The management, mitigation and measurement of model risk in financial risk models

L van Biljon

orcid.org/0000-0002-2378-638X

Dissertation submitted in *partial* fulfillment of the requirements for the *Masters* degree in *BMI Risk Analysis* at the North-West University

Supervisor: Prof H Raubenheimer

Graduation May 2018

29231914

NWU®
NORTH-WEST UNIVERSITY
NOORDWES-UNIVERSITEIT
UNIBESITI YA BOKONE-BOPHIRIMA

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr Leendert Haasbroek for his contribution and support towards the shaping of this dissertation.

I am grateful for the guidance I received from my supervisor Prof Helgard Raubenheimer.

In addition, I thank Dr Kevin Panman for his input.

ABSTRACT

Financial risk models are simplifications of complex real-world phenomena used to better understand the intricate nature of an underlying process such as the loss generating process in financial risk analyses. Due to this simplification the model-representation of the real-world is by definition an approximation which implies a risk that the model may not fully reflect the real-world dynamics it is designed to mimic – this is referred to as model risk. This inherent model risk can probably not be eliminated completely since that would require an exact representation of the real-world which is arguably unattainable due to the complexity thereof. Model risk impacts a number of risk categories within financial risk, including operational risk, credit risk, strategic risk, reputational risk and more. In this dissertation a contribution is made to the management, mitigation and measurement of model risk of financial risk models. The contributions include the proposal of a standardised definition of model risk through the categorisation of model risk types based on definitions of model risk available in literature, summarising model risk mitigating techniques, the development of a practical and repeatable risk assessment method to establish model risk management maturity in a financial institution, the development of a method to measure the impact of model risk due to parameter uncertainty, and the proposal of a subjective scorecard to evaluate the relative model risk of a suite of models.

Key words:

Model risk, model error, financial risk models, model risk management, model risk mitigation, model risk quantification

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Financial risk models are simplifications of complex real-world phenomena used to better understand the intricate nature of an underlying process such as the loss generating process in financial risk analyses. Due to this simplification the model-representation of the real-world is by definition an approximation which implies a risk that the model may not fully reflect the real-world dynamics it is designed to mimic – this is referred to as model risk. This inherent model risk can probably not be eliminated completely since that would require an exact representation of the real-world which is arguably unattainable due to the complexity thereof. Model risk impacts a number of risk categories within financial risk, including operational risk, credit risk, strategic risk, reputational risk and more. Therefore, in this dissertation a contribution is made to the management, mitigation and measurement of model risk of financial risk models.

Models have become increasingly more complex due to an increase in computing power, financial-industry regulators requiring advanced risk models, and the sophistication of the implementation of models. The complexity also emanates from the desire to have optimal risk sensitivity when it comes to financial modelling. Risk sensitivity of a model allows for many benefits, but also comes with a number of potentially adverse consequences. In financial risk management a relatively small error in the model could lead to financial losses, non-optimal capital supply or reputational damage where stakeholders could lose confidence in the organisation due to model risk incidents such as accounting miss-statements or incorrect model-driven management decisions. Globally supervisors and regulators have been showing an increasing interest in model risk management and measurement practices, where some banks are even expected to hold a capital buffer for potential material model risk incidents. It is therefore a necessity that model risk management and measurement practices are leading-edge.

Chapter 2 is devoted to a review of the background of model risk. Specifically, the main types of models used in financial risk management are listed and discussed. The importance of model risk management is demonstrated listing some of the material model risk incidents that are publically known in the financial industry globally. Thereafter the evolution of the awareness of model risk is explored over the past three decades leading up to the present. The contribution of the industry, academia and regulators are demarcated where possible. This background discussion is concluded with speculation about future trends.

In Chapter 3 model risk is decomposed into a number of logical categories. This dissecting of model risk is proposed due to the apparent lack of a standardised definition of model risk across regulators, academia and industry. It is shown that the definition of model risk is not standardised and it can vary from narrow to broad. The divergence of definitions increases the difficulty of model risk management and measurement and comparability between financial institutions. The absence of a standardised definition of model risk creates an additional layer of uncertainty when it comes to model risk management and quantification, as the definitions available at this time range from narrow to very comprehensive. In the proposed research four model risk category types are proposed based on the available definitions of model risk. This enables a standardised approach of defining model risk, which could possibly eliminate the additional layer of uncertainty stemming from the lack of a standardised definition.

In Chapter 4 it is argued that because model risk cannot be eliminated entirely, it is important to have a comprehensive suite of model risk mitigating measures in place. Model risk mitigation includes a wide variety of practical ways to ensure that model risk is minimised effectively. For the purpose of this research an extensive, non-exhaustive list of mitigating techniques are grouped into four mitigation categories namely, governance, controls, testing, and monitoring and assurance.

In Chapter 5 it is argued that, because the measurement of model risk is not yet as mature in comparison with other major risk types, model risk management is a very important ingredient in mitigating model risk. A practical and repeatable model risk assessment method is proposed in this research. The method aligns with relevant regulatory guidance and observed leading practice. It provides a practical way to determine the current model risk management maturity level of a bank as well as determining a targeted state of where the level of model risk management maturity should be.

Measurement of model risk is explored in Chapter 6. The increasing reliance on the wide range of models within financial institutions as well as the diverging definitions of model risk calls for an enhanced focus on both model risk management and measurement. The magnitude of model errors leading to model risk depends amongst other on the complexity of the underlying reality that the models attempt to mimic. It is therefore necessary to measure the model error to appreciate the inherent model risk. The challenge with quantifying model risk is that relying on a model in order to accurately quantify model risk can result in an additional quantum of uncertainty. It can be argued that even though there is a risk involved in estimating model risk, it can be used for materiality-based model risk management to ensure that models with greater exposure are afforded more robust model risk mitigating measures. Some of the research available on model risk quantification is potentially too abstract to use in practice. Globally, some banks indicate that they are expected to quantify model risk for Pillar 2 capital purposes. The proposed research aims to measure model risk in a practical way which could possibly satisfy the expectations of regulators when it comes to the estimation of model risk capital.

In conclusion, the main contribution of this research is briefly summarised as follows:

i) demonstration that model risk is an important financial risk type that requires ongoing focus in the financial industry;
ii) proposal of a standardised definition of model risk through the categorisation of model risk types based on definitions of model risk available in literature;
iii) summarising model risk mitigating techniques based on the literature available;
iv) development of a practical and repeatable risk assessment method to establish model risk management maturity in a financial institution;
v) development of a method to measure the impact of model risk due to parameter uncertainty;
vi) proposal of a subjective scorecard to evaluate the relative model risk of a suite of models; and
vii) identification of a number of topics for future research that could potentially further improve model risk management, mitigation and measurement practices.

Extracts from this dissertation was published in van Biljon and Haasbroek (2017a) and presented at the 2016 Statistical Association of South Africa (SASA) (van Biljon and Panman, 2016) and 2017 SASA (van Biljon and Haasbroek, 2017b).

CHAPTER 2 BACKGROUND OF MODEL RISK

2.1 Introduction

Real-world data is often represented by statistical models, for example representing observed credit or operational loss data using a parametric distribution. These models are by definition simplifications of complex real-world phenomena since these models are used to describe the intricate nature of an underlying process, for example the loss generating process in financial risk analyses. Therefore model risk is knowingly created through this simplification. The magnitude of model errors leading to model risk depends amongst other on the complexity of the underlying truth (or reality) that the models attempt to mimic. Model errors can lead to financial losses, reputational damage and non-optimal decisions being made by management.

In this chapter different model types that can lead to model risk is discussed along with a history of model risk and why it is an important topic.

2.2 Model types

Real-world data is often represented by statistical models. The Fed and OCC (2011) define a model as *"a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates. A model consists of three components: an information input component, which delivers assumptions and data to the model; a processing component, which transforms inputs into estimates; and a reporting component, which translates the estimates into useful business information."* These models are by definition simplifications of complex real-world phenomena since these models are used to describe the intricate nature of an underlying process. Some model types used in financial risk management include:

i) expected loss models to determine provisions for operational and credit losses;
ii) unexpected loss models to determine regulatory and economic capital requirements for risk types such as credit, market and operational risk;
iii) valuation models for financial instruments such as financial derivatives which are dependent on observed market prices;
iv) credit application and behavioural decision-support models such as scorecards,
v) idiosyncratic and macro-economic stress testing models;
vi) models to identify fraudulent credit applications; and
vii) models to identify transactional fraud.

The benefit of using models include automated decision-making, which leads to improved efficiency, objective decision-making and the ability to synthesise complex issues in a financial environment (MANAGEMENT SOLUTIONS, 2014). These models are simplifications of the real-world phenomena they present and are therefore unfortunately imperfect. These imperfections lead to the presence of model errors which, could in turn lead to model risk if not properly measured and managed. The next section illustrates actual model risk events and their impact.

2.3 Why model risk matters

The Fed and OCC (2011) explain that the expanding use of models in all aspects of banking reflects the extent to which models can improve business decisions, but it should be noted that models also come with costs. These costs refer to not only the cost of resources to develop, implement and maintain models, but the potential costs of relying on models that are incorrect or misused. This section elaborates on why model risk should be recognised as an important risk in financial institutions by listing a non-exhaustive list of model risk incidents of which the detail is publicly known.

Gibson et al. (1999) mention some banks that have suffered extensive losses due to undue reliance on faulty models. These examples include: i) Merrill Lynch's pricing biases that lead to a loss of $70 million (Borodovsky and Lore, 2000), ii) JP Morgan's loss of $200 million in the mortgage-backed securities market in 1992 due to inadequate modelling of prepayments, and iii) NatWest Markets' mispricing on sterling interest rate options that cost them £90 million. In addition, De Jongh et al. (2017a) list examples of model risk incidents available in literature. These examples include: i) Morgan Stanley's housing CDO (Collateralised Debt Obligation) error in 2008 that cost $9 billion, (Springer, 2012), ii) JP Morgan's London Whale that ignored control warnings and changed how the Value at Risk (VaR) was measured which cost $6 billion in 2012 (Heineman, 2013), and iii) a spreadsheet error that lead to reputational damage to Reinhart & Rogoff in 2013, (BBC, 2013). Table 2-1: Examples of model risk incidents available in literature adapted from provides a summary of model risk incidents available in literature adapted from De Jongh et al. (2017a).

Table 2-1: Examples of model risk incidents available in literature adapted from De Jongh et al. (2017a)

Institution	Year	Description	Impact
Merryl Lynch (Borodovsky and Lore, 2000)	1970's	Pricing biases.	$70 million
JP Morgan Chase (Gibson et al., 1999)	1992	Inadequate modelling of prepayments in the mortgage-backed securities market.	$200 million
Fidelity Magel-Ian Fund	1995	Computing error due to a	$1.3 billion

(Godfrey, 1995)		spreadsheet blunder.	
NatWest Markets (Gibson et al., 1999)	1997	Mispricing on sterling interest rate options.	£90 million
Bank of Tokyo-Mitsubishi (Gibson et al., 1999)	1997	Inadequate pricing model on its US interest rate swaption book.	$83 million
Real Africa Durolink (West, 2004)	2001	Volatility skew.	R300 million
Fannie Mae (Wailgym, 2007)	2003	Spreadsheet error.	$1.2 billion
Morgan Stanley (Springer, 2012)	2008	Housing CDO error.	$9 billion
US Federal Reserve (IFOA, 2015)	2010	Spreadsheet error in the Consumer Credit calculations.	$4 billion
Welsh NHS (Anon, 2011)	2011	Spreadsheet calculation error for spending cuts.	Overstatement of £130 million
Mouchel Pension Fund (De Jongh et al., 2017a)	2011	Spreadsheet error for scheme valuation.	£8.6 million
Axa Rosenberg (SEC, 2011)	2011	Spreadsheet error overestimated client investment losses.	$242 million fine
JP Morgan (Heineman, 2013)	2012	So-called "London Whale", where control warnings were ignored and the way VaR is measured was changed.	$6 billion
ABSA (Barclays, 2013)	2012	Home Loans credit provisions model underestimated by R2 billion.	Increase in impairments of almost R300 million.
Reinhart & Rogoff (BBC, 2013)	2013	Spreadsheet error.	Reputational damage

In financial risk management model risk can lead to financial losses (as seen in the examples previously mentioned), non-optimal capital supply or reputational damage, where stakeholders could lose confidence in the organisation due to model risk incidents. Therefore it is necessary to estimate the model risk exposure as well as managing model risk.

Model risk management and measurement has evolved through the last two decades and the history thereof is discussed in the next section.

2.4 The history of model risk

In a presentation titled "*The Complete History of Model Risk- Abridged*" it is recognised that model risk has been around for a very long time (Hill, 2011). The author uses examples such as the incorrect construction of a warship, the Vasa in 1628, where an incorrect calculation method for estimating the amount of ballast required for stability lead to the ship sinking. The focus of this section is on the recent history of model risk within the financial industry starting from 1996. In this section the history of model risk is summarised using a timeline. The timeline is chosen and divided into three sections according to what makes sense using the research of the available literature. The choice of the three timeline periods is based on intuitive groupings of emerging trends from the three periods.

The content of the timeline is discussed following each section. The timeline distinguishes between supervisory guidance, industry initiatives and academic research in order to have a more comprehensive summary as seen in Table 2-2. The colours used refer to the three types of literature used. The timeline is then constructed as a non-exhaustive list of all the literature identified as relevant to model risk for each year included in the selected time period. The timeline can be seen in Table 2-3.

Table 2-2: Key to the model risk timeline

	Supervisory Guidance
	Academic Research
	Industry Initiative

Table 2-3: Summary of model risk literature published from 1996 to date (non-exhaustive)

Year	Different publications	Description of period
1996		Rise in financial derivatives and this period is mostly dominated by market risk models.
1997		
1998		
1999		
2000		Increased use of internal models across different risk types driven by Basel II.
2001		
2002		
2003		
2004		
2005		
2006		
2007		
2008		
2009		
2010		
2011		This period is mostly dominated by the reactions after the financial crisis of 2008 to 2010.
2012		
2013		
2014		
2015		
2016		
2017		

The timeline is divided into three periods. The first period is from 1996 to 1999 where the main themes relate to an increase in complexity of models and the rise of financial derivatives. The second period from 2000 to 2010 experienced an increase in the use of internal models and a heightened focus on model validation practices. The final period of interest is from 2011 to 2017 (present) where an intensified focus on model risk management and measurement is present. The three periods identified are discussed in the remainder of this section.

2.4.1 Period 1: 1996 to 1999

The main themes of this period are the increased complexity of models along with the earlier mentions of model risk. This period is mostly dominated by market risk models.

Table 2-4: 1996 to 1999 publications relating to model risk

	Supervisory Guidance	Academic Research	Industry Initiative
1996	BCBS, 1996	Derman, 1996	
		Jorion, 1996	
1999		Gibson *et al.*, 1999	

Increase in complexity of models

This period saw an increase in the complexity of models driven by the Basel Committee on Banking Supervision's (BCBS) introduction of the requirement to measure and apply capital charges in respect of market risk (BCBS, 1996). This directive coupled by the rise of financial derivatives and the increase in complexity of models lead to the increased need for model validation. The aforementioned directive highlights the need for internal validation processes for market risk such as the verification of the consistency, timeliness and reliability of data sources used to run the internal models, including the independence of such data sources.

Early mentions of model risk

This period also includes the earlier literature published acknowledging the existence of model risk. Jorion (1996) explains that VaR is affected by estimation risk and that the recognition of estimation errors can lead to better measurement methods. Derman (1996) acknowledges model risk and provides different sources of model risk that will be discussed in Chapter 3. Gibson et al. (1999) recognise that model risk is becoming increasingly important in financial valuation, risk management and capital adequacy. The authors further define model risk as *"model risk arises as a consequence of incorrect modelling, model identification or specification errors, and inadequate estimation procedures, as well as from the application of mathematical and statistical properties of financial models in imperfect financial markets"*. The authors also define three sources of uncertainty (see Chapter 3). It is also highlighted that *“model risk analysis should not be considered as a tool to find the perfect model, but rather as an instrument and/or methodology that helps to understand the weaknesses and to exploit the strengths of the alternatives at hand”*.

The increased complexity of models and the first mentions of model risk ties in with the themes of the next period where model validation guidance becomes prevalent, as well as the continued focus on model risk.

2.4.2 Period 2: 2000 to 2010

The main themes prevalent in this period are the regulatory guidance on model validation and academic research relating to identification, management and measurement of estimation errors and/or model risk.

Table 2-5: 2000 to 2010 publications relating to model risk

	Supervisory Guidance	Academic Research	Industry Initiative
2000	OCC		
2002		Berkowitz and O'Brien	
		Talay and Zhang	
2003	The Fed		
2005		Christoffersen and Gonçalves	
2006		Mignola and Ugoccioni	
2008		Sibbertsen *et al.*	
2009	BCBS		
2010		Kerkhof *et al.*	

Regulatory guidance on model validation

The first main theme to come out of the second period of interest is the first substantial supervisory guidance regarding model validation and the requirement for sound model validation processes (The Office of the Comptroller of the Currency (OCC, 2000) and The Board of Governors of the Federal Reserve System, (Fed, 2003)). This guidance highlights the supervisor’s requirement for model validation to evaluate model risk. The BCBS includes model risk as part of the supervisory expectation for valuation adjustments or reserves to be considered in paragraph 718(cix) of the revisions to the Basel II market risk framework (BCBS, 2009a). This highlights the concern surrounding model risk from the supervisors.

Academic research relating to model risk

The second theme that is prevalent in this period is that of academic research relating to identification, management and measurement of estimation errors and/or model risk. Most of the academic research focuses on market risk related research. Berkowitz and O'Brien (2002) examine the accuracy of VaR forecasts and one of the conclusions involves the finding that the 99th percentile VaR forecasts are conservative and in some cases inaccurate. Christoffersen and Gonçalves (2005) focus on market risk and assess the precision of dynamic models through the quantification of the magnitude of estimation error using confidence intervals for VaR and Expected Shortfall (ES). For operational risk, Mignola and Ugoccioni (2006) highlight the uncertainties in modelling operational risk losses, such as the absence of a dynamic model for operational risk losses and the high percentile requested for the measure of risk. Kerkhof et al. (2010) provide a quantitative basis for the incorporation of model risk in regulatory capital for trading activities in a market. Model risk is defined as *"the hazard of working with a potentially incorrect model"*. The authors also divide model risk into three types that will be discussed in Chapter 3. Sibbertsen et al. (2008) include the quantification of model risk and explain a Bayesian and a worst-case approach to measuring model risk. The authors further mention that the quantification approaches available in literature is too abstract to use in practice.

On the management of model risk Talay and Zhang (2002) describe a strategy which a trader can follow in order to manage model risk.

The increase in the supervisory guidance and academic literature regarding the management and measurement continues in the next period of interest, along with the addition of industry initiatives.

2.4.3 Period 3: 2011 to date

The most prominent theme of this period is the regulatory/supervisory reaction to the financial crisis (2008 to 2010) and the required remediation for the post-financial crisis era. The second theme identified is the response from industry to the supervisory guidance on model risk. The third theme of this period relates to the rise in academic research relating to model risk management and measurement.

Table 2-6: 2011 to date publications relating to model risk

	Supervisory Guidance	Academic Research	Industry Initiative
2011	The Fed and OCC		
	BCBS		
2012		Morini	North American CRO Council

		Alexander and Sarabia	
2013	The Fed		KPMG
	The European Parliament and the Council		PWC
	BCBS		Numerix
2014	EBA	Boucher *et al*.	Management Solutions
	The Fed	Glasserman and Xu	
2015	The Fed	Bertram *et al*.	Oliver Wyman
	EBA	Embrechts *et al*.	Institute and Faculty of Actuaries (IoFA)
	SARB		
	BCBS		
2016	EBA	Bignozzi and Tsanakas	Operational Riskdata eXchange Association (ORX)
	BCBS (a&b)	Mignola *et al*.	North American CRO Council
		Kellner *et al*.	
		Quell and Meyer	
2017		van Biljon and Haasbroek	
		De Jongh *et al*. (a&b)	

Regulatory response to the financial crisis

The final period of interest in the timeline sees a further increase in publications compared to the previous two periods. The most prominent theme of this period is the regulatory/supervisory reaction to the financial crisis and the required remediation for the post-financial crisis era. In the guidance paper titled *"Supervisory guidance on model risk management"* it is recognised that model risk is inherent in the use of risk models, but that model risk can be mitigated through robust model risk management (Fed and OCC, 2011). The purpose of the document is to provide comprehensive guidance for banks on model risk management. The importance of model validation is highlighted, as well as the importance of sound model development, implementation, use and governance. This is the basis on which a framework for managing model risk is proposed. The definition of model risk used is that it "*includes fundamental errors that may produce inaccurate outputs when viewed against the design objective and intended business uses, and the incorrect or inappropriate use of a model"*. The guidance also emphasises that model risk should be managed like other risk types in a way that they can identify the sources of risk and assess the magnitude.

Different regulatory/supervisory guidance followed after the publication by the Fed and OCC which confirms the concern of the regulators regarding model risk. The introduction of additional safeguards against model risk in response to the financial crisis, such as supplementing the risk-based measure with a simple measure of risk also confirms their concern (BCBS, 2011).

The Fed (2013) describes that an overall margin of conservatism should adequately account for all uncertainties and weaknesses in the capital quantification process. The supervisory expectations regarding a robust assessment of all uncertainty associated with risk-parameter estimates are also listed and includes amongst other, the identification of material instances of statistical uncertainty which confirms that the supervisors are concerned with the uncertainty linked to model risk.

Other supervisory and regulatory authorities also responded with guidance. The official definition of the European Banking Authority (EBA) aligns with the framework proposed by the Fed and OCC: *"model risk means the potential loss an institution may incur, as a consequence of decisions that could be principally based on the output of internal models, due to errors in the development, implementation or use of such models"* (EBA, 2013). One year later the EBA published its own guidance around model risk and stated that model risk can be assessed as part of the relevant risk type concerned when the risk relating to the underestimation of own funds requirements by regulatory approved models, and should be assessed as part of operational risk when the risk of losses relating to the development, implementation or improper use of any other models by the institution for decision-making is concerned (EBA, 2014). More guidance by the EBA was provided through a standard where the authority included a prescribed "Additional Valuation Adjustment" (AVA) for model risk whereby institutions should estimate an AVA by using a range of alternative valuations and estimate a point within the resulting range of valuations where they are 90% confident they could exit the valuation exposure at that price or better (EBA, 2015). The standard also mentions that the AVA can be estimated using expert judgment. EBA (2016) includes guidance to reduce the unjustified variability when risk parameters are estimated. It is further stated that the guidance is necessary in order to achieve improved comparability of risk parameters estimated and also highlights the need for "trust to be restored" in the models. The amount of guidance published by the EBA confirms the concern for model risk by the regulators.

