Statistical Credit Risk Models and Empirical Evidence for Financial Markets and Banks in Japan
日本の金融市場と銀行融資における信用リスク計量化の実証分析

Akihiro Kawada

Graduate School of Engineering

Tokyo University of Science

A thesis submitted for the degree of

Doctor of Philosophy

February 2015

Preface

This thesis is the product of my Ph.D. studies at the Department of Management Science, Graduate School of Engineering, Tokyo University of Science. The thesis has benefited greatly from the help of a number of individuals.

The thesis comprises three issues on credit risk modeling. Issues I and III are joint collaborations between Akihiro Kawada and Takayuki Shiohama. Takayuki Shiohama is a Junior Associate Professor of the Department of Management Science, Faculty of Engineering, Tokyo University of Science. Issue II is a joint collaboration between Akihiro Kawada and Satoshi Yamashita at the Institute of Statistical Mathematics. Satoshi Yamashita is a professor at the Institute of Statistical Mathematics. Their supervising, advice, feedback, and teaching contributed greatly to each issue.

In particular, I am deeply grateful to my thesis advisors, Takayuki Shiohama and Toshikazu Yamaguchi, for all their help and valuable advice during my Ph.D. studies.

Then, I would like to express my sincere gratitude to Satoshi Yamashita for the opportunity to collaborate on the research, as well as his patience, constructive guidance, and useful advice. Finally, I would like to acknowledge the many helpful discussions during the research collaboration at the Institute of Statistical Mathematics are also gratefully acknowledged.

All omissions and errors are my own.

Akihiro Kawada

Contents

List of Figures **iv**

List of Tables **v**

1 Introduction **1**

2 Issue I - Estimating Interest Rate and Risk Premium Term Structures from Credit Ratings and Financial Information **6**

2.1 Introduction 6

2.2 Statistical Model for Evaluating Credit Risk 8

2.2.1 Bond price valuation formula 8

2.2.2 Model formulation 9

2.2.3 Parameter estimation 13

2.2.4 Selecting the number m, variance parameter s, and penalty parameter $(\lambda_\omega, \lambda_\beta)$ for the basis function 15

2.3 Simulation 16

2.4 Empirical Analysis 23

2.4.1 Description of the data 23

2.4.2 Results and observations 23

2.5 Conclusions and Future Topics 30

2.6 Appendix 30

3 Issue II - Forecasting Loss Given Default for Bank Loans : An Empirical Analysis for Japanese Banks **34**

3.1 Introduction 34

3.2 Literature Review 36

3.3 Data . . . 39
3.3.1 The definition of default . . . 39
3.3.2 The definition of LGD . . . 40
3.3.3 Description of the data . . . 41
3.3.4 Data set A (for developing the PD predictive model) . . . 42
3.3.5 Data set B (for the development of the LGD predictive model) 44
3.4 Linear Regression Analysis . . . 47
3.5 The EL Forecasting Model . . . 49
3.5.1 Overview of the EL forecasting model . . . 49
3.5.2 Model coefficients given by regressions . . . 50
3.6 Validation of the Multi-stage LGD Model . . . 56
3.7 Conclusions . . . 57

4 Issue III - Structural Credit Risks with Non-Gaussian and Serially Correlated Innovations 59
4.1 Introduction . . . 59
4.2 Merton's Model with Non-Gaussian and Dependent Innovations . 61
4.3 Examples of Evaluating Credit Risks . . . 67
4.4 Parameter Estimation . . . 69
4.5 Empirical Analysis . . . 71
4.6 Summary and Conclusion . . . 72
4.7 Appendix . . . 77

5 Summary and Conclusion 78

References 80

List of Figures

2.1 Example of Gaussian basis function 11
2.2 Term structure of forward rates for government bonds 17
2.3 Forward rates for government and corporate bonds 26
2.4 Corporate bond risk premiums 27
2.5 Forward rates considering financial ratios 27
2.6 Forward rates considering financial ratios (A) 28
2.7 Forward rates considering financial ratios (Baa) 28
2.8 Forward rates considering financial ratios (Ba) 29

3.1 Histogram of LGD 45
3.2 EL forecasting model 50

4.1 Data for implementation (ANA Holdings Inc.). 73
4.2 The distribution of the logarithmic return of ANA Holdings Inc. The thick line represents the kernel density plot, and the dotted line represents the normal distribution. 74
4.3 The top four panels are time-series plots for the estimates of parameters μ, σ, C_3, and C_4 in both Merton's model and the non-Gaussian model. The fifth panel shows the time-series plots for the estimated asset process. The bottom panel compares the probability of default between Merton's model and the non-Gaussian model. 75
4.4 Plots of the term structures of credit spreads as a function of maturity T. 76
4.5 Plots of the distance to default as a function of maturity T. . . . 76

List of Tables

2.1 Simulation results 1: $RMSE$ of $\hat{\boldsymbol{\beta}}$ 20
2.2 Simulation results 2: proportion of variables eliminated 21
2.3 Variance-covariance matrix 22
2.4 Estimates of $\boldsymbol{\beta}$ coefficients for financial ratios 24
2.5 Estimates of error variance 26

3.1 The LGD levels in published empirical studies 37
3.2 The number of borrowers, records, and defaults 42
3.3 Median, mean, and standard deviation of variables (data set A) . 43
3.4 Correlation matrix of data set A 43
3.5 The number of each type of default resolution 44
3.6 Median, mean, and standard deviation of LGD 45
3.7 Median, mean, and standard deviation of the explanatory variables (data set B) 46
3.8 Length of the workout process (in years) for each bank 46
3.9 Correlation matrix of data set B 47
3.10 Result of the linear regression analysis of the factors influencing LGD 48
3.11 Data used to build each model 50
3.12 The estimation results for (I)PD model 52
3.13 The estimation results for (II)Pr(Recovery) model 53
3.14 The estimation results for (III)Pr($LGD > 0$) model 54
3.15 The estimation results for (IV)LGD regression model 55
3.16 10-fold cross validation 57

4.1 Parameter estimates of Merton's model and the non-Gaussian model 73

1

Introduction

Credit risk refers to the risk that loan borrowers, bond issuers and participates in financial transactions will default. It is crucial for financial institutions to accurately estimate credit risk. Credit risk modeling can be divided broadly into the following three approaches. The first approach is the structural model. The second approach is the reduced form model. The third approach is the statistical model. Intended purposes and calibration data vary between each approach. Hull (2012) provides a comprehensive introduction to credit risk modeling. We provide brief overviews of each approach.

The structural model

In the structural model, the probability of default can be estimated from equity prices. Merton (1974) has proposed to estimate the probability of default (PD) from equity prices. Merton considers a company's equity as a call option on the assets of the company. The Merton's model based on the assumption that the financing of the firm consists of one type of equity and one zero coupon bond that matures at time T for simplicity. If the value of the firm's assets at time T falls below the face value of the bond, the firm fails to repay the debt at time T and defaults. If the value of the firm's assets at time T exceeds the face value of the bond, the firm can repay the debt and do not default. The strike price of the call option on the value of the assets of the firm is the same as the face value of the bond. In the Merton's model, the firm defaults when this call option is not exercised.

Therefore, the probability of default can be obtained from the application of the Black-Sholes options pricing model (Black and Scholes (1973)).

The reduced form model

In the reduced form model, the probability of default can be estimated from credit spreads which are observed in financial markets. The credit spreads is the excess rate of interest. Two types of credit spread can be observed in financial markets. One is the credit default swap (CDS) spread. Another is the bond yield spread. In general, investors require the appropriate interest rate for a credit risk, so the credit spreads in financial market imply the information for the credit risk. Duffie and Singleton (1999) and Jarrow and Turnbull (1995) are representative models of the reduced form model. In the reduced form model, the credit spreads are considered expected loss rate. On this assumption, the probability of default can be estimated implicit in the credit spreads when a certain recovery rate is given.

The statistical model

The statistical model predicts defaults from financial information (e.g., accounting ratios). Altman (1968) developed the statistical model known as the Z-Score model. This model is based on the statistical method known as discriminant analysis. The Z-Score is calculated as a linear combination of accounting ratios. This score indicates the likeliness that the firm will default. The firm with a high score is unlikely to default. The firm with a low score is likely to default. The statistical model is closely related to capital regulations of banking organizations. Basel Committee on Banking Supervision (2006) proposed a framework for capital measurement and capital standards. This framework is known as Basel II. In Basel II, banks are allowed to calculate capital requirement with the internal Ratings-Based (IRB) Approach. Most banks have internal credit rating systems for evaluating loan borrower's creditworthiness. In the IRB approach, key credit risk parameters are the probability of default (PD), the loss given default (LGD) and the exposure at default (EAD). These credit risk parameters are estimated with the forecasting models which are developed by each bank. These models are generally based on statistical methods. The statistical models are calibrated to bank's own loan data. Using the predicted credit risk parameters from the models as inputs, banks calculate the capital requirements.

This study is aimed at contributing extensively to improvements of credit risk modeling. To obtain a wide range of knowledge about credit risk modeling, we focus on the issues in all three approaches of credit risk modeling in this study.

This thesis comprises three issues on credit risk modeling. Each issue can be read independently. However, they are mutually related via dealing with credit risk quantification. Issue I and III of this thesis are joint collaborations between Akihiro Kawada and Takayuki Shiohama. Takayuki Shiohama is a Junior Associate Professor of Department of Management Science, Faculty of Engineering, Tokyo University of Science. Issue II of this thesis is a joint collaboration between Akihiro Kawada and Satoshi Yamashita. Satoshi Yamashita is a Professor, The Institute of Statistical Mathematics. The issues that we focus on in each issue as follows.

Issue I - Estimating Interest Rate and Risk Premium Term Structures from Credit Ratings and Financial Information

This comparative study shows that yields on corporate debt traded in the market are normally higher than those on government debt because corporate debt yields carry higher credit and liquidity risks. This difference in yields is called the risk premium. The risk premium is useful information to measure the credit risk of the corporation issuing the debt, and has thus long been estimated from bond prices. In this paper, we build a statistical model that expands on an inductive model, using it to estimate the term structure of interest rates and risk premiums. Estimating the term structure of the risk premium needs to be done per rating. However, this cannot express the differences between companies with the same rating; thus, we employ a model where the risk premium level varies with the issuing company's financial metrics. Ando and Yamashita (2005) include in their model financial measures pre-selected for their strong relationship to the risk premium in performing their estimations of risk premiums. In this research, we propose a methodology which simultaneously selects variables for financial measures and estimates model parameters, using a penalty term which selects variables, based on Fan and Li (2001).

Issue II - Forecasting Loss Given Default for Bank Loans : An Empirical Analysis for Japanese Banks

PD(Probability of Default) and LGD (Loss Given Default) are key risk parameters in credit risk management. It is crucial for financial institutions to accurately estimate PD and LGD. Most of researches on LGD is based on the corporate bond market. On the other hand, few studies focused on LGD for bank loans. In particularly in Japan, few data on bank loan losses are not publicly available for researchers. Consequently, knowledge about LGD of Japanese bank loans has not sufficiently accumulated. In this study, using the data on Japanese bank loans, we analyze influencing factors on LGD and develop the model for predicting LGD EL (Expected Loss). As a result, we found that collaterals, guarantees, loan sizes have impacts on LGD and that the levels of LGD for Japanese bank loans are significantly lower than the levels of LGD reported in the other countries. It is difficult for traditional linear models to accurately predict LGD because of business practices unique to Japanese banks and characteristics of LGD for Japanese bank loans. Therefore, we propose the multi-stage model for predicting LGD, and confirmed that the multi-stage model performs better compared to traditional linear models.

Issue III - Structural Credit Risks with Non-Gaussian and Serially Correlated Innovations

We expand the structural credit risk model pioneered by Merton (1974) to include an underlying asset process with a non-Gaussian and dependent stochastic nature. By using a standard Edgeworth expansion, we arrive at closed-form analytic expressions for the probability of default, the distance to default, and the term structure of credit spread, which allows us to evaluate a more accurate credit risk incorporating non-normal asset returns. The Moody's KMV analogous procedures for the estimation of the model parameters are proposed. Empirical applications for the credit risk evaluations are illustrated using a large Japanese corporation (ANA Holdings Inc.), which reveal the significant effects on credit risk due to the non-Gaussianity of the underlying firm's asset process.

The remainder of this thesis is structured as follows. Chapter 2 present a proposal to incorporate financial information of bond issuers with the reduced form models. This chapter is based on Kawada and Shiohama (2009). Chapter 3 focuses on the statistical model for the LGD on bank loans. This chapter based on Kawada and Yamashita (2013). Chapter 4 attempts to expand the structural model proposed in Merton (1974) so as to assume a non-Gaussian asset return process. This chapter is based on Kawada and Shiohama (submitted). Finally, Chapter 5 provides summary and conclusion.

2

Issue I - Estimating Interest Rate and Risk Premium Term Structures from Credit Ratings and Financial Information

2.1 Introduction

This study quantifies credit risk by using credit ratings and financial information. Kijima and Komoribayashi (1999) define credit risk as follows: a situation wherein the issuers of financial products or partners in a transaction cannot fulfill their obligations is called a default, and credit risk is the direct or indirect loss suffered in a default. In the case of loans or credit, a default causes all or some portion of the total of principle and unpaid interest to become unrecoverable, and in the case of derivatives, the unrealized gains would be lost. There is also the indirect loss caused by a decline in bond prices and a decrease in liquidity reflected in lower creditworthiness caused by a higher possibility of default.

It is important for financial institutions and companies holding debt to gauge credit risk. Particularly, it is necessary to fairly and accurately value debt by considering credit risk, which necessitates analyzing credit risk in more detail.

According to Kusuoka et al. (2001), credit risk in financial products can be broadly divided into two categories. Issuer risk arises when the issuer defaults on bonds or other financial products. For example, in the case of buying corporate bonds of a certain company from another company, the credit risk stems from that company, the issuer of the bond. Counterparty risk, in contrast, in the context of a derivatives contract, refers to the risk of the other party in the derivatives contract becoming unable to fulfill its obligations before default on the underlying asset.

Most attempts to quantify credit risk are based on corporate debt and ratings. Yields on corporate debt are higher, compared with government debt yields, as the possibility of default on government debt is relatively lower. This difference between yields is called the risk premium. Credit risk is the risk involved with the inability to fulfill obligations, and risk premiums in the corporate debt market are perceived as combining liquidity risk with credit risk. Generally, lower the risk premium, greater the ability of the debt issuer to fulfill its obligation. The ratings assigned by ratings agencies are also important information for evaluating credit risk.

This study uses basis functions to represent term structures. Substantial prior research exists on methodologies for estimating interest term structures from cross sections of interest-bearing debt. McCulloch (1971) provides a paper of particular import on the topic of estimating interest rate term structures using basis functions. Ando and Yamashita (2005) estimate interest rate and risk premium term structures using basis functions, and their model includes the financial ratios believed to have the strongest relationship to credit risk. In this research, in contrast, we select financial ratios from multiple candidates and simultaneously estimate model parameters, using the maximum-likelihood method with a penalty term that selects variables, based on Fan and Li (2001).

This chapter is organized as follows. Section 2.2 describes the mathematical model for the methodology for evaluating credit risk, and proposes an approach to simultaneously estimate parameters and selecting variables. Section 2.3 presents the results of a simulation to validate the effectiveness of the proposed approach. Section 2.4 conducts the estimation of the term structure of interest rates and risk premiums, selecting the variables from pricing data for government and corporate obligations and financial ratios of issuers. Section 2.5 summarizes our results and discusses topics for future research. Section 2.6 presents the derivation of the Bayesian Information Criterion.

2.2 Statistical Model for Evaluating Credit Risk

2.2.1 Bond price valuation formula

The basic concepts of mathematical finance indicate that the value of a financial product is represented by the sum of the cash flows generated by the asset in the present and into the future discounted to a present value at a spot rate, based on the assumption of risk neutrality. Accordingly, let us formulate the theoretical price for government and corporate obligations.

Consider a government bond with coupon value C, and in addition to the coupon a cash flow on maturity of a par value of R. Assuming that from the present to maturity T_L there will be L coupon payments, at times given by $\mathbf{t} = (t_1, ..., t_L)'$.

Expressing the term structure of the risk-free instantaneous forward rate as $r(t)$, the PV (present value) of the government bond maturing at T_L can be expressed as follows:

$$PV(r(\cdot), \boldsymbol{t}) = \sum_{i=1}^{L} C \cdot \exp\left\{-\int_0^{t_i} r(u)du\right\} + R \cdot \exp\left\{-\int_0^{t_L} r(u)du\right\}.$$

According to this formula, the present value of the government bond can be represented as the future cash flows (C, R) discounted to present value by the discount function $d(t) = \exp\left\{-\int_0^t r(u)du\right\}$.

Let us now consider the cash flows for the same corporate bond. The present value of a corporate bond can be expressed as follows, using risk premium $s(t)$,

$$\begin{aligned} PVcp = &\sum_{i=1}^{L} C \cdot \exp\left[-\int_0^{t_i} \{r(u) + s(u)\}\, du\right] \\ &+ R \cdot \exp\left[-\int_0^{t_L} \{r(u) + s(u)\}\, du\right]. \end{aligned}$$

The first term is the coupon, and the second is the face value redemption amount in case of default. For corporate bonds, we discount to present value using the discount function $d(t) = \exp\left[-\int_0^{t_L} \{r(u) + s(u)\}\, du\right]$, which considers the risk premium.

2.2.2 Model formulation

Let us show a model using the bond price valuation formula which simultaneously estimates the term structure for risk-free interest rates and risk premiums. The risk premiums estimated in this study are equal to a hazard in the case of a recovery rate of 0. In other words, we are evaluating credit risk more conservatively. Estimating recovery rates is in many cases more difficult than estimating probabilities of default, largely due to the paucity of published data. Araten et al. (2004) and Altman et al. (2004) make use of actual default data to show that recovery rates are lower during periods of higher probability of default. Since the theoretical price of corporate bonds and CDs depends on the expected value of the loss from default, bond prices differ greatly due to their wide variations in valuation of recovery rates by market participants, the seniority structure, and pertinent circumstances in the economic, financial, and credit markets. Therefore, we can estimate the term structure of recovery rates from bond prices; however, the results of the estimation include a large bias. In this study we have, thus, chosen to estimate only risk premiums, rather than term structures for recovery rates. Jarrow (2001) attempts to simultaneously estimate defaults and recovery rates by considering seniority structure.

Assume, we are given data for a government bond trading price on a specific date y_i^g, and a corporate bond trading price y_i^{cp}. Let the number of government bond data points be n_0. For each corporate bond, let ratings be assigned of $1, ..., J$, with the total number of types of ratings J and the total number of bond data points belonging to each rating being $n_1, ..., n_J$. Denote the financial information observed for the companies issuing each corporate bond as $\boldsymbol{x} = (x_1, ..., x_p)'$. The parameters we wish to estimate are the interest rate term structure $r(t)$ and risk premium term structure by rating $s_j(t)$. We will estimate these simultaneously.

