

4.5.1. Capital requirement to a change in PD and LGD

The change in the capital requirement to a change in PD and LGD is given in Figure 4a.

Figure 4a shows that as LGD increases the capital requirement increases. The capital requirement increases as PD increases until ≈30%. Beyond a PD of 30% the capital requirement decreases with PD. For PD < 30% and for an increasing LGD, the capital requirement increases while for a PD increasing above 30% and an increasing LGD the capital requirement decreases. For example, using a PD of 10% and an LGD of 40% the capital requirement is lower than when the PD is 20% and LGD is 60%. Capital requirements reach a peak when PD=30% and LGD=100%. However, when PD=60% and LGD=100% the capital requirement is higher than when PD=80%.

LGD has an almost proportional relationship with the capital requirement whereas PDs lower than 30% have a steep increasing relationship and thereafter have a decreasing relationship that declines at a slower rate than when increasing. This relationship between PD and LGD is supported by Figure 4a which is

Table 6: Portfolio 3 - Risk measures as a percentage of total EAD (Empirical vs. Vasicek)

Risk measure	Empirical	Vasicek	Empirical/Vasicek
EL	5.61	5.57	1.007
C-VaR	18.87	20.62	0.916
UL	13.26	15.05	0.882

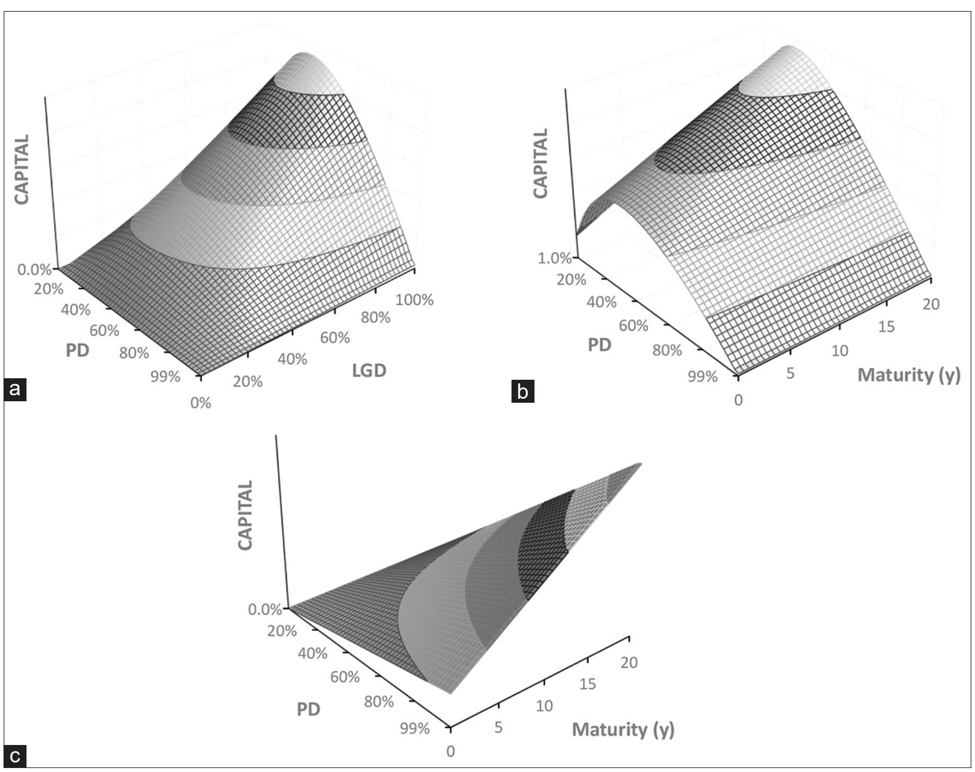
Table 7: Portfolio 4-Risk measures as a percentage of total EAD (Empirical vs. Vasicek)

Risk measure	Empirical	Vasicek	Empirical/Vasicek
EL	8.26	7.73	1.068
C-VaR	43.39	25.56	1.698
UL	35.13	17.82	1.971

Table 8: Ranges and constant value for risk variables

Risk variable	Range	Value
LGD	[0.00001;1]	40%
PD	[0.00001;1]	15%
M	[1;20]	1 year

Figure 4: (a) Capital requirement to a change in PD and LGD, (b) capital requirement to a change in PD and M (c) capital requirement to a change in LGD and M



skewed to the right with lower PDs and an increasing slope for a fixed PD.

