Computational Finance Masters Thesis
Expected Credit Loss Modeling with Macroeconomic Forecasts

Ntshuxeko Charity Mdaka
Student Number: 813796

Supervisor: Dr. Blessing Mudavanhu

Programme in Computational And Applied Mathematics,
School of Computer Science and Applied Mathematics,
University of the Witwatersrand, Johannesburg.

February 7, 2019

Declaration

I declare that this thesis is an original report of my research, has been written by me and has not been submitted for any previous degree. Due references have been provided on all supporting literatures and resources.

Signature:____

Ms. N.C Mdaka

Abstract

The potential that a counterparty will default on their contractual obligations is known as credit risk. Credit risk plays a crucial role in the banking industry, as the primary role of banks is to give out loans. Thus credit risk evaluation, which is also influenced by the macroeconomic environment, is undoubtedly one of the most important activities needed to be carried out by banks.

In this thesis we explore three fundamental parameters used in credit risk management and analysis which are namely, probability of default (PD), loss given default (LGD) and exposure at default (EAD). Focusing specifically on PD, we want to estimate these credit risk parameters throughout the lifetime of a loan under the newly adopted IFRS (International Financial Reporting Standards) 9 accounting standard, whilst incorporating macroeconomic factors. In other words we infer a PD term structure curve, which is a major input, into the modelling of credit risk with inclusion of macroeconomic forecasts that are relevant.

Acknowledgements

I would like to express my earnest appreciation to my supervisor, Dr. Blessing Mudavanhu, for his assistance and introducing me to this relevant topic. I would also like to thank my mother, Ms. A.J Siphuma, as well as everyone from the University of the Witwatersrand who provided support and encouragement.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Introduction of IFRS 9 standard

Credit risk can be simply defined as the risk of a loss resulting from the fact that a borrower or counterparty fails to fulfill their financial obligations under the agreed term [21]. Credit risk is said to be the most important type of risk that has been present in finance, commerce and trade transactions [25]. Thus one can say that credit risk management is a challenging and complex task in the financial industry due to the unpredictable nature of macroeconomic factors [27].

The taking and management of risks forms a big part of banking business. Lending, which is a major activity of banks, carry the risk that the borrower will not repay the loan as promised [34]. It is required that banks retain an adequate capital amount, in order to cover potential losses, that are unexpected. There are regulations that are in place to ensure that banks abide by these requirements. Previously, these requirements for financial instruments were implemented under the international accounting standard 39 (IAS 39) regulation. However the 2007 financial crisis highlighted the systemic costs of delayed recognition of credit losses on the part of banks and other lenders. This meant the IAS 39 had to be replaced as it exhibited some inconsistencies regarding the impairment (credit losses) methodology. In other words a new credit risk modelling paradigm needed to be introduced where the best estimate approaches for projection of credit risk, that are macroeconomic -forward looking, are used.

On the 1^{st} January 2018, IAS 39 was replaced with the international financial reporting standards 9 (IFRS 9) in order to address the "too little too late" recognition of credit losses during the financial crisis. In other words, since the fundamental responsibility of accounting is to ensure the "true and accurate" representation of accounts in the Profit and Loss statement, accountants were uncomfortable with accounting based on "expected" outcomes, unless they have "sufficient proof" to convince them otherwise, hence leading to the "too late" provisioning of loan losses [32].

This transition from IAS 39 to IFRS 9 took place through three components namely, the measurement and classification of financial assets and liabilities, the methodology of impairment as well as hedge accounting [27], as can be seen in the figure below. The identification of business models for asset classification is simplified due to the fact that in emerging markets, many banks have asset allocations that are not complex nor prone to changes that are dynamic [32].

Component 1	Component 2	Component 3
Classification and measurement	***Impairment***	***Hedge Accounting***
Classification of instruments are now based on: - Entity's Business Model - Contractual cash flow characteristics test (SPPI)	IFRS 9 replaces IAS 39 "Incurred Loss" Model with new Forward looking Expected Credit Loss (ECL) model.	Simplified hedge accounting rules to reflect more accurately how an entity manages its risk.
Introduction of new measurement category: "Fair Value through Other Comprehensive Income" (FVOCI)	ECL is applicable for instruments classified under Amortized Cost and FVOCI (debt instruments) category.	Introduction of new hedge accounting model.
"Tainting rule" has ceased to exist in IFRS 9- i.e. sales of "held to maturity" assets under IAS 39 before maturity jeopardize amortized cost accounting for entire portfolio.	Recognizes 12-month loss allowance at initial recognition and lifetime loss allowances on significant increase in credit risk.	The 80-125% hedge effectiveness testing ratio ranges is replaced by an objectives based test that focuses on the economic relationship.
	Requires incorporation of forward looking information in ECL estimates.	More designations of groups of items as the hedged items are possible under new rules

Figure 1: Three structured components of IFRS 9. Adapted from [32].

The impairment methodology however, is the component that has the greatest impact on the most reported characteristics of banking institutions because extensive credit risk modeling is required for impairment calculations, making impairment modeling one of the core areas where banks need to focus on [32]. Hence, the biggest weakness of IAS 39 was in the mechanism of the calculation of impairment of financial assets and accounting their loss allowances [27]. This deficiency of IAS 39 was the strongest motivation for the replacement by IFRS 9.

Under IAS 39, the impairment methodology took place based on the Incurred Loss Model, meaning only after certain adverse events have occurred were loss allowances recognized. The IFRS 9 has thus introduced a change with regards to this problem and with regards to the requirements for the calculation of impairment. Now under the IFRS 9, the impairment methodology is based on the "forward-looking" Expected Credit Losses Model, meaning credit losses and loss allowances have to be recognized based on expectations, which means before a certain potential adverse event takes place [27]. The need for the ECL model was thus established after the global crisis and provides better and advanced information to investors [25].

The Expected Credit Losses (ECL) Model which is the new credit impairment model, depends mainly on the quality and availability of credit risk data. This new model requires considering forward looking information and not only historical data. This forward looking information is related to the entities' own estimate of their customers likelihood of probability of default, expected recovery patterns as well as macroeconomic factors like unemployment and recession [34]. Thus this model provides timely recognition of ECL, assessment of significant increase in credit risk which will provide better disclosure as well as assurance of better financial ratios [40]. And its successful implementation requires a close coordination between Risk and Finance teams within a bank [32].

Financial assets that apply to the ECL model include debtors, inter-corporate loans, any debt investments, loan commitments, financial guarantee contracts, lease receivables amongst others. Based on whether there has been a significant increase in credit risk of a financial asset from its initial recognition, the new ECL model, under IFRS 9, establishes a 3 staged impairment model whereby the impairment of financial assets can either be measured as 12-month or lifetime expected credit losses. These stages determine the amount of impairment to be recognized as ECL at each reporting date.

There are two main goals behind the intention of the three staged assessment procedure. Firstly, it is to prevent inaccurate distribution of income for accounts that have decreased quality since origination, because the interest rate may no longer cover the credit risk premium accordingly. Secondly, to recognize and make allowance for potential losses at an early stage, rather than waiting until the accounts become 90 days past the due date [52]. The diagram below shows these three stages:

Deterioration of credit quality from initial recognition

STAGE 1	STAGE 2	STAGE 3
PERFORMING	**UNDER-PERFORMING**	**CREDIT-IMPAIRED**
- Financial instruments that have no significant deterioration of credit quality since initial recognition OR - Indicate low credit risk at reporting date. - **12 months expected credit losses is recognized for these assets.** - Effective interest rate on the gross carrying amount.	- Financial instruments with significant deterioration of credit quality (not impaired i.e. don't portray any objective evidence of impairment). - **Lifetime expected credit losses is recognized for these assets.** - Effective interest rate on the gross carrying amount.	- Credit quality deteriorated to a level at which credit loss actually occurs (objective evidence of impairment on reporting date). - **Lifetime expected credit losses is recognized for these assets.** - Effective interest rate on net book value.

Figure 2: Three phases of the new ECL model. Adapted from [1].

As we can see on the above figure the requirement for stage assessment is "significant increase in credit risk". The interpretation of such a requirement leads to situations that may not necessarily agree with the traditional way that bankers consider asset quality. Now because of the stage assessment two assets that fall under the same rating may no longer fall in the same stage because it's not only the absolute rating grades that matter , but the change in credit quality i.e. change in rating grades.

This means that the three staged IFRS 9 is a relative reflection of the deterioration of portfolio quality since inception and not an absolute representation of the quality of the portfolio [52].

One of the key factors for stage assessment is lifetime PD as the true risk of default is not being captured for financial assets that have a maturity of more than 12 months, especially if adverse macroeconomic conditions were considered. This points out the need to identify PDs term structure. The feasibility of the of measurement of the lifetime PD at the time of origination remains a matter of interest , even though banks are able to identify the lifetime PD for an account, on the reporting date. In other words, while all non-performing accounts may fall into stage 2, there could be other additions to the stage 2 classification that are based on either a change in the lifetime PD or the triggers of credit deterioration [52].

Traditionally, banks have been depending on the "days-past-due" framework of credit monitoring. However, credit deterioration triggers also need to be considered for this stage assessment. Thus banks need to build a framework that points out scenarios that would cause movement of assets from the first to the second phase. Fortunately, IFSR 9 allows banks to define their own policy for migration and indicators.

Below is a table that shows the 16 classes of indicators which should be considered in stage assessments by banks.[1]. This table indicates that the triggers are not only based on historical facts and information, but also also on forward looking scenarios as emphasized by the word "expected" in the table [52]. Hence stage assessment requires significant amount of information analysis as well as utilization of experienced credit judgement.

[1]They are stated in paragraphs B5.5.17 in the IFRS 9 standards

Classes of credit quality indicators			
1.	Change in internal credit spread or risk premium.	9.	Significant difference in rates or terms of newly issued similar contracts.
2.	CDS spread, equity or debt price.	10.	Actual or expected change in external credit rating.
3.	Actual or expected change in operating results of the borrower.	11.	Significant increase in credit risk on other financial instruments of the same borrower.
4.	Actual or expected significant change in internal credit rating or behavioral score.	12.	Existing or forecast adverse changes in business, financial, or economic conditions.
5.	Regulatory, economic, or technological environment of the borrower.	13.	Collateral value.
6.	Quality of guarantee.	14.	Reductions in financial support from parent entity or credit enhancement quality.
7.	Expected change in loan documentation.	15.	Significant changes in the expected performance and behavior/ borrower/ group.
8.	Changes in bank's credit management approach or appetite in relation to the financial instrument.	16.	30-dpd rebuttable presumption.

Figure 3: Credit quality indicators. Adapted from [52]

In this thesis however, we will assume that the assets are already in the second phase/ stage (for under-performing assets), thus we won't be implementing stage assessments, as we are more concerned with the financial modeling i.e. construction of the term structure of risk parameters in order to compute lifetime expected credit loss.

Expected credit losses should be calculated based on a weighted average of credit losses that can occur within different cases with the certain probability. The expected credit losses should be discounted using the effective interest rate as the time value of money is also taken into account. The implementation of the IFRS 9 requirements for financial assets are expected to cause an increase in the overall level of loss allowances, which is one of its most significant impacts [40]. An increase in overall loss allowances also leads to increase in costs which in turn will affect the profit and loss of accounts of banks.

All relevant information about the historical, current and future conditions that can be obtained without excessive cost, should be taken into account when estimating expected credit losses. Under the Basel framework (internal ratings-based approach- IRBA), the expected losses are greater in general than loss allowances under IAS 39. Now it's expected that the loss allowances under IFRS 9 will be greater than those under the Basel framework (IRBA) [27].

There are however some methodological differences in the estimation of credit risk parameters under both IFRS 9 and Basel. The first difference, as mentioned above is in the calculating of lifetime expected credit losses when there is a significant increase in credit risk, which is compensated for in part by the nature of credit risk parameters compared to Basel, which includes important measures like considering an economic decline [27]. The other methodological difference is the nature of PD estimates which is point-in-time (PIT) or through-the-cycle (TTC). The point-in-time nature means account ratings move over the cycle given more recent economic conditions, whilst through-the-cycle nature predicts average default rate performance over an economic cycle and ignores short run changes to an account's PD [54].

Under the Basel framework, in general the PDs are estimated as TTC because of the low volatility of credit risk capital requirements that are desired [1]. Thus under IFRS 9, the PD estimates should be more real time, and thus PIT, including forward looking information, especially macroeconomic factors. The implementation of IFRS 9 may also have an effect on pricing and other internal processes such as sales and marketing [54].

Hence since macroeconomic stress testing is being officially adopted by many jurisdictions worldwide as a risk management standard, multi-horizon/period credit modelling has become increasingly important. Thus putting everything together the new credit risk modeling requirements introduced by the IFRS 9 framework, which is the new international accounting standard for credit loss recognition, can be summarized into three main characteristics:

(i) Forward looking and based on macroeconomic factors

In the past credit models were based on history and were independent of the macroeconomic environment. This limited focus on the macroeconomic environment is what led the credit modeling to be biased towards the optimistic side, prior to the financial crisis [48]. Thus the introduction of macroeconomic sensitive credit analyses that are based on a forward looking view enables us to now reflect scenarios that might not have existed in the historical economic data that is available [48]. Thus we can now capture the underlying factors for systematic credit effects that were previously overlooked.

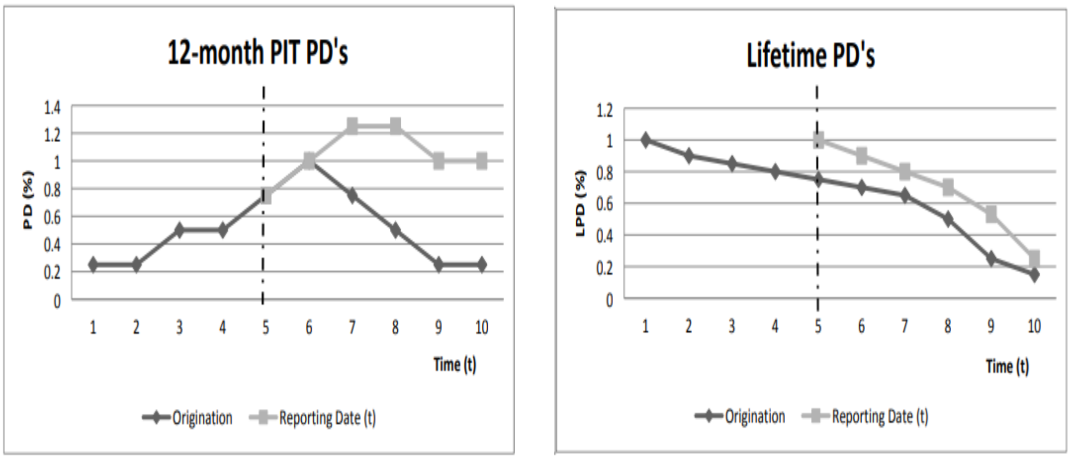
(ii) Multi-period

Banking book credits are mostly held to maturity, and there is seldom a jump-to-default [48]. Actual defaults occur as a consequence of migrations to increasing delinquency, which is a situation where a borrower is late or overdue on a payment, from one period to the next [48]. Thus in the same time frame, delinquent accounts can recover and be recognized as healthy once again [48].

It is thus important for credit deterioration to be recognized long before the actual default [48], as it not only affects the default forecasting but also the modelling of the cash flows in the future. For instance, a delinquency that is forecasted may cause a shortage of cash inflow, which will affect the liquidity and net income of a future period [48], whilst the treatment of a delinquency may bring in more cash flows, accrued interest and fees to the bank [48].

Thus as mentioned above, under IFRS 9 when an exposures credit quality decreases significantly or is impaired, then the exposure is subject to lifetime credit loss calculation which is based on a business-as-usual economic scenario. If the credit exposure's credit quality deterioration is not significant, then a 12 month best estimate for expected credit loss is used [48].

The graphs below illustrate the importance of multi-period forecasts. The graph on the left displays a scenario where the 12 months PD doesn't capture the significant increase in credit risk (at $t = 5$ the PD is the same on the origination and reporting date, meaning the economic downturn has not been captured, it's impact is only recognized at $t = 7$) is not captured, when an economic downturn is expected to occur at a later stage, where t would correspond to the PD time period. We see however that the Lifetime PD does capture this downturn (at $t = 5$ there is a jump, which shows an immediate response to economic downturn) and thus identifies the significant increase in credit risk sooner, this is shown by the graph on the right.



(a) (b)

Figure 4: 12 months vs. Lifetime PDs. Adapted from [24].

Thus the left graph is a proxy for cases where it's not a reasonable approximation to use a 12-month PD to using lifetime PD [49], whereby a lifetime PD refers to a PD over more than one time period, i.e. multiperiod PD, and a 12 month PD refers to PD of time one period , which is a year. Examples of such cases include, bullet repayment loans where the payment obligations of the debtor are not significant during the first 12 months of the loan and also loans where changes in credit-related factors have an impact on the credit risk of the financial instrument only beyond 12 months [49].

(iii) *Granular level*

The introduction of IFRS 9 calls for as much granularity as possible. In other words the level of detail considered when assessing expected credit loss is deep, in order to capture the exposure that reacts differently to the drivers of risk under stress scenarios [48]. In order to optimize modelling performance, this increased granularity of credit risk modelling demands better programming applications whilst at the same time significantly increasing computational requirements [48].

1.2 Parameters for credit risk modelling

Under the IFRS 9 framework, the Impairment assessment requires the computation expected credit losses (ECLs), as we discussed above. These ECLs should reflect a probability-weighted outcome, the time value of money as well as the best forward looking information that's available and can be calculated using a modular approach that uses the three parameters (PD, LGD and EAD). Below we briefly define and describe each parameter[24].

1.2.1 Probability of default (PD)

Before we define PD we need to know what exactly we mean by default. The way financial institutions defined default may vary and depending on these internal definitions, the reasons for default differ. Reasons for default may range from suspicions by the credit holder of not being repaid in full, being late with payments to filing for bankruptcy [36]. There are acts and regulations published set by the European Union and the Basel Committee on the calculation of certain capital requirements as well as concluding when an obligor is in default [36].

According to these regulations[2], an institution should consider a default to have occurred if one of the following scenarios occurs:

(i) The obligor is unlikely to pay their credit obligations to the institution in full, whether it be the principal, the interest or the fees amount.

(ii) The obligor is past due more than 90 days on any material credit obligation to the institution.

(iii) In the context of swaps and derivative contracts, default can be defined by looking at credit events, stated in the International Swap Dealers Association (ISDA), that could trigger a settlement under a Credit Default Swap (CDS) agreement. The three credit events that are most common are [37]:

  (a) *Bankruptcy*: This is when the obligor entity has filed for relief under bankruptcy or equivalent law.

  (b) *Failure to pay*: The obligor entity fails to make due interest or principal payments, after the grace period expires, if applicable.

  (c) *Debt restructuring*: When the arrangement of debt obligations is altered in such a way that the credit holder is negatively affected, such as extension of maturity or reduction of coupons.

Thus stemming from the definitions of default, the probability of default (PD) is said to be the key parameter of credit risk for estimating credit losses. It is defined as the likelihood that a borrower will not be able to make scheduled repayments over a specified period, usually one year [2]. It is a key metric that summarizes the credit worthiness of the obligors of a bank [28]. PD doesn't only depend on the borrower's characteristics but also on the economic environment. This probability is implied by the credit rating for businesses and PDs can also be estimated using historical data and statistical techniques [2].

[2] EU Regulation No.575/ 2013, under Article 178

Within the banking industry, since the vital activities involve granting loans, credit scoring is used to predict the probability that a loan applicant or client will default [27]. Thus credit scoring is a process or statistical approach used to predict PD in banks. Credit scoring is split into two categories which are application and behavioural credit scoring which represent credit risk evaluation of loan applicants and actual clients respectively.

In the past banks mainly focused on application credit scoring governed on a qualitative approach based on the judgement of the credit officer. However, this approach had some obvious problems such as inconsistency, inefficiency, incomprehensiveness and subjectivity. This led to development over the years of a quantitative approach where statistical credit models have been developed and advanced. These credit scoring models thus assist in overcoming the shortfalls of the qualitative approach based on the credit officer's judgement. Thus credit scoring models have become a standard method for estimation of PD.

Credit scoring is said to be the most popular application of logistic regression [27]. Logistic regression is the statistical method for analyzing a data set in which there is one or more independent variables that determine the outcome [3]. Even after more advanced models have been developed, logistic regression is still used as the standard tool because it is simple and intuitive and also provides relatively good results. Below is the logistic regression formula[3]:

$$ln\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta_1 x \quad \Rightarrow \quad \frac{p(x)}{1-p(x)} = e^{\beta_0+\beta_1 x} \tag{1}$$

where:

- ln is the natural logarithm.

- $p(x)$ is the probability that the dependent variable equals a case, given some linear combination of the explanatory variables

- β_0 is the intercept from the linear regression equation.

[3]Formula adapted from [4]

- e denotes the exponential function.

- $\beta_1 x$ is the regression coefficient multiplied by some value of the predictor.

Contrary to the common misconception, as a parameter, PD is quite different from the default rate. As mentioned above, the PD is the predicted probability that a group of obligors will default over a predefined future time horizon. While on the other hand, the default rates are the defaults that are actually realized over a given period [28].

1.2.2 Loss Given Default (LGD)

Loss given default is described as the percentage of exposure which will be not recovered after a counterparty defaults [25]. An accurate forecast of the loss given default (LGD) parameter of loans plays a crucial role for risk-based decision making by banks [37]. However, LGD has not been as extensively studied, and is considered to be a more daunting modeling challenge than other parameters, such as PD. This lack of studies on the LGD modeling is mainly because to the fact that the PD and the LGD are difficult to separate based on the price of single financial instrument [37].

LGD is also directly related to the recovery rate (R) on a defaulted loan. The recovery rate is the proportion of bad debt that may be recovered in the event of default [25]. Thus:

$$R = 1 - LGD$$

LGD affects many areas of banking operations. It influences the economic capital required to support the loans which involves stress testing, the development of risk metrics, and the estimation of loan-loss reserves on bank financial statements [50]. Many smaller banks struggle with these calculations as it's hard to quantify LGD when the bank does not have many defaulted loans, and especially also if there are difficulties in measuring historical LGD for those loans [50].

Because credit models have focused on systematic components of credit risk which attract risk premia, the major focus on PD is understandable as the determinants of LGD have been ascribed to idiosyncratic borrower specific factors [42].

1.2.3 Exposure at Default (EAD)

Exposure at default is defined as the total value a bank is exposed to at the time of a loan's default [5]. The estimation of EAD corresponds to the terms of payment, period of exposure as well as the point of time at which default is expected or occurs [31]. For defaulted accounts, EAD is simply the outstanding amount at the point of default [23].

On the other hand, the computation of EAD for performing accounts under IFRS 9, a number of elements are needed at instrument level [31]. The first element is the time period over which we want to estimate the EAD, followed by the cash flows that are projected until the estimated point of default, and then the residual maturity as well as the the nature of the terms of payment which can be deterministic or non-deterministic. Finally, last element needed is the forward looking macroeconomic factors [31].

EAD is relatively simpler to compute for shorter time horizons (12 months) i.e. phase 1, under IFRS 9. However, for longer time horizons that incorporate forward looking information i.e. phase 2 under IFRS 9, the computation of EAD requires us to simultaneously consider the residual maturity, the uncertainty of cash flow and the prediction of default [34].Prepayment, draw down factor or credit conversion factor (CCF) are some of the forward looking factors and they vary depending on the type of instrument and also depend on the terms of contract as well as the financial institution's business practices [31].

1.3 Macroeconomic variables

As mentioned above in the previous subsection and in the title of this paper, risk parameters are influenced by the macroeconomic environment. Factors like gross domestic product (GDP), inflation and market interest rate have been identified as having significant impact on credit risk.

Also other macroeconomic variables such as the required reserve, the discount rate, inflation, government borrowing, treasury bill and the gross domestic product per capita (GDPPC) can be considered [34]. The level of economic development, cyclical output and stock market capitalization can also be considered as macroeconomic variables [34].

The macro economy is important as it has been observed that lending mistakes of banks are prevalent during upturns than in periods of recession. Favourable macroeconomic conditions impact performance measures positively and thus conditions associated with good economic periods assist in decreasing the credit risk exposure of banks [34].

More specifically, relating to PD, two PD estimates known as TTC (through the cycle) and PIT (point in time) PDs respectively, resulted from the macroeconomic conditions impact [49]. The TTC PD is not significantly affected by the state of the economy, whilst the PIT PD estimate changes with the economic cycle [49].

There are challenges in developing the methodology for the forecasting macroeconomic conditions. IFRS 9 states two flexibilites to reduce these challenges [24]:

- If there are valid justifications, then judgemental override on long-term macroeconomic forecasts is allowed [33].
- We can generate multiple forward looking scenarios and then use them to generate multiple PIT PD estimates for a facility, instead of using only one forward looking scenario.

Though these flexibilities assist in mitigating the long-term macroeconomic forecast challenges, it will still remain a big obstacle to establish a stable relationship between the PIT PD and macroeconomic conditions [24].

In this paper we will use the macroeconomic factors for which the market information and data are accessible, which are the South African GDP growth rate, unemployment rate as well as the consumer price index (CPI).

2 Literature Review

2.1 Early Credit risk models

Credit risk models can be divided into two categories namely, structural and reduced form models. The aim of structural models is to provide an explicit relationship between the credit structure and default events, whereas the reduced form models model the defaults as exogenous events that are driven by a stochastic process [49]. These two models differ in the methods used to incorporate the effect of the macroeconomic environment on PD.

Structural models measure the cyclical impact on PD. They do this by incorporating systematic risk factors into the specification of the process driving the variation in the values of the firm's assets [49], where systematic risk factors are macro factors which influence the direction and volatility of the entire market [35]. Structural models are also based on the framework in Merton (1974). Reduced form models on the other hand model default events as a Poisson process with time varying default intensity, where the intensity function partly comprises of the macroeconomic variables. They introduce separate assumptions on the dynamic of PD and the recovery rate, which are modeled independently from the structural features of the firm [26].