To express the divergence between bond prices at which transactions actually occur and theoretical prices, we write the current transacted price for government bonds PV^g, and current transacted price for corporate bonds PV^{cp} in the form of the following statistical model:

$$\begin{cases} \text{Government bond price model} \\ \quad y_{i0}^g = PV_{i0}^g(r(\cdot), t_{i0}) + \varepsilon_{i0}^g (i_0 = 1, ..., n_0), \\ \text{Corporate bond price model (rating 1)} \\ \quad y_{i1}^{cp} = PV_{i1}^{cp}(r(\cdot), s_1(\cdot), \boldsymbol{t}_{i1}) + \varepsilon_{i1}^1 (i_1 = 1, ..., n_1), \\ \qquad \vdots \\ \text{Corporate bond price model (rating } J) \\ \quad y_{iJ}^{cp} = PV_{iJ}^{cp}(r(\cdot), s_J(\cdot), \boldsymbol{t}_{iJ}) + \varepsilon_{iJ}^J (i_J = 1, ..., n_J), \end{cases} \tag{2.2.1}$$

here we assume the noise factors $\varepsilon_{i0}^g, \varepsilon_{i1}^1, ..., \varepsilon_{iJ}^J$ are mutually independent, following a normal distribution with mean of 0 and variance of $\sigma_0^2, \sigma_1^2, ..., \sigma_J^2$. Since the time $\boldsymbol{t}$ of the arrival of the cash flow differs for each bond, we represent symbolically the time of cash flow of the i_0'th government bond as $\boldsymbol{t}_{i0}$, and the time of cash flow of the ij'th corporate bond, which has been assigned the rating of j, as $\boldsymbol{t}_{ij}$.

In this study, to flexibly estimate the interest term structure $r(t)$ and risk premium term structure by rating $h_j(t)(j = 1, ..., J)$, we assume the following structure,

$$\begin{aligned} r(t, \boldsymbol{\omega}_0) &= \sum_{k=1}^{m_0} \omega_{0k}(t) = \boldsymbol{\omega}_0' \boldsymbol{\phi}_0(t), \\ s_j(t, \boldsymbol{\omega}_j) &= \left[\sum_{k=1}^{m_j} \omega_{jk} \phi_{jk}(t) \right] \cdot \exp \left\{ \sum_{l=1}^{p} \beta_l x_l \right\} = \left[\boldsymbol{\omega}_j' \boldsymbol{\phi}_j(t) \right] \cdot \exp \left\{ \boldsymbol{\beta}' \boldsymbol{x} \right\}. \end{aligned} \tag{2.2.2}$$

Here $\boldsymbol{\phi}_j(t) = (\phi_1(t), ..., \phi_{m_j}(t))'$ is the known basis function vector comprising basis functions, and $\boldsymbol{\omega}_j = (\omega_{j1}, ..., \omega_{jm_j})', \boldsymbol{\beta} = (\beta_1, ..., \beta_p)'$ are the unknown parameters to be estimated. This equation expresses the term structure as a smooth curve. In addition, the risk premium level has a structure which increases or decreases depending on the value of financial ratio $\boldsymbol{x}$, due to $\exp\{\boldsymbol{\beta}' \boldsymbol{x}\}$. Assuming this structure, the interest rate and risk premium term structures are represented by linearly independent basis functions, such that the risk premium term structure is influenced by the financial ratios.

Since McCulloch (1971), the methodology of estimating interest rate term structures by regressing a cross section of interest-bearing instruments against a set of appropriate default functions has been widely used in both academia and businesses. Various types of basis functions have been proposed, but in this study we make use of Gaussian basis functions, which have been used in recent years to model complex natural and social phenomena with non-linear structures. Kawasaki and Ando (2002) also used these basis functions to express interest rate term structures. While we use Gaussian default functions in this study, it is also possible to use P-splines or B-splines. However, when using spline functions a problem of node selection arises. Even when using Gaussian default functions problems arise, of selecting basis functions and smoothing parameters, but we shall estimate these parameters in Section 2.2.4 using information criteria. The Gaussian basis functions $\boldsymbol{\phi}_j(t) = (\phi_1(t), ..., \phi_{m_j}(t))'$ are given by the following formula,

$$\phi_j(t) = \exp\left(-\frac{(t-\mu_j)^2}{2s^2}\right), \qquad j = 1, ..., m_j.$$

Here μ_j is the center of the basis function, and s^2 is a quantity representing the spread of the basis function. An example is shown in Figure 2.1.

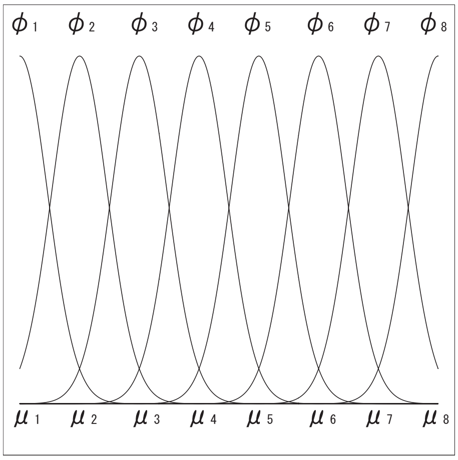

Figure 2.1: Example of Gaussian basis function

From (2.2.2), the formula for bond price valuation, the discounted current value PV^g of a government bond with L interest payments from the present to maturity, including T_L at time of redemption, and the discounted present value PV^{cp} of a corporate bond assigned rating j are, respectively:

$$PV^g(\boldsymbol{\omega}_0, \boldsymbol{t}) = \sum_{i=1}^{L} C \cdot \exp\{-\boldsymbol{\omega}_0' \boldsymbol{\psi}_0(t_i)\} + R \cdot \exp\{-\boldsymbol{\omega}_0' \boldsymbol{\psi}_0(t_L)\},$$

$$\begin{aligned} PV^{cp}(\boldsymbol{x}, \boldsymbol{\omega}_0, \boldsymbol{\omega}_j, \boldsymbol{\beta}, \boldsymbol{t}) = &\sum_{i=1}^{L} C \cdot \exp\left\{-\boldsymbol{\omega}_0' \boldsymbol{\psi}_0(t_i) - \boldsymbol{\omega}_j' \boldsymbol{\psi}_j(t_i) \cdot \exp(\boldsymbol{\beta x})\right\} \\ &+ R \cdot \exp\left\{-\boldsymbol{\omega}_0' \boldsymbol{\psi}_0(t_L) - \boldsymbol{\omega}_j' \boldsymbol{\psi}_j(t_L) \cdot \exp(\boldsymbol{\beta x})\right\}, \quad (2.2.3) \end{aligned}$$

where $\boldsymbol{t} = (t_1, ..., t_L)'$ are the times of interest payments, and $\boldsymbol{\psi}_j(t) = (\psi_{j1}(t), ..., \psi_{jm_j}(t))'$ is a vector of dimension m_j with components of

$$\psi_{jk}(t) = \int_0^t \phi_j k(u) du.$$

Based on the statistical model (2.2.1) and bond price valuation formula (2.2.3), the bond price model can be formulated as the density function of a normal distribution:

$$\begin{cases} \text{Government bond price model} \\ \quad f(y_{i0}^g | \boldsymbol{t}_{i0}; \boldsymbol{\omega}_0, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{y_{i0} - PV_{i0}^g(\boldsymbol{\omega}_0, \boldsymbol{t}_{i0})^2}{2\sigma_0^2}\right], \\ \text{Corporate bond price model (rating 1)} \\ \quad f(y_{i1}^{cp} | \boldsymbol{t}_{i1}, \boldsymbol{x}_{i1}; \boldsymbol{\omega}_0, \boldsymbol{\omega}_1, \boldsymbol{\beta}, \sigma_1^2) \\ \quad = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{y_{i1} - PV_{i1}^{cp}(\boldsymbol{x}_{i1}, \boldsymbol{\omega}_0, \boldsymbol{\omega}_1, \boldsymbol{\beta}, \boldsymbol{t}_{i1}))^2}{2\sigma_1^2}\right], \\ \qquad \vdots \\ \text{Corporate bond price model (rating J)} \\ \quad f(y_{iJ}^{cp} | \boldsymbol{t}_{iJ}, \boldsymbol{x}_{iJ}; \boldsymbol{\omega}_0, \boldsymbol{\omega}_J, \boldsymbol{\beta}, \sigma_J^2) \\ \quad = \frac{1}{\sqrt{2\pi\sigma_J^2}} \exp\left[-\frac{y_{iJ} - PV_{iJ}^{cp}(\boldsymbol{x}_{iJ}, \boldsymbol{\omega}_0, \boldsymbol{\omega}_J, \boldsymbol{\beta}, \boldsymbol{t}_{iJ}))^2}{2\sigma_J^2}\right], \end{cases}$$

where y_{i0}^g, y_{ij}^{cp} is the bond transaction price. The bond pricing model takes the company financial data $\boldsymbol{x}$, current bond price y_{i0}^g, y_{ij}^{cp}, the times of occurrence of each flow t, and rating of the company issuing the bond $j = 1, ..., J$ as inputs. We estimate the parameters $\boldsymbol{\theta} = (\boldsymbol{\omega'}, \boldsymbol{\sigma'}, \boldsymbol{\beta'})'$, $\boldsymbol{\omega} = (\boldsymbol{\omega}_0', ..., \boldsymbol{\omega}_J')'$, $\boldsymbol{\sigma} = (\sigma_0^2, \sigma_1^2, ..., \sigma_J^2)'$.

2.2.3 Parameter estimation

As a method for estimating the parameters of the bond valuation formula based on Gaussian basis functions, we can use the maximum-likelihood method. The estimated maximum-likelihood quantities can be obtained by maximizing the logarithmic maximum-likelihood function,

$$l(\boldsymbol{\theta}) = \sum_{i_0=1}^{n_0} \log f(y_{i_0}^{g} | \boldsymbol{t}_{i_0}; \boldsymbol{\omega}_0, \sigma_0^2) + \sum_{j=1}^{J} \left[\sum_{i_j=1}^{n_j} \log f(y_{i_j}^{cp} | \boldsymbol{t}_{i_j}, \boldsymbol{x}_{i_j}; \boldsymbol{\omega}_0, \boldsymbol{\omega}_j, \boldsymbol{\beta}, \sigma_j^2) \right].$$

However, in the maximum-likelihood method, depending on the number of basic functions the estimated maximum-likelihood values may become unstable, with excessive fluctuations in the estimated curve. In this study, therefore, we estimate based on maximizing a logarithmic maximum-likelihood function with a penalty using a second-order difference of $\boldsymbol{\theta}$, a parameter included in the bond price model. Penalties using second-order differences were studied in Green and Silverman (1993) and elsewhere. In addition, we at the same time select parameter estimations and financial information to be incorporated into the model by adding a penalty that selects variables.

$$l_\lambda(\theta) = l(\theta) - \sum_{j=0}^{J} \frac{n_j \lambda_\omega}{2} \boldsymbol{\omega}_j' \boldsymbol{K}_j \boldsymbol{\omega}_j - n \cdot \sum_{j=1}^{p} p_{\lambda_{\beta_j}}(|\beta_j|). \tag{2.2.4}$$

Here λ_ω is a normalization parameter, the effect of which is to increase the stability of the estimation. The penalty term is as follows:

- Penalty using second-order difference

$$\frac{n\lambda_\omega}{2} \boldsymbol{\omega}' \boldsymbol{K}_j \boldsymbol{\omega} \quad \text{where,} \quad \boldsymbol{\omega}' \boldsymbol{K}_j \boldsymbol{\omega} = \sum_{k=2}^{m_j} (\Delta^2 \omega_{jk})^2.$$

- Hard thresholding penalty

$$p_{\lambda_{\beta_j}}(|\theta|) = \lambda_{\beta_j}^2 - (|\theta| - \lambda_{\beta_j})^2 I(|\theta| < \lambda_{\beta_j}), \hat{\theta} = zI(|z| > \lambda_{\beta_j}).$$

Here z is a value estimated using the maximum-likelihood method with penalty. We let the order of the array $\boldsymbol{K}_j$ be $m_j - r_j$ (for example, see Green (1987)). The penalty using a second-order difference has the effect of applying a penalty to the interest rate/hazard curve, which suppresses extreme fluctuations in the term structure. The other penalty term, the hard thresholding penalty, is a penalty term for variable selection, which takes on a value of zero if the obtained estimated value β_j is less than λ_{β_j}, acting to eliminate spurious financial ratios from the set of explanatory variables.

There is no need for the penalty functions $p_{\lambda_{\beta_j}}$ to be the same for all $\beta_j,\ j = 1, \ldots, p$. For example, it is possible for the penalty to be small for variables believed to be necessary to the model, or large for variables not believed to be necessary. In this study, we used the same penalty function for simulations and data analysis to simplify parameter estimation. Currently, we let the L_q penalty function be $p_\lambda(|\theta|) = \lambda|\theta|^q$. Many penalty functions included in this class of penalty functions have been proposed. The $q = 1$ penalty is the least absolute shrinkage and solution operator (LASSO) proposed by Tibshirani (1996) and Tibshirani (1997), the $q = 2$ penalty is ridge estimation, and there is the elastic net method of Zou and Hastie (2005), combining $q = 1, 2$.

The estimated parameter value $\hat{\boldsymbol{\theta}}$ is estimated by numerical optimization, and the bond pricing model is determined by this parameter. The interest rate and hazard term structures to be estimated are further implicitly derived from bond market data. This means that the estimates are probabilities in a risk-neutral world, and can be applied to pricing derivatives.

However, studies by Altman (1989) and others have shown that the levels of hazard in a risk neutral world derived from bond prices are higher than the real-world levels, as calculated from historical default data because the liquidity of corporate bonds is relatively lower than that of government bonds, so that market participants incorporate a liquidity premium into their transaction prices for corporate bonds, and that buyers and sellers of corporate bonds consider the possibility of a decline in bond prices exceeding anything observed in historical data.

2.2.4 Selecting the number m, variance parameter s, and penalty parameter $(\lambda_\omega, \lambda_\beta)$ for the basis function

When estimating the parameter $\boldsymbol{\theta}$, it is necessary to determine the count m; variance parameter s; and $(\lambda_\omega, \lambda_\beta)$ for the basis function. Conceivable criteria for evaluating the model include the Bayesian information criteria of Konishi et al. (2004) the generalized information criteria of Konishi and Kitagawa (1996); the cross-validation method of Stone (1974); and the bootstrap method of Efron and Tibshirani (1994). In our proposed method, we estimate parameters via numerical optimization, which increases the amount of computation due to repeating the evaluation of (2.2.4) , but by using Bayesian information criteria we are able to appropriately estimate the count, variance parameter, and penalty parameter of the basis function.

The Bayesian information criteria (BIC) has been proposed by Schwarz (1978). BIC is predicated on estimating parameters using the maximum-likelihood method, and is not applicable to a model whose parameters are estimated using a maximum-likelihood method with penalty. We therefore employ a Bayesian information criteria adapted to parameter estimation based on the maximum-likelihood method with penalty, using Laplace approximation (Tierney and Kadane (1986)).

$$\begin{aligned} BIC &= -2l(\hat{\boldsymbol{\theta}}) + n\sum_{j=0}^{J}\lambda_j \hat{\boldsymbol{\omega}}_j' \boldsymbol{K}_j \hat{\boldsymbol{\omega}} + n\hat{\boldsymbol{\beta}}'\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)\hat{\boldsymbol{\beta}}' - \sum_{j=0}^{J}\log|\boldsymbol{K}_j|_+ \\ & \quad - \log|\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)|_+ + \log|J(\hat{\boldsymbol{\theta}})| - \sum_{j=1}^{J}(m_j - r_j)\log\lambda_j, \qquad (2.2.5) \end{aligned}$$

where $J(\hat{\boldsymbol{\theta}}) = -\frac{1}{n}\partial^2 l_\lambda(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'|_{\hat{\boldsymbol{\theta}}}$ and, $m(J) = \sum_{j=0}^{J}(m_j - r_j)$. $\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)$ is defined as follows,

$$\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0) = \mathrm{diag}\{p'_{\lambda_{\beta_1}}(|\beta_{10}|/|\beta_{10}|), \ldots, p'_{\lambda_{\beta_p}}(|\beta_{p0}|/|\beta_{p0}|)\}.$$

$|\boldsymbol{K}_j|_+$ and $|\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)|_+$ are respective products of the non-zero eigenvalues of constant matrix $\boldsymbol{K}_j$, $\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)$ with rank $m_j - r_j$, $p - q_{\lambda_\beta}$, and q_{λ_β} is the number of variables eliminated by variable selection. This minimization of BIC determines the basis functions number $\hat{m}$; variance parameter $\hat{s}^2$; and penalty parameters $(\hat{\lambda}_\omega, \hat{\lambda}_\beta)$.

2.3 Simulation

Henceforth, when simulating and analyzing actual data we shall unify the basis function number m_j into a single m independent of rating. We performed a simulation to validate that the correct variables are selected by the method for proposing financial ratios to be included in the model. We generated bond prices by using the following model, and estimated the parameters,

$$\begin{aligned} y(j) &= \sum_{i=1}^{L} C \cdot \exp\left[- \int_0^{t_{ji}} \{ r(u) + k \cdot \exp(\boldsymbol{\beta}' \boldsymbol{x}_j \} du \right] \\ &\quad + R \cdot \exp\left[- \int_0^{t_{jL}} \{ r(u) + k \cdot \exp(\boldsymbol{\beta}' \boldsymbol{x}_j \} du \right] + e(j), \quad j = 1, ..., n, \end{aligned}$$

here $e(j)$ is a probability variable independently obeying a normal distribution. For $r(t)$, we provided an a priori term structure as in Figure 2.2, estimated from government bond prices. Parameter settings were $k = 0.02$ and $\boldsymbol{\beta} = (0.3, 0, 0, 0, 0, -0.4, 0, 0, 0.2, 0)'$. For financial ratio $\boldsymbol{x}$, we used a multi-dimensional normal random variable, based on the variance-covariance matrix in Table 2.3 , derived from actual financial ratios (normalized to mean of 0 and standard deviation of 1). We also performed an estimation using 55 variables incorporating the interactions between each financial ratio, and an estimation using a contaminated normal distribution with error terms $e(j)$ as follows:

$$(1 - \varepsilon) N(0, 1) + \varepsilon N(0, \kappa^2).$$

For the simulation, we set $\varepsilon = 0.10$ and $\kappa = 5$. For the estimation using 55 variables, we set the parameter vector $\boldsymbol{\beta}$ as follows:

$$\beta_i = \begin{cases} -0.3, & i = 8, 16, 21, \\ -0.2, & i = 7, 17, 27, \\ 0.2, & i = 13.53, \\ 0.4, & i = 35, 41, 46, \\ 0, & other. \end{cases}$$

In other words, of the 55 financial ratio variables including interactions, 11 variables influenced risk premiums, and 44 variables did not. Under these settings, we compared the maximum-likelihood method ML without variable selection, and the maximum-likelihood method $ML + HT$ which included a penalty term for variable selection. The estimation methodology ML and $ML+HT$ estimated parameters which maximized the following likelihood function:

$$ML : l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(y_i|\boldsymbol{t}_i; \boldsymbol{x}_{i_j}; \boldsymbol{\beta}, \sigma^2),$$

$$ML + HT : l_\lambda(\theta) = l(\theta) - n \cdot \sum_{l=1}^{p} p_{\lambda_\beta}(|\beta_l|).$$

We conducted the simulation 1000 times, and selected the model with basis function number $\hat{m}$ which minimized the value of BIC, and variance parameter $\hat{s}^2$, $\hat{\lambda}_\omega$, $\hat{\lambda}_\beta$.