The South African Reserve Bank (SARB) aligns with the BCBS' paragraph 718(cix) of the revisions to the Basel II market risk framework and defines model risk as "the risk associated with using a possibly incorrect valuation methodology, and the risk associated with using unobservable (and possibly incorrect) calibration parameters in the valuation model" (SARB, 2015).

The BCBS also confirms their increased concern for model risk with the paper titled *"The regulatory framework: balancing risk sensitivity, simplicity and comparability"* (BCBS, 2013a). The guidance highlights a number of shortcomings in the financial system's regulatory framework and in response thereof introduce a range of reforms designed to raise resilience of banks against shocks. It is emphasised that undue complexity in the pursuit of risk sensitivity may not always be awarded with high precision, but may increase model risk. The financial crisis is used as an example where the banks whose credit risk was under-estimated by the banks and rating agencies could be due to the quest for precision that lead to modelling errors. The BCBS proposes the use of a leverage ratio to serve as a supplementary measure to the risk-based capital framework which provides a floor to the outcome of risk-based capital requirements which provides protection against model risk. The definition used in the paper aligns with the guidance given by the Fed and OCC (2011), but includes more detail regarding specific causes of model risk: *"Model risk refers to the risk that the limitation of models may lead to material divergence between predicted and actual outcomes. This can be due to risk factors not considered in models, backward-looking nature of parameter estimation, potential underestimation of tail risk due to assumptions with respect to probability distributions, and residual uncertainty".*

The guidance on assessing exposure to model risk, awareness on model risk management and the validation of specifically Advanced Measurement Approach (AMA) models also emphasises the concern for model risk by the regulators (Fed, 2014). This guidance is model type specific and shows that supervisors have a heightened concern surrounding model risk for different models and risk types. They further elaborate on the importance of applying conservatism and benchmarking analysis in models' specifications and calibrations. Benchmarking is also stressed as a key tool to provide an important perspective for the model risk management process. The Fed (2015) highlights the need for independent validation and challenge of models specifically for Large Institution Supervision Coordinating Committee (LISCC) firms as well as large and complex firms. This reiterates their concern regarding model risk, as the guidance is aimed at a specific set of institutions and how they can manage model risk. The guidance is consistent with the existing guidance on model risk management published earlier by the Fed and OCC in 2000 and 2011.

Response from industry to the regulatory guidance

The second theme for this period is the response from industry to the supervisory guidance on model risk. The number of industry initiatives published shows that the industry is cognisant of model risk and that they take into account the guidance provided from the regulators.

The definition for model risk used by the NORTH AMERICAN CRO COUNCIL (2012) aligns with the definition provided by the Fed and OCC: "*the risk that a model is not providing accurate output, that a model is being used inappropriately, or that the implementation of an appropriate model is flawed*." The authors further emphasise the role of model validation as a key mitigant of model risk. This aligns with guidance on model validation as a mitigant from supervisors. Another industry initiative by the NORTH AMERICAN CRO COUNCIL (2016) elaborates on practices and principles relating to model risk management. These practices and principles covered the definition of models, model risk management and other guidance such as model inventories and model risk assessment.

Although there is guidance provided by the supervisors, the lack of a standard methodology for model risk measurement and management is evident from the diverging practices and definitions in the industry and academic literature. A survey by KPMG (2013) on model risk management practices confirms that model risk management practices and methodologies vary across financial institutions. Another survey conducted by the Operational Riskdata eXchange Association ORX (2016) highlights key findings including that ownership of model risk within banks is not as clear as for risk types such as credit, market and operational risk. The survey further reveals that quantification of model risk is not common and when it is performed it is mostly based on expert judgement. This aligns with the earlier survey conducted by KPMG (2013) around varying practices. One finding from the ORX survey is that the participants of the survey indicate that they expect model risk to become a greater priority for their regulators. This is an intriguing view seeing as the regulators are currently moving towards standardised methodologies for modelling. In a consultative document, the BCBS proposes a Standardised Measurement Approach (SMA) for operational risk due to their belief that the current AMA is inherently too complex and does not allow for comparability due to the wide range of modelling practices (BCBS, 2016b). In another consultative document the BCBS proposes a Standardised Approach (SA) for credit risk to be a suitable alternative and complement to the Internal Ratings-Based (IRB) approach (BCBS, 2015). The BCBS also revised the SA for the treatment of market risk, where the objective is to move away from sophisticated treatment for market risk (BCBS, 2016a). The intention of the BCBS for moving towards standardised modelling approaches is to find a balance between simplicity and risk sensitivity and to encourage comparability by reducing variability in estimates.

There are some practices relating to model risk management and measurement that converge. One of these principles is the use of the three-lines-of-defence framework for managing model risk (PWC, 2013). The proposed approach by PWC confirms that there exists the need from industry for a more standardised way of managing model risk. PWC proposes that the first line of defence consists of the model developers, owners and users and the second line involves the model risk management unit, which consists of model validation, annual model review, ongoing model risk monitoring and model risk remediation and mitigation. The governance and oversight consists of the senior management and the Board of Directors. The third line is Internal Audit that oversees the compliance of the other role players. The framework proposed also highlights operational risk management function as part of the model risk management group. De Jongh et al. (2017a) mentions the fourth line of defence which is suggested by the Financial Stability

Institute (FSI) as external audit and supervisors. Another proposed framework for managing model risk by van Biljon and Haasbroek (2017a) involves a practical and repeatable risk assessment method to establish model risk management maturity in an organisation and will be discussed in Chapter 5. NUMERIX (2013) focuses on the best practices for model risk management based on four types of model risk (see Chapter 3).

MANAGEMENT SOLUTIONS (2014) proposes a framework for managing model risk and quantifying the impact of model estimation uncertainty. The authors emphasise model validation, documentation and model inventories as an important part of model risk management. A model risk quantification cycle is also proposed. This involves the identification and classification of model risk sources, which confirms the need to categorise model risk into types as seen in Chapter 3. The Institute and Faculty of Actuaries IFOA (2015) proposes a model risk framework that includes concepts such as model risk appetite, model risk identification, model risk monitoring and the mitigation of model risk. Model risk incidents are also highlighted and the philosophical element of model risk is speculated on the where the quantification of model risk can lead to a second order of model risk or "model risk of model-risk models". A scorecard approach for governance, policy and the model validation process proposed by De Jongh et al. (2017a) aligns with the heightened focus and importance of validation.

Academic research

The increase in academic research on model risk management and measurement is the third theme identified for the period. The literature available is next discussed and some definitions of model risk will be highlighted.

Morini (2012) provides an overview of model risk and guidance on how to manage it. The definition for model risk is given as the *"possibility that a financial institution suffers losses due to mistakes in the development and application of valuation models".* Alexander and Sarabia (2012) propose a methodology for quantifying the model risk in quantile risk estimates. The authors further explain that the term model risk is *"commonly applied to encompass various sources of uncertainty in statistical models, including model choice and parameter uncertainty".*

Boucher et al. (2014) describe that the failures of risk models due to high levels of model risk leading to the under forecasting of risk prior to crisis events, the slow reaction of models as a crisis unfolds and the delayed reduction of risk levels post-crisis. This opinion confirms the reason for concern raised by regulators around model risk. The authors also highlight that there is no standard definition for model risk, but that it generally relates to the uncertainty created by the inability to know the true data generating process and that the uncertainty is made up of estimation error and the use of an incorrect model. The focus of the research is around adjusting risk forecasts (such as VaR) for model risk by their historical performance. Embrecths et al. (2015) confirm the concern for using VaR by showing that the uncertainty spread of VaR is larger than for ES. Kellner et al. (2016) also focus on the impact of model risk when determining VaR and ES, but added an Extreme Value Theory (EVT) element. The authors defined first-order effects of model risk, which consists of the misspecification and estimation risk and second-order effect of model risk which refer to the dispersion of risk measure estimates.

Bignozzi and Tsanakas (2016) discuss the uncertainty and referred to specifically model uncertainty and parameter uncertainty and defined the former as *"uncertainty arising from not knowing the model"* and the latter as *"arising from uncertainty about the true parameters, assuming that the model has been correctly chosen"*. Glasserman and Xu (2014) refer to model risk as errors in modelling assumptions that impact risk measurement, while the definition used by Bertram et al. (2015) is broader: "*every risk induced by the choice, specification and estimation of a statistical model"*. The authors also define three types of model risk which will be discussed in Chapter 3. Quell and Meyer (2016) use a broad classification of model risk (see Chapter 3).

Mignola et al. (2016) provide a more direct response to the regulators where the authors comment on the BCBS's proposal for a SMA for computing operational risk regulatory capital for banks. The findings of the authors include that the SMA is not risk sensitive and appears to be variable across banks. This shows that even when models are simplified and standardised, uncertainty and volatility is unavoidable.

From the increase in the publications and supervisory guidance illustrated in the timeline, it is clear that model risk is a topical subject in the financial industry. OLIVER WYMAN (2015) illustrates the increasing focus on model risk management from the 1990's through the financial crisis and afterwards. The authors highlight that the financial crisis took place before the maturation of model risk management, but it can be argued that given the review of literature in the post-financial crisis, that there is still room for improvement on the maturity of model risk management. Some of the reasons as to why model risk is topical are highlighted in Section 2.3.

Speculation about future model risk trends

Models are being used more frequently to depict complex real-world phenomena. There is also an increase in the sophistication of models, their implementation and the rise of machine learning. Financial industries are also experiencing a drive to have more automated models and model validation practices. Regulators seem cautious surrounding the complexity of models and have been calling for more simple and standardised models such as the standardised approaches proposed for credit risk, market risk and most recently operational risk.

The risk related to these models will not become much less prevalent in the near future and from the current trends and heightened scrutiny from regulators model risk will remain topical.

2.5 Conclusion

This chapter illustrates the way in which model risk has become more topical through the years due to the potential impact model risk can have on financial institutions, as well as the increased focus from regulators and supervisors. From the ORX (2016) survey is it clear that model risk will most likely remain topical and it is expected to receive even more attention from supervisors.

This chapter further illustrates the evolution of model risk in academic research as well as industry initiatives. The main themes identified for the first period (1996-1999) are the increased complexity of models along with the earlier mentions of model risk. The main themes for the second period (2000-2010) are the regulatory guidance on model validation and academic research relating to identification, management and measurement of estimation errors and/or model risk. The first theme identified for the final period (2011 to date) is the regulatory/supervisory reaction to the financial crisis and the required remediation for the post-financial crisis era. The second theme relates to the response from industry to the supervisory guidance on model risk. The third theme of this period relates to the rise in academic research relating to model risk management and measurement.

The evolution of model risk confirms that financial institutions as well as academia are positioning themselves to enable thorough model risk management and substantial research has been conducted surrounding model risk measurement. Two key gaps can be identified in the evolution of model risk. Firstly, most of the available research on model risk measurement is potentially too abstract to use in practice. Therefore it is argued that there is a definite need for simpler approaches to measure the risk associated with model error without creating an extra layer of model risk. The second gap is that there is no standard definition for model risk. This creates an additional layer of uncertainty when it comes to managing model risk and the expectations of supervisors surrounding quantification, and makes comparison of model risk management and quantification practices between industry participants challenging.

In the next chapter model risk is categorised according to the definitions available in literature.

CHAPTER 3 MODEL RISK CATEGORIES

3.1 Introduction

Model risk categories are the categorisation of model risk types based on the different definitions available for model risk. The reason for categorising model risk emanates from the lack of a standardised definition for model risk highlighted in Chapter 2. van Biljon and Haasbroek (2017a) show that the definitions for model risk range from narrow to very comprehensive where the authors define model risk as *"all sources of uncertainty related to the model choice, parameter choice, model application and interpretation of model results"*. This definition is aligned with that provided by the Fed and OCC (2011) and Morini (2012), but is broader than the definition used by Alexander and Sarabia (2012), Glasserman and Xu (2014), Bignozzi and Tsanakas (2016) and Bertram et al. (2015) since they only consider model risk as due to model choice and parameter misspecification. The absence of a standard definition creates an additional layer of uncertainty when it comes to model risk management and quantification. The benefits of using model risk categories include the following:

i) a standardised approach of defining model risk;
ii) less uncertainty on how to classify model risk;
iii) more clarity on how to manage model risk based on the category type;
iv) more clarity on how to mitigate model risk based on the category type;
v) the enablement of comparing model risk management, mitigation and measurement across financial institutions due to the standard categorisation; and
vi) allows for benchmarking due to standardised definition or categorisation.

The benefits of categorising model risk into different types are appreciated to some extent in current literature. Some sources do not explicitly refer to the categorisation, but through their definitions of model risk the different category types can be deduced. Gibson et al. (1999) include uncertainty about the estimates of the model parameters, given the model structure as a source of uncertainty, as well as the uncertainty about the model structure. The authors further include uncertainty about the application of a model in a specific situation, given the model structure and the parameter estimation as another source of uncertainty. For this example three types of model risk can be deduced from the definition given by the authors. Another example is where the authors mention parameter uncertainty and model choice as sources of uncertainty in statistical models (Alexander and Sarabia, 2012). This example also illustrates the authors implicitly categorising model risk into types through the definition. Other sources explicitly categorise model risk into different types. NUMERIX (2013) defines best practices for model risk based on four sources of model risk. MANAGEMENT SOLUTIONS (2014) also lists sources of model risk, which can be interpreted as model risk categories, which include data deficiencies and model misuse.

Following from the need to categorise model risk, the following model risk categories are defined from using available definitions and own experience:

i) Type 1: Model parameter uncertainty;
ii) Type 2: Model misspecification;
iii) Type 3: Change in the dynamics of real-world phenomena; and
iv) Type 4: Incorrect model implementation, misinterpretation of model output; and other errors.

The four category types mentioned above are defined and discussed with examples in the following sections.

3.2 Type 1: Model parameter uncertainty

Model parameter uncertainty arises when an optimal model with no change in the underlying dynamics of the real-world phenomenon has non-optimal parameters. Gibson et al. (1999) include uncertainty about the estimates of the model parameters, given the model structure as part of the sources of uncertainty. Alexander and Sarabia (2012) also mention parameter uncertainty as part of the uncertainty in statistical models. Boucher et al. (2014) explain that parameter estimation error arises from uncertainty in the parameter values of the chosen model. Kellner et al. (2016) define the second-order effect of model risk as the dispersion of risk measure estimates. Bignozzi and Tsanakas (2016) define parameter uncertainty as "*arising from uncertainty about the true parameters, assuming that the model has been correctly chosen".*

Some examples of this model risk type include:

i) when an obvious input variable is not considered in a credit scorecard which then translates to a missing parameter from the model;
ii) if a loss generating process has a certain distribution such as a lognormal distribution, but with non-optimal parameters that specify the lognormal distribution (see for example the Loss Distribution Approach (LDA) in operational risk);
iii) when derivative pricing is dependent on a parameter that is unobservable, which leads to the use of proxies and assumptions being made; and
iv) if the discount rates used for recovery cash flows to calculate the loss given default (LGD) in credit risk is not certain.

3.3 Type 2: Model misspecification

Model misspecification arises when a non-optimal model is selected to represent the current underlying dynamics of a real-world phenomenon. Derman (1996) lists an incorrect model as one of the sources of model risk. Gibson et al. (1999) acknowledge the uncertainty about the model structure as a source of uncertainty. Kerkhof et al. (2010) include misspecification risk as one of the model risk types. Alexander and Sarabia (2012) mention model choice as another source of uncertainty in statistical models. Bertram et al. (2015) include model risk in distribution and model risk in functional form as part of the types of model risk. Bignozzi and Tsanakas (2016) define model uncertainty as *"uncertainty arising from not knowing the model"*. Kellner et al. (2016) define first-order effects of model risk, which consists of misspecification and estimation risk.

Some examples of this model risk type include:

i) when a non-optimal model is used for a credit scorecard, for example using linear regression when logistic regression is the optimal choice;
ii) when a non-optimal distribution is used instead of the optimal choice such as using a Burr instead of the lognormal distribution for a LDA in operational risk;
iii) when the volatility is assumed as a deterministic process instead of a more appropriate stochastic process in financial derivatives;
iv) if the correlation between assets in a multi-asset valuation model for financial derivatives is ignored; and
v) when the correlation of the probability of default (PDs) and LGDs in credit risk are not taken into account.

3.4 Type 3: Change in the dynamics of real-world phenomena

A change in the dynamics of the real-world phenomenon occurs when the model was initially optimally specified with optimal parameters, but due to a recent change in the dynamics of the underlying real-world phenomenon, the model is no longer suitable. Derman (1996) lists the correct model but an incorrect solution as part of the sources of uncertainty. In this context this would mean the model is correct, but due to the change in the underlying dynamic of the real-world phenomenon, the solution the model provides is no longer correct. Kerkhof et al. (2010) include identification risk as part of the model risk types. This relates to the inability to select an econometric model when different models describe the same data. Boucher et al. (2014) explain that model risk generally relates to the inability to know the true data generating process. Glasserman and Xu (2014) include errors in modelling assumptions that impact risk measurement in their definition for model risk. In the context of model risk category Type 3, the underlying phenomenon is assumed to remain unchanged when in reality the dynamics evolved from the dynamics observed when the initial model was developed.

Some examples of this model risk type include:

i) when the model initially had a high Gini coefficient, but deteriorated due to a change in the underlying dynamics of the scored population;
ii) when the loss generating process in operational risk initially is an optimal distribution with the correct parameters, but then changes due to internal or external factors which leads to the model not being suitable anymore;
iii) when the underlying market changes the assumption of constant volatility, see for example the volatility "smile" (West, 2004); and
iv) when the modelling assumption that interest rates will always be positive were violated due to some markets experiencing negative interest rates after the financial crisis in 2008/9.

3.5 Type 4: Incorrect model implementation, misinterpretation of model output, and other errors

This model risk type occurs when an optimally developed model is implemented, used or interpreted incorrectly which leads to inaccurate results or non-optimal conclusions. The correct model, but inappropriate use is a source of model risk (Derman, 1996). Gibson et al. (1999) explain that the uncertainty about the application of a model in a specific situation, given the model structure and the parameter estimation is one of the sources of uncertainty. NUMERIX (2013) defines best practices for model risk based on four sources of model risk of which three fit into the Type 4 category namely bad data, bad implementation and bad usage. The sources of model risk listed by MANAGEMENT SOLUTIONS (2014) align to the classification by (NUMERIX, 2013), which include data deficiencies and model misuse. Quell and Meyer (2016) list the inaccurate implementation of an otherwise fit for purpose model design in their definition of model risk. van Biljon and Haasbroek (2017a) include model application and interpretation of model results as part of the uncertainty relating to model risk.

Some examples of this model risk type include:

i) when a fit-for-purpose model design is implemented incorrectly resulting in the implemented model providing incorrect model results;
ii) when a fit-for-purpose model is used incorrectly due to the user providing invalid model input data resulting in invalid model output; and
iii) when a fit-for-purpose model is used incorrectly due to the user misinterpreting the model output.

3.6 Conclusion

In this chapter model risk is categorised into four category types namely: i) Type 1: model parameter uncertainty, ii) Type 2: model misspecification, iii) Type 3: change in the dynamics of real-world phenomena, and iv) Type 4: incorrect model implementation, misinterpretation of model output, and other errors. The need for the categorisation is due to the apparent lack of a standardised definition for model risk. These classifications are based on the available definitions from supervisory guidance, industry initiatives and academic research. The categorisation assists with the classification and analysis of model risk.

Current literature confirms the need for categorisation of model risk, even though it is sometimes only implicitly categorised in the model risk definitions. The model risk categorisation assists with identifying mitigation measures in order to address the specific model risk category type. In the next chapter model risk mitigation methods are discussed.

CHAPTER 4 MODEL RISK MITIGATION

4.1 Introduction

It is recognised that model risk is unlikely to be entirely eliminated largely because models are simplified representations of complex real-world phenomena. Therefore models are unfortunately imperfect, which implies the presence of model uncertainty and can result in model risk. From the categorisation of model risk types discussed in Chapter 3, it is clear that there exists a broad range of sources of model risk and uncertainty. An extensive suite of model risk mitigating measures can be used in order to aid with the management of model risk for the different model risk category types. For the scope of this dissertation the following non-exhaustive model risk mitigating measures are considered through using research and practical experience:

i) data quality tests;
ii) development and validation standards;
iii) formal approval process;
iv) materiality-based governance;
v) change control;
vi) measured conservatism;
vii) technical specialists;
viii) training and awareness;
ix) model monitoring;
x) model validation;
xi) model audit;
xii) managing model limitations;
xiii) benchmarks;
xiv) sensitivity tests;
xv) stress tests;
xvi) backtesting;
xvii) exposure limit management;
xviii) model inventories; and
xix) ongoing research.

The remainder of this section is dedicated to explaining the aforementioned model risk mitigating measures, discuss their impact and identify the model risk type they affect.

4.2 Model risk mitigating measures

Each model risk mitigating measure is discussed next including the impact of using the measure.

4.2.1 Data quality tests

Data is a key input for models and therefore if the quality of the data is not up to standard the model results can be negatively affected. The credibility of the modelling approach depends on the relevance, integrity, and internal consistency of the underlying data (Fed, 2014). BCBS (1996) emphasises the need for the verification of the consistency, timeliness and reliability of data sources used to run internal models. Fed and OCC (2011) stress that data is of "*critical importance*" to the development of a model. It is further highlighted that the developer of the model should be able to demonstrate the suitability of the data and methodology used in the development of the model. The critical importance of data quality during the development process translates to a lot of time spent on data by model development. Data quality and completeness tests should be performed and improved over time according to the principle set out in *"Principles for effective risk data aggregation and risk reporting"* (BCBS, 2013b).

The impact of this mitigating measure results in improved data quality and completeness, which reduces the risk of model errors due to using incorrect or incomplete data inputs during model development as well during the ongoing use of a model.

4.2.2 Development and validation standards

The minimum standards to which model development and validation functions need to adhere should be formalised in model development and validation policies, respectively. Regulatory and supervisory guidance is available for guidance on the expectations surrounding the minimum standards of model development and validation. The Fed and OCC (2011) state that the design, theory and logic of the model developed should be documented and supported by research and industry. It is further stated that model development should ensure that the model performs as intended by testing the robustness and stability of the model. This process should also be documented. Model validation minimum standards can include for example the requirements that all model components should be subject to validation, that the validation should be independent from development, that the rigour of the validation should align with the exposure of the model, and that validation should occur when material changes are made to existing models (Fed and OCC, 2011).

The impact of formalising minimum standards for model development and validation ensures that model development and validation is performed with the appropriate rigour to meet the supervisory and regulatory requirements. It also improves consistency and standardisation across different modelling focus areas.