Several models have been proposed in literature that incorporate the impact of the macroeconomic environment on the PD. Below we discuss some of these models. They differ from the IFRS 9 credit loss model in that they focus on one year horizons. However, since credit exposures have maturities that are greater than a year, analyses over longer periods are needed in order to ensure proper credit risk pricing and management, which is what the IFRS 9 framework does.

Jarrow–Turnbull model (1995)

This model is referred to as the first reduced-form credit model [6].It is an extension of the Merton (1976) reduced-form model to a random interest rates framework. In order to calculate the probability of default, this model uses multi-factor as well as dynamic analysis of interest rates [7]. It works in a complete market setting or economy that's frictionless and arbitrage-free, and where credit risk classes are allocated to firms [51]. In other words, it uses an arbitrage-free default/bankruptcy process that triggers default. This frictionless economy has a finite time horizon $[0,\tau]$, in which trading can be discrete or continuous. Two types of zero-coupon bonds are traded in this economy namely, default-free and default risky zero coupon bonds.

The default-free zero coupon bonds guarantee a $R1$ payment at their maturity time T i.e. $P(T,T) = 1$ and have a price $P(t,T)$ at time t [46]. The default risky zero coupon bonds on the other hand, promise to a payment of $R1$ but may not be paid in full at their maturity time [46]. If they only pay $\delta < 1$ Rands, δ is known as the recovery rate and it is a fraction [55]. In this model, this recovery rate is a constant that is exogenously given. The constancy of δ is in order to simplify estimation, which means that credit spreads stochastic term structure will be independent of the recovery rate and thus dependent on the spot interest rates stochastic term structure and the default/bankruptcy process only [55].

Wilson (1997 a,b)

This is one of the first credit risk models that were developed that link macroeconomic conditions to PD. This is a reduced form model that explicitly links the impact of the economic state to rating transitions. A rating transition matrix consists of probabilities of moving from one rating class to another, where the last column contains PDs [38].

This model uses a logit model to estimate the PD conditional on the economic state [38]. Then as a proxy variable for the systemic risk factor a macroeconomic index is constructed that is defined as an autoregressive process [38]. This constructed index is built on several macroeconomic variables. Finally empirical application of this model shows that the PD contains cyclical movements, particularly in periods of recession when the PD increases significantly [38].

Belkin, Forest and Suchower (1998)

This model employs the Merton framework to derive a standard normally distributed cycle index and proposes a method to estimate transition probabilities based on this index [38]. This model is also referred to as the one-factor Merton model. The cycle index is constructed to be negative in periods of contraction and positive during expansions.

More specifically, the cycle index is positive when two things happen, namely when the default rate is lower than average default rate and also when the ratio of upgrades to downgrades is higher than average. In periods of contraction the reverse is true when the cycle index is negative [38]. The estimation of the credit cycle index is based on default rates that are provided by Moody's as well as Standard and Poor's.

Kim (1999)

This model is inspired by those of Belkin, Forest and Suchower (1998) and Wilson (1997a,b) in that it applies an ordered probit model, which is a model used to estimate relationships between an ordinal dependent variable and a set of independent variables[26], whereby an ordinal variable is a variable that is categorical and ordered [38].

This ordered probit model is then used to estimate migration probabilities, also known as transition probabilities, conditional on the economic state [38]. This model also creates a credit cycle index that is based on macroeconomic variables such as real GDP growth, unemployment and interest rates.

Nickell, Perraudin and Varotto (2000)

This model is an extension of the ordered probit model introduced by Kim (1999). The extension is incorporated in the fact that the transition probabilities are permitted to be conditional also on the industry and country domicile and not only the economic cycle [38].

Also in this model the economic state variable is modelled as a discrete variable instead of a continuous macroeconomic index that is used in Kim (1999). In this model GDP growth is used to divide the economic state into three states namely, high, medium and low growth periods. For these three different stages of the economy, the transition matrices are estimated separately [38].

By observation, it has been found that the economic cycle variable is the most important factor that explains the variation in transition probabilities, particularly in the movements of PD.

Bangia et al. (2002)

This model proposes a method to estimate transition matrices that are conditional on discrete economic states. However, the difference from Nickell, Perraudin and Varotto (2000) is that in order to derive the systematic shifts in transition probabilities it employs a structural approach [38]. It divides the economic cycle into two stages namely, recession and expansion. It does this by using the classification that is provided by The National Bureau of Economic Research of the economic state of the US.

This division is done by firstly estimating the transition matrix, unconditional of the economic state from the rating statistics that are provided by Standard and Poor's. After this estimation a separate estimation is made for the two economic stages.

So to examine whether the expansion and recession matrices differ, they are compared to the unconditional migration/ transition matrix that was estimated in the beginning [38]. The prominent distinction between these matrices is the PDs, they increase significantly when there is a recession [26].

Koopman, Lucas and Klassen (2005)

This model uses an interesting approach to model the systematic risk factor. It employs macroeconomic tools such as the HP-filter to extract a latent economic cycle component [38]. An HP-filter is defined as a mathematical tool used in macroeconomics, especially in real business cycle theory, to remove the cyclical component of a time series from raw data [38].

Thus this model decomposes U.S business failure into two components namely, an autoregressive and time-varying cycle component [38]. These two components are then interpreted as systematic risk factors. These components are primarily studied to analyse whether cyclical movements exist and ,if they do, whether they correspond to economic cycles [38].

Empirical results from this model show that the default rate is comprised of long-run movements and cycle pattern. Also it is observed that the extracted credit cycle has similar length to the economic cycles.

2.2 Different modeling methodologies in loss forecasting

There are several modeling practices that are used across industries for loss forecasting. Below we discuss some these forecasting techniques that are well-known and their shortfalls compared to the expected loss model that we will be using in this paper.

Roll Rate models

This loss forecasting is done at a portfolio rather than on an account level. The whole portfolio amount is divided across various "buckets". Using current balance examples of these buckets are Current, 1-30 DPD, 31-60 DPD, 61-90 DPD, 91-120 DPD, where DPD stands for days past due. This roll-rate technique produces a forecast in which the movement of outstanding amounts from a lower level to a higher level of delinquency is applied to the current portfolio outstanding mix [41].

Thus it moves from the current balance through all the buckets of delinquency to "charge off". After the historical net roll rates have been calculated by bucket, then their patterns are examined overtime and the future roll rates are estimated. The product of flow rates from the bucket to final 180+ bucket gives the losses [41]. However, limitations of this technique is that it's particularly suited for retail portfolios, it doesn't consider loan specific information and it has heavy assumptions for long term estimations, thus it's used to forecast.

Vintage Loss Models

Under this forecasting modeling technique that also widely used, the portfolio is divided into origination vintages and not delinquency buckets as the roll rate models. For each of the vintages, the loss rate can be tracked overtime through the full life cycle [41]. The vintage model had two advantages. Firstly, the loss trend can be forecast for a longer term, by incorporating the effect of maturation and secondly, to incorporate the current and future market movements, the model takes the economic factors into account. However, it is suited for retail portfolios like the roll rate model, which is one of it's limitations. Another limitation that it doesn't consider loan specific information.

Provision Matrix Model

The provision matrix is also delinquency based model just like the roll rate model. This model is used to estimate bad debt reserves, which are the amount of receivables that the company does not expect to actually collect [8], for short term trade receivables. Usually, as practiced by many entities, when the provision matrix is used to estimate credit losses, the trade receivables are categorized on the basis of different customer qualities as well as different historical loss patterns such as, customer rating, product type , geographical region etc [41]. However, under IFRS 9, historical provision rates need to be updated with current and forward looking macroeconomic factors and thus has the same limitations as roll rate models that we mentioned above.

Discount Cash Flow Method (DCF)

This method based on the concept of the time value. It discounts all the estimated future cash flows. Thus to reach their present values appropriate effective interest is used [41]. This method however is only ideal for the loss estimation of stage 3 instruments. Hence, it's limitation is that it is difficult to implement for a large number of instruments in the banking book [41].

Expected Loss Models (ELM)

These models differ significantly from roll rate and vintage loss models. As mentioned in one of our previous sections, expected loss estimation is based on three risk parameters, namely probability of default (PD), exposure at default (EAD) as well as Loss Given Default (LGD) by incorporating characteristics that are specific to the loan. The ELM can be used for both retail and corporate/ wholesale portfolios. This methodology models each of the three risk factors separately in order to capture the specific behavior of the account or cohort.

In order to directly predict the "time to default" for a loan, survival models can be used [24]. Based on origination variables, the customers are divided across homogeneous groups in such a way that for each of the groups, the historical bad rates across time never intersect each other. Thus the average PD, LGD and EAD for each of these groups is then used to calculate expected losses. One main limitation of this model is that it is heavy on data requirements i.e. it requires large amounts of data. In this project we will be using the ELM methodology.

According to IFRS 9, the estimation of losses shouldn't only incorporate past due information but also all the current and future credit information that is relevant, including forward looking macroeconomic forecasts, irrespective of the choice of the methodology, because of this reason, one can say that the ELM methodology is the best [24].

This ELM approach has four benefits. Firstly, the fact that modeling is done on an instrument level, provides a finer risk profiling for each individual or homogeneous pool of instruments. Secondly, there is a more dynamic view of economic impact, because each risk parameter is driven independently by a separate set of economic factors. Thirdly, this modeling methodology is in line with prevailing statistical techniques in the consumer lending arena and can intuitively be adapted by most model developers. Finally, the lifetime PD that incorporates the forward looking macroeconomic factors can be directly used for the significant increase of credit risk assessment [24].

However, since IFRS 9 is a principle based guideline, it doesn't prescribe a specific methodology for the estimation of lifetime expected loss nor is there a single methodology that suits all portfolios [24]. Thus, the choice of the methodology should be one that is informed and based on data availability and the quality of the portfolio. The table below shows a summary of some of the most used approaches (some that we mentioned above) by different institutions to calculate EL [45] under IFRS 9:

(I) HISTORICAL AVERAGE APPROACH	**(II) DYNAMIC MIGRATION APPROACH**
-Long run average is: Long run PD*LGD*EAD -This approach considers the PD long term structure and doesn't capture the macro – sensitivity. -Macro-sensitivity can be included as qualitative overlay. Long Run PD → LGD*EAD → Un-weighted EL	-Historical data is gathered on PD migration across rating buckets. -This is used to estimate and forecast transitions probability matrices. -This is then used to forecast PD for each year of loan tenor, and hence EL for each year of loan tenor. -This approach is most granular and requires high quality data, also most transparent. Transition matrices → PD forecasts for defined periods → EL computation/ discounting
(III) HAZARD APPROACH	**(IV) REGRESSION MODEL BASED APPROACH**
-A hazard function approach is used to provide estimates of PD conditional on survival i.e. Lifetime PD. A time dependent PD well described by a Weilbull-Pareto distribution function. -Combination with EAD and LGD estimates at default to obtain ECL. -This methodology calculates the Lifetime PD at the loan level, however the calculation can be allocated to the account level. Hazard function → Future lifetime PD → EL computation /discounting	-This model is based on an aggregated default rate that is regressed on various factors, including economic variables. -The regression model incorporates the economic impact into the lifetime PD. -The forecast results are used with the regression model to obtain the forward looking PD term structure. Regression models → Macroeconomic Forecasts → Forward looking values

Figure 5: Different EL modelling approaches. Adapted from [45].

2.3 Modelling of risk parameters

Many models have been developed that can be used to construct a term structure of PDs and LGDs, as well as modelling EAD whilst incorporating macroeconomic factors. In this thesis we will use some of these models to construct a term structure of the credit risk parameters, under IFRS 9.

When measuring expected credit losses, the probability- weighted outcome should be considered, as well as the time value of money and reasonable information that contributes to this calculation. Thus IFRS 9 is a big change for banks and many other financial institutions as now they don't only have to recognize credit losses that have already occurred but also those that are expected to occur in the future [19]. This will thus help to ensure that these entities are properly capitalized for the loans that have been created.

Probability of default modelling

In order to estimate the PIT PD term structure there are two types of methodologies that can be used:

(i) Markov models and transition/ migration matrices

Two broad approaches are used to estimate transition matrices:

- The Cohort approach

  The entries of the transition matrix are given by:

$$P_{ij}(t) = \frac{\sum_{t=1}^{T} N_{ij}(t)}{\sum_{t=1}^{T} N_i(t)} \tag{2}$$

  where:

  - $N_{ij}(t)$ is the number of firms that migrated to j from i by the end of year t

  - $N_i(t)$ is the number of firms in the category i at the beginning of year t

The problem with this approach is that any movements within year t (e.g. within 2 months of the year) are not accounted for, only the period, e.g. one year (t), movements are considered.

– The two variants of duration (Hazard) approach

This approach can be parametric (time-homogeneous) or non-parametric (time- inhomogeneous). It counts all the changes over the course of the year or within the multiple period and divides by the number of firm years spent in each state, in order to obtain a matrix of migration intensities, which are assumed to be time-homogeneous.

This matrix is also known as a generator/ intensity matrix, Γ, and consists of transition intensities given by:

$$\delta_{ij} = \frac{n_{ij}(T)}{\int_0^T Y_i(s)ds} \tag{3}$$

where:

- n_{ij} is the total number of transitions over the period from i to j where $i \neq j$

- $Y_i(s)$ is the number of firms with rating i at time s

The t-year transition matrix is thus given by:

$$P(t) = exp(\Gamma t) \tag{4}$$

(ii) Merton-type model

According to Merton (1974), a borrower defaults if the value of its assets fall below the amount borrowed. This means that PD is the probability that the asset value is below the debt value, d_i, at some given point in time. The one-factor Merton model is built on asset returns, r_i, rather than the asset value itself. Thus the PIT PD is given by:

$$PD_i(z) = P(r_i < d_i | Z = z) = \Phi\left(\frac{d_i - \sqrt{\rho_i} z)}{\sqrt{1-\rho_i}}\right) \tag{5}$$

where:

$$r_i = \sqrt{\rho i} Z + \sqrt{1-\rho_i}\varepsilon_i, \qquad d_i = \phi^{-1}(PD_i) \tag{6}$$

and PD_i is the TTC PD, ρ is the correlation between the asset return and the systematic risk factor , Z is the systematic risk factor and ε is the idiosyncratic risk factor.

Below we discuss two methodologies to the two types of approaches that we have defined above that can be used to develop the term structure of PDs, namely the mapping to external rating agency term structure which falls under the Merton-type model approach, as well as the Markov chain and rating transition approach which falls under the Markov models and transition/ migration matrices approach. We will briefly describe each methodology and select the approach that best suits the data we have:

The mapping to external rating agency term structure

This approach requires banks to map their internal rating grades with those of global rating agencies and then adopt the PD term structures that correspond to the rating grades that are benchmarked with global rating agencies [24].

Since global rating agencies follow a TTC rating philosophy, if the bank's internal rating model also follows a TTC philosophy, then this is a suitable approach to create a TTC PD term structure. However, if the bank's internal rating model follows a PIT rating philosophy, then the mapping between the bank's interbank rating grade and external rating grade will need to be conducted regularly, at least once a year, in order to make it stable over time and avoid over- or under-estimation of the TTC PD term structure, depending on the current state of the business cycle at the mapping.

Thus for IFRS 9 purposes, before banks use the external rating agency's TTC PD term structure, they have to convert the rating agency TTC PD term structure into the PIT PD term structure and then incorporate macroeconomic conditions. One such way involves two steps:

(a) Use historical default rates in the Nelson-Siegel function to estimate the TTC PD

(b) Apply the one-factor Merton model and transform the TTC PD term structure to a PIT PD term structure with macroeconomic adjustments.

However, since the global rating agency's TTC PD term structure is mostly based on the default experience of large publicly rated borrowers (such as firms), that are also mainly located in developed countries, the application of this approach to a portfolio consisting of SME (Small and Medium Enterprise) borrowers as well as borrowers from emerging market economies may not be accurate [24].

The Markov chain and rating transition approach

This approach applies matrix multiplication techniques to develop the PIT PD term structure. Banks may estimate the effect that macroeconomic factors have on the transition matrix in the first year and then use it to create the PIT PD term structure. To generate the PIT PD term using the one-year rating transition matrix information is not valid, as seen in the Binomial approach.

The most important task with regards to the estimation of PD under IFRS 9 is going to be the estimation for more than one period ahead, which is usually referred to as a year, that takes into account the economic forecast. We will assume that the one year PD estimates are available, thus our work will be to extend them over more future time periods and to incorporate future expected economic development.

The multi-period PD estimation can be performed with Markov models. A Markov model is stochastic model that describes a sequence of possible events in which the probability of each event depends on the state attained in the previous event only. Markov models are based on the Markov chains theory. A Markov chain is a discrete random process that starts in a certain state i and moves to state j in the next step with transition probability p_{ij}. Thus the probability that the process will be in a certain state at time $t+1$ will depend only on the state the process is in at time t, and not on the previous states [27]. This process can be shown by the following equation:

$$\begin{aligned} P(X_{t+1} = j | X_0 = i_0, ... X_{t-1} = i_{t-1}, X_t = i) &= P(X_{t+1} = j | X_t = i) \\ &= p_{i,j} \end{aligned} \tag{7}$$

If p_{ij} doesn't depend on t we have a time-homogeneous Markov chain. However, since IFRS 9 requires that future economic conditions be taken into account, the above equation wouldn't be acceptable without further adjustment [27]. Thus the PD would be underestimated, which means time-homogeneity must be incorporated. This means the Markovian property above can be rewritten as:

$$\begin{aligned} P(X_{t+1} = j | X_0 = i_0, ... X_{t-1} = i_{t-1}, X_t = i) &= P(X_{t+1} = j | X_t = i) \\ &= p_{i,j}(t) \end{aligned} \tag{8}$$

The above equation now shows the dependency of the transition probabilities on t. In order to incorporate the economic forecast we will use the decomposition of the "economic adjustment coefficient", which we will estimate by linear regression. The transition matrix is the basis for PD estimation for more than one period ahead [27]. The transition matrix consists transition probabilities between individual states. These transition probabilities can be estimated using either the Cohort or Hazard method described above.

We will use the computational framework used in [27] to extend the one period PD. This computational framework assumes that the default state is absorbing, meaning should an obligor default they cannot recover. It also omits the assumptions that PDs tend to decrease over time.

The outline of this computational framework is as follows:

(i) Obtain the one period or one year PD, in this case the one year transitional matrix, **T**.

(ii) Estimate the economic adjustment coefficient, using a simple linear regression model, which captures the impact of future expected economic development i.e macroeconomic factors on the PD. The general form of such a simple linear regression model is given by:

$$Y = \sum_{i=1}^{n} \beta_i X_i \tag{9}$$

where Y is the dependent variable, which in our case would be the PD proxy, the β's are the coefficients which we refer to as the EACs (Economic Adjustment Coefficients) and the X_i's are the independent variables, which in our case are the macroeconomic variables.

For instance if we made the GDP (gross domestic product) as the macroeconomic factor and a share of non-performing loans (NPL) as a proxy for the PD, then we would have a simple regression model given by:

$$NPL = \alpha + \beta . GDP \tag{10}$$

(iii) Modify or adjust the one year PDs or transition matrix for the future time periods using the decomposing effect in (ii).

(iv) We then calculate the multi-period PDs using the concept of Markov models.

When this logic is used the effects of idiosyncratic risk (specific to a particular group) and systemic risk (influences clients as whole) are more separated [27]. There are three modern or popular approaches to incorporating macroeconomic factors.
Below we use an approach based on Markov models and transition matrices. There are other two approaches that are based on the Merton-type models and the survival analysis framework, which we will explain later which also ensure that the PD estimate is as accurate as possible as IFRS 9 requires these PDs to be lifetime, forward-looking and point in time [39].

Example[4]

To illustrate this methodology, we show a simple example of calculation the multiperiod PD taken from [27]:

Suppose we have a rating system with two grades:

(i) Non-default state

(ii) Default state

Also assume that a probability of default of 4% for a one year time horizon has been assigned at time. Which means we have a state-vector given by $s_t = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and initial transition matrix:

$$Q_{t+1} = \begin{bmatrix} 0.96 & 0.04 \\ 0 & 1 \end{bmatrix} \tag{11}$$

Thus the three year PD estimate would be:

$$\begin{aligned} s_{t+3} &= s_t \cdot Q_{t+1}^3 \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \begin{bmatrix} 0.96 & 0.04 \\ 0 & 1 \end{bmatrix}^3 \\ &= \begin{pmatrix} 0.8847 & 0.1153 \end{pmatrix} \end{aligned} \tag{12}$$

$\Rightarrow$ The 3-year PD estimate is 11.53%

We then use a simple *linear regression model* to capture the impact of the economic development on PD i.e economic adjustment coefficient [27]. We have to choose proxies for the PD and the macroeconomic factors.

[4]Example taken from [27]

For simplicity we'll choose the share of non-performing loans (NPL) and gross domestic product (GDP), to represent PD and macroeconomic factors respectively.

$$NPL = \alpha + \beta \cdot GDP \tag{13}$$

where β is the economic adjustment coefficient. In our exampleβ was estimated to be -0.233, meaning that if GDP increases by one unit, NPL decrease by by 0.233%

We then use ordinary least squares (OLS) method to estimate the parameters. Once β is found then the three year PD estimate, $PD(t+3)$, is performed within the Markov models framework [1]. Thus the transition matrix needs to be adjusted in ***every time period*** by the expected economic conditions. Now:

$$s_{t+n} = s_t \cdot \prod_{k=1}^{\infty} Q_{t+k} \tag{14}$$

In our example we now have:

$$s_{t+3} = s_t \cdot Q_{t+1} \cdot Q_{t+2} \cdot Q_{t+3} \tag{15}$$

Instead of:

$$s_{t+3} = s_t \cdot Q_{t+1}^3 \tag{16}$$

The incorporation of the macroeconomic factors is done the following way:

$$s_{t+n} = s_t \cdot \prod_{k=1}^{\infty} Q_{t+k} = s_t \cdot \prod_{k=1}^{\infty} \begin{bmatrix} q_{1,1} - \delta_{t+k} \cdot \beta & q_{1,2} + \delta_{t+k} \cdot \beta \\ 0 & 1 \end{bmatrix} \tag{17}$$

where δ_{t+k} is the differences between the present economic conditions and the future economic conditions.

The reason we have a minus sign in front of $q_{1,1}$ and plus sign in front of the $q_{1,2}$, in equation 16 above is so that the row sum may be equal to one as $-\delta_{t+k} \cdot \beta$ would cancel with $+\delta_{t+k} \cdot \beta$, and we would expect $q_{1,1}$ and $q_{1,2}$ to sum to one. So:

$$
\begin{aligned}
s_{t+3} &= \begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \begin{bmatrix} 0.9576 & 0.0424 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0.9590 & 0.0410 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0.9591 & 0.0409 \\ 0 & 1 \end{bmatrix} \\
&= \begin{pmatrix} 0.8808 & 0.1192 \end{pmatrix}
\end{aligned}
\tag{18}
$$

Thus we can see the increase of the PD from 11.52% to the adjusted value, 11.92%, after incorporation of macroeconomic factors. The higher estimated PDs means that higher expected credit losses will be recognized. This means that banks will create higher "reserves" and therefore be better prepared to cover potential losses coming from expected macroeconomic factors [27]. If the economic adjustment coefficient was positive, we'd have the opposite effect. This means that the estimated PD would be lower.

The example above was only based on two rating grades, however this framework can be generalized or extended to incorporate an arbitrary number of rating grades. For instance we can have 6 rating grades, where the lower grade represents a better rating and the last grade represents the default [27].

Another issue that can be noted from the above example is that the economic adjustment coefficient (EAC) was only decomposed into two parts, where each part used for the adjustment of the non-default and default rating grade respectively. This approach can also be used in a more generalized manner depending which direction is desired to be emphasized of the adjustment [27].

The diagram below illustrates the PD modeling process:

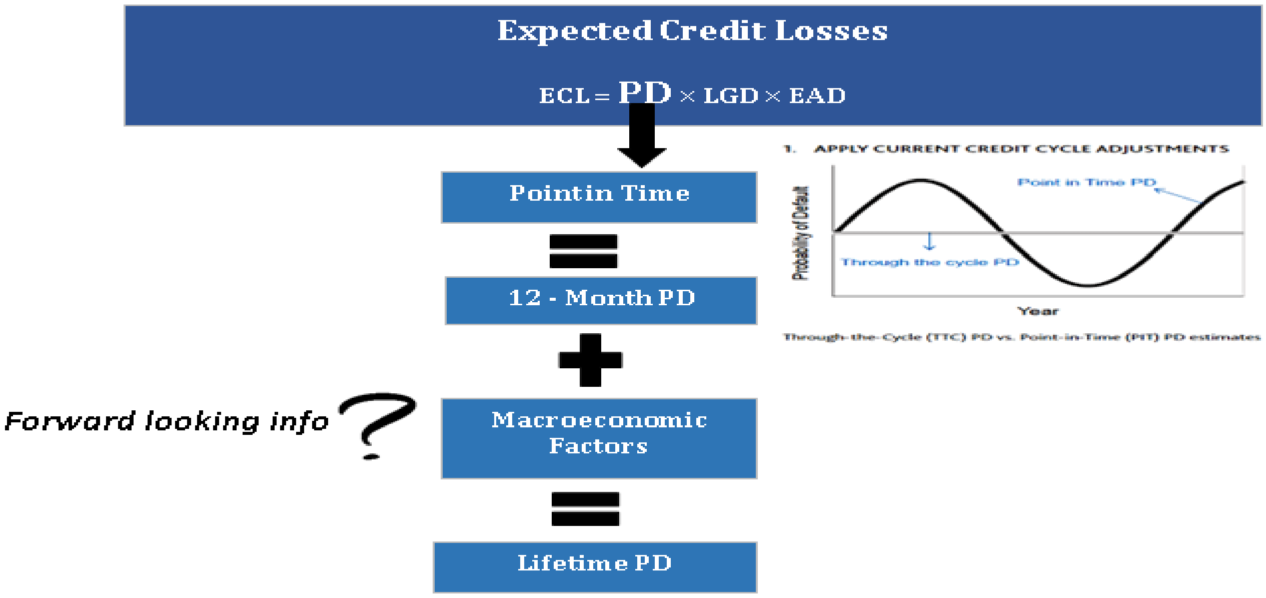


Figure 6: Summary of estimating lifetime PD. Adapted from [43].