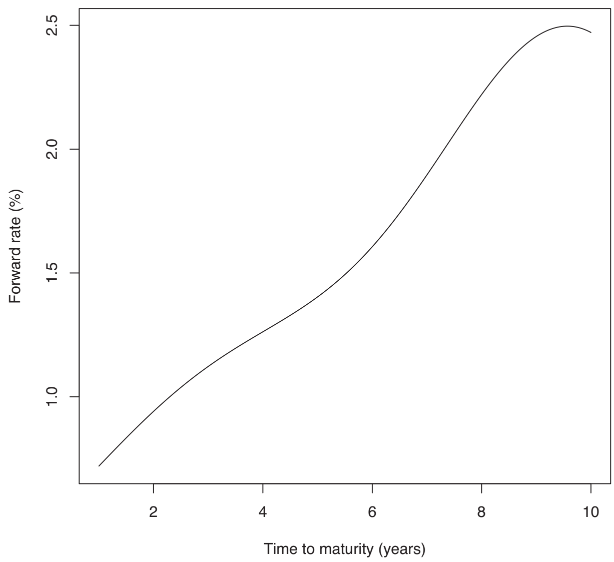

Figure 2.2: Term structure of forward rates for government bonds

Table 2.1 shows $RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K} ||\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}||^2}$ and the average number of explanatory variables eliminated. K is the number of simulations, 1000. Parenthesized numbers are standard deviations. From the root mean square error ($RMSE$) results in Table 2.1, when using either a normal distribution or contaminated normal distribution for the error term, the value of $RMSE$ and its standard deviation is lower when the number of data points $n = 200$, compared with 100, indicating that estimation accuracy rises with the number of samples.

The accuracy of the estimation results is worse for the contaminated normal distribution than for the normal distribution. We identified that estimation accuracy declines when the number of variables is increased from 10 to 55. Observing the differences in estimation methodology, the accuracy of the $RMSE$ value is markedly higher for $ML+HT$ with variable selection. However, considering that variability increased in the case of 55 variables, the problem of the stability of estimation accuracy remains when the number of explanatory variables used in the analysis increases. The fact that the variability in ML is less with 55 variables than with 10 is also due to the parameter settings. Furthermore, in the case of variable selection with $ML + HT$, looking at the number of eliminated parameters, in the case of a sample size of 100 and 10 variables the number of parameters with values of zero is seven, whereas the average number of variables eliminated in the case of using an error term with normal distribution or contaminated normal distribution is 7.06 and 7.03 respectively, showing no effect on the tails of the distribution. We find that the variability in the number of eliminated variables in the case of the contaminated normal distribution is greater than for the normal distribution, and the estimation accuracy is diminished. However, increasing the number of samples to 200 does not meaningfully increase the variability. Whereas with sample size of 100 and 55 variables the number of eliminated variables is 44, the actual average number of eliminated variables as 44.57 and 44.73 for the normal distribution and contaminated normal distribution respectively, showing a slight upward bias. The results thus did not improve even if the sample size was increased to 200. Variability shows the same tendency as with 10 variables.

Table 2.2 shows the proportion of variables eliminated in 1000 simulation runs in the case of 10 explanatory variables. In this table, looking at β_9 in the case of a sample size of 100 for the normal distribution, whereas we would have expected an estimate of 0.2, the proportion of cases where variables were mistakenly selected, and its value was instead 0, was 99% for $n = 100$ and 5% for $n = 200$. This may be due to the fact that its absolute value is close to 0. We see however, that for other variables, the variables are selected nearly 100% of the time. With a sample size of 100 and a contaminated normal distribution, there are 1-3% of cases where a variable that should have been eliminated was not, and 1-14% of cases where a variable which should not have been eliminated was instead mistakenly eliminated; the accuracy of variable selection is thus less than for the normal distribution. Increasing the sample size to 200 improves the ratio of incorrect selections for β_9, and we see that for the contaminated normal distribution as well the correct variables are selected with a probability approaching 1.

The results of this simulation show us that variable selection for financial ratios has a pronounced impact on estimation accuracy and the interpretation of the results of the estimation. This confirms that the variable selection based on the maximum-likelihood method with penalty as proposed in this study allows us to obtain superior estimation results.

Table 2.1: Simulation results 1: $RMSE$ of $\hat{\beta}$

		$RMSE$ Normal distribution	$RMSE$ Contaminated normal distribution	Number of variables eliminated Normal distribution	Number of variables eliminated Contaminated normal distribution
$n = 100$					
10 variables	ML	0.182(0.069)	0.303(0.157)		
	ML+HT	0.053(0.065)	0.106(0.116)	7.06(0.27)	7.03(0.50)
55 variables	ML	0.290(0.022)	0.339(0.032)		
	ML+HT	0.164(0.153)	0.198(0.161)	44.57(1.10)	44.73(1.26)
$n = 200$					
10 variables	ML	0.140(0.054)	0.210(0.087)		
	ML+HT	0.032(0.039)	0.054(0.054)	7.03(0.19)	7.03(0.23)
55 variables	ML	0.286(0.022)	0.336(0.032)		
	ML+HT	0.134(0.134)	0.146(0.132)	44.45(0.88)	44.46(0.92)

Table 2.2: Simulation results 2: proportion of variables eliminated

		β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
Absolute value		0.3	0	0	0	0	-0.4	0	0	0.2	0
$n = 100$	normal distribution	0	1	1	1	0.99	0	1	1	0.08	0.99
	Contaminated normal distribution	0.01	0.98	0.99	1	0.97	0.01	0.98	0.98	0.14	0.98
$n = 200$	normal distribution	0	1	1	1	1	0	1	1	0.04	1
	Contaminated normal distribution	0	1	1	1	1	0	1	1	0.05	0.99

Table 2.3: Variance-covariance matrix

	Revenues	Total assets	Shareholders' equity	Shareholders' equity to total assets	ROA	ROE	Pretax profit to total assets	Operating margin	FCF on revenues	CF margin
Revenues	1.00	0.66	0.64	−0.04	0.21	0.24	0.18	−0.06	0.02	−0.04
Total assets	0.66	1.00	0.90	−0.14	0.17	0.21	0.06	0.19	−0.11	−0.06
Shareholders' equity	0.64	0.90	1.00	0.19	0.21	0.18	0.18	0.14	0.03	0.04
Shareholders' equity to total assets	−0.04	−0.14	0.19	1.00	0.30	0.16	0.47	0.09	0.18	0.11
ROA	0.21	0.17	0.21	0.30	1.00	0.94	0.87	0.83	−0.16	−0.10
ROE	0.24	0.21	0.18	0.16	0.94	1.00	0.80	0.83	−0.23	−0.15
Pretax profit to total assets	0.18	0.06	0.18	0.47	0.87	0.80	1.00	0.70	−0.09	−0.05
Operating margin	−0.06	0.19	0.14	0.09	0.83	0.83	0.70	1.00	−0.26	−0.11
FCF on revenues	0.02	−0.11	0.03	0.18	−0.16	−0.23	−0.09	−0.26	1.00	0.93
CF margin	−0.04	−0.06	0.04	0.11	−0.10	−0.15	−0.05	−0.11	0.93	1.00

2.4 Empirical Analysis

2.4.1 Description of the data

Let us explain the data used in this study. For government bonds and corporate bonds, we used the Japan Securities Dealers Association's over-the-counter purchase and sale reference statistical data from October 25, 2007. We used Moody's data for the same date for credit ratings. We performed our estimates using data for 1,225 issues in segments ranging from manufacturing, trading, finance, communications, and energy to transportation. We used $J = 4$ for the total number of ratings J in the bond pricing model, divided into four sectors: sector one for Aaa and Aa, comprising 494 issues; sector two for A, comprising 336 issues; sector three for Baa, comprising 331 issues; sector four for Ba, considered non-investment grade, comprising 16 issues.

For financial information, we used the 10 measures below, indexes of company size, financial health, and profitability. They are revenue; ROA; ROE; shareholders' equity to total assets; pretax profit to total assets; total assets; shareholders' equity; operating margin; ratio of free cash to sales; and cash flow margin (operating cash flows divided by revenues). We used the data for the most recent closing for each issuing company, taking the data from Nikkei NEEDS.

2.4.2 Results and observations

Table 2.4 shows the estimates of the $\boldsymbol{\beta}$ coefficient for the financial ratios.

For comparison, we show the estimates from PML, without variable selection, and from $PML + HT$, with variable selection.

$$PML : l_\lambda(\theta) = l(\theta) - \sum_{j=0}^{J} \frac{n_j \lambda_\omega}{2} \boldsymbol{\omega}'_j \boldsymbol{K}_j \boldsymbol{\omega}_j,$$

$$PML + HT : l_\lambda(\theta) = l(\theta) - \sum_{j=0}^{J} \frac{n_j \lambda_\omega}{2} \boldsymbol{\omega}'_j \boldsymbol{K}_j \boldsymbol{\omega}_j - n \cdot \sum_{j=1}^{p} p_{\lambda_\beta}(|\beta_j|).$$

When we optimized the basis function number m, variance parameter s, and penalty parameters $(\lambda_\omega, \lambda_\beta)$ based on BIC, the results were $\hat{m} = 8$, $\hat{s} = 3.1$, $\hat{\lambda}_\omega =$

Table 2.4: Estimates of $\boldsymbol{\beta}$ coefficients for financial ratios

	PML	PML+HT
Revenues	-0.176(0.032)	0
Total assets	0.156(0.046)	0
Shareholders' equity	-0.022(0.040)	0
Shareholders' equity to total assets	0.044(0.026)	0
ROA	-0.019(0.017)	0
ROE	-0.332(0.018)	-0.38(0.003)
Pretax profit to total assets	-0.015(0.032)	0
Operating margin	-0.038(0.025)	0
FCF as ratio of sales	-0.124(0.071)	0
CF margin	0.066(0.066)	-0.077(0.006)

0.11, $\hat{\lambda}_\beta = 0.15$. The $PML + HT$ estimation method, which selects variables, chose only ROE and CF margins. Performing variable selection allowed us to obtain a model with improved and more complex characteristics. We also learned through variable solution that ROE and CF margins influence risk premium. The coefficient for both financial ratios is negative, meaning that an increase in ROE and CF margins reduces the risk premium. Both selected financial ratios indicate profitability, and since this indicates that higher profitability reduces risk premium, the results seem plausible. We observe that financial ratios indicating profitability were selected while those indicating company size or financial health were not; market participants emphasized issuer profitability when pricing bonds for issuers with the same credit rating. Let us show the results of estimating the term structure for risk-free interest rates and risk premiums obtained from our proposed method.

Figure 2.3 shows the forward rates $r(t)$ and $r(t)+s(t)$ in (2.2.2) for government and corporate bonds. The figures indicate that the forward rate levels are lowest for government bonds ($r(t)$ in the figure), then grow as the credit rating worsens, from ratings of Aaa and Aa, to A, to Baa, then to Ba. The reliability of the estimates for the Ba rating is low, since there are only three companies with the rating; we show them merely for reference.

Figure 2.4 shows the risk premium $s(t)$ in (2.2.2) for corporate bonds. These figures show low levels of ratings for Aaa, Aa, and A for all terms. For the Baa rating, until around year 5 the values are close for the higher rankings, but then show a tendency to climb in the second half. This represents the fact that the market is assigning value to short-term stability of A-rated bonds, and that higher-rated bonds are assigned more value for their long-term stability.

Figures 2.5-2.8 show forward rates considering financial ratios that are the results from estimating using $PML + HT$. We see that the level risk premium increases or decreases depending on financial ratios. In addition, the lower ratings have wider spreads in their upward and downward movements because lower the rating, greater the variability in financial ratios for companies with the same rating.

Table 2.5 shows the results of estimating the error variance. The table shows that more risky bonds have larger error variances. This results suggests that higher the risk on a bond, greater the variability in market expectations. In addition, the large error variance for Ba ratings may be due to instability in the estimates, since the sample size was only 16.

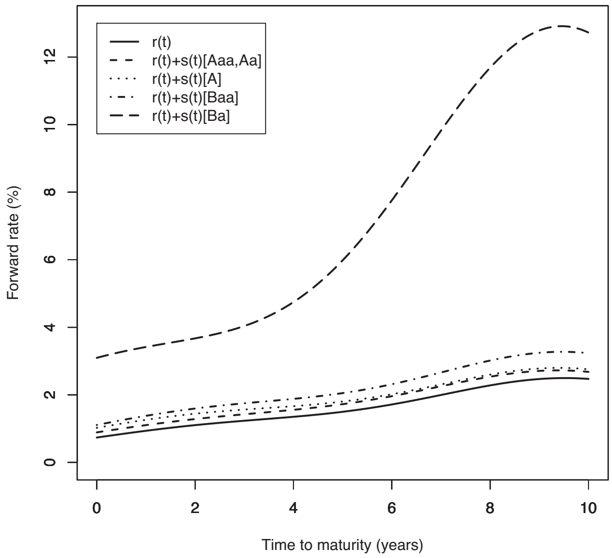

Figure 2.3: Forward rates for government and corporate bonds

Table 2.5: Estimates of error variance

	JGB	Aaa,Aa	A	Baa	Ba
Estimated error variances $\hat{\sigma}^2$	0.065	0.147	0.309	0.570	2.079

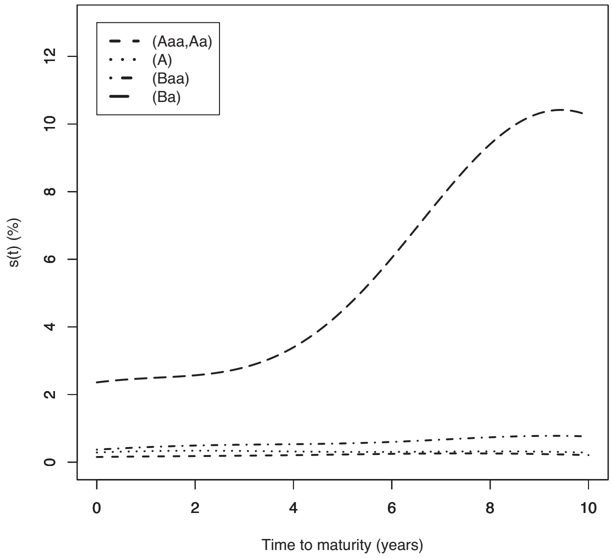

Figure 2.4: Corporate bond risk premiums

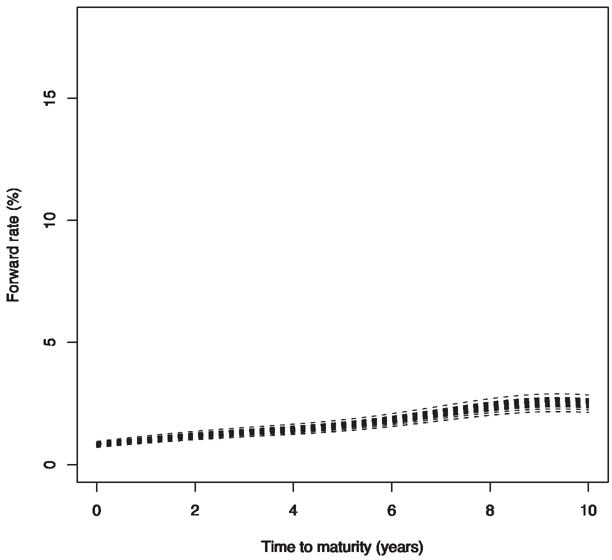

Figure 2.5: Forward rates considering financial ratios

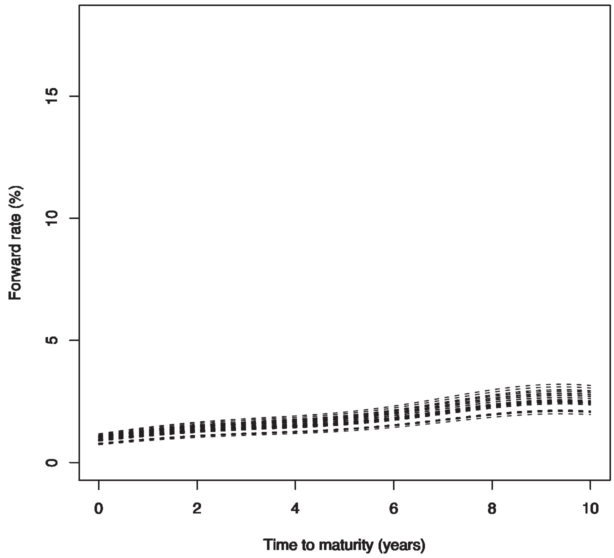

Figure 2.6: Forward rates considering financial ratios (A)

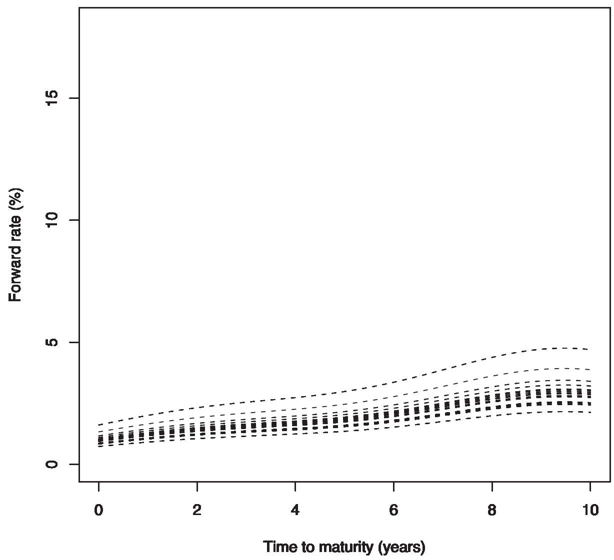

Figure 2.7: Forward rates considering financial ratios (Baa)

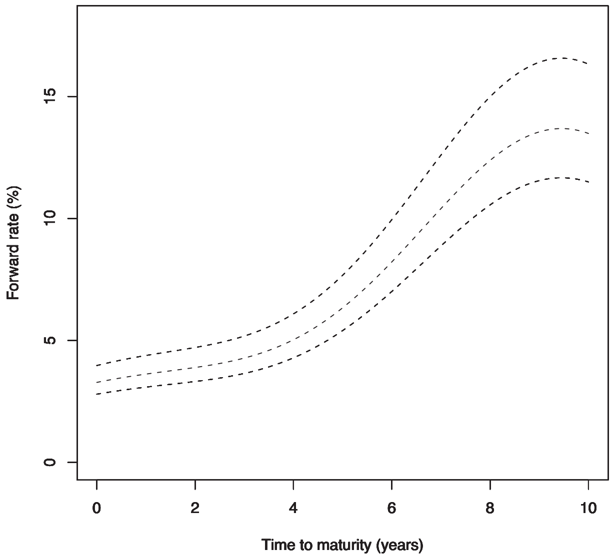

Figure 2.8: Forward rates considering financial ratios (Ba)

2.5 Conclusions and Future Topics

In this study we expressed the term structure of risk-free interest rates and hazards using basis functions. Basis functions have been used since McCulloch (1971) to estimate interest rate term structures; however, this approach was plagued by instability in the estimation curves. This study uses the maximum-likelihood method with penalties to estimate the parameters, allowing for the successful estimation of stable term structures. With regard to hazard term structures, we incorporated financial ratios into the hazards for each rating, thus allowing us to express the different term structures depending on the issuing company, even for bonds with the same rating. Incorporating a penalty term which performs variable selection for the financial information, we were able to eliminate unnecessary financial information from the model.

One area of interest, which we did not address, concerns the quantification of liquidity risk. In terms of topics for future research, our hope is that as the corporate bond market develops and as more default data is gathered, it will become possible to construct more accurate credit risk models.

