Developing Credit Risk Models Using SAS® Enterprise Miner™ and SAS/STAT®

Theory and Applications

Iain L. J. Brown, PhD

Contents

Chapter 1 Introduction

1.1 Book Overview

This book aims to define the concepts underpinning credit risk modeling and to show how these concepts can be formulated with practical examples using SAS software. Each chapter tackles a different problem encountered by practitioners working or looking to work in the field of credit risk and give a step-by-step approach to leverage the power of the SAS Analytics suite of software to solve these issues.

This chapter begins by giving an overview of what credit risk modeling entails, explaining the concepts and terms that one would typically come across working in this area. We then go on to scrutinize the current regulatory environment, highlighting the key reporting parameters that need to be estimated by financial institutions subject to the Basel capital requirements. Finally, we discuss the SAS analytics software used for the analysis part of this book.

The remaining chapters are structured as follows:

Chapter 2 covers the area of sampling and data pre-processing. This chapter defines and contextualizes issues such as variable selection, missing values, and outlier detection within the area of credit risk modeling, and gives practical applications of how these issues can be solved.

Chapter 3 details the theory and practical aspects behind the creation of Probability of Default (PD) models. This focuses on standard and novel modeling techniques, shows how each of these can be used in the estimation of PD, and demonstrates the full development of an application and behavioral scorecard using SAS Enterprise Miner.

Chapter 4 focuses on the development of Loss Given Default (LGD) models and the considerations with regard to the distribution of LGD that have to be made for modeling this parameter. A variety of modeling approaches are discussed and compared in a case study in order to show how improvements over the traditional industry approach of linear regression can be made.

Chapter 5 defines the concept of Exposure at Default (EAD) and how this parameter is formulated and estimated. A full model development process is shown through practical examples. The aim of this chapter is to fully explore the implications of model choice, input variables, and how best to estimate EAD.

Chapter 6 defines and explains the concepts of stress testing under the three pillars of the Basel Capital Accord and what this entails for financial institutions.

Chapter 7 focuses on how model reports can be generated from the procedures and methodologies created throughout this book. This chapter covers the key reporting outputs required within the regulatory framework and shows through SAS Model Manager and example code how these outputs can be created.

By the conclusion of this book, readers will have a comprehensive guide to developing credit risk models both from a theoretical and practical perspective. We also aim to show how analysts can create and implement credit risk models using example code and projects in SAS.

1.2 Overview of Credit Risk Modeling

With cyclical financial instabilities in the credit markets, the area of credit risk modeling has become ever more important, leading to the need for more accurate and robust models. Since the introduction of the Basel II Capital Accord (Basel Committee on Banking Supervision, 2004) over a decade ago, qualifying financial institutions have been able to derive their own internal credit risk models under the advanced internal ratings-based approach (A-IRB) without relying on regulator's fixed estimates.

The Basel II Capital Accord prescribes the minimum amount of regulatory capital an institution must hold so as to provide a safety cushion against unexpected losses. Under the advanced internal ratings-based approach (A-IRB), the accord allows financial institutions to build risk models for three key risk parameters: Probability of Default (PD), Loss Given Default (LGD), and Exposure at Default (EAD). PD is defined as the likelihood that a loan will not be repaid and will therefore fall into default. LGD is the estimated economic loss, expressed as a percentage of exposure, which will be incurred if an obligor goes into default. EAD is a measure of the monetary exposure should an obligor go into default. These topics will be explained in more detail in the next section.

With the arrival of Basel III and as a response to the latest financial crisis, the objective to strengthen global capital standards has been reinstated. A key focus here is the reduction in reliance on external ratings by the financial institutions, as well as a greater focus on stress testing. Although changes are inevitable, a key point worth noting is that with Basel III there is no major impact on underlying credit risk models. Hence the significance in creating these robust risk models continues to be of paramount importance.

In this book, we use theory and practical applications to show how these underlying credit risk models can be constructed and implemented through the use of SAS (in particular, SAS Enterprise Miner and SAS/STAT). To achieve this, we present a comprehensive guide to the classification and regression techniques needed to develop models for the prediction of all three components of expected loss: PD, LGD and EAD. The reason why these particular topics have been chosen is due in part to the increased scrutiny on the financial sector and the pressure placed on them by the financial regulators to move to the advanced internal ratings-based approach. The financial sector is therefore looking for the best possible models to determine their minimum capital requirements through the estimation of PD, LGD and EAD.