4.2.3 Formal approval process

Formal approval through a defined governance process is required before new or amended models are put in use. BCBS (1996) mentions the importance of the approval process for risk pricing models and valuation systems. The Fed and OCC (2011) details the importance of good governance, policies and control over a model risk governance framework.

According to the Fed and OCC (2011) a typical governance hierarchy includes the board of directors and senior management through the establishment of a bank-wide approach to model risk management that includes a model risk governance framework which fits into the rest of the organisation's risk management as well as suitable policies for all components of model risk such as model development, model implementation, model validation and model audit.

Formal approval ensures that all relevant stakeholders (including regulators where applicable) are aware and satisfied that the new or amended model is appropriate before being put in use.

4.2.4 Materiality-based governance

The governance process should be designed to take the materiality of models into account. High-materiality models should be afforded extra scrutiny to further reduce model risk. Extra scrutiny can involve more detailed validation practices and additional levels of approval required before model output is used.

The extra scrutiny that is afforded to high materiality models ensures that the model risk of these models is not overlooked.

4.2.5 Change control

A formal change-control process should be in place for making changes to in-use models or implementing new models into production systems. A change-control process involves approval of model changes, documenting the changes and ensuring that no changes are made by unauthorised individuals.

Controlled model changes ensure that the model owner and users explicitly approve the model changes, and that the model behaves as intended once implemented into the target production system.

4.2.6 Measured conservatism

In dealing with unavailable or uncertain model inputs, assumptions, and/or model methods, a measured degree of conservatism is incorporated into the model to compensate for the model uncertainty. Models should typically contain a degree of conservatism. For example, diversification benefits can be ignored and extreme percentiles of the loss distribution are used to represent unexpected losses.

De Jongh et al. (2017b) mention that Basel II acknowledges that the estimates of parameters such as PDs, LGDs and EADs are likely to involve unpredictable errors and to avoid "over-optimism", banks must add a margin of conservatism to their estimates. The paper discusses the detail surrounding the conservatism requirements.

The BCBS applies a scaling factor of 1.06 to the risk weighted assets for credit risk under the Internal Ratings Based (IRB) approach (BCBS, 2006). This factor can be seen as a factor to account for model risk. The Fed (2014) also confirms the importance of applying conservatism in models.

Models containing measured conservatism compensate for additional uncertainties and model risk.

4.2.7 Technical specialists

A pool of skilled, experienced and technically competent staff should be maintained in the development, validation and audit functions. The training and experience of developers can affect the extent of model risk faced in the development of a model (Fed and OCC, 2011).

Using experienced technical specialists in model development, validation and audit functions improves the likelihood of delivering appropriate models, thorough model validation and robust audit.

4.2.8 Training and awareness

Appropriate awareness of model changes with stakeholders should occur. The appropriate awareness of model changes can be facilitated through formal governance meetings where models are approved.

Training on the correct use of models by users must take place to assist with avoiding the incorrect use of a model.

4.2.9 Model monitoring

The performance of in-use models should regularly be tested against appropriate criteria to ensure the ongoing fitness for use. The minimum requirements for model monitoring can include stability tests used in model development and validation, sensitivity tests to ensure robustness of the model, and industry benchmarks to ensure that model inputs and results are relevant.

Monitoring of models identifies potential model performance deterioration or operational errors, for example data input errors or unexpected changes in the underlying dynamics the model was originally designed to replicate.

4.2.10 Model validation

Model validation should be performed independently of the model development. Independence is obtained through separate reporting lines (i.e. different reporting lines for model development and model validation), separate infrastructure and restrictions to access of code and compensation linked to the quality of the validation performed and not the performance of the model.

Model validation provides additional confirmation that the model is fit-for-purpose and performs as intended. A number of supervisory guidance published refers to the importance of model validation. See for example BCBS (1996), OCC (2000), the Fed (2003), the Fed and OCC (2011), the Fed (2014) and the Fed (2015).

4.2.11 Model audit

Model audit should be performed on model development and model validation functions. Model audit is the third line that oversees the other role players in the model risk three lines-of-defence (PWC, 2013). De Jongh et al. (2017a) explain that internal audit should verify that no models go to production without formal approval and is also responsible for the oversight of model validation and its policies.

Model audit ensures the adherence to minimum standards defined in relevant model development and validation policies.

4.2.12 Managing model limitations

Model limitations highlighted during model monitoring or validation processes should be recorded and addressed. The Fed (2013) includes a comprehensive listing of other sources of uncertainty in the parameter quantification processes, including gaps and limitations in reference data or in underlying processes that potentially impact the accuracy of the risk-parameter estimates and an assessment of materiality of each limitation or gap, as part of the expectations of risk parameter estimates.

Managing model limitations can lead to defining appropriate model development actions and tracking the progress thereof to ensure timely remediation of identified model limitations.

4.2.13 Benchmarks

Model results should be verified by comparing it to available benchmarks to indicate whether the model design and results compare favourably to the results of equivalent models and with industry expectations. Examples of typical benchmark data used for different risk types include: for operational risk, ORX, for credit risk CDG and Mark IT supplies a benchmark data for market risk. The Fed (2014) stresses benchmarking as a key tool to provide an important perspective for the model risk management process.

Using benchmarks can give management some comfort that the model results are in line with industry expectations.

4.2.14 Sensitivity tests

Sensitivity tests involve making moderate changes to model inputs and parameters to confirm the model stability and the intuitive correct model behaviour in respect of changes to the drivers of model results.

Sensitivity tests give an important indication of how stable the models are.

4.2.15 Stress tests

Stress testing is required through the Basel II capital adequacy framework (BCBS, 2006). The BCBS suggests that stress testing is a valuable tool for risk management as it provides, amongst other, a forward-looking assessment of risk and facilitates risk mitigation. Under Pillar 2 of the Basel II framework supervisors examine bank’s stress testing results in order to determine whether the capital and liquidity is adequate.

Macro-economic and idiosyncratic stress tests should be performed periodically on key risk exposures. Macro-economic stresses evaluate the effects of macro-economic factors on the behaviour of key risk exposures. Idiosyncratic stresses assess the sources of vulnerabilities and indicates what risk exposure can be impacted.

Stress testing provides important information regarding the forward-looking assessment of risk and how models could possibly react in stressed conditions.

4.2.16 Backtesting

Backtesting is when model results are compared to subsequent realised events. Some examples of backtesting include actual losses versus expected losses for operational risk and modelled PDs versus actual PDs in credit risk.

Backtesting confirms that the model adequately captures the underlying phenomena it was designed to mimic.

4.2.17 Exposure limit management

Risk exposure limits can be maintained for concentrations in respect of factors such as country, region, industry sector, collateral type and counterparty, which results in the maximum risk exposure being capped.

4.2.18 Model inventories

Model inventories should be maintained that include information pertinent to models including model type, purpose, model owner and review dates. MANAGEMENT SOLUTIONS (2014) lists model inventories as an important model risk mitigant.

Maintaining model inventories assist with improving corporate memory related to in-use models and with preventing the overlooking of models during validation and audit cycles.

4.2.19 Ongoing research

Technical specialists can periodically attend conferences, take part in industry initiatives, and assess recent academic research on relevant modelling and validation techniques.

Research allows staff to keep abreast of emerging best practices, new techniques, and potential changes in the dynamics of the phenomena for which models are in place.

4.3 Conclusion

Table 4-1 summarises the model risk mitigating methods along with the model risk category type which is mitigated through each method.

Table 4-1: Model risk mitigating methods

Mitigating Measure	Model Risk Category Mitigated			
	Type 1	Type 2	Type 3	Type 4

Data quality tests	x	x		x
Development and validation standards	x	x	x	x
Formal approval process	x	x	x	x
Materiality-based governance	x	x		
Change control	x	x	x	x
Measured conservatism	x	x		
Technical specialists	x	x	x	x
Training and awareness				x
Model monitoring	x	x	x	x
Model validation	x	x	x	x
Model audit	x	x	x	x
Managing model limitations	x	x	x	x
Benchmarks	x	x	x	x
Sensitivity tests	x	x		
Stress testing	x	x	x	
Backtesting	x	x	x	
Exposure limit management	x	x	x	
Model inventories				x
Ongoing research	x	x	x	x

While it is recognised that model risk may never fully be eliminated, the mitigating measures explained in this chapter serves as an important tool for the management of model risk. These mitigation measures cover the governance, controls, testing, monitoring and assurance surrounding models. Furthermore, as illustrated in Table 4-2, these mitigation measures are suitable to mitigate model risk arising from all four of the model risk categories defined in Chapter 3.

The effectiveness of model risk management, including the use of the mitigation tools listed here, contributes to how well model risk is managed. Therefore a practical model risk management maturity assessment method is proposed in the next chapter.

CHAPTER 5 MODEL RISK MANAGEMENT

5.1 Introduction

Models are increasing in both complexity and in their range of applications due to the rise in complexity of banking and banking products, more advanced computing power, increased reporting and regulatory requirements and the value-added to decision making. Overreliance on models that are not fit-for-purpose can result in a bank being uncompetitive due to mispricing or holding a non-optimal capital supply.

As mentioned previously, models are by definition simplified representations of intricate real-world phenomena. This simplification leads to the presence of model risk. In contrast to other major risk types, where risk-quantification methods have evolved substantially, the study of model risk quantification is still emerging. While measurement of a risk type is an important factor in managing risk, and while model risk quantification methods are still maturing, model risk still needs to be managed effectively. The increased focus of regulators and supervisory guidance, as well as the absence of standardised model risk quantification methods further justify model risk management as an important element to consider for a bank's risk management practices.

In this chapter model risk management is dissected into a five-step process based on emerging best practices as well as experience in practice. The five-step process is then used to provide a practical solution on how to assess the maturity level of model risk management within a bank.

The work presented here was also published in the Journal of Risk Model Validation, titled *"A practical maturity assessment method for model risk management in banks"* (van Biljon and Haasbroek, 2017a) and presented at the 2017 SASA (van Biljon and Haasbroek, 2017b).

5.2 The five steps of model risk management

Model risk management involves different role players such as:

- i) business users of models;
- ii) model development teams;
- iii) model implementation and maintenance teams;
- iv) model validation teams;
- v) model auditors; and
- vi) operational risk managers.

Other important role players include the banking regulator and the bank's Board of Directors. When looking at the listed players in the context of the commonly used three-lines-of-defence structure, model development, business users and model implementation teams are the first line of defence, the second line of defence is model validation and then model audit serves as the third line of defence. In a different context operational risk management serves as the second line of defence with respect to all of the other role players, with the exception of model audit which plays an oversight role to all the players. Lastly, the regulator assesses the bank's model risk management through tools such as self-assessments, with the bank's Board of Directors ultimately being responsible for model risk management. Based on the above operating structure, regulatory guidance in BCBS (2009b), BCBS (2013a), Fed and OCC (2011) and EBA (2014) as well as experience in practice, model risk management is dissected into a five-step process. The five steps:

i) data quality, extraction and transformation;
ii) model definition, development and documentation;
iii) model validation and approval;
iv) model implementation, change control and usage; and
v) reporting and monitoring, is discussed below.

5.2.1 Data quality, extraction and transformation

Data is a key input for models and is therefore a key driver of model results. In a typical financial institution setting, data used in models is from a variety of different sources. It requires collation from these sources via automated data feeds to the model calculation engine, and reconciliation to reliable sources to test accuracy and completeness. In most cases, some form of data transformation is required before the data can be used as a model input. The robustness with which these data-related steps are performed determines how accurate the model-data input is. It also governs the agility with which the model results can be updated, e.g., during periods of market stress, which may require model results to be updated more frequently than during typical market periods, or by performing investigations to unpack model results, which may require frequent reproduction of model results using different data sets as input. The following practices are required to ensure the robustness of these data-related processes:

i) frequent data validation to ensure ongoing accuracy and completeness of the data;
ii) automation of data feeds and data transformation to limit manual intervention that could introduce operational errors and hamper the nimbleness of model updates;
iii) availability of up-to-date complete data dictionaries that describe key attributes of data elements as well as data sources to allow unambiguous interpretation of the required data;

iv) data transformations applied to data that are clearly documented and included in the model-approval process; and
v) changes to data sources and feeds being done through a robust change-control process to ensure adequate testing and awareness by stakeholders of pending changes.

Typical early warning signs of suboptimal data-related processes are the following:

i) open and overdue audit findings relating to data processes;
ii) data validation that occurs infrequently, and issues identified during these validations not being addressed adequately or not on schedule;
iii) frequent manual workarounds being needed to address data feed, quality or completeness issues;
iv) data dictionaries being outdated, incomplete or not in place;
v) data-transformation procedures not being documented sufficiently;
vi) changes to data sources and feeds occurring without a robust process in place to inform other downstream data users; and
vii) the ability to increase the frequency of model updates being limited.

5.2.2 Model definition, development and documentation

Model development activities primarily involve the development of new models and the enhancement of existing models. The reasons for model development are mainly changing business user needs or changing regulations, the availability of richer input data, improved modelling methods or the discovery of flaws in existing models. Model development requires technically skilled and experienced individuals that have a good understanding of relevant regulation and the business area for which the model is being developed, where applicable. Core competencies include the ability to:

i) distil the model definition from regulation or interaction with business users;
ii) translate the model definition into a computer coding language;
iii) define data input requirements and assist with any additional work needed to deliver new data inputs to the model;
iv) embed the model in a system; and
v) train users and design reports of model results

For continuity and transparency purposes, it is important that the design and logic behind the model, including any model limitations, are clearly and comprehensively documented. The robustness with which these activities are performed is usually determined by the materiality of the model, using materiality measures such as the perceived complexity of the model, income statement impact in the case of, for example, a valuation model, or balance sheet impact in the case of a capital model. The practices that are found to aid model development vigour include the following:

i) the minimum standards that define how model development is undertaken are well documented, reviewed and approved regularly;
ii) the model development team is adequately staffed with appropriately skilled and experienced staff, who keep abreast of emerging model development methods in the industry;
iii) project management principles, including project plans, timelines and dedicated project managers, are applied to model development activities to ensure measurable progress, drive accountability and assist with prioritisation, should conflicting modelling requests be received;
iv) business engagements that lead to key model decisions are clearly documented to assist with ensuring the delivered model is aligned with the user's requirements;
v) the model documentation allows for a competent person to replicate the model independently, which then demonstrates that the documentation is sufficiently comprehensive; and
vi) an up-to-date model inventory is maintained that lists all in-use models along with key dates, such as implementation date, validation date and status, and next validation date, to prevent in-use models from being overlooked during future model monitoring, validations or audits.

Typical early warning signs of suboptimal model development practices are the following:

i) model documentation is not in place, incorrect, incomplete or not approved by the appropriate governance committees;
ii) the model development team has skill and capacity shortages for dealing with the current workload, which may jeopardise the development rigor;
iii) there is limited evidence of business engagements during the model design phase, which could result in the final model not aligning to business requirements;
iv) model inventories are outdated or incomplete; and
v) the model validation's recommendations or audit findings are not tracked or attended to within the agreed time lines, resulting in in-use models remaining non-optimal for longer than initially anticipated.

5.2.3 Model validation and approval

The main objective of model validation is to confirm that "models are performing as expected, in line with their design objectives and business uses" (Fed and OCC, 2011). Model validation operates independently from model development to ensure objectivity. A particular challenge of validation is that the time available to complete validation work is typically limited due to pressures from business users to implement the model; therefore, this team should have sufficiently skilled and experienced analysts, together with a wide range of validation approaches, to speedily conclude on the soundness of a model. Validation approaches can include benchmarking, backtesting and independently rebuilding the model. This wide range of approaches is essential when vendor models are validated due to the often limited disclosures by vendors of their proprietary models. Once validation confirms the appropriateness of a new or amended model, it requires approval before it is put to use; this approval process is driven by the model development team, with assistance from the business users. Occasionally, due to time-conscious considerations such as a time-to-market for a new product valuation model, or correcting a newly discovered flaw of an in-use model, model changes can be made without first completing the governance process. However, these cases are typically rare: the governance process is still completed, albeit after the event, and, at the time these extraordinary model changes are being made, heightened scrutiny is applied. Should the model change result in a financial statement impact, measured conservatism with respect to any profit recognition or capital impact is applied.

Effective model validation practices include the following:

i) the model validation processes and minimum standards are documented in a policy that is regularly reviewed and approved at appropriate governance committees;
ii) the model validation team has appropriately skilled and experienced staff to deliver credible validation outcomes within a constrained time period, which is typically shorter than the time used by model development to develop the model;
iii) model validation operates independently from model development;
iv) the validation scope is sufficiently wide to cover not only model design but also ancillary aspects, such as model data inputs and outputs; and
v) pre-implementation approval of new or amended models takes place at governance committees, with representation from relevant stakeholders.

Signs of suboptimal model validation practices are:

i) model validation documentation is inadequate and/or not approved at appropriate governance committees;

ii) the model validation team has skill and capacity shortages that may negatively affect the thoroughness of the validation effort;

iii) the scope of the validation is too limited, for example it excludes input data by relying on data validation performed outside the validation team, which may not be up to the standard required for model use;

iv) recurring validations are sometimes not performed according to the agreed validation schedule, thereby overlooking the deterioration of a previously fit-for-purpose model; and

v) vendor models are not afforded the same validation rigor as internally developed models.

5.2.4 Model implementation, change control and usage

Model implementation refers to the embedding of the developed model into a system environment where it can be employed by the model users. Properly implemented models enable the automated updating of model input data, automated model calculations and result reports. Often the developed model needs to be translated into the target system's programming language, with additional development work to automate data input fields and result reports. The embedding of models into such system environments is needed to reduce time-consuming manual interventions; limit operational errors; enable access to computing clusters to speed up calculations; keep up with reporting demands; and free up the time of developers or users, who would otherwise produce model results manually, which is non-optimal. It is crucial that testing of the embedded model is thorough in order to ensure it is performing as intended. To this end, the developers are best positioned to provide appropriate testing scenarios. The system environment wherein the model is embedded should be treated as a production-status system environment, with robust change-control procedures being followed for any system amendments. To ensure appropriate use of the model, user training on the model and its limitations is mandatory, together with logical access controls to prevent unauthorised use.

Thorough model implementation, change control and usage consist of the following:

i) new and amended models are implemented only after formal approval to ensure adequate testing before implementation;

ii) model implementation testing is performed and documented to assist in assessing whether the model is fit-for-purpose according to the model user's requirements;

iii) models are implemented in a production-status environment with proper change-control and logical-access management to avoid any unintended effects caused by incorrect implementation and amendments by unauthorised users;
iv) users of the models are appropriately trained to use the model correctly to ensure that models are used as intended;
v) in order to limit the subjectivity of the outcomes, minimal overrides are applied to model results;
vi) a formal process is in place to notify the model development team of changes to product design and input data changes that may affect the in-use model.

Early warning signs of inefficient model implementation, change control and usage include the following:

i) new and amended models are implemented before formal approval has taken place;
ii) model implementation errors are often discovered after implementation, indicating inadequate testing and test scenarios during implementation;
iii) models are not implemented in a production-status environment, forcing manual operation of the model and reducing the ability to apply robust change-control and logical-access management;
iv) model users are not relevantly trained, which may result in the incorrect use of models;
v) user documentation is limited;
vi) overrides are applied regularly to model results, indicating that users are not confident in the model results; and
vii) changes are made to the design or data of in-use products without a formal process in place to notify key players, such as the model development team, beforehand, which may result in previously fit-for-purpose models becoming inadequate.

5.2.5 Reporting and monitoring

Reporting and monitoring refers to the disclosing of model results to end-users and the ongoing verification of the soundness of the model. Reports influence decision making and should therefore contain all key results presented in such a way that they are understandable to nontechnical users, with clear commentary where applicable. Model monitoring is typically performed by model development teams using a variety of tests to confirm that the model continues to perform in line with expectations. A high degree of automation of result reports and model monitoring is possible due to the routine nature of these activities. Automation is also preferred in order to:

i) reduce the time spend by technical resources to produce reports manually (an activity that is typically disliked by these resources);
ii) improve the expandability of the reporting framework so that adding more models does not result in a commensurate increase in effort; and
iii) improve the ability to increase the frequency of these activities during stress periods or for testing and result investigations.

The practices identified to ensure transparent reporting and effective monitoring are the following:

i) minimum requirements for model monitoring and reporting, including the frequency thereof, are defined and clearly documented;
ii) the production of reports is automated for efficiency and to prevent manual errors;
iii) the frequency of reporting can be increased under stress conditions; and
iv) there is demonstrable evidence that monitoring results are used to amend in-use models where needed.

Typical early warning signs of suboptimal reporting and monitoring include the following:

i) minimum monitoring and reporting requirements are not defined;
ii) the monitoring requirements are elementary in nature and do not take into account the different stakeholders;
iii) monitoring reports are manually produced with many errors and delays; and
iv) there is limited evidence that monitoring results are used to amend in-use models.

5.3 Model risk maturity assessment method

In order to improve model risk management practices, a bank can use a repeatable method to determine the current maturity of its model risk management practices. This method is described in this section.

5.3.1 Method

The proposed method to assess the maturity of model risk management practices within a bank uses the five steps of model risk management as described in the previous section:

A. Data quality, extraction and transformation
B. Model definition, development and documentation
C. Model validation and approval
D. Model implementation, change control and usage

E. Reporting and monitoring

In addition these steps are rated according to a five-point rating scale as follows:

1) major improvement required, model risk management not fit-for-purpose;
2) some aspects of model risk management require major improvement;
3) adequate model risk management, some aspects require improvement;
4) good model risk management with minor room for improvement; and
5) leading-edge model risk management.

The above is then used to formulate a two-dimensional matrix with rows from the process steps A, B,…, E and columns from the ratings 1, 2,…, 5, defined above. For every combination of process step and rating, a set of assessment statements was formulated. Collectively, this is referred to as the "model risk management maturity matrix" depicted in Table 5-1. The specific assessment statements can be seen below in Table 5-2 to Table 5-6.

The five-point rating scale is designed to allow sufficient granularity in the rating of each of the model risk management process steps. It is equivalent to the risk-rating scale applied in risk and control self-assessments within operational risk. Namely,

i) a 1-rating is catastrophic risk exposure and translates into the model risk management not being fit-for-purpose and requiring major improvement;
ii) a 2-rating is a high-risk exposure and implies that some aspects of model risk management require major improvement;
iii) a 3-rating is a tolerable risk exposure and indicates that model risk management is adequate, with only some aspects requiring improvement;
iv) a 4-rating is a low-risk exposure and implies good model risk management is in place, with minor room for further improvement;
v) a 5-rating is a very low-risk exposure equivalent to leading-edge model risk management practices being observed.