2.6 Appendix

Here, we derive the BIC for (2.2.5) in Section 2.2. It is impossible to solve the maximum-likelihood method with penalty, since the penalty function for variable selection is not differentiable at the origin, and the second-order differential does not exist. However, it is possible to use a second-degree local approximation for the penalty function for variable selection, following Fan and Li (2001). We explain that approach here. Starting with an initial value $\boldsymbol{\beta}^{(0)}$, which gives a value close to the minimum of the maximum-likelihood function, if $|\beta_j^{(0)}| < \eta$, we let $\hat{\beta}_j = 0$, and if $|\beta_j^{(0)}| \geq \eta$, we use the following equation to approximate the first derivative of the penalty function:

$$[p_{\lambda_{\beta_j}}(|\beta_j|)]' = p'_{\lambda_{\beta_j}} \mathrm{sgn}(\beta_j) \approx \{p'_{\lambda_{\beta_j}}(|\beta_{j0}|/|\beta_{j0}|)\}\beta_j.$$

Using this approach, for $\beta_j \approx \beta_{j0}$, the penalty function is given by the following second-order approximation:

$$p_{\lambda_{\beta_j}}(|\beta_j|)] \approx p_{\lambda_{\beta_j}}(|\beta_{j0}|) + \frac{1}{2}\{p'_{\lambda_{\beta_j}}(|\beta_{j0}|/|\beta_{j0}|)\}(\beta_j^2 - \beta_{j0}^2).$$

Then, it follows that $l(\theta) - n\sum_{j=1}^{d} p_{\lambda_{\beta_j}}(|\beta_j|)$ can be approximated, omitting the constant term, by the following formula:

$$l(\theta) - \frac{n}{2}\boldsymbol{\beta}'\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)\boldsymbol{\beta}.$$

Here, $\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)$ is defined as follows:

$$\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0) = \mathrm{diag}\{p'_{\lambda_{\beta_1}}(|\beta_{10}|/|\beta_{10}|), \ldots, p'_{\lambda_{\beta_p}}(|\beta_{p0}|/|\beta_{p0}|)\}.$$

From this, we can express the logarithmic maximum-likelihood function with penalty from Eq. (2.4) as follows:

$$\begin{aligned} & l_\lambda(\boldsymbol{\theta}) \\ \approx\ & \log l(\boldsymbol{\theta}) + \log\left\{\exp\left(-\sum_{j=0}^{J}\frac{n_j\lambda_\omega}{2}\boldsymbol{\omega}_j'\boldsymbol{K}_j\boldsymbol{\omega}_j\right)\right\} + \log\left\{\exp\left(-\frac{n}{2}\boldsymbol{\beta}'\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)\boldsymbol{\beta}\right)\right\} \\ =\ & \log\left\{l(\boldsymbol{\theta})\exp\left(-\sum_{j=0}^{J}\frac{n_j\lambda_\omega}{2}\boldsymbol{\omega}_j'\boldsymbol{K}_j\boldsymbol{\omega}_j\right)\exp\left(-\frac{n}{2}\boldsymbol{\beta}'\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)\boldsymbol{\beta}\right)\right\}. \end{aligned}$$

Here, if we view the exponential function term as a degenerate normal distribution of dimension $m_j - r_j, j = 0, \ldots, J$ and dimension $p - d_{\lambda_\beta}$, with mean vector $\mathbf{0}$, we can add an error term to make it a density function, and represent it as an a priori distribution of $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}$:

$$\pi(\boldsymbol{\omega}_j|\lambda_\omega) = \left(\frac{n\lambda_\omega}{2\pi}\right)^{(m_j-r_j)/2}|\boldsymbol{K}_j|_+^{1/2}\exp\left(-\frac{n\lambda_\omega}{2}\boldsymbol{\omega}'\boldsymbol{K}_j\boldsymbol{\omega}\right), \tag{2.6.1}$$

$$\pi(\boldsymbol{\beta}|\lambda_\beta) = \left(\frac{n}{2\pi}\right)^{(p-q_{\lambda_\beta})/2}|\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)|_+^{1/2}\exp\left(-\frac{n}{2}\boldsymbol{\beta}'\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)\boldsymbol{\beta}\right). \tag{2.6.2}$$

Here, q_{λ_β} is the number of variables eliminated by variable selection, and is determined by the dependency on λ_β. Then, $|\boldsymbol{K}_j|_+$ and$|\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)|_+$ are the respective products of the non-zero eigenvalues of constant matrix $\boldsymbol{K}_j$ and $\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)$, of rank $m_j - r_j$ and $p - q_{\lambda_\beta}$, respectively.

The marginal likelihood of the model with a priori distributions with $\boldsymbol{\theta}$ of the m_j- dimension normal distribution of (2.6.1) and the p-dimension normal distribution of (2.6.2) can be expressed as follows:

$$
\begin{aligned}
& \int l(\boldsymbol{\theta}) \prod_{j=0}^{J} \pi(\boldsymbol{\omega}_j|\lambda_\omega)\pi(\boldsymbol{\beta}|\lambda_\beta)d\boldsymbol{\theta} \\
= & \int \exp\left[n \times \frac{1}{n}\log\left\{l(\boldsymbol{\theta}) \prod_{j=0}^{J} \pi(\boldsymbol{\omega}_j|\lambda_\omega)\pi(\boldsymbol{\beta}|\lambda_\beta)\right\}\right] d\boldsymbol{\theta} \\
= & \int \exp\{nq(\boldsymbol{\theta}|\lambda_\omega, \lambda_\beta)\}.
\end{aligned}
$$

Here, the mode of

$$
q(\boldsymbol{\theta}|\lambda_\omega, \lambda_\beta) = \frac{1}{n}\left\{\log l(\theta) + \sum_{j=0}^{J} \log \pi(\boldsymbol{\omega}_j|\lambda_\omega) + \log \pi(\boldsymbol{\beta}|\lambda_\beta)\right\}
$$

is equal to the estimated value of the logarithmic maximum-likelihood function with penalty, so that the Laplace approximation of the integral can be used to evaluate the BIC as follows (e.g., see Konishi and Kitagawa (2004), pp. 153):

$$
\begin{aligned}
BIC &= -2\log\left\{\int l(\boldsymbol{\theta}) \prod_{j=0}^{J} \pi(\boldsymbol{\omega}_j|\lambda_\omega)\pi(\boldsymbol{\beta}|\lambda_\beta)\right\} d\boldsymbol{\theta} \\
&= -2\log\left\{\int \exp(nq(\boldsymbol{\theta}))d\boldsymbol{\theta}\right\} \\
&\approx -2\log\left\{\frac{2\pi^{(m(J)+p-d_{\lambda_\beta})/2}}{n^{(m(J)+p-d_{\lambda_\beta})/2}|J(\hat{\boldsymbol{\theta}})|^{1/2}}\exp(nq(\hat{\boldsymbol{\theta}}))\right\} \\
&= -2l(\hat{\boldsymbol{\theta}}) + n\sum_{j=0}^{J} \lambda_j \hat{\boldsymbol{\omega}}_j' \boldsymbol{K}_j \hat{\boldsymbol{\omega}} + n\hat{\boldsymbol{\beta}}'\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)\hat{\boldsymbol{\beta}}' - \sum_{j=0}^{J} \log|\boldsymbol{K}_j|_+ \\
&\quad - \log|\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)|_+ + \log|J(\hat{\boldsymbol{\theta}})| - \sum_{j=1}^{J}(m_j - r_j)\log\lambda_j,
\end{aligned}
$$

where $J(\hat{\boldsymbol{\theta}}) = -\frac{1}{n}\partial^2 l_\lambda(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'|_{\hat{\boldsymbol{\theta}}}$ and $m(J) = \sum_{j=0}^{J}(m_j - r_j)$. When actually using the BIC, $\Sigma_{\lambda_\beta}(\boldsymbol{\beta}_0)$ is approximated by $\Sigma_{\lambda_\beta}(\boldsymbol{\beta}^{(0)})$ or $\Sigma_{\lambda_\beta}(\hat{\boldsymbol{\beta}})$ using the initial value $\boldsymbol{\beta}^{(0)}$. This gives a value close to the minimum of the logarithmic maximum-likelihood function or estimated $\hat{\boldsymbol{\beta}}$.

3

Issue II - Forecasting Loss Given Default for Bank Loans : An Empirical Analysis for Japanese Banks

3.1 Introduction

The Basel II/III Accord allows banks to estimate credit risk capital requirements using an Internal Ratings-Based (IRB) approach. The probability of default (PD) and loss given default (LGD) are the most important credit risk parameters in an IRB approach. If banks select the Foundations Internal Ratings-Based (FIRB) approach, there is no need for a proprietary LGD predictive model. However, if banks select the Advanced Internal Ratings-Based (AIRB) approach, they are required to build a proprietary predictive model for LGD. Studies on LGD can be divided into two groups according to the data they use. One group is based on the corporate bond market, and the other group is based on bank loan losses. The majority of LGD research is based on corporate bond markets. Only a few LGD studies are based on bank loan losses. This is because bank loans are private instruments and, thus, limited bank loan data are publicly available.

The following systems and business practices are unique to Japan.

A unique credit guarantee system: National Federation of Credit Guarantee Corporations are unique to Japan and were established to smooth the financing of small and medium enterprises (SMEs). SMEs typically experience difficulty obtaining financing from banks because SME loans are considered high risk. The National Federation of Credit Guarantee Corporations guarantees the loans for SMEs.

Complex security agreements: Open-ended collateral represents a Japanese bank practice. Normal collateral secures a certain loan, whereas open-ended collateral secures multiple loans, which includes loans to be contracted in the future for a certain borrower. It is difficult to reasonably allocate open-ended collateral to each loan, and analyses of LGD have grown in complexity as a result.

Additional financing for defaulted borrowers: Japanese banks do not start to collect a debt as soon as a borrower defaults. Japanese banks often provide assistance, such as additional financing for defaulted borrowers. Additional financing complicates the determination of cash flow collection and the calculation of loss amounts.

Duration of the workout process: In Japan, the duration of the workout process can vary from months to years. Hence, a considerable amount of censored data exists in a data set, which may cause bias in an analysis. Therefore, the appropriate treatment of censored data is a significant factor in LGD analysis.

When building a predictive model of LGD based on banks' actual workout data, we must consider the aforementioned factors. However, a predictive LGD model that considers these characteristics has not been proposed in previous research, although Itoh and Yamashita (2008) and Miura et al. (2010) conducted LGD studies in the Japanese context and provide knowledge of LGD in Japan.

We analyze LGD using the data provided by three Japanese banks. The aims of this study are as follows. First, we compare the level of LGD in Japan with other countries. Here, the Japanese bank loan LGD is of particular interest. Second, we analyze the factors that influence Japanese LGD. Third, we develop an expected loss (EL) predictive model, composed of the PD predictive model and the LGD predictive model.

The analysis shows that the average value of LGD in Japan is approximately 9%. This is significantly lower than the LGD levels reported in other countries (25 to 58%). Collateral quotas, credit guarantee quotas, exposure at default (EAD), and the duration of the workout process are significant influencers of LGD. The PD predictive model finds that, in addition to borrower characteristics, loan characteristics such as collateral and credit guarantees influence PD.

This study builds on the work of Kawada and Yamashita (2013). Here, we further investigate the factors influencing LGD and improve the LGD predictive model proposed in Kawada and Yamashita (2013).

The remainder of this chapter is structured as follows. Section 3.2 presents a literature review on modeling LGD methods. Section 3.3 describes the data set of the bank loans analyzed in this study. Section 3.4 discusses the factors influencing LGD. In Section 3.5, we propose an EL predictive model that is composed of the PD and the LGD predictive models. Section 3.6 evaluates the predictive accuracy of the LGD predictive model. Finally, Section 3.7 concludes the chapter.

3.2 Literature Review

This section presents a review of related literatures. The majority of the LGD research is based on corporate bond markets rather than bank loan losses. Altman (2006) provides a comprehensive survey of previous studies based on the corporate bond market. Acharya et al. (2003), Altman (1989), Altman and Eberhart (1994), and Nickell et al. (2000) obtained LGD data from the market price of defaulted bonds. Then, Jarrow (2001) and Yamashita and Kihara (2004) proposed calculating an implied LGD from the stock prices of issuers.

Compared to studies based on corporate bond markets, those based on bank loan losses are scarce. This is because bank loans are private instruments and limited data are publicly available. We describe the previous research that focuses on bank loan losses in Table 3.1. The table shows the LGD levels in published empirical studies. Note that the definitions of LGD, countries, and data observations are different in each study.

Here, we summarize the relationship between LGD and the factors influencing LGD reported in the previous studies. Many previous studies reported that

Table 3.1: The LGD levels in published empirical studies

Authors	Country	Observation period	Number of observations	Average LGD
Asarnow and Edwards (1995)	United States	1970-1993	831	35%
Felsovalyi and Hurt (1998)	Latin America	1970-1996	1149	32%
Eales and Bosworth (1998)	Australia	1992-1995	5782	31%
Araten et al. (2004)	United States	1982-1999	3761	40%
Franks et al. (2004)	United Kingdom	1984-2003	1418	25%
Franks et al. (2004)	France	1984-2003	586	47%
Franks et al. (2004)	Germany	1984-2003	276	39%
Dermine and De Carvalho (2006)	Portugal	1995-2000	374	29%
Querci (2005)	Italy	1980-2004	15827	50%
Caselli et al. (2008)	Italy	1990-2004	11649	54%
Grunert and Weber (2009)	Germany	1992-2003	120	28%
Zhang and Thomas (2012)	United Kingdom	1987-2003	18972	58%

collateral, loan size, duration of recovery, risk premiums, and firm size have a significant influence on LGD. Others (e.g., Araten et al. (2004), Dermine and De Carvalho (2006), Grunert and Weber (2009), and Miura et al. (2010)) confirmed that collateral has a reduction effect on LGD. The relationship between LGD and loan size has often been the subject of research. Felsovalyi and Hurt (1998) and Dermine and De Carvalho (2006) reported that loan size has a negative impact on recovery rates. In contrast, Grunert and Weber (2009) confirmed that a greater loan size leads to lower LGD, although Miura et al. (2010) reported that there is no relationship between loan size and LGD. Gürtler and Hibbeln (2011) noted that the longer the duration of the workout process, the greater the LGD. Grunert and Weber (2009) investigated the relationship between LGD and a borrower's creditworthiness, measured by the risk premium in interest rates. The results showed that LGD for loans to borrowers with lower creditworthiness tends to be greater. The relationship between LGD and borrower company size has also been the subject of some studies. Grunert and Weber (2009) and Felsovalyi and Hurt (1998) reported that LGD increases with company size. However, Asarnow and Edwards (1995) found that the size of the borrower company has a positive effect on the recovery rate. The approaches to analyses vary among previous studies, and the effectiveness of each influencing factor depends on the method of analysis. Additionally, previous studies are based on certain bank data that simply capture that bank's idiosyncratic characteristics. Querci (2005) analyzed data on Italian bank loans and reported that none of the explanatory variables (i.e., loan forms, company activity regions, status of securities, and duration of the workout process) have sufficient explanatory power.

The relationship between LGD and business cycles has been the subject of research because the Basel Accord requires estimates of LGD downturn. Caselli et al. (2008) and Bellotti and Crook (2012) noted that macroeconomic variables can explain LGD. In contrast, Grunert and Weber (2009) and Felsovalyi and Hurt (1998) found no influence of business cycles on LGD. Caselli et al. (2008) reported that LGD for household loans has been affected by unemployment rates and household consumption, and that LGD for small and medium enterprise loans has been influenced by employment statistics and GDP growth rate. Bellotti and Crook (2012) analyzed credit card LGD and confirmed that banks' interest rates and unemployment rates impact LGD. Dermine and De Carvalho (2006) found that the explanatory power of year dummy variables is significant. However, the author did not discuss the relationship between business cycles and LGD, and there is, as yet, no consensus on this relationship.

Previous research conducted basic analyses (such as fundamental statistics, distributions, and influencing factors). More recent research has developed predictive LGD models and evaluated the performance of the developed models. LGD typically lies in the interval [0, 1] and is concentrated around 0 and 1. These LGD features imply that a simple linear regression model may not have a high level of predictive power. Thus, some previous studies proposed models that considered these LGD features. Bastos (2010) proposed dividing defaulted borrowers into homogeneous groups by repeating the binary logical determination. This binary logical determination uses loan size, interest rates, and internal ratings. The average LGD value of each group is obtained from the predicted LGD. The authors also confirmed that the proposed model has high explanatory power when compared with direct regression models. Matuszyk et al. (2010) divided defaulted borrowers into groups using the collections strategy used (i.e., in-house, using agent, and selling off the debt) and then built the regression model for each group. Superior LGD predictive models are different according to the employed data and validation criteria. Loterman et al. (2012) applied various parametric or nonparametric models to the data provided by six banks. The results showed that the optimal model depends on the data and performance metrics. The author also reported that nonparametric models and multi-stage models have a high level of overall explanatory power.

Miura et al. (2010), which is a study based on Japanese bank loans, is strongly related to the present study. Considering the extended duration of the workout process in Japan, the authors proposed incorporating elapsed time from a default into the model. Miura et al. (2010) was based on data provided by a single bank. In contrast, we analyze the data of three banks and develop the multi-stage model for predicting LGD using this data. The multistage model proposed in this study consists of binary decision models and a regression model. Applying the multi-stage model to LGD modeling has been proposed in previous studies (e.g., Lucas (2006), Gürtler and Hibbeln (2011), and Bellotti and Crook (2012)). Since these studies relate directly to this study, we introduce them in detail in Section 3.5.

3.3 Data

This section defines a default and LGD and describes the data employed in this study.

3.3.1 The definition of default

We define a default as a downgrade in a bank's internal borrower ratings. Banks typically have internal rating systems as follows:

- Non-default ratings
    1. Normal, performing borrowers.
    2. Performing borrowers with some future concerns.
- Default ratings
    3. Performing borrowers that require monitoring.
    4. Non-performing and probably irrecoverable borrowers.
    5. Practically uncollectible borrowers.

6. Uncollectible borrowers.

According to the Basel Capital Accord (Basel Committee on Banking Supervision (2006)), we set the default boundary between "Performing borrowers with some concerns for the future" and "Performing borrowers that require monitoring." In the actual bank's borrower internal rating system, each rating is subdivided into parts. "Normal performing borrowers" is subdivided into three to six parts. Typically, the bank's internal borrower rating systems have 10 or more ratings altogether. The use of a subdivided internal rating system facilitates a refined analysis of bank loan credit risk.

3.3.2 The definition of LGD

The units used in calculating LGD, the types of default resolutions, and the definition of loss are crucial to the definition of LGD.

Units for calculation of LGD

LGD can be calculated for each borrower or each loan contract. In this study, we calculate LGD for each borrower because open-ended collateral is a common Japanese banking practice. Normal collateral secures a certain loan, whereas open-ended collateral secures multiple loans, which include loans to be contracted in the future for a certain borrower. Since it is difficult to reasonably allocate open-ended collateral to each loan, we calculate LGD for each borrower. However, there are some disadvantages to calculating LGD for each borrower. For example, we cannot obtain estimates of LGD on each loan contract. Consequently, information used to make decisions on loan financing cannot be obtained from these estimates.

Default resolutions

There are two cases of default resolution. One is write-offs, the other is recoveries. Gürtler and Hibbeln (2011) also address default resolution cases.

- Write-offs: The bank gives up on being repaid and writes-off the loan contract from the account.

- Recoveries: The borrower upgrades to a non-default rating from a default rating.

The definition of loss

Loan losses can be defined by the amount of cash collected from defaulted borrowers or the write-off amount. Here, we define losses by the write-off amount. Typically, Japanese banks do not start to collect a debt as soon as a borrower defaults. Japanese banks sometimes provide assistance, such as additional financing for defaulted borrowers. A considerable number of defaulted borrowers upgrade to a non-default rating from a default rating. These characteristics complicate the definition of losses from collected cash flows. In contrast, defining losses by the write-off amount facilitates the reliable observation of realized losses. Thus, for accuracy, we define losses by the write-off amount in this study. The LGD calculation formula is expressed by the following formula:

$$LGD = \frac{\text{Total write-off amount}}{EAD}, \tag{3.3.1}$$

where total write-off amount is the sum of the write-off amount from the default until the end of the default, and EAD represents the exposure at default.