There are also two key approaches for incorporating macroeconomic adjustments for PD, we list and briefly discuss them below:

Macroeconomic adjustments of rating grade migration approach

Under this approach, relations between one-year transition probabilities of each rating grade and macroeconomic factors are established:

$$P_{ij} = f(macroeconomic factors) \tag{19}$$

where:

P_{ij} is the transition probability from a rating grade i to rating grade j within a one year time period

When the internal rating model of the bank is developed based on the PIT rating philosophy, both Markov chain and macroeconomic adjustments of rating grade migration methodologies are more applicable. This is because under the PIT rating philosophy, rating models consist of both idiosyncratic and macroeconomic factors.

Thus rating grades created by these models will change as a result of macroeconomic factors and hence the migration probability of borrowers from one rating grade to another will change depending on the macroeconomic factors i.e. the business cycle.

Markov chain approach

Under this approach a linkage is formed between the one-year transition matrix and the forward looking macroeconomic factors. Historical transition matrices are converted into z-scores. A z-score indicates how many standard deviations an element is from the mean. These z-scores are then linked with forward looking macroeconomic factors.

The entire transition matrix is converted into a single dependent variable and the effect of macroeconomic factors is assessed based on this variable. Thus the effect of migration of the macroeconomic factors will be the same across the rating grades.

Banks can use any of the two approaches mentioned above for macroeconomic adjustments of the PIT PD under IFRS 9, depending on the rating philosophy they follow for the development of their internal rating model.

Loss given default modelling

An accurate forecast of the LGD of loans parameter plays an important role in risk-based decision making process by banks. Thus banks using LGD models that have higher prediction power can generate competitive advantages, whereas weak predictions can lead to adverse selection [50].

The estimation of LGDs for banks is based on the discounted recovering cashflows [37]. There are different routes of LGD related literature. Some studies estimate the distribution of LGDs for credit

portfolio modelling, whilst others analyse factors affecting LGDs, and others deal with the relation between PDs and LGDs [37].
IFRS 9 requires LGDs to be lifetime for Stage 2, best estimate (e.g. no downturn bias, collateral, limits, etc.), forward looking and to include direct costs only [39].

Exposure at default modelling
Exposure-at-default (EAD) is said to be one of the most interesting and most difficult parameters to estimate in counterparty credit risk [56]. Understanding how loan exposures are expected to change over time is crucial to an unbiased measurement of ECLs (Expected credit losses), although IFRS 9 does not explicitly require banks to model EAD [56].

The modelling approach for EAD reflects expected changes in the outstanding balance over the lifetime of the loan exposure that are allowed by the current contractual conditions [53]. IFRS 9 requires EADs to be point in time expected credit exposure and to be the best estimate for drawn and undrawn amount [39].

2.4 TTC (Through-the-cycle) vs. PIT (Point-in-Time)

There is no consensus about the true meaning of the two concepts, TTC and PIT credit rating systems, though they are widely used in banks, by credit rating agencies as well as supervisors [18]. However, consensus only exists in the sense that PIT credit ratings are defined to use all information, as well as both obligor-specific features and macroeconomic conditions available at a particular point in time [49], whereas TTC credit ratings are defined as being adjusted for cyclical effects in macroeconomic conditions [49].

Since financial market participants need accurate measures about the ability of an obligor to fulfil their future financial obligations, creditworthiness plays a major role. The creditworthiness is usually characterised by a credit rating, or credit score, which is typically associated with a PD [18]. As mentioned above PD describes the likelihood that a firm will experience a credit event (unlikeliness to pay, bankruptcy, payment delay, etc.) over a given time period [49].

Thus the different approaches in which PDs can be estimated can be categorized into their PIT vs. TTC orientation. We can see from the consensus definition of PIT and TTC credit ratings that the PIT ratings provide the most accurate as well as timely PD estimates [49]. On the other hand, TTC ratings provide more stability overtime, but this stability comes at the cost of reduced timeliness and reduced accuracy in PD prediction [49].

Under the IAS 39 framework PDs were estimated as TTC. However, under the IFRS 9 framework the PD estimates are PIT and considered more accurate. In other words, under the old regulatory framework standards (IAS 39), the PDs were neutral to the economic cycle [49]. This means they were less sensitive to changes in economic conditions and thus reflected longer-term trends in PD behaviour compared to the PIT PDs,as can be seen in the figure below:

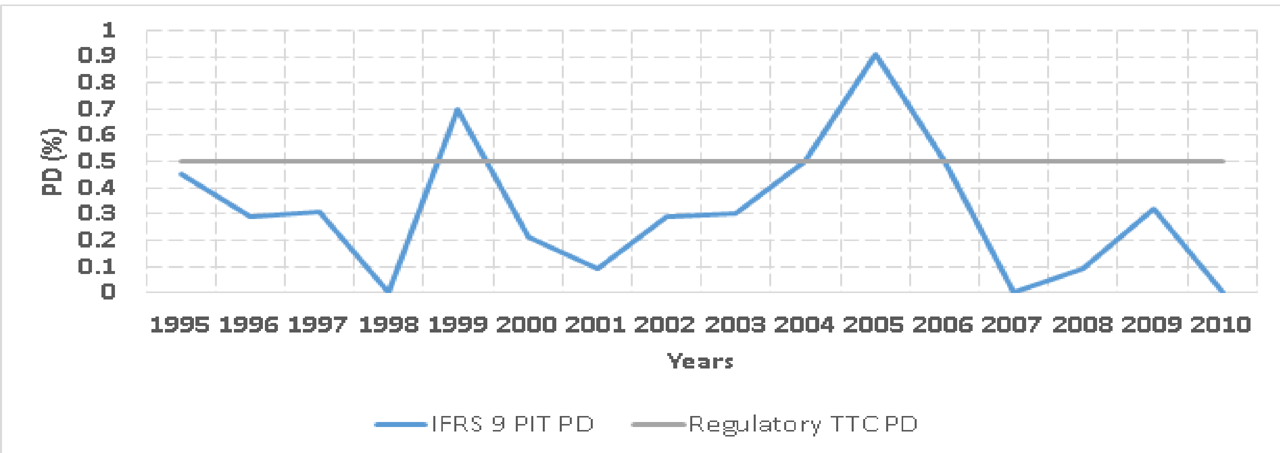


Figure 7: PIT vs. TTC PD estimates. Adapted from [40].

Now since under IFRS 9, the PDs are required to be PIT it means we are looking at PDs in the current economic condition [9]. Thus in this case, as an entity moves through the economic cycle, the PDs must change [9]. This means that during a good credit environment, the IFRS 9 PIT PD will actually be lower than the TTC PD. The opposite would occur during a financial crisis, meaning the PIT PD would be higher than the TTC PD. The figure below [5] illustrates this:

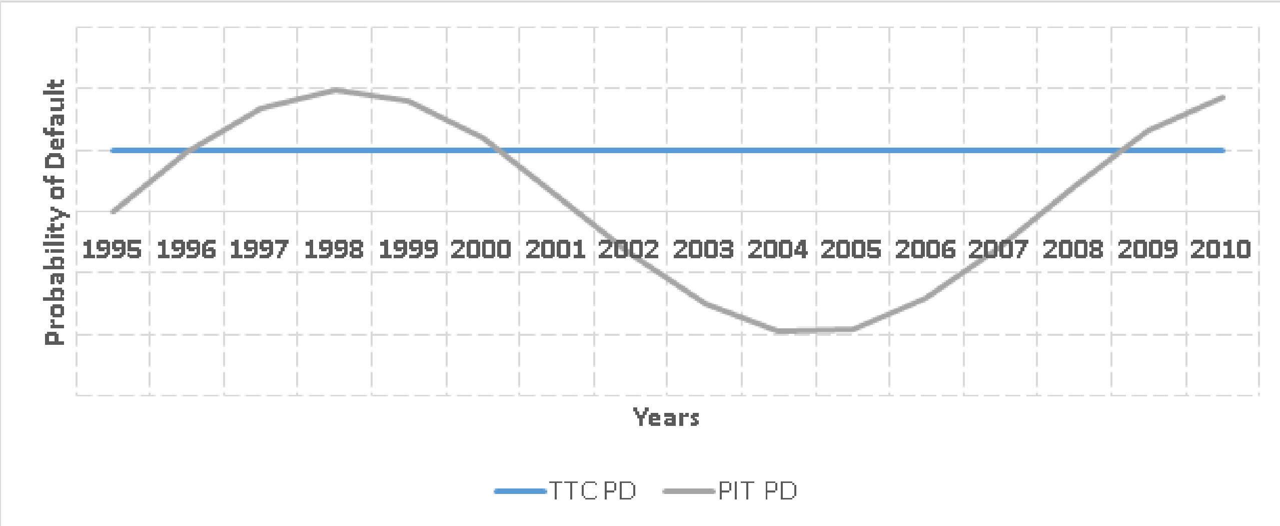


Figure 8: PIT vs. TTC PD estimates with credit cycle adjustment

2.5 IAS 39 vs. IFRS 9

The IAS 39 framework was one of the financial crisis triggers. Thus an improvement of the standard for financial instruments had to be proposed with an aim to increase financial stability. This improvement had to take into account that the standard already in place for financial instruments is complex , the extent to which the financial instrument is subject to fair value, as well as the procedure of recognition and measurement of financial instruments [35].

The accounting under IAS 39 is said to be based on rules, while accounting under IFRS 9 is based on principles [35]. This rules basis approach of IAS 39 allows the decision makers to make decisions that

[5]Note: Values and dates on this figure are random and not real. They are only for illustrative purposes.

are more predictable and stable in an unstable environment. However a shortfall of this rules- based approach is that these rules don't adapt and are actually useless in an environment that has innovative transactions [35]. The IFRS 9 principles approach also has a shortfall in that it lacks operational guidance.

The most significant change that comes from the replacement of IAS 39 with IFRS 9, is the introduction of the expected credit losses model that allows for timely or early recognition of losses that are inevitable, especially in banks. In addition to this significant change, it improves financial reporting in the debt instrument field. The changes in accounting policies brought by the impairment of financial instruments are different but significant and are based on the futures losses model, which gives stakeholders more insight to financial instruments that have increased in credit risk [35].

At the time of implementation of IFRS 9, the costs incurred are high but the benefits actually outweigh the costs at the end [6]. Under IAS 39, several impairment models based on incurred losses are used. However, under IFRS 9, a unified model of impairment for all financial instruments is used, known as the expected losses model [35].

To understand IFRS 9 more, we look at a SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis[7] for it:

Strengths

Other than the fact that IFRS 9 addresses issues that arose from the financial crisis, it first of all reduces the complexity of the classification and measurement of financial instruments. Secondly, it focuses on shareholders and the accounting is in line with business strategy. Finally, IFRS 9 improves reporting consistency and transparency with global rivals and is able to detect losses properly [35].

[6] Marshall (2015, p.1)
[7] Hulan (2012, p.42)

Weaknesses

The first disadvantage of IFRS 9 is that new concepts are introduced based on the business model, that require judgement that is professional and hence can lead to subjectivity. Secondly, a systematic approach for financial liabilities is not provided and hence it doesn't solve impairment of hedge accounting questions [35]. Finally, the fact that we also have to adjust or upgrade the accounting systems already in place to new calculations for IFRS 9 is also disadvantageous[8] [35].

Opportunities

IFRS 9 allows firstly for professional judgement when it comes to accounting decisions. Secondly, it allows for better choices by the standard setter as a result of a slower staging of the completion of the second stages [35].

Threats

Firstly, there's a possibility that IFRS 9 will apply only to the stock-exchange listed organizations, while about 700 000 small and medium organizations are still using the national accounting standards [35]. Thus the earlier adoption of standard by some organizations means that there will be a display of both standards in presentations and disclosures, which makes financial statements less useful.

Secondly, an approach that contains multiple stages may create mismatches due to new requirements as well as other existing rules [35]. Thirdly, a lot of tolerance on several topics may result in choosing a certain option only to meet accounting requirements [25]. And finally, the cost of implementation is relatively not easy to measure [35].

Although the replacement of IAS 39 with IFRS 9 took place in 3 phases namely, the classification and measurement of financial instruments, the impairment methodology and hedge accounting, in this thesis we will focus on the impairment methodology phase of financial instruments. The impairment of a financial instrument is defined as the adjustment of the prices in the financial statements with the

[8](Ghasmi, 2016, pp. 30, 31)

prices on the markets as well as market conditions [35].

As mentioned above, the impairment of financial instruments under IAS 39 is based on incurred losses which is what caused problems that initiated the financial crisis due to a delay in the recognizing of impairment and losses. However, IFRS 9 introduces impairment based on expected losses. Thus, this impairment model under IFRS 9 is in essence a "loss allowances" model as it enables the recognition of provision for expected credit losses on financial assets even before any of these losses have been incurred [35].

This impairment model also updates the amount of expected credit losses that is recognized at each reporting date, thus allowing the reflection of changes in credit risk of financial instruments [35]. To illustrate the impairment of financial assets under IFRS 9 better, we describe the difference between economic and accounting values of loans:

The economic value of a loan is defined as the present value of the borrower's future cash flows. Thus when a loan is recorded based on it's economic value there is no need to recognize and compensate for it's loss allowance because the contractual interests already cover all the expected losses for the whole loan period [35]. Thus when new circumstances are introduced, the economic value is adjusted due to expected PD changes of the borrower as well as interest rate changes.

The formula below can be used to calculate expected loss[9]:

$$EL_t = \sum_{t=1}^{N} PD_t(l_t) \frac{LGD(l_t)}{(1+dr)^t} \tag{20}$$

where EL_t is the expected life loss, $PD_t(l_t)$ is the cumulative probability of default, $LGD(l_t)$ is the loss given default and dr is the rate of discount for the expected cashflows.

[9] (Novotny-Farkas, 2015, p. 11)

All the parameters are then adjusted at the new information that occurs at time $t(l_t)$.

All the expected losses arising from both changes in credit risk, which reflects as a change in the PD, as well as changes in the market interest rates, which are among the major risks that financial organizations are exposed to, should be included in fair value accounting only [35]. Fair value accounting is also consistent with the economic value of loans.

As mentioned earlier, the impairment model under IFRS 9 is based on 3 stages and according to the change in credit risk, the financial instrument can be placed in either stage 1, stage 2 or stage 3. The transfer from stage 2 to stage 3 is for the financial instruments which have reasonable or fair facts for impairment. Thus we can say that a three-staged impairment model which is based on expected credit losses is actually an approximation of fair value accounting and the economic value of the loans [35]. The economic value of loan can thus be defined as the expected loan losses or credit costs subtracted from the discounted present value of the future cash flows of the loan [29]. Whilst the accounting value of a loan is simply the principal value of the loan when default is not probable [58].

There is also different use of the annual effective interest rate for the calculation of future cash flows which depends on the stage the financial instrument is in [35]. Compared to IAS 39 where interests are recognized as income without an adjustment for credit risks at purchase, calculations of future cash flows under IFRS 9 actually take into account, at the time of recognition or purchase of the financial instrument, the effective interest rate.

Since this new model of impairment on the basis of the expected credit losses under IFRS 9 assumes that organizations are able to evaluate the expected credit losses and therefore also verify a significant increase in credit risk on the reporting date, there must be a way in an which organization can define a significant increase in credit risk.

Deloitte conducted a questionnaire[10] on this and we can thus use it as a benchmark to see how an organization defines a significant increase in credit risk [35]. Below is a summary of the responses of banks questioned about what they define a trigger for significant increase in credit risk:

- Missed payments- 41 % of banks.
- Change in rating - 35 % of banks.

In addition, 60 % of banks are still using the existing impairment models which are used for capital adequacy calculations in accordance with Basel [11].

Thus in summary one can say that this replacement is a challenge for organizations, as there is a move from looking back to forward-looking [35]. However, the characteristics of IFRS 9 standards are qualitative, and take into account the importance of the reliability of the presentation, comparability, verifiability, timeliness as well as understandability of the accounting data presented[12].

2.6 Incurred Loss Model (ILM) vs. Expected Loss Model (ELM)

A loss is said to have occurred when the present value of the future cashflows is less than the carrying amount, amounts that the company has on its books for an asset or a liability [19], of credit. It can be awakened by various reasons ranging from the debtor experiencing repayment problems of debts that are due to the deterioration of the debtor's rating, amongst many other reasons. Thus losses have an effect on the Income Statement and hence also on the Balance Sheet.

This means that financial statements and reporting are sensitive to the rules used for loss recognition. Asset's values such as held debt instruments and receivables, amongst others, become impaired in the presence of a loss in the Balance Sheet [19]. ILM and ELM are thus two methods, under different accounting standards (IAS 39 and IFRS 9 respectively), that are used for estimation of future credit

[10] (Deloitte, 2015, p. 6)
[11] (Deloitte, 2015, p. 11)
[12] (International Accounting Standards Board, 2010, p. 16)

losses. Although the ELM is now considered better than the ILM, both models have their own short-comings and advantages which we discuss below. The figure below illustrates their differences in terms of loss recognition:

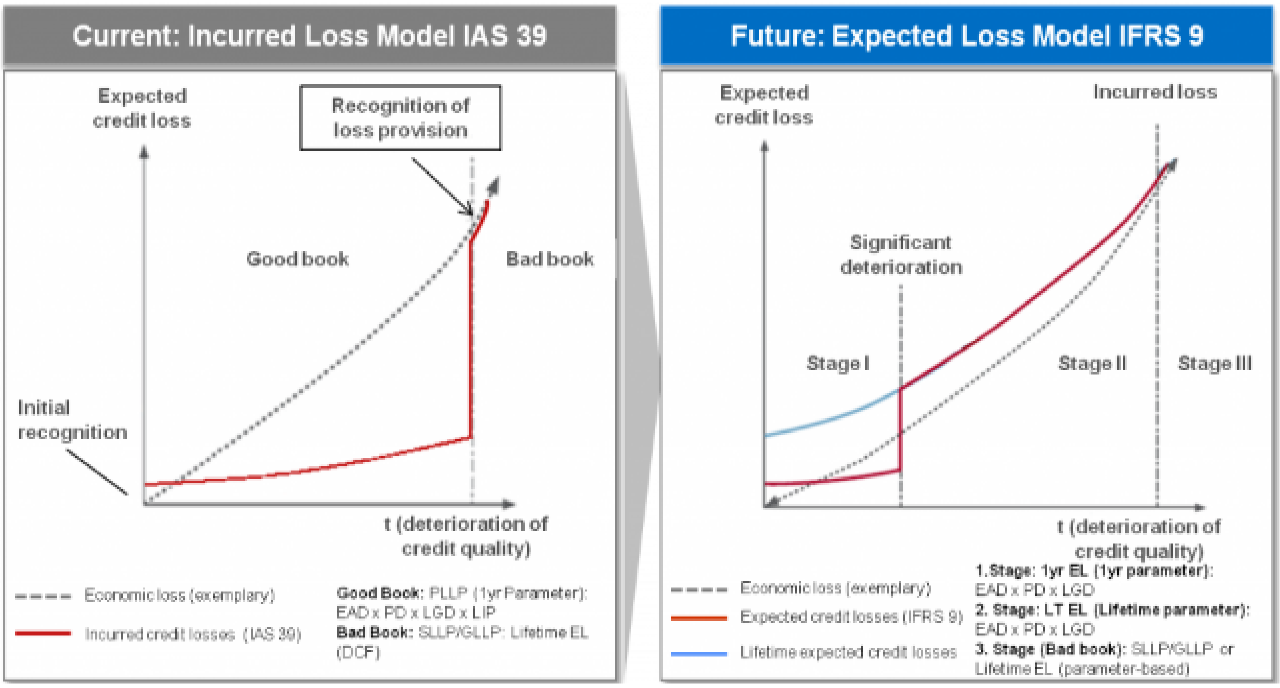


Figure 9: ILM vs. ELM. Adapted from [1]

An overview of the ILM

As mentioned in the previous sections, IAS 39, which has been in place since 1998 as a method for credit loss valuation, was based on the ILM. The assumption that forms the basis of this method is that unless there's evidence that strongly suggests otherwise, loans will be repaid [19]. It's only when there's a clear sign that a loan won't be repaid, that the asset will be written down to a lower value by the reporter [19]. Thus a certain level of optimism is introduced since the reduction in value doesn't have to be recognized although there's a reasonable chance of loss.

The motive behind the prescription of the ILM at its time, by the International Accounting Standards Board (IASB), was to address the need for banks to maintain their profit levels. The recording of losses was thus only possible only in the presence of certain events that indicated that a loss had been incurred. The disadvantage of this however, was that the reporting companies couldn't make allowance or expenses for losses had not yet been incurred but that were highly probable.

The main feature of the ILM is that it makes the recognition of loss dependent on an event that has occurred. Thus it can be seen as a way to constrain banks to refrain from manipulating their costs which causes them to save on taxes and also enables the reduction of payment of dividends basis [19]. This means that before recognizing a loss, reporting companies had to wait for a material event to occur. These material events comprise of a decrease in the amount of money that is expected to be received back by the creditor and hence is classified as a credit loss.

The ILM only recognizes a loss when an event which has led to a loss in credit value or an event which has led to a decrease in the probability to regain the contractually-binding amount has occurred [19]. This means it doesn't allow the recouping company to record an unreal loss even when there is an objective reason behind this recognition. Also the ILM enables earnings management in the sense that the postponing of losses as well as the hiding of risks produced by financial instruments was possible in the reporting entities [19].

However, there's a high probability of the occurrence of a certain amount of losses based on the inevitable and unpredictable evolution of financial markets, the repayment records of a debtor as well as experience, especially when credits are represented by a portfolio of assets [19]. Thus not taking this into consideration and also due to the fact that losses can sometimes not be recognized in Profit and Loss, though they were reasonably expected, led to financial statements appearing much favourable than they actually were.

Hence the misleading financial statements did not display a realistic image of the real state of the company. This problem is said to have induced the financial crisis as the credit worthiness of the debtors continued to reflect credits at their fair value, though it was uncertain in the financial statements of the bank [19].

Thus in conclusion the delayed recognition of loss affects stakeholders greatly, without warning nor giving them the chance to respond, which caused the generation of a procyclical effects as well as overestimation of interest rate revenues [19].

An overview of the ELM

After the financial crisis the International Accounting Standards Board (IASB) began working on adjusting the rules of loss recognition. The first proposal for this was made in 2009[13] which was that bad debt losses can be calculated by applying the effective interest rate of the financial instrument [19]. The basic idea behind this was to take into account the expected credit losses in interest income. However, this proposal was rejected because of the fact that banks seldom checked the valuation of the credit worthiness of the financial instruments in their operational business, though it was acknowledged as conceptually convincing [19].

Thus a second proposal was introduced in 2011[14] by the IASB. This proposal based on a model where the recognition of interest and bad debts were separate as before. It required financial instruments to be divided into two books namely, white and black [19]. The white book was attributed to all financial assets where default was unlikely. The black book on the other hand was for all other credits.

Thus the idea was that as the bad debt credit losses covered the remaining duration of the credit, which was approximately the same as the bad debts recognized under the ILM, the provision for risk for the business cases in the white book would reflect the expected credit losses during the credit period [19].

[13] (ED/2009/12)
[14] (ED/2011/1)

The third proposal was presented by the IASB in 2013[15]. It suggested the recognition of provisions for loss also for financial instruments that were displaying non-performance previously till currently.

The final proposal was made by the IASB on the 24th July 2014 through the publication of IFRS 9 (Financial instruments), and it came into effect on the 1st January 2018. IFRS 9 requires accounting for expected credit losses when financial instruments are initially recognized and also recognition of full time expected losses on a more timely basis by entities [19]. Hence the introduction of the ELM, which requires a more timely recognition of expected credit losses.

With the ELM it's no longer necessary for a trigger event to occur first, what we need now is more timely information about expected credit losses. In other words, expected credit losses are based on the availability of reasonable and supportable information without undue cost or effort and that's inclusive of historical records and forecast information as stated by IASB, 2014.

[15](ED/2013/3)

3 Data and Methodology

3.1 PD modelling data and methodology

PD modelling data

The first step in Markov chain modelling is to have a transition matrix that is constructed based on observed changes of states. Thus in the case of corporate credit risk modeling it generally means an annual transition matrix that reflects the change in rating. This transition matrix can be constructed from either internal or external data. However, since the appropriate internal data for this modeling purpose is possessed by financial institutions only, it's not possible for us to publish models using internal banking data. Thus we will use the long run global corporate annual probabilities of transitions of the Standard and Poor's (S & P) rating agency.

We will attempt to apply the Markov chain and transition matrix methodology using South African data, but because South African data is limited, we will make a few exceptions/ assumptions depending on the available data. For instance, since South Africa falls under emerging markets, in cases where we can't find data specifically for South Africa, we will use emerging markets' data. For this methodology we will use the following external data:

(i) For the initial one year PD matrix, we will use Standard and Poor's (S & P) *2017 One-Year Corporate Transition Rates By Region* for emerging markets [10].

(ii) For the EAC we need a time series that will serve as a proxy for the probability of default. In this thesis we use South African government bond spreads [16] as a proxy for PD as they reflect the relative risks of the bonds being compared and then find the relationship between the credit spreads and PD that we will use in the linear regression. This relationship, obtained from [11] is given by:

[16]The data for the SA government bond spreads and the resultant PD proxy data is attached in the Appendix (Table 16) and were provided by Supervisor, source is unknown

$$CS = (1 - RR) \times PD \quad \Rightarrow \quad PD = \frac{CS}{(1 - RR)} \tag{21}$$

where *CS* is the credit spreads and *RR* is the recovery rate.

Government bond spreads are considered a country's most creditworthy bonds, thus by referring to bond spread, we are comparing the yields on government bonds to the bonds of other issuers, in our case those of South African corporations (corporate bonds).

Certain macroeconomic conditions factor into bond spreads. For instance, when inflation is on the *rise*, the credit risk for a corporate bond *increases* because the likelihood of default is higher due to rising prices, which means we expect a positive EAC. This causes the corporate bond spread to widen /increase. Also, when the economy is experiencing a period of *growth*, corporate bond spreads generally *narrow* because there is a lower likelihood of the company defaulting on its debt obligations during a time of economic expansion, thus we'd expect a negative EAC.

