

# Developing Credit Risk Models Using SAS<sup>®</sup> Enterprise Miner<sup>™</sup> and SAS/STAT<sup>®</sup>

Theory and Applications



Iain L. J. Brown, PhD

---

# Developing Credit Risk Models

Using SAS<sup>®</sup> Enterprise Miner<sup>™</sup>

and SAS/STAT<sup>®</sup>

Theory and Applications

Iain L. J. Brown, PhD



[support.sas.com/bookstore](http://support.sas.com/bookstore)

The correct bibliographic citation for this manual is as follows: Brown, Iain. 2014. *Developing Credit Risk Models Using SAS<sup>®</sup> Enterprise Miner<sup>™</sup> and SAS/STAT<sup>®</sup>: Theory and Applications*. Cary, NC: SAS Institute Inc.

## **Developing Credit Risk Models Using SAS<sup>®</sup> Enterprise Miner<sup>™</sup> and SAS/STAT<sup>®</sup>: Theory and Applications**

Copyright © 2014, SAS Institute Inc., Cary, NC, USA

ISBN 978-1-61290-691-1 (Hardcopy)

ISBN 978-1-62959-486-6 (EPUB)

ISBN 978-1-62959-487-3 (MOBI)

ISBN 978-1-62959-488-0 (PDF)

All rights reserved. Produced in the United States of America.

**For a hard-copy book:** No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

**For a web download or e-book:** Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

**U.S. Government License Rights; Restricted Rights:** The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a) and DFAR 227.7202-4 and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513-2414.

December 2014

SAS provides a complete selection of books and electronic products to help customers use SAS<sup>®</sup> software to its fullest potential. For more information about our offerings, visit [support.sas.com/bookstore](http://support.sas.com/bookstore) or call 1-800-727-0025.