Table 5-1: Model risk management maturity matrix

	Rating Scale				
Model risk management: **Process steps**	1. Major improvement required, model risk management not fit for purpose	2. Some aspects of model risk management requires major improvement	3. Adequate model risk management, some aspects require improvement	4. Good model risk management, with minor room for improvement	5. Leading-edge model risk management
A. Data quality, extraction and transformation	*A.1*	*A.2*	*A.3*	*A.4*	*A.5*
B. Model definition,	*B.1*	*B.2*	*B.3*	*B.4*	*B.5*

development and documentation					
C. Model validation and approval	*C.1*	*C.2*	*C.3*	*C.4*	*C.5*
D. Model implementation, change control and usage	*D.1*	*D.2*	*D.3*	*D.4*	*D.5*
E. Reporting and monitoring	*E.1*	*E.2*	*E.3*	*E.4*	*E.5*

Table 5-2: Assessment statements related to process step A. Data quality, extraction and transformation

A.1	**A.2**	**A.3**	**A.4**	**A.5**
a. Input data has **frequent** accuracy and completeness problems which require **frequent** manual intervention and workarounds.	a. Input data has **frequent** accuracy and completeness problems which require **frequent** manual intervention and workarounds.	a. Input data has **intermitted** accuracy and completeness problems which require **frequent** manual intervention and workarounds.	a. Input data **is** accurate and complete with **some** manual workarounds required to ensure accuracy and completeness.	a. Input data **is** accurate and complete with **no** manual workarounds needed.
b. Intermitted data errors such as data feed problems or missing data are **often** not discovered.	**b.** Intermitted data errors such as data feed problems or missing data are **sometimes** not discovered, but those that are discovered are corrected afterwards which may result in **frequent** restatement of previously published model results.	**b.** Intermitted data errors such as data feed problems or missing data are **always** discovered but not necessarily timeously, and are corrected afterwards which may result in **infrequent** restatement of previously published model results.	**b.** Intermitted data errors such as data feed problems or missing data are **timeously** resolved and the impact thereof **mitigated** with appropriate modeling techniques.	**b.** There are **no** intermitted data errors such as data feed problems or missing data.
c. Data validation is **not** performed.	**c.** Data validation is **not** performed.	**c.** Data validation is done on an ***ad hoc*** basis.	**c. Frequent** data validation is in place.	**c. Frequent** data validation is in place.
d. Model validation in respect of model input data is **unsatisfactory**.	**d.** Model validation in respect of model input data is **unsatisfactory** or **satisfactory** with **limitations**, and has a number of high-priority validation **recommendations**.	**d.** Model validation in respect of model input data is **satisfactory** with room for **improvement**.	**d.** Model validation in respect of model input data is **satisfactory** with room for **improvement**.	**d.** Model validation in respect of model input data is **satisfactory**.
e. Actions to improve data quality are **overdue**, or have **not** been **defined**.	**e.** Actions to improve data quality have been **defined**, but most actions are **overdue**.	**e.** Actions to improve data quality have been **defined**, with only some actions being **overdue**.	**e.** Actions to improve data quality have been **defined**, and are **on schedule** to be completed within the agreed timeframe.	**e.** There is **no** room for **improvement** regarding data quality.
f. Data dictionaries of relevant source data used in models are **not** in place.	**f.** Data dictionaries of relevant source data used in models are **not** in place.	**f.** Data dictionaries of relevant source data used in models **are** in place, but are **incomplete** and/or the	**f.** Data dictionaries **are** in place but **not** necessarily up to date.	**f.** Data dictionaries **are** in place, complete, and up to date.

g. Data manipulations of model input data during extraction, transformation and loading (ETL) is **not** documented. **h.** There are **unsatisfactory** and/or **overdue** audit findings related to data.	**g.** Data manipulations of model input data during extraction, transformation and loading (ETL) is **not** documented. **h.** There are **overdue** audit findings related to data.	content is **not** necessarily up to date. **g.** Data manipulations of model input data during extraction, transformation and loading (ETL) is **not** documented. **h.** There are **overdue** audit findings related to data.	**g.** Data manipulations of model input data during extraction, transformation and loading (ETL) is **partially** documented. **h.** There are **no** overdue audit findings related to data.	**g.** Data manipulations of model input data during extraction, transformation and loading (ETL) **is** documented and **approved**. **h.** There are **no** open audit findings related to data.

Table 5-3: Assessment statements related to process step B. Model definition, development and documentation

B.1	**B.2**	**B.3**	**B.4**	**B.5**
a. Minimum standards that define how model development should be performed are **not** mature, documented or approved at the appropriate committee. **b.** The model design **should** be improved upon. **c.** There is **no** evidence that the model development team keeps abreast of emerging model development methods in the industry. **d.** Model validation is generally **unsatisfactory** with a large number of high priority model validation recommendations. **e.** Model documentation is **mostly** identified by model validation as **unsatisfactory**. **f.** Model inventories are	**a.** Minimum standards that define how model development should be performed have **evolved**, but are **not** documented or approved at the appropriate committee. **b.** The model design **could** be improved upon. **c.** There is **no** evidence that the model development team keeps abreast of emerging model development methods in the industry. **d.** Model validation is generally **unsatisfactory** with a large number of high priority model validation recommendations. **e.** Model documentation is **often** identified by model validation as **unsatisfactory** requiring **improvements**. **f.** Model inventories are **not** in place.	**a.** Minimum standards that define how model development should be performed **are** documented in a model development policy, but are **not** necessarily annually approved at the appropriate committee. **b.** The model design **is** fit for purpose. **c.** There is **limited** evidence that the model development team keeps abreast of emerging model development methods in the industry. **d.** Model validation is generally **satisfactory** with **limitations** with room for **improvement** identified. **e.** Model documentation is **often** identified by model validation as requiring **minor improvements**. **f.** Model inventories **are** in place, but are	**a.** Minimum standards that define how model development should be performed **are** documented in a model development policy that **is** annually approved at the appropriate committee. **b.** The model design **is** fit for purpose. **c.** There is **adequate** evidence that the model development team keeps abreast of emerging model development methods in the industry. **d.** Model validation is generally **satisfactory** with **limitations** with room for **improvement** identified for **some** models. **e.** Model documentation **is** complete, however **requires** minimal interaction between validation and development to replicate the model. **f.** Model inventories **are** complete and	**a.** Minimum standards that define how model development should be performed **are** documented in a model development policy that is annually approved at the appropriate committee. **b.** The model design equals or exceeds known **leading edge**. **c.** There is **demonstrable** evidence that the model development team keeps abreast of emerging model development methods in the industry. **d.** Model validation is **satisfactory** with **no** room for **improvement** identified. **e.** Model documentation **is** complete and of a high standard to allow an apparent competent person to replicate the model independently. **f.** Model inventories **are** complete and up

not in place.		**not** necessarily up to date and/or complete.	up to date.	to date.
g. No evidence of business engagements during the key phases of the model development process is in place.	**g. No** evidence of business engagements during the key phases of the model development process is in place.	**g. Limited** evidence of business engagements during the key phases of the model development process is in place.	**g.** Evidence of business engagements during the key phases of the model development process **is** in place, but is **not** necessarily documented in the model development document.	**g.** Business engagements during the model development process **are** documented in the model development document.
h. The model development team **has** skill and capacity shortages to deal current workload.	**h.** The model development team **has** skill and capacity shortages to deal current workload.	**h.** The model development team **has** appropriately skilled staff, but **lacks** experience and/or requires more human resources to deal current workload.	**h.** The model development team **has** appropriately skilled and experienced staff.	**h.** The model development team **is** adequately staffed with appropriately skilled and experienced staff.
i. Project management principles – including a project plan, timelines and a project manager – are **not** applied to material model development activities.	**i.** Project management principles – including a project plan, timelines and a project manager – are **generally not** applied to material model development activities.	**i.** Project management principles – including a project plan, timelines and a project manager – are **mostly** applied to material model development activities.	**i.** Project management principles – including a project plan, timelines and a project manager – **are** applied to material model development activities.	**i.** Project management principles – including a project plan, timelines and a project manager – **are** applied to model development activities.
j. There are **several** overdue model-related audit findings.	**j.** There are **some** overdue model-related audit findings.	**j.** There are **minimal** overdue model-related audit findings.	**j.** There are **no** overdue model-related audit findings, but there may be **some** open audit findings that are **on** schedule.	**j.** There are **no** open or overdue model-related audit findings.
k. No log is maintained by model development of model validation recommendations.	**k. No** log is maintained by model development of model validation recommendations.	**k.** A log **is** maintained by model development of model validation recommendations, but is **not** updated.	**k.** A log **is** maintained by model development of model validation recommendations.	**k.** A log **is** maintained by model development of model validation recommendations.
l. The risk and control self-assessment (RCSA) of the model development team is **not** in place.	**l.** The RCSA of the model development team is **not** in place.	**l.** The RCSA of the model development team **is** in place, but may **not** be entirely complete.	**l.** The RCSA of the model development team **is** in place.	**l.** The RCSA of the model development team **is** complete and up to date.
m. Model risk incidents are **not** captured onto the internal loss database with root cause analysis **not** being performed.	**m.** Model risk incidents are **not** captured onto the internal loss database with root cause analysis **not** being performed.	**m.** Model risk incidents are **not** captured onto the internal loss database with root cause analysis **not** being performed.	**m.** Model risk incidents **are** captured onto the internal loss database **with** root cause analysis being performed.	**m.** Model risk incidents **are** captured onto the internal loss database **with** root cause analysis.

Table 5-4: Assessment statements related to process step C. Model validation and approval

C.1	C.2	C.3	C.4	C.5
a. Minimum standards that define how model validation should be performed are **not** mature, documented or approved at the appropriate committee.	**a.** Minimum standards that define how model validation should be performed have **evolved**, but are **not** documented or approved at the appropriate committee.	**a.** Minimum standards that define how model validation should be performed **are** documented in a model validation policy, but are **not** necessarily annually approved at the appropriate committee.	**a.** Minimum standards that define how model validation should be performed **are** documented in a model validation policy that **is** annually approved at the appropriate committee.	**a.** Minimum standards that define how model validation should be performed **are** documented in a model validation policy that **is** annually approved at the appropriate committee.
b. There is **no** evidence that the model validation team keeps abreast of emerging model validation methods in the industry.	**b.** There is **no** evidence that the model validation team keeps abreast of emerging model validation methods in the industry.	**b.** There is **limited** evidence that the model validation team keeps abreast of emerging model validation methods in the industry.	**b.** There is **adequate** evidence that the model validation team keeps abreast of emerging model validation methods in the industry.	**b.** There is **demonstrable** evidence that the model validation team keeps abreast of emerging model validation methods in the industry.
c. Model validation of a new or amended model does **not** occur.	**c.** Model validation is typically only performed **once** at inception of the model, with **no** annual review or ongoing validation.	**c.** Model validation **is** performed according to the agreed frequency, with some validations being performed **late** but with appropriate notification to relevant parties.	**c.** Model validation **is** performed according to the agreed frequency.	**c.** Model validation **is** performed according to the agreed frequency.
d. Models are **not** tabled at the appropriate technical committee and/or model approval committee for approval before being put into use.	**d.** Models **are** tabled at the appropriate technical committee and/or model approval committee for approval, sometimes **after** the model was put into use.	**d.** Models **are** tabled **timeously** at the appropriate technical committee and/or model approval committee for approval.	**d.** Models **are** tabled **timeously** at the appropriate technical committee and model approval committee for approval.	**d.** Models **are** tabled **timeously** at the appropriate technical committee and model approval committee for approval.
e. Model validation documentation is **limited** or **not available**.	**e.** Model validation documentation requires **improvement**.	**e.** Model validation documentation requires **minor improvement**.	**e.** Model validation documentation **is** complete and up to standard.	**e.** Model validation documentation **is** complete and of a high standard.
f. The model validation team **is** under resourced to deal current workload.	**f.** The model validation team is **generally** under resourced to deal current workload.	**f.** The model validation team **has** appropriately skilled and experienced staff, but requires **more** human resources to deal current workload.	**f.** The model validation team **has** appropriately skilled and experienced staff.	**f.** The model validation team is **adequately** staffed with appropriately skilled and experienced staff.
g. The risk and control self-assessment (RCSA) of the model validation team is **not** in place.	**g.** The RCSA of the model validation team is **not** in place.	**g.** The RCSA of the model validation team **is** in place, but may **not** be complete.	**g.** The RCSA of the model validation team **is** in place.	**g.** The RCSA of the model validation team **is** complete and up to date.
h. Model risk assessments	**h.** Model risk assessments	**h.** Model risk assessments	**h.** Model risk assessments	**h.** Model risk assessments

performed by model validation to determine the risk-based materiality of the model for validation scope purposes, are **not** performed.	performed by model validation to determine the risk-based materiality of the model for validation scope purposes, are **not** performed.	performed by model validation to determine the risk-based materiality of the model for validation scope purposes, are **not always** performed.	performed by model validation to determine the risk-based materiality of the model for validation scope purposes, are **always** performed.	performed by model validation to determine the risk-based materiality of the model for validation scope purposes, are **always** performed.
i. Model validation scope is **limited** and therefore may **exclude** material aspects that impacts model risk.	**i.** Model validation scope is **limited** and therefore may **exclude** some material aspects that impacts model risk.	**i. Some** material aspects that impacts model risk, e.g. data is excluded from validations.	**i. Minimal** material aspects that impacts model risk, e.g. data is **excluded** from validations.	**i. No** material aspects that impacts model risk, e.g. data is excluded from the validation.
j. There are **several** overdue model validation-related audit findings.	**j.** There are **some** overdue model validation-related audit findings.	**j.** There are **minimal** overdue model validation-related audit findings.	**j.** There are **no** overdue model validation-related audit findings, but there may be some open model-related audit findings that are on schedule.	**j.** There are **no** open or overdue model validation-related audit findings.
k. Most model development actions in respect of model validation recommendations are **overdue**.	**k. Some** model development actions in respect of previous model validation recommendations are overdue.	**k. Minimal** model development actions in respect of previous model validation recommendations are overdue.	**k. All** model development actions in respect of previous model validation recommendations are on schedule.	**k.** There are **no** outstanding model development actions in respect of previous model validation recommendations.

Table 5-5: Assessment statements related to process step D. Model implementation, change control and usage

D.1	**D.2**	**D.3**	**D.4**	**D.5**
a. New models and/or amended models are implemented **before** formal approval is obtained.	**a.** New models are implemented **only** once formal approval is obtained.	**a.** New models are implemented **only** once formal approval is obtained.	**a.** New models are implemented **only** once formal approval is obtained.	**a.** New models are implemented **only** once formal approval is obtained.
b. Model implementation testing requires **significant** improvement.	**b.** Changes to in-use models are made **before** formal approval is obtained.	**b.** Changes to in-use models are **only** made once formal approval is obtained.	**b.** Changes to in-use models are **only** effected once formal approval is obtained.	**b.** Changes to in-use models are **only** made once formal approval is obtained.
c. Models are implemented in **non-production** status environment such as Excel or SAS with **little** or **no** change control and access controls present.	**c.** Model-implementation testing **is** performed, but requires **improvement** and documentation due to frequent discovery of errors only after the model is implemented.	**c.** Model implementation testing **is** performed, but is **not** documented to evidence the testing.	**c.** Model implementation testing **is** performed, but the documentation thereof to evidence the testing requires **improvement.**	**c.** Model implementation testing **is** performed and documented.
d. Users of	**d.** Models are	**d.** Models are	**d.** Models are	**d.** Models are

models are **not** trained to use the model correctly.	implemented in **non-production** status environment such as Excel or SAS, and **some** control deficiencies in respect of change control and access controls are present.	implemented in **non-production** status environments such as Excel or SAS, but **adequate** controls are in place to mimic a production status environment with proper change control and access controls.	implemented in a **production** status environment with **proper** change control and access controls.	implemented in a **production** status environment with **proper** change control and access controls.
e. User documentation and/or a training log to evidence training are **not** in place.	**e.** Users of models are **not necessarily** trained to use the model correctly.	**e.** Users of models are **appropriately** trained to use the model correctly.	**e.** Users of models are **appropriately** trained to use the model correctly.	**e.** Users of models are **appropriately** trained to use the model correctly.
f. A **large number** of overrides are applied to model results.	**f.** User documentation and/or a training log to evidence training are **not** in place.	**f.** User documentation and/or a training log **are** in place, but are **stale**.	**f.** User documentation and a training log **are** maintained, but may be **stale**.	**f. Detailed** user documentation and an up to date training log is maintained.
g. No process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model.	**g.** A **large number** of overrides are applied to model results.	**g. Some** overrides are applied to model results.	**g. Some** overrides are applied to model results.	**g. No** overrides are applied to model results.
h. There are **several** overdue audit findings related to model implementation or usage.	**h. No** process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model.	**h. No** process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model.	**h.** An **imperfect** process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model.	**h.** A **formal** process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model.
	i. There are **several** overdue audit findings related to model implementation or usage.	**i.** There are **some** overdue audit findings related to model implementation or usage.	**i.** There are **no** overdue audit findings related to model implementation or usage, but there may be **open** audit findings of which the remediation is **on schedule**.	**i.** There are **no** open audit findings related to model implementation or usage.

Table 5-6: Assessment statements related to process step E. Reporting and monitoring

E.1	**E.2**	**E.3**	**E.4**	**E.5**
a. The minimum requirements for model monitoring reporting and frequency are **not** defined.	**a.** The minimum requirements for model monitoring reporting and frequency are **not** defined.	**a.** The minimum requirements for model monitoring reporting and frequency are **not** defined.	**a.** The minimum requirements for model monitoring reporting and frequency **are** defined, but **not**	**a.** The minimum requirements for model monitoring reporting and frequency **are** defined and documented.

			formally documented.	
b. No attempt was made to investigate leading practice in respect of model monitoring.	**b.** The **evolved** monitoring reporting is **not** aligned with leading practice, or **no** attempt was made to investigate leading practice.	**b.** The **evolved** monitoring reporting is **not** aligned with leading practice, or **no** attempt was made to investigate leading practice.	**b.** The minimum monitoring requirements **equals leading practice**.	**b.** The minimum monitoring requirements **equals** or **exceeds leading practice**.
c. No monitoring reporting is currently performed.	**c.** Monitoring reports and/or frequency have been identified by business, model development, and audit or model validation as requiring **significant** improvement.	**c.** Monitoring reports and/or frequency have been identified by business, model development, audit or model validation as requiring **some** improvement.	**c.** Monitoring reports and frequency **equals** the agreed minimum defined requirements.	**c.** Monitoring reports and frequency **equals** or **exceeds** the agreed minimum defined requirements.
	d. The production of monitoring reports is **manually** performed.	**d.** The production of monitoring reports is **manually** performed.	**d.** The production of monitoring reports is **semi-automated** to limit manual errors.	**d.** The production of monitoring reports is **automated** to prevent manual errors.
	e. Monitoring reports contains **limited** commentary.	**e.** Monitoring reports **contains** commentary.	**e.** Monitoring reports contains **adequate** commentary to explain model outcomes to a non-technical audience.	**e.** Monitoring reports contains **adequate** commentary to explain model outcomes to a non-technical audience.
	f. No ability exists to increase the frequency of reporting under stress conditions.	**f. Limited** or **no** ability exists to increase the frequency of reporting under stress conditions.	**f. Limited** ability exists to increase the frequency of reporting under stress conditions.	**f.** The ability **exists** to increase the frequency of reporting under stress conditions.
	g. Monitoring results are **infrequently** tabled at the appropriate committees for noting.	**g.** Monitoring results **are** tabled at the appropriate committees for noting, but sometimes the reports are **late** or **stale**.	**g.** Monitoring results **are** tabled **timeously** at the appropriate committees for noting.	**g.** Monitoring results **are** tabled **timeously** at the appropriate committees for noting.
	h. There is **no** demonstrable evidence that monitoring results are used to amend in-use models if needed.	**h.** There is **no** demonstrable evidence that monitoring results are used to amend in-use models if needed.	**h.** There is **demonstrable** evidence that monitoring results are used to amend in-use models if needed.	**h.** There is **demonstrable** evidence that monitoring results are used to amend in-use models if needed.

The specific assessment statements, in the tables above, for each risk rating were formulated from observed practices to align with the defined rating scale. In this context, ratings worse than a 3-rating indicate model risk management practices are not fit-for-purpose and are broadly not compliant with minimum regulatory requirements. A 3-rating indicates compliance with minimum regulatory requirements, and higher ratings exceed minimum regulatory requirements, with a 5-rating indicating leading-edge practices.

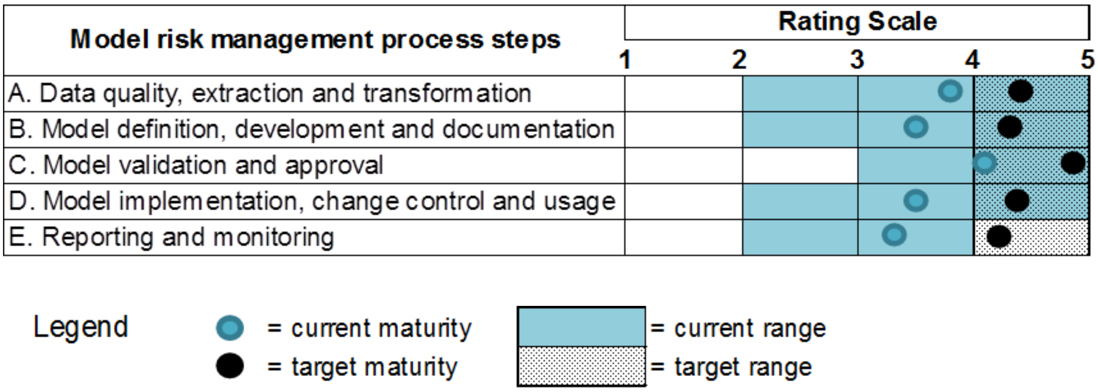
5.3.2 Application

The model risk management maturity matrix can be used in a number of ways. The main application thereof is to determine the current state of model risk management practices. It is recommended that the matrix be applied by key model development areas due to model development teams being arguably the best-positioned to have a holistic view of the model risk management process related to the model types they specialise in. Here the current maturity level of key model types are assessed by critically reviewing compliance with each of the assessment statements related to every model risk management process step listed in Table 5-1. For each process step, the rating-scale column from where most of the assessment statements align with current practices indicates the rating to apply. The combined rating per process step for all assessed model types is then the average for the bank across model types, and represents the maturity level. In the event that the materiality of models is deemed important to consider, a weighted average linked to the materiality of models can be applied.

An additional application of the method is to define a targeted state of model risk management maturity. Once the current state is known, the list of assessment statements (refer to Table 5-1 to Table 5-6) would guide the user on what specific aspects to address to obtain a higher rating. The user then also considers the constraints to improve the maturity rating from the current state. These constraints usually include the allocation of budget, information technology dependencies and data quality and availability limitations, amongst other. Since the assessment statements detail what needs to improve to obtain the targeted maturity rating, it assists with defining the project plan needed to improve current practices to reach the targeted state.