This introduction chapter is structured as follows. In the next section, we give an overview of the current regulatory environment, with emphasis on its implications to credit risk modeling. In this section, we explain the three key components of the minimum capital requirements: PD, LGD and EAD. Finally, we discuss the SAS software used in this book to support the practical applications of the concepts covered.

1.3 Regulatory Environment

The banking/financial sector is one of the most closely scrutinized and regulated industries and, as such, is subject to stringent controls. The reason for this is that banks can only lend out money in the form of loans if depositors trust that the bank and the banking system is stable enough and their money will be there when they require to withdraw it. However, in order for the banking sector to provide personal loans, credit cards, and mortgages, they must leverage depositors' savings, meaning that only with this trust can they continue to function. It is imperative, therefore, to prevent a loss of confidence and distrust in the banking sector from occurring, as it can have serious implications to the wider economy as a whole.

The job of the regulatory bodies is to contribute to ensuring the necessary trust and stability by limiting the level of risk that banks are allowed to take. In order for this to work effectively, the maximum risk level banks can take needs to be set in relation to the bank's own capital. From the bank's perspective, the high cost of acquiring and holding capital makes it prohibitive and unfeasible to have it fully cover all of a bank's risks. As a compromise, the major regulatory body of the banking industry, the Basel Committee on Banking Supervision, proposed guidelines in 1988 whereby a solvability coefficient of eight percent was introduced. In other words, the total assets, weighted for their risk, must not exceed eight percent of the bank's own capital (SAS Institute, 2002).

The figure of eight percent assigned by the Basel Committee was somewhat arbitrary, and as such, this has been subject to much debate since the conception of the idea. After the introduction of the Basel I Accord, more than one hundred countries worldwide adopted the guidelines, marking a major milestone in the history of global banking regulation. However, a number of the accord's inadequacies, in particular with regard to the way that credit risk was measured, became apparent over time (SAS Institute, 2002). To account for these issues, a revised accord, Basel II, was conceived. The aim of the Basel II Capital Accord was to further strengthen the financial sector through a three pillar approach. The following sections detail the current state of the regulatory environment and the constraints put upon financial institutions.

1.3.1 Minimum Capital Requirements

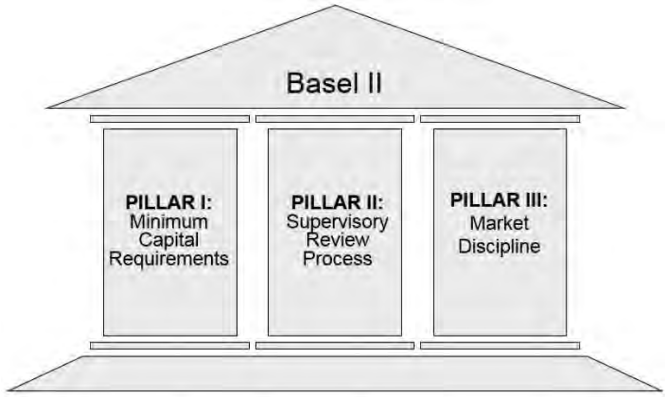
The Basel Capital Accord (Basel Committee on Banking Supervision, 2001a) prescribes the minimum amount of regulatory capital an institution must hold so as to provide a safety cushion against unexpected losses. The Accord is comprised of three pillars, as illustrated by Figure 1.1:

Pillar 1: Minimum Capital Requirements

Pillar 2: Supervisory Review Process

Pillar 3: Market Discipline (and Public Disclosure)

Figure 1.1: Pillars of the Basel Capital Accord



Pillar 1 aligns the minimum capital requirements to a bank's actual risk of economic loss. Various approaches to calculating this are prescribed in the Accord (including more risk-sensitive standardized and internal ratings-based approaches) which will be described in more detail and are of the main focus of this text. Pillar 2 refers to supervisors evaluating the activities and risk profiles of banks to determine whether they should hold higher levels of capital than those prescribed by Pillar 1, and offers guidelines for the supervisory review process, including the approval of internal rating systems. Pillar 3 leverages the ability of market discipline to motivate prudent management by enhancing the degree of transparency in banks' public disclosure (Basel, 2004).