SAS<sup>®</sup> and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

# Contents

---

**About this Book**

**About the Author**

**Acknowledgments**

**Chapter 1 Introduction**

1.1 Book Overview

1.2 Overview of Credit Risk Modeling

1.3 Regulatory Environment

1.3.1 Minimum Capital Requirements

1.3.2 Expected Loss

1.3.3 Unexpected Loss

1.3.4 Risk Weighted Assets

1.4 SAS Software Utilized

1.5 Chapter Summary

1.6 References and Further Reading

**Chapter 2 Sampling and Data Pre-Processing**

2.1 Introduction

2.2 Sampling and Variable Selection

2.2.1 Sampling

2.2.2 Variable Selection

2.3 Missing Values and Outlier Treatment

2.3.1 Missing Values

2.3.2 Outlier Detection

2.4 Data Segmentation

2.4.1 Decision Trees for Segmentation

2.4.2 K-Means Clustering

2.5 Chapter Summary

2.6 References and Further Reading

**Chapter 3 Development of a Probability of Default (PD) Model**

3.1 Overview of Probability of Default

3.1.1 PD Models for Retail Credit

3.1.2 PD Models for Corporate Credit

### 3.1.3 PD Calibration

---

## 3.2 Classification Techniques for PD

### 3.2.1 Logistic Regression

### 3.2.2 Linear and Quadratic Discriminant Analysis

### 3.2.3 Neural Networks

### 3.2.4 Decision Trees

### 3.2.5 Memory Based Reasoning

### 3.2.6 Random Forests

### 3.2.7 Gradient Boosting

## 3.3 Model Development (Application Scorecards)

### 3.3.1 Motivation for Application Scorecards

### 3.3.2 Developing a PD Model for Application Scoring

## 3.4 Model Development (Behavioral Scoring)

### 3.4.1 Motivation for Behavioral Scorecards

### 3.4.2 Developing a PD Model for Behavioral Scoring

## 3.5 PD Model Reporting

### 3.5.1 Overview

### 3.5.2 Variable Worth Statistics

### 3.5.3 Scorecard Strength

### 3.5.4 Model Performance Measures

### 3.5.5 Tuning the Model

## 3.6 Model Deployment

### 3.6.1 Creating a Model Package

### 3.6.2 Registering a Model Package

## 3.7 Chapter Summary

## 3.8 References and Further Reading

# **Chapter 4 Development of a Loss Given Default (LGD) Model**

## 4.1 Overview of Loss Given Default

### 4.1.1 LGD Models for Retail Credit

### 4.1.2 LGD Models for Corporate Credit

### 4.1.3 Economic Variables for LGD Estimation

### 4.1.4 Estimating Downturn LGD

## 4.2 Regression Techniques for LGD

### 4.2.1 Ordinary Least Squares – Linear Regression

### 4.2.2 Ordinary Least Squares with Beta Transformation

### 4.2.3 Beta Regression

### 4.2.4 Ordinary Least Squares with Box-Cox Transformation

### 4.2.5 Regression Trees

### 4.2.6 Artificial Neural Networks

4.2.7 Linear Regression and Non-linear Regression

4.2.8 Logistic Regression and Non-linear Regression

---

4.3 Performance Metrics for LGD

4.3.1 Root Mean Squared Error

4.3.2 Mean Absolute Error

4.3.3 Area Under the Receiver Operating Curve

4.3.4 Area Over the Regression Error Characteristic Curves

4.3.5 R-square

4.3.6 Pearson's Correlation Coefficient

4.3.7 Spearman's Correlation Coefficient

4.3.8 Kendall's Correlation Coefficient

4.4 Model Development

4.4.1 Motivation for LGD models

4.4.2 Developing an LGD Model

4.5 Case Study: Benchmarking Regression Algorithms for LGD

4.5.1 Data Set Characteristics

4.5.2 Experimental Set-Up

4.5.3 Results and Discussion

4.6 Chapter Summary

4.7 References and Further Reading

## **Chapter 5 Development of an Exposure at Default (EAD) Model**

5.1 Overview of Exposure at Default

5.2 Time Horizons for CCF

5.3 Data Preparation

5.4 CCF Distribution – Transformations

5.5 Model Development

5.5.1 Input Selection

5.5.2 Model Methodology

5.5.3 Performance Metrics

5.6 Model Validation and Reporting

5.6.1 Model Validation

5.6.2 Reports

5.7 Chapter Summary

5.8 References and Further Reading

## **Chapter 6 Stress Testing**

6.1 Overview of Stress Testing

6.2 Purpose of Stress Testing

## 6.3 Stress Testing Methods

---

### 6.3.1 Sensitivity Testing

### 6.3.2 Scenario Testing

## 6.4 Regulatory Stress Testing

## 6.5 Chapter Summary

## 6.6 References and Further Reading

# Chapter 7 Producing Model Reports

## 7.1 Surfacing Regulatory Reports

## 7.2 Model Validation

### 7.2.1 Model Performance

### 7.2.2 Model Stability

### 7.2.3 Model Calibration

## 7.3 SAS Model Manager Examples

### 7.3.1 Create a PD Report

### 7.3.2 Create a LGD Report

## 7.4 Chapter Summary

# Tutorial A – Getting Started with SAS Enterprise Miner

## A.1 Starting SAS Enterprise Miner

## A.2 Assigning a Library Location

## A.3 Defining a New Data Set

# Tutorial B – Developing an Application Scorecard Model in SAS Enterprise Miner

## B.1 Overview

### B.1.1 Step 1 – Import the XML Diagram

### B.1.2 Step 2 – Define the Data Source

### B.1.3 Step 3 – Visualize the Data

### B.1.4 Step 4 – Partition the Data

### B.1.5 Step 5 – Perform Screening and Grouping with Interactive Grouping

### B.1.6 Step 6 – Create a Scorecard and Fit a Logistic Regression Model

### B.1.7 Step 7 – Create a Rejected Data Source

### B.1.8 Step 8 – Perform Reject Inference and Create an Augmented Data Set

### B.1.9 Step 9 – Partition the Augmented Data Set into Training, Test and Validation Samples

### B.1.10 Step 10 – Perform Univariate Characteristic Screening and Grouping on the Augmented Data Set

### B.1.11 Step 11 – Fit a Logistic Regression Model and Score the Augmented Data Set

## B.2 Tutorial Summary

# Appendix A Data Used in This Book

## A.1 Data Used in This Book

---

Chapter 3: Known Good Bad Data

Chapter 3: Rejected Candidates Data

Chapter 4: LGD Data

Chapter 5: Exposure at Default Data

**Index**



# About This Book

---

## Purpose

This book sets out to empower readers with both theoretical and practical skills for developing credit risk models for Probability of Default (PD), Loss Given Default (LGD) and Exposure At Default (EAD) models using SAS Enterprise Miner and SAS/STAT. From data processing and sampling, through segmentation analysis and model building and onto reporting and validation, this text aims to explain through theory and application how credit risk problems are formulated and solved.

