A Study on Covariation and Contagion of Credit Risk using Bayesian Statistics

Kensuke Kato

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

Graduate School of International Corporate Strategy (ICS) Finance Program
Hitotsubashi University

Tokyo, Japan

2023

Contents

List of Tables

List of Figures

Abstract

This study proposes the models for covariation and contagion of credit risk by developing and advancing the numerical calculation using the Bayesian statistical inference method, that is, the Markov chain Monte Carlo (MCMC) method based on Bayesian statistics and statistical physics. This study aims to gain a deeper understanding of covariation and contagion of credit risk through the analysis using the models and to apply them in practice.

The concept and method of the structural approach and the reduced-form approach are used to model covariation and contagion of credit risk and also applied to the risk management for credit portfolios and the investment strategy in credit products. The replica exchange Monte Carlo method and the Hamiltonian Monte Carlo method are used to solve the problems that the Metropolis-Hastings algorithm have.

This study consists of three subjects as mentioned below. The first subject constructs the long-range Ising model used in statistical physics for credit risk modeling in heterogeneous credit portfolios. The tail of the loss distribution of the long-range Ising model has characteristics different from that of the standard model (Merton-type model, etc.) in credit risk modeling because the long-range Ising model can capture a second peak in loss distributions. A second peak is unique to the Ising model and appears in large loss areas due to the collective behavior of borrowers, which cannot be captured by the standard model. Therefore, there is a possibility of different evaluations of credit risk.

This study aims to apply methods of statistical physics to finance in the view of econophysics and is an interdisciplinary study targeting the boundary region between finance and physics. Organized as scientific knowledge applying newly techniques of statistical mechanics used in physics to credit risk modeling for credit portfolios in finance, there is a possibility to form a new academic field.

The second subject aims to understand the cointegration between credit default swap (CDS) spreads with multiple tenors through the vector error correction model for minimal modeling of cointegrated hazard rates. We discuss the cointegration of CDSs in the framework of arbitrage-free pricing. We also aim to establish the pairs trading strategy of CDSs based on cointegrated hazard rates. We develop the ordinary differential equation (ODE) based Bayesian inference to estimate the cointegration dynamics for hazard rates because the theoretical spread of cointegrated CDSs depending on such latent hazard rates does not have an analytical formula.

The third subject is to expand the ODE-based Bayesian inference to the partial differential equation (PDE) based Bayesian inference. Corporate credit risks based on its stock values are analyzed by applying the PDE-based MCMC combined with the finite difference method to the constant elasticity of variance dynamics followed by asset values since the PDEs cannot be analytically solved.

As a result, this study brings and contributes the novelty to finance and econophysics from both theoretical and empirical aspects. We will apply the concepts and methods in physics to the problems in finance because financial engineering and physics have many overlapping ideas. We will take two paths, econophysics and finance, and fill the gap between both fields. Concepts that exist in physics but not in finance (temperature, etc.) are the key to develop new academic fields.

Acknowledgements

This dissertation is a collection of research results during my time at Hitotsubashi University, Graduate School of International Corporate Strategy (ICS) Finance Program.

I would like to thank my supervisor Professor Nobuhiro Nakamura for giving me the opportunity to conduct this research and guidance during the running of this project. I was taught carefully the manners of writing a thesis and the thesis structure and written expression that are easy for the reader to understand concisely and accurately in addition to studies. I feel to have built a solid foundation to continue my studies.

I would like to thank all faculty members. The lectures attended and advices received were useful in advancing my research.

Finally, I would like to thank my wife Naomi for warmly supporting and encouraging me for a very long time.

Chapter 1

Introduction

1.1 Background of study

Covariation and contagion of credit risk have concerned researchers and practitioners. Many relevant studies have been conducted and applied in practice. The measurement of credit risk and the analysis of default interdependence between borrowers in credit portfolios have become more important than ever in the practices of financial institutions, as was highlighted by the financial crisis and credit concentration (large-lot borrowers and sector concentration) and same as the measurement of market risk and liquidity risk. The US subprime mortgage crisis of 2007 triggered the global financial crisis of 2008, which negatively impacted the global credit market and equity market leading to a decline in global stock prices. Consequently, strict regulations and market practices have been imposed on financial institutions in the following years. Therefore, the measurement of credit risk is essential for the management of financial institutions.

Alternatively, investment strategies for credit products typified by bonds and credit default swaps (CDSs) are being developed day-to-day, the same as interest rates and stocks, and utilized to make profits in the trading desk of financial institutions. Transactions and products handled by financial institutions have increased complexity and diversification year by year. Compliance has also become more stringent. In these situations, it is important for financial institutions to ensure more secure profit sources without violating regulations, develop new investment strategies, and manage risk better.

The occurrence time of default and the magnitude of loss incurred at that time are the most important and significant characteristics of credit risk. Due to these characteristics, the nature of risk and the method of modeling in credit risk are different from those in market risk (loss caused by price fluctuations in the market) and liquidity risk (funding liquidity risk or market liquidity risk). Although credit risk losses are rare, these can have a large negative impact and cause modeling difficulties because of the lack of empirical data. Furthermore, the evaluation of credit portfolios is required in practice. For modeling complexity and rare event, the model validation is also difficult.

Covariation and contagion of credit risk (credit risk for multiple entities) are a more advanced concept than credit risk for a single entity. Although the evaluation is more difficult, its concept and method are applied to the risk management for credit portfolios and the investment strategy in credit products in financial institutions. It is necessary to take into account the relationship and interaction between reference entities when considering credit risk for multiple entities. Covariation and contagion of credit risk are included in the concept of these relationship and interaction. Covariation means that when one of two variables changes, the other also changes. Credit contagion means the phenomenon that a chain of defaults and declines in prices occurs by a propagation mechanism working between variables. It is important for prac-

titioners how to develop the model, capturing accurately the nature of the problem in reality, for covariation and contagion of credit risk. Realistic modeling leads to generating profits and avoiding losses for financial institutions in the future.

There are two conceptually different approaches in credit risk modeling. They are the structural approach (Merton [53]) and the reduced-form approach (Duffie and Singleton [21]). In the structural approach, we model the occurrence of the default event explicitly to the financial structure of the company based on asset, liability, and equity. A default occurs when the asset value goes below the liability value (i.e., the equity value goes below zero). In the reduced-form approach, we give the distribution of default time exogenously and regard the default of company as an event caused by an exogenous shock. Specifically, the credit event of reference entity is defined as a jump only once, and a hazard rate is defined as the intensity process of the jump under the conditions that a default does not occur by that time.

The Merton model [53] and the copula model [51] are standard models used in credit risk modeling for the structural approach. CreditMetrics [30] and CreditRisk+ [15] are also well-known in this field and are applied in practice. In the reduced-form approach, the pricing models of defaultable bonds and CDSs discussed by Duffie and Singleton [21] and Lando [50] are mentioned and are applied to the investment strategy and risk management for credit products in financial institutions. In this study, the concept and method of both the structural and reduced-form approach are used to model covariation and contagion of credit risk and applied to the risk management for credit portfolios and the investment strategy in credit products.

While many models for credit risk have been studied in the field of finance, the long-range Ising model for credit risk modeling in credit portfolios such as Molins and Vives [54] [55] and Kitsukawa et al. [48] have been studied in the field of econophysics. These studies are based on a statistical mechanical approach that explains macro phenomena (portfolio loss) in terms of numerous micro phenomena (state of borrowers in default or non-default) using a stochastic model. The measurement of credit risk for portfolios and the derivation of loss distributions (Filiz et al. [25] and Brunel [9]) have been studied in the field of information science as approaches to this problem in other fields.

In general, it is rare that the solutions of the credit risk model are obtained in closed form. The solutions may usually be calculated using the numerical calculation method of simulation and numerical integration. Solution means risk values and model parameters estimated and solved under some problem setting. The numerical calculation method (estimation method) can be roughly divided into two categories: the Bayesian inference and non-Bayes methods. The Bayesian inference uses the Markov chain Monte Carlo (MCMC) method based on Bayesian statistics and statistical physics. This study uses the MCMC consistently as the numerical calculation method. The advantages will be explained throughout this paper. In contrast, non-Bayes methods are based on traditional statistics. For example, Duan and Fulop [17] conducted the measurement of credit risk by the Merton model in the structural approach using the maximum likelihood estimation combined with the particle filter.

The MCMC is used to sample points from a probability distribution using random numbers when numerically calculating the solution of model. The most basic method to generate random numbers is a form of random sampling from a probability distribution. However, it is difficult to sample uniformly a high-dimensional space. Therefore, Gibbs sampler and the Metropolis-Hastings (MH) algorithm were invented to sample points selectively from an important region that has a large probability to solve this problem. Although the MH algorithm is still widely used, it is difficult to sample points from a multi-modal distribution, and the drop in acceptance rate is significant in a large number of parameters. Thus, the extended ensemble Monte Carlo method [37] including the replica exchange Monte Carlo (RMC) method [36] was invented to solve the former, and the Hamiltonian Monte Carlo (HMC) method [18] [57] was invented to

solve the latter.

Consequently, it has enormous significance to study the modeling for covariation and contagion of credit risk and the advancement of numerical calculation method using the MCMC for risk management and investment strategy from a practical point of view in financial institutions.

1.2 Purpose of study

This study proposes the models for covariation and contagion of credit risk by developing and advancing the numerical calculation methods using the Bayesian statistical inference method (MCMC) based on Bayesian statistics and statistical physics in credit risk modeling. This study aims to gain a deeper understanding of covariation and contagion of credit risk through the analysis using the models and to apply them in practice. This study consists of three subjects as follows.

Analysis of covariation and contagion of credit risk using the Ising model

We construct the long-range Ising model, which is used to model ferromagnetics in statistical physics, for credit risk modeling in heterogeneous credit portfolios held by financial institutions. The tail of the loss distribution of the long-range Ising model, developed by Molins and Vives [54] [55] and Kitsukawa et al. [48], has characteristics different from that of the standard model (Merton-type model, etc.) in credit risk modeling because the long-range Ising model can capture a second peak in loss distributions. A second peak is unique to the Ising model and appears in large loss areas due to the collective behavior of borrowers, which cannot be captured by the standard model. Therefore, there is a possibility of different evaluations of credit risk.

It is remarkable that the standard credit risk models cannot give a theoretically accurate loss distribution in the limit where the default correlations between all borrowers are one (perfect correlation) because of the non-existence of a second peak in loss distributions. That is, the standard models do not satisfy the obvious boundary condition in this limit. Therefore, there is a possibility that the standard models cannot accurately evaluate a rare event (tail phenomenon), a situation of high default correlations and confidence level, and a default cooperative phenomenon. This fact is a critical defect that could erode the accuracy of the measurement results of credit risk (furthermore, capital adequacy ratio) currently measured and published by financial institutions. In contrast, the credit risk model for credit portfolios using the long-range Ising model satisfies the obvious boundary condition in this limit because it can capture a second peak of loss distributions. This difference between the two models is very important because of one of few clear clues related to credit risk modeling.

However, existing studies in econophysics target only credit portfolios where credit exposures are homogeneous. There are no studies targeting the credit risk model for heterogeneous credit portfolios, which is available in practice. For the heterogeneity of the default probability and the default correlation, the characteristics of a model have not been sufficiently revealed based on numerical calculation. To solve these problems, this study constructs the credit risk model for heterogeneous credit portfolios (practical credit portfolios) held by financial institutions using the ferromagnetic Ising model based on techniques of statistical mechanics used in condensed matter physics from the perspective of econophysics. Then, we evaluate the credit risk and analyze the default interdependence by applying our model to various credit portfolios.

Under the significance of using the Ising model as mentioned above, this study aims to apply the Ising model to credit risk modeling for practical credit portfolios and expands homogeneous credit portfolios targeted by existing studies to heterogeneous credit portfolios from a practical perspective as a subject of study as credit portfolios are heterogeneous in practice.

Existing studies derive the loss distribution of homogeneous credit portfolios as an analytical solution. However, it is necessary to use numerical calculations as it is difficult to derive the loss distribution of heterogeneous credit portfolios analytically. This study numerically calculates the loss distribution using the RMC method because a multi-modal loss distribution (canonical distribution followed by the state of spins) induces the problem of slow relaxation. It is difficult to calculate multi-modal loss distributions using the Metropolis method, which is a typical MCMC.

This study aims to apply methods of statistical physics to finance from the perspective of econophysics and is an interdisciplinary study targeting the boundary region between finance and physics. It is possible to form a new academic field organized as scientific knowledge applying newly techniques of statistical mechanics used in physics to credit risk modeling for credit portfolios in finance. Techniques of statistical mechanics are applied to the various problems of information science, including error-correcting code, image restoration, brain science (associative memory and learning), and combinatorial optimization problem, for non-physical phenomena in the field of information statistical mechanics as explained in Nishimori [58]. As there are some application examples to finance, it could be strongly expected to apply to this subject. Technological innovation would advance rapidly for the acceleration of computers and the advancement of numerical calculation in the future. Therefore, it could be expected that the application value of this model would increase.

Analysis of covariation and contagion of credit risk based on cointegration

We analyze the cointegration of CDSs through hazard rates and apply it to pairs trading. The first aim is to understand the cointegration between the term structures of CDS spreads (CDS spreads with multiple tenors) through the cointegration between hazard rates. In the existing studies related to cointegration, there are financial market analyses based on the cointegration between a bond and its corresponding CDS (Blanco et al. [7] and Guidolin et al. [29]) and the study discussing cointegration in the framework of arbitrage-free pricing (Nakajima and Ohashi [56]). However, existing studies have the following problems: no study has been done on the cointegration analysis between CDSs and the error correction model (ECM) of hazard rates; many approaches select the pairs of financial products ad-hoc and test for presence/absence of cointegration; cointegration has not been discussed in the framework of arbitrage-free pricing although it is possible for bonds and CDSs to be priced unlike stocks. This study discusses the cointegration of CDSs in the framework of arbitrage-free pricing and construct the vector error correction model (VECM) for minimal modeling of cointegrated hazard rates that can explain CDS spreads, which can be priced.

The second aim is to apply the cointegration of hazard rates driving CDSs to the pairs trading of CDSs. Existing studies assess the pairs trading strategy of stocks based on cointegration (Higashide et al. [33] and Tourin and Yan [70]). Although it is not a pairs trading strategy, some studies have been done on the market-neutral strategy related to the term structure of interest rates and CDSs (Bali et al. [3] and Jarrow et al. [39]). However, existing studies have the following problems: there are no studies applying the cointegration of CDSs to an investment strategy and focusing on the cointegration between the term structures of interest rates and CDSs. This study establishes a trading methodology based on cointegration that is consistent with an arbitrage-free pricing framework.

Although the theoretical spread of cointegrated CDSs depending on such latent hazard rates does not have an analytical formula, the partial differential equations (PDEs) related to a theoretical CDS spread can be attributed to the ordinary differential equations (ODEs) by assuming an exponential affine model. However, because the ODEs cannot be analytically solved, we develop the Bayesian inference combined with the Runge-Kutta method to numeri-

cally solve ODEs (ODE-based Bayesian inference) to estimate the cointegration dynamics for hazard rates.

Expansion of ODE-based Bayesian inference to PDE-based Bayesian inference

We analyze corporate credit risks based on the stock value, developing the process followed by the asset value into the constant elasticity of variance (CEV) model [11] [12] in the structural approach of Merton based on assets, liabilities, and stocks. As the PDE satisfied by the stock value cannot be analytically solved and attributed to the ODEs, we focus on developing a Bayesian inference method from a different perspective. Specifically, by expanding the ODE-based Bayesian inference in the previous subject, we develop the Bayesian inference combined with the finite difference method (FDM) to numerically solve PDEs (PDE-based Bayesian inference) to estimate the CEV dynamics for asset values.

Duan and Fulop [17] solve the problem of unobservable asset values that is a main defect in the structural credit risk modeling. The parameters can be obtained since the observation equation can be analytically expressed using the Black-Scholes (BS) model based on the Merton model. However, as the system equation becomes more complex, the parameters cannot be estimated. For example, if the system equation is the CEV model, the parameters cannot be estimated because the option value that represents the observation equation cannot be described as analytical forms.

This study suggests the method to estimate the parameters in the CEV model using the PDE-based Bayesian inference to solve this problem. The CEV model is used rationally to increase the degree of freedom of the model selection by generalizing the BS model in the Merton model and making our model fitted better into the observation data. The range of the elastic constant is expanded from 1 to [0,1]. This study describes the observation equation as the curve of option value obtained by solving PDEs. We measure corporate credit risk by applying this method to the stock price data.

It is easy to use the MCMC in cases where the observation equation can be described as an analytical formula; for example, the BS model of stocks, the Vaciek model [71], and the Cox-Ingersoll-Ross (CIR) model [13] of interest rates. However, it is difficult to use the MCMC because the observation equation cannot be described as an analytical formula; for example, the CEV model and the stochastic volatility model. This study estimates the distribution of parameters using the PDE-based MCMC in the state space model, so that we become able to calculate the parameters of the model in which the observation equation cannot be described as an analytical formula and the parameters cannot be calculated. Our method, which combines the Bayesian inference with the PDEs, has not been suggested in the past.

1.3 Organization of this paper

Figure 1.1 shows the relationship between Chapters 2 to 4 that are the contribution of this study as the overhead view of study. Focusing on the analysis of covariation and contagion of credit risk and the development of numerical calculation method using the Bayesian inference (MCMC), this study works on the following three subjects. It constructs the long-range Ising model for credit risk modeling in heterogeneous credit portfolios and measures the credit risk of credit portfolios with heterogeneous credit exposures; analyzes the cointegration of CDSs through that of hazard rates and applies it to pairs trading; and measures corporate credit risks based on the stock values by using the structural approach expanded to the CEV model. Each subject conducts the numerical calculation from the view of development to application of the model and considers the results.

In this study, Chapter 2 is based on Kato [44]; Chapter 3 is based on Kato and Nakamura [45]; Chapter 4 is based on Kato and Nakamura [46]. The contents of each chapter will be summarized below.

Chapter 2 proposes a finite-size long-range Ising model as a model for heterogeneous credit portfolios held by financial institutions from the perspective of econophysics. This model expresses the heterogeneity of the default probability and the default correlation by dividing a credit portfolio into multiple sectors characterized by credit rating and industry. It also expresses the heterogeneity of the credit exposure, which is difficult to evaluate analytically, by applying the RMC method to numerically calculate loss distributions for a multi-modal distribution followed by spin states generated due to the interaction between spins (states of borrowers). We apply this model to various credit portfolios and evaluate credit risk to analyze the characteristics of loss distributions for credit portfolios with heterogeneous credit exposures. As a result, we show that the tail of loss distributions calculated by this model has characteristics different from that by the standard models used in credit risk modeling. We also show that there is a possibility of different evaluations of credit risk according to the pattern of heterogeneity.

Chapter 3 examines the cointegration relationship between multiple CDS spreads by constructing the cointegrated hazard rate model, which assumes the structure of the VECM in the drift term of hazard rate processes. We merge the cointegration nature into the framework of arbitrage-free pricing and thereby derive the theoretical spread formula of multiple cointegrated CDSs. For the estimation of hazard rate dynamics, we develop a Bayesian statistical inference method combined with the numerical ODE solver because the theoretical CDS spread cannot be expressed in closed form. In the empirical study of Japanese corporate CDSs, we find that the overall term structures of cointegrated CDSs (CDSs with multiple tenors) can be explained by a simple two-dimensional VECM of cointegrated hazard rates. Furthermore, this model can be applied to the pairs trading strategy of CDSs.

Chapter 4 proposes a method to infer the parameters of the CEV model from the market value of stocks after the extension from the asset process of the Merton model in the structural credit risk model to that of the CEV model. The state space model is used, which consists of the asset process (system equation) and the call option pricing a stock value (observation quation), for the inference. However, it is usually difficult to apply the MCMC to the estimation of the parameters of the CEV model because the observation equation of this state space model has no analytical formula. Our method solves this parameter estimation problem by applying the MCMC combined with the FDM of PDEs, where the stock value obtained as a CEV option price is numerically solved. As an empirical analysis, we estimate the parameters from the real stock values of the US financial institutions. Furthermore, we analyze the default probability and measure the credit risk of bank portfolios, while comparing the CEV model and the Merton model.

Chapter 2 studies covariation and contagion of credit risk between borrowers through the Ising model focusing on the measurement of credit risk for heterogeneous credit portfolios in financial institutions. Comparatively, Chapter 3 studies covariation and contagion of credit risk between CDSs through the VECM (cointegration) of hazard rates focusing on the application to the pairs trading strategy of CDSs. Changing our perspective, Chapter 4 expand the ODE-based MCMC developed in Chapter 3 to the PDE-based MCMC in the numerical calculation method using the Bayesian inference focusing on the measurement of corporate credit risks based on assets, liabilities, and equities.

Comparing credit risk modeling in Chapters 2 and 3, Chapter 2 brings the Ising model used for the modeling of spins that compose matter in condensed matter physics into finance and applies the model to the state of borrowers (default or non-default) that compose a credit portfolio. Then, Chapter 2 calculates a value at risk (VaR) and conditional value at risk

(CVaR, expected shortfall), which are macroeconomic indicators. In contrast, in Chapter 3, we understand the behavior of cointegration of CDSs by applying the VECM to the modeling for hazard rates that compose a CDS. A CDS spread, the present value of CDS, and that of its pair correspond to a macroeconomic indicator.

Comparing the numerical calculation methods of the models, Chapter 2 uses the RMC method to solve the problem that the Metropolis method does not work because of the multimodality of loss distributions (probability distributions followed by the states of borrowers). Chapter 3 uses the HMC method to solve the problem that the MH method does not work because of a large number of the model parameters and latent variables. Furthermore, we develop the HMC method combined with the numerical calculation method for the ODEs because the observation equation of the state space model cannot be solved analytically. In contrast, Chapter 4 develops the HMC method combined with the numerical calculation method for the PDEs because the ODEs cannot be even obtained as in Chapter 3.

Comparing the reference entities of credit risk, Chapter 2 deals with the borrowers that compose credit portfolios. We target the reference of CDSs in Chapter 3 and the issuer of stocks in Chapter 4.

Chapter 5 concludes our studies and mentions some concluding remarks. Appendix addresses the basic theory of extended ensemble Monte Carlo method as well as the derivation and specification of the model.

Item	Chapter 2	Chapter 3	Chapter 4
Financial products of target	Credit portfolios	CDSs	Stocks
Portfolio	Borrowers of bank	Companies in the same industry	Banking industry
Application of model	Risk management	Investment strategy	Risk analysis
Risk index	VaR, CVaR	-	Default probability
Credit risk modeling	Structural approach (Merton [1974])	Reduced-form approach (Duffie and Singleton [1999])	Structural approach (Merton [1974])
Basic model in credit risk modeling of covariation and contagion	Long-range Ising model	State space model (VECM and Arbitrage-free pricing)	State space model (Merton model and CEV model)
Intended purpose of model	Capture of second peak	Expression of the cointergration of CDSs through that of hazard rates	Improved model fitting by relaxing the restrictions of elastic constant
Background of model	Statistical physics	Mathematical finance	Mathematical finance
Comparable model	Standard credit risk model (Merton based model)	-	Black-Scholes model
Numerical calculation method	Replica exchange Monte Carlo method	ODE-based Bayesian inference (HMC combined with Runge-Kutta)	PDE-based Bayesian inference (HMC combined with FDM)
Intended purpose of method	Heterogeneous credit exposures, Multimodality of loss distribution	Multidimensional Riccati ODEs, Large number of parameters	Numerical calculation of PDEs, Large number of parameters
Background of method	Statistical mechanics	Analytical mechanics	Analytical mechanics
Implementation/Programming	VBA, C++	R, RStan	R, RStan
Hamiltonian	Potential energy (Internal energy)	Potential energy and kinetic energy	Potential energy and kinetic energy
Stationary distribution of MCMC	Loss distribution (Canonical distribution)	Distribution of parameters and latent variables	Distribution of parameters and latent variables

Figure 1.1: Overhead view of study. The themes of this study are the analysis of covariation and contagion of credit risk and the development of the numerical calculation method using the Bayesian inference (MCMC). The idea within the dotted lines is important.

Chapter 2

Long-range Ising Model for Credit Portfolios with Heterogeneous Credit Exposures

2.1 Introduction

Many models for credit risk have been studied in the field of financial engineering, such as the Merton model [53] which is a standard model used in credit risk modeling. On the other hand, in the field of econophysics, models for credit risk such as Molins and Vives [54] and Kitsukawa et al. [48] have been studied. Both of these studies examine the long-range Ising model for credit risk modeling in homogeneous credit portfolios. Molins and Vives [54] derive the loss distribution for credit portfolios expressing the state of borrowers (default or non-default) by Ising spin variables and using the maximum statistical entropy principle in information theory. They show that when the default correlation increases, a first-order-like transition may occur inducing a sudden risk (VaR) increase. This is because this model can capture the existence of a second peak in the large loss region of the loss distribution, which is unique to the Ising model and is not measured by the standard models. Kitsukawa et al. [48] express the loss distribution for credit portfolios by the superposition of two binomial distributions using a perturbative calculation method and derive explicit expressions for the default probability and the default correlation in terms of the exchange interaction and the external field. They discuss the effect of the default correlation on the average loss rate of a tranche in securitization, which is important for pricing. The model proposed in Molins and Vives [54] is improved and applied to heterogeneous credit portfolios in Molins and Vives [55]. Molins and Vives [55] derive the Jungle model, which is equivalent to the Ising model in the physics literature, as the probability distribution followed by default indicator variables expressing the state of borrowers (default or non-default) by default indicator variables and using the maximum statistical entropy principle. They calculate the loss distribution and the VaR for the Dandelion model (credit portfolios where a central element is connected with the rest of the nodes and no other pair of nodes is connected) and the Diamond model (homogeneous credit portfolios), which are particular cases of the Jungle model, and analyze the relation between the model parameters and the default probability, the default correlation. As a result, they show that the Jungle model generates doubly peaked loss distributions and quasi-phase transitions, and that the Jungle model can model contagion among borrowers. Here, these studies in the field of econophysics are based on a statistical mechanical approach that explains macro phenomena (portfolio loss) in terms of numerous micro phenomena (state of borrowers) using a stochastic model. There are approaches to this problem in other fields. Filiz et al. [25] propose the graphical model for credit risk modeling in credit portfolios where the default probability and the default correlation are

heterogeneous. Brunel [9] derives the loss distribution for credit portfolios, the loss distribution for asset-backed securities (ABS), and the distribution for recovery rates from the maximum statistical entropy principle.

The tail of the loss distribution of the long-range Ising model has characteristics different from that of the standard model because the long-range Ising model can capture the second peak of the loss distribution, which cannot be captured by the standard model. Therefore, there is a possibility of different evaluations of credit risk. The existing studies show approaches considering the heterogeneity of the default probability and the default correlation. These approaches are applied to particular credit portfolios using an analytic formula. Although there are studies mentioning approaches to the heterogeneity of the credit exposure, they have not sufficiently revealed the characteristics of a model because the loss distribution for credit portfolios has never been specifically evaluated. To solve these problems, we construct a finite-size long-range Ising model for credit risk modeling in heterogeneous credit portfolios held by a financial institution. We expand homogeneous credit portfolios to heterogeneous credit portfolios from a practical perspective as a subject of study because credit portfolios are heterogeneous in practice. Specifically, we incorporate the heterogeneity of the default probability, the default correlation, and the credit exposure of borrowers in the long-range Ising model. Our model consists of two parts: the sector model and an algorithm to numerically calculate the loss distribution. In the sector model, a credit portfolio is divided into multiple sectors characterized by credit rating and industry, thereby considering the heterogeneity of the default probability and the default correlation. Kitsukawa et al. [48] suggest such an approach. In the algorithm to numerically calculate the loss distribution, the RMC method [36] is used, which considers the heterogeneity of the credit exposure. The loss distribution for credit portfolios where the credit exposure is heterogeneous is difficult to calculate analytically, whereas the loss distribution for credit portfolios where the credit exposure is homogeneous can be calculated analytically. Therefore, we adopt the method that is not analytical but numerical.

To analyze the characteristics of the loss distribution for credit portfolios with heterogeneous credit exposures, we evaluate the credit risk by applying our model to various credit portfolios. The loss distribution and the VaR calculated by our model are compared with those of the standard model. Additionally, we analyze the variation of the loss distribution and the VaR, especially, the variation of the second peak (shape of the tail of the loss distribution) from homogeneous credit portfolios to credit portfolios with heterogeneous credit exposures. As a result, we show that the tail of the loss distribution calculated by our model has characteristics that are different from that of the standard models used in credit risk modeling and there is a possibility of different evaluations of credit risk according to the pattern of heterogeneity. To evaluate the credit risk for credit portfolios where the default probability and the default correlation are heterogeneous, we have to develop the calibration method of parameters for determining the exchange interactions and the external fields, which are input parameters for the Ising model, from the default probabilities and the default correlations that are given in the future. Therefore, as a reference, we evaluate the credit risk for credit portfolios that are a linear combination of two homogeneous credit portfolios, the parameters of which can be calculated.

The organization of Chapter 2 is as follows. In Section 2.2, we explain the long-range Ising model for credit risk modeling in heterogeneous credit portfolios; specifically, the derivation of this model and the algorithm to numerically calculate the loss distribution is explained. Section 2.3 is devoted to the calculation of credit risk using this model and the analysis of the result. Finally, in Section 2.4, we conclude and mention potential problems.

2.2 The model

The long-range Ising model for credit risk modeling in heterogeneous credit portfolios consists of two parts: the sector model to consider the heterogeneity of the default probability and the default correlation, and the algorithm to numerically calculate the loss distribution to consider the heterogeneity of the credit exposure.

2.2.1 Sector model

General heterogeneous credit portfolios

Let us consider general heterogeneous credit portfolios with N borrowers set to the heterogeneous default probability and default correlation. The default indicator variable $X_i (i = 1, \ldots, N)$ expresses the state of the i-th borrower and takes values 0 or 1 depending on whether the i-th borrower has not defaulted or has defaulted. The Ising spin variable $S_i (i = 1, \ldots, N)$ also expresses the state of the i-th borrower and takes values 1 or -1 depending on whether the i-th borrower has not defaulted or has defaulted. S_i is related to X_i as $S_i = 1 - 2X_i$, and $S = (S_1, \ldots, S_N)$ expresses the state of a credit portfolio (the microscopic state of an Ising spin system). Let us denote the marginal default probability of the i-th borrower as p_i (defined by $\langle X_i \rangle$) and the joint default probability of the i-th and the j-th borrower as $p_{i,j}$ (defined by $\langle X_i X_j \rangle$). $\langle \ldots \rangle$ represents a thermal average value. Therefore, the default correlation between the i-th and the j-th borrower $\rho_{i,j}$ is defined by

$$\rho_{i,j} = \frac{p_{i,j} - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}, \tag{2.1}$$

where the marginal default probability and the default correlation are exogenous variables. From (2.1), the expected value of the Ising spin variable is related to the marginal default probability and the default correlation in the following way:

$$\langle S_i \rangle = 1 - 2\langle X_i \rangle = 1 - 2p_i, \tag{2.2}$$

$$\begin{aligned} \langle S_i S_j \rangle &= 1 - 2\langle X_i \rangle - 2\langle X_j \rangle + 4\langle X_i X_j \rangle \\ &= 1 - 2p_i - 2p_j + 4p_i p_j + 4\rho_{i,j}\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}. \end{aligned} \tag{2.3}$$

According to Molins and Vives [54], Molins and Vives [55], and Brunel [9], we derive the probability distribution $P(S)$ followed by the state of the credit portfolio (Ising spin variables) using the maximum statistical entropy principle in information theory where the default probability and the default correlation are heterogeneous. We derive the probability distribution $P(S)$ that maximizes the entropy functional

$$S[P] = -k_B \sum_{S \in \Omega_S} P(S) \ln P(S) \tag{2.4}$$

with three constraints

$$\sum_{S \in \Omega_S} P(S) = 1, \tag{2.5}$$

$$\sum_{S \in \Omega_S} S_i P(S) = 1 - 2p_i, \quad i = 1, \ldots, N, \tag{2.6}$$

$$\sum_{S \in \Omega_S} S_i S_j P(S) = 1 - 2p_i - 2p_j + 4p_i p_j + 4\rho_{i,j}\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}, \quad i = 1, \ldots, N, j = 1, \ldots, N, i < j, \tag{2.7}$$

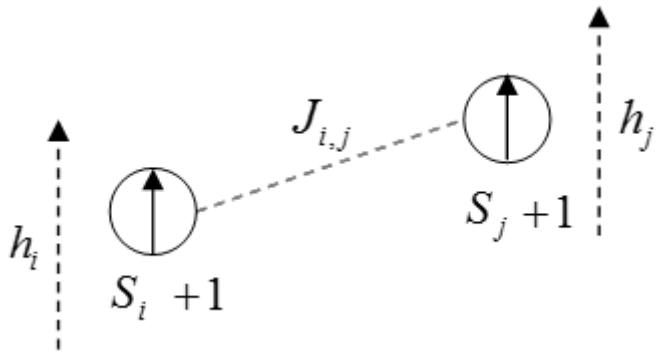


Figure 2.1: Ising model of two spins system of the i-th spin S_i and the j-th spin S_j with the exchange interaction $J_{i,j}$ and the external field h_i.

where the constraints are the normalization constraint (2.5), the expected value constraint (2.6) from (2.2), and the covariance constraint (2.7) from (2.3), respectively. Then, $\Omega_S = \{S\}$ is the set of all states of a credit portfolio (all microscopic states of an Ising spin system), and k_B is the Boltzmann constant. Let us consider the optimization problem the objective function of which is $S[P]/k_B$, which is entropy (2.4) divided by the Boltzmann constant. The probability distribution that maximizes the Lagrangian function

$$\begin{aligned}\Psi[P, \lambda, a_1, \ldots, a_N, b_{1,2}, \ldots, b_{N-1,N}] &= -\sum_{S\in\Omega_S} P(S)\ln P(S) \\ &+ \lambda\{1 - \sum_{S\in\Omega_S} P(S)\} + \sum_{i=1}^{N} a_i\{1 - 2p_i - \sum_{S\in\Omega_S} S_i P(S)\} \\ &+ \sum_{i<j} b_{i,j}\{1 - 2p_i - 2p_j + 4p_ip_j + 4\rho_{i,j}\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)} - \sum_{S\in\Omega_S} S_iS_jP(S)\}\end{aligned} \tag{2.8}$$

is

$$P(S) = \frac{1}{Z(\beta)}e^{-\beta H(S)} = \frac{1}{Z(\beta)}\exp\left[\beta\sum_{i<j} J_{i,j}S_iS_j + \beta\sum_{i=1}^{N} h_iS_i\right], \tag{2.9}$$

where the Lagrange multipliers $a_i, b_{i,j}$ in (2.8) are

$$\begin{cases} a_i = -\beta h_i, \quad i = 1, \ldots, N \\ b_{i,j} = -\beta J_{i,j}, \quad i = 1, \ldots, N, j = 1, \ldots, N, i < j \end{cases}, \tag{2.10}$$

and the probability distribution $P(S)$ of (2.9) is the canonical distribution with the exchange interaction $J_{i,j}$ between the i-th and the j-th spin, the external field h_i applied to the i-th spin, and the inverse temperature β ($\beta = (k_BT)^{-1}$ with the temperature T). As a simple example, Figure 2.1 shows the Ising model of two spins system. We propose an algorithm to numerically calculate the loss distribution for credit portfolios using the RMC method later in Subsection 2.2.2. In the RMC method, we set multiple replicas of the system to different temperatures T (inverse temperatures β), and this method simulates the loss distribution. Therefore, we introduce a superfluous extra multiplicative factor β in equation (2.10). Then, $H(S)$ is the Hamiltonian (energy function), and $Z(\beta)$ is the partition function (normalization factor of the

probability distribution followed by Ising spin variables):

$$H(S) = -\sum_{i<j} J_{i,j} S_i S_j - \sum_{i=1}^{N} h_i S_i, \tag{2.11}$$

$$Z(\beta) = \sum_{S \in \Omega_S} e^{-\beta H(S)} = \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \exp\left[\beta \sum_{i<j} J_{i,j} S_i S_j + \beta \sum_{i=1}^{N} h_i S_i\right]. \tag{2.12}$$

As will be shown later in this section, this probability distribution (2.9) is followed by the state of general heterogeneous credit portfolios which include homogeneous credit portfolios addressed in Molins and Vives [54] and Kitsukawa et al. [48], and is also equivalent to the Jungle model addressed in Molins and Vives [55] and the graphical model addressed in Filiz et al. [25] by changing the variables between the default indicator and the spin.

Sector model

We derive the sector model from the Ising model for general heterogeneous credit portfolios. A general heterogeneous credit portfolio is set to a different default probability for each borrower and a different default correlation for each pair of borrowers. However, it is difficult to calculate the loss distribution in view of the calculation load due to the use of a numerical calculation in our model. Therefore, let us consider the sector model constructed by dividing a credit portfolio into multiple sectors characterized by credit rating and industry. A credit portfolio is divided into groups based on credit rating, wherein the default probability of borrowers in each group is homogeneous. A credit portfolio is also divided into groups based on industry, wherein the default correlation among borrowers in the same business area is homogeneous, and the default correlation among borrowers across the same pair of different business areas is homogeneous. However, in general, the default correlation among borrowers in one business area is not equal to that in another business area. The default correlation among borrowers across different business areas is also not equal to that in the business area. Additionally, the default correlation among borrowers across one pair of different business areas is not equal to that across another pair of areas. Thereafter, a sector is defined as a unique combination of credit ratings and industries. For example, if a credit portfolio is divided into three groups based on credit rating and into two groups based on industry, this credit portfolio should be expressed by a six-sector model. Each borrower in this credit portfolio is characterized by a pair of credit ratings and industries. The approach of the sector model is similar to business practice because we often classify borrowers into groups based on credit rating and industry for the practical measurement of credit risk. To derive the canonical distribution of the sector model from that of general heterogeneous credit portfolios, we set the parameters as follows:

$$\begin{cases} p_i = p^I, & i \in \Omega_I \\ \rho_{i,j} = \rho^I, & i,j \in \Omega_I, i<j \\ \rho_{i,j} = \rho^{I,J}, & i \in \Omega_I, j \in \Omega_J \end{cases} \Leftrightarrow \begin{cases} h_i = h^I, & i \in \Omega_I \\ J_{i,j} = \dfrac{J^I}{N^I}, & i,j \in \Omega_I, i<j \\ J_{i,j} = \dfrac{J^{I,J}}{\sqrt{N^I N^J}}, & i \in \Omega_I, j \in \Omega_J \end{cases}, \tag{2.13}$$

where Ω_I is the set of borrowers in the I-th sector; p^I is the marginal default probability of borrowers in the I-th sector; ρ^I is the default correlation among borrowers in the I-th sector; and $\rho^{I,J}$ is the default correlation among borrowers between the I-th and the J-th sector. Additionally, N^I is the number of borrowers in the I-th sector; J^I/N^I is the exchange interaction among spins in the I-th sector; $J^{I,J}/\sqrt{N^I N^J}$ is the exchange interaction among

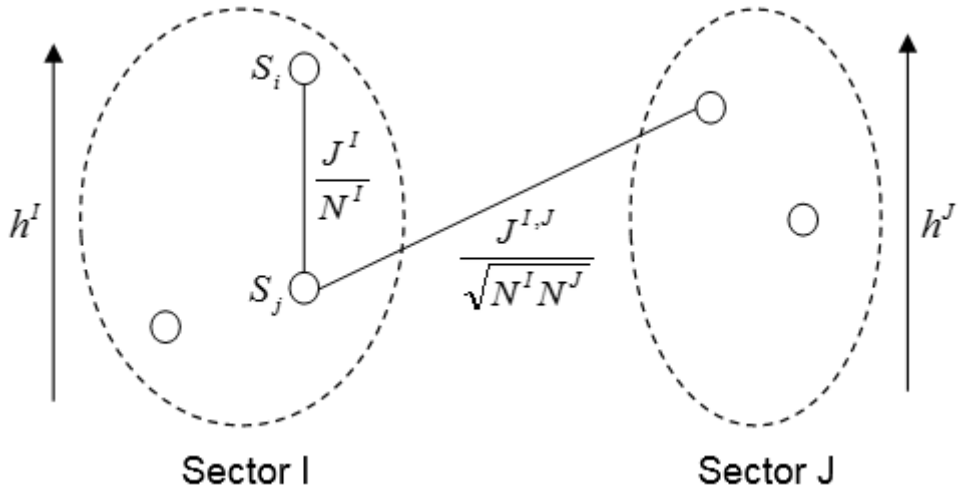


Figure 2.2: Sector model of two sectors portfolio of Sector I and Sector J with the number of borrowers N^I, the exchange interaction J^I/N^I, the external field h^I in the I-th sector, and the exchange interaction $J^{I,J}/\sqrt{N^I N^J}$ between the I-th and the J-th sector.

spins between the I-th and the J-th sector; and h^I is the external field applied to spins in the I-th sector.