Pillar 1 of the Basel II Capital Accord entitles banks to compute their credit risk capital in either of two ways:

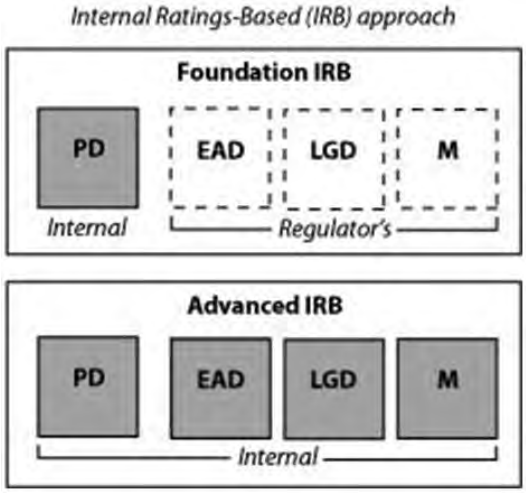
1. Standardized Approach
2. Internal Ratings-Based (IRB) Approach
   a. Foundation Approach
   b. Advanced Approach

Under the standardized approach, banks are required to use ratings from external credit rating agencies to quantify required capital. The main purpose and strategy of the Basel committee is to offer capital incentives to banks that move from a supervisory approach to a best-practice advanced internal ratings-based approach. The two versions of the internal ratings-based (IRB) approach permit banks to develop and use their own internal risk ratings, to varying degrees. The IRB approach is based on the following four key parameters:

1. Probability of Default (PD): the likelihood that a loan will not be repaid and will therefore fall into default in the next 12 months;
2. Loss Given Default (LGD): the estimated economic loss, expressed as a percentage of exposure, which will be incurred if an obligor goes into default - in other words, LGD equals: 1 minus the recovery rate;
3. Exposure At Default (EAD): a measure of the monetary exposure should an obligor go into default;
4. Maturity (M): is the length of time to the final payment date of a loan or other financial instrument.

The internal ratings-based approach requires financial institutions to estimate values for PD, LGD, and EAD for their various portfolios. Two IRB options are available to financial institutions: a foundation approach and an advanced approach (Figure 1.2) (Basel Committee on Banking Supervision, 2001a).

Figure 1.2: Illustration of Foundation and Advanced Internal Ratings-Based (IRB) approach



The difference between these two approaches is the degree to which the four parameters can be measured internally. For the foundation approach, only PD may be calculated internally, subject to supervisory review (Pillar 2). The values for LGD and EAD are fixed and based on supervisory values. For the final parameter, M, a single average maturity of 2.5 years is assumed for the portfolio. In the advanced IRB approach, all four parameters are to be calculated by the bank and are subject to supervisory review (Schuermann, 2004).

Under the A-IRB, financial institutions are also recommended to estimate a ”Downturn LGD”, which ‘cannot be less than the long-run default-weighted average LGD calculated based on the average economic loss of all observed defaults with the data source for that type of facility’ (Basel, 2004).

1.3.2 Expected Loss

Financial institutions expect a certain number of the loans they make to go into default; however they cannot identify in advance which loans will default. To account for this risk, a value for expected loss is priced into the products they offer. Expected Loss (EL) can be defined as the expected means loss over a 12 month period from which a basic premium rate is formulated. Regulatory controllers assume organizations will cover EL through loan loss provisions. Consumers experience this provisioning of expected loss in the form of the interest rates organizations charge on their loan products.

To calculate this value, the PD of an entity is multiplied by the estimated LGD and the current exposure if the entity were to go into default.