---

## Is This Book for You?

Those who will benefit most from this book are practitioners (particularly analysts) and students wishing to develop their statistical and industry knowledge of the techniques required for modelling credit risk parameters. The step-by-step guide shows how models can be constructed through the use of SAS technology and demonstrates a best-practice approach to ensure accurate and timely decisions are made. Tutorials at the end of the book detail how to create projects in SAS Enterprise Miner and walk through a typical credit risk model building process.

---

## Prerequisites

In order to make the most of this text, a familiarity with statistical modelling is beneficial. This book also assumes a foundation level of SAS programming skills. Knowledge of SAS Enterprise Miner is not required, as detailed use cases will be given.

---

## Scope of This Book

This book covers the use of SAS statistical programming (Base SAS, SAS/STAT, SAS Enterprise Guide), SAS Enterprise Miner in the development of credit risk models, and a small amount of SAS Model Manager for model monitoring and reporting.

This book does not provide proof of the statistical algorithms used. References and further readings to sources where readers can gain more information on these algorithms are given throughout this book.

---

## About the Examples

### Software Used to Develop the Book's Content

SAS 9.4

SAS/STAT 12.3

SAS Enterprise Miner 12.3 (with Credit Scoring nodes)

SAS Model Manager 12.3

---

## Example Code and Data

You can access the example code and data for this book by linking to its author page at <http://support.sas.com/publishing/authors>. Select the name of the author. Then, look for the cover thumbnail of this book, and select Example Code and Data to display the SAS programs that are included in this book.

For an alphabetical listing of all books for which example code and data is available, see <http://support.sas.com/bookcode>. Select a title to display the book's example code.

If you are unable to access the code through the website, send e-mail to [saspress@sas.com](mailto:saspress@sas.com).

---

## Additional Resources

SAS offers you a rich variety of resources to help build your SAS skills and explore and apply the full power of SAS software. Whether you are in a professional or academic setting, we have learning products that can help you maximize your investment in SAS.

Bookstore	<a href="http://support.sas.com/bookstore/">http://support.sas.com/bookstore/</a>
Training	<a href="http://support.sas.com/training/">http://support.sas.com/training/</a>
Certification	<a href="http://support.sas.com/certify/">http://support.sas.com/certify/</a>
SAS Global Academic Program	<a href="http://support.sas.com/learn/ap/">http://support.sas.com/learn/ap/</a>
SAS OnDemand	<a href="http://support.sas.com/learn/ondemand/">http://support.sas.com/learn/ondemand/</a>
Support	<a href="http://support.sas.com/techsup/">http://support.sas.com/techsup/</a>
Training and Bookstore	<a href="http://support.sas.com/learn/">http://support.sas.com/learn/</a>
Community	<a href="http://support.sas.com/community/">http://support.sas.com/community/</a>

---

## Keep in Touch

We look forward to hearing from you. We invite questions, comments, and concerns. If you want to contact us about a specific book, please include the book title in your correspondence.

## **To Contact the Author through SAS Press**

---

By e-mail: [saspress@sas.com](mailto:saspress@sas.com)

Via the Web: [http://support.sas.com/author\\_feedback](http://support.sas.com/author_feedback)

---

### **SAS Books**

For a complete list of books available through SAS, visit <http://support.sas.com/bookstore>.

Phone: 1-800-727-0025

Fax: 1-919-677-8166

E-mail: [sasbook@sas.com](mailto:sasbook@sas.com)

---

### **SAS Book Report**

Receive up-to-date information about all new SAS publications via e-mail by subscribing to the SAS Book Report monthly eNewsletter. Visit <http://support.sas.com/sbr>.

---

### **Publish with SAS**

SAS is recruiting authors! Are you interested in writing a book? Visit <http://support.sas.com/saspress> for more information.

---

### **Data Mining with SAS Enterprise Miner**

SAS Enterprise Miner streamlines the data mining process to create highly accurate predictive and descriptive models based on analysis of vast amounts of data from across an enterprise. Data mining is applicable in a variety of industries and provides methodologies for such diverse business problems as fraud detection, customer retention and attrition, database marketing, market segmentation, risk analysis, affinity analysis, customer satisfaction, bankruptcy prediction, and portfolio analysis.