3.3.3 Description of the data

The data employed in this study contain loan information from three Japanese banks (Bank A, Bank B, and Bank C) from the year 2004 to the year 2011. The observation frequency is once every six months. Activity areas vary between banks. These data contain the following fields:

- Exposure: Total amount loaned to each borrower.
- Bank's internal borrower rating: Rating of the borrower.
- Write-off amount: The amount that the bank wrote-off from the account.
- Creditworthiness score: The score that indicates creditworthiness. A synthesized variable from borrower financial information used to estimate the probability of default.

- Collateral quota: The quota for each type of collateral (real estate, commercial bills, deposits, and marketable securities).
- Credit guarantee quota: The quota of credit guarantees from the National Federation of Credit Guarantee Corporations.
- The duration of the workout process: The duration between the default and the end of the workout process.

We compile two data sets from the data provided by the three banks. We use one data set to develop the PD predictive model. We refer to this data set as data set A. The other data set is used to develop the LGD predictive model. We refer to this data set as data set B.

3.3.4 Data set A (for developing the PD predictive model)

Data set A contains 679607 records for 81931 borrowers. This data set includes 8732 default records and contains both default and non-default records. We show the number of borrowers, records, and defaults in Table 3.2.

Table 3.2: The number of borrowers, records, and defaults

Number of borrowers	Number of records	Number of defaults
81931	679607	8732

We show the fundamental statistics of data set A in Table 3.3. The average credit worthiness score is 48.372. This score, which indicates the creditworthiness of the borrower, is adjusted so that the average value is 50 for all borrowers. Since data set A contains both defaulted borrowers and non-defaulted borrowers, this result is natural. Real estate accounts for the majority of the total collateral, with the average collateral quota (real estate) being 0.24. The average credit guarantee quota is 0.501. This credit guarantee contains the guarantee from the National Federation of Credit Guarantee Corporations only. The average exposure is 1.103 hundred million yen. The high standard deviation of exposure indicates that some large-scale borrowers increase the average value of exposures.

Table 3.3: Median, mean, and standard deviation of variables (data set A)

Variable	Median	Mean	SD
Creditworthiness score	49.000	48.372	16.860
Collateral quota (real estate)	0.000	0.240	0.485
Collateral quota (commercial bills)	0.000	0.050	0.174
Collateral quota (deposits)	0.000	0.019	0.127
Collateral quota (marketable securities)	0.000	0.004	0.066
Credit guarantee quota	0.513	0.501	0.439
Exposure in hundred million yen	0.203	1.103	5.502

We show the correlation matrix for data set A in Table 3.4. The creditworthiness score has a relatively high correlation (-0.125) with the default flag. This reflects that borrowers with a high creditworthiness score tend not to default. This is a natural result because the creditworthiness score represents the synthesized variable for the estimation of the probability of default. There is a negative correlation between the creditworthiness score and the credit guarantee quota. This may be because the bank requires loans for the low creditworthy borrower to be secured by credit guarantees.

Table 3.4: Correlation matrix of data set A

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Default flag	1	-0.125	-0.020	-0.013	-0.004	-0.002	0.029	0.020
(2) Creditworthiness score	-0.125	1	0.039	0.040	0.000	-0.008	-0.291	0.053
(3) Collateral quota (real estate)	-0.020	0.039	1	-0.042	0.004	0.006	-0.170	0.047
(4) Collateral quota (commercial bills)	-0.013	0.040	-0.042	1	-0.007	-0.006	-0.197	-0.021
(5) Collateral quota (deposits)	-0.004	0.000	0.004	-0.007	1	0.006	-0.084	0.000
(6) Collateral quota (marketable securities)	-0.002	-0.008	0.006	-0.006	0.006	1	-0.044	0.025
(7) Credit guarantee quota	0.029	-0.291	-0.170	-0.197	-0.084	-0.044	1	-0.423
(8) ln(Exposure)	0.020	0.053	0.047	-0.021	0.000	0.025	-0.423	1

3.3.5 Data set B (for the development of the LGD predictive model)

The management of censored data (workout proceeding data) becomes a problem in the analysis of LGD. Typically, the length of the workout process is over months or years. Consequently, a considerable amount of censored data exists in the data set, and an analysis based only on defaults with completed workout processes may contain bias. To avoid this bias, several techniques have been proposed in previous studies. Zhang and Thomas (2012) proposed applying the survival analysis technique to manage censored data. Gürtler and Hibbeln (2011) suggested using default data only with completed workout processes, but within a certain observation period to avoid the bias from observation period constraints. However, we do not consider this bias because the data observation period in this study is sufficiently long for the duration of the workout process. Thus, we analyze only those defaults with completed workout processes. However, if the observation period is not sufficiently long, it is necessary to consider this bias. We analyze 5664 defaults out of 8732 defaults with completed workout processes that occurred during the observation period. We show the number of each type of default resolution in Table 3.5. Table 3.5 shows that the majority are defaults without loss. More than half of the defaulted borrowers did not cause damage.

Table 3.5: The number of each type of default resolution

Recoveries	Write-offs	
	LGD=0	LGD>0
1334	3214	1116

We show the fundamental statistics of LGD in Table 3.6. The mean value of LGD is 0.089. The average value of LGD in Japan is significantly lower than the levels reported in previous studies in other countries.

We show the histogram of LGD in Figure 3.1. As Figure 3.1 shows, LGD on a large share of defaulted borrowers is concentrated near 0. This result is distinct from the results of many previous studies that have reported a bimodal distribution of LGD (e.g., Asarnow and Edwards (1995), Felsovalyi and Hurt

Table 3.6: Median, mean, and standard deviation of LGD

Median	Mean	SD
0.000	0.089	0.228

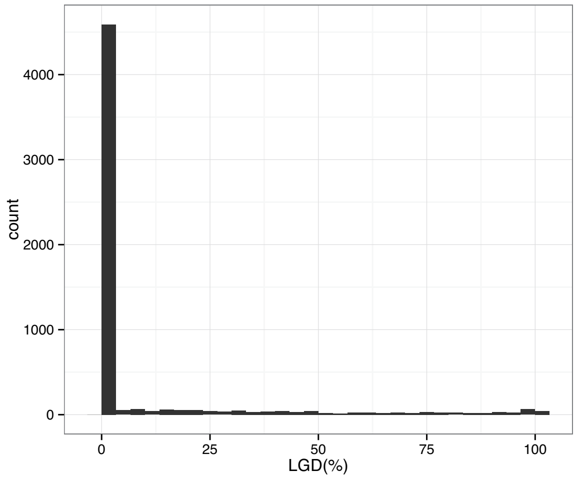

Figure 3.1: Histogram of LGD

(1998), Franks et al. (2004), Araten et al. (2004), Caselli et al. (2008), and Bastos (2010)). This result indicates that LGD cases of 100%, while found in other countries, are rare in Japan.

Table 3.7 shows the fundamental statistics of data set B. The average creditworthiness score is 30.849. The average creditworthiness score is relatively low in comparison with data set A because data set B consists of defaulted borrowers only. The majority of collateral is real estate, and the average value is 0.146. The credit guarantee quota is high (mean 0.628; median 0.824). The average value of EAD is 1.015 hundred million yen. The high standard deviation for EAD indicates that some large-scale borrowers increase the average value. The average duration of the workout process for all three banks is 1.339 years. The mean values for the duration of the workout process vary between banks. Table 3.8 shows the duration of the workout process for each bank. The average durations for the workout process are 1.370 for Bank A, 1.453 for Bank B, and 0.926 for Bank C.

There is an approximate six-month difference between Bank B and Bank C in the duration of the workout process. Therefore, a sufficient observation period for estimating LGD may be different between banks.

Table 3.7: Median, mean, and standard deviation of the explanatory variables (data set B)

Variable	Median	Mean	SD
Creditworthiness score	31.000	30.849	13.812
Collateral quota (real estate)	0.000	0.146	0.338
Collateral quota (commercial bills)	0.000	0.034	0.127
Collateral quota (deposits)	0.000	0.016	0.095
Collateral quota (marketable securities)	0.000	0.003	0.031
Credit guarantee quota	0.824	0.628	0.411
EAD in hundred million yen	0.240	1.015	3.961
Duration of the workout process (in years)	1.000	1.339	1.125

Table 3.8: Length of the workout process (in years) for each bank

	Median	Mean	SD
Bank A	1.000	1.370	1.149
Bank B	1.000	1.453	1.166
Bank C	0.500	0.926	0.797

Table 3.9 shows the correlation matrix of data set B. The quotas for credit guarantee, EAD, and duration of the workout process have a relatively high correlation with LGD. Since the quota for credit guarantee is high, LGD is lower. The large EAD and the extended duration of the workout process results in high LGD. Although the collateral quotas have a negative correlation with LGD, the absolute value itself is small. The correlation matrix could not confirm that collateral has a significant influence on LGD.

Table 3.9: Correlation matrix of data set B

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) LGD	1	0.155	-0.096	-0.003	-0.023	-0.021	-0.452	0.153	0.006
(2) Creditworthiness score	0.155	1	-0.037	0.005	0.022	0.008	-0.254	0.128	-0.042
(3) Collateral quota (real estate)	-0.096	-0.037	1	-0.038	0.019	-0.003	-0.210	0.109	0.076
(4) Collateral quota (commercial bills)	-0.003	0.005	-0.038	1	0.017	0.005	-0.216	0.114	-0.041
(5) Collateral quota (deposits)	-0.023	0.022	0.019	0.017	1	0.076	-0.141	0.062	-0.023
(6) Collateral quota (marketable securities)	-0.021	0.008	-0.003	0.005	0.076	1	-0.085	0.085	0.013
(7) Credit guarantee quota	-0.452	-0.254	-0.210	-0.216	-0.141	-0.085	1	-0.465	-0.075
(8) ln(EAD)	0.153	0.128	0.109	0.114	0.062	0.085	-0.465	1	0.132
(9) Duration of the workout process	0.006	-0.042	0.076	-0.041	-0.023	0.013	-0.075	0.132	1

3.4 Linear Regression Analysis

In this section, we conduct a linear regression analysis to survey the impacts of influencing factors on LGD. LGD has a characteristic that lies in the interval between 0 and 1, which is confirmed in Figure 3.1. In consideration of this characteristic, we use the logit transformed linear regression model. The predicted values from this regression are guaranteed to lie in the unit interval. The logit transformed linear regression is expressed as follows:

$$\log\left(\frac{LGD}{1-LGD}\right) = -\left(\alpha^{(0)} + \sum_j \beta_j^{(0)} x_j^{(0)}\right), \qquad (3.4.1)$$

where $x^{(0)}$ represent the explanatory variables, and $\alpha^{(0)}$ and $\beta^{(0)}$ represent the parameters in the regression. Table 3.10 shows the result of the logit transformed linear regression.

Table 3.10 shows that the creditworthiness score is nonsignificant. This result indicates that there is no relationship between the creditworthiness score and LGD. Grunert and Weber (2009) have confirmed that the LGD for low creditworthy borrowers is high, but we could not find significant influence of creditworthiness in this study. All types of collateral are statistically significant and have reduction effects on LGD. Credit guarantee quotas are significant. Credit guarantee has a reduction effect on LGD. EAD is not statistically significant at the 5% level. This result indicates that there is no relationship between LGD and loan size. This result is not consistent with Felsovalyi and Hurt (1998) and Bastos (2010), who reported that the LGD for large loans tends to be high. The length of the workout process is significant. The extended duration of the workout process leads to high LGD. However, Querci (2005) reported that there is no relationship between the length of the workout process and LGD. LGD is considered to be affected by the business cycle. Altman et al. (2001) confirmed that LGD is affected by economic fluctuations. In this study, some of the year default dummies are significant. However, we are unable to address the relationship between LGD and business cycles because of the difficulty in defining peaks and troughs in business cycles.

Table 3.10: Result of the linear regression analysis of the factors influencing LGD

	Estimate	Std. Error	t-value	Pr(> $\|t\|$)
(Intercept)	1.858	0.091	20.476	0.000
Creditworthiness score	-0.001	0.002	-0.680	0.497
Collateral quota (real estate)	1.160	0.070	16.488	0.000
Collateral quota (commercial bills)	1.660	0.187	8.878	0.000
Collateral quota (deposits)	1.701	0.244	6.984	0.000
Collateral quota (marketable securities)	3.490	0.735	4.747	0.000
Credit guarantee quota	2.651	0.068	38.930	0.000
ln (EAD)	0.029	0.016	1.731	0.084
Duration of the workout process (in year)	-0.055	0.022	-2.451	0.014
Year of default				
2004	0.416	0.083	5.039	0.000
2005	0.439	0.062	7.069	0.000
2006	0.109	0.060	1.819	0.069
2007	0.113	0.054	2.083	0.037
2008	-0.042	0.054	-0.790	0.430
2009	-0.124	0.061	-2.033	0.042
2010	-0.191	0.078	-2.433	0.015
2011	-0.720	0.148	-4.859	0.000
Adjusted R-squared	0.268			
Number of observations	5664			

3.5 The EL Forecasting Model

3.5.1 Overview of the EL forecasting model

EL is typically calculated as PD×LGD. In this study, we build both the PD predictive model and the LGD predictive model. We also obtain estimates of EL from the product of model estimates of PD and LGD. We use the multi-stage model for the LGD predictive model. The multi-stage LGD model consists of three models.

First, we introduce previous studies of the multi-stage model that are associated with this study. Lucas (2006) suggested a two-stage model for modeling LGD associated with mortgages. The author divided a workout process according to whether a property is repossessed and then calculated the loss in the case of a repossession. Gürtler and Hibbeln (2011) found significant differences between the characteristics of recovered and written-off loans. To account for these differences, the authors divided defaults into two cases, namely recoveries and write-offs, using logistic regression. The authors then conducted two separate regressions for each case. Bellotti and Crook (2012) also proposed using a multi-stage model to predict LGD. Considering that LGD is concentrated around 0 and 1, the authors split LGD into three cases ($LGD = 0, LGD = 1, 0 < LGD < 1$) using logistic regression models. The ordinary least squares (OLS) regression model is then used for the case of $0 < LGD < 1$.

As Figure 3.2 shows, the EL forecasting model is composed of (I)the PD model and the multi-stage LGD model. The multi-stage LGD model consists of (II) the Pr(Recovery) model, (III) the Pr($LGD > 0$) model, and (IV) the LGD regression model; (II) the Pr(Recovery) model predicts the probability of recovery for a default borrower; (III) the Pr($LGD > 0$) model predicts the probability of causing a loss if a default borrower is written-off; and (IV) the LGD regression model predicts the LGD when a loss occurs ($LGD_{LGD>0}$). The predicted LGD from the multi-stage LGD model is expressed as follows:

$$LGD = (1 - \Pr(\text{Recovery})) \times \Pr(\text{LGD} > 0) \times \text{LGD}_{\text{LGD}>0}. \qquad (3.5.1)$$

Data sets to be used for model building vary between the models. We show the data used for each model in Table 3.11. We use all the data to build (I) the

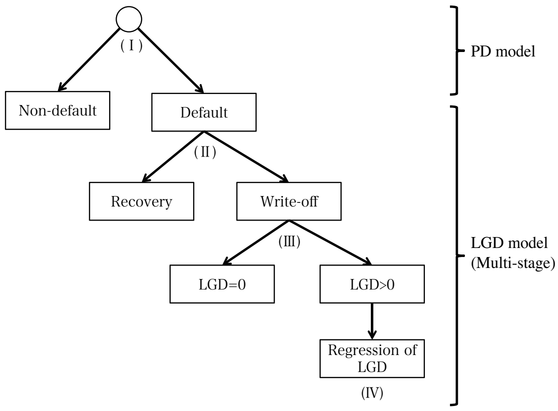

Figure 3.2: EL forecasting model

PD model. We use the default data excluding the censored data to build (II) the Pr(Recovery) model. We use the default data of write-offs to build (III) the Pr($LGD > 0$) model. Then, we use the default data with $LGD > 0$ to build (IV) the LGD regression model.

Table 3.11: Data used to build each model

model	Non-default	Default			
		Recoveries	Write-offs (LGD=0)	Write-offs (LGD>0)	Censored
(I) PD model	○	○	○	○	○
(II) Pr(Recovery) model		○	○	○	
(III) Pr(LGD>0) model			○	○	
(IV) Regression model				○	

3.5.2 Model coefficients given by regressions

We present the model coefficients from the regressions. We use Akaike's information criterion (AIC) (Akaike (1973)) to select the variables.

The estimation results for (I) the PD model

We use the logistic regression model for (I) the PD model. The logistic regression model is expressed as follows:

$$\begin{aligned} PD &= \frac{1}{1+\exp(Z^I)} \\ Z^I &= \alpha^I \sum_k \beta_k^I x_k^I, \end{aligned} \tag{3.5.2}$$

where x^I represents the explanatory variables, and α^I and β^I represent the regression parameters. The result of the regression is shown in Table 3.12. All the explanatory variables are statistically significant. We find that loan characteristics and borrower characteristics have explanatory power in predicting PD. Borrowers with a high creditworthiness score tend not to default. This is a natural result because the creditworthiness score is synthesized for the default probability estimation. All types of collateral are statistically significant. This result indicates that the PD for borrowers with a high collateral quota is low. However, the PD for borrowers with a high credit guarantee quota is high. Therefore, the bank requires that loans for borrowers with low creditworthiness be secured by credit guarantees. We find a negative relationship between exposure and PD. Although it is considered that borrowers with substantial exposure tend to default, we obtain the opposite result in this study.

Table 3.12: The estimation results for (I)PD model

	Estimate	Std. Error	z-value	Pr(> $\|z\|$)
(Intercept)	1.180	0.036	32.900	0.000
Creditworthiness score	0.072	0.001	96.546	0.000
Collateral quota (real estate)	0.637	0.035	18.012	0.000
Collateral quota (commercial bills)	0.714	0.094	7.557	0.000
Collateral quota (deposits)	0.450	0.122	3.681	0.000
Collateral quota (marketable securities)	0.886	0.258	3.427	0.001
Credit guarantee quota	-0.101	0.033	-3.060	0.002
ln (Exposure)	-0.203	0.008	-26.171	0.000
AUC	0.815			
Number of observations	679607			
Number of defaults	8732			

The estimation results for (II) the Pr(Recovery) model

We use the logistic regression model for (II) the Pr(Recovery) model. The logistic regression model is expressed as follows:

$$\begin{aligned} \Pr(\text{Recovery}) &= \frac{1}{1+\exp(Z^{II})} \\ Z^{II} &= \alpha^{II} \sum_{l} \beta_l^{II} x_l^{II}, \end{aligned} \tag{3.5.3}$$

where x^{II} represents the explanatory variables, and α^{II} and β^{II} represent the regression parameters. The result of the regression is shown in Table 3.13. Creditworthiness score, some types of collateral, credit guarantee, and EAD are statistically significant. As the creditworthiness score increases, the probability of recovery increases. Financially stable borrowers experience an easy recovery after defaults. With respect to the quota for each security, borrowers with a high collateral quota (real estate) experience an easy recovery. Borrowers secured by collateral (commercial bills) and credit guarantees experience a more difficult recovery. This result indicates that the impacts on the probability of recovery vary between the types of security. EAD also has an effect on the probability of recovery in that substantial EAD leads to a high probability of recovery.