4.5.2. Capital requirement to a change in PD and M (maturity)

The change in the capital requirement to a change in PD and M is given in Figure 4b. PD maintains the same trend as in Figure 4a and b. The capital requirement increases sharply as PD increases to 30% then decreases gradually as the PD decreases. The shape in Figure 4a remains constant for PD against the capital requirement and multiplied by some factor as M increases. For example, when M = 1 year and PD = 20% the capital requirement is lower than the capital requirement when the PD = 20% and M = 15 years.

Figure 4 supports BCBS (2005) stating that the distribution of the PD and K relationship does not change other than being scaled by the maturity adjustment factor. Thus, it is evident that since the capital requirement is simply multiplied by the maturity factor, a maturity above 5 years would be futile. The maturity factor may be too high or even too low for high maturities and would not be accurate as multiple factors influencing the capital requirement may change in that time.

4.5.3. Capital requirement to a change in LGD and M (maturity)

The change in the capital requirement to a change in LGD and M is given in Figure 4c. As LGD increases and M increases so too does the capital requirement. For example, for an LGD of 20% and a maturity of 1 year, the capital requirement would be lower than for an LGD of 30% and a maturity of 5 years. As LGD increases, the capital requirement increases at a steeper rate as maturity increases when compared to lower LGDs.

LGD has a proportional relationship with the capital requirement and thus as LGD increases, more capital is required in case of counterparty default. When including a maturity adjustment this relationship is multiplied by the maturity factor which only multiplies the capital requirement by this factor. Since the maturity adjustment is only a function of maturity and PD, the maturity factor is only altered slightly by an increasing maturity since PD is kept constant. Figure 4c further supports that the LGD and K relationship is not altered but the distribution is merely scaled by the maturity factor.

To test the robustness of the Vasicek model, a portfolio with a few dominating loans was considered. Results indicate that the Vasicek model does not provide a good fit for the empirical distribution when dominating loans are included in the portfolio. This confirms the conditions of the loan portfolio that are required for the closed-form approximation to converge to the empirical distribution, as suggested by Vasicek (2002). These dominating loans influence the risk measures significantly and result in the Vasicek model underestimating extreme losses. Thus, when the assumptions of the Vasicek model are followed, the model serves to be an accurate approximation of the empirical distribution.

5. CONCLUSION

Credit risk has become increasingly more important in recent years and poses a big risk to financial institutions if not managed adequately. There are several different models available to estimate credit losses within a portfolio. One of these models is the Vasicek (2002) model which provides a closed-form approximation for the distribution of credit losses by making use of average input variables. Vasicek's model (2002) was used to estimate the credit loss distribution of a portfolio and the robustness of the model in comparison to the empirical distribution.

Results indicate that the asymptotic single risk factor Vasicek model generally provides a good fit for the empirical distribution, specifically when considering a homogeneous portfolio. As random PDs, EADs and LGDs are introduced within the portfolio, the accuracy of the Vasicek distribution decreases. Results indicate that the Vasicek model is highly dependent on the PD. The lower the PD, the greater the correlation, resulting in a greater effect from the single risk factor. As PD increases, the EL accounts for more of the overall portfolio losses than the UL. This increases the accuracy of Vasicek's approximation since the EL is predicted with a higher degree of accuracy compared to the UL, since the Vasicek distribution makes use of the average PD and LGD across the portfolio.

REFERENCES

Adamko, P., Kliěstik, T., Birtus, M. (2014), History of credit risk models. Advances in Education Research, 61, 148-153.

Altman, E.I. (1968), Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, 23(4), 589-609.

Altman, E.I., Haldeman, R.G., Narayanan, P. (1977), ZetaManalysis a new model to identify bankruptcy risk of corporations. Journal of Banking and Finance, 1(1), 29-54.