From the equation it can be seen that credit spreads are directly proportional to the probability of default, thus the above mentioned effects on credit spreads have the same impact on the implied PD. In this paper, for no particular reason, we will assume an arbitrary recovery rate of 40%.

(iii) We also need macroeconomic forecasts for the EAC. We will divide our macroeconomic factors into 2 broad categories:

- Those related to general macroeconomic conditions i.e. the **unemployment rate** [12] and **Consumer Price index (CPI)** [13]
- Those related to the direction in which the economy is moving i.e **real GDP growth** [14]

Below are the plots of the historical and forecast data of the above mentioned macroeconomic factors [17]. Where the history (2017 and backward) is represented by the plot before the dotted line and the forecast by the plot after the dotted line (2018 and forward). These forecasts for GDP growth rate, unemployment rate as well CPI were obtained from OECD [18].

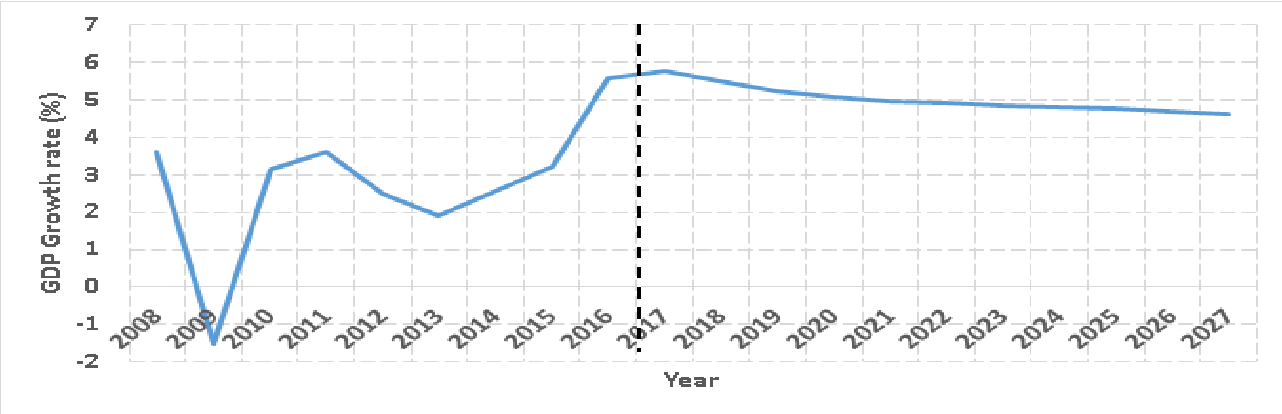


Figure 10: GDP growth history and forecast

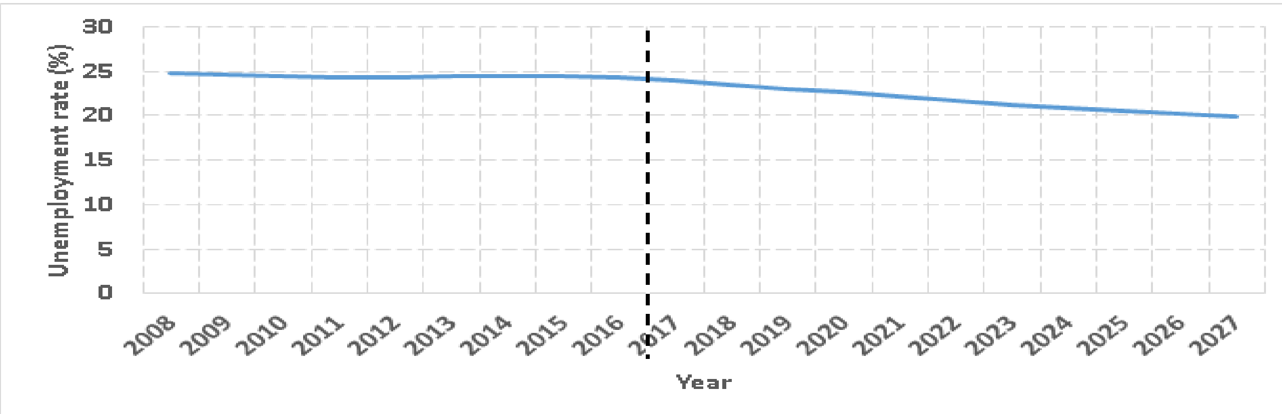


Figure 11: Unemployment rate history and forecast

[17] The data on these macroeconomic factors is attached in the appendix (Table 17

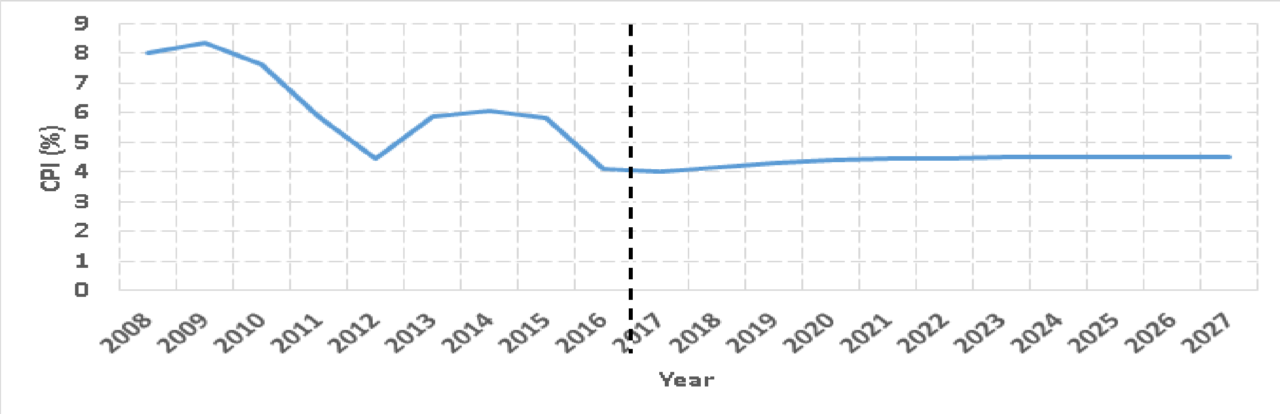


Figure 12: Inflation growth history and forecast

PD modelling methodology: The Markov chain and rating transition approach

We first start by showing a theoretical and straightforward logic based illustration that verifies the PD formula which shows how we incorporate macroeconomic factors (EAC) into the initial one year transition matrix in order to compute the multiple year transition matrices and thus the multiple year PDs. This verification, adapted from [57], forms as a basis for the computational framework, that we'll explain immediately after the proof.

Theoretical PD formula verification

We illustrate the proof by means of an example. We start by assuming that we have 4 rating grades: R_1, R_2, R_3 and R_4, where R4 is the default rate. We know that the probabilities of these default and non-default grades have to sum up to one. Which means PD can be considered as the complement of of PND (probability of non-default). This means that:

$$\begin{aligned} PD &= 1 - PND \\ &= 1 - \sum_{i=1}^{n} p_i \end{aligned} \tag{22}$$

where PD is the probability of default, PND is the probability of non default, p_i are the transition probabilities to rating grades (R_i) and n is the number of rating grades, in this case n=4 (since we have 4 rating grades)

Thus if the one year PD of a client in R_1 is 5% in year t, this implies that the client will default with a probability of 5% and survive with a probability of 95%. Therefore in this case, this means the transition probabilities to the non-default grades, R_1, R_2 and R_3, should sum to 95%.

Assuming the default rate is absorbing and that the one year PD of a client in R_2 and R_3 is 4% and 6% respectively, it would result in the following transition matrix :

$$\begin{bmatrix} 0.95 & 0 & 0 & 0.05 \\ 0 & 0.96 & 0 & 0.04 \\ 0 & 0 & 0.94 & 0.06 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{23}$$

Since the proof is very simplistic, we further assume that the off diagonal entries are zero i.e. the client can either stay at their current non-default grade or default, this is in order to make the PD formula make theoretical sense.

Thus in order to calculate the two year PD, it's assumed that the client will survive the first year, meaning the PND for an p_1 client in the next two years from the first year is given by:

$$\begin{aligned} PND(2) &= (1 - PD(1)) \times (1 - PD(1)) \\ &= (1 - 0.05)^2 \\ &= 0.9025 \end{aligned} \tag{24}$$

In the same way, the PND for an R_2 and R_3 client is 0.9216 and 0.8836 respectively.

Thus based on the described logic, PD equals one minus the probability of non default. The general formula for calculating PND at time t is given by:

$$PND(t) = \prod_{i=2}^{t} (1 - PD(1))^t \tag{25}$$

where $PD(1)$ is the initial one year transition matrix.

It is important to note that equations (24) and (25) only hold true because the client cannot jump in between states, but can only jump into state of default. Thus based on the described logic, PD equals one minus the probability of non default, which gives us the following for an R_1 client:

$$\begin{aligned} PD(2) &= 1 - PND(2) \\ &= 1 - 0.9025 \\ &= 0.0975 \end{aligned} \tag{26}$$

Using the same formula for R_2 and R_3 clients we get a PD of 0.0784 and 0.1164 respectively. Thus the general formula for calculating PD is given by:

$$\begin{aligned} PD(t) &= 1 - PND(t) \\ &= 1 - (1 - PD(1))^t \end{aligned} \tag{27}$$

However, this formula doesn't incorporate macroeconomic adjustments. This means that we have to incorporate the macroeconomic forecasts into the above mentioned formula. We do this by taking the product of the adjusted one year probability of default instead of raising it to the power :

$$\begin{aligned} PD(t+n) &= 1 - PND(t) \\ &= 1 - \prod_{k=1}^{t} (1 - PD(1) \times \Delta(t+k) \times EAC \times w) \end{aligned} \tag{28}$$

where:

- $\Delta(t)$ is the macroeconomic variable forecast in period $t+k$ compared to the base period t

- EAC is the economic adjustment coefficient

- w is the weighting that is places for the economic adjustment effect

The formula above is the PD formula that forms the basis of the formula that we use in our methodology for forward looking PD estimation that is explained below.

Computational mechanism

We will build on the computational framework used in Vanek and Hampel (2017) to extend the one period PD. This computational framework assumes that the default state is absorbing, meaning should an obligor default they cannot recover. It also excludes the assumption that PDs tend to decrease over time.

The outline of this computational framework is as follows:

(i) Obtain the one year transitional matrix,P , which consists of r rating grades and the last column, (rth), is the default column.

(ii) Since we use the concept of Markov chains to estimate the future PD transition matrices, which are time-homogeneous, we introduce time-inhomogeneity by estimating an economic adjustment coefficient, using a simple linear regression model, which captures the impact of future expected economic development i.e macroeconomic factors on the PD.

(iii) Perform a variable selection process on the macroeconomic factors to see which ones will have a significant impact on PD. We will use the correlation and concordance between the macroeconomic variables and PD as well visualization and business intuition to select the most effective macroeconomic factors [30]:

   (a) *Correlation* shortlists the macroeconomic factors that show a linear trend with change in the PD proxy, so the higher absolute value, the strongly related the change in the variable is to the PD proxy.

(b) *Concordance* further shortlists macroeconomic factors that display both a linear and non-linear trend with PD proxy.

(c) *Visualization* of PD proxy with change in macroeconomic factors assists in checking *business intuition*, and thus selecting additional variables based on this business intuition as well as the availability of macroeconomic forecasts. For instance from visualization, we can note that a change in the inflation or interest rates may not affect PD immediately but maybe after a year, thus in this case we used a lagged version of the inflation and interest rate.

(iv) Modify or adjust the one year PDs or transition matrix for the future time periods using the decomposing effect in (ii).

(v) We then calculate the multi-period PDs, P', by using either of the following equations depending on the weighting we want place for the economic adjustment effect or on what direction of the adjustment we desire to emphasize as the setting of the weighting is fully dependent on the practitioner and the application [57]:

(a) Uniform decomposition

The first and second half of the EAC effect is for adjusting the non-default grades and default rating grade respectively. It's desired that the worse the rating grade (further to the right), the bigger the adjustment towards the default grade should be made. A uniform decomposition within non-default grades within both left and right directions is considered. This can be achieved by making the following adjustments to our one year PD transition matrix:

- Last /default column entries:

$$p'_{i,r} = p_{i,r} + \frac{\delta_{t+k} \times EAC}{2(r-1)} \times \frac{2i-1}{r-1} \quad for \quad i = 1, ..., r-1 \tag{29}$$

where:

- the second term equals γ

- EAC is the economic adjustment coefficient

- δ_{t+k} is the difference between forecasted macroeconomic variable value and it's actual observed/current value at the initial time t.

- Above but excluding diagonal entries:

$$p'_{i,r} = p_{i,r} + \frac{\gamma}{r-i-1} \quad for \quad i,j < r-1 \quad and \quad j > i \tag{30}$$

- Under diagonal entries:

$$p'_{i,r} = p_{i,r} - 2 \times \frac{\gamma}{i} \quad for \quad i,j < r-1 \quad and \quad j \leq i \tag{31}$$

- On diagonal entries:

$$p'_{i,r} = p_{i,r} - \frac{\gamma}{i} \quad for \quad i = r-1 \quad and \quad j \leq i \tag{32}$$

(b) Decreasing decomposition

The first and second half of the EAC effect is used for the non-default grades and the default rating grades adjustment respectively. Like in (a) it's desired that the worse the rating grade (further to the right), the bigger the adjustment towards the default grade should be made. However, in this case, the decreasing decomposition from the current grade within non-default grades in both directions is considered. This can be achieved by making the following adjustments to our one year PD transition matrix:

- Last /default column entries: Same as (28)

- Above but excluding diagonal entries:

$$p'_{i,r} = p_{i,r} + \frac{\gamma}{r-i-1} \times \frac{2(r-j-1)+1}{r-i-j} \quad for \quad i,j<r \quad and \quad j>i \tag{33}$$

- Under diagonal entries:

$$p'_{i,r} = p_{i,r} - 2 \times \frac{\gamma}{i} \times \frac{2(i-j)+1}{i} \quad for \quad i,j<r-1 \quad and \quad j \leq i \tag{34}$$

- On diagonal entries:

$$p'_{i,r} = p_{i,r} - \frac{\gamma}{i} \times \frac{2(i-j)+1}{i} \quad for \quad i=r-1 \quad and \quad j \leq i \tag{35}$$

(c) Increasing decomposition

The first and second half of the EAC effect is used for non-default grade grades and default rating grade adjustment respectively. Like (a) and (b) it's desired that the bigger the adjustment towards the default grade should be made. However, in this case an increasing decomposition from the current grade within non-default grades in both directions is considered. This is opposite to (b). This can be achieved by making the following adjustments to our one year PD transition matrix:

- Last /default column entries: Same as (28)

- Above but excluding diagonal entries:

$$p'_{i,r} = p_{i,r} + \frac{\gamma}{r-i-1} \times \frac{2(j-i-1)+1}{r-i-1} \quad for \quad i,j<r \quad and \quad j>i \tag{36}$$

- Under diagonal entries: Same as (6)

- On diagonal entries: Same as (7)

(d) No half effect

The EAC effect is not divided between non-default and default grades. In this case the EAC effect is divided between increasing transition probabilities to worse grades an decreasing transition probabilities to better grades. This can be achieved by making the following adjustments to our one year PD transition matrix:

- Above but excluding diagonal entries:

$$p'_{i,r} = p_{i,r} + \frac{\sigma_i}{r-i} \times \frac{2(r-j)+1}{r-i} \quad for \quad i,j \leq r \quad and \quad j > i \tag{37}$$

where $\sigma_i = \sum_{j=i+1}^{r} \psi_{i,j}$ and

$\psi_{i,j}$ is the "plus" changes associated with probabilities $p_{i,j} \Rightarrow \sigma_i$ is the sum of these changes in each individual row. Changes are calculated according to equations (6), (7) and (9) with $p_{i,j}$ omitted.

Thus the changes in transition probabilities to worse rating grades from the initial state continuously decrease, including in the default grade.

- Under diagonal entries: Same as (33)

- On diagonal entries: Same as (34)

We also include our own approach, which straightforward and the multi-dimensional version of equation (27) i.e. it stems from equation (27), where we weight the whole effect of EAC equally into all rating grades including the default grade. Thus we do the following transformation to obtain the adjusted transition matrix:

$$p'_{i,j} = p_{i,j} + \delta_{t+k} \times EAC \times w \tag{38}$$

where:

- δ_{t+k} is the difference between forecasted macroeconomic variable value and it's actual observed/current value at the initial time t.

- w is the weighting, which is one for the whole effect, a half for half effect etc.

(vi) Then finally, since the proposed adjustments shown in (iv) don't ensure the property that the transition probabilities fall in the interval $(0,1)$ i.e. $0 < p_{i,j} < 1$, we need to correct for this in order to avoid extreme transition probabilities soley because of macroeconomic development as well as to maintain the dynamics of the model. We do this by setting a desired threshold for the transition probabilities. We use the floor of $0,03\%$[18] for all transition probabilities.

We code the above mentioned approaches in Matlab. Below is a summary that explains the algorithm of the Matlab code and how we reach the multiperiod PD term structure:

(i) Given the initial transition matrix, the EAC and yearly deltas (differences between the annual macroeconomic forecasts and initial (base) values of the macroeconomic variable), we adjust the initial transition matrix using the different approaches to obtain the forward looking multiple period transition matrices. The rows in each of these transition matrices must sum to one.

(ii) In order to compute the PD, we need to calculate the state vector. Thus, since we are working with multiple rating grades and each row in the transition matrix represents a specific rating grade, we would have a $(1 \times r)$ state vector, which is dependent on which state we are in at the beginning and where r is the number of rating grades.

[18] The value of $0,03\%$ is set for PD in European Parliament and Council within the framework of credit risk capital requirements calculation

For instance, assume we have 4 rating grades, meaning that if we are in the second rating grade, initially, then our state vector would look like this:

$$S_t = [0,1,0,0] \tag{39}$$

However, since we desire the state vector that contains the probabilities of being in the individual possible grades using the $(r \times r)$ transition matrix, the general formula for the state vector is:

$$S_{t+n} = S_t \times \prod_{k=1}^{n} P_{t+k}. \tag{40}$$

This means that the PD would be the last entry in the resultant state vector, whilst the other entries represent the probability of being in the other non-default grades.

(iii) The multiperiod PDs will thus be extracted from these state vectors for each period. For instance, in our case in order to get the PD for each rating grade, we would have to multiply the product of the transition matrices with an $(r \times r)$ identity matrix, where each row of the identity matrix represents the state vector for each rating grade. This means that the last column of each row of the resultant matrix would be the PD for each rating grade.

Illustrative example

Assume we have 4 rating grades given by A, B, C, D, where D is the default grade, and each row and column represents these 4 rating grades i.e. row 1 represents the transition probabilities of rating grade A to other rating grades. Thus to compute the 3 year PD of each grade, for instance, assume that the one, two and three year transition matrices are given by the matrices below:

$$P_{t+1} = \begin{bmatrix} 0.4662 & 0.3778 & 0.1335 & 0.0225 \\ 0.0003 & 0.5517 & 0.35 & 0.0980 \\ 0.0003 & 0.0003 & 0.2 & 0.7994 \\ 0 & 0 & 0 & 1 \end{bmatrix}, P_{t+2} = \begin{bmatrix} 0.4782 & 0.3768 & 0.1304 & 0.0145 \\ 0.0003 & 5947 & 0.33 & 0.0750 \\ 0.0003 & 0.0003 & 0.23 & 0.7694 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{41}$$

$$\text{and} \quad P_{t+3} = \begin{bmatrix} 0.4905 & 0.3758 & 0.1274 & 0.0063 \\ 0.0003 & 0.6497 & 0.3 & 0.05 \\ 0.0003 & 0.0003 & 0.2097 & 0.7897 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{42}$$

Thus the matrix of state vectors for the 3-year PDs would be given by:

$$S_{t+3} = I \times P_{t+1} \times P_{t+2} \times P_{t+3} \tag{43}$$

where I is the $r \times r$ identity matrix

The result of the matrix multiplication gives:

$$S_{t+3} = \begin{bmatrix} 0.1096 & 0.3440 & 0.1939 & 0.3525 \\ 0.0004 & 0.2135 & 0.1536 & 0.6325 \\ 0.0001 & 0.0003 & 0.0098 & 0.9898 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{44}$$

From matrix above the three year PDs for ratings A, B and C are 35.25%, 63.25% and 98.98% respectively (taken from the last column of S_{t+3}). Thus in general, the nth year PDs for each rating grade is computed by taking the last column of the product of the identity matrix with the the nth year's transition matrix, i.e:

$$S_{t+n} = I \times \prod_{k=1}^{n} P_{t+k} \tag{45}$$

In general, we expect that when an macroeconomic factor is unfavourable to PD, it will cause the PD to increase and vice versa. The figure below shows this expectation of the behaviour of PD:

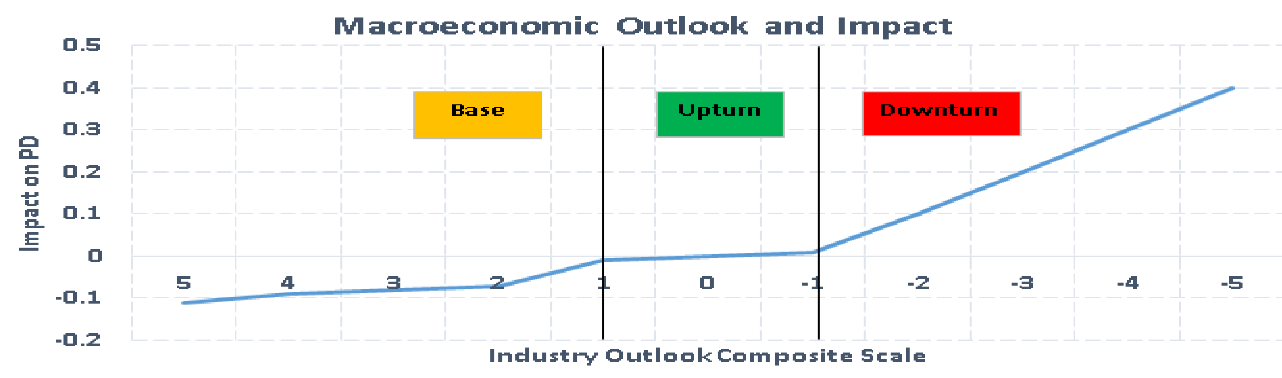


Figure 13: Expected PD behaviour

In our analysis we will be using three macroeconomic scenarios namely, baseline, downturn and upturn as shown by the Figure 13 above. The baseline scenarios are the macroeconomic forecasts that are estimated by OECD that we plotted in Figure 10, 11 and 12.

We assume that the downturn scenario is a continuous and successive decrease of the macroeconomic variable by 1% each year, and take the upturn scenario to be a continuous and successive increase of the macroeconomic variable by 1% each year. In our case the macroeconomic variable can either be the GDP growth rate, the unemployment rate or the CPI. It is also very important to note that assumptions are unrealistic and for illustration only.

These assumptions about the different scenarios will assist us in observing the impact that different economic conditions have on PD . Below are plots[19] of these three macroeconomic scenarios from our data for changes in GDP growth rate, unemployment and CPI, respectively:

[19] Data for plots are obtained from subsection 3.1 (iii)

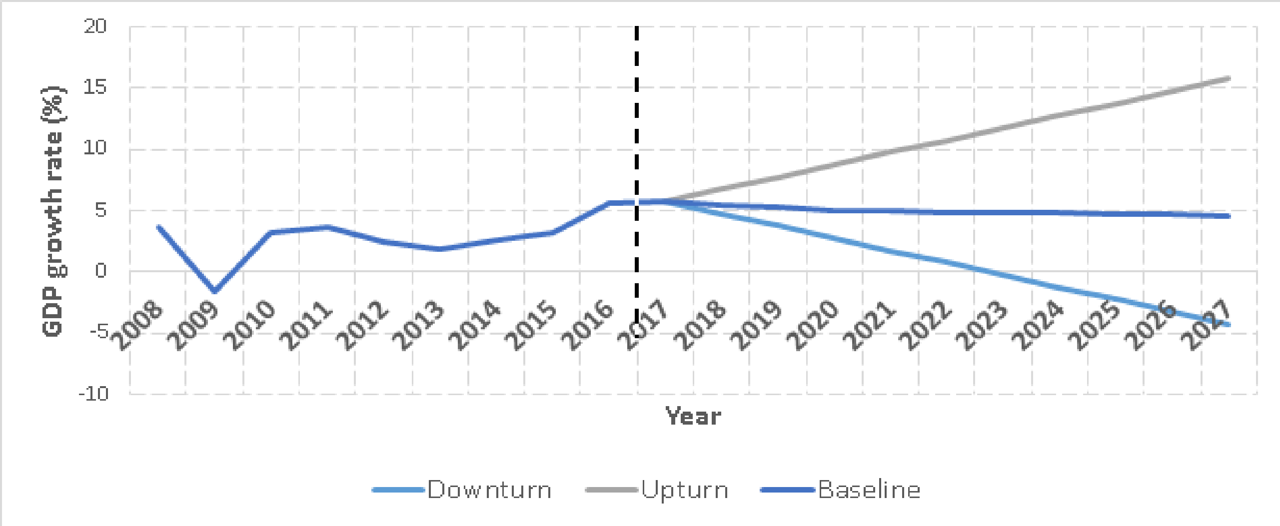


Figure 14: GDP growth rate macroeconomic scenarios

For GDP growth rate, we can see from the above figure that an upturn is characterized by a 1% increase every year, whilst a downturn is characterized by a 1% decrease each year.