From the parameters, PD, LGD and EAD, expected loss (EL) can be derived as follows:

$$EL = PD \times LGD \times EAD \quad (1.1)$$

For example, if PD = 2%, LGD = 40% and EAD = $10,000, then EL would equal $80. Expected Loss can also be measured as a percentage of EAD:

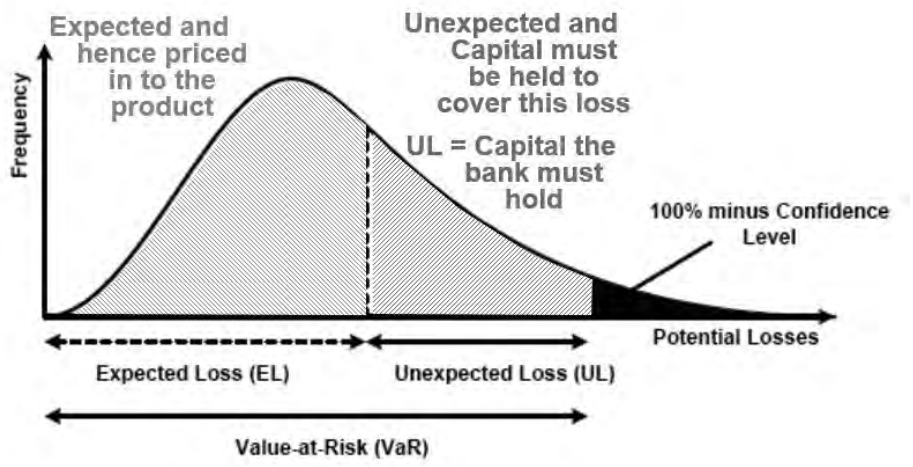
$$EL\% = PD \times LGD \quad (1.2)$$

In the previous example, expected loss as a percentage of EAD would be equal to $EL\% = 0.8\%$.

1.3.3 Unexpected Loss

Unexpected loss is defined as any loss on a financial product that was not expected by a financial organization and therefore not factored into the price of the product. The purpose of the Basel regulations is to force banks to retain capital to cover the entire amount of the Value-at-Risk (VaR), which is a combination of this unexpected loss plus the expected loss. Figure 1.3 highlights the Unexpected Loss, where UL is the difference between the Expected Loss and a 1 in 1000 chance level of loss.

Figure 1.3: Illustration of the Difference between Expected/Unexpected Loss and a 1 in 1000 Chance Level of Loss



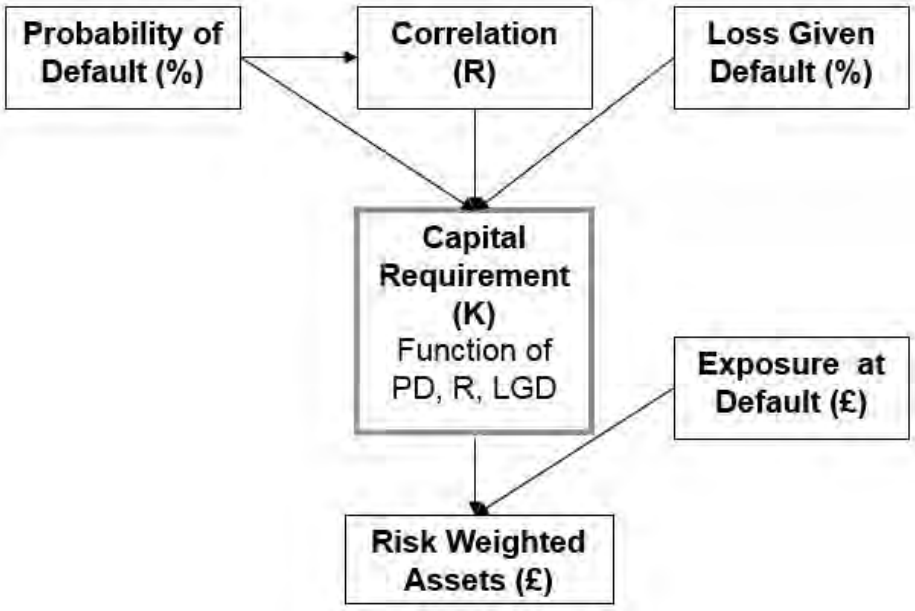
1.3.4 Risk Weighted Assets

Risk Weighted Assets (RWA) are the assets of the bank (money lent out to customers and businesses in the form of loans) accounted for by their riskiness. The RWA are a function of PD, LGD, EAD and M, where K is the capital requirement:

$$RWA = (12.5) \times K \times EAD \quad (1.3)$$

Under the Basel capital regulations, all banks must declare their RWA, hence the importance in estimating the three components, PD, LGD, and EAD, which go towards the formulation of RWA. The multiplication of the capital requirement (K) by 12.5 $\left(\frac{1}{12.5} = 0.08\right)$ is to ensure capital is no less than 8% of RWA. Figure 1.4 is a graphical representation of RWA and shows how each component feeds into the final RWA value.