A fictional example of the application of the proposed method is shown in Figure 5-1 below. In this example a number of key model development areas were tasked with assessing the model risk management maturity of the model types they specialise in. The "current range" shows the dispersion of the individual ratings whilst the "current maturity" is the average rating assessed across all model types. Similarly the "target range" is the distribution of individual model-type target ratings and the "target maturity" is the average target rating. Note that in this example the "target maturity" was chosen to be slightly higher than the "current maturity" across all model risk management process steps which indicates the ambition to improve model risk management practices across model types. Therefore, supporting this assessment would typically be a project plan detailing the work required to progress from the current maturity state to the targeted maturity state.

Figure 5-1: An example of results from a model risk management maturity assessment



For further explanation of the results in Figure 5-1, consider process step A: “Data quality, extraction and transformation”. Here the “current maturity” is 3.8 which is the average rating assessed for individual model types with individual ratings ranging from 2 to 5. The average “target maturity” is higher at 4.4 with the corresponding individual target ratings ranging from 4 to 5. Therefore, in this example the model risk management for some model types intentionally do not target the leading-edge rating of 5. This may be due to considering constraints such as budget allocation which prevents achieving the leading-edge rating over the medium term. Furthermore, because the “current range” includes ratings below a rating of 3 (which is calibrated to coincide with the minimum regulatory expectations) some model types require remediation work to meet regulatory expectations. For this work the detailed assessment statements from the model risk management maturity matrix can be used to formulate remediation plans to advance from the noncompliant model types’ current rating to the respective target ratings. In all cases the “target range” exceeds minimum regulatory expectations which imply that the model risk management of all model types is scheduled to exceed regulatory expectations over the medium term.

5.3.3 Limitations

An awareness of the limitations of the maturity assessment method is key in helping to improve the relevance of results. Due to the recommended application approach detailed above resembling a self-assessment, it is likely that a degree of overconfidence and subjectivity will be present in the assessment results, since self-assessments typically display a degree of overestimation of actual ability, performance, level of control or chance of success. This bias can be limited with controls such as:

- (i) involving an objective central party to coordinate the assessment in order to ensure that the necessary oversight takes place, for example, operational risk management;
- (ii) requiring the assessor to provide sufficient evidence to substantiate ratings; and
- (iii) verifying that the results are consistent with other sources, such as existing model validation recommendations and audit findings.

Another limitation is that, while comparing the current practice to the assessment statements per rating column, it may so happen that the assessment items that are comparable to practice are from different rating columns. In these cases, the rating that corresponds to the majority of valid assessment statements is the final rating.

A further limitation is that the assessment items include words that can possibly be interpreted as vague such as “frequent”, “adequate”, “limited”, etc. To avoid ambiguity it is important that clear interpretations are given to each of the descriptors by the organisation before the assessment commences so that the outcome can be validated against clear guidelines set by the organisation completing the assessment.

5.4 Conclusion

Model risk cannot be entirely eliminated and it is therefore argued that model risk management is a key element in a bank’s risk management practices. The proposed approach to manage model risk is to use a practical and repeatable method to assess the maturity of the bank’s model risk management. This approach is of particular relevance, because it can be argued that the sophistication of model risk management lags behind that of other major risk types such as credit, market and operational risk. This method decomposes model risk management into five process steps as well as a five-point rating scale. For every combination of process step and maturity rating, a set of hands-on assessment statements were formulated to align with emerging best practice and regulatory guidance.

The method provides a view on the current maturity level of model risk and can also be used to define a targeted maturity level. The method provides a practical and relevant approach to advance the management and mitigation of model risk.

In the more mature risk types, such as credit risk and operational risk, a key ingredient to risk management is measurement. The next chapter explores model risk measurement.

CHAPTER 6 MODEL RISK MEASUREMENT

6.1 Introduction

The old adage from management studies that "*you cannot manage what you cannot measure*", suggests that measurement should be a key component of model risk management. Main risk types such as credit and market risk have already matured in their measurement practices. The challenge with model risk measurement is that there is not yet an industry or regulatory standard such as the case in more mature risk types. It can be argued that even though there is an absence of a standard approach to measure model risk, it is still an important facet of model risk management. At this stage it seems like it is not possible to have a generic solution to model risk measurement for all model types. The quantification of model risk should likely be specifically customised for each specific model.

Another challenge with quantifying model risk is that relying on a model in order to accurately quantify model risk can result in an additional quantum of uncertainty. In the paper published by the IFOA (2015), it is recognised that quantifying model risk can lead to a second order of model risk, namely the model risk of model-risk models. It can be argued that even though there is a risk involved in estimating model risk, it can be used for materiality-based model risk management to ensure that models with greater exposure are afforded more robust model risk mitigating measures.

This chapter is dedicated to model risk quantification of the different model risk category types discussed in Chapter 3 as well the proposal of a qualitative scorecard assessment approach.

Extracts of the work presented here were also presented at SASA 2016 (van Biljon and Panman, 2016).

6.2 Quantifying Type 1 model risk

As defined in Chapter 3, Type 1 model risk relates to the uncertainty of the model parameters. In this section the magnitude of model errors relating to parameter uncertainty is estimated. The process followed can be generalised for different types of models, but for the scope of this work models based on parametric distributions are used. The focus of the method is to establish how severe model risk relating to Type 1 can be, given certain aspects such as the type of model, as well as the sample size used.

In this section a method to quantify Type 1 model risk for a parametric distribution is explained and this can be used to demonstrate how severe (or not) model risk of Type 1 can be.

6.2.1 Quantifying Type 1 model risk of a parametric distribution

Parametric distributions are often used to represent loss data in the financial modelling world. The purpose of this section is to estimate the model error, due to parameter uncertainty, of a parametric distribution that represents synthetic loss data. Different parametric distributions are chosen to show the impact of different model types. The sample size is also varied in order to establish what type of influence it has on model risk. The results of this quantification method can be used to inform the risk relating to certain models and how certain aspects impact the severity of Type 1 model risk.

Methodology

The purpose of this study is to estimate the model error of a parametric distribution that represents synthetic loss data.

The proposed approach involves defining a "true" model representation of reality $F_i(x;\theta)$, where F_i is some cumulative distribution function and θ its parameters where i represents a different parametric distribution and its parameters. The assumption of a "true" distribution can seem unrealistic, but for the purpose of this study it is used to estimate the relative differences in order to determine the range of relative model errors. An example of a distribution function would be $Lognorm(\mu,\sigma)$. The parameters of $F_i(x;\theta)$, $\hat{\theta}$ is then estimated using different samples from the assumed "true" population X. The results are obtained using a simulation study.

Simulation design

The aim of the simulation study is to determine the range of model errors produced by models using parametric distributions. Since the truth of any complex real world phenomena cannot be known, the truth has to be assumed. The specific truth that is assumed also impacts the model results, i.e. when assuming a specific distribution and its parameters $F_1(x;\theta)$ it will be different to when assuming that the underlying truth is described by a different distribution $F_2(x;\theta)$.

For the purpose of this study different distributions, sample sizes and quantiles are chosen in order to establish the effect on the model risk quantum.

<u>Sample size</u>

The sample sizes n=100, 200, 500 and 1000 were considered. In loss modelling the sample sizes selected are considered to be small and the expectation is that the model errors produced using these smaller sample sizes will be larger and will therefore better demonstrate the effect of the sample sizes on Type 1 model risk.

The effect of the sample size on relative model error for Type 1 model risk will therefore be deduced from the results.

Parametric distributions

For the purpose of this study the lognormal and Burr type XII distributions are used.

The extreme value index (EVI) is a statistical measure summarising the "heaviness" of a statistical distribution's tails (see e.g. Beirlant et al. (2004) for more on the EVI).

The specific parametrisation of the distributions are chosen to span a range of different levels of tail "heaviness". This is determined by the EVI. For the purpose of this study a wide range of EVIs are chosen (0, 0.25 and 0.5) such that a wide range of distributions in terms of their tail heaviness is considered in order to establish the impact tail heaviness has on model risk.

The lognormal distribution function is given by

$$Lognorm(x;\mu,\sigma) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{ln(x)-\mu}{\sqrt{2}\sigma}\right) = \Phi\left(\frac{ln(x)-\mu}{\sigma}\right), \text{ for } x > 0$$

with location parameter $-\infty < \mu < \infty$, scale parameter $\sigma > 0$ and $\Phi(.)$ the standard Normal distribution function.

The EVI of the lognormal distribution is $EVI = \gamma = 0$.

The Burr type XII distribution function is given by

$$Burr(x;\eta,\kappa,\alpha) = 1 - (1 + (x/\eta)^{\kappa})^{-\alpha}, \text{ for } x > 0$$

with parameters $\eta, \kappa, \alpha > 0$. η is a scale parameter and both κ and α shape parameters.

The EVI is $EVI = \gamma = 1/\kappa\alpha$. The Burr distribution has an infinite mean when the EVI $\gamma \geq 1$ and the variance is infinite when EVI $\gamma \geq 0.5$.

Parametrisation was chosen as such to not include results where the distribution has an infinite mean. The expectation is that where the distribution has an infinite mean, the model risk results will be more severe.

Simulated parameters

The following is used in the simulation study $F_i(x;\theta)$:

$F_1(x) \equiv Lognorm(0,1)$ with $\gamma_1 = 0$ its EVI,

$F_2(x) \equiv Lognorm(0,2)$ with $\gamma_2 = 0$ its EVI,

$F_3(x) \equiv Burr(1,4,1)$ with $\gamma_3 = 0.25$ its EVI, and

$F_4(x) \equiv Burr(1,2,1)$ with $\gamma_4 = 0.5$ its EVI.

Simulation algorithm

The simulation is performed using a Monte Carlo simulation. Other methods such as the Panjer recursion and the Fast Fourier transformation can be used, but for the purpose of this study computational efficiency is not a concern. See for example Panman (2015). The process is explained next:

For the "true" quantile:

Calculate the percentile, q using the inverse of the cumulative distribution function i.e. ${F_i}^{-1}$ with the defined μ and σ for i=1,2 and the defined κ,η and α for i=3,4. For this study the 95th, 97.5th, 99th and 99.9th percentiles are included. The motivation behind the choice of quantiles is given next.

For the "estimated" quantile:

i) Generate n random X values from the distribution F_i.

ii) Estimate the parameters of F_i, i.e. $\hat{\mu}$, $\hat{\sigma}$ for i=1,2 and $\hat{\kappa}, \hat{\eta}$ and $\hat{\alpha}$ for i=3,4.

iii) Use the estimated parameters to estimate the q different quantiles where q is set equal to 95th, 97.5th, 99th and 99.9th percentile of F_i and the estimate is denoted by $\hat{q}$.

iv) Repeat the estimation of S times, where S=1000, denoted by $\hat{q}_s$.

The reason for including the 95th percentile is due to it being a relevant percentile for market risk capital. The 99.9th percentile is a popular choice in financial modelling for regulatory capital estimates. The other two percentiles are included to see the effect of two quantiles between the 95th and 99.9th percentile. For operational risk some research suggests that using a lower percentile than 99.9th can improve the reasonability of the estimates (Mignola et al., 2016).

VaR versus ES:

For the purpose of this study not only the VaR will be estimated, but also the ES. The VaR is simply the quantile (q) of the loss distribution. With the ES, instead of fixing the quantile, the VaR over all levels greater than the quantile are averaged and therefore it looks further into the tail of the loss distribution. It is therefore expected that the ES relative model errors will be greater than for the VaR.

Calculate the relative model error E_s:

$$E_s = \frac{\hat{q}_s}{q} - 1$$

Results

The results are summarised for each distribution $F_i(x; \theta)$.

$\underline{Lognorm(0,1)}$ <u>with EVI of 0</u>

The first set of results can be seen in Table 6-1. The table includes the results for the VaR and ES, as well as each of the quantiles of interest for the scope of this study. Where a number has been rounded to 0, more decimals have been added to illustrate that it is not exactly 0. A list of measures is given, but the main discussions will be around the 75th percentile, 25th percentile, the IQR and the median.

Table 6-1: $\boldsymbol{Lognorm(0,1)}$ relative model error results

Sample Size	**100**	**200**	**500**	**1000**	**100**	**200**	**500**	**1000**
				VaR				
Quantile	**95%**				**97.5%**			
Mean	-0.01	-0.001	0.01	0.0002	-0.01	0.00003	0.01	0.0004
Max	0.52	0.45	0.23	0.18	0.59	0.49	0.26	0.2
Q3	0.07	0.07	0.05	0.03	0.08	0.08	0.05	0.03
Median	-0.02	-0.002	0.004	-0.0002	-0.02	-0.001	0.004	-0.00003
Q1	-0.12	-0.08	-0.04	-0.03	-0.13	-0.09	-0.05	-0.04
IQR	0.2	0.15	0.09	0.06	0.22	0.17	0.1	0.07
Min	-0.42	-0.3	-0.16	-0.14	-0.47	-0.33	-0.18	-0.15
Range	0.95	0.74	0.39	0.32	1.05	0.82	0.44	0.35
Quantile	**99%**				**99.9%**			
Mean	-0.01	0.001	0.01	0.001	-0.01	0.004	0.01	0.002
Max	0.66	0.55	0.31	0.23	0.92	0.67	0.41	0.28
Q3	0.1	0.09	0.06	0.04	0.13	0.11	0.07	0.05
Median	-0.03	0.0003	0.004	-0.0004	-0.04	0.0003	0.004	0.0002
Q1	-0.15	-0.1	-0.05	-0.04	-0.18	-0.13	-0.06	-0.05
IQR	0.25	0.19	0.11	0.08	0.31	0.24	0.14	0.1
Min	-0.51	-0.36	-0.2	-0.16	-0.6	-0.43	-0.25	-0.19
Range	1.18	0.91	0.51	0.39	1.51	1.1	0.65	0.47
				ES				
Quantile	**95%**				**97.5%**			

Mean	-0.01	0.001	0.01	0.001	-0.01	0.002	0.01	0.001
Max	0.65	0.54	0.3	0.22	0.7	0.58	0.33	0.24
Q3	0.1	0.09	0.06	0.04	0.1	0.09	0.06	0.04
Median	-0.03	0.0004	0.004	-0.0001	-0.03	-0.0002	0.004	0.0003
Q1	-0.14	-0.1	-0.05	-0.04	-0.15	-0.11	-0.06	-0.04
IQR	-0.24	-0.18	-0.11	-0.08	0.26	0.2	0.12	0.08
Min	-0.5	-0.35	-0.2	-0.16	-0.53	-0.38	-0.21	-0.17
Range	1.15	0.89	0.49	0.38	1.23	0.95	0.54	0.41
Quantile	**99%**				**99.9%**			
Mean	-0.01	0.003	0.01	0.001	-0.004	0.01	0.01	0.002
Max	0.8	0.62	0.37	0.26	1.09	0.73	0.46	0.31
Q3	0.12	0.1	0.07	0.04	0.14	0.12	0.08	0.05
Median	-0.03	-0.001	0.004	-0.0003	-0.04	-0.002	0.003	-0.0002
Q1	-0.16	-0.12	-0.06	-0.05	-0.19	-0.13	-0.07	-0.06
IQR	0.28	0.22	0.13	0.09	0.33	0.25	0.15	0.11
Min	-0.56	-0.4	-0.23	-0.18	-0.63	-0.46	-0.27	-0.21
Range	1.36	1.03	0.59	0.44	1.72	1.19	0.72	0.51

Table 6-1 summarises the relative model error results for each quantile using VaR and ES. Within each of the four quantiles used, the results show a trend of decreasing relative model error with increasing sample size. It is also noted that the relative model error increases with increasing quantiles.

For example if the 75th percentile (Q_3) of the relative model error for the VaR results is observed, it can be shown that the relative model error for the 97.5th quantile decreases from 0.05 for a sample size of 500 to 0.03 for a sample size of 1000 (37.7% decrease). If the VaR 99.9th quantile results are compared to the 97.5th quantile for a sample size of 500, the Q_3 increases from 0.05 for the 97.5th quantile to 0.07 for the 99.9th quantile (35.8% increase).

If the VaR and ES results are compared, the ES sees more severe relative model error results. For example when only taking the 99th quantile results into account, the result for the IQR relative model error for a sample size of 500 decreases from 0.13 for the ES to 0.11 for the VaR. This translates into a 11.9% decrease when using VaR.

The percentage change results are useful when comparing the different distribution's results.

Figure 6-1 shows the summary of the boxplots for each of the four quantiles for the $Lognorm(0,1)$. The top four boxplots represent the results based on using ES and the bottom four the results based on using VaR. The y-axis shows the relative model error, while the x-axis shows the sample size n.

Figure 6-1: Relative model error for $Lognorm(0,1)$ for VaR and ES and different $\hat{q}$

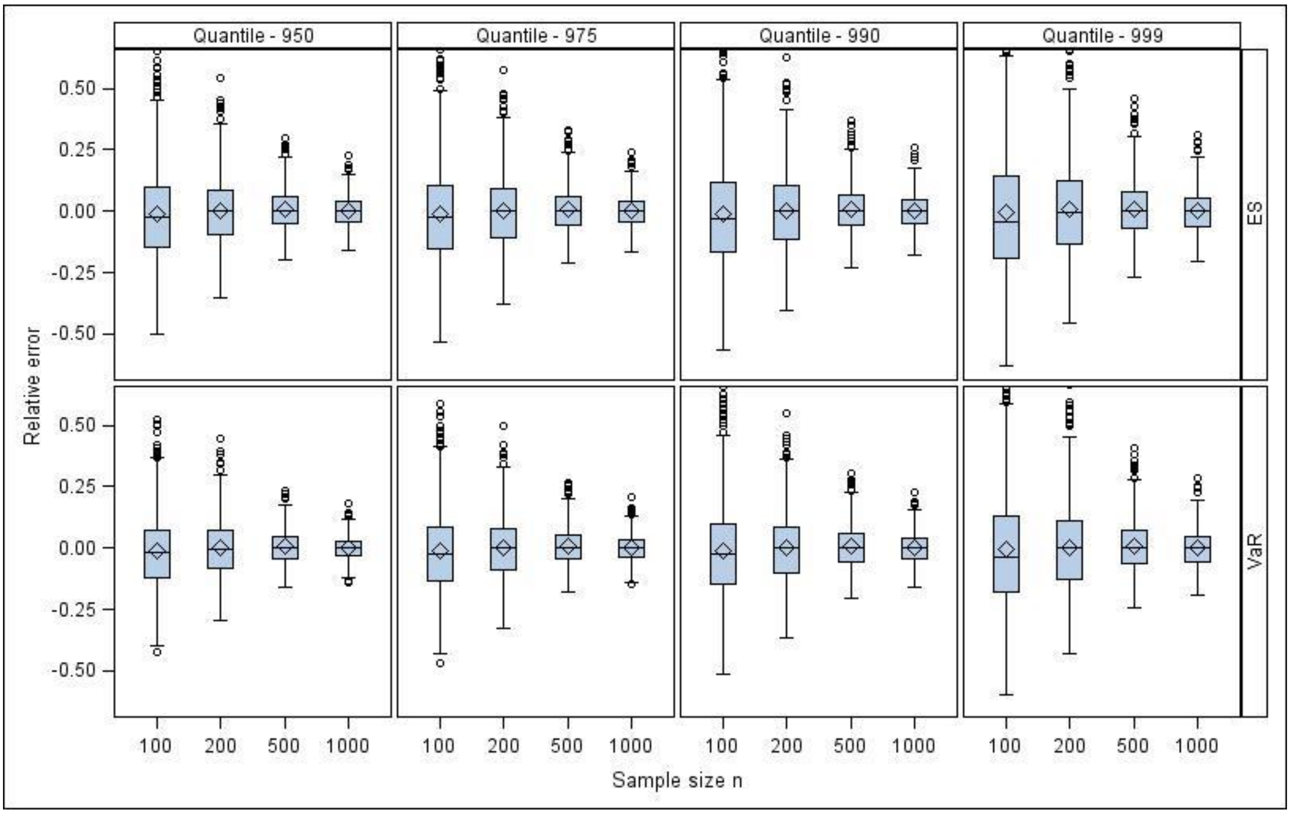


Figure 6-1 confirms the trends highlighted from the Table 6-1. It shows that the relative model error decreases with increasing sample size. This is the case for both the VaR and ES results.

The boxplots also show that the IQR of the relative model error increases with higher quantiles for both the VaR and ES results. For example the IQR for the 95th quantile and a sample size of 200 (0.15) is narrower than the IQR for the 99.9th quantile and a sample size of 200 (0.24).

The final observation that is made is that the relative model error seems to be slightly less severe for the VaR results than for the ES results. For example when still looking at the IQR for the 95th quantile and a sample size of 200, the range is wider for the ES (-0.18) than for the VaR (0.15).

<u>$Lognorm(0,2)$ with EVI of 0</u>

Table 6-2 summarises the results for the $Lognorm(0,2)$.

Table 6-2: $Lognorm(0,2)$ relative model error results

Sample Size	100	200	500	1000	100	200	500	1000

VaR								
Quantile	**95%**				**97.5%**			
Mean	0.03	0.02	0.01	-0.0005	0.04	0.03	0.01	0.0001
Max	1.50	1.01	0.57	0.35	1.85	1.26	0.61	0.39
Q3	0.21	0.15	0.09	0.06	0.23	0.17	0.11	0.07
Median	-0.01	0.004	-0.003	-0.005	-0.01	0.001	-0.003	-0.01
Q1	-0.20	-0.14	-0.09	-0.07	-0.22	-0.15	-0.10	-0.08
IQR	0.41	0.29	0.18	0.13	0.45	0.32	0.21	0.15
Min	-0.58	-0.53	-0.35	-0.25	-0.62	-0.56	-0.38	-0.29
Range	2.08	1.53	0.92	0.60	2.47	1.81	1.00	0.67
Quantile	**99%**				**99.9%**			
Mean	0.06	0.04	0.01	0.001	0.09	0.06	0.02	0.004
Max	2.33	1.63	0.66	0.44	3.61	2.61	0.87	0.61
Q3	0.27	0.20	0.12	0.08	0.35	0.25	0.14	0.10
Median	-0.03	-0.001	-0.01	-0.01	-0.04	-0.004	-0.01	-0.01
Q1	-0.24	-0.17	-0.11	-0.08	-0.30	-0.21	-0.14	-0.10
IQR	0.51	0.37	0.23	0.16	0.64	0.45	0.28	0.20
Min	-0.66	-0.59	-0.43	-0.33	-0.73	-0.65	-0.52	-0.42
Range	2.99	2.22	1.09	0.77	4.34	3.26	1.38	1.03
ES								
Quantile	**95%**				**97.5%**			
Mean	0.07	0.04	0.02	0.002	0.08	0.05	0.02	0.003
Max	2.89	2.05	0.71	0.50	3.18	2.28	0.77	0.54
Q3	0.29	0.21	0.13	0.09	0.31	0.23	0.13	0.09
Median	-0.03	-0.001	-0.003	-0.01	-0.03	-0.003	-0.01	-0.01
Q1	-0.25	-0.18	-0.12	-0.09	-0.27	-0.19	-0.12	-0.09
IQR	0.54	0.40	0.25	0.17	0.57	0.42	0.26	0.18
Min	-0.68	-0.61	-0.45	-0.36	-0.70	-0.62	-0.48	-0.38
Range	3.57	2.66	1.16	0.86	3.88	2.90	1.24	0.93
Quantile	**99%**				**99.9%**			
Mean	0.09	0.06	0.02	0.004	0.12	0.08	0.03	0.01
Max	3.60	2.60	0.85	0.60	4.83	3.54	1.09	0.74
Q3	0.34	0.24	0.14	0.10	0.40	0.28	0.17	0.11
Median	-0.04	-0.004	-0.01	-0.01	-0.05	-0.002	-0.01	-0.01
Q1	-0.29	-0.20	-0.13	-0.10	-0.33	-0.24	-0.16	-0.11
IQR	0.63	0.45	0.28	0.20	0.73	0.52	0.33	0.22
Min	-0.72	-0.64	-0.51	-0.41	-0.77	-0.68	-0.57	-0.47
Range	4.32	3.24	1.36	1.01	5.60	4.22	1.66	1.21

Table 6-2 summarises the relative model error results for each quantile using VaR and ES.