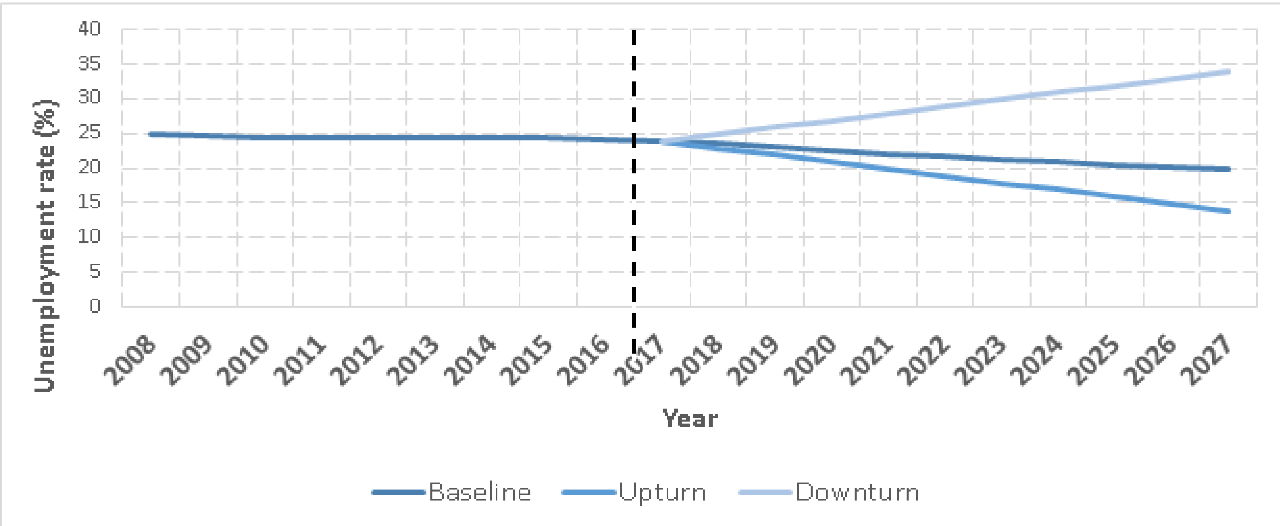


Figure 15: Unemployment rate macroeconomic scenarios

The upturn under unemployment rate is characterized by a yearly decrease of 1% in the unemployment rate, whilst a downturn is characterized by a yearly increase of 1% in the unemployment rate.

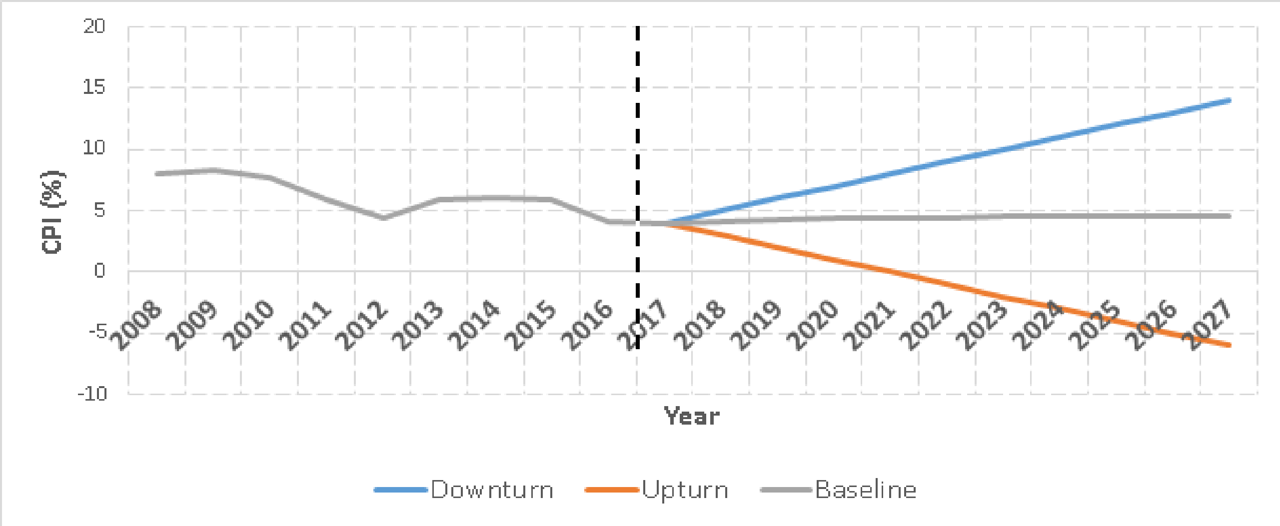


Figure 16: CPI macroeconomic scenarios

The upturn under CPI is also characterized by a yearly decrease of 1% in CPI, whilst a downturn is characterized by a yearly increase of 1% in CPI.

3.2 LGD modelling data and methodology

LGD can be defined as the ratio of losses to exposure at default. The methodology for modelling LGD is similar to that of PD modelling. It follows the following generalized process [47]:

(i) Obtaining the historical LGDs for prior periods. There are various ways to can obtain these historical LGDs such as financial data vendors terminals, structural models, Basel guidelines etc.

(ii) Obtaining historical macroeconomic factors that are related to the sector of the loan portfolio and then forecasting the future macroeconomic variables using different statistical methods.

(iii) Utilize the macroeconomic factors and the historical LGDs to predict the forward looking LGD. One can use the Gaussian copula model, regression analysis, probit model, logit model etc. to predict the forward looking LGD.

There are three kinds of losses that fall LGD namely, the loss of the principal amount, the carrying costs of non-performing loans such as the forgoing interest income and finally workout expenses such as collections, legal etc. [20]. There are four broad ways of modelling the LGD for an instrument [22]:

(a) *Market LGD* : This LGD is estimated from the market prices of bonds that have defaulted or marketable loans right after the event of the actual default and is used in wholesale only.

(b) *Workout LGD* : This is the set of properly discounted cashflows that are estimated, resulting from the workout or collections process and the estimated exposure. It can be used in both retail and wholesale.

(c) *Implied market LGD*: Theses LGDs are derived from bond prices that are non-defaulted, but risky using a theoretical model and is also used in wholesale only.

(d) *Statistical LGD*: These LGDs are estimated using regression on facility characteristics and historical LGDs. It can be used in both retail and wholesale.

Although workout LGD approach is said to give the most precise, flexible, transparent and logical estimate of LGD, it's more calculation intensive [22]. This is because it has explicit structures that represent real-world processes and the probabilities of certain recovery outcomes [20]. There's no single model development strategy for workout LGD that can be used across banks, let alone even in the same bank. The right model choice depends on the following factors [20]:

(a) Policies that are specific to the bank, as internal bank policies actually play as crucial role in determining the LGD methodology that is most suitable.

(b) The region. This is because local regulations in some economies may have a direct impact on the recognizing and managing defaults. Also,in different regions, the costs that are associated with loan recovery and the process of recovery, could be notably different.

(c) The portfolio(s) in the scope of a particular LGD model. For instance, the way in which a corporate LGD is structured could be significantly different from a retail or an SME (Small-to-Medium Enterprise) LGD model. And differing strategies to LGD modelling for each portfolio or sub-portfolio could be used within a specific bank.

(d) The availability of data is also very important in determining the most effective strategy for LGD modelling.

One can either use statistical, judgemental or hybrid methodologies to estimate LGD. In this paper we will only discuss the statistical LGD estimation methodologies. There are four statistical LGD estimation methodologies that one can use, however, they are severely constrained by the availability of data [20]. Thus these methodologies are most preferred when abundant recovery data is available, which with no obvious segmentation, exhibits a relatively smooth, homogeneous structure and represents the portfolio at hand, the relevant collateral being modelled, the recovery processes associated with the relevant collateral divisions as well as future processes [20]. Below we briefly explain these four statistical LGD estimation methodologies [20]:

(i) Econometric methodology:

This methodology is specifically for modelling proportions. For instance, the nonlinear, fractional regression that is estimated using methods of quasi-maximum likelihood methods.

(ii) Nonparametric regression trees for modelling recoveries on bank loans:

The tree models from this methodology resemble look-up tables that contain historical recovery averages, which is easy to interpret and bounded to the unit interval.

(iii) Prediction of loan recoveries with neural networks:

The variables used by the neural network models to derive their output coincide with those that are significant in parametric regression models, to a great extent. Thus neural networks may have better predictive power than parametric regression models, given a sufficiently large number of observations, as suggested by out- of- sample prediction errors.

(iv) Multivariate linear regression:

This is the most straightforward methodology for modelling recoveries. It regresses a continuous recovery rate variable on explanatory variables and it allows the incorporation of forward looking macroeconomic factors, which is an IFRS 9 requirement, as explanatory variables, allowing for accounting of inter dependencies among some of these variables. The predicted values can, however fall outside the range and may have to be moderated based on expert judgement, which is allowed under IFRS 9.

In this thesis we will discuss the modelling of the market LGD, using multivariate linear regression, as it's comparatively less computationally intensive, works well for liquid market instruments and the recovery studies of rating agencies is based on this approach. To estimate LGD, banks are also advised to use either the market or implied LGD approach for instruments that are liquid and traded, and the workout LGD for illiquid instruments that have no market [22]. We will thus look at the market LGD for loans and bonds that have defaulted, as long as trade has actually occurred, which allows one can directly observe prices.

The actual prices are based on par which is equal to 100 cents in the Rand and therefore can easily be translated into a recovery percentage or:

$$LGD = 100\% - Recovery \tag{46}$$

These prices also have properties that are desirable as they are observed early and also reflection of what the market sentiment is at that time [44]. They are also less subject to debate about proper valuation because they are a result of a market transaction. Finally, these prices reflect the suitably discounted recovery expected by investors and therefore they include recoveries of both the discounted and missed interest payments and also the costs of restructuring as well as the uncertainty associated with that restructuring process. For example, in the Moody's dataset, they are observed a month after the first event of default. This price is thus the market's expected present value of eventual recovery [44].

We thus proceed to modelling market LGD under IFRS 9. As mentioned above LGD is defined by the following equation:

$$LGD = \frac{Losses}{EAD} = \frac{EAD - NR}{EAD} = \frac{EAD}{EAD} - \frac{NR}{EAD} = 1 - RR \tag{47}$$

where EAD is the exposure at default, RR is the recovery rate which is the remaining share of a financial asset that we expect to recover when a borrower defaults and NR is the net recoveries and they also include all the recovery cashflows minus the costs involved in recoveries.

Because there is no relevant readily available data for LGDs that we can use to compute forward looking LGDs we will only illustrate the simplest and most transparent LGD modelling method from [44] which uses regression analysis and can used when the required data is available :

Step 1: Definition of the observation period $(t1)$

This is the period from the date of initiation of the loan to the date of observation or the reporting date. This period is usually one year.

Step 2: Definition of events of default within the observation period

Default events can fall in either of the following categories:

(i) Exposures past the due that exceed a particular number of days such as 60 days. These exposures can either be the principal amount or interest amount.

(ii) Exposures which have fallen into the first stage and thus consequently removed from the books.

Step 3: Definition of the LGD period $(t2)$

This is the period from the default date and it's usually 12 months.

Step 4: Collection of data for net recoveries during $t2$

This is inclusive of both the recoveries and costs that are related to the recoveries during $t1$.

Step 5: Discounting to the default time, using a suitable rate to the cashflows and taking the sum

It's required that the NRs be discounted to the default date. Thus the sum of the discounted cashflows implies the MVLD (Market Value of the Loan at the time of Default). The rate used to discount, which is not the contract rate, should resemble the yield to maturity of the loan at the default time. This rate should eventually be greater than the contractual rate as it also includes an enhances premium of default risk.

Step 6: Application of the CCF (Credit Conversion Factors) to the loan exposure

A CCF converts the amount of a free credit line and other off balance sheet transactions to an EAD amount. If the loan exposure is the same as the credit limits, then the CCF is 100%. If it's less than the credit limits , then the portion that is not withdrawn needs to be converted to loan equivalent exposures. This methodology is the simplistic approach of applying the regulatory CCF. A more sophisticated methodology of applying CCF is by building an EAD model. The exposure after the CCF is the loan equivalent (LEQ).

Step 7: Calculation of the RR

The *RR* for the exposure is the dependent variable and is given by the following ratio:

$$RR = \frac{MVLD}{LEQ} \tag{48}$$

The distribution of the *RR* will most likely be bimodal with a large number of the observations distributed either near 0% of 100%. A bimodal distribution is for the beta distribution with parameters $alpha$ (α) and *beta* (β) almost the same. A beta distribution represents all the possible values of the probabilities when the actual probability is not known i.e. it is a distribution of probabilities. It is defined as a family of continuous probability distributions defined on the interval $[0,1]$ parametrized by two positive shape parameters, denoted by α and β, that appear as exponents of the random variable and control the shape of the distribution [15].

Step 8: Identifying the independent variables

These variables should be specific to the loan, to the firm and to macroeconomic variables, like product type, financial ratios and GDP growth respectively, amongst other things. One can identify all the possible variables based on discussion with the business divisions and then statistically select those which have a good predictive power. The widely used method to determine the predictive power of each independent variable is the Accuracy Ratio. Finally, it's important that the correlation among the independent variables is not high.

Step 9: Regression

The recovery rate from step 7, which is a bimodal distribution, is not suitable enough to apply OLS regression as it requires the variables to be normally distributed. Thus the *RR* needs to be transformed into a normally distributed variable before we regress it with the independent variables. This transformation can be done on Matlab or Excel using the *NORMDIST* function in the following way:

$$NORMDIST\,(var, \bar{x}, \sigma) \tag{49}$$

where *var* is the variable that we want to convert to the normal distribution, $\bar{x}$ is the mean of the transformed normally distributed variables and σ is the standard deviation of the transformed normally distributed variables.

The parameters $\bar{x}$ and σ are given by the following equations:

$$\bar{x} = \frac{\alpha}{\alpha+\beta} \times Max \qquad \sigma = Max \times \sqrt{\frac{\alpha \times \beta}{(\alpha+\beta)^2} \times (1+\alpha-\beta)} \tag{50}$$

where *Max* is the maximum value of the variables.

Once the transformation of the bimodal distribution variables to the normal distribution is done the OLS regression can be done using statistical softwares such as R, SAS, STATA as well as Excel. Basic hypothesis testing or statistical significance of regression coefficients need to be followed, and these can be viewed from the ANOVA (Analysis of variance) table [44].

Step 10: LGD Estimation

When the coefficients resulting from the regression is applied to the respective independent variables of a new data set gives us the *RR*. Thus we can calculate LGD using equation (1).

The main advantage of this methodology is it's robustness, simplicity as well as transparency. Out of sample and out of time tests should be used to validate the model. Apart from the initial validation, a periodic validation, preferably annually, should also performed to ensure that the predictive power of the model is satisfactory. This can be analyzed by arriving to the accuracy ratio.

A drawback of this method however, is that the predicted values can be outside the $(0,1)$ range. Thus in such cases, one may restrict the predicted probabilities within the zero and one range by assuming the probability to be zero when the estimated predicted probabilities are negative and one if they are more than one.

Alternative LGD forward-looking adjustment approach

Another approach of making the forward-looking adjustment to the estimated LGD is by using the Frye-Jacobs function [20]. Under this approach, the macroeconomic factor impact is linked to PD, and then to LGD via the Frye-Jacobs function. This approach assumes co-monotonicity between the the asymptotic distributions of default and loss. Essentially, this approach makes an adjustment to an expected LGD based on the expected PD as implied by macroeconomic factors over the lifetime of the instrument [20]. The Frye-Jacobs function is given by [33]:

$$cLGD = \frac{\Phi[\Phi^{-1}[cDR]]}{cDR} \tag{51}$$

where $k = \dfrac{\Phi^{-1}[PD] - \Phi^{-1}[EL]}{\sqrt{1-\rho}}$, $cLGD$ is the condtionally expected LGD, Φ is the cumulative distribution function (CDF) of the normal distribution, Φ^{-1} is the inverse CDF, cDR is the conditionally expected default rate, PD is the probaility of default and EL is the expected loss.

The diagram below shows how the Frye-Jacobs approach works:

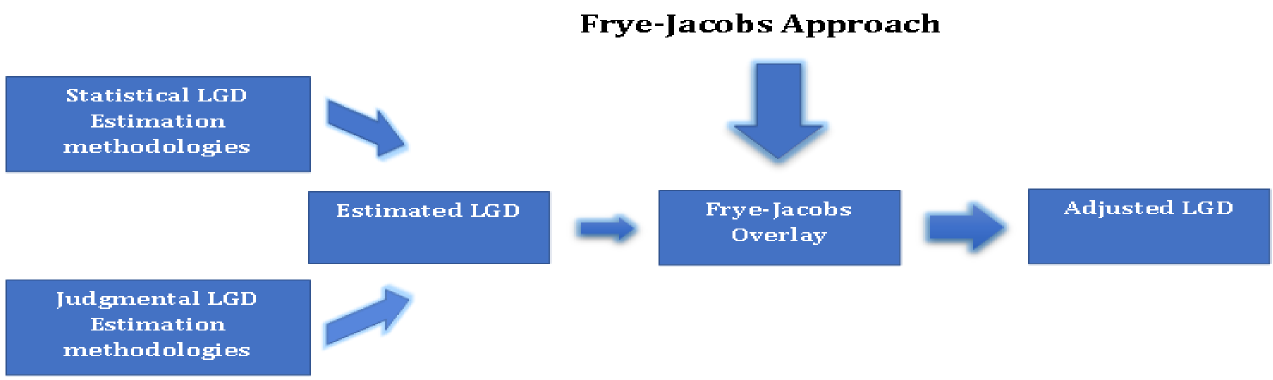


Figure 17: Flow diagram of Frye- Jacobs methodology. Adapted from [20].

3.3 EAD modelling data and methodology

As mentioned in the introduction exposure at default can be defined as the percentage of the nominal amount of a loan or the maximum amount that's available on the credit line [47]. It can also be described as the loss exposure stated as an actual amount like the balance of the loan that's outstanding. EAD can be divided into drawn and undrawn committments, which are known and needs to be estimated respectively [16].

The period of estimation for EAD depends on the type of the instrument exposure we're looking at i.e term loan, overdraft, off-balance sheet etc. For individual assessment of a term loan, it can be the maximum contractual period under consideration of extension options. For collective assessment of a term loan, the average life of the loans or the loan that has the largest life in the bucket can be considered [47].

For an overdraft one can either consider the normal life of the overdraft facilities or the management policy regarding overdraft policies. For off-balance sheet can also either consider the management policy regarding off-balance sheet exposures, the behaviour and the normal life of off-balance sheet exposures or the credit conversion factor of the bank.

There are various ways in which EAD can be estimated. Due to lack of access to sufficient data for EAD estimation, in this thesis we will only name and briefly describe three ways:

(a) Effective interest rate (EIR)

This method involves using the effective interest rate to estimate EAD. For an individual assessment, the effective interest rate will be the effective interest rate of each exposure, whilst it will be the average effective interest rate of exposures in the bucket or the effective interest rate of each exposure in the portfolio for collective assessment [47].

(b) Portfolio risk segmentation

This involves categorizing each exposure accordingly into risk groups that have similar credit risk. The table below shows an example of such a categorization[20].

Sector	Customer	Product	Amount (Rand)	Rating	Stage
Finance and Insurance	Direct Insurance Ltd	Lease	150000	Grade 1: Low Risk	Stage 1
Finance and Insurance	Unite Insurers	Lease	125000	Grade 7: Doubtful	Stage 2
Finance and Insurance	WSB Global Investment Ltd	Term loan	145000	Grade 9: Lost	Stage 3
Agriculture	Freshest Ltd	Advances	165400	Grade 2: Watchlist	Stage 2
Agriculture	Freshest Ltd	Advances	400356	Grade 2: Watchlist	Stage 2
General	Charity Vetezo	Term loan	30000	Grade 2: Watchlist	Stage 2
General	Thamsanqa Mdaka	Advances	50000	Grade 1: Low Risk	Stage 1
Manufacturing	Quality Fix Ltd	Advances	1000000	Grade 7: Doubtful	Stage 2
General	Reach Out Ministry	Lease	185000	Grade 9: Lost	Stage 3
Total			2250756		

Table 1: Portfolio risk segmentation. Adapted from [47].

(c) Credit conversion factor (CCF)

For this method, in order to obtain its On-balance sheet equivalent, we apply the CCF of the client to the Off-balance sheet exposures [47]. The following equation shows how make this application:

$$T_{OnBS} = E_{OnBS} + (E_{OffBS} \times CCF)$$

where:

- T_{OnBS} is the total On-balance sheet exposure,

- E_{OnBS} are the On-balance sheet exposures of the bank,

- E_{OffBS} are the Off-balance sheet exposures, and - CCF is the credit conversion factor.

There may be a change in exposure of loans that have undrawn limits over time as a result of unused

[20]Note: The values and names on the table are arbitrary, were randomly chosen and don't refer to any specific company. The structure of the table was adapted from [47]

limits that are available. Stage 1 assets aren't affected that much by this phenomenon compared to Stage 2 assets because of it's more likely for a drawdown to occur during stress events [47].

Thus in order to consider drawdowns we have to compute the CCF. If the bank provides future commitments, in addition to the current credit for exposures, then as EAD, these exposures contain both on and off balance sheet values. I.e.[21]:

$$
\begin{aligned}
EAD &= \text{Draw line} + CCF \times \text{Undrawn credit line} \\
CCF &= \frac{\text{Increase in exposure over the period}}{\text{Available funds at the start of the period}}
\end{aligned}
$$

Other factors that should be considered for EAD estimation are characteristics that are peculiar to the exposure, prepayment rates, rescheduling and re-negotiations, etc. [47].

[21] Equations obtained from [47]

4 PD Results and Analysis

We follow through the outline discussed in the Section 3.1. Below is the "2017 One-Year Corporate Transition Rates By Region" for emerging markets from S & P that we used as our initial one year PD matrix [22]:

	AAA	AA	A	BBB	BB	B	CCC/C	D	NR
AAA	0.375	0.625	0	0	0	0	0	0	0
AA	0	0.75	0.25	0	0	0	0	0	0
A	0	0.0046	0.945	0.0229	0	0	0	0	0.0275
BBB	0	0	0.0062	0.8868	0.0556	0	0	0	0.0514
BB	0	0	0	0.0268	0.7938	0.0474	0	0	0.133
B	0	0	0	0	0.0511	0.77	0.0288	0.0064	0.1438
CCC/C	0	0	0	0	0	0.3939	0.2121	0.1818	0.2121

Table 2: Initial transition matrix obtained from Standard and Poor's. Adapted from [10].

However, we note that this transition matrix also includes transitions to the 'Not-Rated' (NR) state. This state represents information that was lost about companies that were withdrawn from the rating pool due to mergers or repayment of their debts. Transitions to and from NR states are out of the scope of the ratings migrations under examination. In other words, these types of transitions don't occur within the state space that the Markov model is used for, they are thus seen as non-informative and removed from the data set. Thus by removing the NR column (last column) from the above matrix, we have the following matrix:

	AAA	AA	A	BBB	BB	B	CCC/C	D
AAA	0.375	0.625	0	0	0	0	0	0
AA	0	0.75	0.25	0	0	0	0	0
A	0	0.0046	0.945	0.0229	0	0	0	0
BBB	0	0	0.0062	0.8868	0.0556	0	0	0
BB	0	0	0	0.0268	0.7938	0.0474	0	0
B	0	0	0	0	0.0511	0.77	0.0288	0.0064
CCC/C	0	0	0	0	0	0.3939	0.2121	0.1818

Table 3: Transition matrix with NR column removed

However we note that simply removing the NR column, violates the assumption that the rows of the PD transition matrix sum to one. Thus we need to adjust the matrix accordingly. We can do this by

[22]Note: We exclude the last row (default row) from Tables 2 and 3, which has zero entries for all the non-default rating grades and one for the last column (default grade) due to the absorbing rate assumption

using the following formula to adjust the diagonal elements of the above transition matrix that doesn't satisfy the summation property and transform it into one that does:

$$p'_{i,i} = p_{i,i} + (1 - \sum_{j=1}^{n-1} p_{i,j}) \tag{52}$$

Below is the new transition matrix that we will use as our initial PD transition matrix:

	AAA	**AA**	**A**	**BBB**	**BB**	**B**	**CCC/C**	**D**
AAA	0.375	0.625	0	0	0	0	0	0
AA	0	0.75	0.25	0	0	0	0	0
A	0	0.0046	0.9725	0.0229	0	0	0	0
BBB	0	0	0.0062	0.9382	0.0556	0	0	0
BB	0	0	0	0.0268	0.9258	0.0474	0	0
B	0	0	0	0	0.0511	0.9137	0.0288	0.0064
CCC/C	0	0	0	0	0	0.3939	0.4243	0.1818
D	0	0	0	0	0	0	0	1

Table 4: New adjusted PD transition matrix

Now the row sums are equal to one. Also note that the default rates themselves are not adjusted. Another important note to observe is that both the above matrices have diagonal elements that are dominant. They represent probabilities that a counterparty's credit rating stays the same during a one year period. Thus, high value diagonal entries suggest that the level of credit migration is low.

Next we calculate the EAC in order to introduce time-inhomogeneity into our multiple year PD estimation. As we mentioned above we will be using South African real GDP growth, inflation and the CPI as our macroeconomic factors. In order to calculate this EAC we use the simple linear regression model estimated by the ordinary least squares method with heteroscedasticity and standard errors. Initially, our macroeconomic model will take the following form:

$$dPD = \beta_0 + \beta_1 \cdot gGDP + \beta_2 \cdot dUNE + \beta_3 \cdot dCPI + \varepsilon. \tag{53}$$

where:

dPD - is the changes/differences in the PD proxy, $gGDP$ - is changes/differences in the real GDP growth rate,

$dUNE$ - is the changes/differences in unemployment rate,

$dCPI$ - is the changes/differences in the consumer price index.

We will then do a backward elimination on equation (53) by running the three different regression models, using Excel, consisting of each of macroeconomic variables separately, to see which of them are statistically significant i.e. has the most effective influence on the PD.