By applying the parameters of the sector model (2.13) to the Ising model for general heterogeneous credit portfolios (2.9), we obtain the probability distribution (canonical distribution) $P(S)$ followed by the state of credit portfolios of the sector model (Ising spin variables) as follows:

$$\begin{aligned} P(S) &= \frac{1}{Z(\beta)} e^{-\beta H(S)} \\ &= \frac{1}{Z(\beta)} \exp\left[\beta \sum_{I=1}^{K} \frac{J^I}{N^I} \sum_{i,j\in\Omega_I, i<j} S_i S_j + \beta \sum_{I<J} \frac{J^{I,J}}{\sqrt{N^I N^J}} \sum_{i\in\Omega_I} \sum_{j\in\Omega_J} S_i S_j + \beta \sum_{I=1}^{K} h^I \sum_{i\in\Omega_I} S_i\right], \end{aligned} \tag{2.14}$$

where K is the number of sectors. As a simple example, Figure 2.2 shows the sector model of two sectors portfolio. From (2.11) and (2.12), the Hamiltonian $H(S)$ and the partition function $Z(\beta)$ are:

$$H(S) = -\sum_{I=1}^{K} \frac{J^I}{N^I} \sum_{i,j\in\Omega_I, i<j} S_i S_j - \sum_{I<J} \frac{J^{I,J}}{\sqrt{N^I N^J}} \sum_{i\in\Omega_I} \sum_{j\in\Omega_J} S_i S_j - \sum_{I=1}^{K} h^I \sum_{i\in\Omega_I} S_i, \tag{2.15}$$

$$\begin{aligned} Z(\beta) = \sum_{S\in\Omega_S} e^{-\beta H(S)} &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \\ &\exp\left[\beta \sum_{I=1}^{K} \frac{J^I}{N^I} \sum_{i,j\in\Omega_I, i<j} S_i S_j + \beta \sum_{I<J} \frac{J^{I,J}}{\sqrt{N^I N^J}} \sum_{i\in\Omega_I} \sum_{j\in\Omega_J} S_i S_j + \beta \sum_{I=1}^{K} h^I \sum_{i\in\Omega_I} S_i\right], \end{aligned} \tag{2.16}$$

respectively. We are able to set the heterogeneity of the default probability and the default correlation more precisely. Theoretically, we can set a default probability for every borrower and a default correlation for every pair of borrowers. The identical sector specified by credit rating and industry is a homogeneous credit portfolio.

Homogeneous credit portfolios

We derive the Ising model for homogeneous credit portfolios. A homogeneous credit portfolio is obtained by homogenizing the parameters of a general heterogeneous credit portfolio and is simultaneously equivalent to the credit portfolio in a sector model which has just one sector. To derive the canonical distribution of homogeneous credit portfolios from that of general heterogeneous credit portfolios, we set the parameters as follows:

$$\begin{cases} p_i = p, & i \in \Omega \\ \rho_{i,j} = \rho, & i,j \in \Omega, i < j \end{cases} \Leftrightarrow \begin{cases} h_i = h, & i \in \Omega \\ J_{i,j} = \dfrac{J}{N}, & i,j \in \Omega, i < j \end{cases}, \tag{2.17}$$

where Ω is the set of all borrowers in a credit portfolio; p is the marginal default probability of borrowers; ρ is the default correlation among borrowers; J/N is the exchange interaction among spins, and h is the external field applied to spins. All parameters are homogeneous in the credit portfolio.

By applying the parameters of a homogeneous credit portfolio (2.17) to the Ising model for general heterogeneous credit portfolios (2.9), we obtain the probability distribution (canonical distribution) $P(S)$ followed by the state of homogeneous credit portfolios (Ising spin variables) as follows:

$$P(S) = \frac{1}{Z(\beta)} e^{-\beta H(S)} = \frac{1}{Z(\beta)} \exp\left[\frac{\beta J}{N} \sum_{i<j} S_i S_j + \beta h \sum_{i=1}^{N} S_i \right], \tag{2.18}$$

where the Hamiltonian $H(S)$ and the partition function $Z(\beta)$ are:

$$H(S) = -\frac{J}{N} \sum_{i<j} S_i S_j - h \sum_{i=1}^{N} S_i, \tag{2.19}$$

$$Z(\beta) = \sum_{S \in \Omega_S} e^{-\beta H(S)} = \sum_{S_1 = \pm 1} \cdots \sum_{S_N = \pm 1} \exp\left[\frac{\beta J}{N} \sum_{i<j} S_i S_j + \beta h \sum_{i=1}^{N} S_i \right], \tag{2.20}$$

respectively.

This probability distribution coincides with that followed by the state of homogeneous credit portfolios addressed in Molins and Vives [54] and Kitsukawa et al. [48], and is also equivalent to the Diamond model addressed in Molins and Vives [55] by changing the variables between the default indicator and the spin.

2.2.2 Algorithm to numerically calculate loss distributions

As discussed in the previous subsection, we consider the heterogeneity of the default probability and the default correlation using the sector model. In this subsection, we consider the heterogeneity of the credit exposure using a numerical calculation by the RMC method. The loss distribution for credit portfolios where the credit exposure is heterogeneous is difficult to calculate analytically while that for credit portfolios where the credit exposure is homogeneous can be calculated analytically. This is because, if the credit exposure is heterogeneous, the loss amount in a credit portfolio depends on the state of spins even if the credit portfolio has the same number of defaulting borrowers. The difference in the credit exposure cannot be considered by the canonical distribution analytically. On the other hand, if the credit exposure is homogeneous, the loss amount in a credit portfolio depends only on the number of defaulting borrowers. The distribution of the number of defaulting borrowers can be calculated from the canonical distribution analytically. In other words, the default probability and the default correlation that are parameters of the Hamiltonian or the canonical distribution determine the

behavior of the system. The credit exposure does not determine the behavior of the system; it determines the size of loss from the structure of the credit exposure in a credit portfolio after the system behavior is determined. In addition, if the heterogeneity of the default probability and the default correlation is strong (e.g., the number of sectors is numerous), it is difficult to calculate the distribution of the number of defaulting borrowers analytically. Therefore, we develop the algorithm to numerically calculate the loss distribution by constructing a method that is not analytical but numerical. In this algorithm, we distinguish two parts by dividing the entire calculation process into the probability with which the state of spins emerges and the structure of credit exposures. Then, we calculate the loss distribution by calculating the loss amount from the state of spins sampled from the canonical distribution of the sector model and the structure of credit exposures.

In this algorithm, we adopt the RMC method, which is one of the extended ensemble Monte Carlo methods, and accelerate the relaxation. The reason is as follows. If credit portfolios are heterogeneous, the canonical distribution has numerous peaks because the minimum energy state or the states near it exist numerously away from each other in the high-dimensional space. As a result, a multi-modal loss distribution is observed. If the competitive complicated interaction induces the probability distribution to become multi-modal, the transitions from one peak to the other peak occur very rarely in the Metropolis method. Therefore, it induces the problem of slow relaxation, that is, samples keep staying in a local subspace in the high-dimensional space. In the RMC method, we set multiple replicas to different temperatures T (inverse temperatures β). This method simulates the probability distribution by updating the states of every replica using the Metropolis method and exchanging the states between replicas adjacent to each other. Under a low temperature ($T = \text{low}$, $\beta = \text{high}$), it is difficult for the state to transit into the high energy state, whereas, under a high temperature ($T = \text{high}$, $\beta = \text{low}$), the state can easily transit into the high energy state. The temperature T (inverse temperature β) is the index that expresses the degree of stochastic restrictions. As the state of each replica is updated at its given temperature, it moves around the temperature axis by the exchange of states. Therefore, even if the state is in the meta-stable state in the low temperature, the relaxation of the system is accelerated by heating once and escaping the meta-stable state. The explanation of the RMC method can be found in Appendix A.1. In this paper, although we set the loss given default (LGD) of borrowers to a uniform rate of 100%, it is easy to expand our model to take heterogeneous LGD into consideration.

The algorithm to numerically calculate loss distributions for the sector model and the credit portfolios where the credit exposure is heterogeneous is as follows:

Step 1: Generation of initial state

Let us denote the number of replicas as R; $\Theta = (\Theta_1, \ldots, \Theta_R)$ is the group of the states of spins; $\Theta_l = (S_1^l, \ldots, S_N^l)$ is the state of the spins in the l-th replica; and $\beta = (\beta_1, \ldots, \beta_R)$ is the group of inverse temperatures. Generate the group of the initial states of spins Θ randomly. Sample a uniform random number x from [0,1] and $S_i^l = +1$ if $x \geq 0.5$, $S_i^l = -1$ if $x < 0.5$. $(i = 1, \ldots, N, l = 1, \ldots, R)$

Step 2: Choice of the spin inverted virtually

Choose the spin S_k^l inverted virtually. The inverted candidate varies in turn, such as $k = 1, 2, \ldots, N, 1, 2, \ldots$. The spin S_k^l inverted virtually is the state of the k-th borrower, where Φ is the sector to which the k-th borrower belongs. $(\Phi = 1, \ldots, K, l = 1, \ldots, R)$

Step 3: Calculation of the field applied to the spin inverted virtually

Calculate the field (internal field and external field) applied to the spin inverted virtually

$$V_k^l = \frac{J^\Phi}{N^\Phi} \sum_{j \in \Omega_\Phi, j \neq k} S_j^l + \sum_{J=1, J \neq \Phi}^{K} \frac{J^{\Phi,J}}{\sqrt{N^\Phi N^J}} \sum_{j \in \Omega_J} S_j^l + h^\Phi, \tag{2.21}$$

where Ω_I is the set of borrowers in the I-th sector.

Step 4: Execution of spin flip
If the field V_k^l of (2.21) applied to the spin inverted virtually faces

- in the same direction as the inversion $S_k'^l(=-S_k^l)$, execute the spin flip, and
- in the opposite direction of the inversion $S_k'^l(=-S_k^l)$, sample a uniform random number x from [0,1] and compare it with the transition probability $e^{-\beta_l \Delta H_l}$. Execute the spin flip if $x < e^{-\beta_l \Delta H_l}$, and leave it as it is if $x \geq e^{-\beta_l \Delta H_l}$,

where $\Delta H_l = -2S_k'^l V_k^l$.

Step 5 Execute the state transition from Step 3 to Step 4 for all replicas independently ($l = 1, \ldots, R$).

Step 6: Choice of the pair of replicas adjacent to each other exchanged virtually
Choose the pair of replicas (Θ_l, Θ_{l+1}) adjacent to each other that are exchanged virtually.

- If the step number of replica exchange is odd, $(\Theta_l, \Theta_{l+1}) \in \{(\Theta_l, \Theta_{l+1}) : l = 1, \ldots, R-1, l \text{ is odd}\}$.
- If the step number of replica exchange is even, $(\Theta_l, \Theta_{l+1}) \in \{(\Theta_l, \Theta_{l+1}) : l = 1, \ldots, R-1, l \text{ is even}\}$.

Step 7: Calculation of Δ accompanied by exchanging the replicas virtually
Calculate the value

$$\Delta_{l,l+1} = -(\beta_{l+1} - \beta_l)\{H(\Theta_{l+1}) - H(\Theta_l)\} \tag{2.22}$$

accompanied by exchanging the replicas virtually, where the Hamiltonian $H(\Theta_l)$ is

$$H(\Theta_l) = -\frac{1}{2}\sum_{I=1}^{K}\frac{J^I}{N^I}(\sum_{i\in\Omega_I} S_i^l)^2 + \frac{1}{2}\sum_{I=1}^{K} J^I - \sum_{I<J}\frac{J^{I,J}}{\sqrt{N^I N^J}}(\sum_{i\in\Omega_I} S_i^l)(\sum_{j\in\Omega_J} S_j^l) - \sum_{I=1}^{K} h^I \sum_{i\in\Omega_I} S_i^l.$$

Step 8: Execution of replica exchange
With regard to the value $\Delta_{l,l+1}$ of (2.22) accompanied by exchanging the replicas virtually,

- if $\Delta_{l,l+1} \leq 0$, execute the replica exchange, and
- if $\Delta_{l,l+1} > 0$, sample a uniform random number x from [0,1] and compare it with the transition probability $e^{-\Delta_{l,l+1}}$. Execute the replica exchange if $x < e^{-\Delta_{l,l+1}}$, and leave it as it is if $x \geq e^{-\Delta_{l,l+1}}$.

Step 9 Execute the state transition from Step 7 to Step 8 for all pairs of replicas that are adjacent to each other.

Step 10: Calculation of loss amount
Calculate the loss amount for the chosen state:

$$L^l = \frac{1}{2}\sum_{i=1}^{N} A_i(1 - S_i^l), \quad 0 \leq L^l \leq \sum_{i=1}^{N} A_i, l = 1, \ldots, R, \tag{2.23}$$

where A_i is the credit exposure of the i-th borrower.

Step 11 Repeat the state transition from Step 2 to Step 10 sufficiently.

Step 12 Adopt the loss distribution $\{L^l\}$ of the needed inverse temperature ($\beta_l = 1$).

The state transition from Step 2 to Step 5 is updating the states of every replica independently (the Metropolis method in the narrow sense), and that from Step 6 to Step 9 is exchanging the states between replicas adjacent to each other. This algorithm can calculate the loss distribution for an arbitrary credit portfolio that can be described using the sector model where the exchange interaction, the external field, and the credit exposure are heterogeneous. This also includes the numerical calculation of the loss distribution for homogeneous credit portfolios. We use XorShift [52] for random number sequences. The setting of the RMC method (e.g., temperature parameters) in this paper can be found in Appendix A.2.

2.2.3 Model I proposed in this study

We propose the long-range Ising model for credit risk modeling in heterogeneous credit portfolios. This model consists of the two parts: the sector model and the algorithm to numerically calculate loss distributions. This model is a practical credit risk model because the heterogeneity of the default probability and the default correlation is considered in the sector model, and the heterogeneity of the credit exposure is considered in the algorithm to numerically calculate loss distributions. We summarize the credit risk model of (2.14), (2.15), (2.16), and (2.23) proposed in this study as follows.

- Model I
  - Sector model

$$\begin{cases} P(S_1, \ldots, S_N) = \dfrac{1}{Z(\beta)} \exp\left[-\beta H(S_1, \ldots, S_N)\right] \\ H(S_1, \ldots, S_N) = -\displaystyle\sum_{I=1}^{K} \frac{J^I}{N^I} \sum_{i,j \in \Omega_I, i<j} S_i S_j - \sum_{I<J} \frac{J^{I,J}}{\sqrt{N^I N^J}} \sum_{i \in \Omega_I} \sum_{j \in \Omega_J} S_i S_j - \sum_{I=1}^{K} h^I \sum_{i \in \Omega_I} S_i \end{cases}$$

  - Algorithm to numerically calculate loss distributions
    Sample the state of spins $(S_1, \ldots, S_N)$ randomly from the canonical distribution of the sector model $P(S_1, \ldots, S_N)$ using the RMC method.

$$\text{Loss distribution: } (L^{(1)}, \ldots, L^{(\Lambda)}), \quad L^{(k)} = \frac{1}{2} \sum_{i=1}^{N} A_i (1 - S_i), \quad k = 1, \ldots, \Lambda$$

$$\text{Expected loss: } \langle L \rangle \approx \frac{1}{\Lambda} \sum_{k=1}^{\Lambda} L^{(k)}$$

The trial number of simulation is denoted by Λ.

As mentioned above, this model can calculate the loss distribution for an arbitrary credit portfolio that can be described using the sector model where the exchange interaction, the external field, and the credit exposure are heterogeneous. Therefore, this model can calculate the loss distribution for an arbitrary credit portfolio where the credit exposure is heterogeneous. However, to evaluate the loss distribution for credit portfolios where the default probability and the default correlation are heterogeneous, we have to develop the calibration method of parameters for determining the exchange interactions and the external fields, which are input parameters for the Ising model, from the default probabilities and the default correlations that are given in the future.

2.3 Numerical calculation

To analyze the characteristics of the loss distribution for credit portfolios with heterogeneous credit exposures, we evaluate the credit risk by applying our model to various credit portfolios. We set the cases of credit portfolios used in the calculation of credit risk and calculate the loss distribution and the VaR numerically.

2.3.1 Cases of credit portfolios

Let us explain the cases of credit portfolios used in the calculation of credit risk. By varying the credit exposure, we enumerate four types of cases, Case 1 to Case 4. As a reference, we enumerate Case 5 as a heterogeneous default correlation case, which is a credit portfolio consisting of two independent homogeneous sectors, the parameters of which can be calculated. For comparison, all cases are set such that the number of borrowers and the total number of credit exposures in the credit portfolio are equal to each other. We describe the Hamiltonian, the loss amount, and the parameter because each case is characterized as the Hamiltonian and the loss amount.

Case 1: Homogeneous credit portfolio

This case considers the default probability, the default correlation, and the credit exposure to be homogeneous.

Hamiltonian: $H(S) = -\frac{J}{2N}(\sum_{i=1}^{N} S_i)^2 + \frac{J}{2} - h\sum_{i=1}^{N} S_i$

Loss amount: $L = \frac{1}{2}\sum_{i=1}^{N} A(1 - S_i) = \frac{1}{2}A(N - \sum_{i=1}^{N} S_i)$

Parameters: $N = 500, p = 0.02, \rho = 0.01, A = 1$

Case 2: Credit portfolio with one large exposure

In this case, the credit exposure of one borrower is large and that of the other borrowers is small.

Hamiltonian: Same as Case 1

Loss amount: $L = \frac{1}{2}A^1(N^1 - S_1) + \frac{1}{2}A^2(N^2 - \sum_{i=2}^{N} S_i)$

Parameters: $N = 500, N^1 = 1, N^2 = 499, p = 0.02, \rho = 0.01, A^1 = 250.5, A^2 = 0.5$

Case 3: Credit portfolio with two large exposures

In this case, the credit exposures of two borrowers are large and that of the other borrowers is small.

Hamiltonian: Same as Case 1

Loss amount: $L = \frac{1}{2}A^1(N^1 - \sum_{i=1}^{2} S_i) + \frac{1}{2}A^2(N^2 - \sum_{i=3}^{N} S_i)$

Parameters: $N = 500, N^1 = 2, N^2 = 498, p = 0.02, \rho = 0.01, A^1 = 125.5, A^2 = 0.5$

Case 4: Credit portfolio in which credit exposures follow a power-law distribution

In this case, the credit exposures are distributed according to a power-law distribution (credit exposure distribution close to that of existing business practices).

Hamiltonian: Same as Case 1

Loss amount: $L = \frac{1}{2}\sum_{i=1}^{N} A_i(1 - S_i)$

Parameters: $N = 500, p = 0.02, \rho = 0.01, A_i (i = 1, \ldots, N) \sim$ power-law distribution

Case 5: Credit portfolio consisting of two independent homogeneous sectors

In this case, the homogeneous credit portfolio is divided into two sectors, and the default correlation among borrowers across the two sectors is set to zero.

$$\text{Hamiltonian: } H(S) = -\frac{J^1}{2N^1}(\sum_{i=1}^{N^1} S_i)^2 - \frac{J^2}{2N^2}(\sum_{i=N^1+1}^{N} S_i)^2 + \frac{1}{2}(J^1 + J^2) - \frac{J^{1,2}}{\sqrt{N^1 N^2}}(\sum_{i=1}^{N^1} S_i)(\sum_{i=N^1+1}^{N} S_i) - h^1 \sum_{i=1}^{N^1} S_i - h^2 \sum_{i=N^1+1}^{N} S_i$$

Loss amount: Same as Case 1

Parameters: $N = 500, N^1 = 100, N^2 = 400, p^1 = p^2 = 0.02, \rho^1 = \rho^2 = 0.01, \rho^{1,2} = 0, A = 1$

Case 1 is homogeneous and is used for comparison with Case 2 to Case 5 and for confirmation that the loss distribution calculated by our model converges to that calculated by the analytical formula in existing studies. In Case 2, Case 3, and Case 4, the credit exposure of the credit portfolios is heterogeneous, but the default probability and the default correlation are homogeneous. In Case 2 and Case 3, the credit exposure in the credit portfolios is concentrated to a few borrowers. In Case 2, it is expected that the second peak is split into two peaks because the total loss amount of the credit portfolio varies drastically depending on whether or not one borrower who has the large credit exposure is included in a large-scale chain of defaults corresponding to the second peak. In Case 3, it is expected that the second peak is split into three peaks for the same reason. These cases are also used to explain the loss distribution for the credit portfolios where the credit exposures are distributed according to a power-law distribution. In Case 4, the credit exposures in the credit portfolios are distributed according to a power-law distribution. The shape of the substantial credit exposure distribution of financial institutions is similar to that of a power-law distribution. Case 5 is a credit portfolio consisting of two independent homogeneous sectors. We divide a credit portfolio into two sectors such that the default correlation among borrowers in a particular sector is homogeneous and equal to the default correlation among borrowers in the other sector, and the default correlation among borrowers across the two sectors is set to zero. The default probability and the credit exposure are homogeneous. In that sense, Case 5 is a credit portfolio with heterogeneous default correlations as a whole whereas it is a linear combination of two homogeneous credit portfolios. We enumerate this case as a reference to consider credit portfolios that have an intermediate default correlation among borrowers across the two sectors ($0 < \rho^{1,2} < 0.01$) in the future because Case 1 is $\rho^{1,2} = 0.01$ and Case 5 is $\rho^{1,2} = 0$. When the two sectors are independent ($J^{1,2}/\sqrt{N^1 N^2} = 0$ in the Hamiltonian in Case 5), the Hamiltonian is

$$H(S) = -\frac{J^1}{2N^1}(\sum_{i=1}^{N^1} S_i)^2 + \frac{J^1}{2} - h^1 \sum_{i=1}^{N^1} S_i - \frac{J^2}{2N^2}(\sum_{i=N^1+1}^{N} S_i)^2 + \frac{J^2}{2} - h^2 \sum_{i=N^1+1}^{N} S_i.$$

Both Sectors 1 and 2 constitute homogeneous credit portfolios. Therefore, we calculate the exchange interaction and the external field from the default probability and the default correlation for each homogeneous credit portfolio using the following relational expressions for homogeneous credit portfolios shown in Kitsukawa et al. [48]:

$$\begin{cases} p = \dfrac{\langle Y \rangle}{N} \\ \rho = \dfrac{\dfrac{\langle Y^2 \rangle - \langle Y \rangle}{N(N-1)} - p^2}{p(1-p)} \end{cases},$$

where Y is the total number of defaulting borrowers in a credit portfolio, and $P(Y)$ is the probability distribution followed by the number of defaulting borrowers which can be calculated by substituting $S_i = 1 - 2X_i$ for (2.18), (2.19), and (2.20):

$$Y = \sum_{i=1}^{N} X_i,$$

$$P(Y) = \frac{\exp\left[\dfrac{\beta N J}{2} - \dfrac{\beta J}{2} + \beta N h\right]}{Z(\beta)} {}_N\mathrm{C}_Y \exp\left[\frac{2\beta J}{N} Y^2 - 2\beta(J+h)Y\right],$$

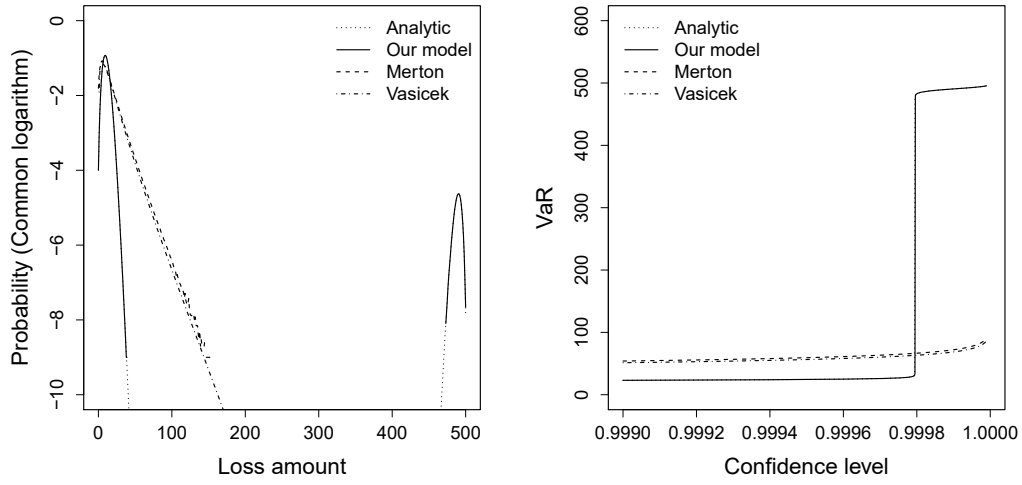
respectively, where J and h are calculated based on $\beta = 1$.

2.3.2 Analysis of loss distributions and VaRs

We show the numerical calculation results of each case described in the previous subsection. We compare our model with the standard model in terms of the loss distribution and the VaR. We adopt a one-factor Merton-type model as the standard model. We also analyze the variation of the loss distribution and the VaR from homogeneous credit portfolios to credit portfolios with heterogeneous credit exposures, especially, the variation of the second peak (shape of the tail of the loss distribution).

Figure 2.3 shows the results of our numerical calculation of the loss distribution and the VaR for the homogeneous credit portfolio in Case 1. Our model is compared with the standard model. Since this case consists of a homogeneous credit portfolio, we also show the analytical formula of the loss distribution in Molins and Vives [54] or Kitsukawa et al. [48] and the limiting loss distribution in Vasicek [72]. VaR means the maximum value of the loss amount that can occur with a certain probability (confidence level) in the future, and the probability that the loss amount will become more than VaR is equal to the difference between one and the confidence level. The left peak in the loss distribution is called the first peak, and the part, which is observed in the standard model, has a high probability. The right peak is called the second peak, and the part, which is not observed in the standard model, emerges in the large loss region and is characteristic of the Ising model. Our model can measure the second peak of the loss distribution that cannot be measured using the standard model and that has been measured in existing studies only by using the analytical formula of the loss distribution. In our model, the VaR increases drastically as the confidence level rises because of the effect of the second peak as shown in Molins and Vives [54]. Additionally, we can confirm from the results of our numerical calculation that the loss distribution calculated by our model converges sufficiently, and it reproduces the loss distribution calculated by the analytical formula in existing studies sufficiently.

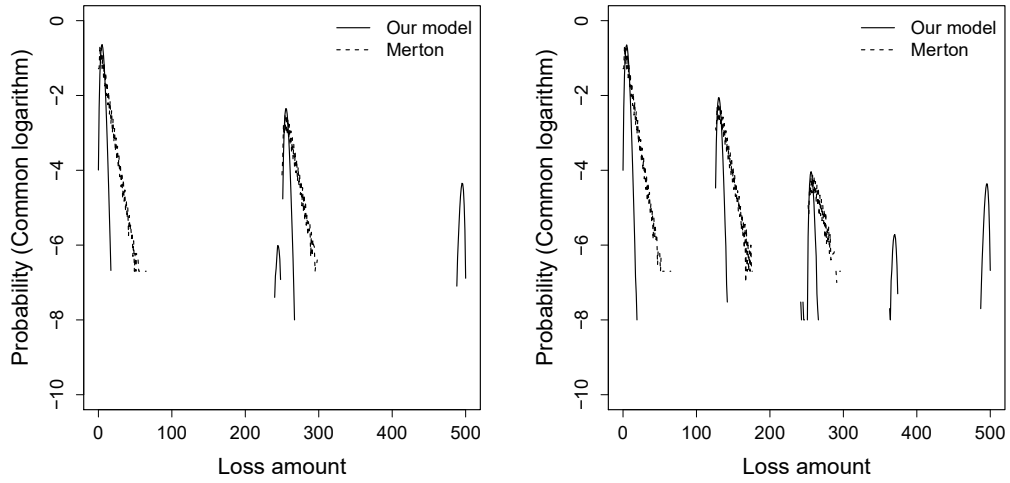
Figure 2.4 shows the results of our numerical calculation of the loss distribution for the credit portfolio in which the credit exposure of one borrower is large and that of the other borrowers



(a) Loss distribution (b) VaR

Figure 2.3: Loss distribution and VaR for the homogeneous credit portfolio. The dotted line shows the loss distribution from the analytical formula and the solid line shows the loss distribution from our model. The dashed line and the dot-dash line show the loss distribution from the one-factor Merton-type model and the limiting loss distribution, respectively.

is small in Case 2 (left graph) and the credit portfolio in which the credit exposures of two borrowers are large and that of the other borrowers is small in Case 3 (right graph). Our model is compared with the standard model. In the credit portfolio where only the credit exposure of one borrower is large, four peaks are measured by our model (solid line). The second peak in the homogeneous credit portfolio splits into the smallest peak to the left of 250 losses and the peak around 500 losses; these peaks correspond to default chains. The smallest peak to the left of 250 losses represents the case where the borrower with a large credit exposure has not defaulted while the peak around 500 losses represents the case where the borrower with a large credit exposure has defaulted. There are no corresponding peaks in the standard model. On the other hand, the first peak in the homogeneous credit portfolio splits into the peak around 0 loss and the peak to the right of 250 losses; these peaks do not correspond to default chains. The peak around 0 loss represents the case where the borrower with a large credit exposure has not defaulted while the peak to the right of 250 losses represents the case where the borrower with a large credit exposure has defaulted. There do exist corresponding peaks in the standard model (dashed line). In the credit portfolio where the credit exposures of two borrowers are large, six peaks are measured by our model (solid line). The second peak in the homogeneous credit portfolio splits into the smallest peak to the left of 250 losses, the peak around 375 losses, and the peak around 500 losses; these peaks correspond to default chains. The smallest peak to the left of 250 losses represents the case where both borrowers with large credit exposures have not defaulted, the peak around 375 losses represents the case where one of them has defaulted, and the peak around 500 losses represents the case where both of them have defaulted. There are no corresponding peaks in the standard model. On the other hand, the first peak in the homogeneous credit portfolio splits into the peak around 0 loss, the peak around 125 losses, and the peak to the right of 250 losses; these peaks do not correspond to default chains. The peak around 0 loss represents the case where both borrowers with large credit exposures have not defaulted, the peak around 125 losses represents the case where one of them has defaulted, and the peak to the right of 250 losses represents the case where both of them have defaulted. There do exist corresponding peaks in the standard model (dashed line). In these cases, the peaks mean that default chains emerge in our model but do not emerge in the standard model, similar to Case 1. As the credit exposure of borrowers is specified more precisely as being distributed



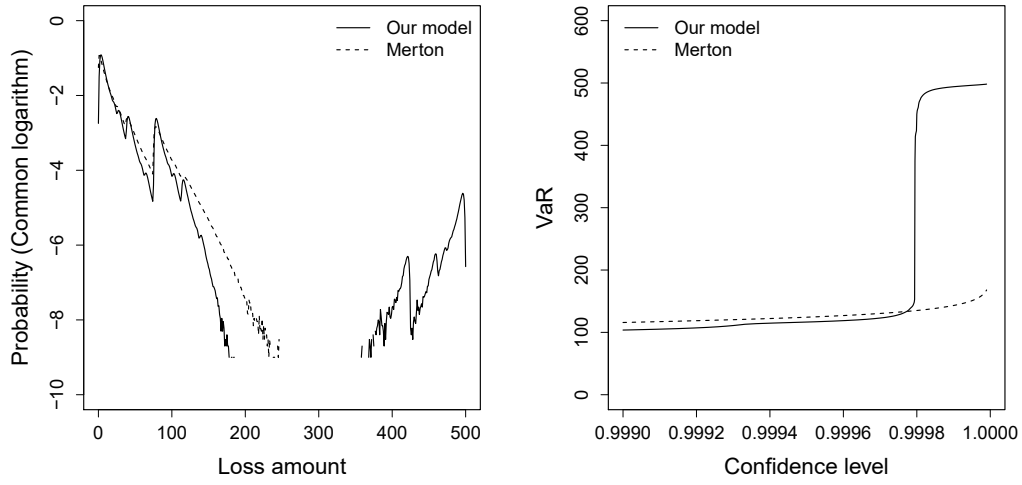
(a) Loss distribution in Case 2 (b) Loss distribution in Case 3

Figure 2.4: Loss distribution for the credit portfolio in which the credit exposure of one borrower is large and that of the other borrowers is small (left graph) and in which the credit exposures of two borrowers are large and that of the other borrowers is small (right graph). In both graphs, the solid line shows the loss distribution from our model, and the dashed line shows the loss distribution from the one-factor Merton-type model.

according to a power-law distribution, it is believed that the loss distributions converge to the loss distribution for the credit portfolio in which the credit exposures are distributed according to a power-law distribution.

Figure 2.5 shows the results of our numerical calculation of the loss distribution and the VaR for the credit portfolio in which the credit exposures are distributed according to a power-law distribution in Case 4. Our model is compared with the standard model. If the credit exposures are distributed according to a power-law distribution, the shape of the loss distribution reflects the construction of credit exposures. Both our model and the standard model have small peaks in the position corresponding to the default of borrowers that have large credit exposure. The left peak means no default chain while the right peak means a default chain. In the case of a power-law distribution, both peaks emerge outside in comparison to the homogeneous credit portfolio. Both peaks also decrease relatively slowly towards the center and a certain probability remains. Therefore, the VaR is larger for the low confidence level because of the effect of the fat tail in the loss distribution. In our model, the second peak is observed, and the VaR increases drastically as the confidence level rises because of the effect of the second peak, similar to Case 1.

Figure 2.6 shows the results of our numerical calculation of the loss distribution and the VaR for the credit portfolio consisting of two independent homogeneous sectors in Case 5. We divide the credit portfolio of 500 borrowers into two sectors of 100 and 400 borrowers. Here, we compare two setups of the credit portfolio with each other. In Setup 1, the default correlation among borrowers across the two sectors is equal to the default correlation among borrowers in the sector, that is, Setup 1 is the homogeneous credit portfolio of 500 borrowers as a whole and is equivalent to Case 1. In Setup 2, the default correlation among borrowers across the two sectors is zero, that is, Setup 2 is the two independent homogeneous credit portfolios of 100 and 400 borrowers and represents Case 5. Setup 1 and two sectors of Setup 2 are homogeneous credit portfolios where the default probability is 0.02 and the default correlation is 0.01. Figure 2.6a shows the positions of the second peaks of the loss distribution based on the difference of the default correlation among borrowers across the two sectors. It is observed that there is one peak that corresponds to a large-scale chain of defaults around 500 losses if Setup 1 is

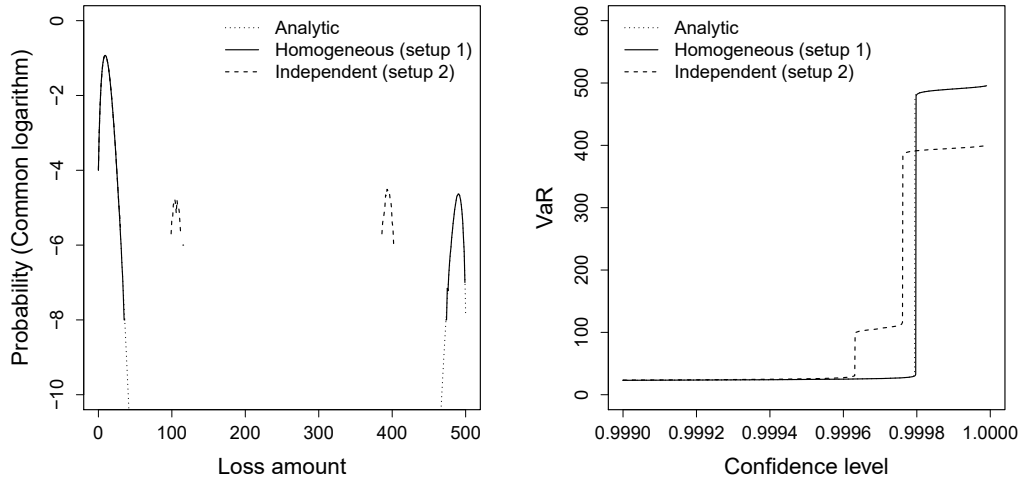


(a) Loss distribution (b) VaR

Figure 2.5: Loss distribution and VaR for the credit portfolio in which the credit exposures are distributed according to a power-law distribution. In both graphs, the solid line shows the loss distribution from our model, and the dashed line shows the loss distribution from the one-factor Merton-type model.

the credit portfolio while there are two peaks that correspond to a large-scale chain of defaults around 100 and 400 losses if Setup 2 is the credit portfolio. When the default correlation among borrowers across the two sectors decreases from Setup 1 ($\rho^{1,2} = 0.01$) to Setup 2 ($\rho^{1,2} = 0$), the second peak transits from one peak around 500 losses into two peaks around 100 and 400 losses. Although we have not yet solved how path the second peak trace, the intermediate default correlation cases ($0 < \rho^{1,2} < 0.01$) exist on the track connecting both extremes (Setup 1 and Setup 2). Here, the peak around 0 loss in the loss distribution means that default chains do not occur in both sectors of 100 and 400 borrowers, and the peak around 500 losses means that default chains occur in both sectors. The peak around 100 losses also means that default chains occur in the sector of 100 borrowers but not in the sector of 400 borrowers. Conversely, the peak around 400 losses means that default chains do not occur in the sector of 100 borrowers but in the sector of 400 borrowers. Moreover, the peak around 100 losses is observed at a slightly larger loss amount than 100 losses. This is likely to be the case because the peak is a compound of the second peak of the sector of 100 borrowers (96 to 100 losses) and the first peak of the sector of 400 borrowers (6 to 9 losses). In the same way, the peak around 400 losses is observed at a slightly smaller loss amount than 400 losses. This is probably because the peak is a compound of the first peak of the sector of 100 borrowers (0 to 4 losses) and the second peak of the sector of 400 borrowers (390 to 394 losses). Additionally, the behavior of the VaR changes significantly because the tail of the loss distribution exhibits a drastic change. Figure 2.6b shows that the magnitude relationship of the VaR (risk measure) is switched in the homogeneous credit portfolio of Setup 1 and the two independent homogeneous credit portfolios of Setup 2 as the confidence level varies. When the default correlation among borrowers across the two sectors decreases from Setup 1 ($\rho^{1,2} = 0.01$) to Setup 2 ($\rho^{1,2} = 0$), the VaR transits from Setup 1 into Setup 2. The intermediate default correlation cases ($0 < \rho^{1,2} < 0.01$) exist on the track connecting both extremes (Setup 1 and Setup 2) for the same reason as the loss distribution.