The results are more severe when compared to the previous $Lognorm(0,1)$ results. For example if the previous VaR Q_3 relative model error result for the 99.9th quantile and a sample size of 1000 is compared to the same result for the $Lognorm$(0,2), the Q_3 increases from 0.05 for the $Lognorm(0,1)$ to 0.10 for the $Lognorm$(0,2). This translates into a percentage increase of 108.1%.

Within each quantile set of results the trend of decreasing relative model error with increasing sample size is clear. It is also noted that the relative model error increases with increasing quantiles. For example if the Q_3 of the relative model error for the VaR results is observed, it can be seen that the relative model error for the 97.5th quantile decreases from 0.11 for a sample size of 500 to 0.07 for a sample size of 1000 (31.3% decrease). If the 99.9th quantile results are compared to the 97.5th quantile for a sample size of 500, the Q_3 increases from 0.11 for the 97.5th quantile to 0.14 for the 99.9th quantile (34.7% increase).

If the VaR and ES results are compared, the ES sees more severe relative model error results. For example when only taking the 99th quantile results into account, the result for the IQR relative model error for a sample size of 500 decreases from 0.28 for the ES to 0.23 for the VaR. This translates into a 16.3% decrease when using VaR. For the $Lognorm(0,1)$ results this percentage decrease was less at 11.9%.

Figure 6-2 shows the results for the $Lognorm(0,2)$.

Figure 6-2: Relative model error for $Lognorm(0, 2)$ for VaR and ES and different $\hat{q}$

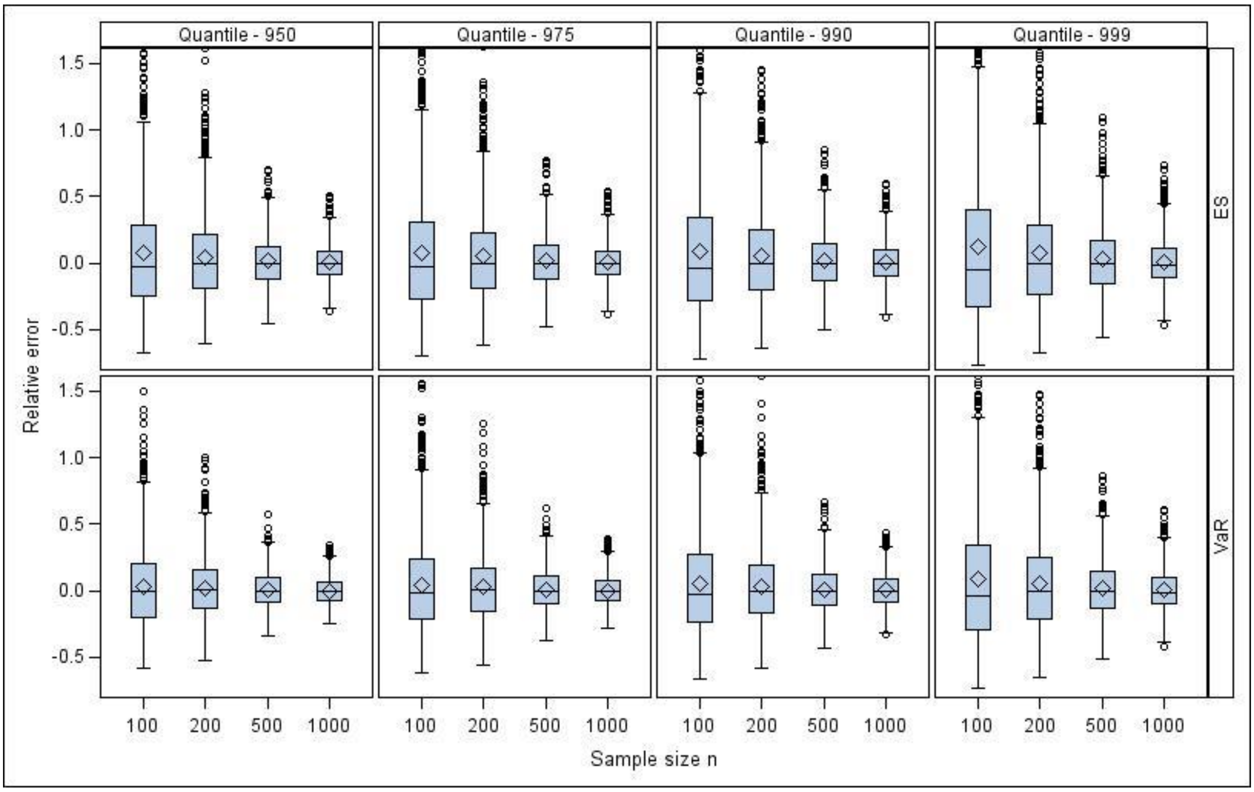


Figure 6-2 confirms a similar trend when comparing to the $Lognorm(0,1)$. The trend is similar, but the $Lognorm(0,2)$ sees more severe relative model errors.

The relative model error still decreases with increasing sample size and increases for higher percentiles.

The relative model errors are more severe for the ES results than the VaR. This can be illustrated by using the IQR for the 95th quantile and a sample size of 200, where the range for the ES (0.54) is clearly wider than for the VaR results (0.41).

When comparing the $Lognorm(0,1)$ to the $Lognorm(0,2)$, it can be seen that using the IQR for the 95th quantile and a sample size of 200, the range for the $Lognorm(0,2)$ is not only wider than the range for the $Lognorm(0,1)$ (0.41 versus 0.15), but the values across which it spans are more severe, i.e. the Q_1 (-0.14 versus -0.08) and Q_3 (0.15 versus 0.07) are more severe for the $Lognorm(0,2)$.

$\boldsymbol{Burr(1,4,1)}$ with EVI of 0.25

Table 6-3 shows the results for the $Burr(1,4,1)$ which has an EVI of 0.25.

Table 6-3: $\boldsymbol{Burr(1,4,1)}$ relative model error results

Sample Size	100	200	500	1000	100	200	500	1000
VaR								
Quantile	**95%**				**97.5%**			
Mean	-0.001	0.001	-0.002	-0.001	-0.001	0.002	-0.003	-0.001
Max	0.41	0.22	0.16	0.10	0.62	0.29	0.20	0.14
Q3	0.06	0.04	0.02	0.02	0.08	0.06	0.03	0.02
Median	-0.01	-0.004	-0.01	-0.001	-0.02	-0.01	-0.01	-0.002
Q1	-0.07	-0.05	-0.03	-0.02	-0.09	-0.06	-0.04	-0.03
IQR	0.13	0.09	0.06	0.04	0.18	0.13	0.08	0.05
Min	-0.23	-0.18	-0.14	-0.08	-0.28	-0.22	-0.17	-0.11
Range	0.64	0.40	0.30	0.18	0.90	0.50	0.37	0.25
Quantile	**99%**				**99.9%**			
Mean	0.001	0.004	-0.003	-0.002	0.02	0.01	-0.003	-0.002
Max	0.95	0.42	0.28	0.20	2.16	0.85	0.54	0.37
Q3	0.12	0.09	0.05	0.03	0.21	0.15	0.08	0.06
Median	-0.03	-0.01	-0.01	-0.003	-0.06	-0.01	-0.01	-0.01
Q1	-0.13	-0.08	-0.06	-0.04	-0.22	-0.14	-0.11	-0.07
IQR	0.25	0.17	0.11	0.07	0.43	0.30	0.19	0.13
Min	-0.38	-0.28	-0.22	-0.15	-0.59	-0.45	-0.34	-0.26

Range	1.33	0.70	0.50	0.36	2.76	1.30	0.88	0.63
ES								
Quantile	**95%**				**97.5%**			
Mean	0.003	0.004	-0.003	-0.001	0.01	0.01	-0.003	-0.001
Max	0.94	0.40	0.27	0.19	1.24	0.51	0.33	0.24
Q3	0.11	0.08	0.04	0.03	0.13	0.09	0.05	0.04
Median	-0.02	-0.01	-0.01	-0.003	-0.03	-0.01	-0.01	-0.005
Q1	-0.11	-0.08	-0.06	-0.04	-0.14	-0.10	-0.07	-0.04
IQR	0.23	0.16	0.10	0.07	0.28	0.19	0.12	0.08
Min	-0.34	-0.26	-0.21	-0.14	-0.41	-0.31	-0.24	-0.17
Range	1.28	0.66	0.47	0.33	1.65	0.82	0.58	0.41
Quantile	**99%**				**99.9%**			
Mean	0.01	0.01	-0.003	-0.002	0.04	0.03	0.000	-0.001
Max	1.71	0.68	0.43	0.30	3.39	1.18	0.72	0.49
Q3	0.16	0.12	0.07	0.05	0.26	0.20	0.10	0.08
Median	-0.04	-0.01	-0.01	-0.01	-0.07	-0.01	-0.02	-0.01
Q1	-0.18	-0.12	-0.09	-0.06	-0.27	-0.18	-0.13	-0.08
IQR	0.34	0.24	0.15	0.10	0.53	0.38	0.24	0.16
Min	-0.50	-0.38	-0.29	-0.22	-0.68	-0.53	-0.41	-0.32
Range	2.21	1.06	0.72	0.52	4.08	1.71	1.12	0.81

Table 6-3 summarises the relative model error results for each quantile using VaR and ES.

The results seem less severe than seen for the previous $Lognorm(0,2)$ results. For example if the previous VaR Q_3 relative model error result for the 99.9th quantile and a sample size of 1000 is compared to the same result for the $Burr(1,4,1)$ the Q_3 decreases from 0.10 for the $Lognorm(0,2)$ to 0.06 for the $Burr(1,4,1)$. This translates into a percentage decrease of 39.8%.

The trend of decreasing relative model error with increasing sample size is clear and seems to happen at an increased tempo when comparing to the previous results. For example if Q_3 of the relative model error for the VaR results is observed, it can be shown that the relative model error for the 97.5th quantile decreases from 0.06 for a sample size of 500 to 0.03 for a sample size of 1000. This is a percentage decrease of 47.0%, where the $Lognorm(0,2)$ only showed a 31.2% decrease.

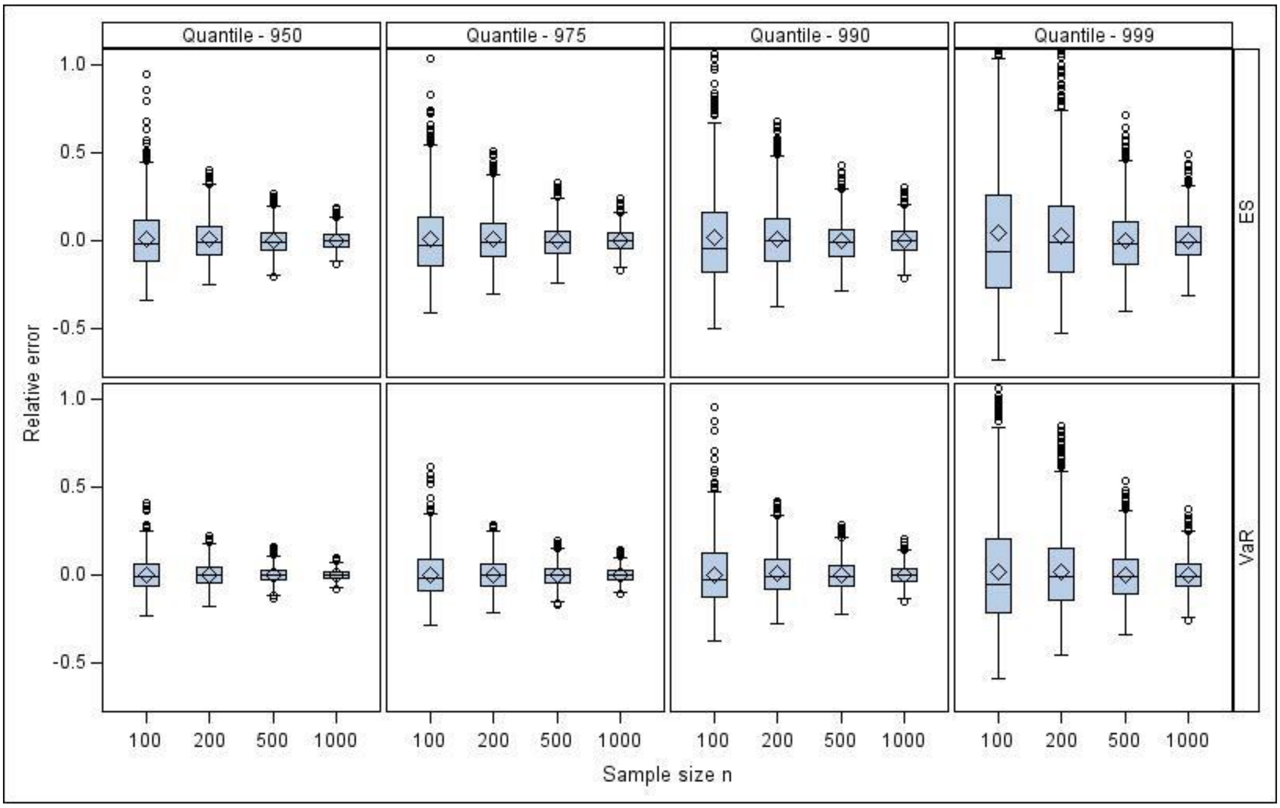
It is also noted that the relative model error increases with increasing quantiles. If the 99.9th quantile results are compared to the 97.5th quantile for a sample size of 500, the Q_3 increases from 0.06 for the 97.5th quantile to 0.08 for the 99.9th quantile. This is a 29.5% increase, which is less than the 34.7% increase seen for the $Lognorm(0,2)$, but still a high percentage change.

If the VaR and ES results are compared, the ES sees more severe relative model error results. For example when only taking the 99th quantile results into account, the result for the IQR relative model error for a sample size of 500 decreases from 0.10 for the ES to 0.07 for the VaR. This translates into a 44.3% decrease when using VaR. For the $Lognorm(0,2)$ results this percentage decrease was less at 16.3%.

From this set of results it can be seen that the results are less severe than for the $Lognorm(0,2)$. Important conclusions that can be made for the results are that the increased EVI results in a higher tempo decrease in the relative model error when the sample size is increased than for a distribution with an EVI of 0. The increase in relative model error for the distribution with a higher EVI when the quantiles are increased is less pronounced than for the distribution with an EVI of 0. And lastly when a distribution has a higher EVI the difference in using ES or VaR (VaR is less than ES) is more prominent than for the distributions with an EVI of 0.

Figure 6-3 shows the boxplots for the $Burr(1,4,1)$ which introduces a higher EVI than the lognormal distributions (which have an EVI of 0).

Figure 6-3: Relative model error for $\boldsymbol{Burr(1,4,1)}$ for VaR and ES and different $\hat{\mathbf{q}}$



The relative model error can be seen to decrease as the sample size increases. The decrease is more prominent than for the previous two distributions. I.e. the tempo at which the relative model error decreases is higher than for the $Lognorm(0,1)$ and $Lognorm(0,2)$.

The increase in the relative model error for higher quantiles is also clear from the boxplots when comparing the VaR IQR relative model error results for a sample size of 200. The 95th quantile gives 0.09, the 97.5th quantile 0.13, the 99th quantile 0.17, and the 99.9th quantile the highest at 0.30.

It can be seen that the ES has a more severe relative model error than the VaR. This difference is also more pronounced than for the previous distributions. It can imply that when using a distribution with a higher EVI, to avoid or limit potential model errors, the VaR is likely to be preferred over the ES.

$\boldsymbol{Burr(1,2,1)}$ with EVI of 0.5

Table 6-4 summarises the results for the $Burr(1,2,1)$.

Table 6-4: $\boldsymbol{Burr(1,2,1)}$ relative model error results

Sample Size	**100**	**200**	**500**	**1000**	**100**	**200**	**500**	**1000**
VaR								
Quantile	**95%**				**97.5%**			
Mean	0.01	-0.01	0.003	0.001	0.02	-0.005	0.01	0.001
Max	1.33	0.70	0.37	0.23	1.96	0.96	0.56	0.32
Q3	0.11	0.07	0.05	0.04	0.16	0.09	0.07	0.05
Median	-0.02	-0.02	-0.001	-0.001	-0.02	-0.03	-0.001	-0.003
Q1	-0.13	-0.10	-0.05	-0.04	-0.17	-0.14	-0.07	-0.05
IQR	0.24	0.17	0.11	0.08	0.33	0.23	0.15	0.11
Min	-0.42	-0.34	-0.23	-0.21	-0.51	-0.41	-0.28	-0.26
Range	1.75	1.04	0.60	0.44	2.47	1.37	0.84	0.58
Quantile	**99%**				**99.9%**			
Mean	0.05	0.003	0.01	0.002	0.18	0.05	0.04	0.01
Max	3.11	1.38	0.88	0.45	9.78	3.75	2.01	0.87
Q3	0.22	0.14	0.11	0.07	0.40	0.26	0.20	0.12
Median	-0.03	-0.04	-0.0003	-0.01	-0.04	-0.06	0.002	-0.01
Q1	-0.23	-0.18	-0.10	-0.08	-0.38	-0.30	-0.16	-0.13
IQR	0.44	0.32	0.20	0.15	0.78	0.55	0.36	0.26
Min	-0.61	-0.52	-0.37	-0.32	-0.79	-0.72	-0.57	-0.49
Range	3.72	1.90	1.24	0.77	10.57	4.47	2.57	1.35
ES								
Quantile	**95%**				**97.5%**			
Mean	0.03	0.02	0.02	0.01	0.08	0.04	0.03	0.01
Max	1.92	1.98	1.30	0.56	9.68	2.49	1.65	0.68
Q3	0.23	0.15	0.13	0.08	0.29	0.20	0.15	0.10

Median	-0.05	-0.04	-0.0005	-0.01	-0.04	-0.05	0.001	-0.01
Q1	-0.25	-0.20	-0.11	-0.08	-0.29	-0.23	-0.13	-0.10
IQR	0.48	0.35	0.24	0.16	0.58	0.43	0.28	0.20
Min	-0.62	-0.53	-0.38	-0.33	-0.68	-0.60	-0.45	-0.37
Range	2.54	2.51	1.69	0.90	10.36	3.09	2.09	1.05
Quantile	**99%**				**99.9%**			
Mean	0.14	0.06	0.04	0.01	0.10	0.07	0.08	0.03
Max	4.39	3.38	2.19	0.86	3.37	3.50	1.93	1.40
Q3	0.38	0.25	0.19	0.12	0.47	0.34	0.29	0.18
Median	-0.04	-0.06	0.001	-0.01	-0.11	-0.09	0.01	-0.02
Q1	-0.35	-0.28	-0.15	-0.13	-0.49	-0.39	-0.21	-0.18
IQR	0.73	0.53	0.34	0.24	0.96	0.72	0.50	0.36
Min	-0.75	-0.67	-0.52	-0.44	-0.87	-0.81	-0.68	-0.61
Range	5.14	4.06	2.71	1.30	4.25	4.31	2.60	2.01

Table 6-4 summarises the relative model error results for each quantile using VaR and ES. The $Burr(1,2,1)$ has a higher EVI than the $Burr(1,4,1)$ (i.e. 0.5>0.25).

The results are more severe than seen for the previous $Burr(1,4,1)$ results. For example if the previous VaR Q_3 relative model error result for the 99.9th quantile and a sample size of 1000 is compared to the same result for the $Burr(1,2,1)$ the Q_3 increases from 0.06 for the $Burr(1,4,1)$. to 0.12 for the $Burr(1,2,1)$. This translates into a percentage increase of 104.4%. If the same result is compared to the $Lognorm(0,1)$ and the $Lognorm(0,2)$ result, an increase of 156.0% and 23.0% is obtained. It is clear that the results for the $Burr(1,2,1)$ is the most severe.

Within each quantile set of results the trend of decreasing relative model error with increasing sample size is clear. For example if the Q_3 of the relative model error for the VaR results is observed, it can be seen that the relative model error for the 97.5th quantile decreases from 0.07 for a sample size of 500 to 0.05 for a sample size of 1000. This is a percentage decrease of 32.0%, which is less than the 47.0% decrease seen for the $Burr(1,4,1)$.

It is also noted that the relative model error increases with increasing percentiles. If the VaR 99.9th quantile results are compared to the 97.5th quantile for a sample size of 500, the Q_3 increases from 0.09 for the 97.5th quantile to 0.20 for the 99.9th quantile. This is a 168.6% increase, which is greater than the 29.5% increase seen for the $Burr(1,4,1)$. This implies that the very high EVI leads to a faster tempo of increase in the relative model error when the quantiles increase.