Before we do the backward elimination, running the multi-variable regression including all the macroeconomic variables (equation 53) gives us the following results:

	Coefficients	**Standard Error**	**t Stat**	**P-value**
Intercept	0.646	0.968	0.667	0.529
gGDP	-0.004	0.227	-0.019	0.985
dUnemp	-0.027	0.043	-0.614	0.562
dCPI	0.072	0.492	0.146	0.889
R Square	0.061			
Adjusted R Square	-0.409			

Table 5: Multiple linear regression regression analysis

From the table above we can see that the R-square, which is the significance level, is only 6.1$\%$ which is significantly low, meaning the the variation in the changes in PD is explained by the model. The signs of the coefficients are as expected, except for changes in the unemployment rate, which has a negative sign. We expect an increase in the unemployment rate to lead to an increase in PD, i.e. a positive sense. Therefore, in this case it wouldn't make economic sense to use this multi-variable regression to obtain the EAC.

However, if even if all the signs made economic sense, what we ultimately want is to select the variable(s) that show the most impact on PD in order to avoid an over-fit model. This can be done by doing simple (one-variable) linear regression and performing different variable selection tests as done below. Alternatively, the multi-variable regression model would assist in terms of showing us how the different explanatory variables compare to each other in terms of the effect they have on PD.

Thus in order to estimate the PD term structure, we would substitute the macroeconomic forecasts' values into the multi-variable regression model. In our case we choose to use the simple linear regression model for each macroeconomic factor because our multiple linear regression model did not make economic sense in terms of the coefficient signs of one of the macroeconomic variables.

GDP Growth rate

The macroeconomic regression model for the GDP growth rate is given by the following equation:

$$dPD = \beta_0 + \beta_1 \cdot gGDP + \varepsilon. \tag{54}$$

We use a plot, a correlation matrix as well as a simple regression analysis of the changes GDP growth rate and changes in PD in order to see the relationship between these two variables. Since the GDP growth rate is given quarterly, we will show the impact of these quarterly GDP growth rates on the changes in PD.

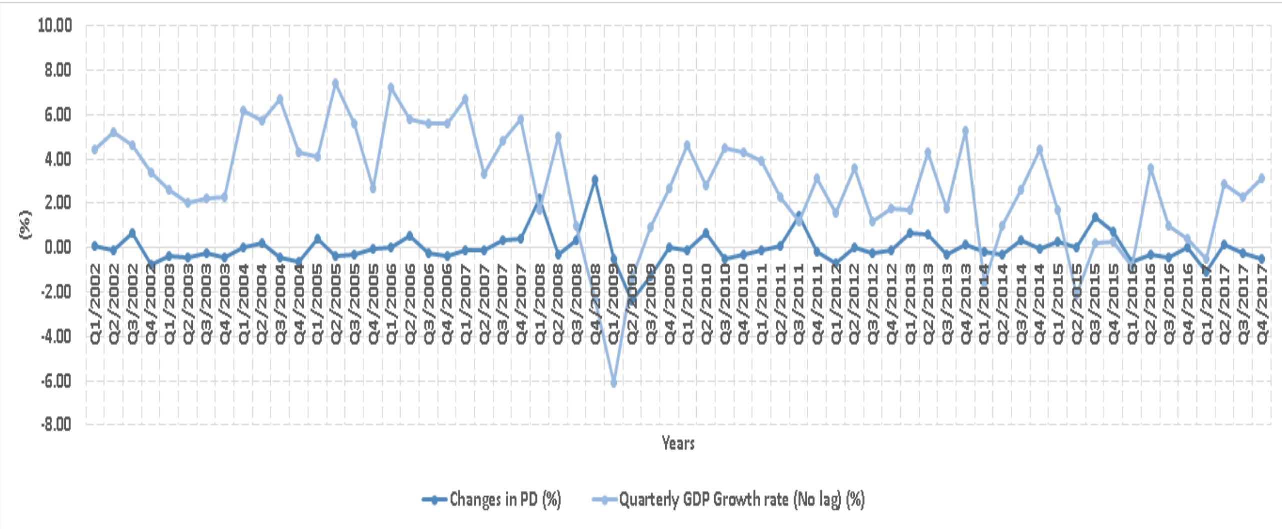


Figure 18: Quarterly GDP growth rate vs. PD

We expect that when GDP growth increases, bond spreads will decrease leading to a subsequent decrease in PD. In other words we expect a negative (opposite) relationship i.e. when there is a postive economic growth (increase in GDP growth rate), the risk of default ad thus the PD should decrease.

The plot above in Figure 18, though not clearly, shows this negative relation, meaning that we should also get a negative EAC. From visualization of the plot we see that the relationship in the plot is not perfectly negative, but that a lagged version of the changes in GDP growth rate would be moving almost perfectly opposite to the changes in PD.

We thus examine the correlation between the one, two, three and four quarter lags of changes in GDP growth rate with the changes in PD respectively. We then select the lag that has the highest correlation, which we will then use to run the regression analysis.

	Changes in PD (%)	**1Q lag**	**2Q lag**	**3Q lag**	**4Q lag**	**No lag**
Changes in PD (%)	1	-0.18	-0.20	-0.21	-0.31	-0.06
1Q lag	-0.18	1	0.58	0.45	0.29	0.56
2Q lag	-0.20	0.58	1	0.58	0.45	0.42
3Q lag	-0.21	0.45	0.58	1	0.58	0.26
4Q lag	-0.31	0.29	0.45	0.58	1	0.22
No lag	-0.06	0.56	0.42	0.26	0.22	1

Table 6: GDP growth rate correlation matrix

From the correlation matrix (Table 6) above we see that the quarterly changes in the GDP growth rate only have a significant effect on the changes in PD after a four quarter lag (i.e. after one year lag). This is expected as a change in the growth rate would take some time to fully materialize and subsequently impact PD. In other words, we don't expect an immediate response of PD to GDP growth rate.

Below is the plot of the four quarter lag of the GDP growth rate with changes in PD, which shows the negative relationship between more clearly in each quarter, compared to figure 18.

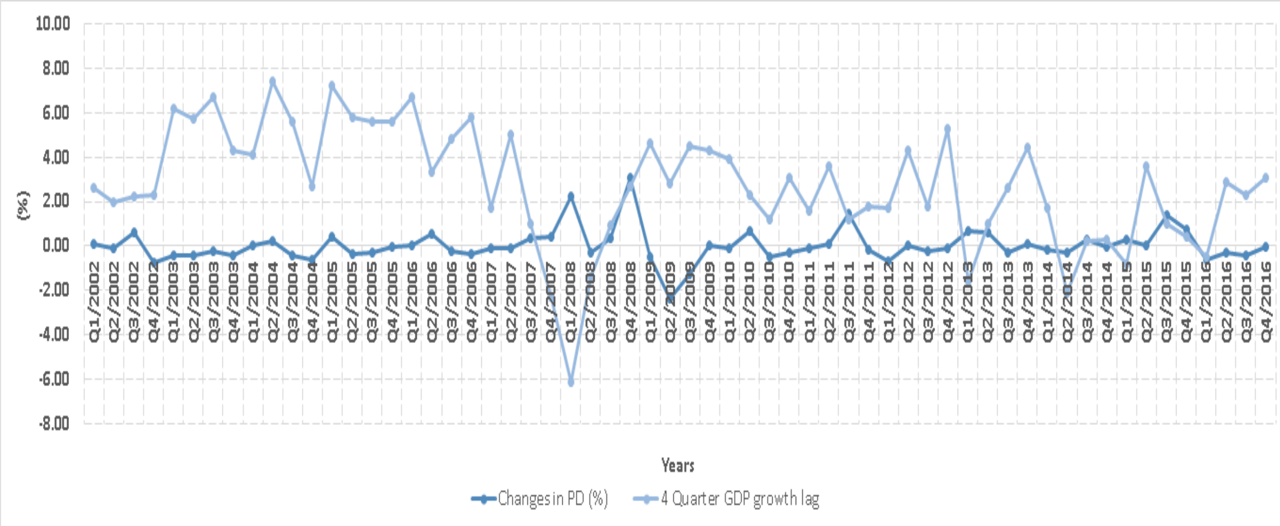


Figure 19: 4Q lag in GDP growth rate vs. PD

The figures below show the results from our regression analysis for no lag and four quarter lagged version of the GDP growth rate with changes in PD, respectively:

	Coefficients	Standard Error	t Stat	P-value
Intercept	0.012	0.140	0.083	0.934
Quarterly gGDP (No lag)	-0.008	0.037	-0.225	0.822
R Square	0.001			
Adjusted R Square	-0.015			

Table 7: GDP regression analysis with no lag

	Coefficients	Standard Error	t Stat	P-value
Intercept	0.270	0.135	2.008	0.049
gGDP 4Q lag	-0.093	0.036	-2.612	0.011
R Square	0.105			
Adjusted R Square	0.090			

Table 8: GDP regression analysis with 4Q lag

From above tables we see a notable increase in the level of significance (R Square) from 0.1% to 10.5%. Thus from Table 7 we see that EAC is given by -0.093, this means that a 1% positive change (increase) in the GDP growth rate leads to a 0.093% negative change (decrease) in PD.

Now that we have our EAC from the macroeconomic model given by equation 13 we can incorporate it into the initial one year PD transition matrix in order to estimate the multi-year PDs (Table 4). Since our initial transition matrix is for the year 2017, we want to estimate a 10-year PD term structure i.e from 2018 to 2027. Thus we will have to estimate 10 PD transition matrices using the 10-year macroeconomic forecasts that we plotted in the previous section.

We use the approaches (a), (c) as well as our own approach which are stated in the outline in Section 3.1 (v), in order to incorporate the EAC for GDP growth rate. As mentioned earlier, we will be using three macroeconomic scenarios namely, baseline, downturn and upturn. Below we show the resulting PD term structures[23]:

Approach (a) - GDP growth rate

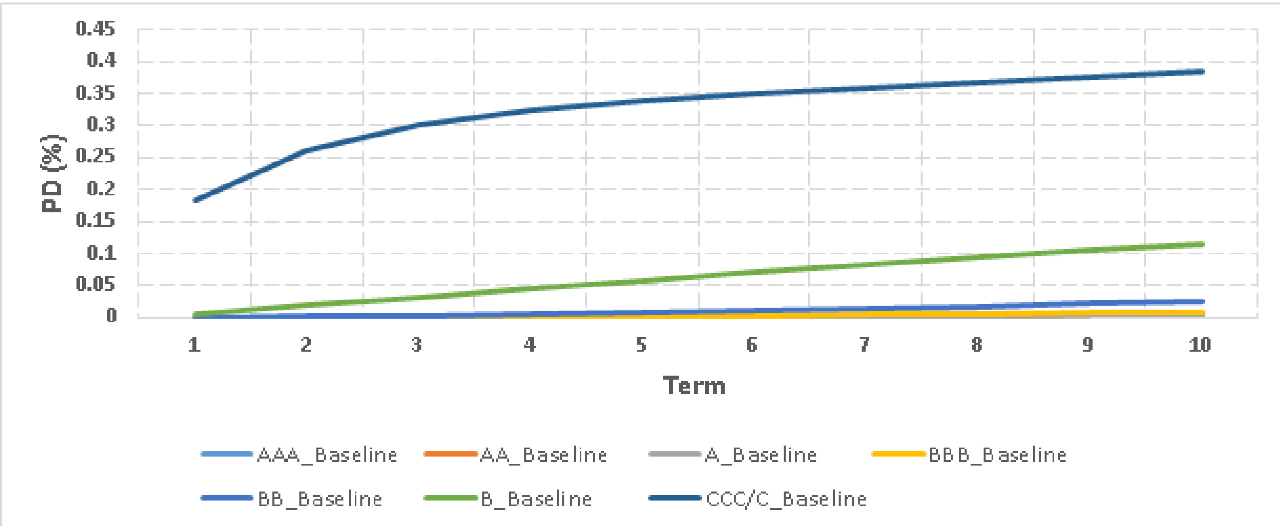


Figure 20: Approach (a) for baseline GDP growth

[23]Note: The values on y-axis for figures 20 to 31 are multiples of 100, e.g. 0.15 represents 15%

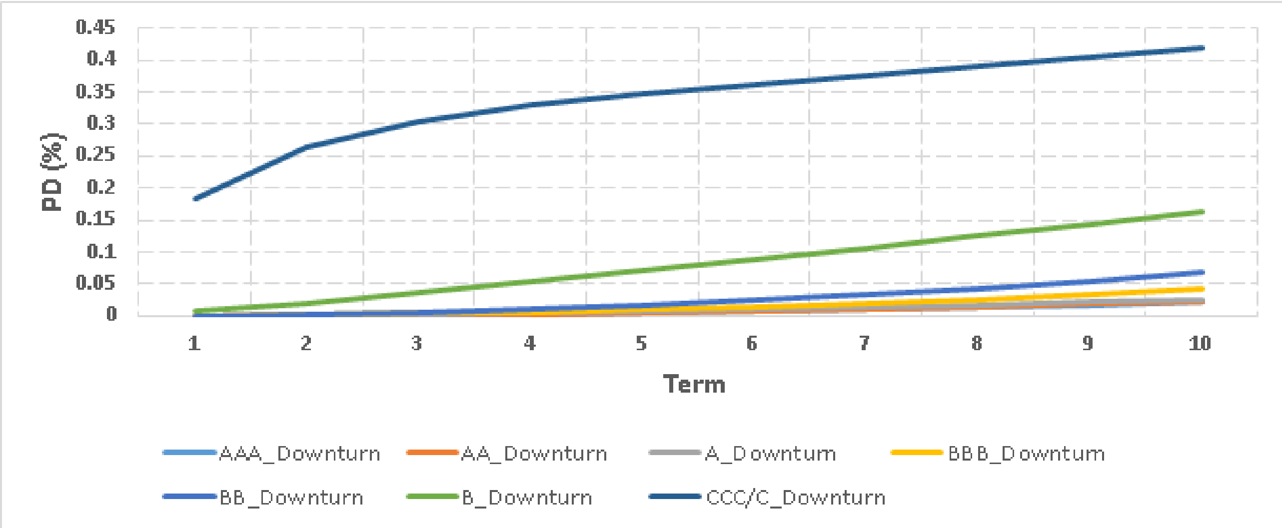


Figure 21: Approach (a) for downturn GDP growth

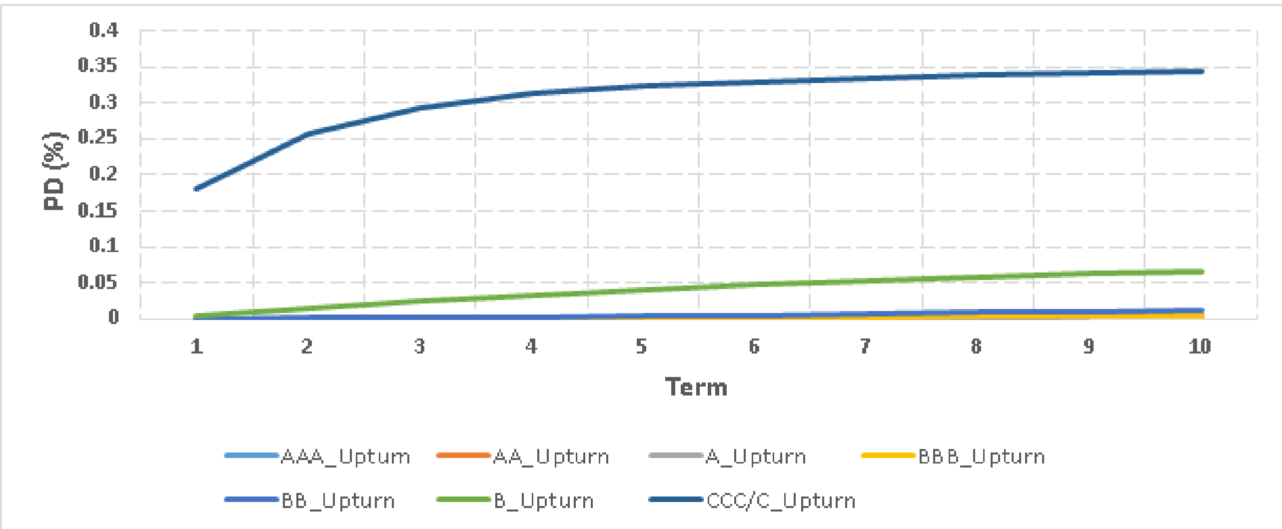


Figure 22: Approach (a) for upturn GDP growth

From the PD term structures above we can see that when there is a downturn, the PD is higher than the baseline PD. For example by just looking at the "CCC/C" rating in isolation, we see that the baseline PD is between 35% and 40%, whilst the downturn PD is between 40% and 45%.

The opposite can also be seen, i.e. when there is an upturn the PD is lower than the baseline, which is what we expect. Again by looking at the "CCC/C" rating only, the PD lies between 30% and 35%, which has a lower range than the baseline PD range which is between 35% and 40%.

This can clearly be seen in the figure below where we have isolated the PDs for the "CCC/C" rating grade under the three different macroeconomic scenarios.

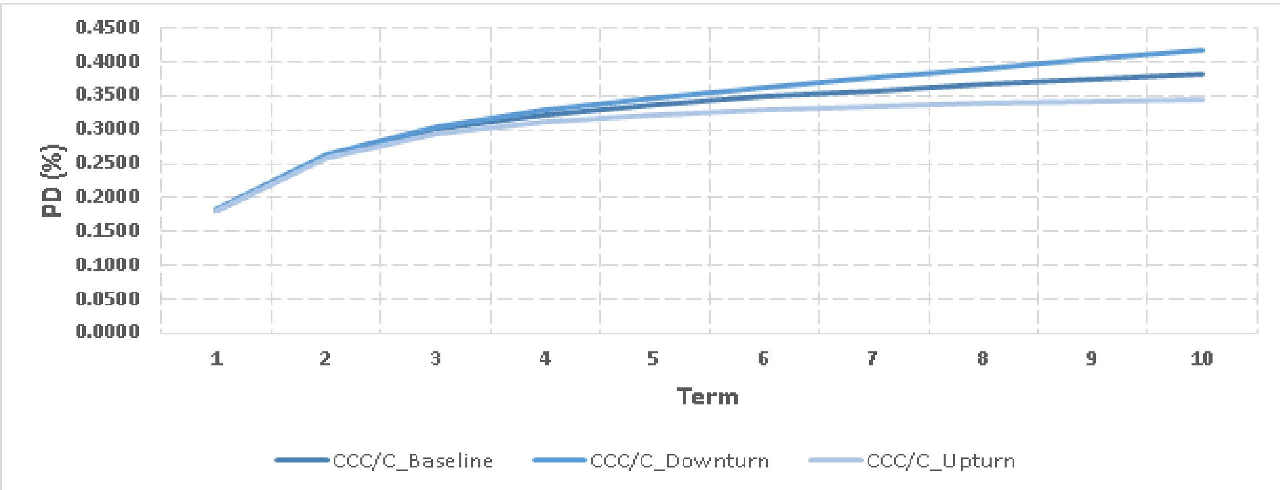


Figure 23: Approach (a) for CCC/C rating grade

Approach (b) *- GDP growth rate*

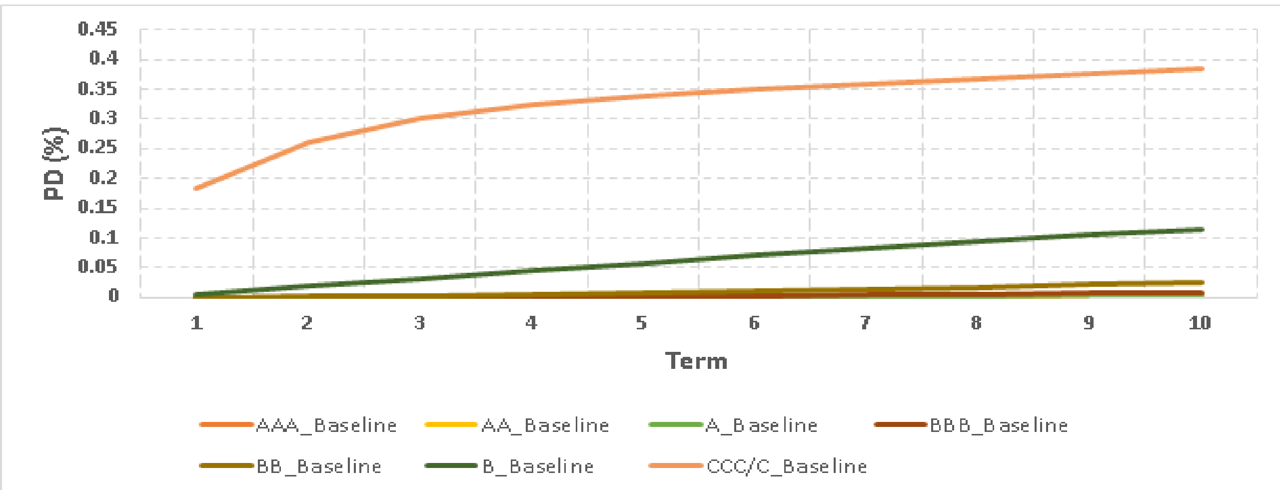


Figure 24: Approach (b) for baseline GDP growth

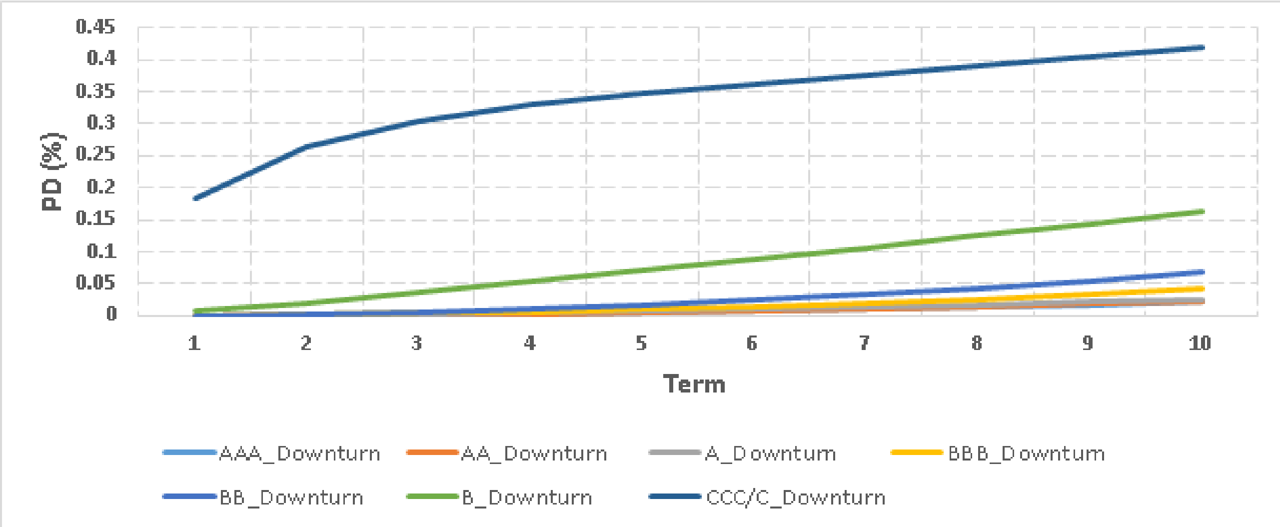


Figure 25: Approach (b) for downturn GDP growth

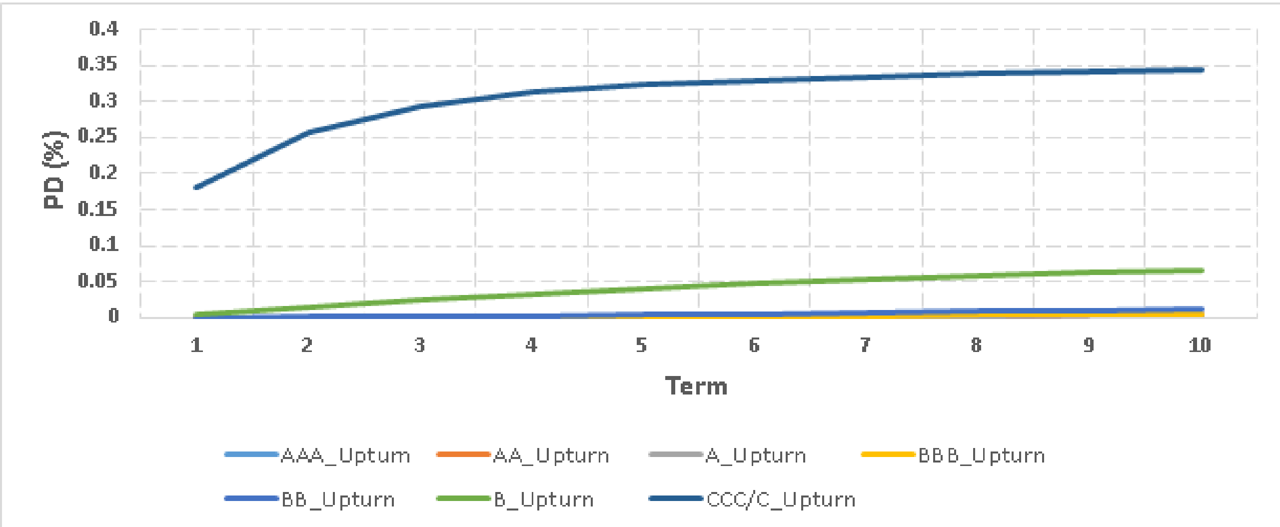


Figure 26: Approach (b) for upturn GDP growth

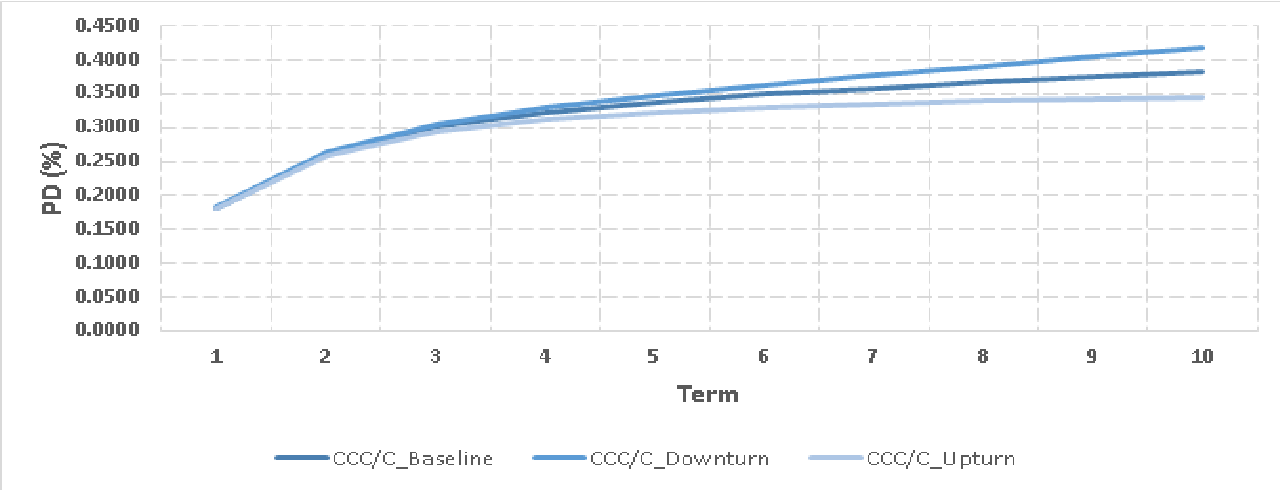


Figure 27: Approach (b) for CCC/C rating grade

The same analysis used in approach (a) holds for approach (b) as well with regards to the level of PD in the three different macroeconomic scenarios. The plots look very similar and in fact only differ by a few decimal places, but they are not the same. There is a small difference between the PD values which can't be seen from the plot because of rounding errors. To show that the PD values between approach a and b we have attached the PD term structure tables with the values for these approaches in the Appendix.