Figure 1.4: Relationship between PD, LGD, EAD and RWA



The Capital Requirement (K) is defined as a function of PD, a correlation factor (R) and LGD

$$K = LGD \times \left(\phi\left(\sqrt{\frac{1}{1-R}}\phi^{-1}(PD) + \sqrt{\frac{R}{1-R}}\phi^{-1}(0.999) \right) - PD \right) \quad (1.4)$$

where ϕ denotes the normal cumulative distribution function and ϕ^{-1} denotes the inverse cumulative distribution function. The correlation factor (R) is determined based on the portfolio being assessed. For example, for revolving retail exposures (credit cards) not in default, the correlation factor is set to 4%. A full derivation of the capital requirement can be found in Basel Committee on Banking Supervision (2004).

In practice, how do estimations of PD, LGD and EAD impact the overall capital requirements? If we take PD as 0.03, LGD as 0.5, and EAD as \$10,000, then $K(0.03, 0.5) \times (10000) = \$34.37$. If an overestimate of 10% was made on PD, then the resulting capital required would then be $K(0.033, 0.5) \times (10000) = \$36.73$, requiring an increase of 6.9% in capital (\$2.36). However if an overestimate of 10% was made on LGD, then the resulting capital required would be $K(0.03, 0.55) \times (10000) = \$37.80$, requiring an increase of 10% in capital (\$3.43).

Because LGD and EAD enter the Risk Weight Function in a linear way, it is of crucial importance to have models that estimate LGD and EAD as accurately as possible, as LGD and EAD errors are more expensive than PD errors.

1.4 SAS Software Utilized

Throughout this book, examples and screenshots aid in the understanding and practical implementation of model development. The key tools used to achieve this are Base SAS programming with SAS/STAT procedures, as well as the point-and-click interfaces of SAS Enterprise Guide and SAS Enterprise Miner. For model report generation and performance monitoring, examples are drawn from SAS Model Manager. Base SAS is a comprehensive programming language used throughout multiple industries to manage and model data. SAS Enterprise Guide (Figure 1.5) is a powerful Microsoft Windows client application that provides a guided

mechanism to exploit the power of SAS and publish dynamic results throughout the organization through a point-and-click interface. SAS Enterprise Miner (Figure 1.6) is a powerful data mining tool for applying advanced modeling techniques to large volumes of data in order to achieve a greater understanding of the underlying data. SAS Model Manager (Figure 1.7) is a tool which encompasses the steps of creating, managing, deploying, monitoring, and operationalizing analytic models, ensuring the best model at the right time is in production.

Typically analysts utilize a variety of tools in their development and refinement of model building and data visualization. Through a step-by-step approach, we can identify which tool from the SAS toolbox is best suited for each task a modeler will encounter.