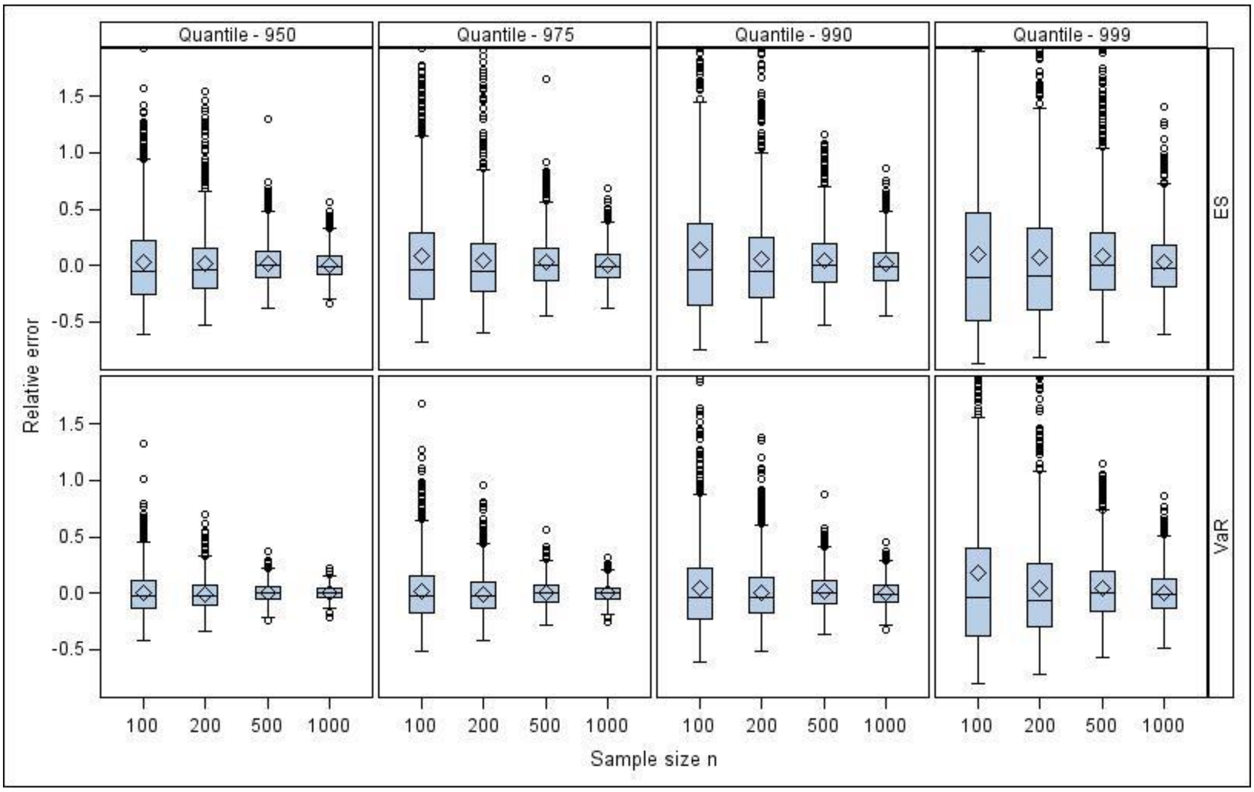
If the VaR and ES results are compared, the ES sees more severe relative model error results. For example when only taking the 99th quantile results into account, the result for the IQR

relative model error for a sample size of 500 decreases from 0.34 for the ES to 0.20 for the VaR. This translates into a 40.1% decrease when using VaR. This is similar to the result seen for the $Burr(1,4,1)$.

From this set of results it can be seen that the results are more severe than the other distributions in scope of the study. Important conclusions that can be made for the results are that a very high EVI results in a higher tempo increase in the relative model error when the quantiles increase. When a distribution has a high EVI the difference in using ES or VaR (VaR is less than ES) is more prominent than for the distributions with an EVI of 0.

Figure 6-4 shows the boxplots for the results of this heavier tailed distribution.

Figure 6-4: Relative model error for $\boldsymbol{Burr(1,2,1)}$ for VaR and ES and different $\hat{\mathbf{q}}$



It can immediately be seen that the results summarised in Figure 6-4 are more severe than the previous three sets of results. There are more outliers in the results and the scale is more severe when comparing to the previous results.

The same trends can be noted regarding the decrease in relative model error with an increase in sample size and the increase in relative model error when using higher quantiles.

It can be seen that the relative model error increases with an increasing EVI. When comparing the VaR IQR for the 99.9th quantile and a sample size of 200 for the $Burr(1,2,1)$ it is wider at 0.55 than for the $Burr(1,4,1)$ which has a value of 0.38.

As seen with the $Burr(1,4,1)$, the increase in relative model error when using ES instead of VaR is very pronounced. The IQR is narrower for the VaR estimates of all quantiles than for the ES estimates. This can imply that the VaR is potentially a safer bet when it comes to distributions with high EVI values.

Summary of results

The results from the previous section is next summarised in order to illustrate the effect of the sample size, EVI and quantile on the model risk. The results are shown for Q_3, the median and Q_1 of the VaR.

Table 6-5: Q_3 Results

Sample Size		100	200	500	1000	100	200	500	1000
Quantile		95%				97.5%			
Distribution	F_1	0.07	0.07	0.05	0.03	0.08	0.08	0.05	0.03
	F_2	0.21	0.15	0.09	0.06	0.23	0.17	0.11	0.07
	F_3	0.06	0.04	0.02	0.02	0.08	0.06	0.03	0.02
	F_4	0.11	0.07	0.05	0.04	0.16	0.09	0.07	0.05
Quantile		99%				99.9%			
Distribution	F_1	0.10	0.09	0.06	0.04	0.13	0.11	0.07	0.05
	F_2	0.27	0.20	0.12	0.08	0.35	0.25	0.14	0.10
	F_3	0.12	0.09	0.05	0.03	0.21	0.15	0.08	0.06
	F_4	0.22	0.14	0.11	0.07	0.40	0.26	0.20	0.12

Table 6-5 shows that if the 75th percentile is used as a proxy for model risk, the results are mostly over-estimated. The results also show the relative model error decreases with increasing sample size, increases with higher quantiles and that the distribution with the highest EVI, shows more severe results. The highest value is 0.40, for the $Burr(1,2,1)$ for the 99.9th quantile and a sample size of 100.

Table 6-6: Median Results

Sample Size		100	200	500	1000	100	200	500	1000
Quantile		95%				97.5%			
Distribution	F_1	-0.02	-0.002	0.004	-0.0002	-0.02	-0.001	0.004	-0.00003
	F_2	-0.01	0.004	-0.003	-0.005	-0.01	0.001	-0.003	-0.01
	F_3	-0.01	-0.004	-0.01	-0.001	-0.02	-0.01	-0.01	-0.002

	F_4	-0.02	-0.02	-0.001	-0.001	-0.02	-0.03	-0.001	-0.003
Quantile		99%				99.9%			
Distribution	F_1	-0.03	0.0003	0.004	-0.0004	-0.04	0.0003	0.004	0.0002
	F_2	-0.03	-0.001	-0.01	-0.01	-0.04	-0.004	-0.01	-0.01
	F_3	-0.03	-0.01	-0.01	-0.003	-0.06	-0.01	-0.01	-0.01
	F_4	-0.03	-0.04	-0.0003	-0.01	-0.04	-0.06	0.002	-0.01

Table 6-6 shows that if the median is used as a proxy for model risk, the results are mostly close to zero, which can be interpreted as minimal (immaterial) model risk.

Table 6-7: Q_1 Results

Sample Size		100	200	500	1000	100	200	500	1000
Quantile		95%				97.5%			
Distribution	F_1	-0.12	-0.08	-0.04	-0.03	-0.13	-0.09	-0.05	-0.04
	F_2	-0.20	-0.14	-0.09	-0.07	-0.22	-0.15	-0.10	-0.08
	F_3	-0.07	-0.05	-0.03	-0.02	-0.09	-0.06	-0.04	-0.03
	F_4	-0.13	-0.10	-0.05	-0.04	-0.17	-0.14	-0.07	-0.05
Quantile		99%				99.9%			
Distribution	F_1	-0.15	-0.10	-0.05	-0.04	-0.18	-0.13	-0.06	-0.05
	F_2	-0.24	-0.17	-0.11	-0.08	-0.30	-0.21	-0.14	-0.10
	F_3	-0.13	-0.08	-0.06	-0.04	-0.22	-0.14	-0.11	-0.07
	F_4	-0.23	-0.18	-0.10	-0.08	-0.38	-0.30	-0.16	-0.13

Table 6-7 shows that if the 25^{th} percentile is used as a proxy for model risk, the results are mostly under-estimated. The results also show the relative model error decreases with increasing sample size, increases with higher quantiles and that the distribution with the highest EVI, shows more severe results. The most severe Q_1 relative model error value is -0.38, for the $Burr(1,2,1)$ for the 99.9^{th} quantile and a sample size of 100.

It can therefore be concluded that the median relative model error is not as severe as the 75^{th} or 25^{th} percentile values.

From the results of the study the following contributing factors to model risk can be made:

i) The relative model error increases when using ES instead of VaR.
ii) The relative model error increases with very high values of EVI.
iii) The relative model error increases with decreasing values of sample size.
iv) The relative model error increases with increasing quantiles.

The further investigation into the effect on model risk of using ES instead of VaR is proposed as future research.

Concluding remarks

The focus of this section is to quantify the model risk related to Type 1 model risk, or the risk associated with parameter uncertainty. Four different distributions were used to illustrate the effect of parameter uncertainty related to sample size, EVI, different quantiles, and VaR versus ES.

As expected the relative model error decreases with increasing sample size. When the sample size is 100 and VaR is used, the relative model error ranges from 7% to 40%, when using the Q_3 as a proxy for model risk and looking at each $F_i(x;\theta)$. If the sample size increases to 1000, the relative model error ranges from 2% to 12%, which is much lower than for the sample of 100. It is also noted that the tempo at which the relative model error decreases is enhanced for distributions with a higher EVI. The $Burr(1,4,1)$ with an EVI of 0.25 showed a decrease of close to 50% in the relative model error when the sample size was increased from 500 to 1000. This is higher than the $Lognorm(0,2)$ which only showed decrease close to 30%.

It can also be seen that the magnitude of the model errors increased with an increase in the quantiles for all four of the distributions analysed. The results show that this increase has a higher tempo for distributions with a very high EVI. The $Burr(1,2,1)$, which has an EVI of 0.5 saw a 168.6% increase in relative model error for the Q_3 and a sample size of 500 when moving from the 97.5th quantile to the 99.9th quantile.

It is also noted that the relative model errors for the ES is more severe than for the VaR. For the case of the $Burr(1,4,1)$, the decrease in relative model error when using VaR instead of ES translated into a 44% decrease.

Type 1 model errors can range from immaterial to more than 100% based on the parameters of this simulation study, as well as the definition of model error and using different measures for example the IQR, Q_3, maximum relative error and the median.

Globally there have been cases where regulators require banks to estimate model risk capital. In the survey conducted by ORX, a key finding was that one third of the banks that took part in the survey was expected to quantify model risk for Pillar 2 capital purposes (ORX,2016). If at a point a regulatory requirement for quantifying model risk capital is established, this method can possibly be used to potentially satisfy expectations from regulators.

It can be argued that even though there is a risk involved in estimating model risk, it can be used for materiality based model risk management to ensure that models with greater exposure are afforded more robust model risk mitigating measures. It is therefore a valuable tool in model risk management.

6.3 Quantifying Type 2 model risk

In Chapter 3 Type 2 model risk, or model misspecification, is explained as arising when a non-optimal model is selected to represent the current underlying dynamics of a real-world phenomenon.

This dissertation does not detail the methodology of how to quantify Type 2 model risk, but it is however suggested as a future research topic and will work similarly to the quantification process of Type 1 model risk, where the simulation approach will be used to measure the relative error when a non-optimal model is specified. I.e. in the case of Type 1 model risk it was assumed to be the most optimal model (Type 2 model risk assumed to be non-existent) and the parameters were varied. In the case of Type 2 model risk, the model types will be varied.

The expectation is that Type 2 model risk will have a greater effect when compared to Type 1 model risk. The reason for this is that Type 2 model risk also includes Type 1 model risk.

6.4 Quantifying Type 3 model risk

Type 3 model risk is defined in Chapter 3 as a change in the dynamics of the real-world phenomenon and occurs when the model was initially optimally specified with optimal parameters, but due to a recent change in the dynamics of the underlying real-world phenomenon, the model is no longer suitable.

This is a challenging model risk category to quantify. It is not always possible to identify a change in the underlying real-world phenomenon. It is proposed to use some of the mitigating methods detailed in Chapter 4 to assist with the identification and mitigation of this model risk type. Model monitoring can be a key tool to use in order to identify when the model is no longer optimal due to a change in the dynamics of the underlying real-world phenomenon. It is also proposed to investigate the quantification of this model risk type as future research.

6.5 Quantifying Type 4 model risk

Type 4 model risk is defined as the incorrect model implementation, misinterpretation of model output, and other errors occurs when an optimally developed model is implemented, used or interpreted incorrectly which leads to inaccurate results or non-optimal conclusions.

This model risk type is also challenging to quantify as it involves human error and subjectivity. It is also challenging quantifying the magnitude of a model error caused by human error. The proposal to quantify this model risk type is to use an approach similar to that of operational risk loss modelling and insurance risk modelling. This involves using a LDA model.

The LDA model requires loss data. It is proposed that the loss data for this model risk type will be gathered by a model validation function. The role of model validation will be to estimate a value for each error that is identified during a validation of each model. The loss data will have to be gathered over a couple of years in order to be used for loss modelling.

This method can also be used without using loss data. Scenario data can be used instead of loss data. This would involve the use of expert judgment to come up with suitable scenarios that can be used in the LDA model.

The model separates the loss (actual loss or scenario data) into a frequency and severity component. A statistical distribution of the frequency of loss events is then estimated from the model risk loss data. A statistical distribution of the severity of the model risk loss data is also estimated. A single loss distribution is then calculated using the estimated frequency and severity distributions and simulation such as Monte Carlo. The Type 4 model risk quantum can then be taken as a quantile from the compound loss distribution. See more on LDA modelling for example Frachot et al. (2001) and Chernobai et al. (2007).

This is a high level overview of the method for quantifying Type 4 model risk and a more detailed version is proposed for future research. The mitigating efforts detailed in Chapter 4 are key in reducing the impact of Type 4 model risk.

6.6 Model risk scorecard

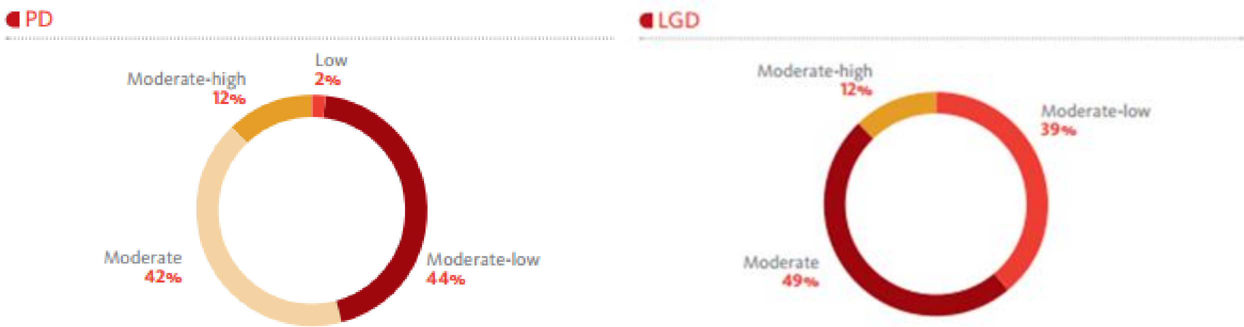
A qualitative assessment approach based on expert judgment can be used to measure the model risk relating to a specific model. This approach does not quantify model risk, but rather gives an indication of how risky a model is relative to the rest of the models in scope.

6.6.1 Background on the use of model risk scorecards

De Jongh et al. (2017a) propose using a model validation scorecard based on best practices and comprising of three components: model validation governance scorecard, model validation policy scorecard and the model validation process scorecard. The scorecards use a four grade scale (1 to 4) which is in line with what is used by the Regulatory Consistency Assessment Programme (RCAP) (BCBS, 2016c). The scorecards have no weights added as the purpose is not to obtain an aggregate score. Therefore their proposal resembles a dashboard rather than a scorecard. However, it can easily be constructed as a scorecard by defining appropriate weights for each dashboard element. Some of the elements that are rated in the scorecards overlap to what is proposed in this paper. The main difference is that in this paper the purpose of the scorecard is to obtain an overall score for a model in order to compare the model risk relative to other models.

Malan (2017) summarises the public disclosure of selected banks regarding model risk and shows that Grupo Santander uses a scoring based assessment to establish the level of model risk for each model type, for example the proportion of low, moderate-low, moderate, moderate high or high risk is illustrated for each model type. See Figure 6-5 where the PD and LGD model risk results are shown. This aligns with the objective of the scorecard proposed in this paper to give a relative model risk score for each model. The main difference here is that the proposed scorecard will be more granular than the one for Grupo Santander. I.e. each model will have a score and this will be relative to the other models in scope.

Figure 6-5: Grupo Santander model risk assessment results for PD and LGD models



(Santander, 2016)

FirstRand (2016) mentions a scorecard approach is followed to measure model risk. It states that risk factors based on model risk management principles are used for the measurement of model risk. The scorecard is customised for each risk type through the use of weighting and refining of considerations used. The ratings used are low, medium or high risk. This mostly aligns to the scorecard proposed in this dissertation as it also aligns with model risk management principles.

The scorecard proposed for the work in this dissertation will be discussed next.

6.6.2 Proposed qualitative assessment scorecard

A qualitative approach to score model risk for each model in scope using model risk management principles detailed in Chapter 5, is proposed.

Assessment items

In Chapter 5, model risk management is dissected into a five-step process based on the paper by van Biljon and Haasbroek (2017a). These five steps are:

A. Data quality, extraction and transformation
B. Model definition, development and documentation
C. Model validation and approval
D. Model implementation, change control and usage
E. Reporting and monitoring

The assessment items of the proposed scorecard are based on the model risk management five step process. Each assessment item has sub sections that are evaluated using a five-point rating scale.

Assessment scale

As explained in Chapter 5 the five-point rating scale is designed to allow sufficient granularity in the rating of each of the model risk management process steps. De Jongh et al. (2017a) propose the use of a “four-grade” rating scale in order to eliminate respondents selecting the mid-point or average score that can be seen with a 3-point or a 5-point scale. The five-point rating scale is chosen for the purpose of this research to align with the model risk management process in Chapter 5.

What differs here from Chapter 5 is that a rating of 5 in this scorecard translates into “very high risk” and a rating of 1 translates to “very low risk” i.e.:

i) a 5-rating is “very high risk” exposure and translates into the model risk requiring major improvement;
ii) a 4-rating is “high risk” exposure and implies that some aspects of model risk require major improvement;

iii) a 3-rating is "moderate risk" exposure and indicates that the model risk adequate, with only some aspects requiring improvement;

iv) a 2-rating is "low risk" exposure and implies the level of model risk is more than adequate, with minor room for further improvement;

v) a 1-rating is "very low" risk exposure and is equivalent to leading-edge model risk practices being observed.

Scorecard

Based on research, perceived best practices and experience as summarised in Chapter 5, the scorecard shown in Table 6-8 is proposed. Note that the assessment items and weights shown in this scorecard contain a degree of subjectively and should be further customised for the specific purpose it will be used.

Table 6-8: Model risk scorecard

	Assessment item	Weight	1	2	3	4	5
1	**Data quality, extraction and transformation**	**0.2**					
A	Any known data issues have a work-around or is scheduled to be fixed	0.3	All known data issues have been fixed	All known data issues have a work-around or is scheduled to be fixed	All data issues have a work-around, but are not scheduled to be fixed	Some data issues have a work-around, but are not scheduled to be fixed	Most data issues do not have a work-around and are not scheduled to be fixed
B	Automation of data feeds and transformation	0.2	All data feeds and transformation is automated		All data feeds are automated, but the transformation is not		There is no automation of the data feeds and transformation
C	Documentation of data source and transformation	0.5	The documentation includes a high level of detail regarding the data source and transformation	The documentation includes a high level of detail regarding the data source but less information on	The documentation includes a moderate level of detail regarding the data source but less information on transformation	The documentation includes a low level of detail regarding the data source but less information on transformation	The documentation does not include any information on data or transformation

				transformation			
2	**Model definition, development and documentation**	**0.3**					
A	Minimum standards that define the development is documented	0.4	A high level of detail that define the development standards are documented		Minimum standards that define the development is documented		The documentation on the standards for development is low quality
B	Level of experience and skill of developer	0.2	Highly experienced analyst		Analyst with some experience relating to the model that is developed, i.e. two to three years		Analyst with less than two years of experience
C	Model inventory is in place	0.2	A very detailed and up to date model inventory is in place	An up to date model inventory containing the minimum required information is in place	An outdated model inventory containing the minimum required information is in place	A model inventory is planned, but yet to be finalised	No model inventory is in place with no plans to construct one
D	Complexity of real-world phenomenon being modelled	0.2	Simple problem and industry practice is used to develop the model		Relatively complex problem, but commonly used industry assumptions are available		Very complex problem that is not well known in industry
3	**Model validation and approval**	**0.2**					
A	Model validation is independent of model development	0.3	Model validation is completely independent of model development with their own independent	Model validation is independent of model development, but certain aspects of	Model validation is independent of model development, but large portions of the code comes	Model validation is not fully independent of model development and relies on model	Model validation is part of model development

			process/code used to validate models	the code is sourced from developmen t	from model development	development in order to complete the validation	
B	Pre-implementation approval of new or amended models takes place at governance committees	0.4	Model validation is performed according to the agreed frequency and is always taken through governance on time	Model validation is performed according to the agreed frequency and is taken through governance , even though it might go through late	Model validation is performed according to the agreed frequency, with some validations being performed late but with appropriate notification to relevant parties and the models are taken though governance	Model validation is typically only performed once at inception of the model, with no annual review or ongoing validation and does not always go through governance	Model validation of a new or amended model does not occur and is not taken through governance
C	Model validation documentation	0.3	Minimum standards that define how model validation should be performed are documented in a model validation policy that is annually approved at the appropriate committee and model validation documentation is complete and of a high standard	Minimum standards that define how model validation should be performed are documente d in a model validation policy that is annually approved at the appropriate committee and model validation documentati on is complete and up to standard	Minimum standards that define how model validation should be performed are documented in a model validation policy, but are not necessarily annually approved at the appropriate committee and model validation documentation requires minor improvement	Minimum standards that define how model validation should be performed have evolved, but are not documented or approved at the appropriate committee and model validation documentation requires improvement	Minimum standards that define how model validation should be performed are not mature, documented or approved at the appropriate committee and model validation documentation is limited or not available
4	**Model**	**0.2**					

	implementation, change control and usage						
A	Models are implemented in a production-status environment with proper change-control and access management	0.3	Models are implemented in a production status environment with proper change control and access controls	Models are implemented in a production status environment with proper change control and some access controls	Models are implemented in non-production status environments such as Excel or SAS, but adequate controls are in place to mimic a production status environment with proper change control and access controls	Models are implemented in non-production status environment such as Excel or SAS, and some control deficiencies in respect of change control and access controls are present	Models are implemented in non-production status environment such as Excel or SAS with little or no change control and access controls present
B	Model overrides that are applied to model results	0.3	No overrides are applied to model results	Some overrides are applied to model results		A large number of overrides are applied to model results	
C	A process is in place to notify the model development team of changes to product design and input data changes that may affect the in-use model	0.4	A formal process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model		An imperfect process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model		No process is in place to notify the model development team of changes to product design and data changes that may impact the in-use model
5	**Reporting and monitoring**	**0.1**					
	Minimum requirements for model monitoring and reporting, including the frequency thereof	0.5	The minimum requirements for model monitoring reporting and frequency are defined and		The minimum requirements for model monitoring reporting and frequency are defined, but not		The minimum requirements for model monitoring reporting and frequency are

			documented		formally documented		not defined
	The use of monitoring results	0.5	There is demonstrable evidence that monitoring results are used to amend in-use models if needed		There is no demonstrable evidence that monitoring results are used to amend in-use models if needed		There is no monitoring results that can be used to amend in-use models if needed