From the above results, it is clear that the tail of the loss distribution calculated by our model has characteristics that are different from that of the standard models used in credit risk modeling, and there is a possibility of different evaluations of credit risk according to the pattern of the heterogeneity of the credit exposure. The same might be said of the default



(a) Loss distribution (b) VaR

Figure 2.6: Loss distribution and VaR for the credit portfolio divided into two sectors. The dotted line shows the loss distribution from the analytical formula. The other lines show the loss distribution from our model. The solid line shows the loss distribution for the homogeneous credit portfolio, and the dashed line shows the loss distribution for the two independent homogeneous credit portfolios.

correlation in the future.

2.4 Conclusion

The contribution of this study is the construction of the long range Ising model for credit risk modeling in heterogeneous credit portfolios and the actual measurement of the credit risk for credit portfolios with heterogeneous credit exposures. The credit risk model using the long-range Ising model can measure a second peak that cannot be measured by the standard model, and the model for heterogeneous credit portfolios has been proposed in the existing studies. However, in practice, credit portfolios are set to the default probability and the default correlation based on credit rating and industry, and the credit exposure is heterogeneous. Therefore, we take a step towards modeling business practices by considering the heterogeneity of credit portfolios from a practical perspective in our model. The introduction of heterogeneity to credit portfolios consists of two parts when it is distributed greatly. First, we introduce heterogeneity to the default probability and the default correlation by dividing a credit portfolio into multiple sectors based on credit rating and industry. Next, we introduce heterogeneity to the credit exposure by using an approach based on a numerical calculation to calculate the loss distribution, but not an analytical method. In the numerical calculation, we calculate the loss distribution efficiently using the RMC method. Although only the loss distribution of the credit portfolios that consist of borrowers whose credit exposure is homogeneous, that is, the distribution of the number of defaulting borrowers, has been calculated in the existing studies, we calculate the loss distribution of the credit portfolios that consist of borrowers whose credit exposure is heterogeneous owing to our model.

To analyze the characteristics of the loss distribution of credit portfolios with heterogeneous credit exposures, we apply our model to various credit portfolios and evaluate the credit risk. Consequently, we show that the tail of the loss distribution calculated by our model has characteristics different from that of the standard models used in credit risk modeling, and there is a possibility of different evaluations of credit risk according to the pattern of heterogeneity. From the result of our numerical calculation in the homogeneous credit portfolio, it

is confirmed that our model can measure a second peak of loss distributions that cannot be measured by the standard model as well as the existing studies, and that the VaR increases drastically due to the effect of the second peak as the confidence level increases. From the result of our numerical calculation in the credit portfolios with heterogeneous credit exposure, we find that a loss distribution has a shape that reflects the structure of the credit exposure. In the case of one borrower with a large credit exposure and the case of two borrowers with large credit exposures, the second peak observed in the case of the homogeneous credit portfolio splits into multiple peaks based on whether the defaulting borrower has large credit exposure or not. When the structure of these credit exposures is more complicated, we have the case where the credit exposures are distributed according to a power-law distribution. As the credit exposure is set more precisely to be distributed according to a power-law distribution, the loss distribution will approach the case of a power-law distribution. Also, in credit portfolios with heterogeneous credit exposures, the peaks corresponding to default chains, which emerge in the long range Ising model, do not emerge in the standard model as well as homogeneous credit portfolios. From the comparison between the credit portfolio that is homogeneous as a whole and the credit portfolio consisting of two homogeneous sectors that are independent of each other, we find that both the loss distribution and the VaR show different behavior according to the default correlation among borrowers across the two sectors. The numerical solutions of these portfolios become the starting point and the motivation for understanding the credit portfolios that have the intermediate default correlation among borrowers across the two sectors in the future. In the above cases, a similar phenomenon of the VaR can be observed for the CVaR. From the above-mentioned results, it is believed that by changing the credit portfolio structure, that is, the distribution condition of the default probability, the default correlation, and the credit exposure (the exchange interaction, the external field, and the credit exposure) in the portfolio, there is a possibility of expressing default chains less than medium scale.

There are many potential problems. First, we expand the long-range Ising model for credit portfolios with heterogeneous credit exposures developed in this study to the really heterogeneous credit portfolios where the default probability and the default correlation are heterogeneous. To do that, it is the most important issue to develop the calibration method of parameters for determining the exchange interactions and the external fields, which are input parameters for the Ising model, from the default probabilities and the default correlations that are given for heterogeneous credit portfolios described using the sector model. This method is the algorithm for solving multidimensional nonlinear system of equations (inverse problem). To measure the credit risk of more complicated credit portfolios, especially, practical credit portfolios, calibration of parameters is required. Our numerical calculation method is appropriate for solving credit portfolios with heterogeneous credit exposures, and it might be appropriate in the future for solving the really heterogeneous credit portfolios where the default probability and the default correlation are heterogeneous when a calibration method is developed. When solving multidimensional nonlinear system of equations, the numerical calculation is necessary. In this method, we will consider Newton's method and the Bayesian inference combined with the simulated annealing. Because it is difficult to calculate a thermal average, which appears as the terms of the system of equations, as the dimensionality increases, we will try to improve the computational efficiency by combining the RMC method and the HMC method using Stan. We will also try to utilize the approximation formula using a perturbation calculation in Kitsukawa et al. [48], the maximum likelihood estimation in Filiz et al. [25], and the iterative scaling method used for the parameter estimation of the maximum entropy model.

Second, as a conclusion of this study, in the case of the long-range Ising model for homogeneous credit portfolios, the two peaks of the first peak, which is observed in the standard model and has a high probability, and the second peak, which is not observed in the standard model and is characteristic of the Ising model, in the loss distribution are observed the same as the

previous studies. In the case of credit portfolios with heterogeneous credit exposures, the loss distribution has the shape with the structure of credit exposures added to the loss distribution of homogeneous credit portfolios. Due to the second peak, the VaR increases rapidly as the confidence level is increased. The second peak has a big impact on the measurement of credit risk. Therefore, we will quantitatively elucidate the measurement deviation (tail phenomenon) in the standard model by clarifying theoretically the problems with the assumptions of the standard model and comparing them with the Ising model. We might need a new risk index different from the VaR and CVaR because these indices could not accurately measure credit risks. Furthermore, we will try to improve the computational efficiency by performing parallel computation using the general-purpose computing on graphics processing units (GPGPU) for the MCMC and an independent iterative computation in order that the loss distributions converge.

Third, it is conceivable that the credit risk model for heterogeneous credit portfolios using the Ising model of ferromagnet constructed in this study is expanded to the Ising model of spin glass. Specifically, the credit risk model is developed using the spin glass model (model of magnetic materials that the interactions between spins are randomly distributed pairwise and that the directions of spins are randomly frozen), the random magnetic field Ising model (model to which a magnetic field (external field) is applied in a random direction per spin while the interaction between spins is uniform), and the Potts model (model that generalizes the Ising model to the spin variables of three states or more) to set the default probability and default correlation more precisely because this model becomes the elaboration of the setting for the default probability and the default correlation in the Ising model of ferromagnet. We construct the credit risk model that takes the spatial randomness into account for the default probability and default correlation (exchange interaction and external field) using the spin glass model and the random magnetic field Ising model. We expand two states of borrowers, defaulted or non-defaulted, at the risk horizon into multiple states according to credit ratings by applying the spin variables into more than three states using the Potts model. Even if we refer to default events in the past, a large-scale default chain, indicating default by most of the borrowers in the credit portfolio, never occurs. However, there is a possibility of expressing a large-scale chain of deterioration of credit conditions as a phenomenon that can occur. This can be done by using more than three spin states and expanding the model to account for the various credit conditions of borrowers.

Fourth, the nature of the tail of loss distributions and a default cooperative phenomenon should be analyzed by applying this model to various credit portfolios. Specifically, we want to understand the underlying mechanism and the meaning of a second peak (shape of the tail of loss distributions) by analyzing the variation of the shape of loss distributions or the risk measures (e.g., VaR, CVaR) according to the variation in the parameters (e.g., default probability, default correlation, and confidence level). We clarify the mechanism of phase transition in credit portfolios with a clue of the mechanism of phase transition in the Ising model from the phase diagram of the default probability and default correlation (exchange interaction and external field), Nishimori's line in the phase diagram, and the rapid change (phase transition phenomenon) in the order parameters (risk indices of VaR and CVaR) with respect to change in the parameters (e.g., default probability, default correlation, and confidence level). Furthermore, we elucidate the mechanism of the tail of loss distributions, a default cooperative phenomenon, and the universal nature as large scale, many body, and complex systems. There is a possibility to describe a default cooperative phenomenon (collective behavior of borrowers) for which a sufficient explanation is not always able to be provided in the standard model as a large chain of defaults by using the theory of phase transition in the Ising model. We determine whether a second peak can explain a default cooperative phenomenon or not, as it is believed that a second peak indicates a large-scale default chain based on the collective

behavior of borrowers. Additionally, we provide the economic and financial meanings and interpretations to the objects of energy (Hamiltonian), temperature, exchange interaction, external field, magnetization, specific heat, entropy, enthalpy, Helmholtz free energy, Gibbs free energy, and the other macroscopic quantity in thermodynamics, which are a distribution, function, variable, and parameter. We also consider the economic and financial meanings and interpretations of the theorems and concepts in statistical mechanics and thermodynamics. The theory of statistical mechanics, which deals with unobservable objects, is built to match observable quantities in thermodynamics. The observable quantities in thermodynamics are clues to show correctness of the theory in statistical mechanics. Therefore, the theory of covariation and contagion of credit risk, which deals with unobservable (and observable) objects, should be also built to match observable quantities in markets. This model has been studied from various angles (e.g., statistical mechanics, graph theory, and information theory). However, there is no general framework related to credit risk modeling. Therefore, we will generalize and unify the framework of the model by organizing the previous studies and this study because these models can be uniformly described using the Ising model. An approach using network theory and information geometry should be also considered.

Fifth, it is standard practice to measure the credit risk of practical credit portfolios in financial institutions. Although the model and algorithm constructed and implemented in this study can manage more complicated and practical credit portfolios, we want to measure credit risk on the basis of conditions faced by an actual business practice. To do that, we divide a credit portfolio into multiple segments, such as AAA-C credit ratings, and set each segment to a default probability. We also divide the credit portfolio into multiple segments, such as stock price indices by industry, and set each segment and each segment pair to a default correlation. Furthermore, heterogeneous credit exposures are taken into account such as a power-law distribution. Thereafter, it can be thought of as an approach to set sectors by the combination of them. We also need to manage credit portfolios that consist of a larger number of borrowers in accordance with the number of borrowers in practical credit portfolios. If calculation time satisfies practical requirements, we could suggest the credit risk model applied to the practical credit portfolios in financial institutions. We will develop the program to numerically calculate loss distributions using this model and measure the credit risk of various credit portfolios because they can be realized if the development of a calibration method of parameters and the additional reduction of computation time based on the acceleration and increased efficiency of calculation are conducted.

Chapter 3

Cointegration Analysis of Hazard Rates and CDSs: Applications to Pairs Trading Strategy

3.1 Introduction

The CDS spreads of companies in the same industry and group companies tend to behave similarly. This is a phenomenon frequently seen in real CDS markets. This is partly because these companies share common credit risk factors. If those time series are recognized as I(1)-processes, then they are referred to as cointegration, the concept developed by Engle and Granger [23]. Its background and some applications are described in detail, for example, by Hamilton [31]. The cointegration of stocks, commodities, and so on is being actively studied and used in practical investment strategies.

The first aim of our research is to gain a deeper understanding of the characteristic behavior of cointegrated CDS spreads with different tenors according to the reduced-form approach (Duffie and Singleton [21]) using minimal modeling of cointegrated hazard rates. Merging the cointegration structure into the framework of arbitrage-free pricing and deriving the theoretical spread of CDSs, we explain why cointegration appears among CDSs with different tenors by assuming the structure of the ECM of instantaneous hazard rates.

The cointegration test for CDS spreads is quite similar to the cointegration test for bond yields. For a set of CDS spreads (bond yields) with ad hoc tenors, an investor searches for cointegration pairs. Next, they examine whether such a pair stays in the cointegration state or not when taking the other combinations of tenors. Thus, depending on the number of available tenors and the size of a search set of CDS spreads (bond yields), there are many combinatorial numbers of cointegration tests, unlike in the stock case. Another observation is that the strength of cointegration between the CDS spreads (bond yields) with a far expiry date tends to become weak.

These issues and phenomena have not been studied enough. In the previous analyses of cointegration, the viewpoint has been disregarded such that bonds and CDS spreads, unlike stocks, can be priced according to the rule of no-arbitrage pricing. We take this viewpoint into account and aim to elucidate these issues. This critical difference stems from a difference in formulation, with bond prices and CDS spreads following the backward stochastic differential equation (BSDE) [60] with several terminal conditions, while stock prices follow the forward stochastic differential equation (FSDE) with some initial conditions.

Although several studies analyze the cointegration between a bond and its corresponding CDSs in the existing literature, there are no studies analyzing the cointegration between CDSs as well as treating the ECM of hazard rates. For example, Blanco et al. [7] and Guidolin et

al. [29] analyze the arbitrage relation between the CDS and bond markets by investigating the cointegration between a credit spread (yield to maturity of a bond minus risk-free rate) and its corresponding CDS spread. One of the reasons why it is difficult to take CDSs as the target of research is that corporate CDSs targeted in this study are less liquid than bonds. Since CDSs have a shorter history than bonds, the fact that there are not enough studies for them is also a factor. Furthermore, as there are multiple CDSs with different maturity for the same reference company unlike stocks and commodities, there are many "ad hoc combinations of CDSs" for the cointegration test conducted to search cointegrated CDSs. This leads to the motivation for research that we give a law to briefly explain the diversity of the cointegration between CDSs with different maturity based on hazard rates under an arbitrage-free pricing framework.

In the field of commodities, some studies discuss cointegration in the arbitrage-free framework. Benmenzer et al. [5] model the market price of risk to connect the physical probability and the risk-neutral probability for the cointegration between the future prices of natural gas and crude oil. Nakajima and Ohashi [56] derive the price formula for the future and the call option to incorporate the cointegration between commodity spot prices and show that cointegration can affect the price of commodity derivatives. These models are constructed consistently with the arbitrage-free theory while considering the cointegration of commodity prices. Our study focuses on the cointegration of hazard rates instead of commodities.

The second aim of our research is to explore how the pairs trading strategy of cointegrated CDSs is constructed based on our VECM of cointegrated hazard rates. As is well known, pairs trading is an investment strategy that uses the mean reversion of the price differentials of two cointegrated assets (Refer to [69] and [73] for details). Most papers on pairs trading have concentrated on studying the patterns of stocks that are not priced (i.e., follow the FSDEs), as stated earlier. A novel point of our research is that we attempt to construct the pairs trading strategy for CDSs that can be priced (i.e., follow the BSDEs) in the arbitrage-free pricing framework. In other words, we aim to establish a trading methodology based on cointegration that is consistent with an arbitrage-free pricing framework.

Higashide et al. [33] define the time required from the position building of the pair of two cointegrated stocks to the position closing by reversing trade (that is, the time it takes to leave the level at which the spread of stock prices has widened sufficiently and reach the level at which it has narrowed sufficiently for the first time) as the first passage time (random variable) for a pairs trading strategy. Then, they define the difference of spread in this period as a profit. They consider a portfolio optimization problem of stock pairs to reduce the expectation and variance of the weighted sum of first passage time by fund allocation to each pair. The shorter the first passage time is, the more profit opportunities ones get because of the increase of the opportunity to reconstruct the pair. Tourin and Yan [70] suggest a pairs trading strategy for stocks by solving the Hamilton-Jacobi-Bellman (HJB) equation in the framework of the ECM. Our study focuses on the ECM of hazard rates instead of stocks. Bali et al. [3] study the market-neutral strategy of interest rate swap contracts, in which they utilize the mean reversion of mispricing. Jarrow et al. [39] apply the same approach as Bali et al. [3] to the term structure of CDSs. These two papers address a trading strategy that is based on the resolution of mispricing and is subject to model risk. In contrast, our study utilizes the property of the mean reversion of cointegrated CDS spreads.

We finally emphasize that the ODE-based Bayesian inference is necessary to estimate the cointegration dynamics for hazard rates because the theoretical spread of cointegrated CDSs depending on such latent hazard rates does not have an analytical formula. We explore the hazard rate processes that can be fitted to entire term structures of cointegrated CDS spreads. We recognize the connection between the cointegration of hazard rates, which is the main problem considered in this study, and concepts or methods of statistical mechanics from the point of view of econophysics. Our ECM of hazard rates is thought of as a financial analog of

Table 3.1: Companies in three industry sectors in the target of cointegration test to CDS spreads. These industry sectors include electricity (7 companies), construction (5 companies), and automotive (7 companies). We apply a cointegration test to these companies in the same industry.

Industry	Company
Electricity	Panasonic Corp., Sony Corp., Fujitsu Ltd., Hitachi Ltd., NEC Corp., Toshiba Corp., Mitsubishi Electric Corp.
Construction	Kajima Corp., Obayashi Corp., Taisei Corp., Shimizu Corp., Nishimatsu Construction Co., Ltd.
Automotive	Toyota Motor Corp., Nissan Motor Co., Ltd., Honda Motor Co., Ltd., Mazda Motor Corp., Subaru Corp., Suzuki Motor Corp., Mitsubishi Motors Corp.

constrained Langevin diffusion in non-equilibrium statistical mechanics.

The remainder of Chapter 3 is organized as follows. Section 3.2 analyzes cointegrated corporate CDSs in the Japanese CDS market. Section 3.3 derives the general pricing formula of a cointegrated CDS based on the VECM of multivariate hazard rates and elaborates on the ODE-based Bayesian inference method. Section 3.4 provides the estimation results of our empirical study on Japanese cointegrated corporate CDSs and explains how to construct the pairs trading strategy for CDSs. Section 3.5 concludes our studies and mentions some concluding remarks. Appendix B addresses the derivation of the ODEs involving the general formula of cointegrated CDSs.

3.2 Cointegration analysis for CDSs

In this section, we analyze the cointegration of CDSs that will motivate us to develop the hazard rate based model in Section 3.3 and its empirical analysis in Section 3.4. Specifically, we confirm the existence of several cointegrated CDS pairs in the market and extract a set of cointegrated pairs that will be analyzed closely in Section 3.4.

3.2.1 Identification of cointegrated pairs

A practical method to identify cointegrated pairs is to search for pairs within the same industry. This is simply because cointegrated pairs belonging to different industry sectors may have unimmunized sector-specific risks, whereas CDS spreads belonging to the same industry sector tend to make the same move. Thus, we restrict our search for cointegrated pairs to only CDS spreads from within the same industry sector. For later empirical studies, we focus on three industry sectors: electricity, construction, and automotive. The electricity sector consists of seven companies; the construction sector consists of five companies; and the automotive sector consists of seven companies. The CDS spreads that belong to each industry are shown in Table 3.1. We first extract typical pairs among a set of several cointegrated pairs identified by a cointegration test. Next, we investigate whether the cointegration relationship persists when changing the tenor of CDS in such cointegrated pairs. We take five CDS tenors of one, three, five, seven, and ten years, and the observation period of daily data is 251 business days from November 20, 2020, to November 30, 2021, obtained from Thomson Reuters. To investigate the cointegration of the pair of CDS spreads, we apply the Phillips-Ouliaris test [61]. Then, to test the existence of the unit root, we use the Augmented Dickey-Fuller (ADF) test.

First, to investigate whether the CDS spread of each company is a unit root process, we conduct the ADF test on the five-year CDS spreads of each company. As a result, the null hypothesis of the unit root process is rejected regarding Taisei and Nishimatsu in the construction

Table 3.2: Results of the Phillips-Ouliaris test on the five-year CDS spreads in the same industry. The test statistics are shown with * if the null hypothesis of no cointegration is rejected at the significance level of 5% or ** if it is rejected at the significance level of 1%. The rows represent the explained variable in the regression analysis of the Phillips-Ouliaris test, and the columns represent the explanatory variable. We describe both because swapping the explained variable and the explanatory variable change the test results.

Electricity

	Panasonic	Sony	Fujitsu	Hitachi	NEC	Toshiba	Mitsubishi
Panasonic		7.64	9.78	10.00	7.24	14.21	6.99
Sony	11.06		36.33*	16.43	25.06	10.43	10.60
Fujitsu	17.36	48.08**		10.02	15.21	14.63	12.35
Hitachi	67.46**	73.48**	36.81*		60.03**	69.90**	49.05**
NEC	19.38	56.64**	26.37	25.87		22.78	21.43
Toshiba	52.21**	23.01	29.52	33.42	27.10		24.97
Mitsubishi	45.79*	44.59*	45.43*	48.84**	45.48*	48.69**	

Construction

	Kajima	Obayashi	Shimizu
Kajima		15.99	29.40
Obayashi	4.22		6.59
Shimizu	67.06**	54.97**	

Automotive

	Nissan	Honda	Mazda	Subaru	Suzuki	Mitsubishi
Nissan		4.14	3.54	5.63	5.83	3.23
Honda	38.03*		22.32	44.92*	39.87*	44.30*
Mazda	22.09	16.38		16.91	15.22	17.03
Subaru	32.95	28.69	15.24		48.20**	14.42
Suzuki	48.27**	31.96	17.16	57.53**		28.82
Mitsubishi	21.85	36.65*	18.28	18.38	30.26	

industry sector and Toyota in the automotive industry sector at the significance level of 5%. Accordingly, we exclude these companies from the cointegration test's target. Next, to investigate whether each pair of CDS spreads in the same industry meets the cointegration conditions, we conduct the Phillips-Ouliaris test on the five-year CDS spreads. As shown in Table 3.2, there are several pairs of CDS spreads in which the null hypothesis of no cointegration is rejected by the Phillips-Ouliaris test. For the following three reasons, we find a remarkable pair of Sony and Fujitsu in the electricity industry sector: First, even after swapping the explained variable and the explanatory variable in the Phillips-Ouliaris test, the null hypothesis of no cointegration is rejected for the pair of Sony and Fujitsu in the electricity industry sector and the pair of Subaru and Suzuki in the automotive industry sector. There are many pairs in which the null hypothesis of no cointegration are not rejected. Second, as shown in Table 3.3, many CDS spread pairs between the tenors of Sony and Fujitsu are contegrated by the Phillips-Ouliaris test. Many pairs between the tenors of Subaru and Suzuki have no contegration. Third, the CDS spreads of Sony and Fujitsu are stable. That is, they do not include the period during which CDS spreads are extremely high, such as for Toshiba, Nissan, and Mitsubishi Motors. More importantly, no matter which pair we choose between any CDS spread with a tenor of one, three, or five years, it has a cointegration property to some degree, as will be examined more closely in Section 3.4. Consequently, we select Sony and Fujitsu as the target of our analysis to deduce the common features of cointegration in CDS spreads.

Table 3.3: Results of the Phillips-Ouliaris test on all terms CDS spreads in the pair of Sony and Fujitsu. The test statistics are shown with * if the null hypothesis of no cointegration is rejected at the significance level of 5% or ** if it is rejected at the significance level of 1%. The rows represent the explained variable in the regression analysis of the Phillips-Ouliaris test, and the columns represent the explanatory variable. We describe both because swapping the explained variable and the explanatory variable change the test results.

	Fujitsu 1Y	Fujitsu 3Y	Fujitsu 5Y	Fujitsu 7Y	Fujitsu 10Y
Sony 1Y	28.85	29.21	37.81*	35.46*	36.15*
Sony 3Y	26.71	36.19*	45.85*	45.76*	42.72*
Sony 5Y	21.57	43.34*	36.33*	58.27**	53.29**
Sony 7Y	16.84	27.52	24.61	34.42*	33.61
Sony 10Y	15.60	25.82	23.09	33.07	33.10

	Sony 1Y	Sony 3Y	Sony 5Y	Sony 7Y	Sony 10Y
Fujitsu 1Y	59.95**	95.56**	120.61**	103.88**	103.95**
Fujitsu 3Y	18.02	34.21*	63.94**	49.38**	50.48**
Fujitsu 5Y	15.19	33.56	48.08**	38.40*	39.57*
Fujitsu 7Y	6.02	11.48	22.43	17.63	18.45
Fujitsu 10Y	7.30	12.80	24.09	19.89	21.26

3.2.2 Analysis of CDS spreads

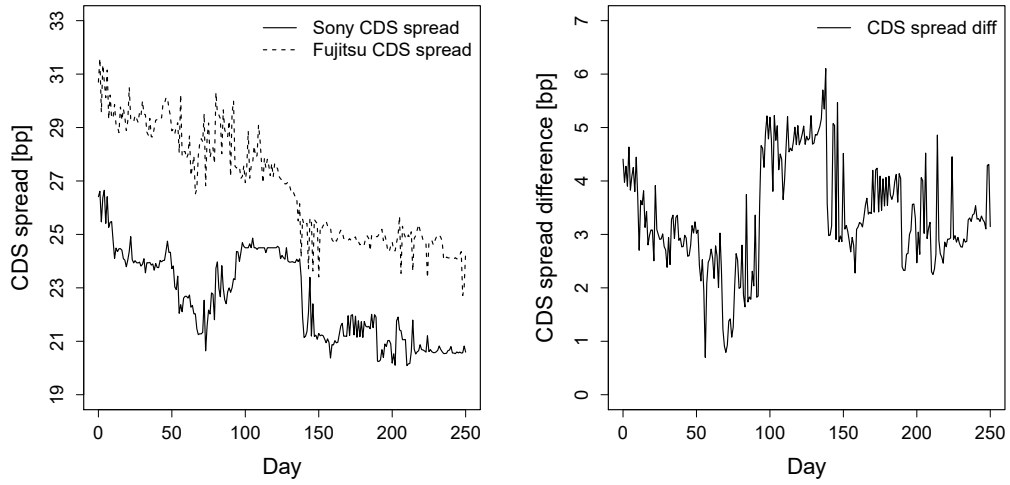
In the first half of this section, we investigate the nature of the CDS spread pair of Sony and Fujitsu selected for the empirical analysis in Section 3.4. In Figure 3.1, we show the five-year market CDS spreads of Sony and Fujitsu, and the difference between both CDS spreads weighted with the cointegrated vector $(1, -0.7167)$, which is obtained by estimating the VECM of both CDS spreads. The left figure shows the CDS spreads of Sony and Fujitsu. Both CDS spreads move in a similar manner at different levels. The right figure shows the difference between both CDS spreads weighted with the cointegrated vector $(1, -0.7167)$, which implies that Sony CDS spreads - 0.7167 $\times$ Fujitsu CDS spreads are a stationary process. Although the difference can sometimes grow, it seems to be moving at a constant value on average. Accordingly, the process seems to be a mean reversion. This analysis result will be compared with that in Section 3.4.

In the second half of this section, we analyze the cointegration of three CDSs. We will analyze the cointegration of two CDSs with the pair of Sony and Fujitsu in Section 3.4, which is the case of one cointegration vector. However, the model suggested in Section 3.3.1 can analyze the cointegration of three CDSs or more because it will be a general multidimensional model. That is, it can treat the case of two cointegration vectors or more. Therefore, we investigate the cointegration of three CDSs to support the usefulness of our general multidimensional model provided in Section 3.3.1.

Since cointegration exists between two CDS spreads in the same industry, we apply the Johansen cointegration test [42] at the significance level of 5% to three CDS spreads in the same industry and investigate the cointegration rank (number of cointegration vectors) in the set of three CDS spreads. However, we only investigate the cointegration rank of three CDS spreads in the same tenor as there are too many combinations of CDS spreads. In Table 3.4, we show the cointegration rank observed from the sets of three CDS spreads in the same tenor in the electricity sector. There are many sets of three CDS spreads in which the cointegration rank is one. It should be noted that there are enough sets in which the cointegration rank is two. We also observe a similar tendency in the automotive sector. Accordingly, the case of having two cointegration vectors (or more) can be analyzed if the problem of having a higher computational load than two CDSs is solved.

Table 3.4: Cointegration rank (number of cointegration vectors) observed from the sets of three CDS spreads in the same tenor in the electricity sector. We conduct the Johansen cointegration test at the significance level of 5%. The rows represent the cointegration rank in each set of three CDS spreads, and the columns represent the tenor.

Set of three CDS spreads			Tenor				
CDS 1	CDS 2	CDS 3	1Y	3Y	5Y	7Y	10Y
Panasonic	Sony	Fujitsu	1	1	1	0	0
Panasonic	Sony	Hitachi	0	1	1	1	1
Panasonic	Sony	NEC	0	0	0	0	0
Panasonic	Sony	Toshiba	0	0	1	1	1
Panasonic	Sony	Mitsubishi	0	0	0	0	1
Panasonic	Fujitsu	Hitachi	1	2	2	1	1
Panasonic	Fujitsu	NEC	1	1	1	0	1
Panasonic	Fujitsu	Toshiba	1	1	1	0	1
Panasonic	Fujitsu	Mitsubishi	2	1	1	0	1
Panasonic	Hitachi	NEC	0	1	1	0	0
Panasonic	Hitachi	Toshiba	0	1	1	1	1
Panasonic	Hitachi	Mitsubishi	0	1	1	1	1
Panasonic	NEC	Toshiba	0	0	1	0	1
Panasonic	NEC	Mitsubishi	1	0	0	0	1
Panasonic	Toshiba	Mitsubishi	0	0	1	0	1
Sony	Fujitsu	Hitachi	1	1	2	2	1
Sony	Fujitsu	NEC	1	1	1	0	0
Sony	Fujitsu	Toshiba	2	1	2	1	1
Sony	Fujitsu	Mitsubishi	1	1	2	2	2
Sony	Hitachi	NEC	0	1	1	1	1
Sony	Hitachi	Toshiba	1	1	2	2	2
Sony	Hitachi	Mitsubishi	0	1	1	1	1
Sony	NEC	Toshiba	0	0	1	1	1
Sony	NEC	Mitsubishi	0	0	0	0	0
Sony	Toshiba	Mitsubishi	1	1	1	2	2
Fujitsu	Hitachi	NEC	1	2	1	0	1
Fujitsu	Hitachi	Toshiba	1	2	1	2	1
Fujitsu	Hitachi	Mitsubishi	1	2	1	0	0
Fujitsu	NEC	Toshiba	1	1	1	0	1
Fujitsu	NEC	Mitsubishi	1	1	1	0	0
Fujitsu	Toshiba	Mitsubishi	1	2	2	1	1
Hitachi	NEC	Toshiba	0	1	1	0	1
Hitachi	NEC	Mitsubishi	0	1	0	0	0
Hitachi	Toshiba	Mitsubishi	0	2	1	1	1
NEC	Toshiba	Mitsubishi	0	0	1	0	0



(a) CDS spreads of Sony and Fujitsu (b) Difference between CDS spreads

Figure 3.1: Market CDS spreads of Sony and Fujitsu and difference between both CDS spreads. In the left figure, the solid line shows the CDS spreads of Sony, and the dashed line shows those of Fujitsu, respectively. In the right figure, the solid line shows the difference between both CDS spreads weighted with the cointegration vector $(1, -0.7167)$, which implies that Sony CDS spreads - 0.7167 $\times$ Fujitsu CDS spreads are a stationary process. The x axis represents the number of days between November 20, 2020, and November 30, 2021.

3.2.3 Principal component analysis

We next conduct the principal component analysis (PCA) of the term structure of the CDS spreads of Sony and Fujitsu. Schneider et al. [66] indicate that about two factors are needed to explain the term structure of corporate CDSs, unlike sovereign CDSs, where it is said that one factor is enough, although it depends on the CDS. Therefore, we investigate the number of factors on which the CDS spreads of Sony and Fujitsu depend using the PCA in order to select the number of hazard rates required in later analysis. The target of the PCA is the one, three, and five-year CDS spreads of each company, which will be analyzed in Section 3.4.

In Table 3.5, we show the results of the PCA for the CDS spreads of Sony and Fujitsu. The contribution ratios of the largest eigenvalue are 84% for Sony and 69% for Fujitsu, and the cumulative contribution ratios to the second largest eigenvalue are close to 100% for both companies. That is, for these two Japanese corporate CDSs, we can only explain 70-80% of all the fluctuations of the CDS spreads by one factor, and we can explain almost any fluctuation of CDS spreads by two factors. Eigenvector 1 represents a level; Eigenvector 2 represents a slope; and Eigenvector 3 represents curvature. We need two factors to explain almost any fluctuation, as pointed out in Schneider et al. [66]. However, for the following three reasons, we assume that the number of factors is one for the hazard rates to explain the term structure of CDS spreads in the next section: First, the main fluctuations, which account for 70 to 80% of the total, can be explained by one factor. Second, the analyses become too complicated if we set two factors or more. Third, we cannot obtain the results of analysis in the realistic calculation time because the numerical calculation load goes up dramatically.

By contrast, we understand that it is necessary to investigate the number of factors required to describe the whole term structure of observable CDS spreads using larger and more various datasets although the purpose of this study is to propose the one factor hazard rate model, which is the most basic model. Accordingly, we conduct the PCA for the two types of term structure, which consists of the one, three, and five-year CDS spreads and the one, three, five, seven, and ten-year CDS spreads, of all companies in Table 3.1. The three observation periods,

Table 3.5: Results of the PCA for one, three, and five-year CDS spreads of Sony and Fujitsu. The left part shows the results for Sony, which consist of the correlation matrix of log return (difference of the logarithms of CDS spreads), eigenvalues, and eigenvectors. The right part shows the corresponding results for Fujitsu.

(a) PCA for Sony 1,3,5Y CDS

Correlation matrix

	1Y	3Y	5Y
1Y	1	0.85	0.59
3Y	0.85	1	0.82
5Y	0.59	0.82	1

Eigenvalue

	1	2	3
Eigenvalue	2.51	0.41	0.08
Contribution ratio	84%	14%	3%
Cumulative ratio	84%	97%	100%

Eigenvector

1	2	3
0.56	0.68	0.47
0.61	0.03	-0.79
0.55	-0.73	0.40

(b) PCA for Fujitsu 1,3,5Y CDS

Correlation matrix

	1Y	3Y	5Y
1Y	1	0.98	0.20
3Y	0.98	1	0.24
5Y	0.20	0.24	1

Eigenvalue

	1	2	3
Eigenvalue	2.07	0.91	0.02
Contribution ratio	69%	30%	1%
Cumulative ratio	69%	99%	100%

Eigenvector

1	2	3
0.68	-0.22	0.70
0.68	-0.17	-0.71
0.28	0.96	0.03

which include the major events of COVID-19 (March 2, 2020, to March 10, 2021) and Russia attack on Ukraine (November 19, 2021, to November 30, 2022) in addition to the analysis period in this study (November 20, 2020, to November 30, 2021), are adopted.

In Table 3.6, the row "First eigenvalue" shows what percentage of the 57 first eigenvalues that correspond to the combinations of 19 companies and 3 periods has how much contribution ratio. That is, the value "53%" on the row "First eigenvalue" and the column "80-90%" in the upper table means that the contribution ratios of 53% of the 57 first eigenvalues are in the range of 80 to 90%. The rows "Second eigenvalue" and "Third eigenvalue" represent how much the cumulative contribution ratio up to each eigenvalue is, respectively. That is, the value "95%" on the row "Second eigenvalue" and the column "95-100%" in the upper table means that the cumulative contribution ratios of 95% of the 57 second eigenvalues are in the range of 95 to 100%. All cumulative contribution ratios for the industry and period are gathered in Table 3.6 because the distribution of cumulative contribution ratios has no tendency for the industry and period. However, the number of factors may be smaller in an emergency. In the one, three, and five-year CDS term structure, the upper table in Table 3.6 shows that 81% of the combinations of companies and periods can explain the variation of 80% or more, and 95% of them can explain the variation of 70% or more, with one factor. 95% of them can explain the variation of 95% or more with two factors. On the other hand, in the one, three, five, seven, and ten-year CDS term structure, the lower table in Table 3.6 shows that 84% of the combinations of companies and periods can explain the variation of 70% or more, and 91% of them can explain the variation of 60% or more, with one factor. 84% of them can explain the variation of 90% or more with two factors.

From the above analysis results of multiple companies and periods, it is shown that a one factor hazard rate model can explain a lot of 70-80% of the variation in the term structure of CDS spreads. These results enhance the relevance of using a one factor hazard rate model to describe the whole term structure of observable CDS spreads in this study.

However, it is clear that a hazard rate model of two factors or more is more elaborate.

Table 3.6: The row "First eigenvalue" shows what percentage of the 57 first eigenvalues that correspond to the combinations of 19 companies and 3 periods has how much contribution ratio. The rows "Second eigenvalue" and "Third eigenvalue" represent how much the cumulative contribution ratio up to each eigenvalue is, respectively.

1,3,5Y CDS term structure

Cumulative ratio	0-60%	60-70%	70-80%	80-90%	90-95%	95-100%
First eigenvalue	2%	4%	14%	53%	25%	4%
Second eigenvalue	0%	0%	0%	4%	2%	95%

1,3,5,7,10Y CDS term structure

Cumulative ratio	0-60%	60-70%	70-80%	80-90%	90-95%	95-100%
First eigenvalue	9%	7%	32%	42%	9%	2%
Second eigenvalue	0%	0%	4%	12%	12%	72%
Third eigenvalue	0%	0%	0%	2%	2%	96%

Accordingly, we describe the approach and problem to extend our model to two factors model. It has been shown that the term structure of sovereign CDS spreads can be sufficiently explained by one factor in Pan and Singleton [59]. It is also generally thought that in the field of finance. The reason might be because the sovereign CDS spreads reflect the sovereign credit. In contrast, it is shown that two factors are required to sufficiently explain the term structure of corporate CDS spreads in Schneider et al. [66], as mentioned in the above. The reason might be because the corporate CDS spreads reflect the specific circumstances of company in addition to the corporate credit.

Therefore, we analyze the term structure of sovereign CDS spreads (US, UK, Germany, Singapore, and Australia in the same three periods as the corporate CDS spreads) in the same way based on that of corporate CDS spreads in the above. However, the trend obtained is not significantly different from the corporate CDS spreads. Accordingly, although it depends on the case, we will be able to improve the interpretability of term structure by increasing factors and refining the model. In contrast, it is necessary to investigate how much does it affect the accuracy improvement of pairs trading. To extend our model to two factors model, it is necessary to try not only the improvement of algorithm but also the various approaches including the use of supercomputers and the parallel computing using the GPGPU etc.