Table 3.13: The estimation results for (II)Pr(Recovery) model

	Estimate	Std. Error	z-value	Pr(> $\|z\|$)
(Intercept)	0.959	0.115	8.315	0.000
Creditworthiness score	-0.013	0.002	-5.167	0.000
Collateral quota (real estate)	-0.323	0.095	-3.418	0.001
Collateral quota (commercial bills)	1.702	0.339	5.015	0.000
Collateral quota (deposits)				
Collateral quota (marketable securities)				
Credit guarantee quota	0.353	0.095	3.730	0.000
ln (EAD)	-0.382	0.025	-15.568	0.000
AUC	0.701			
Number of observations	5664			
Number of recoveries	1334			

The estimation results for (III) the $\Pr(LGD > 0)$ model

We use the logistic regression model for the (III)$\Pr(LGD > 0)$ model. The logistic regression model is expressed as follows:

$$\begin{aligned} \Pr(LGD > 0) &= \frac{1}{1 + \exp(Z^{III})} \\ Z^{III} &= \alpha^{III} \sum_m \beta_m^{III} x_m^{III}, \end{aligned} \tag{3.5.4}$$

where x^{III} represents the explanatory variables, and α^{III} and β^{III} represent the parameters in the regression. The result of the regression is shown in Table 3.14. All the variables, except the creditworthiness score, are statistically significant. A loss is less likely to occur because collateral and credit guarantee quotas are high. Borrowers with large EAD are likely to cause damage.

Table 3.14: The estimation results for (III)Pr($LGD > 0$) model

	Estimate	Std. Error	z-value	Pr($> \|z\|$)
(Intercept)	-2.205	0.107	-20.701	0.000
Creditworthiness score				
Collateral quota (real estate)	2.590	0.208	12.455	0.000
Collateral quota (commercial bills)	2.487	0.295	8.440	0.000
Collateral quota (deposits)	3.221	0.589	5.466	0.000
Collateral quota (marketable securities)	5.384	2.237	2.407	0.016
Credit guarantee quota	3.421	0.122	28.052	0.000
ln (EAD)	-0.654	0.037	-17.621	0.000
AUC	0.904			
Number of observations	4330			
Number of LGD>0s	1116			

The estimation results for (IV) the LGD regression model

We use the logit transformed OLS model for (IV) the LGD regression model. The logit transformed OLS model is expressed as follows:

$$\log\left(\frac{LGD_{LGD>0}}{1-LGD_{LGD>0}}\right) = -\left(\alpha^{IV} + \sum_{n} \beta_n^{IV} x_n^{IV}\right), \qquad (3.5.5)$$

where x^{IV} represents the explanatory variables, and α^{IV} and β^{IV} represent the regression parameters. The result of the regression is shown in Table 3.15. All the variables, except the creditworthiness score, are statistically significant. As collateral and credit guarantee quotas increase, $LGD_{LGD>0}$ decreases. Substantial EAD leads to low $LGD_{LGD>0}$. This result is distinct from the result in Section 3.5.2 because the data set employed is different. In the latter section, the analysis included the categories of write-offs (LGD$>$ 0) and write-offs (LGD = 0). However, in this section, we use the data set that contains only LGD$>$ 0. Although borrowers with large EAD are likely to cause damage, LGD is relatively small if a loss occurs.

Table 3.15: The estimation results for (IV)LGD regression model

	Estimate	Std. Error	t-value	Pr($>|t|$)
(Intercept)	-1.523	0.072	-21.256	0.000
Creditworthiness score				
Collateral quota (real estate)	2.142	0.207	10.334	0.000
Collateral quota (commercial bills)	2.786	0.312	8.935	0.000
Collateral quota (deposits)	4.414	0.640	6.897	0.000
Collateral quota (marketable securities)	6.690	2.465	2.715	0.007
Credit guarantee quota	4.431	0.133	33.198	0.000
ln (EAD)	0.213	0.031	6.788	0.000
Adjusted R-squared	0.539			
Number of observations	1116			

3.6 Validation of the Multi-stage LGD Model

We assess the predictive accuracy of the multi-stage LGD model developed in Section 3.5 using several performance measures: R-squared, Spearman's rho, the mean absolute error (MAE), the root mean squared error (RMSE), and the relative absolute error (RAE). The MAE, RMSE, and RAE are defined as follows:

$$MAE = \frac{1}{n}\sum_i |y_i - \hat{y}_i|, \tag{3.6.1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2}, \tag{3.6.2}$$

$$RAE = \frac{\sum_i |y_i - \hat{y}_i|}{\sum_i |y_i - \bar{y}_i|}, \tag{3.6.3}$$

where y_i and $\hat{y}_i$ represent the actual LGD and the predicted LGD on borrower i, respectively, and n represents the number of observations in the sample. Models with higher R-squared and Spearman's rho have superior predictive accuracy. Models with lower MAE, RMSE, and RAE have superior predictive accuracy. To avoid overfitting the data, we conducted a 10-fold cross validation. Bastos (2010) also implemented a 10-fold cross validation to evaluate predictive accuracy of the LGD model. In a 10-fold cross validation, the entire sample is divided into 10 subsets. Nine of the subsets are used to build the model, and the remaining single subset is used to assess the model. This procedure is repeated 10 times. All subsets are used once as a single test subset. Table 3.16 shows the in-sample and out-of-sample goodness of fit. For comparison, the OLS and the OLS (logit transformed) models are also shown.

The performance of the multi-stage model is superior to other traditional linear regression models in terms of the R-squared, Spearman's rho, and RMSE. Only slight differences are apparent between the OLS (logit transformed) and the multi-stage model in the MAE and RAE.

Table 3.16: 10-fold cross validation

Model	R-squared	Spearman's rho	MAE	RMSE	RAE
In-sample					
OLS	0.272	0.452	0.119	0.195	0.820
OLS (logit transformed)	0.305	0.471	0.100	0.213	0.686
Multi-stage	0.318	0.480	0.103	0.190	0.711
Out-of-sample					
OLS	0.271	0.450	0.119	0.195	0.822
OLS (logit transformed)	0.305	0.469	0.100	0.213	0.686
Multi-stage	0.320	0.478	0.104	0.190	0.714

3.7 Conclusions

In this chapter, we investigated the factors influencing LGD and developed an EL forecasting model, including the multi-stage model, to predict LGD. We confirmed that Japanese bank LGD is significantly smaller than the LGD levels of other countries reported in published empirical research. We calculated the fundamental data statistics and analyzed the relationships between the influencing factors and LGD. Using Japanese bank loan data, we built the LGD and EL forecasting model and proposed methods to estimate LGD and EL. We obtained the following results. Japanese bank LGD levels are lower than those suggested by the FIRB approach and the LGD levels reported in other countries. The duration of the workout process varies between banks. The relationship between LGD and the duration of the workout process is also different between banks. Therefore, sufficient observation periods for the estimation of LGD would differ between banks. Collateral, credit guarantees, and EAD are important factors that influence LGD. We confirmed that the multi-stage LGD model is superior in terms of predictive accuracy than traditional linear models.

Future research tasks concerning bank loan credit risks are as follows. This study used a considerable amount of censored data. This may cause bias in estimations if the duration of the workout process significantly impacts LGD. Modeling the workout process itself would be necessary to include censored data in the analysis. If the number of banks providing data increases, we can further investigate the differences between banks, determine whether the business sectors of borrowers have an effect on LGD, and consider the downturn LGD estimates that are required by the Basel Accord. This study showed that the LGD level differs from year to year. However, we could not address the relationship between business cycles and LGD. If sufficient data are accumulated, we may be able to do so in future.

4

Issue III - Structural Credit Risks with Non-Gaussian and Serially Correlated Innovations

4.1 Introduction

The credit risk problem concerns the probability of financial losses owing to changes in market participants' credit quality. Central to credit risk is a potential default event, which occurs if the debt-holding entity cannot meet its legal obligations according to the debt contract. Since the end of the 1990s, banks have become sophisticated in their risk management. They conduct statistical evaluations of defaultable events and analyze credit risk under the Basel II Accord. Credit risk has become an increasing concern since the recession of the last decade, and is a topic of critical importance in the banking industry, as it affects a variety of stakeholders, institutions, consumers, and regulators.

In credit risk literature, there are two primary classes of modeling credit risk that attempt to describe the default processes: structural and reduced-form models. Structural models, pioneered by Merton (1974), use the evolution of firms' structural variables (e.g., asset and debt values) to determine the time of default, and employ modern option pricing theory in corporate debt valuation. In Merton's model, a firm defaults if, at the time of servicing the debt, its assets are below its outstanding debt.

The beauty of Merton's model lies in the intuition of treating company equity as a call option on its assets. Therefore, it has been widely used in theoretical and empirical analyses. However, the disadvantages of this model are that a very simple stochastic process on assets and an unrealistic capital structure are assumed, implying that default can only happen when the zero-coupon bond hits maturity. For empirical analyses based on Merton's framework, we refer the reader to Tarashev (2008), Kealhofer (2003a,b), Hull et al. (2004), and the references therein.

Many empirical finance studies report that the returns of financial assets are often skewed and have large kurtosis and serially correlated heteroskedasticity, which are commonly known to be stylized facts of asset returns. However, credit risk models in the literature are mainly based on the assumption that the asset value follows a geometric Brownian motion, and few studies incorporate models with non-Gaussian and dependent assumptions. With regard to the modeling of asset returns and option pricing, incorporating volatility clustering and the non-normality of returns, we refer to Jondeau et al. (2007). Accurate measures of credit risk are key prerequisites for sound risk management.

This study develops Merton's framework to evaluate credit risk under the assumption that the underlying asset processes have non-Gaussian and serially correlated innovations. Using a higher-order asymptotic expansion of the distributions of the underlying asset returns, we obtain the closed-form expressions for the probability of default, the distance to default, and the term structure of the credit spread. The expressions obtained allow us to ensure accuracy by investigating how the non-Gaussian and dependent nature of asset returns affect the credit risk evaluation. Thus, we use real-world data to investigate our model with a more realistic depiction of the credit risk evaluation than that of Merton's usual framework.

A number of studies have examined financial models with non-Gaussian and serially correlated innovations using the asymptotic expansion approach. Honda et al. (2010) and Shiohama and Tamaki (2012) consider higher-order asymptotic valuations for zero-coupon bonds and European call options on zero-coupon bonds using single-factor, discretely observed Vasicek models, with non-Gaussian and dependent error structures. With regard to the reduced-form approach to modeling the default event, Miura et al. (2013) develop a closed-form valuation for pricing defaultable bonds that incorporates a stochastic, risk-free interest rate and defaultable intensity processes with non-Gaussian and dependent processes.

The remainder of this chapter is organized as follows. Section 4.2 explains the assumptions and the models with non-Gaussian innovations, and evaluates the probability of default, the distance to default, and the credit spread term structures. Then, numerical examples are provided in Section 4.3. Section 4.4 estimates the model parameters using an extension of Moody's KMV approach. Section 4.5 considers empirical applications, and Section 4.6 concludes the chapter. Proofs are provided in the Appendix.

4.2 Merton's Model with Non-Gaussian and Dependent Innovations

In Merton's model, the firm's capital structure is assumed to be composed of equity and a zero-coupon bond D_t, with maturity T and face value K. We define S_t as the value of the firm's equity and V_t as the value of its asset at time t. Then, a capital structure is given by the balance sheet relationship, $V_t = S_t + D_t$.

If the firm's asset value exceeds the promised payment at time T—that is, $V_T > K$—the lenders are paid the promised amount, and the shareholders receive the residual asset value. On the other hand, if the firm defaults on its debt at T with an asset value V_t, the shareholders are left with nothing. The firm's equity is simply a European call option with maturity T and strike price K on the asset value. Therefore, the firm's debt value is simply the asset value less the equity value.

We assume that the stochastic process $\{V_j\}$ is discretely sampled with interval Δ such that V_j is sampled at times $t, t+\Delta, t+2\Delta, \ldots, t+n_t\Delta (\equiv T)$ over $[t, T]$. According to the Euler approximation, for $j = 1, 2, \ldots, n_t$, the discrete scheme for Merton's model is then

$$\ln V_{t+j\Delta} - \ln V_{t+(j-1)\Delta} = r\Delta + \Delta^{1/2}\sigma_V \varepsilon_j,$$

where r denotes the continuously compounded risk-free interest rate, σ_V is the asset volatility, and ε_j is a standard normal random variable. Following Honda et al. (2010), we extend the discretized Merton model to possess non-Gaussian and dependent innovations, defined as

$$\ln V_{t+j\Delta} - \ln V_{t+(j-1)\Delta} = r\Delta + \Delta^{1/2} X_j, \tag{4.2.1}$$

where $\{X_j\}$ is the fourth-order stationary process defined by the following assumptions.

Assumption 1. The process X_j is fourth-order stationary in the sense that

1. $E[X_j] = 0$,
2. $\mathrm{cum}(X_j, X_{j+u}) = c_X(u)$,
3. $\mathrm{cum}(X_t, X_{j+u_1}, X_{j+u_2}) = c_X(u_1, u_2)$,
4. $\mathrm{cum}(X_t, X_{j+u_1}, X_{j+u_2}, X_{j+u_3}) = c_X(u_1, u_2, u_3)$.

Assumption 2. The k-th order cumulants $c_X(u_1, \ldots, u_{k-1})$ of X_t for $k = 2, 3, 4$ satisfy

$$\sum_{u_1, \cdots, u_{k-1} = -\infty}^{\infty} |c_X(u_1, \ldots, u_{k-1})| < \infty.$$

Assumptions 1 and 2 are satisfied by a wide class of time-series models containing the usual ARMA and GARCH processes.

From (4.2.1), the firm's asset value at maturity $T = t + n_t\Delta$ is expressed as

$$\ln V_{t+n_t\Delta} = \ln V_t + rn_t\Delta + \Delta^{1/2} \sum_{j=1}^{n_t} X_j.$$

Define $Y_{n_t} = \frac{1}{\sqrt{n_t}} \sum_{j=1}^{n_t} X_j$. Then, we observe

$$V_{t+n_t\Delta} = V_T = V_t \exp\left(r(T-t) + \sqrt{T-t}Y_{n_t}\right). \tag{4.2.2}$$

The following lemma evaluates the cumulants of $\{Y_{n_t}\}$.

Lemma 1. *Under Assumptions 1 and 2, the cumulants of Y_{n_t} are evaluated as follows:*

1. $E[Y_{n_t}] = 0$,
2. $\mathrm{Var}[Y_{n_t}] = \sigma^2_{n_t}$,
3. $\mathrm{cum}(Y_{n_t}, Y_{n_t}, Y_{n_t}) = n_t^{-1/2} C_3^{(n_t)}$,
4. $\mathrm{cum}(Y_{n_t}, Y_{n_t}, Y_{n_t}, Y_{n_t}) = n_t^{-1} C_4^{(n_t)}$,

where σ_{n_t}, $C_3^{(n_t)}$, *and* $C_4^{(n_t)}$ *are bounded for* n_t.

To derive the Edgeworth expansion of Y_{n_t}, we need the following assumption.

Assumption 3. The J-th order ($J \geq 5$) cumulants of Y_{n_t} are of order $O(n_t^{-J/2+1})$.

We then arrive at the following theorem.

Theorem 1. *Under Assumptions 1–3, the third-order Edgeworth expansion of the density function of* $Y = Y_{n_t}/\sigma_{n_t}$ *is given by*

$$g(y) = \phi(y)\left\{1 + \frac{\tilde{C}_3^{(n_t)}}{6\sqrt{n_t}} H_3(y) + \frac{\tilde{C}_4^{(n_t)}}{24 n_t} H_4(y) + \frac{(\tilde{C}_3^{(n_t)})^2}{72 n_t} H_6(y)\right\} + o(n_t^{-1}), \tag{4.2.3}$$

where $\phi(\cdot)$ *is the standard normal density function,* $H_k(\cdot)$ *is the* k*-th order Hermite polynomial,* $\tilde{C}_3^{(n_t)} = C_3^{(n_t)}/\sigma^3_{n_t}$, *and* $\tilde{C}_4^{(n_t)} = C_4^{(n_t)}/\sigma^4_{n_t}$.

Next, we consider a martingale measure such that Process (4.2.2) becomes martingale under the probability measure given in Theorem 1. Let $m = r - \frac{\sigma^2_{n_t}}{2} - \frac{\sqrt{T-t}\sigma^3_{n_t}\tilde{C}_3^{(n_t)}}{6\sqrt{n_t}} - \frac{(T-t)\sigma^4_{n_t}\tilde{C}_4^{(n_t)}}{24 n_t}$. With the identity

$$\int_{-\infty}^{\infty} e^{\sigma_{n_t}\sqrt{T-t}z} H_k(z)\phi(z)dz = (\sigma_{n_t}\sqrt{T-t})^k e^{\sigma^2_{n_t}(T-t)/2}$$

and with Theorem 1, we observe that the process $\tilde{V}_T = V_t \exp\{m(T-t) +$

$\sqrt{T-t}\sigma_{n_t}Y\}$ is asymptotically martingale, such that

$$\begin{aligned}
& e^{-r(T-t)}E_t[\tilde{V}_T] = e^{-r(T-t)}V_t E_t[e^{m(T-t)+\sqrt{T-t}\sigma_{n_t}Y}] \\
= \;& V_t e^{(m-r)(T-t)} \exp\left(\frac{(T-t)\sigma_{n_t}^2}{2}\right) \\
\times \;& \left\{1 + \frac{(T-t)^{3/2}\sigma_{n_t}^3 \tilde{C}_3^{(n_t)}}{6\sqrt{n_t}} + \frac{(T-t)^2\sigma_{n_t}^4 \tilde{C}_4^{(n_t)}}{24 n_t} + \frac{(T-t)^3\sigma_{n_t}^6 (\tilde{C}_3^{(n_t)})^2}{72 n_t}\right\} + o(1) \\
\approx \;& V_t e^{(m-r)(T-t)} \exp\left(\frac{(T-t)\sigma_{n_t}^2}{2}\right) \exp\left(\frac{(T-t)^{3/2}\sigma_{n_t}^3 \tilde{C}_3^{(n_t)}}{6\sqrt{n_t}} + \frac{(T-t)^2\sigma_{n_t}^4 \tilde{C}_4^{(n_t)}}{24 n_t}\right) \\
= \;& V_t.
\end{aligned}$$

Hereafter, we consider the process $\tilde{V}_T = e^{m(T-t)+\sqrt{T-t}\sigma_{n_t}Y}$ and note that $e^{-rt}\tilde{V}_T$ is a martingale sequence. The equity value S_t at earlier times $t < T$ can be derived using a similar argument to that discussed in Tamaki and Taniguchi (2007). Let

$$d_1 = \frac{\ln(V_t/K) + (m + \sigma_{n_t}^2/2)(T-t)}{\sigma_{n_t}\sqrt{T-t}}, \quad \text{and} \quad d_2 = d_1 - \sigma_{n_t}\sqrt{T-t}. \quad (4.2.4)$$

Applying the result from Tamaki and Taniguchi (2007), the value of equity at time $t (0 \le t \le T)$ is given by the following theorem.

Theorem 2. *Suppose that Assumptions 1–3 hold. Then, the price of a European call option is given by*

$$S_t = V_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) + \frac{\tilde{C}_3^{(n_t)}G_3}{6\sqrt{n_t}} + \frac{\tilde{C}_4^{(n_t)}G_4}{24 n_t} + \frac{(\tilde{C}_3^{(n_t)})^2 G_6}{72 n_t} + o(1),$$

where

$$G_k = V_t \left\{\sum_{j=1}^{k-1} (\sigma_{n_t}\sqrt{T-t})^j H_{k-j-1}(-d_2)\phi(d_1)\right\}, \quad (4.2.5)$$

for $k = 3, 4$ and

$$G_6 = V_t \left[\sum_{j=1}^{2} (\sigma_{n_t}\sqrt{T-t})^j \left\{H_{5-j}(-d_2) - \sigma_{n_t}^2 (T-t) H_{3-j}(-d_2)\right\}\right] \phi(d_1). \quad (4.2.6)$$

Under the proposed model framework, a credit default at time T occurs when the value of a firm's asset falls below its default points. The default probability is given by

$$P(\tilde{V}_T < K) = P[V_t e^{m(T-t)+\sqrt{T-t}\sigma_{n_t} Y} < K].$$

The next theorem shows the default probability of our proposed non-Gaussian and dependent asset process under the martingale probability measure.