In SAS Enterprise Miner, the data mining process has the following (SEMMA) steps:

- Sample the data by creating one or more data sets. The sample should be large enough to contain significant information, yet small enough to process. This step includes the use of data preparation tools for data importing, merging, appending, and filtering, as well as statistical sampling techniques.
- Explore the data by searching for relationships, trends, and anomalies in order to gain understanding and ideas. This step includes the use of tools for statistical reporting and graphical exploration, variable selection methods, and variable clustering.
- Modify the data by creating, selecting, and transforming the variables to focus the model selection process. This step includes the use of tools for defining transformations, missing value handling, value recoding, and interactive binning.

- Model the data by using the analytical tools to train a statistical or machine learning model to reliably predict a desired outcome. This step includes the use of techniques such as linear and logistic regression, decision trees, neural networks, partial least squares, LARS and LASSO, nearest neighbor, and importing models defined by other users or even outside SAS Enterprise Miner.
- Assess the data by evaluating the usefulness and reliability of the findings from the data mining process. This step includes the use of tools for comparing models and computing new fit statistics, cutoff analysis, decision support, report generation, and score code management.

You might or might not include all of the SEMMA steps in an analysis, and it might be necessary to repeat one or more of the steps several times before you are satisfied with the results.

After you have completed the SEMMA steps, you can apply a scoring formula from one or more champion models to new data that might or might not contain the target variable. Scoring new data that is not available at the time of model training is the goal of most data mining problems.

Furthermore, advanced visualization tools enable you to quickly and easily examine large amounts of data in multidimensional histograms and to graphically compare modeling results.

Scoring new data that is not available at the time of model training is the goal of most data mining exercises. SAS Enterprise Miner includes tools for generating and testing complete score code for the entire process flow diagram as SAS Code, C code, and Java code, as well as tools for interactively scoring new data and examining the results. You can register your model to a SAS Metadata Server to share your results with users of applications such as SAS Enterprise Guide that can integrate the score code into reporting and production processes. SAS Model Manager complements the data mining process by providing a structure for managing projects through development, testing, and production environments and is fully integrated with SAS Enterprise Miner.

---

## About Credit Scoring for SAS Enterprise Miner

The additional add-in of credit scoring for SAS Enterprise Miner facilitates analysts in building, validating, and deploying credit risk models. It enables organizations to create credit scorecards using in-house expertise and resources to decide whether to accept an applicant (application scoring); to determine the likelihood of defaults among customers who have already been accepted (behavioral scoring); and to predict the likely amount of debt that the lender can expect to recover (collection scoring). Three key additional nodes are made available to analysts: the Interactive Grouping node, Scorecard node and Reject Inference node. **Note:** The Credit Scoring for SAS Enterprise Miner solution is not included with the base version of SAS Enterprise Miner. If your site has not licensed Credit Scoring for SAS Enterprise Miner, the credit scoring node tools do not appear in your SAS Enterprise Miner software.

SAS Enterprise Miner and the Credit Scoring nodes will be used extensively throughout this book to demonstrate the data exploration and modelling processes discussed. A tutorial section is also located at the end of this book which gives a step-by-step walkthrough of typical tasks using SAS Enterprise Miner.

---

## **About SAS/STAT**

SAS/STAT software provides comprehensive statistical tools for a wide range of statistical analyses, including analysis of variance, categorical data analysis, cluster analysis, multiple imputation, multivariate analysis, nonparametric analysis, power and sample size computations, psychometric analysis, regression, survey data analysis, and survival analysis. A few examples include nonlinear mixed models, generalized linear models, correspondence analysis, and robust regression. The software is constantly being updated to reflect new methodology. In addition to more than 80 procedures for statistical analysis, SAS/STAT software also includes the Power and Sample Size Application (PSS), an interface to power and sample size computations.

SAS/STAT code is used extensively throughout this book to demonstrate data manipulation and model development coded tasks.

# 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](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

# Acknowledgments

---

This book would not have been possible without the support and guidance of a number of important people whom I would like to take this opportunity to acknowledge and thank.

Among the many people I would like to thank are all those people at SAS involved in making this book possible, including my acquisition editor Shelley Sessoms, technical editors Naeem Siddiqi and Jim Seabolt, whose input has greatly enhanced the text, and a special thanks to Stephenie Joyner and Brenna Leath, my developmental editors, for keeping me on track. I am also greatly thankful to the SAS UK Analytics Practice, in particular Dr. Laurie Miles, John Spooner, and Colin Gray, for their support and guidance throughout my career at SAS.