Results

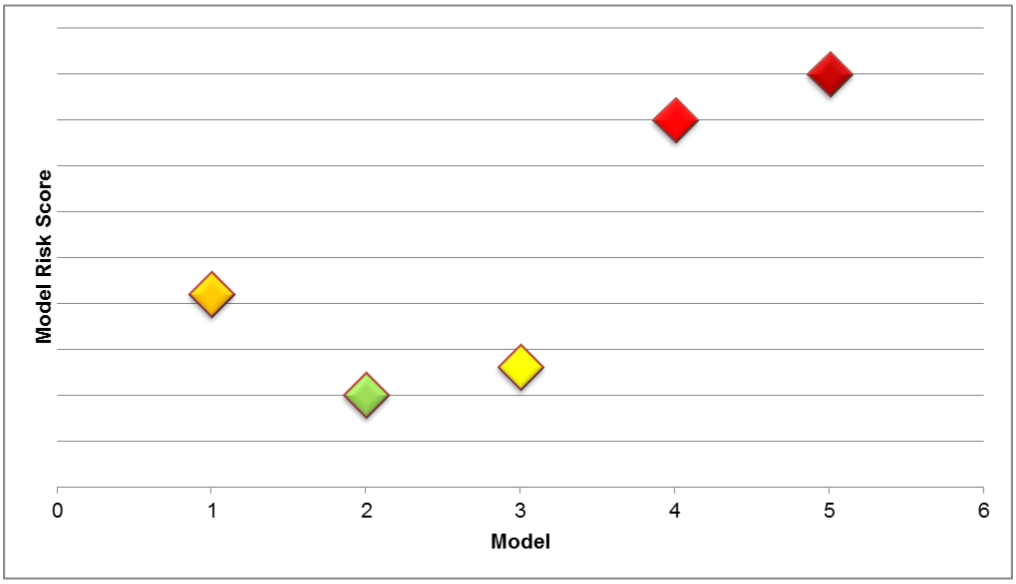
Each model is then rated according to the scorecard. The score is then interpreted according to the five-point scale:

Table 6-9: Model risk scorecard results scale

Score range	Risk Scale	Colour
1	“very low risk”	
1.1 to 2	“low risk”	
2.1 to 3	“moderate risk”	
3.1 to 4	“high risk”	
4.1 to 5	“very high risk”	

Once each model has a score it can be used to gauge the model risk relative to other models. When the rating is re-performed the model can then be compared to its previous results. A fictional example can be seen below:

Figure 6-6: **Model risk scorecard example**



From Figure 6-6 it can be seen that five models were assessed and they range from very low risk to very high risk.

Summary and future improvements

The scorecard can be further enhanced by including model materiality. This can be done by adding a third dimension to the scorecard results, where currently it is model by model risk score and then it would also include the materiality of each model. This can be included by changing the size of the shapes in Figure 6-6 to represent the materiality of the model.

The model risk scorecard proposed provides a practical way to establish the relative level of model risk based on model risk management principles within a bank.

6.7 Conclusion

This chapter is dedicated to model risk quantification of different model risk category types defined in Chapter 3. Even though it can be argued that model risk quantification can lead to an additional unknown quantum of model risk, it can be a value-adding tool in model risk management.

A method was proposed to quantify Type 1 model risk which deals with parameter uncertainty. The proposed method of quantification can show how severe Type 1 model risk can be given a certain set of assumptions such as sample size, EVI, quantile, and using VaR or ES. The result of this quantification can inform model users how risky a model can be and this in turn can result in better model risk management for models that can have more severe model risk. A simulation method was used and results indicate that Type 1 model errors can range from immaterial to more than 100% (depending on the risk measure used) based on the parameters of this simulation study, as well as the definition of model error. The contributing factors to model risk were identified, namely i) the relative model error increases when using ES instead of VaR, ii) the relative model error increases with very high values of EVI (heaviness of tail), iii) the relative model error increases with decreasing values of sample size, and iv) the relative model error increases with increasing quantiles.

A method to quantify Type 2 model risk (model misspecification) was not detailed in this document, but it was proposed as future research and it was also proposed to follow a similar method to that of the simulation method used for Type 2 model risk.

Type 3 model risk is a challenging model risk category to quantify. It can be argued that it is not always possible to identify a change in the underlying real-world phenomenon. It was proposed to use some of the mitigating methods detailed in Chapter 4 to assist with the identification and mitigation of this model risk type. Model monitoring can be a key tool to use in order to identify when the model is no longer optimal due to a change in the dynamics of the underlying real-world phenomenon. It was proposed to investigate the quantification of this model risk type as future research.

The quantification method proposed for Type 4 model risk involves the use of a LDA model for the modelling of model risk losses. The method is briefly described and it is also proposed that a model validation function be responsible for the gathering of model risk loss data relating to Type 4 model risk or that scenario data is used. A more detailed description of this method is proposed for future research.

Lastly a subjective scorecard approach was proposed to measure model risk within an organisation. This method is practical and can assist with the management of model risk. This scorecard approach aligns with the model risk maturity assessment method proposed in Chapter 5.

Even though model risk measurement is not as mature as for other risk types, it is demonstrated in this chapter that there are useful methods available to explore further to quantify model risk.

CHAPTER 7 CONCLUSION

Model risk has increasingly become more topical in the last two decades. There are many reasons for the rise in awareness and interest in the topic. Among these reasons are the financial-industry regulators requiring advanced risk models, the increase in complexity of models used and the wide range of risk types that models impact. The financial crisis also added to the topic becoming more prevalent in recent times. Models are used to represent intricate real-world phenomena. The model risk related to these models is intentionally created through the simplified version of reality. Therefore it is almost certain that the model will contain a degree of model risk.

A couple of key questions arise from the topic of model risk. The first question relates to what is model risk? The answer here is not that simple. It can be seen from this paper that the definition of model risk is not standardised and ranges from narrow, where only parameter- and model misspecification is taken into account, to very broad where the incorrect use of an otherwise fit-for-purpose model is included. The problem with the divergence in definitions is that it adds an additional layer of uncertainty into the already unknown model risk. In Chapter 3 a solution to this issue is proposed by categorising model risk into four types based on the available literature. The four types are: i) model parameter uncertainty, ii) model misspecification, iii) the change in dynamics of the real-world phenomena, and iv) the incorrect model implementation, misinterpretation of model output and other errors. Categorising the different model risk types will enable a standardised way of managing and measuring model risk. This will also aid in the comparability across different institutions, as they will then categorise model risk in the same way.

The second question regarding model risk is the “why?”. Why does model risk matter and why should anyone care? In Chapter 2 different publicly available model risk incidents are discussed. Model risk can have major impacts, such as financial losses, reputational damage, regulatory fines as well non-optimal decision making. AXA Rosenberg was fined $242 million in 2011 due to a spreadsheet error which overestimated client investment losses. It is clear that model risk can have severely adverse impacts on financial organisations. This then leads to the next question of “how?”.

The next important question regarding model risk is, how to manage model risk? Model risk management has been evolving and more and more regulatory guidance have been seen in the last decade. This was mainly due to the rise in complexity of models as well as the response from regulators after the financial crisis. In Chapter 5 a practical and repeatable model risk maturity assessment method is proposed. Model risk management is dissected into a five-step process which consist of: i) data quality, extraction and transformation, ii) model definition, development and documentation, iii) model validation and approval, iv) model implementation, change control and usage, and v) reporting and monitoring. Each of these steps is then used to assess the maturity of model risk management by using a set of assessment items under each process step. The method can also be used to identify a target maturity level which an organisation wants to achieve and then setting up an action plan in order to do so. The method is practical and repeatable and assists with achieving leading edge model risk management.

The next question also relates to "how?", but refers to how model risk can be mitigated. In Chapter 4 a non-exhaustive list of model risk mitigating measures are detailed. The list can be grouped together in four groups namely i) governance, which includes the development and validation standards, ii) controls, which include change control and ongoing research, iii) testing, which include data quality tests, and iv) monitoring and assurance, which include model monitoring and model audit. The main reason for the extensive suite of model risk mitigating measures is that is it can be acknowledged that it is unlikely for model risk to be completely eliminated. This again relates back to the fact that models are simplified representations of intricate real-world phenomena.

The final key question relates to the severity of model risk. Model risk measurement is an evolving field in the financial modelling world. Quantification methods have not yet reached the same maturity as some of the other main risk types, such as credit and market risk. It can be argued that measuring model risk can lead to an additional quantum of model risk, i.e. the model risk of model risk-models. It can also be argued that measuring the model risk can assist with improved model risk management. In Chapter 6 the Type 1 model risk (parameter uncertainty) of a parametric distribution is estimated using a simulation study. Different parametric distributions are used to illustrate the impact of the heaviness of the distribution's tail, the sample size, the quantile and using VaR or ES on the quantum of model risk. The contributing factors to model risk are then identified as namely: i) the relative model error increases when using ES instead of VaR, ii) the relative model error increases with very high values of EVI (heaviness of tail), iii) the relative model error increases with decreasing values of sample size, and iv) the relative model error increases with increasing quantiles. It can also be seen that the relative model error for this set of results range from immaterial to over 100% given the sample size, distribution, quantile and depending on which measure was used as a proxy for model error (25^{th} percentile, 75^{th} percentile, the range, the IQR, the median etc.). The quantification for the other model risk types are proposed as future research.

A qualitative scorecard approach is also proposed as a way to give an indication of how risky a model is relative to the rest of the models within scope. The assessment is also based on the five-step process of model risk management explored in Chapter 5.

It is also highlighted that some regulators globally require banks to estimate model risk capital for Pillar 2 capital purposes. The quantification methods proposed in this research can possibly assist with meeting the expectations of regulators.

Although model risk has been present for an extended period of time, it is clear that it is evolving along with the way in which models are developed and managed. The level of sophistication in which model risk is managed and measured is increasing and will most probably continue to grow.

While model risk evolves into a vital risk category it is important to continually improve on the level of model risk management maturity, increase the scope of model risk quantification of different types of model risk and continue to mitigate model risk effectively, because as George Box said in 1976: *"All models are wrong, but some are useful"* (Box, 1976).

BIBLIOGRAPHY

ALEXANDER, C. & SARABIA, J. M. 2012. Quantile uncertainty and value-at-risk model risk. *Risk Analysis*. **32**(8), pp.1293-1308.

ANON. 2011. *King's Fund apologise for "error" in figures on health spending in Whales*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <http://leftfootforward.org/2011/05/kings-fund-apologise-for-error-on-health-spending-in-wales/>

BARCLAYS. 2013. *Barclays Africa*. [online]. [Accessed 30 Oct 2017]. Available from World Wide Web: <https://www.barclaysafrica.com/content/dam/barclays-africa/bagl/pdf/presentation/annual/full-year-2012-results-investor-presentation-speaker-notes.pdf>

BBC. 2013. *BBC News*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <http://www.bbc.com/news/magazine-22223190 >

BCBS. 1996. *Amendment to the capital accord to incorporate market risks*. Basel: Bank of International Settlements.

BCBS. 2006. *Basel II: International convergence of capital measurement and capital standards: a revised framework*. Basel: Bank of International Settlements.

BCBS. 2009a. *Revisions to the Basel II market risk framework*. Basel: Bank of International Settlements.

BCBS. 2009b. *Supervisory guidance for assessing banks' financial instrument fair value practices*. Basel: Bank of International Settlements.

BCBS. 2011. *Basel III: A global regulatory framework for more resilient banks and banking systems*. Basel: Bank of International Settlements.

BCBS. 2013a. *The regulatory framework: balancing risk sensitivity, simplicity and comparability*. Basel: Bank of International Settlements.

BCBS. 2013b. *Principles for effective risk data aggregation and risk reporting*. Basel: Bank of International Settlements.

BCBS. 2015. *Revisions to the Standardised Approach for credit risk*. Basel: Bank of International Settlements.

BCBS. 2016a. *Minimum capital requirements for market risk*. Basel: Bank of International Settlements.

BCBS. 2016b. *Standardised Measurment Approach for operational risk*. Basel: Bank of International Settlements.

BCBS. 2016c. *Regulatory Consistency Assessment Programme (RCAP): Handbook for jurisdictional assessments*. Basel: Bank of International Settlements.

BEIRLANT, J., GOEGEBEUR, Y., SEGERS, J., TEUGELS, J., DE WAAL, D., & FERRO, C. 2004. *Statistics of extremes: theory and application*. West Sussex: Wiley.

BERKOWITZ, J. & O'BRIEN, J. 2002. How accurate are value-at-risk models at commercial banks? *Journal of Finance*. **57**(3), pp.1093-1111.

BERTRAM, P., SIBBERTSEN, P., & STAHL, G. 2015. The impact of model risk on capital reserves: a quantitative analysis. *Journal of Risk*. **17**(5), pp.1-29.

BIGNOZZI, V. & TSANAKAS, A. 2016. Model uncertainty in risk capital measurement. *Journal of Risk*. **18**(3), pp.1-24.

BORODOVSKY, L. & LORE, M. 2000. *Professional's Handbook of Financial Risk Management*. Oxford: Elsevier Butterworth Heineman.

BOUCHER, C., DANIELSSON, J., & KOUONTCHOU, P. S. 2014. Risk models-at-risk. *Journal of Banking and Finance*. **44**, pp.72-92.

BOX, G. E. P. 1976. Science and statistics. *Journal of the American Statistical Association*. **71**, pp.791-799.

CHERNOBAI, A., TACHEV, S. T., & J., Fabozzi. F. 2007. *Operational Risk: A guide to Basel II capital requirements, models and analysis*. West-Sussex: Wiley.

CHRISTOFFERSEN, P. & GONÇALVES, P. 2005. Estimation risk in financial risk management. *The Journal of Risk*. **7**(3), pp.1-28.

DE JONGH, P. J., LARNEY, J., MARÉ, E., VAN VUUREN, G., & VERSTER, T. 2017a. A proposed best practice model validation framework for banks. *South African Journal of Economic and management Sciences*. **20**(1), pp.1-15.

DE JONGH, P.J., VERSTER, T., REYNOLDS, E., JOUBERT, M., & RAUBENHEIMER, H. 2017b. A critical review of the Basel margin of conservatism requirement in a retail credit context. *International Business & Economics Research Journal*. **16**(4), pp.257-274.

DERMAN, E. 1996. Model risk. *Risk*. **9**(0), pp.34-37.

EBA. 2013. *On access to the activity of credit institutions and the prudential supervision of credit institutions and investment*. Brussels: The European Parliament and the Council.

EBA. 2014. *Guidelines on common procedures and methodologies for the supervisory review and evaluation process (SREP)*. London: European System of Financial Supervision.

EBA. 2015. *Technical Standards on prudent valuation under Article 105(14) of Regulation (EU) No 575/2013 (Capital Requirements Regulation — CRR)*. London: European System of Financial Supervision.

EBA. 2016. *Guidance on PD estimation, LGD estimation and the treatment of defaulted exposures*. London: European System of Financial Supervision.

EMBRECTHS, P., WANG, P., & WANG, R. 2015. Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance Stochastics*. **19**, pp.763-790.

FED. 2003. *Trading and capital-markets activities manual*. Washington: U.S. Presidency.

FED. 2013. *Applying the requirement for conservatism to the parameters in the advanced approaches*. Washington: U.S. Presidency.

FED. 2014. *Supervisory guidance for data, modelling, and model risk management under the operational risk advanced measurement approaches*. Washington: U.S. Presidency.

FED. 2015. *Supervisory Assessment of Capital Planning and Positions for LISCC Firms and Large and Complex Firms*. Washington: U.S. Presidency.

FED & OCC. 2011. *Supervisory guidance on model risk management SR 11-7*. Washington: U.S. Presidency and U.S. Department of the Treasury.

FIRSTRAND. 2016. *FirstRand*. [online]. [Accessed 13 June 2017]. Available from World Wide Web:
<www.firstrand.co.za/InvestorCentre/basel%20II%20Pillar%203%20Disclosure/FirstRand%20Basel%20Pillar%203%20disclosure%20-%20June%202016.pdf>

FRACHOT, A., GEORGES, P., & RONCALLI, T. 2001. *Loss Distribution Approach for operational risk*. Lyon.

GIBSON, R., L'HABITAND, F. S., PISTRE, N., & TALAY, D. 1999. Interest rate model risk: an overview. *The Journal of Risk*. **10**(5), pp.37-62.

GLASSERMAN, P. & XU, X. 2014. Robust risk measurement and model risk. *Quantitative Finance*. **14**(1), pp.29-58.

GODFREY, K. 1995. Computing error at Fidelity's Magellan Fund. *Risk Digest*. **16**(72).

HEINEMAN, B. W. 2013. *The JP Morgan "Whale" report and the ghosts of the financial crisis*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <https://hbr.org/2013/01/the-jp-morgan-whale-report-and>

HILL, J. R. 2011. The complete history of model risk- abridged. *In*: *Measuring and controlling model risk seminar*. London, pp.1-97.

IFOA. 2015. *Model risk daring to open up the black box*.

JORION, P. 1996. Risk: measuring the risk in value-at-risk. *Financial Analysts Journal*. **52**(6), pp.47-56.

KELLNER, R., RÖSCH, D., & SCHEULE, H. 2016. The role of model risk in extreme value theory for capital adequacy. *Journal of Risk*. **18**(6), pp.39-70.

KERKHOF, J., MELENBERG, B., & SCHUMACHER, H. 2010. Model risk and capital reserves. *Journal of Banking and Finance*. **34**(1), pp.267-279.

KPMG. 2013. *KPMG*. [online]. [Accessed 17 November 2016]. Available from World Wide Web: <https://www.kpmg-institutes.com/institutes/advisory-institute/events/2014/02/results-of-2013-model-risk-mgmt-survey.html>

MALAN, G. R. H. 2017. *An overview of emerging best-practices in model risk management*. North-West University, Potchefstroom.

MANAGEMENT SOLUTIONS. 2014. *Model risk management quantitative and qualitative aspect*.

MIGNOLA, G. & UGOCCIONI, R. 2006. Sources of uncertainty in modeling operational risk. *The Journal of Operational Risk*. **1**(2), pp.33-50.

MIGNOLA, G., UGOCCIONI, R., & COPE, E. 2016. Comments on the Basel Committee on Banking Supervision proposal for a new standardized approach for operational risk. *The Journal of Operational Risk*. **11**(3), pp.51-69.

MORINI, M. 2012. *Understanding and managing model risk: a practical guide for quants, traders and validators*. Chichester: Wiley.

NORTH AMERICAN CRO COUNCIL. 2012. *Model validation principles applied to risk and capital models in the insurance industry*. Delaware: North American CRO Council.

NORTH AMERICAN CRO COUNCIL. 2016. *Model risk practices and principles*. Delaware: North American CRO Council.

NUMERIX. 2013. *Mastering model risk: assessment, regulation and best practices*. Numerix.

OCC. 2000. *Model Validation OCC 2000-16*. Washington: U.S. Department of the Treasury.

OLIVER WYMAN. 2015. *Model risk and the great financial crisis: the rise of modern model risk management*.

ORX. 2016. *How do banks manage model risk in 2016?* Operational Riskdata eXchange.

PANMAN, K. 2015. *A comparison of VaR estimates in risk management*. Potchefstroom.

PWC. 2013. *The evolution of model risk management*. URL: fic.wharton.upenn.edu/fic/papers/15/15-01.pdf.

QUELL, P. & MEYER, C. 2016. *Risk Model Validation*. London: Risk Books.

SANTANDER. 2016. *Grupo Santander*. [online]. [Accessed 13 June 2017]. Available from World Wide Web: <https://www.santander.com/csgs/StaticBS?blobcol=urldata&blobheadername1=content-type&blobheadername2=Content-Disposition&blobheadername3=appID&blobheadervalue1=application%2Fpdf&blobheadervalue2=inline%3Bfilename%3D404%5C36%5CPillar+III+2016.pdf&blobheade>

SARB. 2015. *Directive 4/2015: Amendments to the Regulations relating to Banks, and matters re-lated thereto.* [online]. [Accessed 10 February 2018]. Available from World Wide Web: <https://www.resbank.co.za/Lists/News%20and%20Publications/Attachments/6664/D4%20of%202015.pdf>

SEC. 2011. *SEC charges AXA Rosenberg entities for concealing error in quantitative investment model*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <https://www.sec.gov/news/press/2011/2011-37.htm>

SIBBERTSEN, P., STAHL, G., & LUEDTKE, C. 2008. Measuring model risk. *Journal of Risk Model Validation*. **2**(4), pp.65-81.

SPRINGER, K. 2012. *Top 10 biggest trading losses in history by Time Magazine*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <http://newsfeed.time.com/2012/05/11/top-10-biggest-trading-losses-in-history/slide/morgan-stanley-9b/>

TALAY, D. & ZHANG, Z. 2002. Worst case model risk management. *Finance and Stochastics*. **6**(4), pp.517-537.

VAN BILJON, L. & HAASBROEK, L. J. 2017a. A practical maturity assessment method for model risk management in banks. *The Journal of Risk Model Validation*. **11**(4), pp.1-17.

VAN BILJON, L. & HAASBROEK, L.J. 2017b. A practical maturity assessment method for model risk management in banks. *In*: *SASA*. Bloemfontein.

VAN BILJON, L. & PANMAN, K. 2016. On the quantification of model risk in statistical models of empirical data. *In*: *SASA*. Cape Town.

WAILGYM. 2007. *Eight of the worst spreadsheet blunders*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <http://www.cio.com/arti-cle/2438188/enterprise-software/eight-of-the-worst-spreadsheet-blunders.html>

WEST, G. 2004. *Calibration of the SABR model in illiquid markets*. [online]. [Accessed 14 February 2017]. Available from World Wide Web: <http://www.riskworx.com/insights/sabr/sabrilliquid.pdf>