Theorem 3. *Suppose that Assumptions 1–3 hold. Then, the current default probability is expressed as*

$$PD = P(\tilde{V}_T < K) = \Phi(-d_2) + PD_1 + PD_2 + PD_3 + o(n^{-1}), \qquad (4.2.7)$$

where d_2 is given in (4.2.4),

$$\begin{aligned} PD_1 &= -\frac{\tilde{C}_3^{(n_t)}}{6\sqrt{n_t}}\Phi^{(2)}(-d_2), \\ PD_2 &= -\frac{\tilde{C}_4^{(n_t)}}{24 n_t}\Phi^{(3)}(-d_2), \\ PD_3 &= -\frac{(\tilde{C}_3^{(n_t)})^2}{72 n_t}\Phi^{(5)}(-d_2), \end{aligned}$$

where $\Phi(x)$ is the standard normal distribution function of x, and $\Phi^{(a)}(x) = \frac{\partial^a}{\partial^a x}\Phi(x)$.

The distance to default can measure how far a limited-liability firm is from default, which is the distance between the expected value of the asset and the default point. This measure is obtained from the next theorem.

Theorem 4. *Suppose that Assumptions 1–3 hold. The distance to default at time t is evaluated as*

$$DD = d_2 + \frac{\tilde{C}_3^{(n_t)}}{6\sqrt{n_t}}(d_2^2 - 1) + \frac{\tilde{C}_4^{(n_t)}}{24 n_t}(d_2^3 - 3d_2)^2 + \frac{(\tilde{C}_3^{(n_t)})^2}{36 n_t}(2d_2^3 - 5d_2) + o(1),$$

where d_2 is given in (4.2.4).

Finally, we derive the evaluation of the term structure of the firm's credit spread. Although debt holders are exposed to default risk, they can completely hedge their position by purchasing a European put option written on the same underlying asset V_t with strike price K. Such a put option is worth $K - V_t$ if $V_t < K$ and nothing if $V_T > K$.

Combining the debt positions and the put option, we guarantee a payoff of K for debt holders at time T, forming a risk-free position, $D_t + P_t = Ke^{-r(T-t)}$. Here, P_t denotes the put option price at time t, which is evaluated as

$$P_t = Ke^{-r(T-t)}\Phi(-d_2) - V_t\Phi(-d_1) + G_3 + G_4 + G_6 + o(1), \tag{4.2.8}$$

where G_3, G_4, and G_6 are defined as in (4.2.5) and (4.2.6) of Theorem 2. Corporate debt is a risky bond and, thus, should be valued as a credit spread. Let $s = s(t,T)$ denote the continuously compounded credit spread. Then, bond price D_t can be written as

$$D_t = Ke^{-(r+s)(T-t)}. \tag{4.2.9}$$

Using the relation $D_t + P_t = Ke^{-r(T-t)}$, we observe that

$$D_t = Ke^{-r(T-t)} - P_t = Ke^{-r(T-t)}\Phi(d_2) - V_t\Phi(-d_1) - G_3 - G_4 - G_6 + o(1).$$

Then, from (4.2.9), we have a closed-form formula for the credit spread yield to maturity s, defined as

$$s(t,T) = -\frac{1}{T-t}\ln\left(\Phi(d_2) - \frac{e^{r(T-t)}}{K}\left[V_t\Phi(-d_1) + G_3 + G_4 + G_6\right]\right) + o(1). \tag{4.2.10}$$

This is a function of maturity T, asset volatility, skewness, and kurtosis. This information can be used as a proxy to derive the prices of credit default contracts, such as credit default swaps.

To our knowledge, this is the first attempt to obtain the closed-form expressions for credit risk valuation incorporating a firm's non-Gaussian and dependent asset value returns. The closed-form expressions obtained in this section seems complex, but we show our estimation of the model parameters in the next section, as well as how non-Gaussianity affects the credit risk valuation via an empirical analysis.

4.3 Examples of Evaluating Credit Risks

This section presents numerical examples based on our proposed non-Gaussian and serially correlated asset returns.

Example 1. Suppose that $\{X_j\}$ is a sequence of i.i.d. random variables, each with mean zero, variance σ^2, skewness β_1, and kurtosis β_2. Then,

$$c_X(u) = \begin{cases} \sigma^2 & \text{if } u = 0, \\ 0 & \text{otherwise,} \end{cases} \qquad c_X(u_1, u_2) = \begin{cases} \sigma^3 \beta_1 & \text{if } u_1 = u_2 = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$c_X(u_1, u_2, u_3) = \begin{cases} \sigma^4(\beta_2 - 3) & \text{if } u_1 = u_2 = u_3 = 0, \\ 0 & \text{otherwise.} \end{cases}$$

These cumulants yield the third- and fourth-order joint cumulants Y_n, as follows:

$$\tilde{C}_3(n) = \beta_1/\sigma^3 \quad \text{and} \quad \tilde{C}_4(n) = (\beta_2 - 3)/\sigma^4. \tag{4.3.1}$$

Substituting (4.3.1) into (4.2.7), (4.2.8), and (4.2.10), we obtain the probability of default, the distance to default, and the term structure of the credit spread, respectively.

Example 2. Let $\{X_j\}$ be the AR(1) process,

$$X_j = \alpha X_{j-1} + \varepsilon_j,$$

where $|\alpha| < 1$ and $\{\varepsilon_j\}$ is a sequence of i.i.d. random variables, each with mean zero, variance σ^2, skewness γ_1, and kurtosis γ_2. It is easy to see that Assumptions 1 and 2 hold. Indeed, since X_j has the MA(∞) representation

$$X_j = \sum_{i=0}^{\infty} \alpha^i \varepsilon_{j-i},$$

and the k-th order cumulant spectral densities $f_{X,k}(\lambda)$ for $k = 2, 3, 4$ are

$$f_{X,2}(\lambda) = \frac{\sigma^2}{2\pi} A(\lambda) A(-\lambda),$$
$$f_{X,3}(\lambda_1, \lambda_2) = \frac{\sigma^3 \gamma_1}{(2\pi)^2} A(\lambda_1) A(\lambda_2) A(-\lambda_1 - \lambda_2),$$

and

$$f_{X,4}(\lambda_1, \lambda_2, \lambda_3) = \frac{\sigma^4(\gamma_2 - 3)}{(2\pi)^3} A(\lambda_1)A(\lambda_2)A(\lambda_3)A(-\lambda_1 - \lambda_2 - \lambda_3),$$

where $A(\lambda) = \sum_{j=0}^{\infty} \alpha^j e^{-\boldsymbol{i}j\lambda} = 1/(1-\alpha e^{-\boldsymbol{i}\lambda})$ is the transfer function (e.g., Brillinger, 2001, Chapter 2). Thus, we have

$$c_X(u_1, \ldots, u_{k-1}) = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f_{X,k}(\lambda_1, \ldots, \lambda_{k-1}) e^{-\boldsymbol{i}\sum_{j=1}^{k-1} \lambda_j u_j} d\lambda_1 \cdots d\lambda_{k-1},$$

for $k = 2, 3, 4$. Then the variance and third- and fourth-order cumulants of Y_n can be evaluated as

$$\sigma_n^2 = \frac{1}{n} \sum_{k,\ell=1}^{n} c_X(k - \ell), \tag{4.3.2}$$

$$\sigma_n^3 \tilde{C}_3^{(n)} = \frac{1}{n^{3/2}} \sum_{k,\ell,m=1}^{n} c_X(k - \ell, k - m), \tag{4.3.3}$$

and $$\sigma_n^4 \tilde{C}_4^{(n)} = \frac{1}{n^2} \sum_{j,k,\ell,m=1}^{n} c_X(j - k, j - \ell, j - m), \tag{4.3.4}$$

respectively. Here, for example, the autocovariance function $c_X(u)$ at lag u is

$$c_X(u) = \sigma^2 \frac{\alpha^{|u|}}{1 - \alpha^2}.$$

Substituting (4.3.2)–(4.3.4) into (4.2.7), (4.2.8), and (4.2.10), we obtain the probability of default, the distance to default, and the term structure of the credit spread, respectively.

Example 3. Let $\{X_j\}$ follows a GARCH(1,1) process

$$X_j = h_j^{1/2} \varepsilon_j, \quad h_j = \omega + \alpha X_{j-1}^2 + \beta h_{j-1},$$

where $\{\varepsilon_j\}$ is a sequence of i.i.d. standard normal random variables. The parameter values must satisfy $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$, and $1 - 2\alpha^2 - (\alpha + \beta)^2 > 0$ for the existence of the fourth-order moment of $\{X_j\}$. Accordingly, σ_X^2 should be

$$\sigma_X^2 = \frac{\omega}{1 - \alpha - \beta}.$$

Some tedious calculation yields $C_3^{(n)}$ and $C_4^{(n)}$ should become

$$\begin{aligned} C_3^{(n)} &= 0, \\ C_4^{(n)} &= \frac{3}{n}\int_{-\pi}^{\pi} f_{X^2}(\lambda)\mathrm{d}\lambda - 2\frac{3\{(1-(\alpha+\beta)^2)\}}{1-(\alpha+\beta)^2-2\alpha^2}, \end{aligned}$$

where

$$f_{X^2}(\lambda) = \frac{\sigma_\nu^2}{2\pi}\frac{1+\beta^2-2\beta\cos\lambda}{1+(\alpha+\beta)^2-2(\alpha+\beta)\cos\lambda}$$

with

$$\sigma_\nu^2 = \frac{2\omega^2(1+\alpha+\beta)}{\{1-(\alpha+\beta)\}\{1-2\alpha^2-(\alpha+\beta)^2\}}.$$

Using this result, we can evaluate structural credit risks under GARCH(1,1) innovation processes.

4.4 Parameter Estimation

The results obtained in the previous section are under a risk-neutral probability measure. Hence, the drift term of the asset value process is given by the risk-free rate $r\Delta$, as in (4.2.1). In this section, parameter estimation procedures are proposed under a physical probability measure. That is, we use $\mu\Delta$ instead of the $r\Delta$ used in (4.2.1) to represent the mean rate of return on assets.

In Merton's model with non-Gaussian and dependent innovation, a default occurs when the option is not exercised. The probability of default is given in Theorem 3, and the term structure of the corporate credit spread is given by (4.2.10). To evaluate these quantities under non-Gaussian driven Merton models, we need V_t, σ_{n_t}, and $C_3^{(n_t)}$ and $C_4^{(n_t)}$ (i.e., the third- and fourth-order cumulants of the underlying asset return processes). All of these variables are unobservable. For simplicity, we assume that the process $\{X_j\}$ is an independent random variable with skewness β_1 and kurtosis β_2.

Existing literature provides several ways to calibrate V_t and σ_V in Merton's framework. One of the approaches, developed by Duan (1994), is a transformed-data maximum likelihood estimation.

Here we adopt Moody's KMV method to estimate the unobserved asset value and unknown parameters, including the third- and fourth-order cumulants of the underlying processes. The KMV method is an iterative algorithm that infers the value of the firm's unobserved total assets and the unknown expected return and volatility from the firm's equity price series. The latter are required to compute the credit spread and default probability.

The equity price is a function of the unknown asset value with unknown volatility, skewness, and kurtosis—that is, $S_t = f(V_t; \sigma, \beta_1, \beta_2)$, the inverse of which gives the asset value V_t. That is,

$$V_t = f^{-1}(S_t; \sigma, \beta_1, \beta_2),$$

where the function $f(\cdot)$ is given in Theorem 2. Using this inverse relationship, we conduct a simple iterative algorithm that begins with arbitrary values of the model parameters and repeats until the solution converges:

Step 1: Set current time $t = 0$ and compute the implied asset values $\{\hat{V}_{j\Delta}(\hat{\sigma}^{(k)}, \hat{\beta}_1^{(k)}, \hat{\beta}_2^{(k)})\}$ for $j = 1, \dots, n$, which correspond to the observed equity values such that $\hat{V}_{j\Delta}(\hat{\sigma}^{(k)}, \hat{\beta}_1^{(k)}, \hat{\beta}_2^{(k)}) = f^{-1}(S_{j\Delta}; \hat{\sigma}^{(k)}, \hat{\beta}_1^{(k)}, \hat{\beta}_2^{(k)})$.

Step 2: Compute the implied asset returns $\{\hat{R}_{j\Delta}\}$ for $j = 1, \dots, n$, where $\hat{R}_j^{(k)} = \ln\left(\hat{V}_{j\Delta}/\hat{V}_{(j-1)\Delta}\right)$, and update the asset parameters as follows:

$$\begin{aligned}
\bar{R}^{(k)} &= \frac{1}{n}\sum_{j=1}^{n} \hat{R}_j^{(k)}, \quad \left(\hat{\sigma}^{(k+1)}\right)^2 = \frac{1}{n}\sum_{j=1}^{n}\left(\hat{R}_j^{(k)} - \bar{R}^{(k)}\right)^2, \\
\hat{\beta}_1^{(k+1)} &= \frac{1}{n}\sum_{j=1}^{n}(\hat{R}_j^{(k)} - \bar{R}^{(k)})^3/(\hat{\sigma}^{(k+1)})^3, \\
\hat{\beta}_2^{(k+1)} &= \frac{1}{n}\sum_{j=1}^{n}(\hat{R}_j^{(k)} - \bar{R}^{(k)})^4/(\hat{\sigma}^{(k+1)})^4, \\
\hat{\mu}^{(k+1)} &= \frac{1}{n}\bar{R}^{(k)} + \frac{(T\hat{\sigma}^{(m+1)})^2}{2} + \frac{(T\hat{\sigma}^{(m+1)})^3\hat{\beta}_1^{(m+1)}}{6\sqrt{n}} \\
&\quad + \frac{(T\hat{\sigma}^{(m+1)})^4(\hat{\beta}_2^{(m+1)} - 3)}{24n}.
\end{aligned}$$

Step 3: Repeat Step 1 with the updated model parameters, unless convergence has been achieved.

4.5 Empirical Analysis

In this section, we conduct an empirical study using Japanese equity data in order to investigate how the usual credit risk valuation may be affected by incorporating non-Gaussianity in equity returns.

We use equity price data sourced from ANA Holdings Inc., a large Japanese corporation, to analyze our model. The data cover the period from March 11, 2010, to April 10, 2014. ANA Holdings Inc. has a stable outlook and a BBB+ credit rating by R&I, a Japanese rating corporation, and we use this firm as a representative firm to investigate credit risk. Total liabilities are obtained quarterly from balance sheets, and debt is calculated by dividing total liabilities by the outstanding numbers of shares. We use the yield on the Japanese 1-year government bond as the risk-free rate. Figure 4.1 shows the time-series plots for the stock prices of ANA Holdings Inc., as well as its logarithm returns, total liabilities, and risk-free rates on the observed sample period. Figure 4.2 plots the kernel density for the log return of ANA Holdings Inc., along with the normal density and its mean and variance calculated from the log returns. Figure 4.2 shows that the kernel density plot of the log return reveals highly non-Gaussian properties with negative skewness and larger values of kurtosis, which are sufficiently statistically significant to reject the null hypothesis of the normally distributed assumption. Figure 4.3 shows the estimated time-varying parameters of Merton's model and our proposed non-Gaussian Merton model for μ, σ, C_3, and C_4, with the estimated asset value processes $\hat{V}_t$ and the probability of default for $T = 1$. These time-varying parameters are calculated in the following manner. At a particular time t, the model parameters are estimated using the 100 samples prior to day t, resulting in $\hat{\mu}_t$ and $\hat{\sigma}_t$ for both the Gaussian and non-Gaussian models, and $\hat{C}_3$ and $\hat{C}_4$ for the non-Gaussian models. Then, using the KMV method explained in the previous section, we obtain $\hat{V}_t$ and the probability of default. According to Figure 4.3, we see large-value losses, such as -11.3% and -14.9% on March 15, 2011, and July 3, 2012, respectively. Around these days, we have statistically significant values of negative skewness and kurtosis, which caused higher estimated probabilities of default than in the usual Merton model. We see that the effects of non-Gaussianity on the asset values and the probability of default should not be ignored. That is, they require careful attention when estimating default risks.

Next, we consider the shape of the term structures of the credit spreads and the distance to default for Merton's model and our non-Gaussian model. For this comparison, we use the data on December 29, 2011, and July 23, 2012. The former date displays the smallest differences between Merton's model and the non-Gaussian model in the default probability estimates, whereas the latter date displays the largest differences. The estimated model parameters are given in Table 4.1.

Figure 4.4 plots the term structures of the credit spreads, which slope upward for both dates and models. This figure shows that the negative skewness and large kurtosis cause a higher term structure of credit spread for July 23, 2012. However, the difference between Merton's model and the non-Gaussian model is small and the two models have similar credit spread term structures for December 29, 2011.

Figure 4.5 plots the distance to default of the estimates obtained from Merton's model and the non-Gaussian model. The distance to default shortens for July 23, 2012, especially with a maturity of one year or less (i.e., because of the effects of non-Gaussianity, the probability of default is large). The difference is clearly large for shorter maturities of distance to default, because as T increases, the underlying asset return density approaches a normal distribution, from the central limit theorem.

4.6 Summary and Conclusion

In this chapter, we focused on a discretized version of Merton's model to analyze structural credit risk with non-Gaussian and dependent innovations. We obtain the closed-form expression of the default probability, the distance to default, and the corporate credit spread term structures using Edgeworth series expansions. This information is useful for both risk managers and policymakers when evaluating credit risk and deciding on banking regulations. The proposed model can help banks determine whether a firm has a stronger risk of default by comparing the risk to that obtained from Merton's framework.

Table 4.1: Parameter estimates of Merton's model and the non-Gaussian model

Model	V_t	$\hat{\mu}$	$\hat{\sigma}_t$	PD_t	$\hat{\beta}_1$	$\hat{\beta}_2$
December 29, 2011						
Merton	764.5	-0.210	0.088	0.264		
non-Gaussian	764.5	-0.238	0.088	0.382	-3.010	21.847
July 23, 2012						
Merton	828.8	-0.101	0.049	0.000		
non-Gaussian	828.8	-0.097	0.049	0.000	0.335	0.588

Note: For December 29, 2011, we observe $S_t = 181$, $r = 0.101(\%)$, and $D_t = 584.1$. For July 23, 2012, we observe $S_t = 211$, $r = 0.129(\%)$, and $D_t = 618.6$.