3.3 Cointegration analysis for hazard rates

In this section, we describe the ODE-based Bayesian inference method of cointegration dynamics for the hazard rates driving the CDS pairs that consists of two parts: the state space model to describe CDS spreads using hazard rates and the numerical algorithm using ODEs to estimate the parameter posterior distributions adapted to market CDS spreads, where the theoretical CDS spreads cannot be expressed in analytical form. Using the Bayesian inference, we deduce the VECM dynamics of the latent hazard rate processes.

3.3.1 The model

Our model follows the reduced-form approach of Duffie and Singleton [21]. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let $(W_{1t}^P, \cdots, W_{nt}^P), 0 \leq t \leq T$ be an n-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$. The real and risk-neutral probabilities are denoted by P and Q, respectively. The number of cointegrated hazard rates (dimensionality) is denoted by n, and the number of cointegrated vectors (cointegration rank) is denoted by m $(m < n)$.

Conventional hazard rates are ineffective for explaining CDS cointegration. Therefore, we need to assume the model has error correction terms. We consider the state space model, where market CDS spreads are taken to be observables, and hazard rates are treated as continuous latent processes. First, we provide a system of state equations (hereafter referred to as a system equation) for the latent hazard rates. The drift terms of n hazard rate processes have the structure of an n-dimensional VECM, and the diffusion terms have a square root diffusion coefficient since a hazard rate is not a negative value or zero. We write the system of state equations for latent hazard rates as

$$d\lambda_{it} = \left(v_i + \sum_{j=1}^{m} \alpha_{ij} \sum_{k=1}^{n} \beta_{jk} \lambda_{kt} \right) dt + \sigma_i \sqrt{\lambda_{it}} dW_{it}^P, \quad i = 1, \cdots, n \tag{3.1}$$

which is a continuous limit of the discrete-time VECM, where the $m \times 1$ vector $\boldsymbol{B}\boldsymbol{\lambda}_t$ is stationary by the Granger representation theorem when $\boldsymbol{b}_j \equiv (\beta_{j1}, \cdots, \beta_{jn})', j = 1, \cdots, m$ are m linearly independent $n \times 1$ vectors; $\boldsymbol{B} \equiv (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m)'$ is an $m \times n$ matrix; $\boldsymbol{\lambda}_t \equiv (\lambda_{1t}, \cdots, \lambda_{nt})'$ is an $n \times 1$ vector. Then the cointegrated hazard rates are denoted by $(\lambda_{1t}, \cdots, \lambda_{nt})'$; the constant terms are denoted by $(v_1, \cdots, v_n)'$; the adjustment speeds of the error correction terms are denoted by $\alpha_{ij}, i = 1, \cdots, n, j = 1, \cdots, m$; the cointegrated vectors (n-dimensional row vectors that the number is m) are denoted by $\beta_{ji}, i = 1, \cdots, n, j = 1, \cdots, m$; and the volatilities are denoted by $(\sigma_1, \cdots, \sigma_n)'$. The lag of the VECM is 1, and the rank of the $n \times n$ coefficient matrix of $\sum_{j=1}^{m} \alpha_{ij}\beta_{jk}, i = 1, \cdots, n, k = 1, \cdots, n$ is m. We can consider a more general setting of $dW_{it}^P dW_{jt}^P = \rho_{ij} dt, \forall i, j, i \neq j$. However, for simplicity, we set the correlation coefficients to $\rho_{ij} = 0, \forall i, j, i \neq j$, because we focus on the analysis of $\alpha_\bullet, \beta_\bullet$. Approximating (3.1) by Euler's method, the representation in a discrete-time model is

$$\Delta\lambda_{it} = v_i + \sum_{j=1}^{m} \alpha_{ij} \sum_{k=1}^{n} \beta_{jk} \lambda_{k,t-1} + \sqrt{\lambda_{i,t-1}} \epsilon_{it}, \quad i = 1, \cdots, n, \tag{3.2}$$

where white noises are denoted by $(\epsilon_{1t}, \cdots, \epsilon_{nt})'$, whose average vector is an n-dimensional zero vector and covariance matrix is an $n \times n$ diagonal matrix where the diagonal elements are $\sigma_i^2, i = 1, \cdots, n$.

Next, we provide the observation equation. Following Lando [50], we derive the theoretical formula of CDS spreads for the hazard rate model with the error correction term of (3.1). In single-name CDSs, which have one reference entity, a protection buyer pays a periodic premium (insurance fee) to a protection seller. Conversely, a protection seller takes credit risk on a reference entity by assuring to compensate for the loss of reference obligation suffered if a reference entity defaults to a protection buyer. We call the payoff of a buyer and a seller a "fixed leg" and a "protection leg", respectively. Let a default time τ be the first time of occurrence in the Poisson process of the probability intensity (hazard rate) λ_t. We define the information at time t as

$$\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t, \quad 0 \leq t \leq T,$$

where $\mathcal{G}_t$ is a market filtration, and $\mathcal{H}_t = \sigma\{\mathbb{1}_{\{\tau \leq s\}} : 0 \leq s \leq t\}$ is a default filtration. The hazard rates are $\mathcal{G}_t$ adaptable. The indicator function is denoted by $\mathbb{1}_A$, and it represents 1 if A is true, and 0 if false. Then, the fixed leg value at time t is calculated as

$$E^Q \left[\sum_{j=1}^{gl} e^{-\int_t^{t_j} r_u du} Nc(1/g) \mathbb{1}_{\{\tau > t_j\}} | \mathcal{F}_t \right] = \mathbb{1}_{\{\tau > t\}} Nc(1/g) \sum_{j=1}^{gl} D(t, t_j) E^Q \left[e^{-\int_t^{t_j} \lambda_u du} | \mathcal{G}_t \right], \tag{3.3}$$

where $E^Q[\bullet]$ represents an expectation under the risk-neutral probability measure Q; N is a notional amount; c is a premium; g is the number of time points to pay a premium per year;

l is a term of CDS; and $t_j, j = 1, \cdots, gl$ is time points to pay a premium. We provide the structure of risk-free rates with a spot rate process r_t. A default-free discount bond price under the martingale measure Q is given by

$$D(t,s) = E^Q\left[e^{-\int_t^s r_u du}|\mathcal{F}_t\right]. \tag{3.4}$$

In this study, assuming that the behavior between short-term interest rates and default events is independent of each other, we avoid estimating the term structure model of interest rates. On the other hand, the protection leg value at time t is calculated as

$$E^Q\left[e^{-\int_t^\tau r_u du}N(1-\theta)|\mathcal{F}_t\right] = \mathbb{1}_{\{\tau>t\}}N(1-\theta)\int_t^T D(t,s)E^Q\left[\lambda_s e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right]ds, \tag{3.5}$$

where θ is a recovery rate, and T is the maturity of CDS. Following in general quotes in the market, we take $\mathbb{1}_{\{\tau>t\}} = 1$ and $c = CDS_t^{mod,i,l}$. Then, as a fixed leg (3.3) and a protection leg (3.5) match, we obtain the theoretical formula for CDS spreads. Accordingly, we have the theoretical formula for CDS spreads and the observation equation as follows:

$$\begin{cases} CDS_t^{mkt,i,l} = CDS_t^{mod,i,l} + e_t^{i,l} \\ CDS_t^{mod,i,l} = \dfrac{(1-\theta)\int_t^T D(t,s)V_p^i(t,s)ds}{1/g\sum_{j=1}^{gl} D(t,t_j)V_f^i(t,t_j)}, \quad i = 1, \cdots, n, \end{cases} \tag{3.6}$$

where $CDS_t^{mkt,i,l}$ is a market CDS spread of l year standard CDS of company i at time t, $CDS_t^{mod,i,l}$ is the same theoretical CDS spread, and $e_t^{i,l}$ is an observation error. We define

$$\begin{cases} V_f^i(t,s) = E^Q\left[e^{-\int_t^s \lambda_{iu}du}|\mathcal{G}_t\right] \\ V_p^i(t,s) = E^Q\left[\lambda_{is}e^{-\int_t^s \lambda_{iu}du}|\mathcal{G}_t\right], \end{cases} \quad i = 1, \cdots, n \tag{3.7}$$

on a measurable set where a default does not occur.

Since a hazard rate λ_t represents a default intensity, $V_f^i(t,s)$ is a survival function of company i, and $V_p^i(t,s)$ is a density function at the default of company i. We assume that the number of factors is one for the hazard rates to explain the term structure of CDS spreads, as explained in the PCA in Section 3.2.3. A more detailed analysis of two factors or more is a future issue.

The system equation (3.1) and observation equation (3.6) compose a state space model. Because the theoretical value of CDS spread must be calculated under the risk-neutral measure Q as described in (3.6), the Girsanov theorem is used to define the risk-neutral measure Q following Shreve [67]. Let us assume the market prices of credit risk of n entities as

$$\psi_{it} = \phi_{i1}\sqrt{\lambda_{it}} + \frac{\phi_{i2}}{\sqrt{\lambda_{it}}}, \quad i = 1, \cdots, n. \tag{3.8}$$

This slightly general constrained functional form of the market price of credit risk (3.8) guarantees that the form of the mean-reverting square-root process (3.1) is unchanged after a Girsanov transformation defined in (3.9). In order to change the probability measure using the Girsanov theorem, we define the Radon-Nikodym derivative of the equivalent martingale measure Q with respect to P denoting theoretically the maximum time horizon as T as follows:

$$\begin{cases} Z_T = \dfrac{dQ}{dP} = \exp\left[-\sum_{i=1}^n \int_0^T \psi_{iu}dW_{iu}^P - \frac{1}{2}\int_0^T \sum_{i=1}^n \psi_{iu}^2 du\right]. \\ W_{it}^Q = W_{it}^P + \int_0^t \psi_{iu}du, \quad i = 1, \cdots, n \end{cases} \tag{3.9}$$

In $0 \leq t \leq T$, the Radon-Nikodym derivative process is $Z_t = E^P[Z_T|\mathcal{G}_t]$. Then $(W_{1t}^Q, \cdots, W_{nt}^Q), 0 \leq t \leq T$ is an n-dimensional Brownian motion under the probability measure Q.

Accordingly, the stochastic differential equation (SDE) of hazard rates represented in Q is written as

$$\begin{aligned} d\lambda_{it} &= \left(v_i + \sum_{j=1}^{m} \alpha_{ij} \sum_{k=1}^{n} \beta_{jk} \lambda_{kt} \right) dt + \sigma_i \sqrt{\lambda_{it}} (dW_{it}^Q - \psi_{it} dt) \\ &= \left(s_i + \sum_{k=1}^{n} \gamma_{ik} \lambda_{kt} \right) dt + \sigma_i \sqrt{\lambda_{it}} dW_{it}^Q, \quad i = 1, \cdots, n, \end{aligned} \tag{3.10}$$

where

$$\begin{cases} s_i = v_i - \sigma_i \phi_{i2} \\ \gamma_{ik} = \sum_{j=1}^{m} \alpha_{ij} \beta_{jk} - \sigma_k \phi_{k1} \delta_{ik} \end{cases}, \quad i = 1, \cdots, n, k = 1, \cdots, n.$$

The symbol δ_{ij} represents the Kronecker delta and is a two-variable function $\delta_{ij} : \mathbb{N} \times \mathbb{N} \to \{0, 1\}$ defined by

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

regarding the member i, j in the set $\mathbb{N}$. Changing the probability measure from P to Q, the coefficients of the SDE of hazard rates under Q alter as s_i and γ_{ik} in (3.10). This implies that there is a possibility that the SDE in Q may extend beyond the class of VECM, that is, vector autoregression (VAR), depending on the rank of γ_{ik}. In that case, the cointegration introduced in P disappears in Q. This point will be elaborated on later in Section 3.4.2 from the standpoint of the strong linearity of the CDS spread with respect to hazard rates.

The key expectations of (3.7) in the CDS spread formula (3.6) are the solution $V^i(t, \lambda_1, \cdots, \lambda_n)$ of the PDEs

$$\begin{aligned} & V_t^i(t, \lambda_1, \cdots, \lambda_n) \\ & + \sum_{k=1}^{n} \{ (s_k + \sum_{j=1}^{n} \gamma_{kj} \lambda_j) V_{\lambda_k}^i(t, \lambda_1, \cdots, \lambda_n) + \frac{1}{2} \sigma_k^2 \lambda_k V_{\lambda_k \lambda_k}^i(t, \lambda_1, \cdots, \lambda_n) \} \\ & = \lambda_i V^i(t, \lambda_1, \cdots, \lambda_n), \quad i = 1, \cdots, n \end{aligned} \tag{3.11}$$

derived from no-arbitrage arguments, where the corresponding boundary condition is provided as

$$V^i(T, \lambda_i) = \begin{cases} 1 & \text{(fixed leg)} \\ \lambda_i & \text{(protection leg)} \end{cases}, \quad i = 1, \cdots, n,$$

and we introduced a shorthand notation of the partial derivatives of $V^i(t, \lambda_1, \cdots, \lambda_n)$ as

$$\begin{cases} V_t^i(t, \lambda_1, \cdots, \lambda_n) = \dfrac{\partial}{\partial t} V^i(t, \lambda_1, \cdots, \lambda_n) \\ V_{\lambda_k}^i(t, \lambda_1, \cdots, \lambda_n) = \dfrac{\partial}{\partial \lambda_k} V^i(t, \lambda_1, \cdots, \lambda_n) \\ V_{\lambda_k \lambda_k}^i(t, \lambda_1, \cdots, \lambda_n) = \dfrac{\partial^2}{\partial \lambda_k^2} V^i(t, \lambda_1, \cdots, \lambda_n) \end{cases}.$$

These PDEs (3.11) reduce the ODE systems in a standard manner. As for the expectation of the fixed leg, we assume the exponential affine form with respect to the hazard rates:

$$V_f^i(t,s) = E^Q\left[e^{-\int_t^s \lambda_{iu}du}|\mathcal{G}_t\right] = e^{A_{i0}(\tau)-\sum_{k=1}^n A_{ik}(\tau)\lambda_{kt}}, \quad \tau = s-t. \tag{3.12}$$

Substituting the partial derivatives for the PDEs (3.11) and arranging the terms with respect to the terms proportional to $\lambda_k, k = 1, \cdots, n$ and the constant terms, we obtain the ODEs of A's:

$$\begin{cases} A'_{i0}(\tau) = -\sum_{k=1}^n s_k A_{ik}(\tau), \quad A_{i0}(0) = 0 \\ A'_{ij}(\tau) = \sum_{k=1}^n \gamma_{kj} A_{ik}(\tau) - \frac{1}{2}\sigma_j^2 A_{ij}^2(\tau) + \delta_{ij}, \quad A_{ij}(0) = 0 \end{cases} \tag{3.13}$$

for $i = 1, \cdots, n, j = 1, \cdots, n$, where $'$ denotes a derivative with respect to τ. Similarly, as for the expectation of the protection leg, we assume

$$\begin{aligned} V_p^i(t,s) &= E^Q\left[\lambda_{is}e^{-\int_t^s \lambda_{iu}du}|\mathcal{G}_t\right] \\ &= e^{A_{i0}(\tau)-\sum_{k=1}^n A_{ik}(\tau)\lambda_{kt}}\{a_{i0}(\tau) + \sum_{k=1}^n a_{ik}(\tau)\lambda_{kt}\}, \quad \tau = s-t. \end{aligned} \tag{3.14}$$

In the same way as the calculation of the expectation of the fixed leg, we obtain the ODEs of a's:

$$\begin{cases} a'_{i0}(\tau) = \sum_{k=1}^n s_k a_{ik}(\tau), \quad a_{i0}(0) = 0 \\ a'_{ij}(\tau) = \sum_{k=1}^n \gamma_{kj} a_{ik}(\tau) - \sigma_j^2 A_{ij}(\tau) a_{ij}(\tau), \quad a_{ij}(0) = \delta_{ij} \end{cases} \tag{3.15}$$

for $i = 1, \cdots, n, j = 1, \cdots, n$, where A's are the solutions of (3.13). The derivation of the ODEs can be found in Appendix B.1.

3.3.2 Algorithm to numerically calculate hazard rate parameters

General setting for the two-dimensional case

In the previous section, we set the theoretical framework of the model by using the state space model. In this section, we describe the method to estimate the parameters of the model and the time series of latent hazard rates using the numerical calculation of Bayesian inference. To estimate the parameters of a statistical model using the MCMC method, we use Stan [68], which is a free software for the MCMC. In Stan [68], the HMC method [18] [57] has been adopted as the sampling method of random numbers. In the HMC, we regard the distribution we want to sample as the potential energy and introduce the kinetic energy. We increase the transition distance and the acceptance rate of candidates by utilizing the properties that Hamiltonian, which is the sum of both energies, saves. We adopt the HMC to overcome the drawback that the MH algorithm becomes impractical because the drop in acceptance rate is significant in a large number of parameters and to efficiently generate random numbers following the distribution we want to sample. The explanation of the HMC method can be found in Appendix A.3.

We deal with the estimation of a two-dimensional cointegrated hazard rate model since it is enough to illustrate our numerical algorithm. In this case, the cointegrated vector is one. We

can obtain the two-dimensional state space model substituting $n = 2$ and $m = 1$ for the model (3.1) provided in the previous section. The VECM for hazard rates (system equation) under the real probability P is as follows:

$$\begin{cases} d\lambda_{1t} = \{v_1 + \alpha_{11}(\beta_{11}\lambda_{1t} + \beta_{12}\lambda_{2t})\}dt + \sigma_1\sqrt{\lambda_{1t}}dW_{1t}^P \\ d\lambda_{2t} = \{v_2 + \alpha_{21}(\beta_{11}\lambda_{1t} + \beta_{12}\lambda_{2t})\}dt + \sigma_2\sqrt{\lambda_{2t}}dW_{2t}^P \end{cases}, \tag{3.16}$$

where we note that for λ_{1t} and λ_{2t} to keep a mean-reverting property, the eigenvalues of 2×2 matrices of the coefficients of $\lambda_{1t}, \lambda_{2t}$ in the drift terms $\alpha_{i1}\beta_{1j}, i, j = 1, 2$ must be negative. Therefore, $\alpha_{11}\beta_{11} + \alpha_{21}\beta_{12} < 0$. Then, $v_1 > 0, v_2 > 0, \sigma_1 > 0, \sigma_2 > 0$. We obtain the Euler discretization formulae for (3.16) discretizing the formula at the observation time points, $t_x, x = 0, \cdots, \Lambda$ for market CDS spreads $CDS_{t_x}^{mkt,i,l}$ (Λ: the number of the observation time points), from which we can conduct sequentially the Monte Carlo sampling of the cointegrated hazard rates as

$$\begin{cases} \lambda_{1,t_x} \sim N(\lambda_{1,t_{x-1}} + \{v_1 + \alpha_{11}(\beta_{11}\lambda_{1,t_{x-1}} + \beta_{12}\lambda_{2,t_{x-1}})\}(t_x - t_{x-1}), \sigma_1^2\lambda_{1,t_{x-1}}(t_x - t_{x-1})) \\ \lambda_{2,t_x} \sim N(\lambda_{2,t_{x-1}} + \{v_2 + \alpha_{21}(\beta_{11}\lambda_{1,t_{x-1}} + \beta_{12}\lambda_{2,t_{x-1}})\}(t_x - t_{x-1}), \sigma_2^2\lambda_{2,t_{x-1}}(t_x - t_{x-1})) \end{cases}, \tag{3.17}$$

which corresponds to (3.2), where $N(m, s^2)$ represents a normal distribution of average m and variance s^2.

The observation equation (3.6) is slightly modified from the standpoint of the efficiency and tractability of the Bayesian statistical inference:

$$\begin{cases} CDS_t^{mkt,i,l} = CDS_t^{mod,i,l} + e_t^{i,l} \\ CDS_t^{mod,i,l} = \dfrac{(1-\theta)\Delta\sum_{k=1}^{\xi l} D(t, s_k)V_p^i(t, s_k)}{1/g\sum_{j=1}^{gl} D(t, t_j)V_f^i(t, t_j)}, \quad i = 1, 2. \end{cases} \tag{3.18}$$

Although the protection leg of (3.6) is represented in a continuous-time integral until the maturity of the CDS contract, we approximate this as a discretized summation as

$$\int_t^T D(t, s)V_p^i(t, s)ds \approx \sum_{k=1}^{\xi l} \Delta D(t, s_k)V_p^i(t, s_k), \tag{3.19}$$

where ξ denotes the number of time points for the discrete approximation per year; Δ is a time interval of the discrete approximation ($= 1/\xi$); and $s_k, k = 1, \cdots, \xi l$ is time points for the discrete approximation. For example, the discrete approximation (3.19) can be found in Duffie and Singleton [22]. The form of the market price of credit risk (3.8) deserves consideration because the term proportional to $\sqrt{\lambda_t}$ influences the cointegrated vector, and the term proportional to $1/\sqrt{\lambda_t}$ influences the constant term but not the cointegrated vector. We assume leaving only the term proportional to $\sqrt{\lambda_t}$, that is, equivalently $\phi_{i2} = 0, i = 1, 2$.

To calculate the theoretical CDS spreads (3.18), we need to calculate the formulae obtained by substituting $n = 2$ for (3.12) and (3.14) at each time point. We can obtain the ODEs by substituting $n = 2$ and $m = 1$ for (3.13) and (3.15). If we could analytically calculate the ODEs, we could speedily estimate the parameters by the MCMC using the concerned analytical solution. However, because the second formula in the ODEs of the fixed leg (3.13) is the multidimensional Riccati ODE in two dimensions or more, it is difficult to solve these ODEs analytically in general. Accordingly, we also cannot analytically calculate the theoretical CDS spread and cannot estimate the parameters. Because it is difficult to solve $A_{ij}(\tau)$ analytically in the second formula in the ODEs of the protection leg (3.15), we cannot analytically solve these ODEs in the same way. For this reason, we obtain the theoretical CDS spread (3.18)

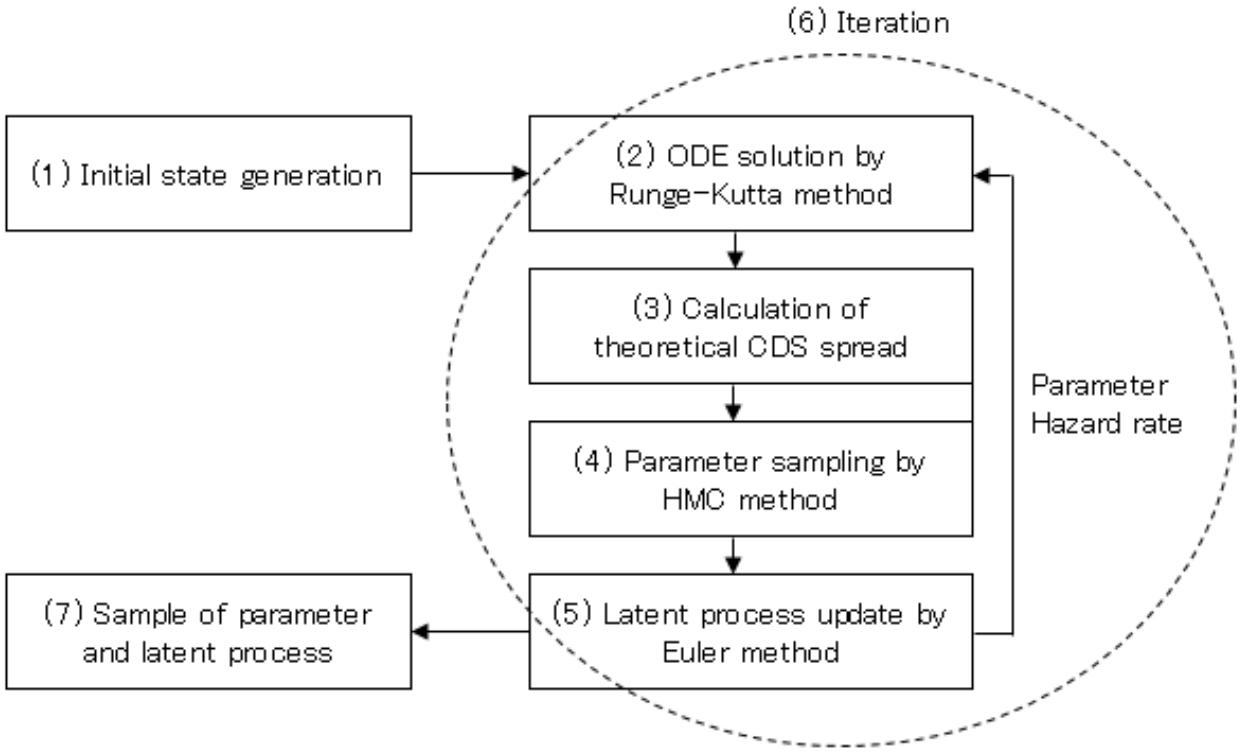


Figure 3.2: Flowchart of the estimation in the ODE-based MCMC. The details of each step are shown in the algorithm described in the text.

by solving the ODEs using the numerical calculation of the Runge-Kutta method [26] in this method.

The flowchart in Figure 3.2 illustrates our estimation algorithm by Bayesian statistical inference (HMC) combined with the ODE solver. Our algorithm is shown step-by-step in a sequential procedure:

Step 1 Generate the initial states of parameters $(v_1, v_2, \alpha_{11}, \alpha_{21}, \beta_{11}, \beta_{12}, \sigma_1, \sigma_2, \phi_{11}, \phi_{21})$ and latent variables $(\lambda_{1,t_0}, \lambda_{2,t_0})$.

Step 2 Solve the ODEs, which $A_{i0}, A_{i1}, A_{i2}, a_{i0}, a_{i1}, a_{i2}, i = 1, 2$ satisfy, at the time points to pay a premium $t_j, j = 1, \cdots, gl$ and the time points for the discrete approximation in the integral of the protection leg $s_k, k = 1, \cdots, \xi l$ using the Runge-Kutta method.

Step 3 Calculate $V_f^i(t, t_j), j = 1, \cdots, gl$ and $V_p^i(t, s_k), k = 1, \cdots, \xi l$, and find the theoretical CDS spread (3.18).

Step 4 Take samples of parameters from the posteriors substituted the market CDS spreads for using the HMC method.

Step 5 Update latent processes $(\lambda_{1,t_x}, \lambda_{2,t_x})$ at the observation time points $t_x, x = 1, \cdots, \Lambda$ for CDS spreads using the formulae (3.17) for the path of hazard rates approximated by the Euler's method.

Step 6 Repeat the flow from Step 2 to Step 5 sufficiently until the parameters and latent variables converge.

Step 7 Aggregate the posterior samples of parameters and latent variables.

Incidentally, we can analytically solve the first formula in the ODEs of the fixed leg (3.13) and that in the ODEs of the protection leg (3.15), respectively. Although we solve all ODEs of both legs using the numerical calculation of the Runge-Kutta method in this study, we can reduce the computational cost using the analytical solutions in the calculation for $A_{i0}(\tau)$ and $a_{i0}(\tau)$ in Step 2.

Approach for solving the two-dimensional case

In the previous subsection, we explained the general setting for estimating the case of two dimensions ($n = 2$). The algorithm updates the parameters and latent variables simultaneously and calculates them all at once. However, because this algorithm requires a high computational load, it is difficult to attain the convergence of parameters and latent variables in a realistic amount of calculation time. The computational cost is expensive because this algorithm numerically calculates the ODEs using the Runge-Kutta method within the HMC sampling phase. In particular, the computational load of latent processes is expensive because two-time series $\lambda_{1,t_x}, \lambda_{2,t_x}, x = 0, \cdots, \Lambda$ consist of $2(\Lambda + 1)$ variables. For example, if the daily data of CDS spreads for one year is observed, $\Lambda = 250$, and the number of latent variables is about 502.

Conversely, we can obtain the analytical solutions of the ODEs in the conventional standard model. In that case, the system equation under the real probability P is as follows:

$$d\lambda_t = (v + \alpha\lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t^P, \tag{3.20}$$

where $v > 0, \alpha < 0, \sigma > 0$. The observation equation is the same as in the cointegrated model. The system equation under the risk-neutral probability Q is as follows:

$$d\lambda_t = (s + \gamma\lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t^Q, \tag{3.21}$$

where $s = v, \gamma = \alpha - \sigma\phi$. We can obtain the ODEs with respect to the coefficients A_0, A_1, a_0, and a_1 of the exponential affine model

$$\begin{cases} V_f(t,s) = E^Q\left[e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right] = e^{A_0(\tau)-A_1(\tau)\lambda_t} \\ V_p(t,s) = E^Q\left[\lambda_s e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right] = e^{A_0(\tau)-A_1(\tau)\lambda_t}\{a_0(\tau) + a_1(\tau)\lambda_t\} \end{cases}, \quad \tau = s - t, \tag{3.22}$$

in a standard manner. In the fixed leg, the ODEs are as follows:

$$\begin{cases} A_0'(\tau) = -sA_1(\tau), \quad A_0(0) = 0 \\ A_1'(\tau) = \gamma A_1(\tau) - \dfrac{1}{2}\sigma^2 A_1^2(\tau) + 1, \quad A_1(0) = 0 \end{cases}. \tag{3.23}$$

In the protection leg, the ODEs are as follows:

$$\begin{cases} a_0'(\tau) = sa_1(\tau), \quad a_0(0) = 0 \\ a_1'(\tau) = \gamma a_1(\tau) - \sigma^2 A_1(\tau)a_1(\tau), \quad a_1(0) = 1 \end{cases}. \tag{3.24}$$

The analytical solutions of these ODEs (3.23) and (3.24) are as follows:

$$\begin{cases} A_0(\tau) = \dfrac{2s}{\sigma^2}\ln\dfrac{\xi e^{-\frac{\gamma}{2}\tau}}{\xi\cosh\xi\tau - \frac{\gamma}{2}\sinh\xi\tau} \\ A_1(\tau) = \dfrac{\sinh\xi\tau}{\xi\cosh\xi\tau - \frac{\gamma}{2}\sinh\xi\tau} \\ a_0(\tau) = sA_1(\tau) \\ a_1(\tau) = \dfrac{\xi^2}{(\xi\cosh\xi\tau - \frac{\gamma}{2}\sinh\xi\tau)^2} \end{cases}, \tag{3.25}$$

where

$$\xi = \frac{1}{2}\sqrt{\gamma^2 + 2\sigma^2}. \tag{3.26}$$

The derivation of the ODE analytical solutions (3.25) can be found in Appendix B.2.

Based on the above, we apply the estimation for the conventional standard model to each of the two CDSs. That is, we consider speedily estimating the time series of latent hazard rates $\lambda_{1,t_x}, \lambda_{2,t_x}$ using the analytical solutions of ODEs (3.25). In the estimation for the conventional standard model, two CDSs are not cointegrated. However, considering that the estimated hazard rates represent the instantaneous default probability, which implies the starting point in the term structure (curve) of CDS spreads at each time point $t_x, x = 0, \cdots, \Lambda$, even if we estimate as the separate problems, we can expect to obtain values close to the estimated values of hazard rates in the cointegrated model. The specification of the conventional standard CDS model can be found in Appendix B.3. Then, given the estimated values of $\lambda_{1,t_x}, \lambda_{2,t_x}, x = 0, \cdots, \Lambda$ obtained using the estimation method for the conventional standard model, we estimate the parameters of the model using the estimation method excluding Step 5 from the algorithm for two dimensions ($n = 2$) in the previous subsection. Although it seems that the parameters and latent processes obtained so far are close to the values estimated simultaneously, the values are different from the results estimated simultaneously. So finally, setting the prior distributions, which have the average and variance close to those of the posterior distributions of the latent variables obtained in the conventional standard model and the parameters obtained in the two-dimensional estimation excluding Step 5, we estimate the parameters and latent processes simultaneously. From the above, we estimate the parameters and latent processes by a three-step estimation as follows:

Stage 1 Conduct the estimation of the conventional standard model described in (3.20)-(3.26) separately for each of the two CDSs for ex-ante identification of the time series of latent hazard rates $\lambda_{1,t_x}, \lambda_{2,t_x}, x = 0, \cdots, \Lambda$ using the analytical solutions of the ODEs (3.25).

Stage 2 Given the estimated values of $\lambda_{1,t_x}, \lambda_{2,t_x}, x = 0, \cdots, \Lambda$, estimate the parameters of model $(v_1, v_2, \alpha_{11}, \alpha_{21}, \beta_{11}, \beta_{12}, \sigma_1, \sigma_2, \phi_{11}, \phi_{21})$ using the algorithm for cointegrated CDSs excluding that of latent variables (Step 5) in two dimensions ($n = 2$).

Stage 3 Set the prior distributions which have the average and variance close to those of the posterior distributions of the latent variables obtained in Stage 1 and the parameters obtained in Stage 2, and estimate the parameters and latent processes simultaneously using the algorithm for cointegrated CDSs in two dimensions ($n = 2$) (including Step 5).

In Stage 3, we aim to speed up convergence to a stationary distribution of samples by estimating the sampling range around the estimated value obtained in Stages 1 and 2 simultaneously. Accordingly, Stage 3 is the essence of estimation, which is explained in the previous subsection, and Stages 1 and 2 are the stages of preparation to speed up convergence in Stage 3. In the numerical calculation in Section 3.4.1, we confirm the significant reduction in estimated time by estimating the above three processes. Incidentally, as we use constant maturity CDSs at each time point $x(= 0, \cdots, \Lambda)$, once we solve the ODEs until the maximum maturity, we can evaluate the CDSs of all time points. That relieves the computational load. If we had the same CDS at each time point, as the maturity became shorter, we would need to solve the ODEs at each time point. In that case, the computational load would be very expensive.

3.3.3 Connection with statistical mechanics

We recognize the connection between the cointegration of n hazard rates, which is the main problem considered in this study, and concepts or methods of statistical mechanics from the point of view of econophysics.

Equation (3.1), which is a continuous limit of the ECM of hazard rates considered in this study, is the special case of the Langevin equation in non-equilibrium statistical mechanics:

$$\frac{d\lambda_{it}}{dt} = v_i + \sum_{j=1}^{m} \alpha_{ij} \sum_{k=1}^{n} \beta_{jk} \lambda_{kt} + f_i(t, \lambda_{1t}, \ldots, \lambda_{nt}), \quad i = 1, \ldots, n, \tag{3.27}$$

where a fluctuation force is denoted by $f_i(t, \lambda_{1t}, \ldots, \lambda_{nt})$. The following holds for a fluctuation force:

$$\begin{cases} \langle f_i(t, \lambda_{1t}, \ldots, \lambda_{nt}) \rangle = 0, \quad i = 1, \ldots, n \\ \langle f_i(t, \lambda_{1t}, \ldots, \lambda_{nt}) f_j(t', \lambda_{1t'}, \ldots, \lambda_{nt'}) \rangle = \sigma_i^2 \lambda_{it} \delta(t - t') \delta_{ij} \end{cases}, \quad i = 1, \ldots, n, j = 1, \ldots, n, \tag{3.28}$$

where $\delta(t-t')$ is the Dirac delta function. Each symbol has an analogy with statistical mechanics as follows: the hazard rate λ_{it} corresponds to the i-th component of the velocity vector of a particle following a Brownian motion; the first term in the RHS in (3.27) corresponds to a viscous friction; the quantity $\sigma_i^2 \lambda_{it}$ in the second formula in (3.28) corresponds to a fluctuation size.

The dynamical system of an n-dimensional hazard rate $(\lambda_{1t}, \ldots, \lambda_{nt})'$ evolves over time under a certain number of constraints depending on the rank of error correction term. If the rank is m, that is,

$$\text{rank}[AB'] = m, \quad 0 < m < n,$$

where

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \cdots & \beta_{mn} \end{pmatrix},$$

when there appears to be $n - m$ constraints among $\lambda_{1t}, \ldots, \lambda_{nt}$ that are represented in their stationary linear combinations with $n - m$ cointegration vectors belonging to the kernel space of the transpose of AB'. Physically, this means that there are $n - m$ directions in the variable space orthogonal to the external force (negative of gradient of latent potential energy). Then, the discrete-time model (3.2) of equation (3.1) is an ECM. If $m = n$, equation (3.2) is a VAR model since all $\lambda_{it}, i = 1, \ldots, n$ are recognized as I(0)-processes. In contrast, if $m = 0$, the difference of $\lambda_{it}, i = 1, \ldots, n$ is a VAR model since no cointegration vector exists.

Identifying the fluctuation force $f_i(t, \lambda_{1t}, \ldots, \lambda_{nt})dt$ with Brownian fluctuation $\sigma_i \sqrt{\lambda_{it}} dW_{it}^P$, we can obtain the SDE of equation (3.1):

$$f_i(t, \lambda_{1t}, \ldots, \lambda_{nt})dt = \sigma_i \sqrt{\lambda_{it}} dW_{it}^P, \quad i = 1, \ldots, n.$$

Our ECM of hazard rates is thought of as a financial analog of constrained Langevin diffusion in statistical mechanics.

3.3.4 Model II proposed in this study

We propose the ODE-based Bayesian inference for cointegration dynamics of hazard rates and CDSs. This model consists of the two parts: the state space model for cointegration and the algorithm to numerically calculate parameters and latent processes. This model can be applied to pairs trading strategies consistent with the framework of arbitrage-free pricing. We summarize the credit cointegration model proposed in this study as follows.

- Model II

– State space model

$$\begin{cases} d\lambda_{it} = \left(v_i + \sum_{j=1}^{m} \alpha_{ij} \sum_{k=1}^{n} \beta_{jk} \lambda_{kt} \right) dt + \sigma_i \sqrt{\lambda_{it}} dW_{it}^P \\ CDS_t^{mkt,i,l} = CDS_t^{mod,i,l} + e_t^{i,l} \\ CDS_t^{mod,i,l} = \dfrac{(1-\theta)\Delta \sum_{k=1}^{\xi l} D(t,s_k) V_p^i(t,s_k)}{1/g \sum_{j=1}^{gl} D(t,t_j) V_f^i(t,t_j)} \end{cases}, \quad i = 1, \ldots, n$$

– Algorithm to numerically calculate parameters
Sample the state of parameters and latent processes randomly from the parameter distribution of the state space model using the HMC method combined with the Runge-Kutta method to numerically solve the ODEs and Euler's method.

Parameter distribution: $\{(\Theta^{(1)}, \Psi^{(1)}), \ldots, (\Theta^{(K)}, \Psi^{(K)})\}$,

$$\begin{cases} \Theta^{(k)} = \{v_i, \alpha_{ij}, \beta_{ji}, \sigma_i, \phi_{i1}\} \\ \Psi^{(k)} = \{\lambda_{i,t_0}, \ldots, \lambda_{i,t_x}, \ldots, \lambda_{i,t_\Lambda}\} \end{cases}, \quad i = 1, \ldots, n, j = 1, \ldots, m, k = 1, \ldots, K$$

The trial number of simulation is denoted by K. These parameters and latent processes can be calculated according to the flowchart in Figure 3.2.

As mentioned above, this model can calculate the parameter distributions for an arbitrary CDS portfolio that can be described using the state space model where the dimensionality n and cointegration rank m are arbitrary. Actually, we show the calculation results in the case of $n = 2$ and $m = 1$. However, to calculate the parameters and latent processes and apply them to pairs trading in the case of $n \geq 3$ or $m \geq 2$, we have to develop an efficient method to reduct computational load in the future.