Own approach - GDP growth rate

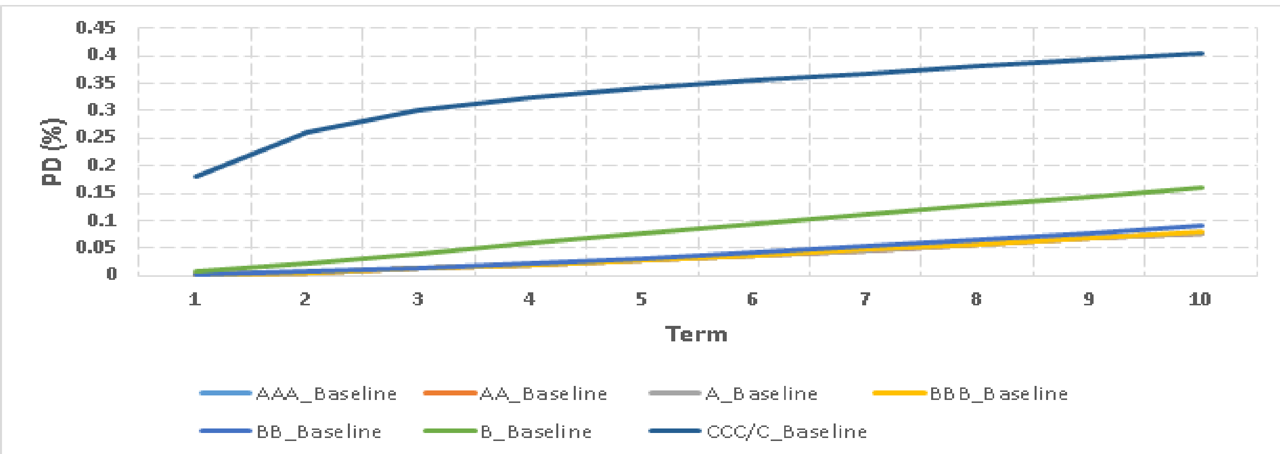


Figure 28: Own approach for baseline GDP growth

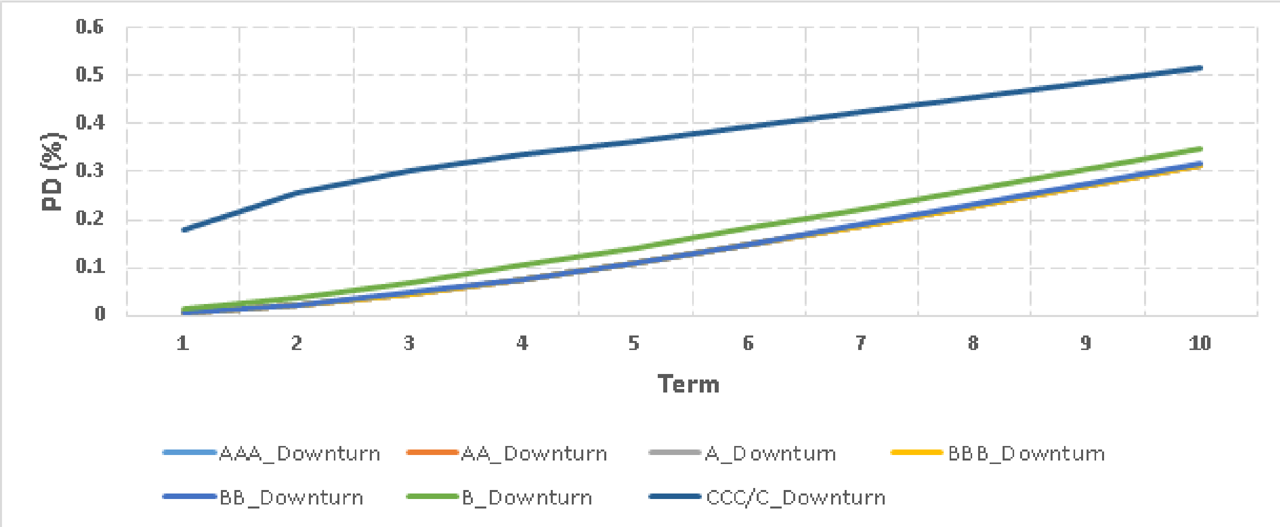


Figure 29: Own approach for downturn GDP growth

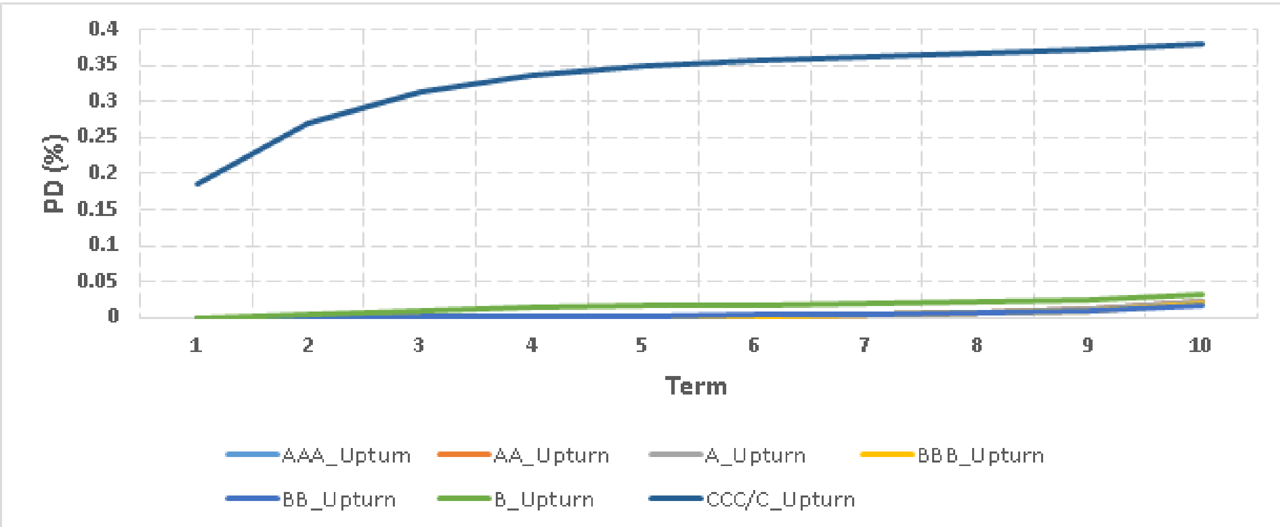


Figure 30: Own approach for upturn GDP growth

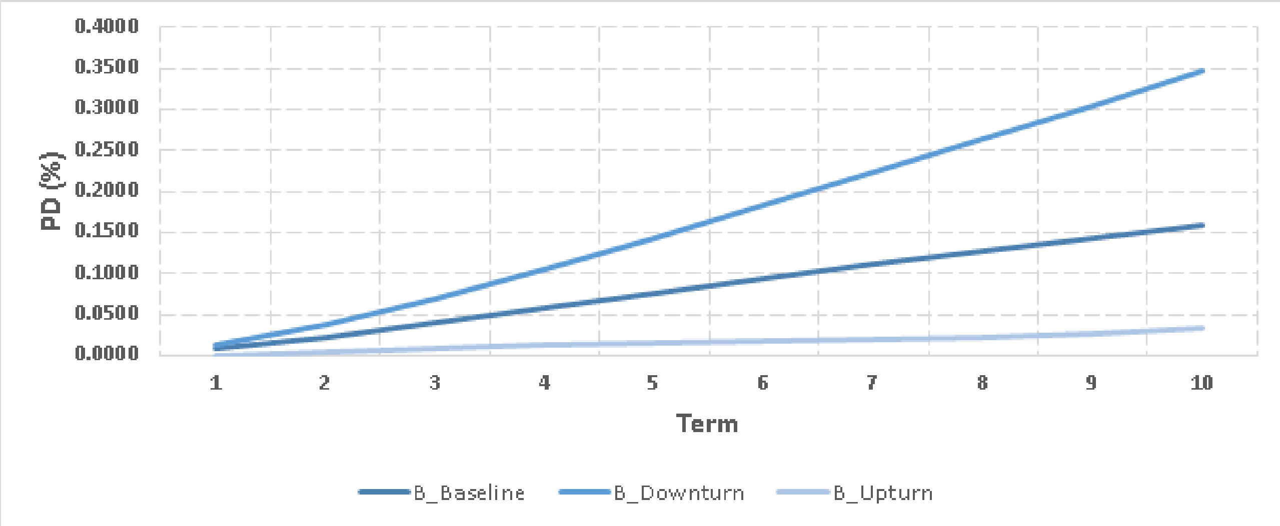


Figure 31: Own approach for CCC/C rating grade

We see that our approach also yields the desired results of the impact of the changes in GDP growth rate on changes in PD. Since our approach uses the whole effect of the EAC with no weighting, the impact on PD is higher than the other approaches. For instance, if we compare the downturn PD term structure of our approach with that of approach *a* and *b*, we note that the one from our approach is more steeper. This shows how important the weighting of the effect of EAC. As mentioned earlier, the choice of the direction of the weighting is fully dependent on the practitioner and the application.

We do the same thing for changes in the unemployment and CPI using our own approach and obtain the following results:

Own approach - Unemployment rate

The macroeconomic regression model for changes in the unemployment rate is given by:

$$dPD = \beta_0 + \beta_1 \cdot dUnemp + \varepsilon. \tag{55}$$

Below is the plot that shows the relation between the annual changes in the unemployment rate and changes in PD:

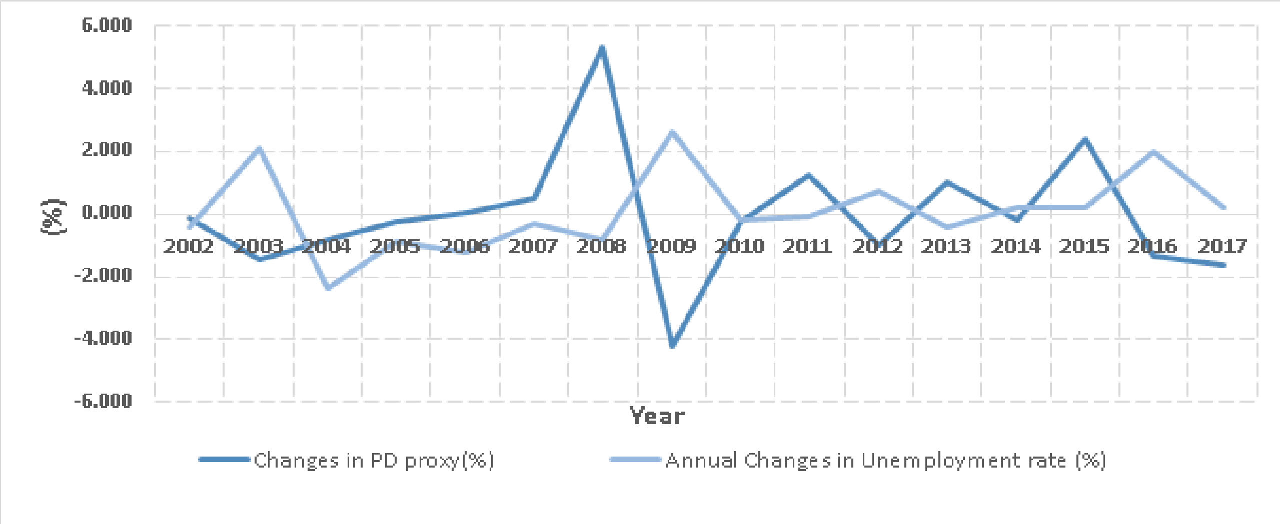


Figure 32: Changes in unemployment rate vs. changes PD

We expect a postive relationship between changes in unemployment rate and changes in PD. This means that an increase in unemployment rate would lead to an increase in risk of default by obligors and thus an increase in the PD. However, the plot shows a negative relationship, which doesn't make economic sense. From visualization, we can see that a one year lag in the changes in unemployment rate would yield the expected relation.

We thus do a correlation analysis with the one, two and three year lags of the changes in unemployment rate.

	Changes in PD	1 year lag	2 year lag	3 year lag	No lag
Changes in PD	1	0.607	0.069	0.038	-0.526
1 year lag	0.607	1	-0.185	0.106	-0.244
2 year lag	0.069	-0.185	1	0.104	0.127
3 year lag	0.038	0.106	0.104	1	-0.132
No lag	-0.526	-0.244	0.127	-0.132	1

Table 9: Unemployment rate correlation matrix

The correlation matrix shows that the one year lag of changes in the unemployment rate has the strongest and expected impact on changes in PD. This can be seen by the high and positive correlation coefficient of 0.607. We thus expect a postive EAC. The plot of the one year lag in changes in the unemployment rate against changes in PD is shown below:

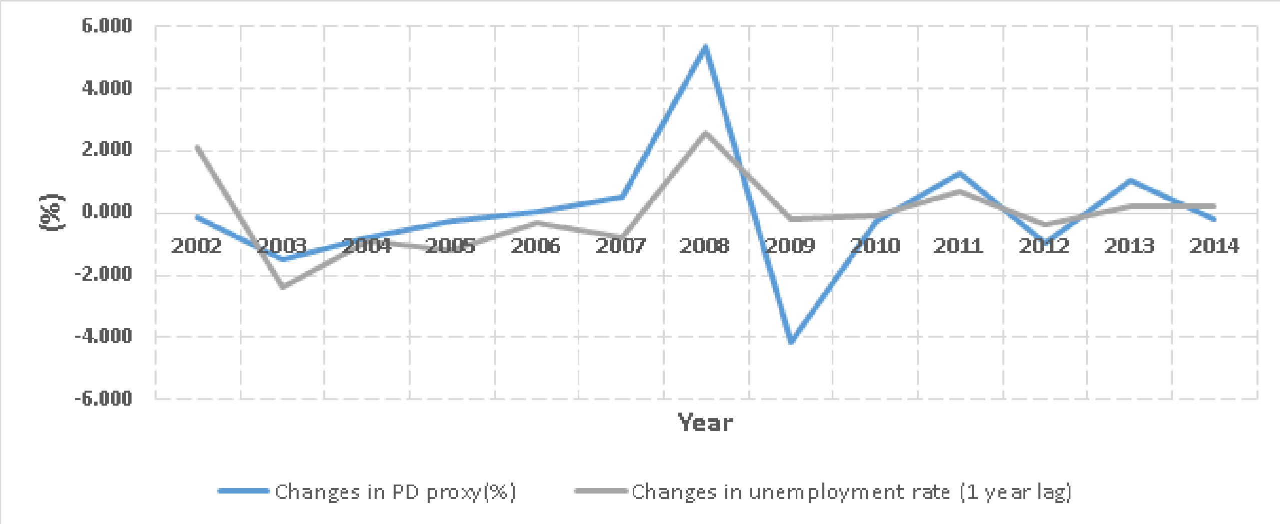


Figure 33: Changes in unemployment rate vs. changes PD (1 year lag)

Now the postive relation can be seen i.e. the changes in the unemployment rate and the changes in PD seem to move in the same direction most of the time. Thus when unemployment increases (decreases), the PD also increase (decrease). Thus we will use the EAC of the one year lag in changes in the unemployment rate because it makes economic sense and has the highest correlation with changes in PD.

The figures below show the results from our regression analysis for no lag as well as one year lagged version of the changes in the unemployment rate with PD, respectively:

	Coefficients	Standard Error	t Stat	P-value
Intercept	-0.085	0.519	-0.164	0.873
No lag	-0.841	0.410	-2.050	0.065
R Square	0.276			
Adjusted R Square	0.211			

Table 10: Regression analysis for unemployment rate (No lag)

We see from Table 9 above that the EAC doesn't have the correct sign, thus it wouldn't make sense to use it.

	Coefficients	Standard Error	t Stat	P-value
Intercept	0.024	0.484	0.049	0.962
1 year lag	0.971	0.384	2.530	0.028
R Square	0.368			
Adjusted R Square	0.310			

Table 11: Regression analysis for unemployment rate (1 year lag)

However, for the one year lag in changes in the unemployment rate, the EAC is postive and the level of significance (R Square) is 36.8%, which is relatively fair.

In the case of unemployment rate, we know that when the unemployment rate increase (decreases), PD also increases (decreases).Thus an upturn would be characterized by a decrease in unemployment rate and a downturn by an increase in unemployment rate.

Below is the PD term structure for a "AAA" rated company under the three different scenarios:

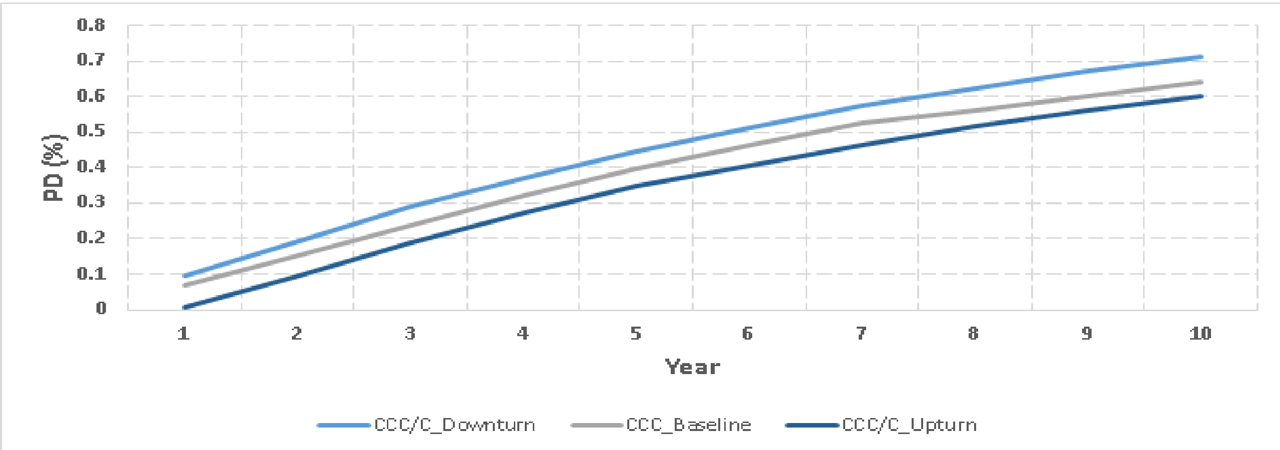


Figure 34: PD term structure using unemployment EAC

Own approach CPI

The macroeconomic regression model for changes in the CPI is given by:

$$dPD = \beta_0 + \beta_1 \cdot dCPI + \varepsilon. \tag{56}$$

Below is the plot that shows the relation between annual changes in the CPI and changes in PD:

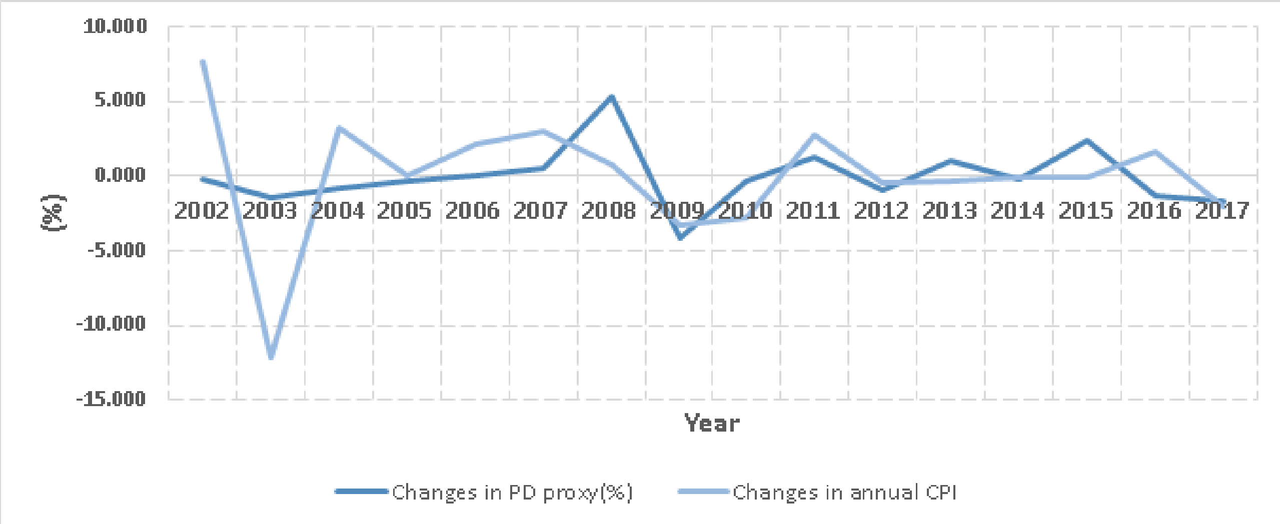


Figure 35: Changes in CPI vs. PD

The Consumer Price Index (CPI) measures monthly changes in prices for a range of consumer products meaning that changes in the CPI record the rate of inflation [17]. Thus we expect that when CPI increases, bond spreads will also increase, which leads to an increase in PD. This is because an increase in CPI is representative of inflation, thus leading to an increase in default risk. This positive relationship can be observed from Figure 31 above. We perform a correlation analysis to check if no lag on the changes in CPI has the highest impact on changes in PD.

	Changes in PD	**1 year lag**	**2 year lag**	**3 year lag**	**No lag**
Changes in PD	1	-0.103	-0.608	0.223	0.326
1 year lag	-0.103	1	-0.293	-0.119	-0.559
2 year lag	-0.608	-0.293	1	0.067	0.110
3 year lag	0.223	-0.119	0.067	1	-0.248
No lag	0.326	-0.559	0.110	-0.248	1

Table 12: CPI correlation matrix

The correlation matrix above shows us that there is no need to lag the changes in CPI as the correlation between changes in CPI with no lag and changes in PD is the highest. Therefore, we will use the EAC that we get from the regression of changes in CPI, with no lag, with changes in PD. Below is the regression results:

	Coefficients	Standard Error	t Stat	P-value
Intercept	-0.021	0.576	-0.036	0.972
No lag	0.147	0.129	1.144	0.277
R Square	0.106			
Adjusted R Square	0.025			

Table 13: Regression analysis for CPI (No lag)

From Table 12 we can see that the level of significance (R square) is 10.6% and the EAC is 0.147, which means that a percentage increase in changes in the CPI would lead to a 0.147% in changes in PD.

CPI, like the unemployment rate is also positively related with PD. This means that an upturn would be characterized by a decrease in the CPI and a downturn by an increase in CPI.

Below is the PD term structure for a "CCC/C" rated company under the three different scenarios:

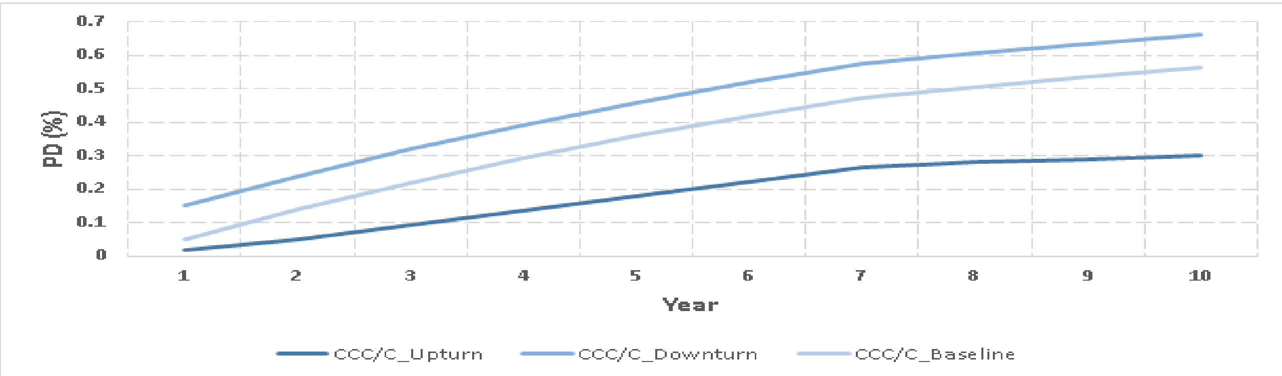


Figure 36: PD term structure using CPI EAC

As expected, as can be seen on the above figure, when there is an upturn (decrease in unemployment), the PD is lower (decreases) than the baseline PD and vice versa for a downturn.

Thus in summary, we see that all our term structures (values obtained from Tables 18 to 19) seem to be monotonically increasing, this is because our baseline forecast data for the macroeconomic factors, were either monotonically increasing or decreasing which causes a subsequent increase or decrease in the upturn or downturn PD values. For instance, the baseline GDP growth rate forecast was decreasing every year, thus leading to an increase in GDP each year shown by the the upward sloping PD term structure curves shown by figures 20 to 31. The same logic holds for the other macroeconomic factor forecast data.