I would also like to pay a special thanks to my supervising team during my doctoral research: Dr. Christophe Mues, Prof. Lyn Thomas, and Dr. Bart Baesens. Without their joint tutorage and expert knowledge in the field of credit risk modelling, I could not have achieved much of the work conducted in this book. I have also had the great pleasure of working alongside a number of well-established and flourishing academics in the field of credit risk and credit scoring. I would like to thank the team I worked alongside on the LGD benchmarking case study referenced in [Chapter 4](#): Dr. Gert Loterman, Dr. David Martens, Dr. Bart Baesens, and Dr. Christophe Mues.

Finally, I would like to express my utmost thanks to my wife, parents, sister, and whole family, without whose support I would not have achieved any of the goals I have set out to attain.

---

# Chapter 1 Introduction

## 1.1 Book Overview

## 1.2 Overview of Credit Risk Modeling

## 1.3 Regulatory Environment

### 1.3.1 Minimum Capital Requirements

### 1.3.2 Expected Loss

### 1.3.3 Unexpected Loss

### 1.3.4 Risk Weighted Assets

## 1.4 SAS Software Utilized

## 1.5 Chapter Summary

## 1.6 References and Further Reading

---

## 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, number of the accord's inadequacies, in particular with regard to the way that credit risk wa

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

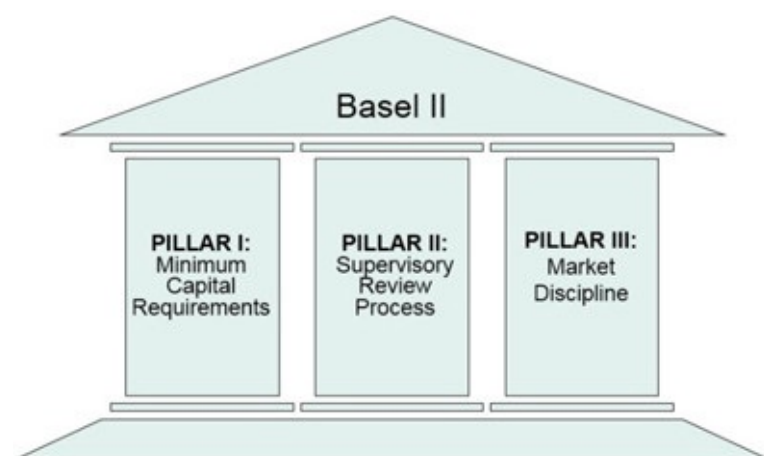
The Basel Capital Accord (Basel Committee on Banking Supervision, 2001a) prescribes the minimum amount of regulatory capital an institution must hold so as to provide a safety cushion against unexpected losses. The Accord is comprised of three pillars, as illustrated by [Figure 1.1](#):

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

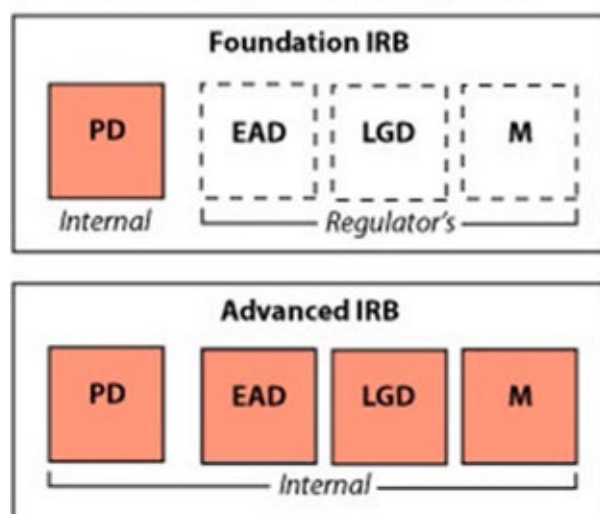
Under the standardized approach, banks are required to use ratings from external credit rating agencies to quantify required capital. The main purpose and strategy of the Basel committee is to offer capital incentives to banks that move from a supervisory approach to a best-practice advanced internal 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**

*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:

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

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:

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

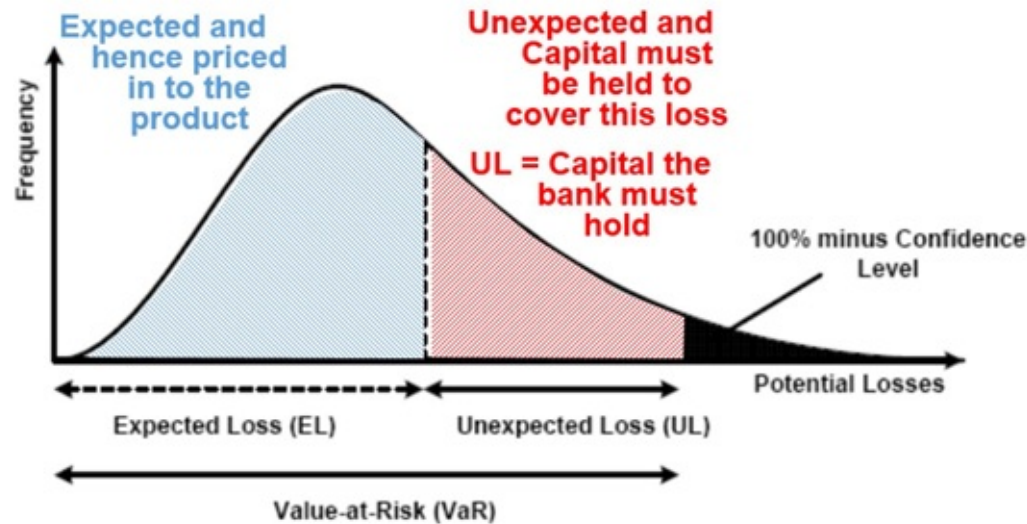
In the previous example, expected loss as a percentage of EAD would be equal to  $EL\% = 0.8\%$ .

---

### 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

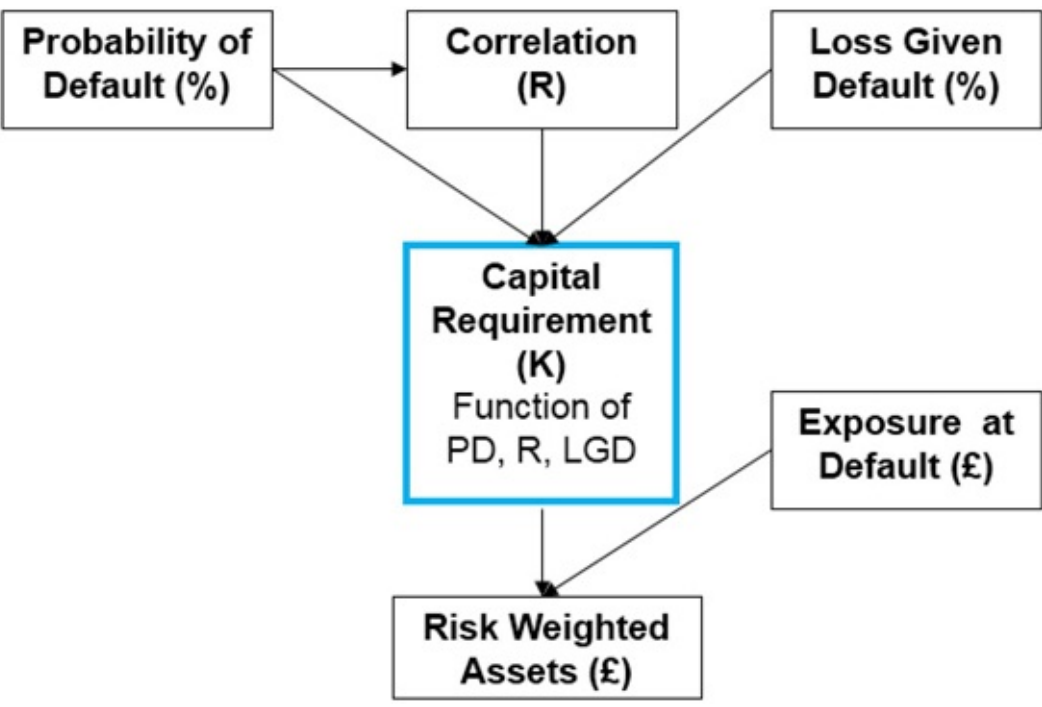
Risk Weighted Assets (RWA) are the assets of the bank (money lent out to customers and businesses in the form of loans) accounted for by their riskiness. The RWA are a function of PD, LGD, EAD and M, where K is the capital requirement:

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

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  $\phi$  denotes the normal cumulative distribution function and  $\phi^{-1}$  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.