3.4 CDS pairs trading strategy

To analyze the cointegration of hazard rates and CDSs, we estimate the parameters and latent variables by applying our method to real data. We set the data used in the parameter estimation using our method and show the results of the numerical calculation. We consider a pairs trading strategy based on the estimated results.

3.4.1 Empirical analysis

We provide the data description and the numerical setting for estimation in our empirical studies. The CDS cointegrated pair of the analysis target is the pair of Sony ($i = 1$) and Fujitsu ($i = 2$) analyzed in Section 3.2. The three representative terms of one, three, and five years ($l = 1, 3, 5$) from the entire term structure data are chosen due to computational load. We confirm that the CDS pair of Sony and Fujitsu has the cointegration property to some degree in terms of one, three, and five years. The time series of all terms of Sony and Fujitsu CDS spreads are daily data spanning 251 business days ($\Lambda = 250$) from November 20, 2020, to November 30, 2021, obtained from Thomson Reuters. The time series of the term structure of risk-free interest rates are the daily data of Japanese yen discount factor curves calculated from the swap rates of overnight index swap (OIS) obtained from Bloomberg. Then, the number of time points to pay a premium per year g is 2; the recovery rate θ is 0.4; and the number of time points for the discrete approximation per year in the integral of the protection leg ξ is 12.

We show the numerical calculation results estimated under the above setting. Table 3.7 shows the estimation results of the parameters of the VECM for hazard rates (system equation)

Table 3.7: Estimation results of the parameters of the VECM for hazard rates (system equation) and the latent variables at the start and end of the observation period under the real probability P. This table reports the mean (expected a posteriori estimate) and standard deviation of the posterior distributions by the MCMC. The Rhat indicates a convergence test index in the MCMC. If the number of chains is three or more and the Rhat of all parameters is below 1.1, it suggests that the sampled values converge to a stationary distribution.

	v_1	α_{11}	β_{11}	σ_1	ϕ_{11}	λ_{1,t_0}	$\lambda_{1,t_{250}}$
Mean	0.00065	-0.598	0.329	0.040	-2.361	0.00062	0.00035
Standard deviation	0.00001	0.125	0.097	0.001	0.908	0.00004	0.00004
Rhat	1.01	1.00	1.00	1.00	1.00	1.00	1.00
	v_2	α_{21}	β_{12}	σ_2	ϕ_{21}	λ_{2,t_0}	$\lambda_{2,t_{250}}$
Mean	0.00066	-0.012	-0.590	0.048	-7.326	0.00066	0.00025
Standard deviation	0.00001	0.080	0.120	0.001	0.518	0.00002	0.00002
Rhat	1.01	1.01	1.00	1.00	1.01	1.00	1.00

and the latent variables at the start and end of the observation period under the real probability P. These are the converged values by the three-stage estimation in the previous subsection. These are the estimates of constant terms v_1, v_2, adjustment speeds of the error correction terms α_{11}, α_{21}, cointegrated vectors β_{11}, β_{12}, volatilities σ_1, σ_2, market prices of credit risk ϕ_{11}, ϕ_{21}, and latent variables $\lambda_{1,t_0}, \lambda_{1,t_{250}}, \lambda_{2,t_0}, \lambda_{2,t_{250}}$. The numerical values in Table 3.7 represent the mean (expected a posteriori estimate) and standard deviation of the posterior distributions by the MCMC.

The quantity Rhat indicates a convergence test index in the MCMC. If the number of chains is three or more and the Rhat of all parameters is below 1.1, it suggests that the sampled values converge to a stationary distribution. The index Rhat is defined as follows:

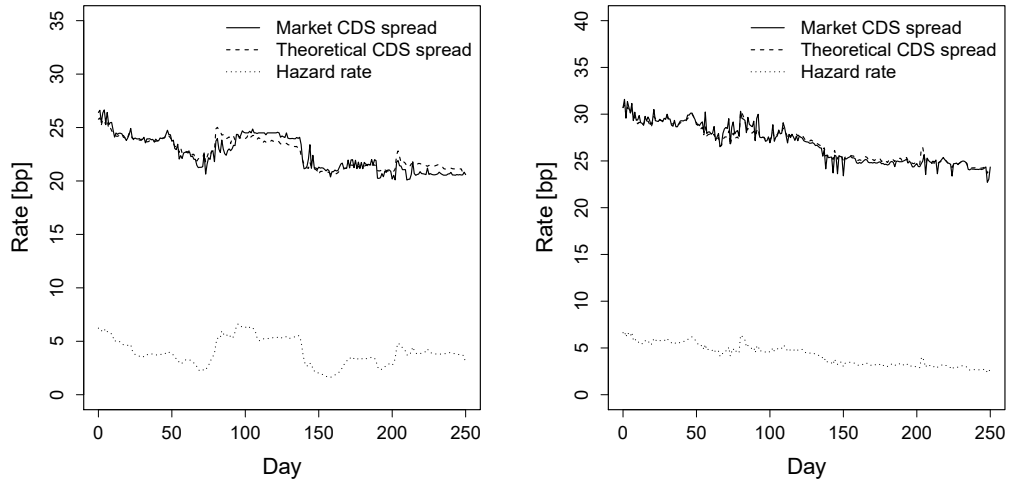
$$\text{Rhat} = \sqrt{\frac{V^+(\theta|y)}{W}},$$

where

$$\begin{cases} V^+(\theta|y) = \dfrac{n-1}{n}W + \dfrac{1}{n}B \\ B = \dfrac{n}{m-1}\displaystyle\sum_{j=1}^{m}(\mu_j - \nu)^2, \quad \mu_j = \dfrac{1}{n}\sum_{i=1}^{n}\theta_{ij}, \quad \nu = \dfrac{1}{m}\sum_{j=1}^{m}\mu_j \\ W = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m}s_j^2, \quad s_j^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(\theta_{ij} - \mu_j)^2 \end{cases}.$$

We generate multiple sequences and discard the warm-up period. Then, we split each of these chains into the first and second halves. Let m be the number of chains (after splitting) and n be the length of each chain. For each scalar estimand θ, we regard the samples of chains as $\theta_{ij}, i = 1, \cdots, n, j = 1, \cdots, m$. Therefore, B is the between-sequence variance, and W is the mean of the within-sequence variances. The quantity $V^+(\theta|y)$ overestimates the marginal posterior variance, but the Rhat declines to one as $n \to \infty$. The index Rhat is explained in Gelman et al. [27]. In Table 3.7, we can judge that they converge to a stationary distribution because the Rhat of all parameters and latent variables is below 1.1. Incidentally, the condition for λ_{1t} and λ_{2t} to follow a mean-reverting process is satisfied because $\alpha_{11}\beta_{11} + \alpha_{21}\beta_{12} < 0$ on average.

Figure 3.3 shows the fitting results of the five-year CDS spreads of Sony and Fujitsu. Because we aim at fitting the term structure of CDS spreads of one, three, and five years for both Sony and Fujitsu, they cannot be fitted with every detail. However, the levels can be reproduced,



(a) Fitting result for Sony 5Y (b) Fitting result for Fujitsu 5Y

Figure 3.3: Fitting results of the five-year CDS spreads of Sony and Fujitsu. In the left figure, the solid line shows the market CDS spreads of Sony, the dashed line shows the estimated theoretical CDS spreads, and the dotted line shows the estimated hazard rates, respectively. In the right figure, the corresponding time series are plotted for Fujitsu. The x axis represents the number of days between November 20, 2020, and November 30, 2021. The y axis represents the rates expressed in basis point unit.

and the general trend is captured. If two factors or more are applied to the hazard rates, we can obtain better results for fitting the entire term structure of CDS spreads. In the estimated values of one and three years, the levels are also reproduced at the market values. However, the estimated values of five years are the best results for fitting the entire term structure of CDS spreads.

From Figure 3.3, we observe that the path of market CDS spreads has some spikes in time series due to rapid fluctuations (overreaction to the market evaluation of an entity's credit risk). In contrast, the path of theoretical CDS spreads is smooth and cannot catch the spikes on the whole. If the hazard rate process follows only a diffusion process, it might not be possible to reproduce these spikes. If a jump process term is added to the diffusive SDE of hazard rate, this phenomenon might be captured in the model. A simple and viable hazard rate model may be the square-root process with the Poisson jump [19]:

$$\begin{cases} d\lambda_{1t} = \{v_1 + \alpha_{11}(\beta_{11}\lambda_{1t} + \beta_{12}\lambda_{2t})\}dt + \sigma_1\sqrt{\lambda_{1t}}dW_{1t}^P + J_{\lambda_1}dN_{1t}^P \\ d\lambda_{2t} = \{v_2 + \alpha_{21}(\beta_{11}\lambda_{1t} + \beta_{12}\lambda_{2t})\}dt + \sigma_2\sqrt{\lambda_{2t}}dW_{2t}^P + J_{\lambda_2}dN_{2t}^P \end{cases},$$

where $(N_{1t}^P)_{t\geq 0}, (N_{2t}^P)_{t\geq 0}$ are a Poisson jump process under the real probability P whose constant intensities are L_1, L_2, and jump size $J_{\lambda_1}, J_{\lambda_2}$ follow an exponential distribution of the average μ_{1J}, μ_{2J} ($J_{\lambda_1} \sim \mathcal{E}(\mu_{1J}), J_{\lambda_2} \sim \mathcal{E}(\mu_{2J})$). The density of $J_{\lambda_1}, J_{\lambda_2}$ is $f(J) = 1/\mu_J \exp[-J/\mu_J]$. This is a concern for the future.

In Section 3.3.1, we assumed that the behavior between short-term interest rates and default events is independent of each other. Therefore, we empirically validate the assumption of independence by measuring the correlation between the variation of risk-free rate and that of hazard rate. Having checked the 152 correlation coefficients between the daily variations of one-year OIS rate and those of one-year CDS spreads of 19 companies in Table 3.1 in addition to between the daily variations of five-year OIS rate and those of five-year CDS spreads of the same companies in the two periods from November 20, 2020, to November 30, 2021, and from November 19, 2021, to November 30, 2022, the mean absolute value of correlation coefficients

is 0.046, and the standard deviation of absolute value of them is 0.033. Then, the maximum and minimum values of them are 0.125 and 0.001, respectively. Accordingly, it can be judged that the assumption of independence is appropriate.

3.4.2 Pairs trading strategy

In this study, we consider a pairs trading strategy utilizing the cointegration of hazard rates, where we use the results of the analysis for the term structure of CDSs in the previous section. There is no study regarding the pairs trading of CDSs in the literature, although there are many studies about pairs trading that utilize the cointegration of stocks, and they are applied in practice.

Pairs trading is a trading method that makes a profit using the variation of the difference between two financial product prices that are highly correlated (cointegrated). In stocks, the stock prices in the same industry often make similar movements. Therefore, we have two stocks that behave very similarly, and we buy the cheap stock and sell the expensive stock. Then, we execute the cross trade at the timing when the stock prices approach each other and fix profit. More information about pairs trading is found in Vidyamurthy [73].

We explain how to build a pairs trading strategy in our model for corporate CDSs. To this end, when we invest in CDS 1 and CDS 2 by the investment weight based on the cointegration of hazard rates λ_1, λ_2, we need to confirm that the linear sum (spread) in the CDS cointegrated pair has a property of mean reversion. Suppose that we invest in cointegrated CDS pairs, CDS 1 (tenor l_1) and CDS 2 (tenor l_2), with investment weight $(\xi_{1t}^{l_1}, \xi_{2t}^{l_2})$. The contract dollar value $Y_t^{l_1,l_2}$ is written in terms of each contract value $X_t^{i,l_i}, i = 1, 2$:

$$Y_t^{l_1,l_2} = \xi_{1t}^{l_1} X_t^{1,l_1} + \xi_{2t}^{l_2} X_t^{2,l_2}. \tag{3.29}$$

The condition that $Y_t^{l_1,l_2}$ has a property of mean reversion may be equivalently represented under the real measure P as

$$dY_t^{l_1,l_2} = \kappa(\mu - Y_t^{l_1,l_2})dt + \text{martingale}, \tag{3.30}$$

where "martingale" means some martingale process written in, e.g., Brownian motions and Poisson processes; $\kappa(> 0)$ denotes a mean reversion speed of $Y_t^{l_1,l_2}$; and $\mu(> 0)$ denotes a mean reversion level under the real probability P.

Next, we show how to determine the investment weight $(\xi_{1t}^{l_1}, \xi_{2t}^{l_2})$ of the CDS cointegrated pair. The VECM for hazard rates under the risk-neutral probability Q (3.10) is as follows:

$$\begin{cases} d\lambda_{1t} = (s_1 + \gamma_{11}\lambda_{1t} + \gamma_{12}\lambda_{2t})dt + \sigma_1\sqrt{\lambda_{1t}}dW_{1t}^Q \\ d\lambda_{2t} = (s_2 + \gamma_{21}\lambda_{1t} + \gamma_{22}\lambda_{2t})dt + \sigma_2\sqrt{\lambda_{2t}}dW_{2t}^Q \end{cases}. \tag{3.31}$$

Since the model formula (3.6) of CDS spreads depends not only on the hazard rate but also on the interest rate, we must introduce the dynamics of the interest rate to quantify the interest rate risk in addition to the credit risk. We consider a stochastic process of a risk-free rate. To this end, we take a simple and popular interest rate model, that is, the CIR model [13], in which the short rate dynamics under the pricing measure Q are given by

$$dr_t = k_0(\theta_0 - r_t)dt + \sigma_0\sqrt{r_t}dW_{0t}^Q, \tag{3.32}$$

where $k_0(> 0)$ denotes a mean reversion speed of r_t and $\theta_0(> 0)$ denotes its mean reversion level. $\sigma_0(> 0)$ denotes volatility of r_t, and we assume that the Brownian motion under Q; W_{0t}^Q is independent of W_{1t}^Q, W_{2t}^Q. Then, a default-free discount bond price (3.4) is given by

$$D(t,s) = E^Q\left[e^{-\int_t^s r_u du}|\mathcal{F}_t\right] = e^{B_0(\tau) - B_1(\tau)r_t}, \quad \tau = s - t. \tag{3.33}$$

The specification of the interest rate model can be found in Appendix B.4.

We derive the SDEs followed by the dollar value of the CDS contract of each company. From the theoretical formula of CDS spreads (3.6), the market value of the CDS contract (type of receiving premium) can be written as follows:

$$\begin{aligned} X_t^{i,l} &= x^{i,l}(t, r_t, \lambda_{1t}, \lambda_{2t}) \\ &= 1/gC_i \sum_{j=1}^{gl} D(t,t_j) V_f^i(t,t_j) - (1-\theta)\Delta \sum_{k=1}^{\xi l} D(t,s_k) V_p^i(t,s_k), \quad i=1,2, \end{aligned} \tag{3.34}$$

where C_i is the CDS premium of company i. Applying Ito's formula to $X_t^{i,l}$ of (3.34) and substituting (3.31) and (3.32) for it, we obtain

$$\begin{aligned} dX_t^{i,l} &= \sigma_0\sqrt{r_t}x_r^{i,l}dW_{0t}^Q + \sigma_1\sqrt{\lambda_{1t}}x_{\lambda_1}^{i,l}dW_{1t}^Q + \sigma_2\sqrt{\lambda_{2t}}x_{\lambda_2}^{i,l}dW_{2t}^Q \\ &= (\sigma_0\phi_{01}r_t x_r^{i,l} + \sigma_1\phi_{11}\lambda_{1t}x_{\lambda_1}^{i,l} + \sigma_2\phi_{21}\lambda_{2t}x_{\lambda_2}^{i,l})dt \\ &\quad + \sigma_0\sqrt{r_t}x_r^{i,l}dW_{0t}^P + \sigma_1\sqrt{\lambda_{1t}}x_{\lambda_1}^{i,l}dW_{1t}^P + \sigma_2\sqrt{\lambda_{2t}}x_{\lambda_2}^{i,l}dW_{2t}^P, \quad i=1,2. \end{aligned} \tag{3.35}$$

Note that in the RHS of the first line of (3.35), the coefficient of dt is zero due to $X_t^{i,l}$ being a Q-martingale.

Let us introduce a notation of the sensitivity of the contract value:

$$x_\bullet^{i,l}(t, r_t, \lambda_{1t}, \lambda_{2t}) = x_\bullet^{i,l}(t, r, \lambda_1, \lambda_2)|_{t=t, r=r_t, \lambda_1=\lambda_{1t}, \lambda_2=\lambda_{2t}}.$$

We define the market price of risk-free interest rate risk $\psi_{0t} = \phi_{01}\sqrt{r_t}$ as with the case of credit risk. From Ito's calculus, we note the assumption that r and λ_1, λ_2 are independent, and λ_1, λ_2 are also independent of each other. Considering these assumptions, we calculate the duration with respect to a risk-free interest rate:

$$\begin{aligned} x_r^{i,l}(t, r_t, \lambda_{1t}, \lambda_{2t}) &= \frac{\partial x^{i,l}}{\partial r}(t, r_t, \lambda_{1t}, \lambda_{2t}) = 1/gC_i \sum_{j=1}^{gl} \{\frac{\partial}{\partial r} D(t,t_j)\} V_f^i(t,t_j) \\ &- (1-\theta)\Delta \sum_{k=1}^{\xi l} \{\frac{\partial}{\partial r} D(t,s_k)\} V_p^i(t,s_k), \quad i=1,2, \end{aligned}$$

where (3.33) readily yields

$$\frac{\partial D}{\partial r} = -B_1(\tau) D.$$

Next, the durations with respect to hazard rates are calculated as

$$\begin{aligned} x_{\lambda_u}^{i,l}(t, r_t, \lambda_{1t}, \lambda_{2t}) &= \frac{\partial x^{i,l}}{\partial \lambda_u}(t, r_t, \lambda_{1t}, \lambda_{2t}) = 1/gC_i \sum_{j=1}^{gl} D(t,t_j) \frac{\partial}{\partial \lambda_u} V_f^i(t,t_j) \\ &- (1-\theta)\Delta \sum_{k=1}^{\xi l} D(t,s_k) \frac{\partial}{\partial \lambda_u} V_p^i(t,s_k), \quad i=1,2, u=1,2, \end{aligned}$$

where (3.12) and (3.14) readily yield

$$\begin{cases} \dfrac{\partial V_f^i}{\partial \lambda_u} = -A_{iu}(\tau) V_f^i \\ \dfrac{\partial V_p^i}{\partial \lambda_u} = -A_{iu}(\tau) V_p^i + a_{iu}(\tau) V_f^i \end{cases}.$$

Thus, the SDE followed by the dollar value of the CDS cointegrated pair is given by

$$\begin{aligned}
dY_t^{l_1,l_2} &= \xi_{1t}^{l_1} dX_t^{1,l_1} + \xi_{2t}^{l_2} dX_t^{2,l_2} \\
&= \{\sigma_0\phi_{01} r_t(\xi_{1t}^{l_1} x_r^{1,l_1} + \xi_{2t}^{l_2} x_r^{2,l_2}) + \sigma_1\phi_{11}\lambda_{1t}(\xi_{1t}^{l_1} x_{\lambda_1}^{1,l_1} + \xi_{2t}^{l_2} x_{\lambda_1}^{2,l_2}) \\
&\quad + \sigma_2\phi_{21}\lambda_{2t}(\xi_{1t}^{l_1} x_{\lambda_2}^{1,l_1} + \xi_{2t}^{l_2} x_{\lambda_2}^{2,l_2})\}dt \\
&\quad + \sigma_0\sqrt{r_t}(\xi_{1t}^{l_1} x_r^{1,l_1} + \xi_{2t}^{l_2} x_r^{2,l_2})dW_{0t}^P + \sigma_1\sqrt{\lambda_{1t}}(\xi_{1t}^{l_1} x_{\lambda_1}^{1,l_1} + \xi_{2t}^{l_2} x_{\lambda_1}^{2,l_2})dW_{1t}^P \\
&\quad + \sigma_2\sqrt{\lambda_{2t}}(\xi_{1t}^{l_1} x_{\lambda_2}^{1,l_1} + \xi_{2t}^{l_2} x_{\lambda_2}^{2,l_2})dW_{2t}^P
\end{aligned} \tag{3.36}$$

from (3.29) and (3.35). Therefore, the condition (3.30) that the investment weight $(\xi_{1t}^{l_1}, \xi_{2t}^{l_2})$ must satisfy reads

$$\begin{aligned}
&\kappa\{\mu - (\xi_{1t}^{l_1} X_t^{1,l_1} + \xi_{2t}^{l_2} X_t^{2,l_2})\} \\
&\quad = \sigma_0\phi_{01} r_t(\xi_{1t}^{l_1} x_r^{1,l_1} + \xi_{2t}^{l_2} x_r^{2,l_2}) + \sigma_1\phi_{11}\lambda_{1t}(\xi_{1t}^{l_1} x_{\lambda_1}^{1,l_1} + \xi_{2t}^{l_2} x_{\lambda_1}^{2,l_2}) \\
&\qquad + \sigma_2\phi_{21}\lambda_{2t}(\xi_{1t}^{l_1} x_{\lambda_2}^{1,l_1} + \xi_{2t}^{l_2} x_{\lambda_2}^{2,l_2})
\end{aligned} \tag{3.37}$$

by comparing with (3.36).

We have the investment weight so that a CDS cointegrated pair has a mean reversion. However, it is difficult to satisfy (3.37) because the degree of freedom of the investment weight are low in the case of two CDSs. Therefore, in general, we can only say that the CDSs can be described using the hazard rates which have cointegration. If the behavior of CDSs is similar to that of hazard rates, the CDS pair may have a mean reversion. This will be explained later. Accordingly, in this study, we conduct the numerical calculation and ask for the cointegration vector without asking for $(\xi_{1t}^{l_1}, \xi_{2t}^{l_2})$ as follows:

$$\begin{aligned}
Y_t^{l_1,l_2} &= \xi_{1t}^{l_1} X_t^{1,l_1} + \xi_{2t}^{l_2} X_t^{2,l_2} \\
&\propto X_t^{1,l_1} + \frac{\xi_{2t}^{l_2}}{\xi_{1t}^{l_1}} X_t^{2,l_2} = X_t^{1,l_1} - \eta X_t^{2,l_2},
\end{aligned} \tag{3.38}$$

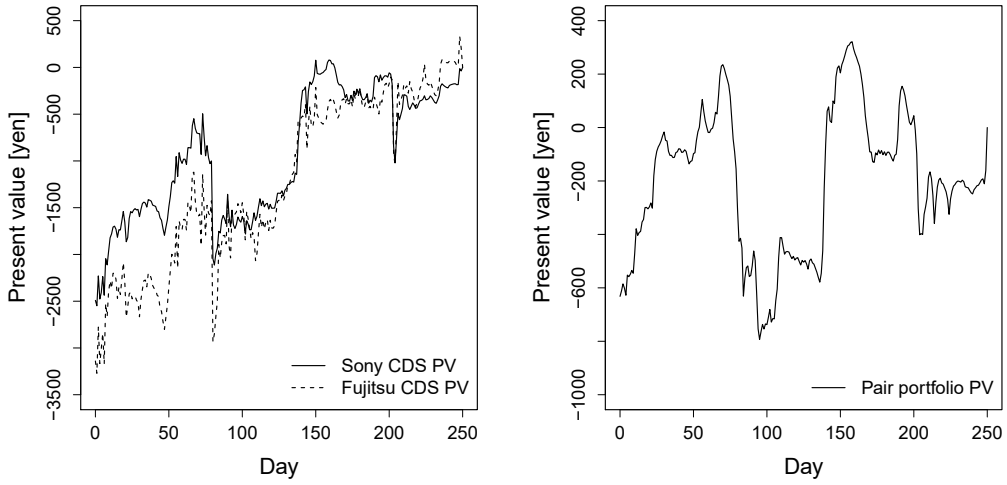
where we assume $\eta = -\xi_{2t}^{l_2}/\xi_{1t}^{l_1}$. Calculating η by regressing $\Delta X_t^{1,l_1}$ over $\Delta X_t^{2,l_2}$ which are the daily fluctuations of the value process of CDS, we conduct a pair trading with the cointegration vector $(1, -\eta)$.

Using the five-year CDSs of Sony and Fujitsu analyzed in Section 3.4.1, we construct the pairs trading strategy. As we have discussed so far, the cointegration investment to a CDS pair is realized when we sell (buy) η units of the CDS of Fujitsu ($i = 2$) if $\eta > 0$ or buy (sell) $|\eta|$ units of it if $\eta < 0$ against buying (selling) one unit of the CDS of Sony ($i = 1$). The value processes of CDS can be calculated using $V_f^i(t_x, t_j), j = 1, \cdots, gl, x = 0, \cdots, \Lambda$ and $V_p^i(t_x, s_k), k = 1, \cdots, \xi l, x = 0, \cdots, \Lambda$ obtained in Section 3.4.1 as follows:

$$\begin{aligned}
&X_{t_x}^{i,l} = N\{1/gC_i \sum_{j=1}^{gl} D(t_x, t_j) V_f^i(t_x, t_j) - (1-\theta)\Delta \sum_{k=1}^{\xi l} D(t_x, s_k) V_p^i(t_x, s_k)\}, \\
&\quad x = 0, \cdots, \Lambda, i = 1, 2,
\end{aligned} \tag{3.39}$$

where the CDS premiums C_i are set so that the present values of each CDS are zero on November 30, 2021, which is the end date of the observation period. The premium of Sony is 20.76bp, and that of Fujitsu is 24.37bp. To make the unit of counting easy to understand, we set the notional amount to one million yen.

In the above setting, calculating $\Delta X_{t_x}^{i,l_i} = X_{t_x}^{i,l_i} - X_{t_{x-1}}^{i,l_i}, i = 1, 2$, the coefficient, which is calculated by regressing $\Delta X_{t_x}^{1,l_1}$ over $\Delta X_{t_x}^{2,l_2}$, is $\eta = 0.5923$. Figure 3.4 shows the results of



(a) CDS value process

(b) Pair portfolio value process

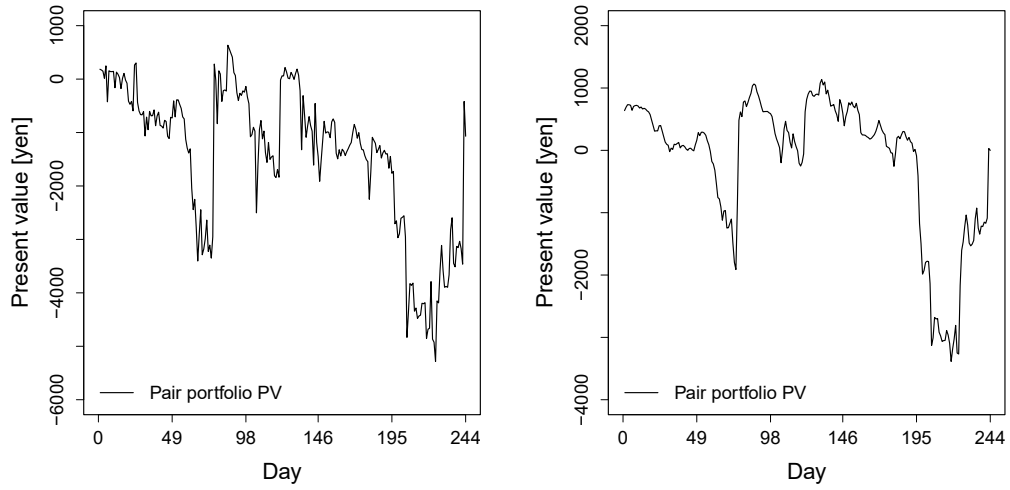
Figure 3.4: Estimation results of CDS value processes and pair portfolio value process for Sony and Fujitsu. In the left figure, the solid line shows the present value of Sony CDS, and the dashed line shows that of Fujitsu CDS, respectively. In the right figure, the solid line shows the present value of the pair portfolio, which is Sony CDS - 0.5923 $\times$ Fujitsu CDS. The x axis represents the number of days between November 20, 2020, and November 30, 2021. The y axis represents the present values denominated in yen.

our numerical calculation of the CDS value processes and pair portfolio value process for Sony and Fujitsu calculated from the last equal sign of (3.38) and (3.39). The left figure shows the present values of Sony CDS and Fujitsu CDS. Both appear to be in a similar transition. Therefore, under the pairs trading strategy, we construct the pair portfolio of the investment weight $(1, -0.5923)$, which is Sony CDS - 0.5923 $\times$ Fujitsu CDS. The right figure shows the present value of this pair portfolio. Although the present value can sometimes grow, it seems to be moving at a constant value on average. Accordingly, when a spread larger than the average is observed, we take a long position of Fujitsu CDS at the same time as taking a short position of Sony CDS. Then, when the spread returns to its original level, we unwind a trade by selling Fujitsu CDS at the same time as buying back Sony CDS. As a result, the position is liquidated by the cross trade, and it is determined as a positive profit to an investor. Notably, the cointegration vector (1,-0.5923) calculated from the estimated values of value processes of Sony CDS and Fujitsu CDS is close to the cointegration vector (1,-0.7167) of market CDS spreads of Sony and Fujitsu calculated in Section 3.2.2. Accordingly, we can say that this model works well and that the term structure of CDSs can be reproduced well. The coefficient is calculated from each martingale part, and it can be theoretically calculated by the formula:

$$\hat{\eta} = \frac{\mathrm{Cov}[\Delta X_t^{1,l_1}, \Delta X_t^{2,l_2}]}{\mathrm{Var}[\Delta X_t^{2,l_2}]}, \tag{3.40}$$

where $\mathrm{Var}[A]$ represents a variance of A, and $\mathrm{Cov}[A, B]$ represents a covariance of A and B. That is, we can calculate it by dividing the quadratic covariation $d\langle X_t^{1,l_1}, X_t^{2,l_2}\rangle$ by the quadratic variation $d\langle X_t^{2,l_2}, X_t^{2,l_2}\rangle$. We can calculate $\hat{\eta} = 0.5899$ using (3.40). This calculation result is about equal to the value of the coefficient $\eta = 0.5923$ calculated from the estimated values of value processes of Sony CDS and Fujitsu CDS.

We conduct the performance verification (backtesting) of the model estimated using the data between November 20, 2020, and November 30, 2021. The purpose of backtesting is to check to what extent the constructed pair realizes the mean reversion in the observation period between December 1, 2021, and November 30, 2022, as out-of-sample testing.



(a) Out-of-sample test (b) In-sample test

Figure 3.5: Result of out-of-sample testing (backtesting) and in-sample testing for the pairs trading of Sony and Fujitsu. In the left figure, the solid line shows the result of out-of-sample testing with $\eta = 0.5923$. In the right figure, the solid line shows the result of in-sample testing with $\eta = 0.6758$. The x axis represents the number of days between December 1, 2021, and November 30, 2022. The y axis represents the present values denominated in yen.

The backtesting is designed as follows. Solving the theoretical CDS spread formula of (3.18) for the protection leg and substituting it for (3.39), the CDS present value in the observation period of backtesting $x = \Lambda + 1, \dots$ can be written using the CDS spreads on the parameter estimation date and backtesting date as follows:

$$\begin{aligned} X_{t_x}^{i,l} &= (C_i - CDS_{t_x}^{mod,i,l})N/g \sum_{j=1}^{gl} D(t_x, t_j) V_f^i(t_x, t_j) \\ &= (C_i - CDS_{t_x}^{mod,i,l}) Z_{t_x}^{i,l}, \quad x = \Lambda + 1, \dots \end{aligned} \tag{3.41}$$

where $Z_{t_x}^{i,l}$ represents the fixed leg value. Substituting (3.41) for (3.38), we obtain the present value of the pair portfolio as follows:

$$Y_{t_x}^{l_1,l_2} \propto (C_1 - CDS_{t_x}^{mod,1,l_1}) Z_{t_x}^{1,l_1} - \eta (C_2 - CDS_{t_x}^{mod,2,l_2}) Z_{t_x}^{2,l_2}, \quad x = \Lambda + 1, \dots \tag{3.42}$$

where $Z_{t_x}^{1,l_1}, Z_{t_x}^{2,l_2}, x = \Lambda + 1, \dots$ are estimated by applying the estimation method in Section 3.3.2 to the data in the observation period. Originally, the fixed leg values should be estimated from the data before the backtesting date, but this is approximately done due to computational load.

On the left side of Figure 3.5, the result of out-of-sample backtesting is shown using the investment weight $\eta = 0.5923$ estimated in the above for the five-years CDS pair of Sony and Fujitsu using (3.42). For comparison, on the right side of Figure 3.5, the result of in-sample testing is shown using the investment weight $\eta = 0.6758$ estimated from the data in the observation period. For out-of-sample testing, we verify the performance, which is whether the CDS pair realizes a mean reversion, between December 1, 2021, and November 30, 2022, using the investment weight $\eta = 0.5923$ estimated from the market CDS spreads between November 20, 2020, and November 30, 2021 (prior period). In contrast, for in-sample testing, we verify it between December 1, 2021, and November 30, 2022, using the investment weight $\eta = 0.6758$ estimated from the CDS spreads between November 19, 2021, and November 30, 2022 (same period).

The result of out-of-sample backtesting shows that the mean reversion is produced to some extent in the posterior one year observation period by using the prior estimated investment weight. The graph of out-of-sample testing is similar to that of in-sample testing. Therefore, our pairs trading strategy seems to work well. It should be noted that both of out-of-sample and in-sample testings are completely different. The difference is the investment weight. While the investment weight of out-of-sample testing is 0.5923, that of in-sample testing is 0.6758. Although the update frequency of investment weight is one year, it should be remarkable that the investment weights estimated from the one year periods before and after November 30, 2021, are not significantly different. It can be said because there are no significant changes in the property of cointegration of Sony CDS and Fujitsu CDS between the one year periods before and after November 30, 2021.

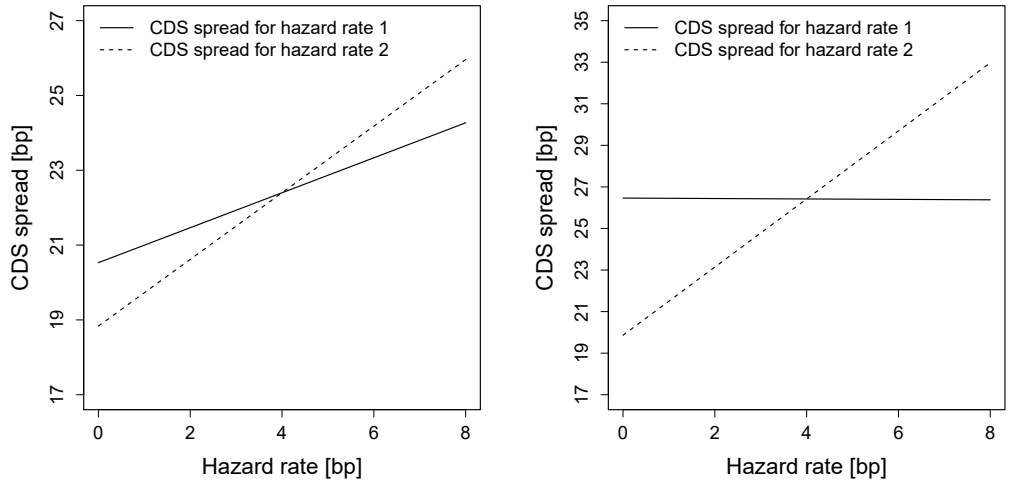
From this backtesting (out-of-sample testing), there is a possibility that we do not obtain enough performance (mean reversion) since the model has room for refinement, and the performance of testing depends on the update frequency of investment weight. In the future, stronger mean reversion will be realized in pairs trading by refining the model and increasing the update frequency (monthly, biweekly, or weekly etc.) with the increase of computational efficiency. Furthermore, the improvement of backtesting has also room for consideration. The validation of operational performance of the CDS pairs trading should be considered for in-sample and out-of-sample testings based on the previous studies for that of the stock pairs trading. In the case of CDSs but not stocks, there are various quote methods in the market such as a CDS spread and an up-front fee etc. However, as the target of validation is a CDS value process, it is necessary to construct the validation method taking the matters specific to CDSs into account. It is the future problem.

We explain why we assume the VECM for hazard rate processes (3.1) in this study. The theoretical formula of CDS spreads (3.6) is a non-linear function of fraction type for hazard rates. Accordingly, even if the hazard rates follow the ECM, the form of the ECM does not directly appear in the theoretical formula of CDS spreads. Furthermore, since the theoretical formula of CDS spreads must be calculated under the risk-neutral probability Q, the ECM nature of hazard rates will disappear as soon as we change the measure. Nonetheless, as demonstrated in Section 3.4.1, we can fit the theoretical CDS spread formula to the contigrated market CDS spreads by conducting the Bayesian inference in such a way that the former approaches the latter. We consider that two theoretical CDS spreads can be fitted to the market values in the real market because the converted CDS spreads fairly maintain the ECM nature.

We partially understand why the ECM nature is maintained after we change the measure and convert it to the non-linear fraction from Figure 3.6. This figure plots the variation in CDS spreads for the variation in hazard rates using the theoretical formula of five-year CDS spreads (3.43) to which the parameters estimated in Table 3.7 are applied. Figure 3.6 indicates that the relationship between the CDS spread and the hazard rates maintain drastic linearity. This is partly because the linear formula of $\lambda_{1t}, \lambda_{2t}$ included in the protection leg (numerator) $a_{i0}(\tau) + a_{i1}(\tau)\lambda_{1t} + a_{i2}(\tau)\lambda_{2t}$ is dominant compared with the discounting factors with respect to the hazard rates, the effect of which could be made offset between the numerator and denominator:

$$
\begin{aligned}
&CDS_t^{mod,i,l} \\
&= \frac{(1-\theta)\Delta \sum_{k=1}^{\xi l} D(t,s_k) e^{A_{i0}(\tau_k) - A_{i1}(\tau_k)\lambda_{1t} - A_{i2}(\tau_k)\lambda_{2t}} \{a_{i0}(\tau_k) + a_{i1}(\tau_k)\lambda_{1t} + a_{i2}(\tau_k)\lambda_{2t}\}}{1/g \sum_{j=1}^{gl} D(t,t_j) e^{A_{i0}(\nu_j) - A_{i1}(\nu_j)\lambda_{1t} - A_{i2}(\nu_j)\lambda_{2t}}}, \quad (3.43) \\
&\nu_j = t_j - t, \tau_k = s_k - t, i = 1,2, l = 5.
\end{aligned}
$$

This is the reason why the cointegration relationship assumed in our hazard rate model is also preserved in the CDS spread. This consideration reinforces our empirical evidence that the



(a) Linearity for Sony (b) Linearity for Fujitsu

Figure 3.6: Linearity of hazard rates and CDS spread for Sony and Fujitsu. In both figures, the solid line shows the CDS spread value for the hazard rate λ_1, and the dashed line shows that for the hazard rate λ_2, respectively. In both figures, the center of x axis is 4 bp, which is the average of estimated latent processes.

cointegration vector for market CDS spreads and that for estimated dollar value processes is fairly close.

Our study is mainly focused on static trading for a cointegrated pair. This can be extended to the portfolio management of multiple cointegrated pairs as in Higashide et al. [33] and Roncalli [63], and also to the dynamic trading of those pairs as in Tourin and Yan [70].