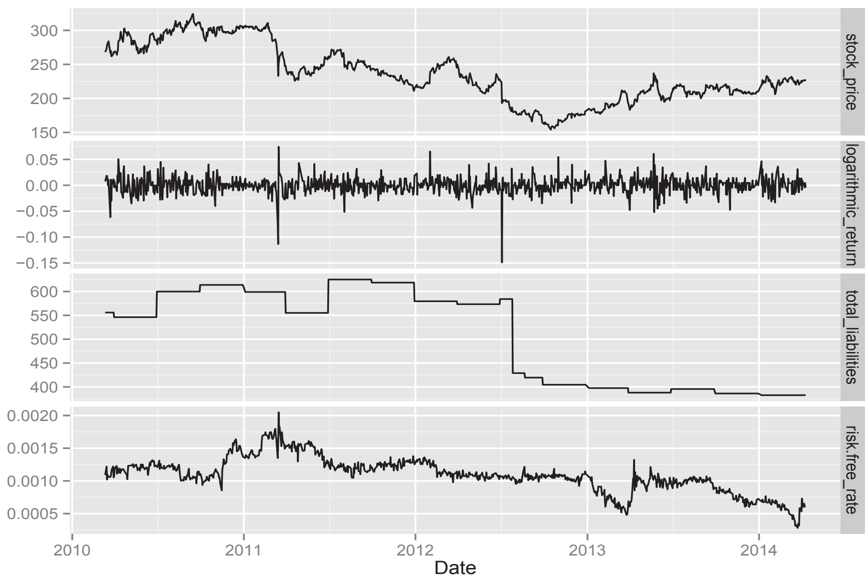

Figure 4.1: Data for implementation (ANA Holdings Inc.).

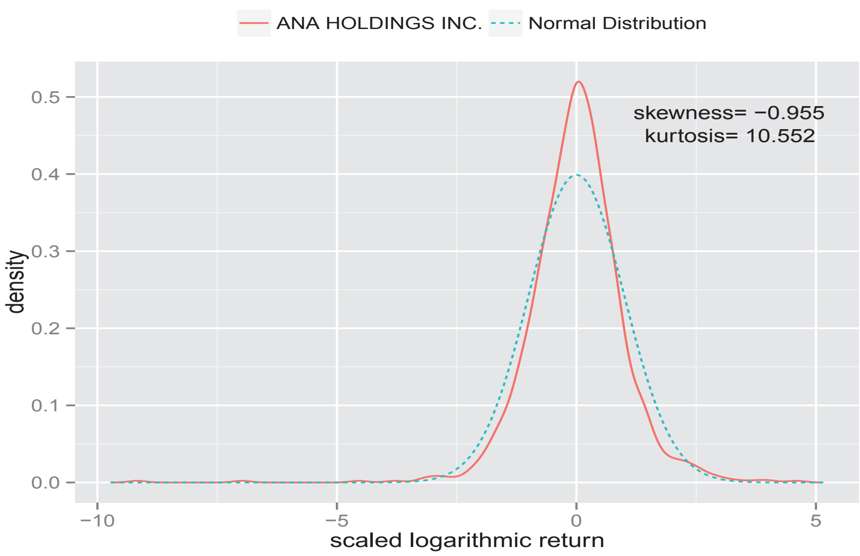

Figure 4.2: The distribution of the logarithmic return of ANA Holdings Inc. The thick line represents the kernel density plot, and the dotted line represents the normal distribution.

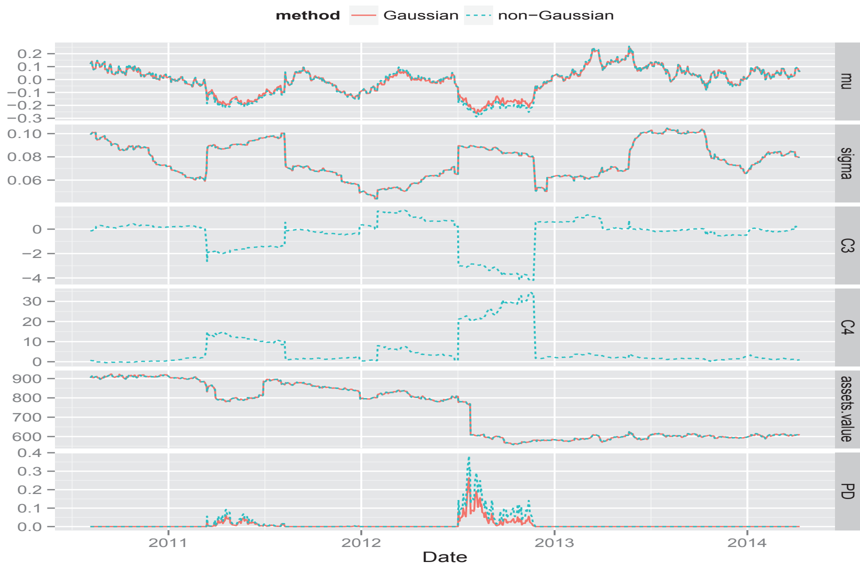

Figure 4.3: The top four panels are time-series plots for the estimates of parameters μ, σ, C_3, and C_4 in both Merton's model and the non-Gaussian model. The fifth panel shows the time-series plots for the estimated asset process. The bottom panel compares the probability of default between Merton's model and the non-Gaussian model.

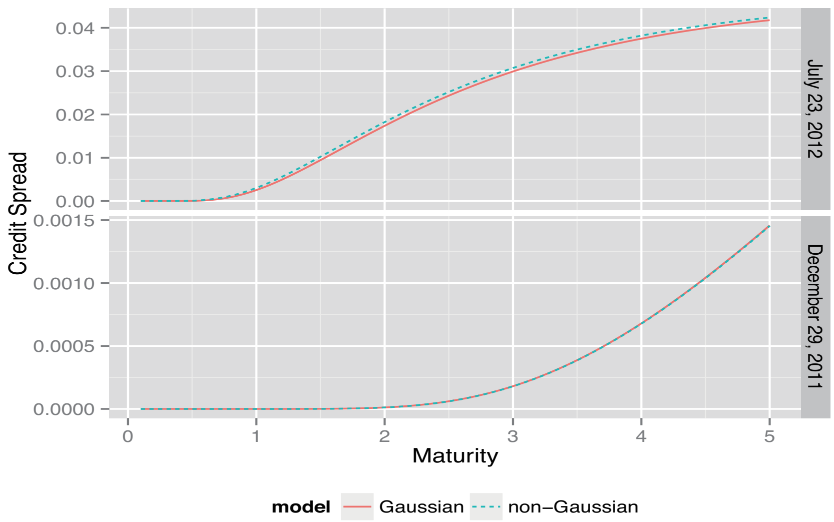

Figure 4.4: Plots of the term structures of credit spreads as a function of maturity T.

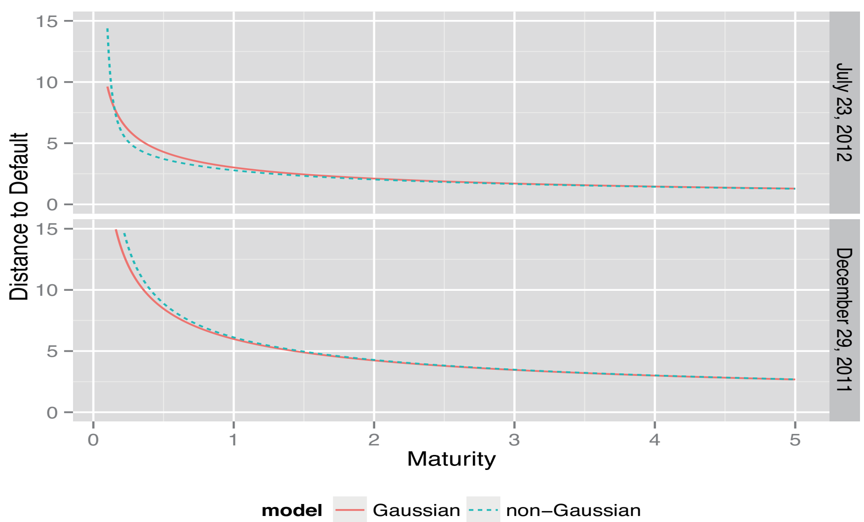

Figure 4.5: Plots of the distance to default as a function of maturity T.

4.7 Appendix

The proof of Lemma 1 is easily confirmed using Assumption 1 and, hence, it is omitted. The proofs of Theorems 1 and 2 are the same as those of Theorem 1 in Honda et al. (2010) and Theorem 3 of Tamaki and Taniguchi (2007), respectively. In this appendix, we provide the proofs for Theorems 3 and 4.

Proof of Theorem 3 The second-order Edgeworth expansion of the distribution of Y is given by

$$G(y) = \Phi(y) - \phi(y)\left(\frac{\tilde{C}_3^{(n_t)}H_2(y)}{6\sqrt{n_t}} + \frac{\tilde{C}_4^{(n_t)}H_3(y)}{24n_t} + \frac{(\tilde{C}_3^{(n_t)})^2H_5(y)}{72n_t}\right) + o(n_t^{-1}). \tag{4.7.1}$$

See, for example, Taniguchi and Kakizawa (2000). We observe

$$\begin{aligned} & P(\tilde{V}_T < K) = P(\ln \tilde{V}_T < \ln K) \\ = \;& P(\ln V_t + m(T-t) + \sqrt{T-t}\sigma_{n_t}Y < \ln K) = P(Y < -d_2) = G(-d_2), \end{aligned}$$

where d_2 is given by (4.2.4), which is a statement of Theorem 3.

Proof of Theorem 4 DD is a quantile function $G^{-1}(\cdot)$ with the distribution function $G(\cdot)$ given in (4.7.1). The theorem follows by the second-order Cornish–Fisher expansion around the Gaussian quantile $-d_2$. See, for example, Chernozhukov et al. (2010).

5

Summary and Conclusion

This thesis comprises three issues that investigate diverse topics in credit risk modeling. Credit risk modeling can be generally divided into the following models: the structural model, the reduced-form model, and the statistical model. These issues focus on the structural model, the reduce the form model and the statistical model, respectively.

In the first issue (Chapter 2), we express the term structure of risk-free rates and hazard rates using basis functions. Basis functions have been broadly used to estimate the term structures of interest rates and hazard rates. Our approach uses the maximum-likelihood method, with penalties, to estimate the parameters, allowing for the successful estimation of stable term structures. We incorporated financial ratios into the hazard rates for each rating, thus allowing us to express the different term structures depending on the issuer company, even for bonds with the same credit rating. By incorporating a penalty term that performs variable selection for the financial ratios, we can eliminate unnecessary financial ratios from the model.

In the second issue (Chapter 3), we investigate the drivers of LGD and develop the EL multi-stage forecasting model to predict LGD. We found that Japanese banks LGD are significantly lower than in other countries. We analyzed the relationships between the influencing factors and LGD. As documented in the second issue, the most important drivers of LGD for bank loans are likely to be collateral quota, credit guarantee quota, and year of default. Using Japanese bank loan data, we built the LGD and EL forecasting model and proposed the methods to estimate LGD and EL. The EL forecasting model proposed in this issue allows us to precisely estimate PD, LGD, and EL.

In the third issue (Chapter 4), we investigate a discretized version of Merton's model to analyze structural credit risk with non-Gaussian and dependent innovations. To demonstrate the appropriateness of our method, we provide an empirical analysis of a discretized version of Merton's model, which allows non-Gaussian and dependent innovations. The results show that the distance to default and the corporate credit spread term structures are directly governed by the skewness and kurtosis of the assets return. The proposed model can help banks to determine whether a firm has a stronger risk of default by comparing the findings to Merton's frameworks.

References

Acharya, V. V., Bharath, S. T., and Srinivasan, A. (2003). *Understanding the Recovery Rates on Defaulted Securities*. Centre for Economic Policy Research.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second international symposium on information theory*, pages 267–281. Akademinai Kiado.

Altman, E. (2006). Default recovery rates and lgd in credit risk modeling and practice: an updated review of the literature and empirical evidence. *New York University, Stern School of Business*.

Altman, E., Resti, A., and Sironi, A. (2004). Default recovery rates in credit risk modelling: a review of the literature and empirical evidence. *Economic Notes*, 33(2):183–208.

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

Altman, E. I. (1989). Measuring corporate bond mortality and performance. *The Journal of Finance*, 44(4):909–922.

Altman, E. I. and Eberhart, A. C. (1994). Do seniority provisions protect bondholders' investments? *The Journal of Portfolio Management*, 20(4):67–75.

Altman, E. I., Resti, A., and Sironi, A. (2001). Analyzing and explaining default recovery rates. *ISDA Report*.

Ando, T. and Yamashita, S. (2005). Simultaneous estimation of term structure of default probabilities and recovery rates by credit ratings and financial information (in Japanese). *FSA Institute Discussion Paper Series*.

Araten, M., Jacobs, M., and Varshney, P. (2004). Measuring lgd on commercial loans: an 18-year internal study. *The RMA Journal*, 86(8):96–103.

Asarnow, E. and Edwards, D. (1995). Measuring loss on defaulted bank loans: A 24-year study. *The Journal of Commercial Lending*, 77(7):11–23.

Basel Committee on Banking Supervision (2006). International convergence of capital measurement and capital standards. June 2006.

Bastos, J. a. A. (2010). Forecasting bank loans loss-given-default. *Journal of Banking & Finance*, 34(10):2510–2517.

Bellotti, T. and Crook, J. (2012). Loss given default models incorporating macroeconomic variables for credit cards. *International Journal of Forecasting*, 28(1):171–182.

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

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory*, volume 36. SIAM.

Caselli, S., Gatti, S., and Querci, F. (2008). The sensitivity of the loss given default rate to systematic risk: new empirical evidence on bank loans. *Journal of Financial Services Research*, 34(1):1–34.

Chernozhukov, V., Fernández-Val, I., and Galichon, A. (2010). Rearranging edgeworth–cornish–fisher expansions. *Economic theory*, 42(2):419–435.

Dermine, J. and De Carvalho, C. N. (2006). Bank loan losses-given-default: A case study. *Journal of Banking & Finance*, 30(4):1219–1243.

Duan, J.-C. (1994). Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance*, 4(2):155–167.

Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720.

Eales, R. and Bosworth, E. (1998). Severity of loss in the event of default in small business and larger consumer loans. *Journal of Lending & Credit Risk Management*, 80(9):58–65.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*, volume 57. CRC press.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360.

Felsovalyi, A. and Hurt, L. (1998). Measuring loss on latin american defaulted bank loans: a 27-year study of 27 countries.

Franks, J., de Servigny, A., and Davydenko, S. (2004). A comparative analysis of the recovery process and recovery rates for private companies in the uk. *France, and Germany, Standard and Poors Risk Solutions*.

Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models. *International Statistical Review/Revue Internationale de Statistique*, 55(3):245–259.

Green, P. J. and Silverman, B. W. (1993). *Nonparametric regression and generalized linear models: a roughness penalty approach*. CRC Press.

Grunert, J. and Weber, M. (2009). Recovery rates of commercial lending: Empirical evidence for german companies. *Journal of Banking & Finance*, 33(3):505–513.

Gürtler, M. and Hibbeln, M. (2011). Pitfalls in modeling loss given default of bank loans. Technical report, Working Paper, Technische Universität Braunschweig.

Honda, T., Tamaki, K., and Shiohama, T. (2010). Higher order asymptotic bond price valuation for interest rates with non-gaussian dependent innovations. *Finance Research Letters*, 7(1):60–69.

Hull, J. (2012). *Risk Management and Financial Institutions*. John Wiley & Sons, 3rd edition.

Hull, J., Nelken, I., and White, A. (2004). Merton's model, credit risk, and volatility skews. *Journal of Credit Risk*, 1(1):3–27.

Itoh, Y. and Yamashita, S. (2008). Empirical research on bank loan recovery rates for SMEs (in Japanese). *FSA Research Review*, 2007:189–218.

Jarrow, R. (2001). Default parameter estimation using market prices. *Financial Analysts Journal*, 57(5):75–92.

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

Jondeau, E., Poon, S.-H., and Rockinger, M. (2007). *Financial Modeling Under non-Gaussian Distributions*. Springer.

Kawada, A. and Shiohama, T. (2009). Modeling term structure of corporate bond and modeling term structure of corporate bond and credit risk spreads (in Japanese). *Japanese journal of Applied Statistics*, 38(2):69–86.

Kawada, A. and Shiohama, T. (2014). Structural credit risks with non-gaussian and sirially correlated innovations (submitted). *American Journal of Mathematical and Management Sciences*.

Kawada, A. and Yamashita, S. (2013). The multi stage model for LGD and EL : Empirical research for Japanese banks (in Japanese). *FSA Research Review*, 7:1–42.

Kawasaki, Y. and Ando, T. (2002). Nonlinear regression models with regularization and their application to yield curve estimation (in Japanese). *Proceedings of the Institute of Statistical Mathematics*, 50(2):149–164.

Kealhofer, S. (2003a). Quantifying credit risk I: default prediction. *Financial Analysts Journal*, 59(1):30–44.

Kealhofer, S. (2003b). Quantifying credit risk II: Debt valuation. *Financial Analysts Journal*, 59(3):78–92.

Kijima, M. and Komoribayashi, K. (1999). *Mathematical Models for Credit Risk Valuation (in Japanese)*. Asakura shoten.

Konishi, S., Ando, T., and Imoto, S. (2004). Bayesian information criteria and smoothing parameter selection in radial basis function networks. *Biometrika*, 91(1):27–43.

Konishi, S. and Kitagawa, G. (1996). Generalised information criteria in model selection. *Biometrika*, 83(4):875–890.

Konishi, S. and Kitagawa, G. (2004). *Information Criteria (in Japanese)*. Asakura shoten.

Kusuoka, S., Aonuma, K., and Nakagawa, H. (2001). *Credit Risk Models (in Japanese)*. Kinzai Institute for Financial Affairs,Inc.

Loterman, G., Brown, I., Martens, D., Mues, C., and Baesens, B. (2012). Benchmarking regression algorithms for loss given default modeling. *International Journal of Forecasting*, 28(1):161–170.

Lucas, A. (2006). Basel ii problem solving. *In QFRMC workshop and conference on Basel II & credit risk modelling in consumer lending.*

Matuszyk, A., Mues, C., and Thomas, L. C. (2010). Modelling LGD for unsecured personal loans: Decision tree approach. *Journal of the Operational Research Society*, 61(3):393–398.

McCulloch, J. H. (1971). Measuring the term structure of interest rates. *Journal of Business*, 44(1):19–31.

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

Miura, K., Yamashita, S., and Eguchi, S. (2010). The statistical model for recovery rates and expected losses in the internal ratings based approach (in Japanese). *FSA Research Review*, (6):174–205.

Miura, M., Tamaki, K., and Shiohama, T. (2013). Asymptotic expansion for term structures of defaultable bonds with non-gaussian dependent innovations. *Asia-Pacific Financial Markets*, 20(4):311–344.

Nickell, P., Perraudin, W., and Varotto, S. (2000). Stability of rating transitions. *Journal of Banking & Finance*, 24(1):203–227.

Querci, F. (2005). Loss given default on a medium-sized italian bank's loans: an empirical exercise. *European Financial Management Association. Milan, Italy.*

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Shiohama, T. and Tamaki, K. (2012). Asymptotic expansion for interest rates with non-gaussian dependent innovations. In Nishiyama, Y., editor, *Interest Rates: Term Structure Models, Monetary Policy, and Prediction*, chapter 2, pages 19–61. Nova Science Publishers.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):111–147.

Tamaki, K. and Taniguchi, M. (2007). Higher order asymptotic option valuation for non-gaussian dependent returns. *Journal of Statistical Planning and Inference*, 137(3):1043–1058.

Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*. Springer New York.

Tarashev, N. A. (2008). An empirical evaluation of structural credit-risk models. *International Journal of Central Banking*, 4(1):1–53.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Tibshirani, R. (1997). The lasso method for variable selection in the cox model. *Statistics in Medicine*, 16(4):385–395.

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81(393):82–86.

Yamashita, S. and Kihara, T. (2004). Simultaneous estimation of implied PD and LGD (in Japanese). *ISM Research Memorandum*, 911.

Zhang, J. and Thomas, L. C. (2012). Comparisons of linear regression and survival analysis using single and mixture distributions approaches in modelling lgd. *International Journal of Forecasting*, 28(1):204–215.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.