Figure 1.5: Enterprise Guide Interface

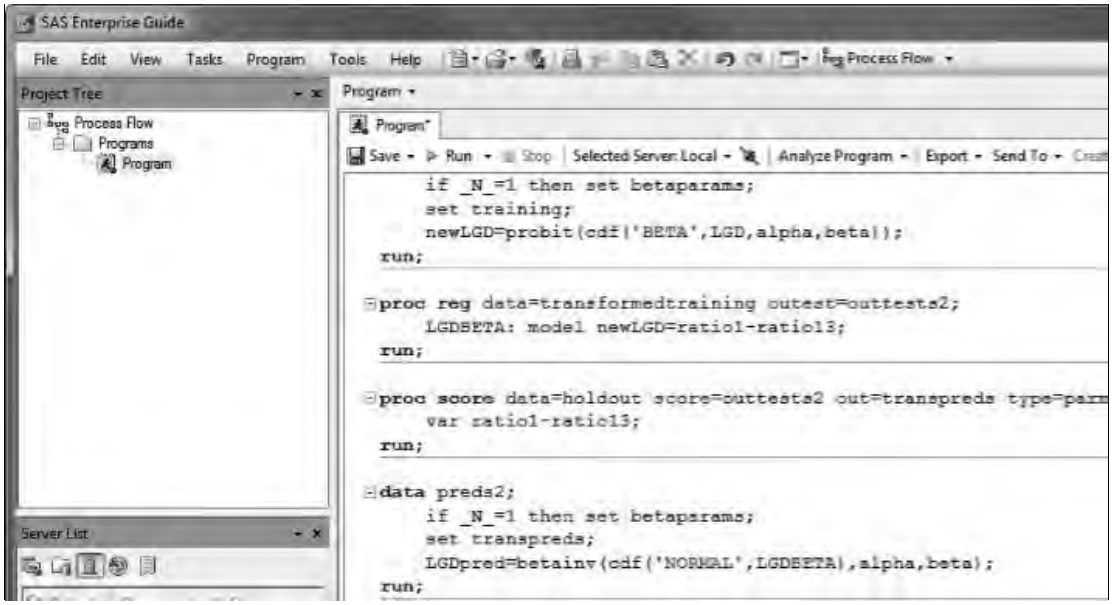


Figure 1.6: Enterprise