Risk parity is an asset management method that reduces risk by distributedly holding each asset occupying the portfolio in such a way that the percentage of the risk of each asset is even. This is the method widely utilized by foreign institutional investors, including pension funds, in Europe and the United States. Since we align the different asset holding risks, including stocks, bonds, and commodities, we change and adjust the weighting of each asset each time according to the movement of the volatility in each market. Risk parity is detailed in Roncalli [63]. In the case of the pairs trading in this study, we can suggest a method to distribute risk evenly to each pair since we can express the risk of each pair and the overall risk using the duration.

3.5 Conclusion

This study develops the ODE-based Bayesian inference of cointegration dynamics for hazard rates. In the empirical study of Japanese cointegrated corporate CDSs, we have demonstrated that our statistical method is workable for estimating parameters and latent processes in the cointegrated hazard rate model.

The contributions of this study are as follows: first, a simple fundamental model for the cointegrated hazard rates can create a diversity of cointegration relationships between various combinations of relevant CDS spreads with different maturities; second, our cointegrated hazard model can be applied to the investment strategy of CDSs.

As for the first point, we can effectively merge the cointegration structure into the framework of arbitrage-free pricing and calculate the theoretical spread of CDSs, so that we clarify why the cointegration appears among CDSs with different maturities. Furthermore, we understand how the cointegration among the term structures of CDSs occurs through hazard rates. As for the second point, by applying our model to the cointegrated CDS spread data of Sony and

Fujitsu, we can obtain the investment weight of each individual CDS.

It is the main idea in this study that we have analyzed the cointegration of CDS spreads by focusing on that of hazard rates. As we use the model that a CDS spread is almost determined by hazard rates, the cointegration between CDS spreads naturally occurs from that between hazard rates. Exactly by utilizing this property of model, we give the structure (rule) how the cointegration of CDS spreads is determined from that of hazard rates.

We explain the image of this rule a little more intuitively by contrasting with natural science. We describe the relationship such as "periodic table of elements" by comparing cointegrated multiple hazard rates to "elementary particles (proton and electron)" and a cointegrated CDS spread consisting of multiple hazard rates to "element". The CDS spread (element) identified by a reference company and a maturity can be described by the combination of a different number of hazard rates (protons and electrons) that compose it through the theoretical CDS spread formula. This study theoretically formulates this relationship with respect to n CDSs in general. Therefore, the dynamics of cointegrated CDS spreads of all maturities can be explained based on the property of cointegrated hazard rates.

The theoretical CDS spread formula shows the following property: the shorter the maturities of CDSs are, the stronger the cointegration is; it becomes weaker as the maturities are longer. It should be theoretically observed that the strength of cointegration between the CDS spreads (bond yields) with a far expiry date tends to become weak. However, having investigated the cointegration rank between companies in the same industry for the industry sectors in Table 3.1 (See Table 3.4 for the electricity sector), this relationship does not always hold. This point should be clarified, for example, whether there is the case that the CDS spreads are not cointegrated although the cointegration of hazard rates is strong.

There are several concluding remarks. For a more precise analysis of the term structure of CDS spreads, we may take into account the correlation in diffusion terms of hazard rate processes and in some cases, may introduce jump terms. We may assume two factors or more for the hazard rates. As an extension to the case of three or more cointegrated CDSs, the estimation and analysis of trading strategies should be performed separately for each rank of cointegration. In particular, as the estimation of three or more cointegrated CDSs is computationally expensive, we need to develop more efficient algorithms. For a dynamic pairs trading strategy, we can consider the derivation of an investment strategy using the HJB equation as in Tourin and Yan [70]. For the optimal investment problem of pairs portfolio, the first hitting time of the spread process of cointegrated pairs should be studied, as in Higashide et al. [33]. From the standpoint of expanding the scope of application of this analysis, further studies should be conducted, for instance, between bonds, or between bond and CDS. We might apply this model to the analysis of the cointegration between CDS and bond which is the subject of Blanco et al. [7] and Guidolin et al. [29]. Incidentally, it is interesting to employ the cointegrated hazard rate model to measure the default probability and credit risk.

Furthermore, there are some theoretical issues. First, we deal with the dependence of the credit risks of multiple companies. In the case of two companies, we consider the model in which the theoretical CDS spread of one company must be calculated using the hazard rates of the other company in a pair. Therefore, we will formulate the theoretical CDS spread where the case that the company in a pair has defaulted during the contract period is taken into account. If the company in a pair has defaulted during the contract period, the error correction terms disappear. Accordingly, it is necessary to calculate the CDS spread taking into account how the structure of the SDE of hazard rates changes after default. When considering the pairs trading of CDSs, this becomes the investment problem that has three stopping times of the first passage time from widening the spread of a pair to narrowing it, mentioned in Higashide et al. [33], and the default times of both companies.

Next, it should be also proved that the CDS spreads of multiple companies can be related

linearly, that is, the cointegration between multiple CDSs. In this study, the condition that an investment weight should satisfy for a CDS pair to have a mean reversion is derived as the condition that the drift of the SDE followed by the value of a CDS pair should satisfy. However, the investment weight that satisfies this condition cannot be identified in the degree of freedom of two CDSs. It should be considered whether we can theoretically derive the investment weight when increasing the number of cointegrated CDSs or under the assumption that a CDS spread and hazard rates have linearity because it is considered that a mean reversion appears when some approximation holds that the dynamics of CDS spread is close to that of hazard rate. The assumption should be specifically formulated.

For the pairs trading of CDSs, we will study the specific techniques in overall strategy including the start and end timing of position. The consistent investment weight of CDSs has been derived based on hazard rates in this study. Since this consistent investment weight is different from the investment weight (cointegrated vector) calculated directly from CDS spreads, we will analyze the strategic differences resulting from it. In the practical aspect of CDS pairs trading, it is necessary to prepare the margin practices on a central clearing counterparty (CCP). When conducting the pairs trading of CDSs through CCPs, we have positive or negative exposure as the value of a CDS pair reverts to the mean. Therefore, the party that has a loss due to daily variations deposits the loss amount as a variation margin into the CCP. Actually, as the exposure is reduced by netting, the pairs trading is positioned as two transactions included in the portfolio of many transactions unless there are special practices as the pairs trading of CDSs. The initial margin practices are also required. We believe that these issues require further research.

Chapter 4

PDE-based Bayesian Inference of CEV Dynamics for Credit Risk in Stock Prices

4.1 Introduction

The measurement and analysis of credit risk have been as important as those of market risk and liquidity risk among the best practices of financial institutions. The US subprime mortgage crisis of 2007 triggered the global financial crisis of 2008, which had a negative impact on the global credit market and equity market leading to a decline in global stock prices. Consequently, strict regulations and market practices have been imposed on financial institutions in the following years. Therefore, the measurement of credit risk is essential for the management of financial institutions.

Although credit risk losses are rare, these can have a large negative impact and cause modeling difficulties because of the lack of empirical data. There are two conceptually different approaches in credit risk modeling. They are the structural approach (Merton [53]) and the reduced-form approach (Duffie and Singleton [21]). In the structural approach, we model the occurrence of the default event explicitly to the financial structure of the company based on asset, liability, and equity. A default occurs when the asset value goes below the liability value (i.e., the equity value goes below zero). In the reduced-form approach, we give the distribution of default time exogenously and regard the default of company as an event caused by an exogenous shock. Specifically, the credit event of reference entity is defined as a jump only once, and a hazard rate is defined as the intensity process of the jump under the conditions that a default does not occur by that time.

The Merton model [53] is a well-known structural approach. In the Merton model, the asset value of a company follows a geometric Brownian motion, and the financial structure of a company consists of an interest-free debt and a common stock. Then, we regard the stock value as the call option of asset values. Moody's KMV [16] and CreditMetrics [30] are the other famous structural credit risk models, which are practical. We can obtain persuasive interpretations about the credit risk of company owing to the structural approach. The studies in finance have focused extensively on estimating the parameters of the Merton model because the financial structure is simple and the equity value is described as a closed form. However, its implementation is complicated in the Merton model. Jarrow and Turnbull [40] indicate that it is difficult to appropriately estimate the parameters (expected return and volatility) of the Merton model because the asset values are unobservable while the equity values are observable. It is believed that the unobservability of asset values is a major drawback in the structural credit risk model.

Many models for the inference of parameters have been studied in the structural approach. Duan and Fulop [17] develop a particle filter-based maximum likelihood estimation method for the structural credit risk model of Merton. They use the state space model in which the system equation is the geometric Brownian motion for asset values and the observation equation is the BS model [6] for stock values and conduct the measurement of credit risk by the Merton model using the maximum likelihood estimation combined with the sampling importance resampling algorithm of the particle filter [28]. The resampling is conducted using the smooth bootstrap procedure of Pitt [62] because the likelihood function evaluated via a particle filter is discontinuous with respect to the parameters. Therefore, it is fitted to the gradient-based optimization. Duan and Fulop [17] estimate the parameters of the model using this method on a condition that asset values are unobservable and empirically analyze the stock data from the 30 Dow Jones firms and 100 firms randomly selected.

The method of Duan and Fulop [17] solves the problem of the unobservable asset values. Since their method is based on the Merton model, it can analytically describe the observation equation as the BS model and estimate the parameters. However, if the system equation becomes more difficult, the parameters cannot be estimated using the maximum likelihood estimation. For example, if the system equation is the CEV model [11] [12], which is the developed form of the Merton model, the option value, which represents the observation equation, cannot be described as an analytical form. Therfore, the parameters cannot be estimated.

To solve these problems, this study suggests the method to estimate the parameters in the CEV model using the MCMC method combined with PDEs. The reason for using the CEV model is to increase the degree of freedom of model selection by generalizing the BS model and making our model fitted better into the observation data. The range of the elastic constant is expanded from 1 to $[0, 1]$. Beckers [4] shows that the CEV model could give a better description of the actual stock price behavior than the BS model used traditionally while discussing an inverse relationship between the level of the stock price and its variance of return that fits a practical sense and can be expressed by the CEV model. This study describes the observation equation as the curve of option values obtained by solving PDEs. The FDM is used in solving PDEs. We measure corporate credit risks by applying the this method to the stock data.

In cases where the observation equation can be described as an analytical formula, for example, the BS model of stocks, the Vaciek model [71], and the CIR model [13] of interest rates, it is easy to use the MCMC. However, when the observation equation cannot be described as an analytical formula, it is difficult to use the MCMC, for example, the CEV model and the stochastic volatility model. This study infers the distribution of parameters using the PDE-based MCMC in the state space model, so that we become able to calculate the parameters of the model in which the observation equation cannot be described as an analytical formula, and the parameters cannot be calculated.

The numerical calculation method (estimation method) can be roughly divided into two categories: the Bayesian inference and non-Bayes methods. The Bayesian inference uses the MCMC based on Bayesian statistics and statistical physics. This study uses the MCMC consistently as the numerical calculation method. In contrast, non-Bayes methods are based on traditional statistics. The difference between Duan and Fulop's [17] method and our method is that their method is a non-Bayes method that is the combination of the particle filter and the maximum likelihood estimation, while our method is a Bayesian estimation.

Jones [43] uses the MCMC to estimate the parameters of the CEV model. His study involves the modeling of short rate process, which has a non-linear drift and a CEV diffusion. He uses one-week euro-dollar rate as the proxy variable for the short rate and estimates the parameters of the model by combining the Euler approximation, the Gibbs sampler, and the MH algorithm. However, he do not use a state space model and therefore a pricing model as the observation equation. In contrast, Aboulaich et al. [1] apply the FDM to the pricing of derivative when

the underlying asset price follows a CEV process. They conduct the pricing of european call option, which has the diffusion term of stochastic volatility following the CEV model. They implement the pricing model using the FDM of PDEs for some values of the elastic constant. However, they do not conduct the estimation of parameters. A parameter estimation is in the problem at a higher level than a derivative pricing. The studies that include the pricing of derivative in the parameter estimation problem are greatly limited.

A method of the Bayesian inference combined with the FDM of PDEs like our method has not been suggested in the existing literature. We emphasize that our study works at a problem at a higher level than the above two studies because we estimate the parameters and latent variables of the system equation (dealt with in Jones [43]) and conduct the pricing of derivative consistent with the framework of arbitrage-free pricing in the observation equation (dealt with in Aboulaich et al. [1]) simultaneously by using a state space model. In that sense, our model is distinguished from the models of a mere parameter estimation for stochastic process and the models for mere pricing of derivatives.

Our model consists of two parts: the state space model and the algorithm to numerically calculate parameter distributions. In the state space model, the system equation is described as the process of assets in the CEV model, and the observation equation is described as the option value curve obtained by solving the PDEs of stocks using the FDM. Thus, in the structural approach, we extend a state space model to the model where the observation equation cannot be analytically described. In the algorithm to numerically calculate parameter distributions, we solve the PDE, which is the observation equation, using the FDM and estimate the parameters and the latent process of asset values using the MCMC. As an empirical analysis, we estimate the parameters of the CEV model and the latent process of asset values using the real stock prices of US financial institutions in the period of the subprime mortgage crisis and COVID-19. Furthermore, we analyze the default probability and measure the credit risk of bank portfolios while comparing the CEV model with the Merton model.

The organization of this paper is as follows: Section 4.2 explains the PDE-based Bayesian inference of CEV dynamics. Specifically, the derivation of this model and the algorithm to numerically calculate parameter distributions are explained. Section 4.3 is devoted to the numerical calculation using this model and the analysis of the results. Furthermore, we measure the credit risk of bank portfolios. Section 4.4 describes the extension of our model. Section 4.5 concludes our studies and mentions some concluding remarks.

4.2 The model

In this section, we derive the PDE-based Bayesian inference of CEV dynamics. The model consists of two parts: the state space model to describe stock values through asset values in the structural approach and the algorithm to numerically calculate parameter distributions in the case that stock values cannot be described analytically.

4.2.1 Modeling for CEV dynamics

General model

Our model follows the structural approach of Merton [53]. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let $W_t^P, 0 \leq t \leq T$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{F}_t, 0 \leq t \leq T$ be a market filtration for this Brownian motion. The real and risk-neutral probabilities are denoted by P and Q, respectively. We consider a state space model, where market stock values are taken to be observables, and asset values are treated as continuous latent processes. First, we provide the system equation (state equation) for latent asset values.

The drift term of asset value processes is the same form as the Merton model, and the diffusion term has the structure of the CEV model to increase the degree of freedom of model selection, where an asset process is not negative value or zero. We write the system equation for latent asset values as

$$dA_t = \mu A_t dt + \sigma A_t^{\gamma} dW_t^P, \tag{4.1}$$

where the drift of asset values is denoted by $\mu \in \mathbb{R}$ under the real probability P; the volatility of asset values is denoted by $\sigma > 0$; and the elastic constant of the CEV model is denoted by $0 < \gamma < 1$. Let us indicate an asset value as A_t. With $\gamma = 0$, this model is the Vaciek model, while with $\gamma = 0.5$, it is the CIR model. With $\gamma = 1$, it is the BS model. A detailed investigation of this class of models is found in Chan et al. [10].

Next, we provide the observation equation. We derive the theoretical formula of stock values for the asset value model with the CEV term of (4.1). In the structural approach, the stock value is calculated as the call option on the asset value under the martingale measure Q. The call option has a payoff $\max[A_T - D, 0]$ at the maturity T, where D represents a liability value (strike of call option). Accordingly, we have the theoretical formula for stock values and the observation equation as follows:

$$\begin{cases} S_t^{mkt} = S^{mod}(t, A_t) + e_t \\ S^{mod}(t, A_t) = E^Q \left[e^{-r(T-t)} \max[A_T - D, 0] \mid \mathcal{F}_t \right] \end{cases}, \tag{4.2}$$

where S_t^{mkt} is a market stock value; $S^{mod}(t, A_t)$ is the same theoretical stock value; e_t is an observation error. A risk-free rate is denoted by r that is assumed to be constant, and $E^Q[\bullet]$ represents an expectation under the risk-neutral probability measure Q. The observation period is described by $T - t$.

The system equation (4.1) and observation equation (4.2) compose a state space model. Since the theoretical stock value must be calculated under the risk-neutral measure Q as described in (4.2), the Girsanov theorem is used to define the risk-neutral measure Q following Shreve [67]. Let us assume the market prices of risk as

$$\psi_t = \phi A_t^{-\gamma+1}. \tag{4.3}$$

This slightly general constrained functional form of the market price of risk (4.3) guarantees that the form of the CEV process (4.1) is unchanged after a Girsanov transformation defined in (4.4). In order to change the probability measure using the Girsanov theorem, we define the Radon-Nikodym derivative of the equivalent martingale measure Q with respect to P denoting theoretically the maximum time horizon as T as follows:

$$\begin{cases} Z_T = \dfrac{dQ}{dP} = \exp\left[-\displaystyle\int_0^T \psi_u dW_u^P - \frac{1}{2} \int_0^T \psi_u^2 du \right] \\ W_t^Q = W_t^P + \displaystyle\int_0^t \psi_u du \end{cases}. \tag{4.4}$$

In $0 \le t \le T$, the Radon-Nikodym derivative process is $Z_t = E^P[Z_T | \mathcal{F}_t]$. Then, $W_t^Q, 0 \le t \le T$ is a Brownian motion under the probability measure Q.

Accordingly, the SDE of asset values represented in Q is written as

$$dA_t = \mu A_t dt + \sigma A_t^{\gamma}(dW_t^Q - \psi_t dt) = r A_t dt + \sigma A_t^{\gamma} dW_t^Q, \tag{4.5}$$

where the following must hold in order to meet the no arbitrage conditions:

$$\mu A_t - \psi_t \sigma A_t^{\gamma} = (\mu - \phi\sigma) A_t = r A_t \quad \Longleftrightarrow \quad \phi = \frac{\mu - r}{\sigma}.$$

The key expectations of the stock value formula (4.2) are the solution $V(t,x)$ of the PDE:

$$\frac{\partial V}{\partial t} + rx\frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2 x^{2\gamma}\frac{\partial^2 V}{\partial x^2} = rV \tag{4.6}$$

derived from no-arbitrage arguments, where the corresponding boundary condition is provided as

$$V(T,x) = \max[x - D, 0]. \tag{4.7}$$

For convenience, hereafter, this model derived as the general model in this section is referred to as the CEV model.

Analytical model

In the case of $\gamma = 1$, the PDE (4.6) and corresponding boundary condition (4.7) reduce the analytical formula of the BS model in a standard manner. Then, substituting $\gamma = 1$ for equations (4.1) and (4.5), the state space model is described as

$$\begin{cases} dA_t = \mu A_t dt + \sigma A_t dW_t^P = rA_t dt + \sigma A_t dW_t^Q \\ S_t^{mkt} = S^{mod}(t, A_t) + e_t \\ S^{mod}(t, A_t) = A_t\Phi(d_1) - De^{-r(T-t)}\Phi(d_2) \end{cases}, \tag{4.8}$$

where

$$d_1 = \frac{\ln(A_t/D) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = \frac{\ln(A_t/D) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \tag{4.9}$$

and $\Phi(\bullet)$ is a distribution function of standard normal distribution. The market price of risk is described as

$$\psi_t = \phi.$$

This analytical models will be used to compare with the CEV model (general model derived in the previous section) in Section 4.3. For convenience, hereafter, this model derived as the analytical model in this section is referred to as the Merton model.

In the case of $\gamma = 0$ and $\gamma = 0.5$, the analytical formulae can be derived as two special cases of the CEV class as discussed in Beckers [4]. Cox [11] derives an option pricing formula that holds if the stock price follows a CEV diffusion and obtains the following call option price of the CEV model:

$$\begin{aligned} S^{mod}(t, A_t) = A_t \sum_{n=0}^{\infty} g(n+1, x) G\left(n + 1 + \frac{1}{2(1-\gamma)}, kD^{2(1-\gamma)}\right) \\ - De^{-r(T-t)} \sum_{n=0}^{\infty} g\left(n + 1 + \frac{1}{2(1-\gamma)}, x\right) G(n+1, kD^{2(1-\gamma)}), \end{aligned} \tag{4.10}$$

where $g(m,v) = e^{-v}v^{m-1}/\Gamma(m)$ is a gamma density function ($\Gamma(m)$ is a gamma function); $G(m,v)$ is a complimentary gamma distribution;

$$k = \frac{2r}{2\sigma^2(1-\gamma)(e^{2r(1-\gamma)(T-t)} - 1)} \quad \text{and} \quad x = kA_t^{2(1-\gamma)}e^{2r(1-\gamma)(T-t)}. \tag{4.11}$$

This model depends on only five data inputs: the asset value A_t, liability value D, risk-free rate r, asset volatirity σ, and time to maturity $T - t$.

Cox and Ross [12] derive the following call option price, which is the observation equation, of $\gamma = 0$:

$$S^{mod}(t, A_t) = (A_t - De^{-r(T-t)})\Phi(y_1) + (A_t + De^{-r(T-t)})\Phi(y_2) + v\{\phi(y_1) - \phi(y_2)\}, \tag{4.12}$$

where $\phi(\bullet)$ is a density function of standard normal distribution,

$$v = \sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}, \quad y_1 = \frac{A_t - De^{-r(T-t)}}{v}, \quad \text{and} \quad y_2 = \frac{-A_t - De^{-r(T-t)}}{v},$$

which is a special case of (4.10) and (4.11). The system equation is written as

$$dA_t = \mu A_t dt + \sigma dW_t^P = rA_t dt + \sigma dW_t^Q, \tag{4.13}$$

where (4.12) and (4.13) compose the state space model of $\gamma = 0$ (hereafter referred to the absolute model).

Similarly, Cox [11] develop an approximative formula for the call option price, which is the observation equation, of $\gamma = 0.5$:

$$S^{mod}(t, A_t) = A_t\Phi(q(4)) - De^{-r(T-t)}\Phi(q(0)), \tag{4.14}$$

where

$$q(w) = \frac{1 + h(h-1)\frac{w+2y}{(w+y)^2} - h(h-1)(2-h)(1-3h)\frac{(w+2y)^2}{2(w+y)^4} - \left(\frac{z}{w+y}\right)^h}{\sqrt{2h^2\frac{w+2y}{(w+y)^2}\left(1-(1-h)(1-3h)\frac{w+2y}{(w+y)^2}\right)}}$$

and

$$h(w) = 1 - \frac{2(w+y)(w+3y)}{3(w+2y)^2};$$

w is a parameter that takes on the values 0 or 4, where

$$y = \frac{4rA_t}{\sigma^2(1-e^{-r(T-t)})} \quad \text{and} \quad z = \frac{4rD}{\sigma^2(e^{r(T-t)}-1)},$$

which is the other special case of (4.10) and (4.11). The system equation is written as

$$dA_t = \mu A_t dt + \sigma\sqrt{A_t} dW_t^P = rA_t dt + \sigma\sqrt{A_t} dW_t^Q, \tag{4.15}$$

where (4.14) and (4.15) compose the state space model of $\gamma = 0.5$ (hereafter referred to the square-root model). The derivation of this approximative formula is based on Sankaran [64] and [65].

4.2.2 Algorithm to numerically calculate CEV parameters

In the previous section, we set the theoretical framework of the model by using the state space model. In this section, we describe the method to estimate the parameters of the model and the time series of latent asset values using the numerical calculation of Bayesian inference. To estimate the parameters of a statistical model using the MCMC, we use Stan [68] as in Section 3.3.2.

We deal with the estimation of parameters and latent variables in the state space model, which consists of the system equation (4.1) and observation equation (4.2). Applying $X_t = \ln A_t/D$ and Ito's formula to the model (4.1), the system equation can be rewritten to the SDE with respect to X_t under the real probability P as follows:

$$dX_t = \left(\mu - \frac{1}{2}\sigma^2 D^{2(\gamma-1)}e^{2(\gamma-1)X_t}\right)dt + \sigma D^{\gamma-1}e^{(\gamma-1)X_t}dW_t^P, \tag{4.16}$$

where $A_t = De^{X_t}$ should be noted. The Euler discretization formula for (4.16) is obtained by discretizing the formula at the observation time points, $t_i, i = 0, \ldots, n$ for market stock values $S_{t_i}^{mkt}$ (n: the number of the observation time points) using Euler's method as follows:

$$\Delta X_{t_i} = \left(\mu - \frac{1}{2}\sigma^2 D^{2(\gamma-1)} e^{2(\gamma-1)X_{t_{i-1}}}\right)(t_i - t_{i-1}) + \sigma D^{\gamma-1} e^{(\gamma-1)X_{t_{i-1}}} \epsilon_{t_i}, \tag{4.17}$$

where $\Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}$; a white noise is denoted by ϵ_{t_i}, whose average is zero and variance is $t_i - t_{i-1}$, from which we can conduct sequentially the Monte Carlo sampling of the asset values as

$$X_{t_i} \sim N\left(X_{t_{i-1}} + \left(\mu - \frac{1}{2}\sigma^2 D^{2(\gamma-1)} e^{2(\gamma-1)X_{t_{i-1}}}\right)(t_i - t_{i-1}), \sigma^2 D^{2(\gamma-1)} e^{2(\gamma-1)X_{t_{i-1}}}(t_i - t_{i-1})\right) \tag{4.18}$$

which corresponds to (4.16) and (4.17), where $N(m, s^2)$ represents a normal distribution of average m and variance s^2.

Since the stock value can be analytically calculated with the BS model, the parameters and latent variables of the state space model for the Merton model of (4.8) can be estimated by the MCMC in a standard manner. However, the stock value cannot be analytically calculated with the CEV model. Therefore, those of the state space model for the CEV model of (4.1) and (4.2) cannot be estimated by the MCMC in general. To solve this problem, we calculate the curve of the call option values of the CEV model, which represents the theoretical values of stock, by numerically solving the PDE (4.6) using the Crank-Nicholson method [14] in FDMs. It should be noted that the theoretical values of stock calculated by the FDM are regarded as the observation equation. To obtain the difference equation of the Crank-Nicholson method, we take the linear combination of the difference equations of the PDE (4.6) by the implicit method and the explicit method. We derive the method to calculate the theoretical stock value from the PDE (4.6) with the boundary condition (4.7) using the Crank-Nicholson method. The difference equation by the implicit method is described as follows:

$$\frac{V_{i+1,j} - V_{i,j}}{\Delta t} = -rx_j \frac{V_{i,j+1} - V_{i,j-1}}{2\Delta x} - \frac{1}{2}\sigma^2 x_j^{2\gamma} \frac{V_{i,j+1} - 2V_{i,j} + V_{i,j-1}}{(\Delta x)^2} + rV_{i,j}, \tag{4.19}$$

where $j = 1, \ldots, m-1$ and $i = n-1, \ldots, 0$. In contrast, the difference equation by the explicit method is described as follows:

$$\frac{V_{i+1,j} - V_{i,j}}{\Delta t} = -rx_j \frac{V_{i+1,j+1} - V_{i+1,j-1}}{2\Delta x} - \frac{1}{2}\sigma^2 x_j^{2\gamma} \frac{V_{i+1,j+1} - 2V_{i+1,j} + V_{i+1,j-1}}{(\Delta x)^2} + rV_{i+1,j}. \tag{4.20}$$

For the constant $0 \leq \theta \leq 1$, we take the weighted sum of the both sides of the difference equation of the implicit method (4.19) multiplied by θ and those of the difference equation of the explicit method (4.20) multiplied by $(1-\theta)$. From the above, we obtain the difference equation of the Crank-Nicholson method:

$$\begin{aligned}\frac{V_{i+1,j} - V_{i,j}}{\Delta t} &= \theta\left\{-rx_j \frac{V_{i,j+1} - V_{i,j-1}}{2\Delta x} - \frac{1}{2}\sigma^2 x_j^{2\gamma} \frac{V_{i,j+1} - 2V_{i,j} + V_{i,j-1}}{(\Delta x)^2} + rV_{i,j}\right\} \\ &+ (1-\theta)\left\{-rx_j \frac{V_{i+1,j+1} - V_{i+1,j-1}}{2\Delta x} - \frac{1}{2}\sigma^2 x_j^{2\gamma} \frac{V_{i+1,j+1} - 2V_{i+1,j} + V_{i+1,j-1}}{(\Delta x)^2} + rV_{i+1,j}\right\}.\end{aligned} \tag{4.21}$$

In this paper, we take $\theta = 0.5$ to set the ratio of the implicit method and the explicit method as one-to-one. Arranging the both sides of the difference equation (4.21) with respect to $V_{i,j-1}, V_{i,j}, V_{i,j+1}$, we have

$$a_j V_{i,j-1} + b_j V_{i,j} + c_j V_{i,j+1} = d_{i+1,j}, \quad j = 1, \ldots, m-1, i = n-1, \ldots, 0, \tag{4.22}$$

where the coefficients of the equation (4.22) are described as follows:

$$\begin{cases} a_j = \theta\left(\dfrac{r\Delta t}{2\Delta x}x_j - \dfrac{\sigma^2\Delta t}{2(\Delta x)^2}x_j^{2\gamma}\right) \\ b_j = 1 + \theta\left(\dfrac{\sigma^2\Delta t}{(\Delta x)^2}x_j^{2\gamma} + r\Delta t\right) \\ c_j = -\theta\left(\dfrac{r\Delta t}{2\Delta x}x_j + \dfrac{\sigma^2\Delta t}{2(\Delta x)^2}x_j^{2\gamma}\right) \\ d_{i+1,j} = f_j V_{i+1,j-1} + g_j V_{i+1,j} + h_j V_{i+1,j+1} \end{cases} \tag{4.23}$$

and

$$\begin{cases} f_j = (1-\theta)\left(-\dfrac{r\Delta t}{2\Delta x}x_j + \dfrac{\sigma^2\Delta t}{2(\Delta x)^2}x_j^{2\gamma}\right) \\ g_j = 1 + (1-\theta)\left(\dfrac{\sigma^2\Delta t}{(\Delta x)^2}x_j^{2\gamma} + r\Delta t\right) \\ h_j = (1-\theta)\left(\dfrac{r\Delta t}{2\Delta x}x_j + \dfrac{\sigma^2\Delta t}{2(\Delta x)^2}x_j^{2\gamma}\right) \end{cases} \tag{4.24}$$

with the boundary conditions:

$$\begin{cases} V_{n,j} = \max[A_{n,j} - D, 0], \quad j = 0, \ldots, m \\ V_{i,0} = V_{i,1}, \quad V_{i,m} = 2V_{i,m-1} - V_{i,m-2}, \quad i = 0, \ldots, n \end{cases}. \tag{4.25}$$

The boundary conditions with respect to time of the first formula in (4.25) are because the stock values at the maturity are the payoff of call option. In contrast, for the boundary conditions with respect to asset values of the second formula in (4.25), the left end has the Dirichlet condition which gives directly the values of solution on the boundary condition as the slope of call option curve is almost zero, and the right end has the Neumann condition which gives the values of derivative of solution on the boundary condition as the slope of call option curve is almost one.

The novelty of our method is to use the MCMC combined with the FDM of PDE. First, after fixing the parameters, we conduct the backward calculation of the difference equation using the Crank-Nicholson method. Thereby, the surface of theoretical stock values is obtained at $i = 0$. This problem is a solution for simultaneous linear equations that are triple diagonal, and it results in the calculation of the triple diagonal matrix. The triple diagonal matrix can be efficiently calculated using the lower-upper (LU) decomposition. Therefore, the numerical calculation speed is fast. Accordingly, although this method uses the MCMC conbined with the FDM, the computational load is not relatively high. The comprehensive explanation for the FDM of PDEs including the details of the method and the proof of the convergence can be found, for example, in In't Hout [38].

Next, for the path of asset values $A_{t_i}(= De^{X_{t_i}})$ at the time points $t_i, i = 0, \ldots, n$ generated by Euler method, we calculate the theoretical stock values with respect to the asset values by conducting the linear interpolation for the ranges $x_{j-1} \leq A_{t_i} \leq x_j, j = 1, \ldots, m$ using the following formula:

$$S^{mod}(t_i, A_{t_i}) = \frac{V_{0,j} - V_{0,j-1}}{x_j - x_{j-1}}(A_{t_i} - x_{j-1}) + V_{0,j-1} \quad i = n-1, \ldots, 0, \tag{4.26}$$

where the maturity of liability is assumed to be constant. That is, the call options always have the maturity of the same length for the asset values at all time points t_i. Therefore, it is enough to interpolate the surface of theoretical stock values at $i = 0$ for all the asset values at each

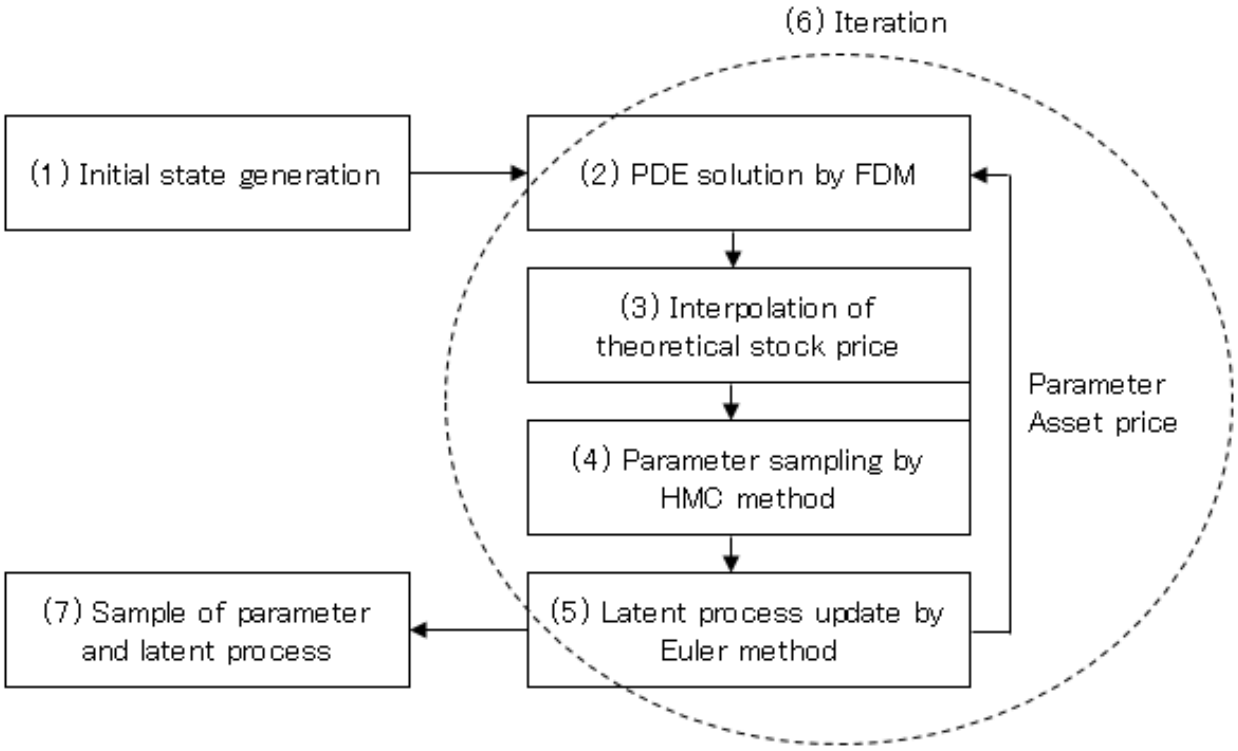


Figure 4.1: Flowchart of the estimation in the PDE-based MCMC. The details of each step are shown in the algorithm described in the text.

time point t_i. We use the call options of constant maturity at each time point $t_i(= 0, \ldots, n)$. Once the PDE is solved from the maturity $i = n - 1$ until $i = 0$, we can evaluate the stock values of all time points. That relieves the computational load. If we had the same call option at each time point, as the maturity became shorter, we would need to solve the PDE at each time point. In that case, the computational load would be very expensive.

The flowchart in Figure 4.1 illustrates our estimation algorithm by the Bayesian statistical inference (HMC) combined with the FDM for solving the PDE. Our algorithm is shown step-by-step in a sequential procedure:

Step 1 Generate the initial states of parameters (μ, σ, γ) and latent variable X_{t_0}.

Step 2 Solve the difference equation (4.22)-(4.24) with (4.25), which $V_{i,j}$ satisfies, corresponding to the PDE (4.6) at all points of asset value $x_j, j = 1, \ldots, m - 1$ and all time points $t_i, i = n - 1, \ldots, 0$ using the Crank-Nicholson method.

Step 3 Conduct the linear interpolation (4.26) for the ranges $x_{j-1} \leq A_{t_i} \leq x_j, j = 1, \ldots, m$, and find the theoretical stock value.

Step 4 Take samples of parameters from the posteriors substituted the market stock values for using the HMC method.

Step 5 Update the latent process X_{t_i} at the observation time points $t_i, i = 1, \cdots, n$ for stock values using the formula (4.18) for the path of asset values approximated by the Euler's method.

Step 6 Repeat the flow from Step 2 to Step 5 sufficiently until the parameters and latent variables converge.

Step 7 Aggregate the posterior samples of parameters and latent variables.

In the Merton model, the theoretical stock values are speedily calculated using the discretized expression of the BS model (4.8) and (4.9) instead of Steps 2 and 3 in this algorithm:

$$S_{t_i}^{mod} = A_{t_i}\Phi(d_1) - De^{-r(T-t_0)}\Phi(d_2),$$

where

$$d_1 = \frac{\ln(A_{t_i}/D) + (r + \frac{\sigma^2}{2})(T-t_0)}{\sigma\sqrt{T-t_0}}, \quad d_2 = \frac{\ln(A_{t_i}/D) + (r - \frac{\sigma^2}{2})(T-t_0)}{\sigma\sqrt{T-t_0}}.$$

4.2.3 Default probability

In this section, we explain the method to calculate the default probability of companies under the real probability P that will be analyzed and used to measure the credit risk of bank portfolios in Section 4.3.3. The CEV model needs the numerical calculation using the FDM to calculate the default probability. Specifically, we can calculate it by replacing the variables $V_{i,j}$ with $p_{i,j}$, which represents the default probability on the grids of the difference equation (4.22), where $j = 1, \ldots, m-1, i = n-1, \ldots, 0$. Additionally, the boundary conditions (4.25) are replaced with the following condition:

$$\begin{cases} p_{n,j} = \mathbb{1}_{\{A_{n,j} \le D\}}, & j = 0, \cdots, m \\ p_{i,0} = p_{i,1}, \quad p_{i,m} = p_{i,m-1}, & i = 0, \cdots, n \end{cases}. \tag{4.27}$$

Furthermore, it is required to substitute the drift μ under P for the risk-free rate. It is noted that for the boundary conditions with respect to time of the first formula in (4.27), the default probabilities at the maturity are the indicator function of default or non-default, and that for the boundary conditions with respect to asset values of the second formula in (4.27), both left and right ends have the Dirichlet condition. Accordingly, the default probability of the CEV model can be calculated by linearly interpolating the curve on $i = 0$ with respect to the asset value in the solution surface.

In the Merton model, the default probability under the real probability P can be analytically calculated by the following formula:

$$PD_t^P = P\{A_T < D\} = \Phi\left(-\frac{\ln(A_t/D) + (\mu - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right), \tag{4.28}$$

where PD_t^P is the default probability of the Merton model, and $P\{A\}$ represents the probability that event A is true under the real probability P.

4.2.4 Model Ⅲ proposed in this study

We propose the PDE-based Bayesian inference for CEV dynamics of assets and stocks. This model consists of the two parts: the state space model for credit risk in stock prices and the algorithm to numerically calculate parameters and latent processes. This model can be applied to the analysis and measurement of credit risk included in the issuers of stocks. We summarize the structural model of credit risk proposed in this study as follows.