5 ECL calculation

After the modelling of the credit risk parameters, PD, LGD as well as EAD, we want to be able to compute the lifetime ECL. We know that ECL is the product of these three credit risk parameters i.e. it is the present value of the amount expected to be lost on a financial asset [47] :

$$ECL = PV(PD \times LGD \times EAD) \tag{57}$$

where PV is the present value.

The tables[24] below show how we would go about computing the lifetime ECL under Stage 2 and Stage 3 using the three above mentioned credit risk parameters compared to the 12 month ECL under Stage 1, if we had the required data:

No.	T	PD	LGD	EAD (Rands)	Survival Prob.	PD(t-1 < D<=t)	EL	PV (EL)	Lifetime ECL (%)
1	1	0.12	0.0542	3448871246	1	0.12	23024023	21437638	100
								21437638	
	Interest rate	0.074							

Table 14: Example of Stage 1 ECL

No.	T	PD	LGD	EAD (Rands)	Survival Prob.	PD(t-1 < D<=t)	EL	PV (EL)	Lifetime ECL (%)
1	1	0.12	0.0542	1023390175	1	1	6831962	6361231	98.38
2	0.75	0.08	0.0008	767542631	0.88	0.05	50277	47656	0.74
3	0.5	0.06	0.0012	511695087	0.81	0.02	37131	35829	0.55
4	0.25	0.05	0.0016	255847543	0.76	0.01	21659	21276	0.33
								6465992	
	Interest rate	0.074							

Table 15: Example of Stage 2 ECL

No.	T	PD	LGD	EAD (Rands)	Survival Prob.	PD(t-1 < D<=t)	EL	PV (EL)	Lifetime ECL (%)
1	1	1	0.0542	41465378	1	1	2247423	2092573	97.04
2	0.75	1	0.0008	31099033	1	0.75	24879	23582	1.09
3	0.5	1	0.0012	20732689	1	0.50	24879	24007	1.11
4	0.25	1	0.0016	10366344	1	0.25	16586	16293	0.76
								2156455	
	Interest rate	0.074							

Table 16: Example of Stage 3 ECL

[24] Tables 13 to 15 were adapted from [47] and they are meant to illustrate how to compute ECL if we had all the three risk factors (PD, LGD and EAD).

The values in the above tables was obtained from the following formulas.

- Survival probability was calculated using the following formula:

$$\text{Survival Prob.} = \prod_{t=1}^{T}(1 - PD(t-1)) \tag{58}$$

- $PD(t-1 < D \leq t)$ is given by the following equation:

$$PD(t-1 < D \leq t) = PD \times SurvivalProb. \times T \tag{59}$$

- EL is calculated by the following equation:

$$EL = PD \times LGD \times EAD \tag{60}$$

- The present value of EL is given by the following equation:

$$PV(EL) = EL \times DF \tag{61}$$

  where DF is the discount factor. We choose it to be $(1+IR)^{-T}$.

- Finally, the ECL percentage is calculated by taking the specific $PV(EL)$ as a fraction of the total $PV(EL)$.

In summary, it is important to note some key differences between Tables 14 to 16 with regards to values of the PD. We can see that in stage 1 (Table 14), there is only one value for PD, which is the 12 month PD and is relatively low because assets in this category are considered to be performing and have the least risk.

In stage 2 (Table 15), we see that now there is more than one time period (T = $1, 0.75, 0.5, 0.25$). There is a term structure of PDs actually increases as time goes on, showing that the assets under this cateory are under-performing and thus lead to increased ECL.

Finally, in stage 3 (Table 16), it's clear that the assets are credit impaired as the PD for all time periods is 1. This then also leads to an increase in EL over time as can be seen by the values in the table.

6 Conclusion and further work

The new IFRS 9 guidelines have a significant impact on the modelling of ECL of financial instruments. The incorporation macroeconomic factors in the credit risk parameters to make ECL forward looking leads to an overall increase in ECL and thus helps banks to be better prepared and make provisions for future losses. In this thesis we were able to compute the PIT term structure of PDs and we observed how certain macroeconomic factors affect the PD level.

We were able to note that PD is sensitive to the macroeconomic forecasts i.e. under unfavourable macroeconomic conditions (i.e. a decrease in GDP or increase in the unemployment rate) the PD is higher and lower under favourable economic conditions (i.e. an increase in GDP or decrease in the unemployment rate. Though this new IFRS 9 framework requires a lot of data, given PD, LGD, and EAD,calculating expected credit losses is a trivial matter.

The choice of which modelling approach to use for ECL estimation is unique to financial institution, and won't be the same for all financial institutions as they have different ways of allocating their portfolios to different stages of credit deterioration under the ECL model and selecting relevant economic scenarios.

With regards to PD modelling, further work would be relaxing the assumption of the default state being absorbing. This means that should an obligor default they can recover. Another further work would be including the assumption that PDs tend to decrease over time. For instance, the closer we are to the maturity date, the lower that PD of a client is, given the client continuously fulfills their contractual obligations.

With regards to LGD and EAD modelling, given real data, further work would be to test the methodologies mentioned in this paper and determining which is most accurate, and thus together with PD computing the ECL.

7 Appendix

7.1 Data

Year	Quarters	Quarterly Gov Bond Spreads (bps)	Quarterly Gov Bond Spreads (%)	PD proxy(%)	Changes in PD (%)
2007	31-03-07	36.48	0.365	0.608	-0.097
	30-06-07	28.56	0.286	0.476	-0.132
	30-09-07	49.77	0.498	0.829	0.353
	31-12-07	72.91	0.729	1.215	0.386
2008	31-03-08	207.43	2.074	3.457	2.242
	30-06-08	190.70	1.907	3.178	-0.279
	30-09-08	210.68	2.107	3.511	0.333
	31-12-08	393.56	3.936	6.559	3.048
2009	31-03-09	362.42	3.624	6.040	-0.519
	30-06-09	218.53	2.185	3.642	-2.398
	30-09-09	142.10	1.421	2.368	-1.274
	31-12-09	142.41	1.424	2.374	0.005
2010	31-03-10	135.74	1.357	2.262	-0.111
	30-06-10	174.64	1.746	2.911	0.648
	30-09-10	145.08	1.451	2.418	-0.493
	31-12-10	126.12	1.261	2.102	-0.316
2011	31-03-11	120.38	1.204	2.006	-0.096
	30-06-11	123.87	1.239	2.064	0.058
	30-09-11	211.51	2.115	3.525	1.461
	31-12-11	202.07	2.021	3.368	-0.157
2012	31-03-12	160.39	1.604	2.673	-0.695
	30-06-12	163.00	1.630	2.717	0.043
	30-09-12	149.25	1.493	2.488	-0.229
	31-12-12	142.82	1.428	2.380	-0.107
2013	31-03-13	181.33	1.813	3.022	0.642
	30-06-13	216.33	2.163	3.605	0.583
	30-09-13	197.23	1.972	3.287	-0.318
	31-12-13	204.11	2.041	3.402	0.115
2014	31-03-14	194.80	1.948	3.247	-0.155
	30-06-14	176.63	1.766	2.944	-0.303
	30-09-14	195.32	1.953	3.255	0.312
	31-12-14	191.79	1.918	3.197	-0.059
2015	31-03-15	208.74	2.087	3.479	0.282
	30-06-15	209.39	2.094	3.490	0.011
	30-09-15	292.49	2.925	4.875	1.385
	31-12-15	335.07	3.351	5.585	0.710
2016	31-03-16	298.90	2.989	4.982	-0.603
	30-06-16	281.04	2.810	4.684	-0.298
	30-09-16	256.34	2.563	4.272	-0.412
	31-12-16	255.67	2.557	4.261	-0.011
2017	31-03-17	191.02	1.910	3.184	-1.077
	30-06-17	200.48	2.005	3.341	0.158
	30-09-17	186.33	1.863	3.106	-0.236
	31-12-17	156.97	1.570	2.616	-0.489

Table 17: PD proxy data (2002 - 2027)

Year	GDP growth rate (%)	Unemployment rate (%)	Unemployment rate (% change)	CPI (%)	CPI (% change)
2007		22.3		8.8	
2008	3.62	24.83	2.53	8.01	-0.79
2009	-1.53	24.59	22.06	8.37	0.36
2010	3.14	24.43	2.37	7.65	-0.72
2011	3.6	24.36	21.99	5.87	-1.78
2012	2.47	24.36	2.37	4.46	-1.41
2013	1.89	24.41	22.04	5.85	1.39
2014	2.55	24.41	2.37	6.06	0.21
2015	3.23	24.41	22.04	5.84	-0.22
2016	5.56	24.22	2.18	4.12	-1.72
2017	5.75	23.9	21.72	4.02	-0.1
2018	5.5	23.5	1.78	4.17	0.15
2019	5.24	23.05	21.27	4.3	0.13
2020	5.07	22.59	1.32	4.39	0.09
2021	4.97	22.13	20.81	4.44	0.05
2022	4.9	21.68	0.87	4.47	0.03
2023	4.85	21.25	20.38	4.49	0.02
2024	4.8	20.86	0.48	4.5	0.01
2025	4.75	20.5	20.02	4.51	0.01
2026	4.68	20.17	0.15	4.51	0
2027	4.6	19.88	19.73	4.51	0

Table 18: Historical data for macroeconomic factors

Term Structure Approach A (Baseline)										
Term	1	2	3	4	5	6	7	8	9	10
AAA	0.0003	0.0007	0.0010	0.0015	0.0019	0.0024	0.0029	0.0034	0.0040	0.0047
AA	0.0003	0.0007	0.0010	0.0015	0.0019	0.0024	0.0029	0.0035	0.0041	0.0048
A	0.0003	0.0007	0.0010	0.0015	0.0020	0.0025	0.0031	0.0037	0.0043	0.0050
BBB	0.0003	0.0007	0.0012	0.0018	0.0026	0.0035	0.0046	0.0058	0.0073	0.0089
BB	0.0003	0.0010	0.0025	0.0045	0.0071	0.0102	0.0137	0.0177	0.0219	0.0265
B	0.0066	0.0182	0.0314	0.0447	0.0577	0.0703	0.0822	0.0937	0.1047	0.1153
CCC/C	0.1821	0.2622	0.3010	0.3227	0.3372	0.3486	0.3583	0.3673	0.3756	0.3836
D	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Term Structure Approach A (Downturn)										
Term	1	2	3	4	5	6	7	8	9	10
AAA	0.0003	0.0008	0.0017	0.0030	0.0048	0.0071	0.0099	0.0134	0.0173	0.0219
AA	0.0003	0.0009	0.0020	0.0035	0.0055	0.0081	0.0111	0.0147	0.0189	0.0236
A	0.0004	0.0013	0.0027	0.0045	0.0069	0.0097	0.0131	0.0170	0.0215	0.0267
BBB	0.0006	0.0018	0.0038	0.0065	0.0100	0.0144	0.0197	0.0260	0.0332	0.0415
BB	0.0008	0.0027	0.0060	0.0107	0.0169	0.0245	0.0334	0.0437	0.0552	0.0680
B	0.0073	0.0204	0.0360	0.0527	0.0700	0.0878	0.1059	0.1245	0.1435	0.1629
CCC	0.1831	0.2646	0.3052	0.3294	0.3470	0.3621	0.3763	0.3903	0.4043	0.4185
D	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Term Structure Approach A (Upturn)										
Term	1	2	3	4	5	6	7	8	9	10
AAA	0.0003	0.0007	0.0010	0.0014	0.0018	0.0022	0.0026	0.0031	0.0035	0.0039
AA	0.0003	0.0007	0.0010	0.0014	0.0018	0.0022	0.0026	0.0031	0.0035	0.0039
A	0.0003	0.0007	0.0010	0.0014	0.0018	0.0022	0.0026	0.0031	0.0035	0.0039
BBB	0.0003	0.0007	0.0010	0.0015	0.0019	0.0024	0.0029	0.0035	0.0041	0.0048
BB	0.0003	0.0009	0.0017	0.0029	0.0042	0.0056	0.0071	0.0087	0.0103	0.0120
B	0.0055	0.0146	0.0242	0.0331	0.0407	0.0472	0.0528	0.0578	0.0625	0.0667
CCC/C	0.1805	0.2583	0.2941	0.3123	0.3228	0.3296	0.3345	0.3387	0.3422	0.3454
D	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 19: Approach A PD term structure with GDP EAC

Term Structure Approach B (Baseline)										
Term	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AAA	0.0003	0.0007	0.0010	0.0015	0.0019	0.0024	0.0029	0.0034	0.0040	0.0047
AA	0.0003	0.0007	0.0010	0.0015	0.0019	0.0024	0.0029	0.0035	0.0041	0.0048
A	0.0003	0.0007	0.0010	0.0015	0.0020	0.0025	0.0031	0.0037	0.0043	0.0050
BBB	0.0003	0.0007	0.0012	0.0018	0.0026	0.0035	0.0046	0.0058	0.0073	0.0089
BB	0.0003	0.0010	0.0025	0.0045	0.0071	0.0102	0.0138	0.0177	0.0220	0.0266
B	0.0066	0.0182	0.0314	0.0447	0.0578	0.0703	0.0823	0.0938	0.1048	0.1154
CCC/C	0.1821	0.2623	0.3010	0.3227	0.3373	0.3486	0.3584	0.3674	0.3757	0.3837
D	1	1	1	1	1	1	1	1	1	1
Term Structure Approach B (Downturn)										
Term	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AAA	0.0003	0.0008	0.0017	0.0030	0.0048	0.0071	0.0099	0.0132	0.0171	0.0215
AA	0.0003	0.0009	0.0020	0.0035	0.0055	0.0080	0.0110	0.0145	0.0185	0.0231
A	0.0004	0.0013	0.0027	0.0045	0.0068	0.0096	0.0129	0.0167	0.0211	0.0260
BBB	0.0006	0.0018	0.0038	0.0064	0.0098	0.0141	0.0192	0.0252	0.0322	0.0402
BB	0.0008	0.0026	0.0059	0.0105	0.0166	0.0240	0.0328	0.0429	0.0543	0.0669
B	0.0073	0.0204	0.0360	0.0527	0.0701	0.0879	0.1062	0.1249	0.1440	0.1636
CCC/C	0.1831	0.2647	0.3053	0.3296	0.3473	0.3625	0.3768	0.3909	0.4051	0.4194
D	1	1	1	1	1	1	1	1	1	1
Term Structure Approach B (Upturn)										
Term	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AAA	0.0003	0.0007	0.0010	0.0014	0.0018	0.0022	0.0026	0.0031	0.0035	0.0039
AA	0.0003	0.0007	0.0010	0.0014	0.0018	0.0022	0.0026	0.0031	0.0035	0.0039
A	0.0003	0.0007	0.0010	0.0014	0.0018	0.0022	0.0026	0.0031	0.0035	0.0039
BBB	0.0003	0.0007	0.0010	0.0015	0.0019	0.0024	0.0029	0.0035	0.0041	0.0047
BB	0.0003	0.0009	0.0017	0.0029	0.0042	0.0056	0.0070	0.0086	0.0102	0.0118
B	0.0055	0.0146	0.0242	0.0330	0.0407	0.0471	0.0526	0.0577	0.0622	0.0664
CCC/C	0.1805	0.2583	0.2940	0.3121	0.3225	0.3293	0.3342	0.3383	0.3418	0.3450
D	1	1	1	1	1	1	1	1	1	1

Table 20: Approach B PD term structure with GDP EAC

Own Approach term Structure (Baseline)										
Term	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AAA	0.0020	0.0065	0.0127	0.0202	0.0284	0.0373	0.0467	0.0566	0.0670	0.0781
AA	0.0020	0.0065	0.0127	0.0202	0.0284	0.0373	0.0467	0.0566	0.0670	0.0781
A	0.0020	0.0065	0.0127	0.0202	0.0284	0.0373	0.0467	0.0566	0.0670	0.0781
BBB	0.0020	0.0065	0.0127	0.0202	0.0286	0.0377	0.0473	0.0575	0.0683	0.0798
BB	0.0020	0.0068	0.0137	0.0223	0.0321	0.0427	0.0540	0.0658	0.0783	0.0913
B	0.0083	0.0233	0.0408	0.0589	0.0768	0.0942	0.1110	0.1274	0.1434	0.1592
CCC/C	0.1809	0.2604	0.2999	0.3235	0.3407	0.3551	0.3680	0.3802	0.3919	0.4034
D	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Own Approach term Structure (Downturn)										
Term	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AAA	0.0078	0.0235	0.0466	0.0756	0.1094	0.1468	0.1868	0.2283	0.2705	0.3130
AA	0.0078	0.0235	0.0466	0.0756	0.1094	0.1468	0.1868	0.2283	0.2705	0.3130
A	0.0078	0.0235	0.0466	0.0756	0.1094	0.1468	0.1868	0.2283	0.2706	0.3130
BBB	0.0078	0.0235	0.0466	0.0756	0.1095	0.1470	0.1870	0.2285	0.2709	0.3133
BB	0.0078	0.0238	0.0474	0.0771	0.1116	0.1496	0.1900	0.2317	0.2741	0.3165
B	0.0138	0.0386	0.0699	0.1051	0.1429	0.1824	0.2229	0.2639	0.3051	0.3460
CCC/C	0.1783	0.2576	0.3017	0.3347	0.3649	0.3949	0.4251	0.4555	0.4860	0.5163
D	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Own Approach term Structure (Upturn)										
Term	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AAA	0.0003	0.0007	0.0011	0.0017	0.0023	0.0033	0.0047	0.0071	0.0117	0.0216
AA	0.0003	0.0007	0.0011	0.0017	0.0023	0.0033	0.0047	0.0071	0.0117	0.0218
A	0.0003	0.0007	0.0011	0.0016	0.0023	0.0033	0.0047	0.0071	0.0118	0.0221
BBB	0.0003	0.0007	0.0011	0.0016	0.0023	0.0032	0.0046	0.0067	0.0108	0.0193
BB	0.0003	0.0007	0.0012	0.0019	0.0026	0.0035	0.0047	0.0067	0.0104	0.0179
B	0.0003	0.0048	0.0098	0.0137	0.0161	0.0181	0.0199	0.0223	0.0261	0.0334
CCC/C	0.1858	0.2708	0.3132	0.3358	0.3485	0.3565	0.3621	0.3669	0.3722	0.3798
D	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 21: Own approach PD term structure with GDP EAC

References

[1] URL: https://www.bankinghub.eu/banking/finance-risk/challenges-ifrs-9-impairment.

[2] URL: https://www.investopedia.com/terms/d/defaultprobability.asp.

[3] URL: https://www.medcalc.org/manual/logisticregression.php.

[4] URL: https://en.wikipedia.org/wiki/Logistic_regression.

[5] URL: https://www.investopedia.com/terms/e/exposureatdefault.asp.

[6] URL: https://en.wikipedia.org/wiki/Jarrow%E2%80%93Turnbull_model#cite_note-1.

[7] URL: https://www.investopedia.com/terms/j/jarrowturnbull.asp.

[8] URL: https://investinganswers.com/financial-dictionary/financial-statement-analysis/bad-debt-reserve-4646.

[9] URL: https://www.pwc.com/gx/en/audit-services/ifrs/publications/ifrs-9/ifrs-in-depth-expected-credit-losses.pdf.

[10] URL: https://www.spratings.com/documents/20184/774196/2017+Annual+Global+Corporate+Default+Study/a4cffa07-e7ca-4054-9e5d-b52a627d8639.

[11] URL: http://www.kamakuraco.com/Blog/tabid/231/EntryId/700/Credit-Spreads-and-Default-Probabilities-A-Simple-Model-Validation-Example.aspx.

[12] URL: https://knoema.com/vwkvxbe/south-africa-unemployment-forecast-2015-2020-and-up-to-2060-data-and-charts.

[13] URL: https://knoema.com/sgdhjif/south-africa-inflation-forecast-2015-2020-and-up-to-2060-data-and-charts.

[14] URL: http://comstat.comesa.int/egagikg/south-africa-gdp-growth-forecast-2015-2020-and-up-to-2060-data-and-charts.

[15] URL: https://en.wikipedia.org/wiki/Beta_distribution.

[16] URL: https://en.wikipedia.org/wiki/Exposure_at_default.

[17] URL: http://www.statssa.gov.za/?p=955.

[18] Economic Outlook No 95. *Long-term Baseline Projections*. Organisation for Economic Co-operation and Development, 2014.

[19] Helmut Siller & Pietro Andrea. *Incurred losses vs. expected losses: a critical comparison*. 2015.

[20] *Cash shortfall & LGD- Two sides of the same coin*. Aptivaa Blog.

[21] *Chapter 5 Credit risk*. Department of Applied Mathematics and Statistics. URL: www.iam.fmph.uniba.sk/institute/jurca/qrm/Chapter5.pdf.

[22] *Credit risk estimation techniques*. CRISIL (Global research and analytics).

[23] Mindy Leow & Jonathan Crook. *Estimation of Credit Card Exposure at Default (EAD)*. Credit Research Centre, 2013.

[24] "Crystal Gazing– Estimating Lifetime PDs". In: *Aptivaa Blog* 10 (2016).

[25] Erika Spuchl'akova & Juraj Cug. "Credit Risk and LGD Modelling". In: *Procedia Economics and Finance* 23 (2014), pp. 439–444.

[26] Tomas Kliestik & Juraj Cug. "Comparison of selected mmodel of credit risk". In: *Elsevier Procedia Economic and Finance* 23 (2015).

[27] Vaněk Tomáš & Hampel David. "The Probability of Default Under IFRS 9: Multi-period Estimation and Macroeconomic Forecast". In: *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunenslo* 65.2 (2017), pp. 759–776.

[28] "Demystifying PD Terminologies". In: *Aptivaa Blog* 8 (2016).

[29] *Evaluating the economic value of loans and the implications: Toward transformation of the business model of the banks and nonbank firms*. Bank of Japan quarterly bulletin, 2003.

[30] *Expected credit loss calculation methodology & provisioning according to IFRS9*. Deloitte, 2017.

[31] "Exposure at Default: IFRS 9 Ramifications". In: *Aptivaa Blog* 12 (2016).

[32] "From IAS 39 to IFRS 9". In: *Aptivaa Blog* 1 (2016).

[33] Jon Frye. *Loss given default as a function of the default rate*. 2013.

[34] David Kwashie Garr. "Determinants of Credit Risk in the Banking Industry of Ghana". In: *ISSN 2224-607X* 3.11 (2013).

[35] Mojca Gornjak. "Comparison of IAS 39 and IFRS 9: The Analysis of Replacement". In: *International Journal of Management, Knowledge and Learning* 6.1 (2017), pp. 115–130.

[36] Rickard Gunnval. *Estimating Probability of Default Using Rating Migrations in Discrete and Continuous Time*.

[37] Martin Gürtler Marc & Hibbeln. "Improvements in loss given default forecasts for bank loans". In: *Journal of Banking & Finance* 37.7 (2013), p. 2354. DOI: 10.1016/j.jbankfin.2013.01.031.

[38] Jessica Hallblad. "The multi-year through- the- cycle and point- in- time of probability of default". MA thesis. Umea University, 2014.

[39] *IFRS 9 – Credit Modelling and Implementation*. PWC, 2015.

[40] *IFRS 9: Expected credit losses, In depth: A look at current financial reporting issues*. PWC, 2014. DOI: INT2014-06.

[41] "Impairment Modelling". In: *Aptivaa Blog* 6 (2016).

[42] Michael Jacobs Jr. "An option theoretic model for ultimate loss-given-default with systematic recovery risk and stochastic returns on defaulted debt". In: *BIS Papers chapters,in: Bank for International Settlements (ed.), Portfolio and risk management for central banks and sovereign wealth funds* 58 (2011), pp. 257–285.

[43] Dr. Ahmed Fouad Khalil. "Economic adjustment for probability of default". MA thesis. Union of Arab Banks, 2017.

[44] Halan Manoj Kumar. *Modelling loss given default - A practical methology*. 2016.

[45] Amal Merzouk. *Modeling challenges in IFRS 9 Standard*. 2016.

[46] Tomas Kliestik & Peter Adamko & Maria Misankova. *Reduced – form models used for the calculation of credit risk*. Research Gate, 2014. URL: https://www.researchgate.net/publication/284725818.

[47] Ferdinand Othieno. *IFRS 9: Financial Intruments (ECL Practical session)*. 2018.

[48] Jimmy Skoglund & Wei Chen & Martim Rocha. “Credit Risk Modelling in a new era”. In: *SAS Institute Inc.* (2016). DOI: PaperSAS6685-2016.

[49] Manuel Mayer & Florian Resch & Stephan Sauer. “Validating Point-in-Time vs. Through-the-Cycle Credit Rating Systems”. 2017.

[50] E.J Ross & Lynn Shibut. “What Drives Loss Given Default? Evidence From Commercial Real Estate Loans at Failed Banks”. In: *FDIC CFR WP 2015-03* (2015).

[51] Manfred Fruhwirth & Leopold Sogner. *The Jarrow/Turnbull default risk model :Evidence from the German Market*. 2001.

[52] “Stage Assessment- Devil is in the Detail”. In: *Aptivaa Blog* 5 (2016).

[53] *The implementation of IFRS 9 impairment requirements by banks*. EY.

[54] Ben Begin & Susie Thomas. *Basel II retail modelling approaches PD Models*. PWC, 2012.

[55] Robert A. Jarrow & David Lando & Stuart M. Turnbull. “A Markov Model for the Term Structure of Credit Risk Spreads”. In: *The Review of Financial Studies* 10.2 (1997). URL: http://www.jstor.org/stable/2962353.

[56] Kimmo Turunen. *Exposure-at-default models for counterparty credit risk under Basel frameworks*. Mat-2.4108 Independent Research Projects in Applied Mathematics, 2010.

[57] Tomas Vanek. “Economic adjustment of default probabilities”. In: *European Journal of Business science and technology* 2.2 (2016), pp. 122–130.

[58] George J. Benston & Larry D. Wall. *How should banks account for loan losses?* Federal reserve bank of Atlanta, 2004.