Miner Interface

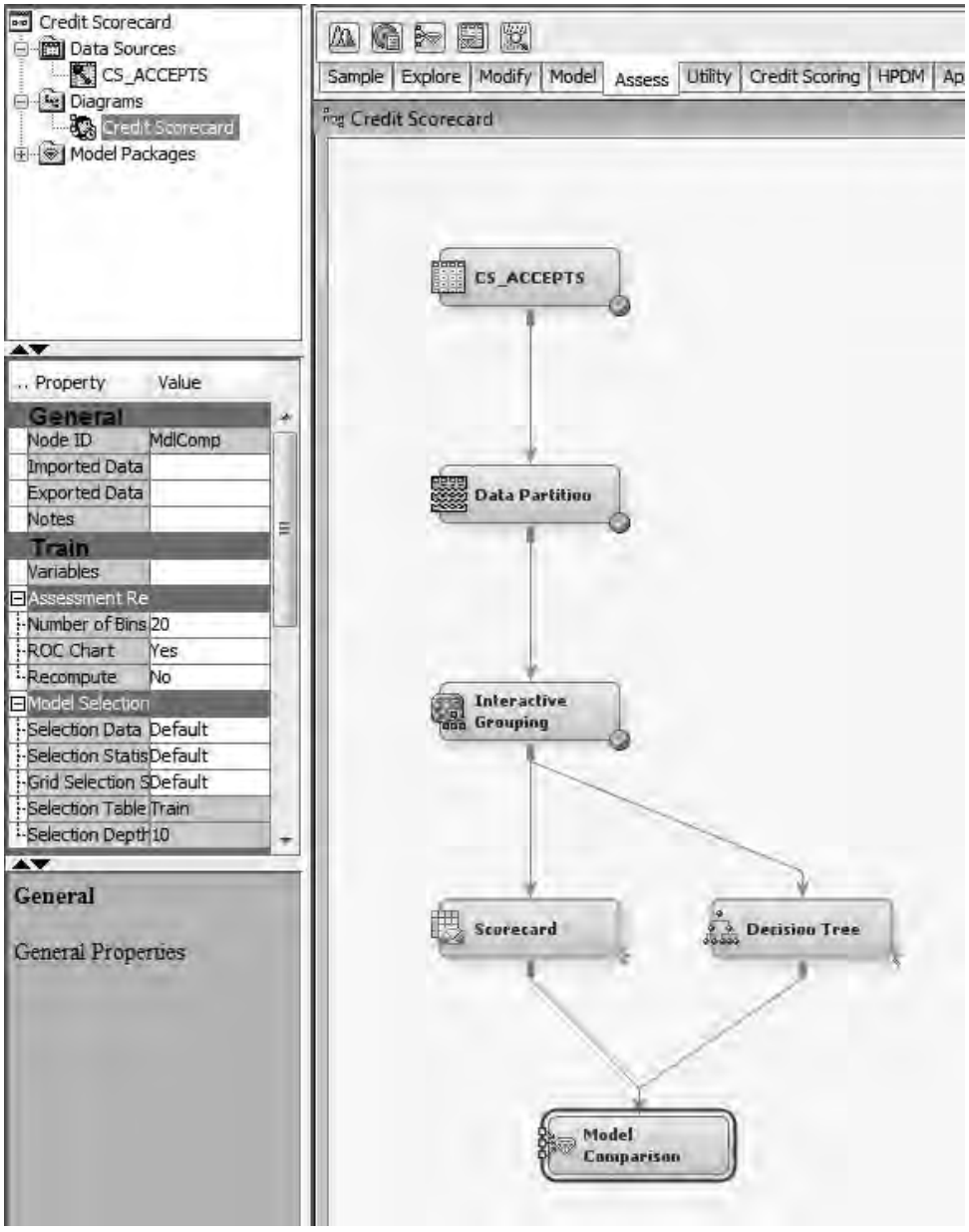
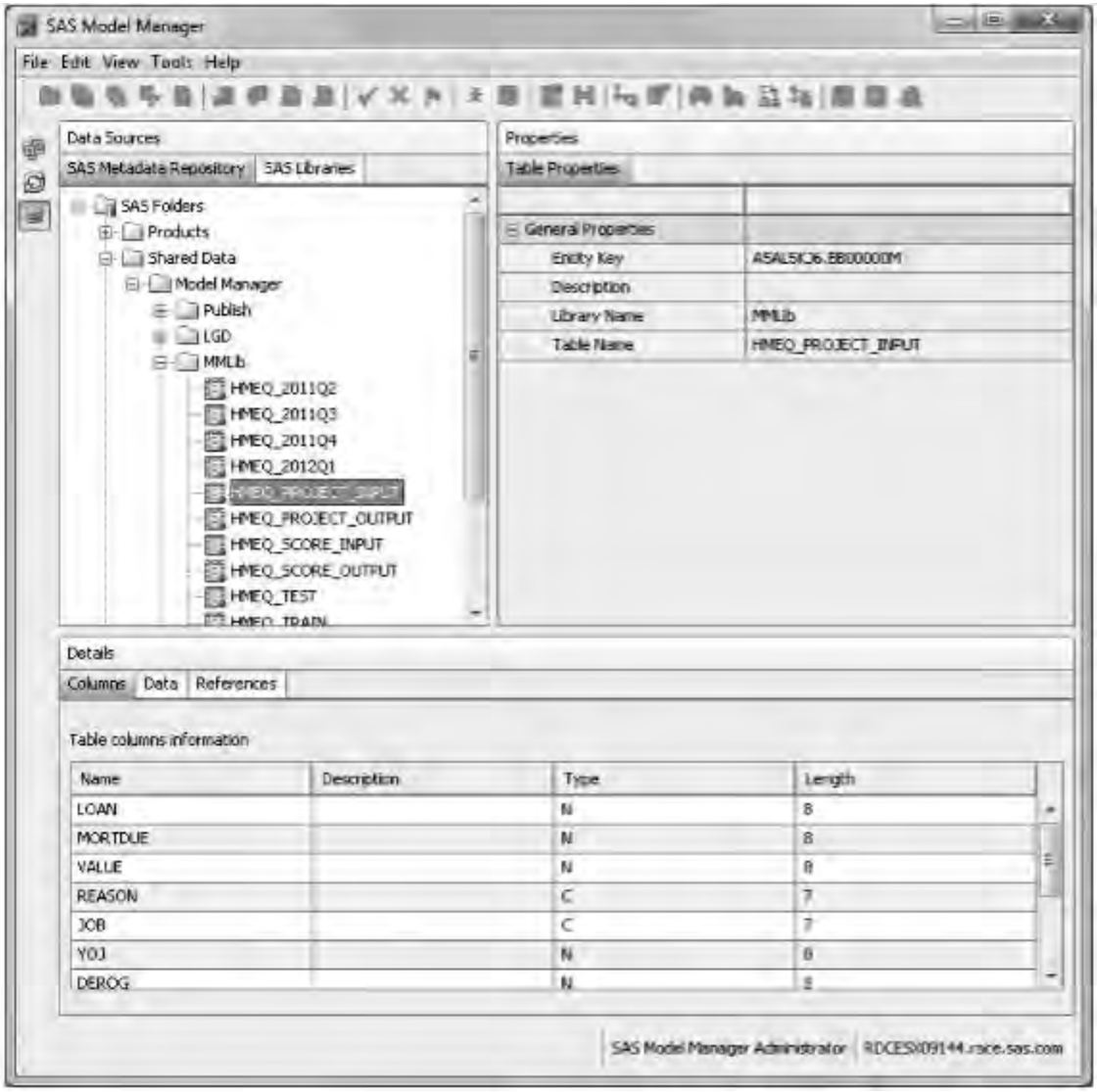