- Model
  - State space model

$$\begin{cases} dA_t = \mu A_t dt + \sigma A_t^{\gamma} dW_t^P \\ S_t^{mkt} = S^{mod}(t, A_t) + e_t \\ S^{mod}(t, A_t) = E^Q\left[e^{-r(T-t)}\max[A_T - D, 0] \mid \mathcal{F}_t\right] \end{cases}$$

– Algorithm to numerically calculate parameters
  Sample the state of parameters and latent processes randomly from the parameter distribution of the state space model using the HMC method combined with the FDM to numerically solve the PDEs and Euler's method.

$$\text{Parameter distribution: } \{(\Theta^{(1)}, \Psi^{(1)}), \ldots, (\Theta^{(K)}, \Psi^{(K)})\},$$
$$\begin{cases} \Theta^{(k)} = \{\mu, \sigma, \gamma\} \\ \Psi^{(k)} = \{A_{t_0}, \ldots, A_{t_i}, \ldots, A_{t_n}\} \end{cases}, \quad k = 1, \ldots, K$$

The trial number of simulation is denoted by K. These parameters and latent processes can be calculated according to the flowchart in Figure 4.1.

As mentioned above, this model can calculate the parameter distributions for an arbitrary stock that can be described using the state space model where the system equation is described by the CEV model. Actually, we show the numerical calculation results in the case of bank portfolios. However, to evaluate the system equation described by more complex models and in two dimensions or more, we have to develop a method to solve the problems and reduct computational load in the future.

4.3 Numerical calculation

To analyze the property of credit risk based on the CEV model, we estimate the parameters and latent variables by applying our method to real data. We set the data used in the parameter estimation using our method and show the results of the numerical calculation. We consider the measurement of credit risk based on the estimated results.

4.3.1 Implementation confirmation

In this section, we confirm the implementation of the PDE-based Bayesian inference for CEV dynamics. Specifically, we confirm the implementation of the FDM (Crank-Nicholson method) coded as a user-defined function in Stan. The following three models are prepared as the verification models for comparison:

Model 1: Stan program implemented as the Merton model where the observation equation is analytically solved using the BS model,

Model 2: Stan program implemented as the Merton model where the observation equation is numerically calculated by the FDM, and

Model 3: Stan program implemented as the CEV model where the elastic constant is close to one ($\gamma \approx 1$ and $\gamma < 1$).

Since Model 1 and Model 2 are theoretically the same, they should give the same estimated values of parameters and latent variables. However, these models have completely different implementation. Therefore, if the calculation results of these models by Stan are equal, it implies that the implementations of Model 2 are accurate assuming the analitical formula of Model 1 is accurate. In contrast, in the limits as γ approaches one, the parameters and latent variables of Model 3 should be equal to those of Model 2 as a boundary condition. Therefore, if the calculation results of these models by Stan are close, it suggests that the implementations of Model 3 are accurate since it is shown that Model 2 is accurate. Accordingly, if all of the estimated parameters μ and σ in Models 1 to 3 are equal, the PDE-based Bayesian inference

Table 4.1: Results of implementation confirmation in the PDE-based MCMC. The details of Models 1 to 3 are described in the text. The estimated parameters of Models 1 to 3 for JPMorgan Chase in the FY 2006 are compared at the mean, 2.5 percentile, and 97.5 percentile of posterior distribution. All of them are expected to be equal.

	Model 1			Model 2			Model 3		
	2.5%	mean	97.5%	2.5%	mean	97.5%	2.5%	mean	97.5%
μ	-0.018	0.023	0.064	-0.019	0.024	0.067	-0.019	0.025	0.067
σ	0.019	0.021	0.023	0.020	0.022	0.025	0.020	0.023	0.025

of (4.18) and (4.22)-(4.26) proposed in previous section is at least accurately implemented in the limit of $\gamma = 1$.

Table 4.1 shows the results of implementation confirmation by comparing the estimated parameters of Models 1 to 3 by Stan for JPMorgan Chase in the FY 2006. The estimated parameters of Models 1 to 3 are almost the same at the mean, 2.5 percentile, and 97.5 percentile of posterior distribution. Therefore, the PDE-based Bayesian inference of (4.18) and (4.22)-(4.26) is appropriately implemented on the boundary of $\gamma = 1$ with the Merton model. In contrast, there is no problem in the implementation of the Merton model. Furthermore, the estimated parameters of the CEV model are almost equal to those of the square-root model and the absolute model on the boundary of $\gamma = 0.5$ and $\gamma = 0$, respectively.

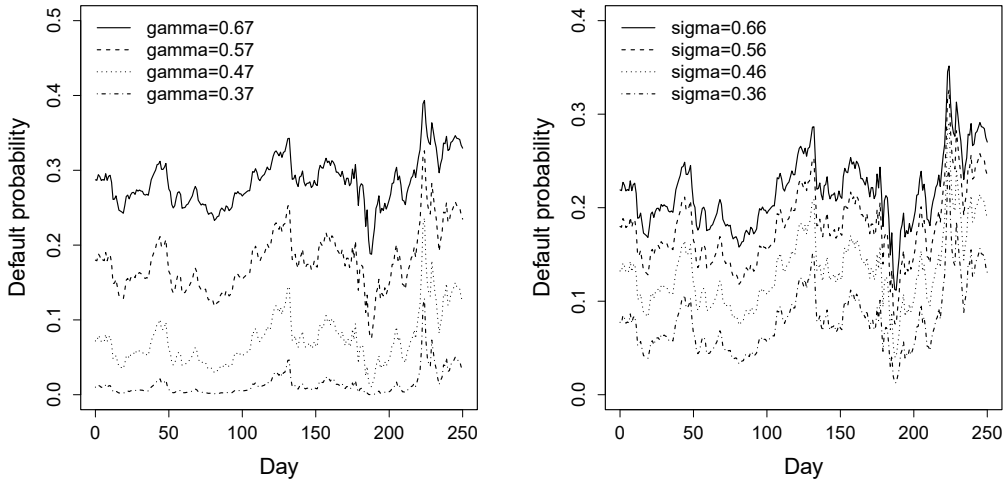
4.3.2 Data settings

We provide the data description and the numerical setting for estimation in our empirical studies. To estimate the parameters and measure the default probabilities using this method, this study uses the daily stock data of the six megabanks in the US: JPMorgan Chase, Goldman Sachs, Citi Group, Morgan Stanley, Bank of America, and Wells Fargo. The observation period of the data is the four periods of the FY 2006 (normal period), the FY 2008 (market deterioration period for the global financial crisis), and the FY 2020 and 2021 (market deterioration period for COVID-19), so that we can compare the normal period with the market deterioration periods. We regard the product of the closing stock price and the total number of issued shares as the market value of stock. As the stock data of each bank, the 251 business days ($\Lambda = 250$) from the end of the fiscal year are used for the estimation using the CEV model and the Merton model. The liability value at the end of the previous year is used. The average value of the daily data of 12 month US dollar LIBOR in the fiscal year is used for the risk-free rate. The maturity of liability is one year and is fixed to the constant maturity. Therefore, all call options, which represent the observation equation, have the maturity of one year at all time points of the market stock price. All data of stocks, liabilities, and interest rates are obtained from Bloomberg. The stock values are calculated on a scale with the liability values up to 100. Then, the number of references m is 6, and the trial number of simulation K is 1 million in the measurement of credit risk of bank portfolios.

4.3.3 Analysis of default probabilities and bank portfolios

Empirical analysis

We show the numerical calculation results estimated under the setting in the previous section. Table 4.2 shows the estimation results of the parameters of the CEV model and the Merton



(a) Default probability for γ (b) Default probability for σ

Figure 4.2: Default probability for elastic constant and asset volatility. In the case of JPMorgan Chase in 2008, the estimated value of elastic constant γ is 0.57, and that of asset volatility σ is 0.56 using the CEV model. The left figure reports the default probability calculated in the cases adding +0.1, -0.1, and -0.2 to $\gamma = 0.57$: 0.67 (solid line), 0.57 (dashed line), 0.47 (dotted line), and 0.37 (dot-dash line). When the elastic constant changes, the other parameters remain fixed. The right figure reports the corresponding time series for the asset volatility $\sigma = 0.56$: 0.66 (solid line), 0.56 (dashed line), 0.46 (dotted line), and 0.36 (dot-dash line). The x axis represents the number of days between January 4, 2008, and December 31, 2008.

model for asset values (system equation) under the real probability P. These are the estimates of average asset return μ, asset volatility σ, and elastic constant γ of the six megabanks in the US in the FY 2006, 2008, 2020, and 2021. The numerical values in Table 4.2 represent the mean (expected a posteriori estimate) of posterior distribution, and those in the parentheses represent the standard deviation of it by the MCMC. We can judge that the sampled values converge to a stationary distribution because the Rhat of all parameters in Table 4.2 is below 1.1. Table 4.2 shows the comparison of the CEV model and the Merton model for the data of the fiscal year in the normal period of 2006, the market deterioration for the global financial crisis of 2008, and that for COVID-19 of 2020 and 2021. The index Watanabe-Akaike information criterion (WAIC) [74] represents how a model fits data.

In common to both the CEV model and the Merton model, it is suggested that the average asset return of most megabanks is positive in 2006 and negative in 2008. Both models reflect the conditions in the normal period and the market deterioration period. The asset volatility is higher in 2008 than 2006. In COVID-19 of 2020 and 2021, the tendency of the market deterioration is similarly read from the parameters. The market in 2021 seems to be recovering from the market decline compared to 2021. Accordingly, it can be said that both models correctly capture the market situation in the deterioration of returns and volatile instability. The elastic constant of the CEV model also shows the same tendency, but it increases significantly compared to the volatility.

Figure 4.2 shows the linkage between the market deterioration and the increase in the elastic constant and the volatility. In the left figure, we show the variation in the default probability with respect to the variation in the elastic constant since the estimated elastic constant was high in the year when the economic environment deteriorated in Table 4.2. Specifically, the cases adding +0.1, -0.1, and -0.2 to the elastic constant in the default probability of JPMorgan Chase in 2008 are compared with each other. When the elastic constant changes, the other parameters remain fixed. As shown in Figure 4.2, the default probability rises when the elastic

Table 4.2: Estimation results of the parameters of the CEV model and the Merton model for asset values (system equation) under the real probability P and the WAIC, which is an evaluation criteria for model selection in Bayesian inference, for the six megabanks in the US in the FY 2006, 2008, 2020, and 2021. This table reports the mean (expected a posteriori estimate) and the standard deviation, which is described in the parentheses, of the posterior distributions by the MCMC. As the Rhat, which indicates a convergence test index in the MCMC, of all parameters is below 1.1, it is suggested that the sampled values of all parameters converge to a stationary distribution. The elastic constant γ of the Merton model is fixed to 1 for all cases.

		JPMorgan Chase		Goldman Sachs		Citi Group	
		CEV	Merton	CEV	Merton	CEV	Merton
2006	μ	0.03 (0.02)	0.02 (0.02)	0.04 (0.02)	0.04 (0.02)	0.02 (0.02)	0.02 (0.02)
	σ	0.50 (0.22)	0.02 (0.00)	0.47 (0.22)	0.02 (0.00)	0.48 (0.21)	0.02 (0.00)
	γ	0.35 (0.11)	1	0.38 (0.12)	1	0.39 (0.12)	1
	WAIC	-937.58	-1,031.26	-850.27	-909.21	-919.21	-997.83
2008	μ	-0.02 (0.07)	-0.01 (0.07)	-0.04 (0.03)	-0.03 (0.03)	-0.06 (0.04)	-0.07 (0.05)
	σ	0.56 (0.21)	0.07 (0.01)	0.54 (0.21)	0.03 (0.00)	0.51 (0.21)	0.05 (0.00)
	γ	0.57 (0.09)	1	0.41 (0.10)	1	0.45 (0.11)	1
	WAIC	68.70	46.50	-670.78	-503.96	-193.99	-377.45
2020	μ	-0.02 (0.06)	-0.02 (0.05)	0.01 (0.03)	0.01 (0.02)	-0.03 (0.04)	-0.02 (0.03)
	σ	0.57 (0.21)	0.05 (0.00)	0.54 (0.21)	0.03 (0.00)	0.53 (0.21)	0.03 (0.00)
	γ	0.53 (0.09)	1	0.42 (0.10)	1	0.46 (0.10)	1
	WAIC	-2.32	2.17	-335.21	-313.17	-383.10	-351.59
2021	μ	0.03 (0.03)	0.02 (0.03)	0.03 (0.03)	0.03 (0.02)	-0.01 (0.02)	0.00 (0.01)
	σ	0.52 (0.21)	0.03 (0.00)	0.51 (0.22)	0.03 (0.00)	0.52 (0.16)	0.02 (0.00)
	γ	0.43 (0.10)	1	0.42 (0.11)	1	0.29 (0.07)	1
	WAIC	-838.59	-964.70	-963.40	-1,059.12	-983.85	-1,095.01
		Morgan Stanley		Bank of America		Wells Fargo	
		CEV	Merton	CEV	Merton	CEV	Merton
2006	μ	0.02 (0.01)	0.02 (0.01)	0.02 (0.02)	0.02 (0.02)	0.03 (0.03)	0.03 (0.03)
	σ	0.46 (0.22)	0.01 (0.00)	0.49 (0.21)	0.02 (0.00)	0.50 (0.22)	0.03 (0.00)
	γ	0.27 (0.12)	1	0.36 (0.11)	1	0.42 (0.11)	1
	WAIC	-1,273.72	-1,446.28	-943.34	-1,053.49	-748.75	-789.75
2008	μ	-0.04 (0.03)	-0.04 (0.03)	-0.08 (0.08)	-0.08 (0.08)	0.03 (0.09)	0.03 (0.08)
	σ	0.49 (0.22)	0.03 (0.00)	0.58 (0.21)	0.08 (0.01)	0.56 (0.21)	0.09 (0.01)
	γ	0.39 (0.11)	1	0.59 (0.09)	1	0.61 (0.09)	1
	WAIC	-231.42	-355.29	-63.28	31.59	176.86	123.38
2020	μ	0.05 (0.05)	0.05 (0.03)	-0.02 (0.05)	-0.02 (0.04)	-0.06 (0.04)	-0.05 (0.04)
	σ	0.56 (0.22)	0.04 (0.00)	0.59 (0.21)	0.04 (0.00)	0.51 (0.21)	0.04 (0.00)
	γ	0.49 (0.10)	1	0.49 (0.08)	1	0.48 (0.10)	1
	WAIC	-259.56	-202.45	-153.48	-117.92	-358.59	-475.00
2021	μ	0.05 (0.04)	0.04 (0.04)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)
	σ	0.49 (0.21)	0.04 (0.00)	0.52 (0.21)	0.03 (0.00)	0.49 (0.22)	0.03 (0.00)
	γ	0.50 (0.10)	1	0.43 (0.10)	1	0.44 (0.11)	1
	WAIC	-690.46	-759.56	-695.60	-803.66	-577.66	-644.27

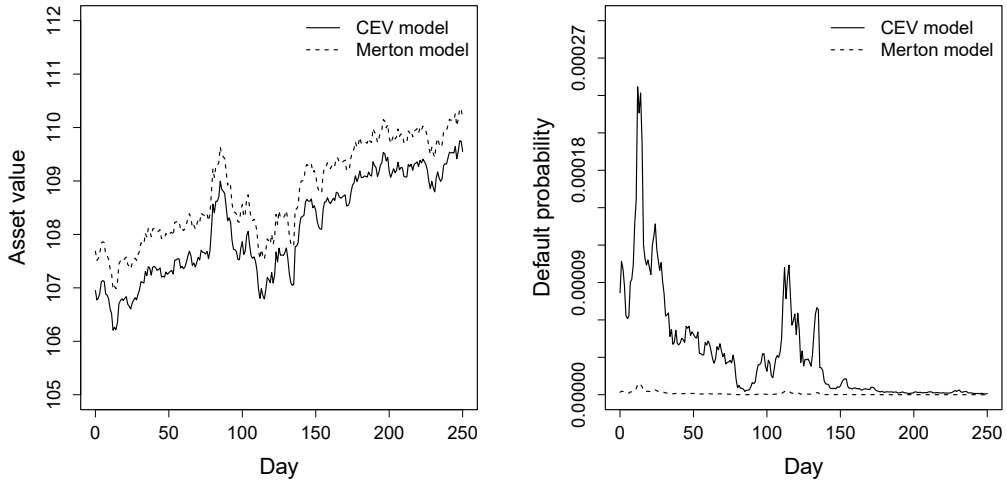
constant rises. In the right figure, we observe a similar trend with respect to the sensitivity to the asset volatility. As shown in comparison of both drifts in the system equation of the CEV model (4.1) and the Merton model (4.8), the average asset returns are almost the same as both models in Table 4.2. Furthermore, as suggested in comparison of both diffusion efficients σA_t^γ and $\hat{\sigma}\hat{A}_t$ in the system equation of both models of (4.1) and (4.8), the volatility of the CEV model multiplied by the asset value to the power of elastic constant is generally equal to the volatility of the Merton model multiplied by the asset value as a standard. Therefore, they are equal on the level of diffusion coefficient. Since the CEV model includes the Merton model, which is in the limit of $\gamma = 1$, it could be said that the estimated parameters of the CEV model are more appropriately selected than those of the Merton model. In the CEV model, since γ relaxes from 1 to $[0, 1]$, two parameters σ and γ are optimized while, in the Merton model, one parameter $\hat{\sigma}$ is only optimized under the condition $\gamma = 1$. The WAIC is different from year to year, therefore the index does not have the tendency which model is lower. Although the CEV model should theoretically fit the data better than the Merton model, the calculation results of the WAIC do not show the property. The estimated theoretical stock values can be reproduced fairly accurately with respect to the market stock values. Incidentally, in this problem, since the observable is only call option, there is a possibility that the convergent solutions exist in local maximum points as the solution of (A_t, σ). If a put option is also an observable in addition to a call option, there is no such problem. However, the observables such as a put option do not exist in the structural approach of credit risk. In such a situation, it is worth analyzing the problem using the the replica exchange Monte Carlo method [36] to escape from local maximum points.

Table 4.3 shows the estimation results of the parameters of the CEV model, the Merton model ($\gamma = 1$), the square-root model ($\gamma = 0.5$), and the absolute model ($\gamma = 0$) for JPMorgan Chase under the real probability P in the FY 2006, 2008, 2020, and 2021. In common to all models, the average asset returns are the same level and have the same tendency as Table 4.2 reflecting the market situation in the deterioration. As seen from the asset volatilities σ and the elastic constant γ of all models, as the elastic constant decreases, the volatility increases, which also shows the same characteristics as Table 4.2. Although the WAIC indicates the worst fitting of the absolute model to the stock data, which is valid as the volatility level is too high, it does not have a tendency for the remaining three models.

Figures 4.3 and 4.4 shows the estimated asset values and default probabilities under P of JPMorgan Chase in the same period using the CEV model and the Merton model. In the left-hand side of all figures, the asset values make the same move with the market stock values in the CEV model and the Merton model during the whole sample time periods. Because the estimated theoretical values of stock are almost similar to the market values, the reproducibility of stock values is very good. In the right-hand side of all figures, the default probabilities move in the opposite directions of the asset values. Therefore, when the graph of the asset values is turned over, it seems to depict the default probabilities. It is noted that there is a difference only in the levels of the asset values and default probabilities between the CEV model and the Merton model although the shape of the graphs is the same. It arises for the difference of the combination of the volatility σ and the elastic constant γ consisting of the diffusion efficient in the system equation, which is the only difference in the CEV model and the Merton model. Beckers [4] shows that the CEV model yields prices which are higher than the BS model prices for in-the-money and at-the-money options. Since, in the structural approach, we consider in-the-money oprions and calculate backwards asset values from the same stock value for the CEV model and the Merton model (BS model), the CEV model should have a lower asset value than the Merton model as shown in Figures 4.3 and 4.4, which is consistent with Beckers [4]. Therefore, the CEV model, whose asset level is closer to the liability level than the Merton model, has a higher default probability as described in Figures 4.3 and 4.4. We will need to

Table 4.3: Estimation results of the parameters of the CEV model, the Merton model ($\gamma = 1$), the square-root model ($\gamma = 0.5$), and the absolute model ($\gamma = 0$) for JPMorgan Chase under the real probability P and the WAIC, which is an evaluation criteria for model selection in Bayesian inference, in the FY 2006, 2008, 2020, and 2021. This table reports the mean (expected a posteriori estimate) and the standard deviation, which is described in the parentheses, of the posterior distributions by the MCMC. As the Rhat, which indicates a convergence test index in the MCMC, of all parameters is below 1.1, it is suggested that the sampled values of all parameters converge to a stationary distribution.

		CEV	Merton	Square-root	Absolute
2006	μ	0.03 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
	σ	0.50 (0.22)	0.02 (0.00)	0.22 (0.01)	1.90 (0.12)
	γ	0.35 (0.11)	1	0.5	0
	WAIC	-937.58	-1,031.26	-1,048.92	-640.65
2008	μ	-0.02 (0.07)	-0.01 (0.07)	-0.02 (0.07)	-0.02 (0.02)
	σ	0.56 (0.21)	0.07 (0.01)	0.72 (0.07)	2.53 (0.17)
	γ	0.57 (0.09)	1	0.5	0
	WAIC	68.70	46.50	83.46	351.00
2020	μ	-0.02 (0.06)	-0.02 (0.05)	-0.02 (0.05)	-0.02 (0.03)
	σ	0.57 (0.21)	0.05 (0.00)	0.52 (0.04)	2.83 (0.14)
	γ	0.53 (0.09)	1	0.5	0
	WAIC	-2.32	2.17	0.58	167.43
2021	μ	0.03 (0.03)	0.02 (0.03)	0.02 (0.03)	0.02 (0.02)
	σ	0.52 (0.21)	0.03 (0.00)	0.29 (0.01)	2.31 (0.13)
	γ	0.43 (0.10)	1	0.5	0
	WAIC	-838.59	-964.70	-929.05	-385.78



(a) Asset value in 2006 (b) Default probability in 2006

Figure 4.3: Asset value and default probability of JPMorgan Chase in FY 2006. In the left figure, the solid line shows the asset values calculated by the CEV model, and the dashed line shows those by the Merton model. In the right figure, the corresponding time series for the default probability. The x axis represents the number of days between January 3, 2006, and December 29, 2006.

analyze the mechanism how to have influence on them and its economic meaning, and which model is more valid should be verified in the future.

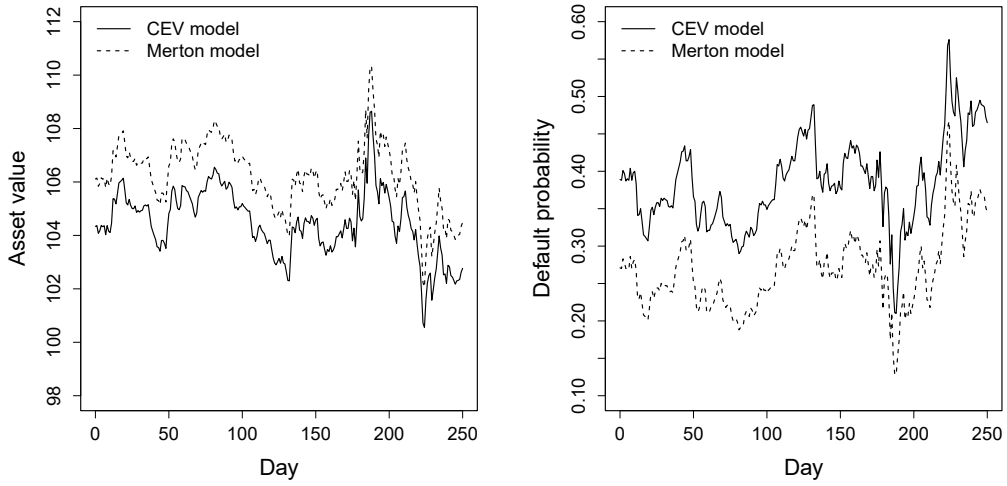
Figure 4.3 shows the estimated results in the normal period of 2006 before the global financial crisis. The default probability in 2006 is almost zero, and it is consistent with the market perception. Figure 4.4 shows the estimated results during the global financial crisis, which depicts the period of financial and economic collapse in 2008. The level of default probabilities is much higher in 2008 than 2006. When the asset value comes close to the level of liability value (equal to 100), the default probability increases. When the global financial crisis happens, the default probability shows a peak. The abovementioned trends are the same in other megabanks and other time periods including COVID-19.

In this method, the maturity of liability is fixed to 1 year, and this setting is consistent with the risk horizon of the credit risk management in practice in view of the financial regulation and the internal rule in banks. However, we can choose the other time-to-maturity. Duan and Fulop [17] chose 10 years. The option value is quite different on the maturity in the observation equation of the CEV model and the Merton model. Accordingly, although this parameter estimation is a problem given the option value, which is the market stock value, it should be recognized that the maturity of liabilities could have a potential impact on the estimation of the parameters.

Measurement of credit risk in bank portfolios

In this study, we consider a credit risk measurement utilizing the default probability of the six megabanks in the US estimated in the previous section. This modeling and analysis lead to the motivation for the advanced modeling and analysis that will be discussed in Section 4.4.3. We adopt a one-factor Merton-type model for the measurement of credit risk. The expected credit loss will be calculated in the portfolio of six banks. Although VaR and CVaR may be used for a risk index in general, we consider an expected loss because the number of reference composing the portfolio is small (six banks) and the VaR and CVaR are unstable as an index that represents a percentile.

The specification of the model is explained as follows. The asset return of reference i is



(a) Asset value in 2008 (b) Default probability in 2008

Figure 4.4: Asset value and default probability of JPMorgan Chase in FY 2008. In the left figure, the solid line shows the asset values calculated by the CEV model, and the dashed line shows those by the Merton model. In the right figure, the corresponding time series for the default probability. The x axis represents the number of days between January 4, 2008, and December 31, 2008.

denoted by $X_i \sim N(0,1)$. The asset return X_i implies the default event of reference i and is described as a one-factor model through a common factor $Z \sim N(0,1)$ and a specific factor $\varepsilon_i \sim N(0,1)$:

$$X_i = \rho_i Z + \sqrt{1-\rho_i^2}\varepsilon_i, \quad i = 1, \ldots, m, \tag{4.29}$$

where m is the number of reference. The efficient ρ_i is the correlation between an asset return X_i and a common factor Z. Assuming that Z and $\varepsilon_i, i = 1, \ldots, m$ are independent of each other, the asset correlation between references can be calculated from (4.29) as follows:

$$\mathrm{Corr}[X_i, X_j] = \rho_i\rho_j + \sqrt{1-\rho_i^2}\sqrt{1-\rho_j^2}\frac{\mathrm{Cov}[\varepsilon_i, \varepsilon_j]}{\mathrm{Var}[Z]} = \begin{cases} 1, & i = j \\ \rho_i\rho_j, & i \neq j \end{cases},$$

where $\mathrm{Corr}[A, B]$ represents a correlation of A and B. In this study, we substitute ρ_i for the correlation between a stock return and a stock market index return as it is conducted in practice. The default threshold, which the default occurs when the asset return goes below, can be written as $c_i = \Phi^{-1}(p_i)$, where p_i is the default probability, and $\Phi^{-1}(\bullet)$ is an inverse function of $\Phi(\bullet)$ because $p_i = P(X_i < c_i) = \Phi(c_i)$. Then, the portfolio loss of the scenario j in simulation by defaults is described as

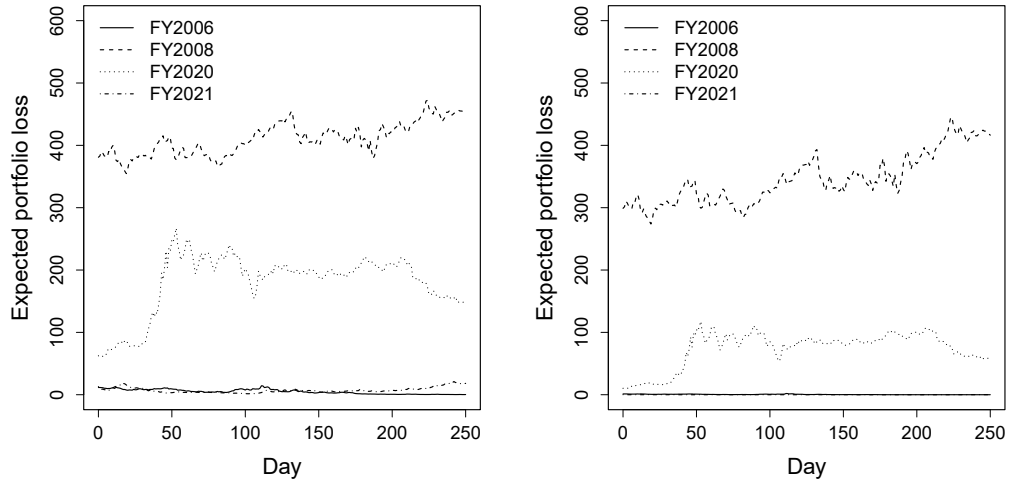
$$PortPL_j = \sum_{i=1}^{m} PL_i, \quad j = 1, \ldots, n,$$

where the loss of reference i is

$$PL_i = \begin{cases} e_i l_i, & X_i < c_i \\ 0, & X_i \geq c_i \end{cases}, \quad i = 1, \ldots, m;$$

the credit exposure is denoted by e_i; the LGD is denoted by l_i; and n is the number of simulation. Accordingly, the expected loss of the portfolio is expressed as

$$EL = \frac{1}{n}\sum_{j=1}^{n} PortPL_j. \tag{4.30}$$



(a) Expected loss for CEV model (b) Expected loss for Merton model

Figure 4.5: Expected loss in the credit risk of the portfolios consisting of the six megabanks in the US. In the left figure, the solid line represents the daily expected losses in the FY 2006; the dashed line represents them in the FY 2008; the dotted line represents them in the FY 2020; and the dot-dash line represents them in the FY 2021 for the CEV model. In the right figure, the corresponding time series for the Merton model. The x axis represents the number of days in each FY year.

We measure the credit risk of the portfolios consisting of the six megabanks in the US as follows. Specifically, we measure the expected loss (4.30) daily in the FY 2006, 2008, 2020, and 2021. The correlation between the asset return of each bank and the common factor is replaced by that between the stock return of each bank and the return of the S&P500 index in each fiscal year. The daily default probability estimated in the previous section is used. The credit exposure is all 100, and the LGD is all 100%. Therefore, the maximum loss of portfolio is 600.

Figure 4.5 shows the numerical calculation results of the expected loss of bank portfolio for the CEV model and the Merton model. In the left figure, we show the expected losses of bank portfolio for the CEV model in the FY 2006, 2008, 2020, and 2021. In the right figure, the corresponding expected losses are discribed for the Merton model. Both figures show the size of expected loss according to the market deterioration. However, the expected losses of the CEV model are totally larger than those of the Merton model since the default probability is larger as shown in Figures 4.3 and 4.4. Because of the small portfolio consisting of the six banks and the invalidity of the evaluation of tail such as the VaR and CVaR, we measure the expected loss as the index of credit risk in FY 2006, 2008, 2020, and 2021. However, the large portfolios of many banks in the banking industry sector should be analyzed in the future.

4.4 Extension of model

We consider the direction of development of this study in the future from three perspectives.

4.4.1 Extension to sophisticated models

In this study, we consider the CEV model, instead of the BS model, in the structural approach, while there are a lot of sophisticated models that have already been proposed. In this section, we discuss the potential of extending our proposed model to other popular models for the parameter estimation using pricing models that is more advanced challenge than the pricing of derivatives.

It could be considered to add a jump term to our diffusion term. Kitabayashi and Nakamura [47] develop a statistical methodology to estimate the jump-diffusion model, specifically, CIR interest rate model with Poisson jump to empirically study the US interest rate market. They assume the state space model as follows:

$$\begin{cases} dr_t = \kappa(\bar{r} - r_t)dt + \sigma_r\sqrt{r_t}dB_t + \xi_t^r dN_t \\ y_t^\tau = -\dfrac{\ln P(\tau, r_t)}{\tau} + \eta_{r,t} \end{cases}, \tag{4.31}$$

where r_t is a short interest rate at time t; κ is a mean-reversion rate; $\bar{r}$ is a long-run mean; σ_r is the standard deviation of a short rate; $\xi_t^r \sim \mathcal{E}(\mu_J)$ is a jump size to follow an exponential distribution of the average μ_J; and $(N_t)_{t\geq 0}$ stands for a Poisson process with a constant arrival intensity λ, independent of a Brownian motion $(B_t)_{t\geq 0}$ under the real probability P. In contrast, y_t^τ is a bond yield remaining τ years; $P(\tau, r_t)$ is the fair value of a bond; and $\eta_{r,t} \sim N(0, \sigma_p^2)$ is an observation error to follow a normal distribution of average 0 and variance σ_p^2. For the market price of interest rate risk, they assume $\phi\sqrt{r_t}$. According to the no-arbitrage argument, the discount bond price in (4.31) is the solution $P(t, r)$ of the integro-PDE (IPDE):

$$\frac{\partial P}{\partial t} + \{\kappa^Q(\bar{r}^Q - r) - \mu_J^Q\lambda\}\frac{\partial P}{\partial r} + \frac{1}{2}\sigma_r^2 r\frac{\partial^2 P}{\partial r^2} + \lambda\int_0^\infty \{P(t, r+\xi_r) - P(t, r)\}f(\xi_r)d\xi_r = rP \tag{4.32}$$

with the terminal condition $P(T, \bullet) = 1$, where $\kappa^Q \coloneqq \kappa + \sigma_r\phi$; $\bar{r}^Q \coloneqq (\kappa\bar{r} + \mu_J^Q\lambda)/\kappa^Q$; $f(\xi_r)$ is the probability density of an exponential distribution of the average μ_J^Q allowing random positive jump sizes $\xi_t^{r,Q} \sim \mathcal{E}(\mu_J^Q)$ under the risk-neutral probability Q. In the equations of (4.31), the system equation is the latent process of short interest rate, and the observation equation is the bond yield calculated from a discount bond price. To estimate the latent short rate process, they employ the Bayesian inference method (MCMC) combined with a FDM to numerically calculate the IPDE (4.32), to the yield curve data (term structure data of discount bond prices).

It could be also considered to expand our model to stochastic volatility models. The Heston model [32] expresses the stock price and its variance by SDEs. In this case, since the stock volatility is also formulated as an observable and it is necessary to numerically solve a two-dimensional PDE, the solution is not easy. Specifically, this problem can be solved using an alternating direction implicit (ADI) scheme, explained in In't Hout [38], for a two-dimensional FDM. Then, Jiao et al. [41] extend the CIR model with a poisson jump to the α-CIR model, which adds the jump part driven by α-stable Levy process. This model characterizes a jump process with heavy tails and a self-exciting structure implying that the jump frequently increases or decreases with the value of the short rate process.

4.4.2 FBSDE of coupling type

Both the CEV model and the Merton model developed in this study are a forward-backward stochastic differential equation (FBSDE) that the "asset" is described as a FSDE with some initial conditions, and the "stock" is described as a BSDE [60] with several terminal conditions, which is a price process of the securities obtained by discounting the payoff given at the maturity. However, in this case, while the asset that is the variable following a FSDE (F-variable) has an effect on the stock that is the variable following a BSDE (B-variable), vice versa does not happen on the structure of this FBSDE. It should be emphasized, because it is so important, that these models are the FBSDE of decoupled F-B variables. We have approached this problem using the Bayesian inference, and Duan and Fulop [17] have approached this problem using the non-Bayes method. Then, both have been able to solve this FBSDE of decoupling type and calculate the parameters of the model and the latent processes.

In contrast, when the credit risk premium depends on the default probability, the asset value and the stock value have influence on each other. Consequently, it is necessary to be formulated by the FBSDE of coupling type. In general, when the F-variable depends on the B-variable in the drift term (that is, there is the coupling between the F-variable and the B-variable), it is difficult to solve the FBSDE. Although Duan and Fulop [17] rely on the forward particle sampling, their method does not work in the case that there is the coupling of B-variable. Therefore, the only valid method is the HMC sampling conbined with the PDE that characterizes the FBSDE as shown in this study. This is a very important point where our Bayesian method dominates their non-Bayes method.

In the case that there is a positive correlation between the market price of credit risk and the default probability (that is, if the default probability is high (low), the market price of credit risk is high (low)), it can be judged that the relationship between the two parameters has a structure. We are interested in whether the market price of credit risk is constant or depends on a default probability. An excess return can be obtained in compensation for default risk when depending on a default probability.

If there is a linear dependence structure from the market price of credit risk to the default probability, the Merton model should be extended to the FBSDE of coupling type. Then, the market price of credit risk is rewritten as follows:

$$\frac{\mu - r}{\sigma} = \phi = \alpha + \beta \cdot PD_t^Q,$$

where the rightmost side represents a linear dependence structure from the market price of credit risk to the default probability. Accordingly, the state space model of the Merton model is rewritten as follows:

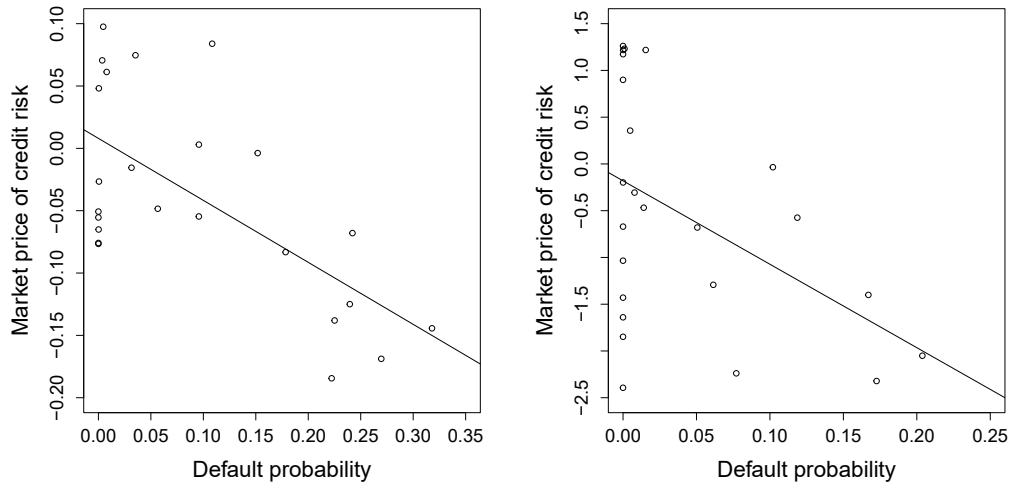
$$\begin{cases} dA_t = (r + \sigma\phi)A_t dt + \sigma A_t dW_t^P = \{r + \sigma(\alpha + \beta \cdot PD_t^Q)\}A_t dt + \sigma A_t dW_t^P \\ dS_t = (r + Z_t\phi)S_t dt + Z_t S_t dW_t^P = \{r + Z_t(\alpha + \beta \cdot PD_t^Q)\}S_t dt + ZS_t dW_t^P \end{cases},$$

where the bundary condition is described as $S_T = \max[A_T - D, 0]$; S_t is the stock value; $Z_t = \sigma\partial S(A_t)/\partial A$ is the stock volatility; and PD_t^Q is the default probability of the Merton model under the risk-neutral probability Q written as

$$PD_t^Q = Q\{A_T < D\} = \Phi\left(-\frac{\ln(A_t/D) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right)$$

obtained by substituting the risk-free rate r for the drift μ in (4.28), and $Q\{A\}$ represents the probability that event A is true under the risk-neutral probability Q. In the existing literature, Duffie et al. [20] and Hogan [35] study a consol rate, which is the yield on a perpetual annuity, in relation to a FBSDE. Similarly, the CEV model can be extended to the FBSDE of coupling type. In that case, the default probability is calculated using the FDM with the boundary conditions (4.27) in Section 4.2.3 substituting the drift r under Q for the risk-free rate.