BCBS. (2001), The Internal Ratings-Based Approach. Available from: https://bis.org/publ/bcbsca05.pdf

BCBS. (2004), International Convergence of Capital Measurement and Capital Standards. Available from: https://bis.org/publ/bcbs107.pdf

BCBS. (2005), An Explanatory Note on the Basel II IRB Risk Weight Functions. Available from: https://bis.org/bcbs/irbriskweight.pdf

BCBS. (2006), International Convergence of Capital Measurement and Capital Standards (Revised Framework). Available from: https://bis.org/publ/bcbs128.pdf

BCBS. (2011), Treatment of Trade Finance Under the Basel Capital Framework. Available from: https://bis.org/publ/bcbs205.pdf

Beaver, W.H. (1966), Financial ratios as predictors of failure. Journal of Accounting Research, 4(1), 71-111.

Black, F., Scholes, M. (1973), The pricing of options and corporate liabilities. The Journal of Political Economy, 81(3), 637-654.

Fitch. (2022), Issuer Default Ratings. Available from: https://your.fitch.group/rating-definitions.html [Last accessed on 2022 Sep 01].

Genest, B., Brie, L. (2013), Basel II IRB risk weight functions: Demonstration & analysis. Available from: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2578936_code2371644.pdf [Last accessed on 2023 Dec 06].

Gordy, M.B. (2003), A risk-factor model foundation for ratings-based bank capital rules. Journal of Financial Intermediation, 12(3), 199-232.

Gregory, J. (2012), Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets. 2nd ed. United States: Wiley.

Harjans, L. (2018), A Comparison of Altman's z-Score and the j-model in Assessing Corporate Failure: Evidence from the USA. PhD Thesis, University of Twente. Available from: https://essay.utwente.nl/75423

Jarrow, R.A., Turnbull, S.M. (1995), Pricing derivatives on financial securities subject to credit risk. The Journal of Finance, 50(1), 53-85.

Jorion, P. (2011), Financial Risk Manager Handbook. 6th ed. United States: Wiley.

Kiliç, E. (2007), Simulation of Credit Risk and Modeling pd/lgd Linkage, PhD Thesis, Istanbul Bilgi U¨ niversitesi.

Kliĕstik, T., Cu´g, J. (2015), Comparison of selected models of credit risk. Procedia Economics and Finance, 23, 356-361.

Kupiec, P. (2008), A generalized single common factor model of portfolio credit risk. The Journal of Derivatives, 15, 25-40.

Lee, J., Wang, J., Zhang, J. (2009), The Relationship between Average Asset Correlation and Default Probability. United States: Moody's KMV Company.

Merton, R.C. (1974), On the pricing of corporate debt: The risk structure of interest rates. The Journal of Finance, 29(2), 449-470.

Moody's Investor Service. (2022), Rating Symbols and Definitions. Available from: https://moodys.com/sites/products/productattachments/ap075378_1_1408_ki.pdf

Pykhtin, M. (2004), Portfolio credit risk multi-factor adjustment. Risk, 17(3), 85-90.

Python. (2022), Random - Generate Pseudo-Random Numbers. Available: https://docs.python.org/3/library/random.html [Last accessed on 2022 Sep 01].

S&P Global. (2021), Corporate Credit Risk Trends in Developing Markets: An LGD Perspective. United States: S&P Global.

S&P Global. (2022), 2021 Annual Global Corporate Default and Rating Transition Study. New York: S&P Global.

Vasicek, O.A. (1984), Credit Valuation. KMV Corporation. Available from: https://ressources-actuarielles.net/ext/isfa/1226.nsf/0/c181fb77ee99d464c125757a00505078/$file/credit_valuation.pdf

Vasicek, O.A. (1987), Probability of Loss on Loan Portfolio. KMV Corporation. Available from: https://moodysanalytics.com/-/media/whitepaper/before-2011/02-12-87-probability-of-loss-on-loan-portfolio.pdf

Vasicek, O.A. (1991), Limiting loan loss probability distribution. In: Finance, Economics and Mathematics. United States: Wiley. p147-148.

Vasicek, O.A. (2002), The distribution of loan portfolio value. Risk, 15(12), 160-162.