Figure 1.7: Model Manager Interface



1.5 Chapter Summary

This introductory chapter explores the key concepts that comprise credit risk modeling, and how this impacts financial institutions in the form of the regulatory environment. We have also looked at how regulations have evolved over time to better account for global risks and to fundamentally prevent financial institutions from over exposing themselves to difficult market factors. To summarize, Basel defines how financial institutions calculate:

- **Expected Loss (EL)** - the means loss over 12 months
- **Unexpected Loss (UL)** - the difference between the Expected Loss and a 1 in 1000 chance level of loss
- **Risk-Weighted Assets (RWA)** - the assets of the financial institution (money lent out to customers & businesses) accounted for by their riskiness
- How much **Capital** financial institutions hold to cover these losses

Three key parameters underpin the calculation of expected loss and risk weighted assets:

- **Probability of Default (PD)** - the likelihood that a loan will not be repaid and will therefore fall into default in the next 12 months
- **Loss Given Default (LGD)** - the estimated economic loss, expressed as a percentage of exposure, which will be incurred if an obligor goes into default - in other words, LGD equals: 1 minus the recovery rate
- **Exposure At Default (EAD)** - a measure of the monetary exposure should an obligor go into default

The purpose of these regulatory requirements is to strengthen the stability of the banking system by ensuring adequate provisions for loss are made.

We have also outlined the SAS technology which will be used through a step-by-step approach to apply the theoretical information given into practical examples.

In order for financial institutions to estimate these three key parameters that underpin the calculation of EL and RWA, they must begin by utilizing the correct data. Chapter 2 covers the area of sampling and data pre-processing. In this chapter, issues such as variable selection, missing values, and outlier detection are defined and contextualized within the area of credit risk modeling. Practical applications of how these issues can be solved are also given.

1.6 References and Further Reading

Basel Committee on Banking Supervision. 2001a. *The New Basel Capital Accord*. Jan. Available at: http://www.bis.org/publ/bcbsca03.pdf.

Basel Committee on Banking Supervision. 2004. *International Convergence of Capital Measurement and Capital Standards: a Revised Framework.* Bank for International Settlements.

SAS Institute. 2002. "Comply and Exceed: Credit Risk Management for Basel II and Beyond." A SAS White Paper.

Schuermann T. 2004. "What do we know about loss given default?" Working Paper No. 04-01, Wharton Financial Institutions Center, Feb.

Index

A

B

C

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

About The Author

Dr. Iain Brown is an Analytics Specialist Consultant at SAS, specializing in Credit Risk. Prior to joining SAS in 2011, he worked as a Credit Risk Analyst at a major UK retail bank where he built and validated PD, LGD, and EAD models using SAS software. He has spoken at a number of internationally renowned conferences and conventions and has published papers on the topic of credit risk modeling in the International Journal of Forecasting and the Journal of Expert Systems with Applications. In 2011, he won the SAS Student Ambassador award for his doctoral research, which recognizes and supports students who use SAS technologies in innovative ways to benefit their respective industries and fields of study.

Iain has a BBA in Business from the University of Kent, an MSc in Operational Research from the London School of Economics and Political Science (LSE), and a PhD in Credit Risk from the University of Southampton. Iain is also an active member of the Operational Research (OR) Society; in July 2014, he was awarded the title of Associate Fellow of the OR Society (AFORS) for his contribution to the field of OR. His research interests include data mining, credit scoring, credit risk modeling, and Basel compliancy.

Learn more about this author by visiting his author page at http://support.sas.com/publishing/authors/brown_iain.html. There, you can download free book excerpts, access example code and data, read the latest reviews, get updates, and more.