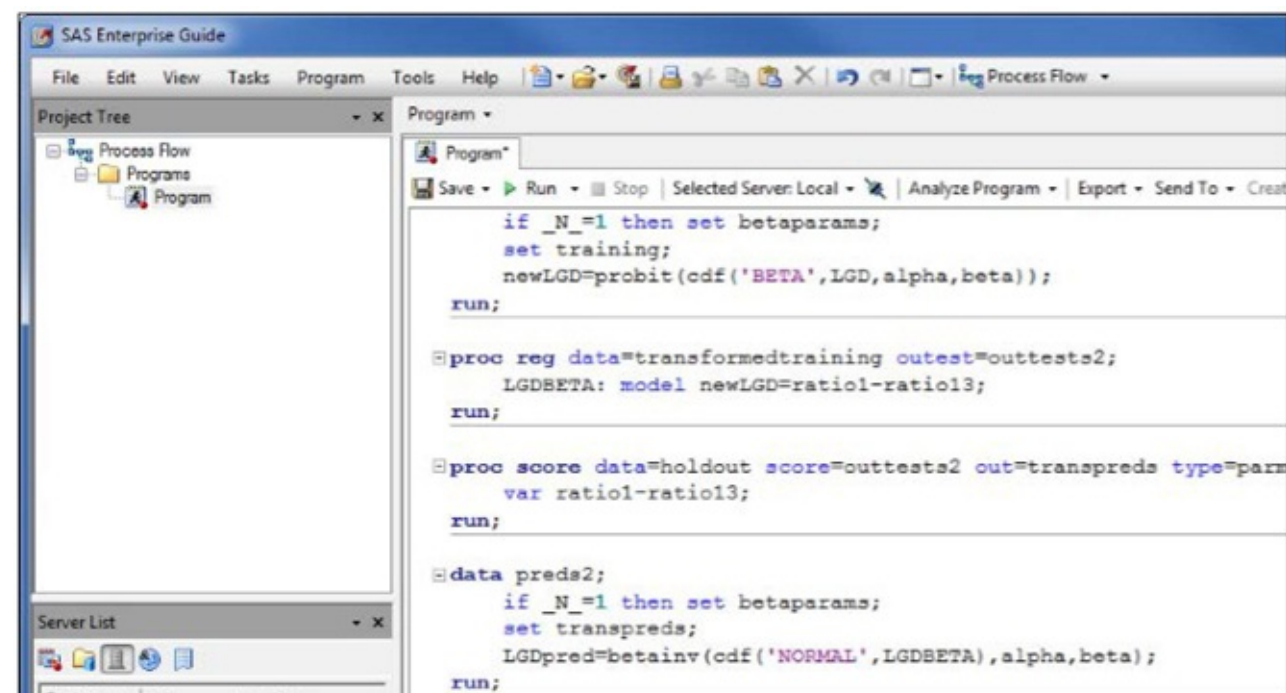
## 1.4 SAS Software Utilized

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**





---

sample content of Developing Credit Risk Models Using SAS Enterprise Miner and SAS/STAT:  
Theory and Applications

- [download The Manual of Scientific Style: A Guide for Authors, Editors, and Researchers for free](#)
- [download online America's Tea Parties: Not One but Four! Boston, Charleston, New York, Philadelphia book](#)
- **[click The Philosophy of Sartre book](#)**
- [read online Early Twentieth-Century Continental Philosophy \(Studies in Continental Thought\) online](#)
  
- <http://www.freightunlocked.co.uk/lib/Term-Limits.pdf>
- <http://creativebeard.ru/freebooks/America-s-Tea-Parties--Not-One-but-Four--Boston--Charleston--New-York--Philadelphia.pdf>
- <http://qolorea.com/library/Is-the-Internet-Changing-the-Way-You-Think---The-Net-s-Impact-on-Our-Minds-and-Future.pdf>
- <http://studystategically.com/freebooks/Early-Twentieth-Century-Continental-Philosophy--Studies-in-Continental-Thought-.pdf>