Figure 4.6 analyzes the relationship by plotting the default probability and the market price of credit risk in the daily stock data of the six megabanks in the FY 2006, 2008, 2020, and 2021. Both the CEV model in the left figure and the Merton model in the right figure seem to have the negative linear relationship between the default probability and the market price of credit risk. The coefficients of determination are 0.423 and 0.211, respectively. Although both variables appear to be related, the sign of the slope of approximation line is opposite to the assumption. However, the two variables may be related. It seems to be worth considering the other formulations between the default probability and the market price of credit risk. We will research this matter in the future with more data by expanding the CEV model developed in this study and the other sophisticated models to a FBSDE.



(a) PD vs risk premium in CEV (b) PD vs risk premium in Merton

Figure 4.6: Relationship of the annual average default probability and the market price of credit risk in the daily stock data of the six megabanks in the FY 2006, 2008, 2020, and 2021. In the left figure, the plots shows the relationship between both variables calculated by the CEV model. In the right figure, the corresponding plots for the Merton model.

4.4.3 Systemic risk in banking industry

It is conceivable to apply our method to the estimation of systemic risk in banking industry. The purpose of this subject is as follows. Since the structure of lending and borrowing create a network of banks in the financial system, the default risk of one bank propagates the entire financial system. Central banks are concerned about the contagion of shock to the entire financial system due to the default of one bank. Accordingly, we will predict the occurrence of systemic risk in a global financial crisis and bubble burst using the time series of bank stock price and realized volatility calculated from the high frequency data of bank stock price. We analyze the relationship between the stock return and the default risk by extracting the information related to the default from stock prices. As the approach of study, it is conceivable to estimate the asset value and volatility of banks using the HMC method, the asset correlation using the dynamic conditional correlation (DCC) [24], and the systemic risk in banking industry based on the conditional VaR (CoVaR) [2] calculated using the quantile regression [49].

This subject links to the study on the Ising model in credit risk modeling in Kato [44]. The common points are to deal with credit risk (phenomenon that the occurrence probability is small and the loss is large), to consider the contagion from individual to whole in credit risk (default), and to capture the signs in the system and environmental degradation as the occurrence probability is small. The credit risk of bank portfolios could be measured using the long-range Ising model in credit risk modeling.

4.5 Conclusion

The contributions of this study are to develope the PDE-based Bayesian inference of CEV dynamics for asset values and to apply the model to the real stock data of the six US megabanks. This PDE-based MCMC enables us to solve the estimation problem of parameters when the observation equation includes the model price that cannot be represented by the analytical formula. The results of the numerical calculation are applied to the analysis of the default probability and the measurement of credit risk in bank portfolios.

We face the following challenge in future, as discussed in Section 4.4. First, it could be

considered the extension of our model for the parameter estimation using pricing models to the sophisticated models for the pricing of derivatives. We will try to add a jump term to our diffusion term and to expand our model to stochastic volatility models such as the Heston model. Second, it could be considered the extension of our model to the FBSDE of coupling type. In the case that the coupling term is included in the FBSDE, the HMC sampling conbined with the FDM of PDEs that characterizes the FBSDE proposed in this study is the only valid method, while the forward particle sampling, which is a non-Bayes method, proposed by Duan and Fulop [17] does not work. This is a very important point where our Bayesian method dominates their non-Bayes method. Third, it is conceivable to apply our method to the estimation of systemic risk in banking industry sector. We will predict the occurrence of systemic risk in a financial crisis using the time series of bank stock price and its volatility. As the approach of study, it is conceivable to estimate the asset values and volatilities of banks using the HMC method conbined with the FDM of PDEs and to measure the systemic risk in banking industry sector based on the CoVaR.

Chapter 5

Conclusion

This study contributes to model and analyze covariation and contagion of credit risk and apply them in practice. It also develops the numerical calculation methods for the credit risk modeling using the Bayesian statistical inference method (MCMC). Focusing on credit risk modeling using Bayesian statistics and statistical physics, this study works on the following three subjects. It constructs the long-range Ising model for credit risk modeling in heterogeneous credit portfolios and measures the credit risk of practical heterogeneous credit portfolios from the point of view in econophysics; analyzes the cointegration structure between the term structures of CDSs through the VECM (cointegration model) of hazard rates and applies it to the pairs trading strategy from the point of view in finance; and measures corporate credit risks based on the stock values in the structural model expanded to the CEV model followed by asset values. Furthermore, this study contributes to developing the following three new models: the Ising model for credit risk modeling (Model I); the ODE-based Bayesian inference (Model II); and the PDE-based Bayesian inference (Model III). As a result, this study contributes the novelty to finance and econophysics from both theoretical and empirical aspects for the modeling and analysis of covariation and contagion of credit risk in each subject mentioned.

Chapter 2 constructs the long-range Ising model for credit risk modeling in heterogeneous credit portfolios. This model expresses the heterogeneity of the default probability and the default correlation by dividing a credit portfolio into multiple sectors characterized by credit rating and industry. It also expresses the heterogeneity of the credit exposure, which is difficult to evaluate analytically, by applying the RMC method to numerically calculate loss distributions. To analyze the characteristics of loss distributions for credit portfolios with heterogeneous credit exposures, we apply this model to various credit portfolios and evaluate credit risk. As a result, we show that the tail of loss distributions calculated by this model has characteristics different from that by the standard models used in credit risk modeling. We also show that there is a possibility of different evaluations of credit risk according to the pattern of heterogeneity.

Chapter 3 reveals the follow two: first, a simple fundamental model for the cointegrated hazard rates can create a diversity of cointegration relationships between various combinations of relevant CDS spreads with different maturities; second, our cointegrated hazard model can be applied to the investment strategy of CDSs. As for the first point, we can effectively merge the cointegration structure into the framework of arbitrage-free pricing and calculate the theoretical spread of CDSs, so that we clarify why the cointegration appears among CDSs with different maturities. Furthermore, we understand how the cointegration among the term structures of CDSs occurs through hazard rates. As for the second point, by applying our model to the Japanese cointegrated corporate CDS spread data, we obtain the investment weight of each individual CDS. We develop the ODE-based Bayesian inference of cointegration dynamics for hazard rates. In the empirical study, we demonstrate that our statistical method is workable for estimating parameters and latent processes in the cointegrated hazard rate model.

Chapter 4 develops the PDE-based Bayesian inference of CEV dynamics for asset values and applies the model to the real stock data of the six US megabanks. This PDE-based MCMC enables us to solve the estimation problem of parameters when the observation equation includes the model price that cannot be described by the analytical formula and even when the ODEs are not obtained. The results of the numerical calculation are applied to the analysis of the default probability and the measurement of credit risk in bank portfolios.

There are many concluding remarks. We will continuously apply the concepts and methods in physics to the problems in finance in the interdisciplinary area between finance and physics because financial engineering and physics have many overlapping ideas. Information statistical mechanics and information science will be helpful. For future research, we will take two paths, econophysics and finance, and fill the gap between both fields from two directions. Concepts that exist in physics but not in finance (temperature, etc.) are the key to developing new academic fields. In econophysics, we analyze and predict the behavior of the target in finance focusing on a default cooperative phenomenon, a phase transition in credit portfolios, and the universal nature as large scale, many body, and complex systems by developing the Ising model for credit risk modeling (Model I). In finance, we study the cointegration of credit products and its application to investment strategies focusing on a dynamic pairs trading strategy, the optimal investment problem of pairs portfolio, and the measurement of the default probability and credit risk by developing the ODE-based Bayesian inference (Model Ⅱ). Furthermore, we study the FBSDE of coupling type and the systemic risk in banking industry using the CoVaR by developing the PDE-based Bayesian inference (Model Ⅲ). We believe that it contributes to the advancement of finance and econophysics to conduct further research for these problems.

Appendix A

Extended Ensemble Monte Carlo Method

A.1 Replica exchange Monte Carlo method

The simplest method of sampling of points from a probability distribution (spin states from a canonical distribution) using random numbers is a form of random sampling that samples points from a probability distribution randomly. However, in high-dimensional and multivariate sampling, it is difficult to sample uniformly the region (spin states) targeted in a high-dimensional space. If the probabilities are small at most points of a probability distribution, there is a high possibility of missing the important points (low energy states) that have a small region (states are few) but have high associated probabilities. In contrast, the Metropolis method and the MH method sample points selectively from an important region that has a large probability but not uniformly from a probability distribution. Specifically, it starts from an initial value (spin initial state) generated randomly and selects a state that varies a little from the present state. Furthermore, it repeats a state transition that randomly adopts the candidate state or not. As a result, it realizes the sampling as samples are proportional to the probability distribution. However, as mentioned in this paper, the transitions from one peak to the other peak for a multi-modal distribution occur only rarely. Therefore, there is a problem in that the sampling points stay on a local subspace in a high-dimensional space, and sampling efficiency deteriorates (the problem of slow relaxation). The RMC method sets multiple replicas of the system to different temperatures (parameters that denote the degree of stochastic state transitions). Further, it sets a targeted distribution (low temperature and slow convergence), a distribution where a constraint is relaxed sufficiently (high temperature and fast convergence), and distributions to interpolate them. It updates the states of each replica in the temperatures using the Metropolis method and exchanges the states between adjacent replicas. Hereby, each replica escapes from local states under low temperature, and the relaxation becomes fast because each replica moves around on a temperature axis (parameter direction) with the temperatures rising and falling automatically.

Let us assume a targeted distribution $P(x)$. The Metropolis method samples as $P(x)$ becomes a stationary distribution. In order that a targeted distribution becomes a stationary distribution, the conditions a transition probability $\pi(x \to x')$ should satisfy are detailed balance and ergodicity. Detailed balance is the condition wherein the inflow and outflow of probability between two arbitrary points are balanced. In the Metropolis method, regarding arbitrary states x, x',

$$P(x)\pi(x \to x') = P(x')\pi(x' \to x) \tag{A.1}$$

is satisfied. Ergodicity is the condition wherein the transition probability between two arbitrary points x and x' is not zero or is expressed by a product of a finite number of transition

probabilities that are not zero. That is, any two points in a space are connected by a finite transition probability, and the state can come and go between any two points in finite steps. In this case, the Markov chain constructed from an arbitrary initial state converges to the only stationary distribution $P(x)$. Since

$$\frac{\pi(x \to x')}{\pi(x' \to x)} = \frac{P(x')}{P(x)} \tag{A.2}$$

from (A.1), detailed balance of (A.2) is satisfied:

$$\pi(x \to x') = \min\left[1, \frac{P(x')}{P(x)}\right]. \tag{A.3}$$

The algorithm of the Metropolis method is defined by executing a transition as a targeted distribution $P(x)$ becomes a stationary distribution. Given a present state x and a selected candidate x', this algorithm accepts the candidate x' by probability $\min[1, r]$, otherwise it stays at the present state x from (A.3), where

$$r = \frac{P(x')}{P(x)}.$$

The RMC method samples using the MCMC as a joint distribution consisting of multiple distributions $P(x \mid \beta_k)$, which have different values of parameter $\beta_k, k = 1, \ldots, R$,

$$P(x_1, \ldots, x_R) = \prod_{k=1}^{R} P(x_k \mid \beta_k) \tag{A.4}$$

becomes a stationary distribution. If $P(x \mid \beta_k), k = 1, \ldots, R$ in (A.4) are canonical distributions, we consider a series of distributions

$$P(x \mid \beta_k) = \frac{1}{Z(\beta_k)} \exp[-\beta_k H(x)], k = 1, \ldots, R \tag{A.5}$$

such that a class of inverse temperatures is $\beta = (\beta_1, \ldots, \beta_R)$. Detailed balance for the RMC method implies that

$$\begin{aligned} &P(x_1, \ldots, x_k, x_{k+1}, \ldots, x_R)\pi(\{x_1, \ldots, x_k, x_{k+1}, \ldots, x_R\} \to \{x_1, \ldots, x_{k+1}, x_k, \ldots, x_R\}) \\ &= P(x_1, \ldots, x_{k+1}, x_k, \ldots, x_R)\pi(\{x_1, \ldots, x_{k+1}, x_k, \ldots, x_R\} \to \{x_1, \ldots, x_k, x_{k+1}, \ldots, x_R\}) \end{aligned} \tag{A.6}$$

is held for arbitrary replicas x_k, x_{k+1}. Since

$$\frac{\pi(\{x_k, x_{k+1}\} \to \{x_{k+1}, x_k\})}{\pi(\{x_{k+1}, x_k\} \to \{x_k, x_{k+1}\})} = \frac{P(x_{k+1} \mid \beta_k)P(x_k \mid \beta_{k+1})}{P(x_k \mid \beta_k)P(x_{k+1} \mid \beta_{k+1})} \tag{A.7}$$

from (A.6), detailed balance of (A.7) holds:

$$\pi(\{x_k, x_{k+1}\} \to \{x_{k+1}, x_k\}) = \min\left[1, \frac{P(x_{k+1} \mid \beta_k)P(x_k \mid \beta_{k+1})}{P(x_k \mid \beta_k)P(x_{k+1} \mid \beta_{k+1})}\right]. \tag{A.8}$$

The algorithm of the RMC method is defined by executing two types of transitions alternately such that the joint distribution (A.4) becomes a stationary distribution. Its algorithm is as follows:

Step 1: Updating the state of each replica independently

Update the state of each replica at the temperature using the Metropolis method. For each distribution $P(x_k \mid \beta_k)$ of (A.5), update state x_k in parallel by the Metropolis method, which has a stationary distribution $P(\cdot \mid \beta_k)$.

Step 2: Exchanging the states between replicas

Every several steps, exchange the adjacent states x_k, x_{k+1} by probability $\min[1, r]$ from (A.8), where

$$r = \frac{P(x_1, \ldots, x_{k+1}, x_k, \ldots, x_R)}{P(x_1, \ldots, x_k, x_{k+1}, \ldots, x_R)} = \frac{P(x_{k+1} \mid \beta_k)P(x_k \mid \beta_{k+1})}{P(x_k \mid \beta_k)P(x_{k+1} \mid \beta_{k+1})}. \tag{A.9}$$

Because $(x_1, \ldots, x_R)$ sampled as a consequence of the above two types of transitions is a sample from $P(x_1, \ldots, x_R)$, x_k can be regarded as a sample from $P(x_k \mid \beta_k)$.

The RMC method is one of the extended ensemble Monte Carlo methods, like the multicanonical method and the simulated tempering method. The comprehensive explanation for the extended ensemble Monte Carlo method including the RMC method can be found in Iba [37]. We adopt the RMC method, which is practically tractable and suited for parallel computing.

A.2 Setting of replica exchange Monte Carlo method

The computation of the RMC method is given in Table A.1. We set temperature parameters according to Hukushima and Nemoto [36] and Iba [37]. With regard to the temperature range, the maximum temperature (minimum inverse temperature) is set to a high temperature layer where the relaxation of the system is sufficiently fast and the distribution of the number of defaulting borrowers becomes uni-modal. On the other hand, the minimum temperature (maximum inverse temperature) is set to a low temperature layer close to the probability distribution of the needed inverse temperature. The probability distribution of the needed inverse temperature is set to $\beta = 1$, and it is the fifth in order of increasing temperature. With regard to the number of temperature values needed in the temperature range and the difference between the adjacent inverse temperatures, in order that each replica wanders over the whole temperature range, the exchange probabilities of every pair of replicas at different temperatures have to be of the order of one and nearly constant. For exchange probability $\min[1, r]$ of the adjacent inverse temperatures β_k, β_{k+1}, the expectation of the logarithm of the ratio (A.9) is

$$\begin{aligned} E[\ln r] &= \sum_{x_k} \sum_{x_{k+1}} P(x_k \mid \beta_k)P(x_{k+1} \mid \beta_{k+1}) \ln \frac{P(x_{k+1} \mid \beta_k)P(x_k \mid \beta_{k+1})}{P(x_k \mid \beta_k)P(x_{k+1} \mid \beta_{k+1})} \\ &= -\left\{ \sum_{x_k} P(x_k \mid \beta_k) \ln \frac{P(x_k \mid \beta_k)}{P(x_k \mid \beta_{k+1})} + \sum_{x_{k+1}} P(x_{k+1} \mid \beta_{k+1}) \ln \frac{P(x_{k+1} \mid \beta_{k+1})}{P(x_{k+1} \mid \beta_k)} \right\} \\ &= -\left\{ D(P(x_k \mid \beta_k) \parallel P(x_k \mid \beta_{k+1})) + D(P(x_{k+1} \mid \beta_{k+1}) \parallel P(x_{k+1} \mid \beta_k)) \right\}, \end{aligned} \tag{A.10}$$

where

$$D(P(x_k \mid \beta_k) \parallel P(x_k \mid \beta_{k+1})) = \sum_{x_k} P(x_k \mid \beta_k) \ln \frac{P(x_k \mid \beta_k)}{P(x_k \mid \beta_{k+1})} \tag{A.11}$$

is Kullback-Leibler divergence. For Fisher information

$$\begin{aligned} I(\beta_k) &= -E\left[\frac{\partial^2 \ln P(x_k \mid \beta_k)}{\partial \beta_k^2} \right] = -\sum_{x_k} P(x_k \mid \beta_k) \frac{\partial^2 \ln P(x_k \mid \beta_k)}{\partial \beta_k^2} \\ &= \frac{\partial^2 \ln Z(\beta_k)}{\partial \beta_k^2} = \frac{N}{\beta_k^2} \frac{\partial}{\partial T_k} \left(\frac{\langle H \rangle}{N} \right) = \frac{N}{\beta_k^2} c_k, \end{aligned} \tag{A.12}$$

Table A.1: Setting of the RMC method. In statistical physics, it has been said that about 50% acceptance probability is good. Although the number 50% is not necessarily reliable, it should be noted that it becomes inefficient when the acceptance probability is too large.

Item	Setting
Temperature	The inverse temperature is in the range of 0.4 to 1.12. Because the number of spins in the system is 500, the number of temperature values is $25(> \sqrt{500})$, and the difference between the adjacent inverse temperatures is set to $0.03(< 1/\sqrt{500})$ at equal intervals.
Exchange probability of replicas	About 63∼64% in the total system
Number of trials	The number of samples composing the loss distribution is 1 billion (the number of generated samples is 1 trillion). However, the number of samples in the case of converging fast due to the parameter setting is less than 1 billion.
Removal of scenarios as initial relaxation	Removing the initial half of the generated samples
Sampling frequency of scenarios	Every 500 times, which is the system size
Exchange frequency of replicas	Every 500 times, which is the system size
Random number	XorShift

when $\beta_k \approx \beta_{k+1}$, by using the relational expression between Kullback-Leibler divergence (A.11) and Fisher information (A.12)

$$D(P(x_k \mid \beta_k) \parallel P(x_k \mid \beta_{k+1})) \approx \frac{1}{2} I(\beta_k)(\beta_{k+1} - \beta_k)^2, \tag{A.13}$$

we obtain

$$E[\ln r] \approx -I(\beta_k)(\beta_{k+1} - \beta_k)^2 = -\frac{N}{\beta_k^2} c_k (\beta_{k+1} - \beta_k)^2 \tag{A.14}$$

from (A.10)-(A.13), where c_k represents specific heat per component. Because $E[\ln r] \sim O(1)$, the difference between the adjacent inverse temperatures to maintain consistency is

$$\beta_{k+1} - \beta_k \sim O\left(\frac{1}{\sqrt{N}}\right) \tag{A.15}$$

from (A.14). On the other hand, the number of temperature values needed in the temperature range is

$$R \sim O(\sqrt{N})$$

from (A.15). We need more replicas where the specific heat per component takes a larger value.

For the setting of Table A.1, it is confirmed that the adjacent replicas exchange sufficiently and each replica reaches from one edge of the temperature range to another by several hundred MCS (Monte Carlo Sweep). Accordingly, mixing time is deemed sufficiently short, and the loss distribution converges sufficiently as indicated in the simulation from the reproduced result of the loss distribution for the homogeneous credit portfolio. Therefore, residual errors can be said to exist because the size of the second peak was originally small.

A.3 Hamiltonian Monte Carlo method

The MH algorithm becomes impractical when it is applied to the case of a large number of parameters, and the drop in acceptance rate is significant. To make sampling using the MH

algorithm successful, it is necessary to control the degree of how far to generate a candidate from a present state appropriately. When this degree is large, the average travel distance is long in one transition. However, if the travel distance becomes too long, the sample often jumps out of the high density area in the targeted distribution in the transitions. Therefore, the problem is caused that the acceptance rate becomes low. In contrast, when this degree is small, the average travel distance is short in one transition. However, if the travel distance becomes too short, the sample has to transition in the narrow area of parameter space. Therefore, it takes a long time to obtain the samples representative of the nature of the targeted distribution. In the case of a large number of parameters, it is difficult to control this degree.

In the HMC method, we put the distribution we want to sample as potential energy and introduce kinetic energy. Then, taking advantage of the nature that the Hamiltonian, which is the sum of both energies, saves and increasing transition distance and candidate acceptance rate at the same time, we obtain the random numbers following the distribution we want to sample. Let us assume a targeted distribution $P(x)$. The MH algorithm samples as $P(x)$ becomes a stationary distribution. In order that a targeted distribution becomes a stationary distribution, the conditions a transition probability $\pi(x \to x')$ should satisfy are detailed balance

$$P(x)\pi(x \to x') = P(x')\pi(x' \to x), \tag{A.16}$$

regarding arbitrary states x, x', and ergodicity as explained in Appendix A.1. In this case, the Markov chain constructed from an arbitrary initial state converges to the only stationary distribution $P(x)$. However, as it is difficult to find a transition probability $\pi(x \to x')$ immediately with which the distribution we want to sample $P(x)$ satisfies detailed balance, we use a proposal distribution instead of it, which is a suitable probability distribution from which it is easy to generate random numbers. Let the probability with which x' is selected as a candidate be $Q(x, x')$ when staying a state x. Then, detailed balance of (A.16) is satisfied:

$$\pi(x \to x') = \min\left[1, \frac{P(x')Q(x', x)}{P(x)Q(x, x')}\right]. \tag{A.17}$$

The algorithm of the MH method is defined by executing a transition as a targeted distribution $P(x)$ becomes a stationary distribution. Given a present state x and a selected candidate x', this algorithm accepts the candidate x' by the probability $\min[1, r]$, otherwise it stays at the present state x from (A.17), where

$$r = \frac{P(x')Q(x', x)}{P(x)Q(x, x')}.$$

In the HMC method, we take advantage of the nature $H(x, p) = U(x) + K(p) = mgh(x) + p^2/2m$ that the Hamiltonian, which is the sum of potential energy $U(x) = mgh(x)$ and kinetic energy $K(p) = p^2/2m$, saves and generate the random numbers following the distribution we want to sample $P(x)$, where x is position; p is momentum; h is height; m is mass; and g is gravitational acceleration. Denoting a time subscript as t, since the Hamiltonian is always constant regardless of changes in time, the Hamilton equation of motion is derived as follows:

$$\begin{cases} \dfrac{dp(t)}{dt} = -\dfrac{dU(x(t))}{dx(t)} = -h'(x(t)) \\ \dfrac{x(t)}{dt} = \dfrac{dK(p(t))}{dp(t)} = p(t) \end{cases}, \tag{A.18}$$

where $h'(x(t))$ is the derivative of $h(x(t))$ with respect to $x(t)$. We assume $m = 1$ as interested in the motion path and experiments on the planet of $g = 1$ instead of the earth in (A.18). We

solve the equation of motion using the leapfrog method, which is a method to conserve the Hamiltonian with good accuracy and predict the path of motion:

$$\begin{cases} p\left(t+\frac{1}{2}\right) = p(t) - \frac{\epsilon}{2}h'(x(t)) \\ x(t+1) = x(t) + \epsilon p\left(t+\frac{1}{2}\right) \\ p(t+1) = p\left(t+\frac{1}{2}\right) - \frac{\epsilon}{2}h'(x(t+1)) \end{cases} \quad , \quad t = 1, \cdots, L, \tag{A.19}$$

where the number of updates L represents exercise time by the leapfrog method and the step size ϵ represents accuracy. The equations (A.19) converges to the Hamilton equation of motion (A.18) in $\epsilon \to 0$. In the leapfrog method, we change the momentum by half in the first formula and identify the position with good accuracy using half-changed momentum in the second formula. Then, we determine total momentum with precise positioning in the third formula. Incidentally, although a subscript t represents real time in the physical world so far, it is time to make one transition in a probability density function in the HMC method. Therefore, we do not display t below. Then, the Hamiltonian is the function of position and momentum as

$$H(x,p) = U(x) + K(p) = h(x) + \frac{1}{2}p^2. \tag{A.20}$$

It can be said that the leapfrog method is an implementation procedure to predict the flow of the Hamiltonian along the contour in the phase space (space with coordinates position and momentum).

In the HMC method, we make a proposal distribution solving the Hamilton equation of motion using the leapfrog method and accept or reject it by the way like the MH algorithm. Let us consider to generate random numbers from the joint probability distribution

$$F(x,p) = P(x)f(p) \tag{A.21}$$

consisting of the distribution we want to sample $P(x)$ and the standard normal distribution $f(p)$ independent of it. The joint probability distribution is

$$\begin{aligned} F(x,p) &= \exp[\ln F(x,p)] = \exp[\ln P(x) + \ln f(p)] \propto \exp\left[\ln P(x) + \ln \exp\left[-\frac{1}{2}p^2\right]\right] \\ &= \exp\left[\ln P(x) - \frac{1}{2}p^2\right] = \exp\left[-h(x) - \frac{1}{2}p^2\right] = \exp[-H(x,p)] \end{aligned} \tag{A.22}$$

from (A.20) and (A.21), where we put $\ln P(x) = -h(x)$. The targeted distribution $P(x)$ is put as minus potential energy $-h(x)$. It means that we regard the parameter x as position and the size of log probability as the low in physical space. The last equal sign in (A.22) means that we can regard x, p in the joint probability distribution as the position coordinates in the phase space, then the kernel in the joint probability distribution is the Hamiltonian itself. The transition probability from the state (x,p) to the state (x',p') at transition destination after time L satisfies

$$\pi(x,p \to x',p') = \pi(x',p' \to x,p) \tag{A.23}$$

as it is reversible, and its volume is preserved. Since the Hamiltonian is constant, the following holds:

$$F(x,p) = F(x',p'). \tag{A.24}$$

Accordingly, the detailed equilibrium condition

$$F(x,p)\pi(x,p \to x',p') = F(x',p')\pi(x',p' \to x,p) \tag{A.25}$$

holds from (A.23) and (A.24), where

$$\pi(x, p \to x', p') = \min\left[1, \frac{F(x', p')}{F(x, p)}\right] \tag{A.26}$$

satisfys the detailed equilibrium condition (A.25). Since the path is predicted using ϵ instead of 0 in the leapfrog method, a numerical calculation errors occur in the Hamiltonian. Thus, setting

$$r = \frac{F(x', p')}{F(x, p)} = \exp[H(x, p) - H(x', p')], \tag{A.27}$$

the Hamiltonian is theoretically conserved, so only numerical errors exist in (A.26). Therefore, the content of exp in (A.27) is almost 0, and r is close to 1. For this reason, the acceptance rate is very high, which is almost 100%.

The algorithm of the HMC method is defined by executing the following transition repeatedly such that the joint probability distribution (A.22) is a stationary distribution:

Step 1 Generate independent standard normal random numbers p.

Step 2 Transition x, p using the leapfrog method (A.19), then take the candidate point x', p'.

Step 3 Accept the state x' suggested by the probability $\min[1, r]$ of (A.26) and (A.27), otherwise stay current state x.

Because (x, p) sampled by the transition of Steps 1 to 3 is the sample from $F(x, p)$, and $P(x)$ and $f(p)$ are independent of each other, we can regard the random number x from $F(x, p)$ as the sample from $P(x)$.

The comprehensive explanation for the HMC method can be found in Duane et al. [18] and Neal [57]. In a practical analysis by the HMC method, we speed up convergence adjusting automatically the variance when generating p, the number of updates L, and the step size ϵ by the NUTS (No-U-Turn Sampler) method [34]. The NUTS method is a method to adjust automatically L and ϵ in process of the HMC method.

Appendix B

Derivation and Specification of Model

B.1 Derivation of ODEs

We derive the ODEs in the case that the hazard rates are n-dimensional. First, let us consider the expectation involving the fixed leg,

$$V_f^i(t,T) = E^Q\left[e^{-\int_t^T \lambda_{is}ds}|\mathcal{G}_t\right] = e^{A_{i0}(\tau)-\sum_{k=1}^n A_{ik}(\tau)\lambda_{kt}} = V^i, \quad \tau = T-t. \tag{B.1}$$

The partial derivatives of (B.1) with respect to each variable are calculated as

$$\begin{cases} V_t^i = V^i\{-A_{i0}'(\tau) + \sum_{k=1}^n A_{ik}'(\tau)\lambda_k\} \\ V_{\lambda_k}^i = V^i\{-A_{ik}(\tau)\} \\ V_{\lambda_k\lambda_k}^i = V^i A_{ik}^2(\tau) \end{cases}. \tag{B.2}$$

Substituting each formula in (B.2) for the PDEs,

$$\begin{aligned} &- A_{i0}'(\tau) + \sum_{k=1}^n A_{ik}'(\tau)\lambda_k \\ &+ \sum_{k=1}^n [(s_k + \sum_{j=1}^n \gamma_{kj}\lambda_j)\{-A_{ik}(\tau)\} + \frac{1}{2}\sigma_k^2\lambda_k A_{ik}^2(\tau)] \\ &- \lambda_i = 0, \quad i = 1,\dots,n, \end{aligned} \tag{B.3}$$

and arranging the terms in (B.3) with respect to the terms proportional to $\lambda_j, j = 1,\dots,n$ and the constant terms, we have

$$\begin{aligned} &- A_{i0}'(\tau) - \sum_{k=1}^n s_k A_{ik}(\tau) \\ &+ \sum_{j=1}^n \{A_{ij}'(\tau) - \sum_{k=1}^n \gamma_{kj}A_{ik}(\tau) + \frac{1}{2}\sigma_j^2 A_{ij}^2(\tau) - \delta_{ij}\}\lambda_j = 0, \quad i = 1,\dots,n. \end{aligned} \tag{B.4}$$

As the coefficients of each term in (B.4) need to be zero, we obtain the ODEs:

$$\begin{cases} A_{i0}'(\tau) = -\sum_{k=1}^n s_k A_{ik}(\tau) \\ A_{ij}'(\tau) = \sum_{k=1}^n \gamma_{kj}A_{ik}(\tau) - \frac{1}{2}\sigma_j^2 A_{ij}^2(\tau) + \delta_{ij} \end{cases}, \quad i = 1,\dots,n, j = 1,\dots,n. \tag{B.5}$$

Next, let us consider the expectation involving the protection leg,

$$\begin{aligned} V_p^i(t,T) &= E^Q\left[\lambda_{iT} e^{-\int_t^T \lambda_{is} ds} | \mathcal{G}_t\right] \\ &= e^{A_{i0}(\tau) - \sum_{k=1}^n A_{ik}(\tau)\lambda_{kt}} \{a_{i0}(\tau) + \sum_{k=1}^n a_{ik}(\tau)\lambda_{kt}\} \\ &= B^i\{a_{i0}(\tau) + \sum_{k=1}^n a_{ik}(\tau)\lambda_{kt}\}, \quad \tau = T - t. \end{aligned} \tag{B.6}$$

The partial derivatives of (B.6) with respect to each variable are calculated as

$$\begin{cases} V_t^i = B^i\{-A'_{i0}(\tau)a_{i0}(\tau) - A'_{i0}(\tau)\sum_{k=1}^n a_{ik}(\tau)\lambda_k + a_{i0}(\tau)\sum_{k=1}^n A'_{ik}(\tau)\lambda_k \\ \quad + \sum_{k=1}^n\sum_{l=1}^n A'_{ik}(\tau)a_{il}(\tau)\lambda_k\lambda_l - a'_{i0}(\tau) - \sum_{k=1}^n a'_{ik}(\tau)\lambda_k\} \\ V_{\lambda_k}^i = B^i\{-A_{ik}(\tau)a_{i0}(\tau) - A_{ik}(\tau)\sum_{l=1}^n a_{il}(\tau)\lambda_l + a_{ik}(\tau)\} \\ V_{\lambda_k\lambda_k}^i = B^i\{A_{ik}^2(\tau)a_{i0}(\tau) + A_{ik}^2(\tau)\sum_{l=1}^n a_{il}(\tau)\lambda_l - 2A_{ik}(\tau)a_{ik}(\tau)\} \end{cases}. \tag{B.7}$$

Substituting each formula in (B.7) for the PDEs,

$$\begin{aligned} &- A'_{i0}(\tau)a_{i0}(\tau) - A'_{i0}(\tau)\sum_{k=1}^n a_{ik}(\tau)\lambda_k + a_{i0}(\tau)\sum_{k=1}^n A'_{ik}(\tau)\lambda_k \\ &+ \sum_{k=1}^n\sum_{l=1}^n A'_{ik}(\tau)a_{il}(\tau)\lambda_k\lambda_l - a'_{i0}(\tau) - \sum_{k=1}^n a'_{ik}(\tau)\lambda_k \\ &+ \sum_{k=1}^n[(s_k + \sum_{j=1}^n \gamma_{kj}\lambda_j)\{-A_{ik}(\tau)a_{i0}(\tau) - A_{ik}(\tau)\sum_{l=1}^n a_{il}(\tau)\lambda_l + a_{ik}(\tau)\} \\ &+ \frac{1}{2}\sigma_k^2\lambda_k\{A_{ik}^2(\tau)a_{i0}(\tau) + A_{ik}^2(\tau)\sum_{l=1}^n a_{il}(\tau)\lambda_l - 2A_{ik}(\tau)a_{ik}(\tau)\}] \\ &- \lambda_i\{a_{i0}(\tau) + \sum_{k=1}^n a_{ik}(\tau)\lambda_k\} = 0, \quad i = 1,\ldots,n, \end{aligned} \tag{B.8}$$

and substituting $A'_{i0}(\tau)$ and $A'_{ij}(\tau)$ in (B.5) for the formula (B.8), arranging the terms in (B.8) with respect to the terms proportional to $\lambda_j, j = 1,\ldots,n$ and the constant terms, we have

$$\begin{aligned} &- a'_{i0}(\tau) + \sum_{k=1}^n s_k a_{ik}(\tau) \\ &+ \sum_{j=1}^n\{-a'_{ij}(\tau) + \sum_{k=1}^n \gamma_{kj}a_{ik}(\tau) - \sigma_j^2 A_{ij}(\tau)a_{ij}(\tau)\}\lambda_j \\ &+ a_{i0}(\tau)\sum_{j=1}^n \delta_{ij}\lambda_j - a_{i0}(\tau)\lambda_i = 0, \quad i = 1,\ldots,n. \end{aligned} \tag{B.9}$$

Incidentally, all second-order terms of λ offset and disappear. As

$$\sum_{j=1}^n \delta_{ij}\lambda_j = \lambda_i$$

and the coefficients of each term in (B.9) need to be zero, we obtain the ODEs:

$$\begin{cases} a'_{i0}(\tau) = \sum_{k=1}^{n} s_k a_{ik}(\tau) \\ a'_{ij}(\tau) = \sum_{k=1}^{n} \gamma_{kj} a_{ik}(\tau) - \sigma_j^2 A_{ij}(\tau) a_{ij}(\tau) \end{cases}, \quad i = 1, \ldots, n, j = 1, \ldots, n.$$

B.2 Derivation of ODE analysis solutions

We derive the ODE analysis solutions in the case that the hazard rate is one-dimensional. In the fixed leg, the ODEs are as follows:

$$\begin{cases} A'_0(\tau) = -sA_1(\tau), \quad A_0(0) = 0 \\ A'_1(\tau) = \gamma A_1(\tau) - \frac{1}{2}\sigma^2 A_1^2(\tau) + 1, \quad A_1(0) = 0 \end{cases}. \tag{B.10}$$

The ODE with respect to $A_1(\tau)$ in the second formula of (B.10)

$$\frac{dA_1}{d\tau} = -\frac{\sigma^2}{2} A_1^2 + \gamma A_1 + 1 \tag{B.11}$$

is the Riccati type ODE. Since the coefficients of each term in the RHS of (B.11) are constants, it can be solved as the separation of variables as follows:

$$\frac{2}{\sigma^2} \int \frac{dA_1}{(A_1 - d_+)(A_1 - d_-)} = -\int d\tau, \tag{B.12}$$

where $d_\pm$ are the solutions of $\frac{\sigma^2}{2} A_1^2 - \gamma A_1 - 1 = 0$:

$$d_\pm = \frac{\gamma \pm \sqrt{\gamma^2 + 2\sigma^2}}{\sigma^2}.$$

Decomposing the LHS of (B.12) into partial fractions and integrating the RHS of (B.12), we have

$$\frac{1}{\sqrt{\gamma^2 + 2\sigma^2}} \int \left(\frac{1}{A_1 - d_+} - \frac{1}{A_1 - d_-} \right) dA_1 = -\tau + C_1. \tag{B.13}$$

Then, integrating the LHS of (B.13), we have

$$\ln \left| \frac{A_1 - d_+}{A_1 - d_-} \right| = -\sqrt{\gamma^2 + 2\sigma^2}\tau + C_2, \tag{B.14}$$

where $C_\bullet$ represents a constant of integration. Solving (B.14) for $A_1(\tau)$ and applying $A_1(0) = 0$, we have

$$A_1(\tau) = \frac{\sinh \xi\tau}{\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau}, \tag{B.15}$$

where

$$\xi = \frac{1}{2}\sqrt{\gamma^2 + 2\sigma^2}.$$

In contrast, $A_0(\tau)$ in the first formula of (B.10) can be found by integrating $A_1(\tau)$ in (B.15) with the current maturity τ:

$$\begin{aligned} A_0(\tau) &= -s \int A_1(\tau) d\tau \\ &= -s \left\{ \int \frac{e^{2\xi\tau}}{\xi + \frac{\gamma}{2} + (\xi - \frac{\gamma}{2}) e^{2\xi\tau}} d\tau - \int \frac{1}{\xi + \frac{\gamma}{2} + (\xi - \frac{\gamma}{2}) e^{2\xi\tau}} d\tau \right\} \\ &= -s [\frac{1}{2\xi(\xi - \frac{\gamma}{2})} \ln \left| \xi + \frac{\gamma}{2} + (\xi - \frac{\gamma}{2}) e^{2\xi\tau} \right| \\ &\quad - \frac{1}{2\xi(\xi + \frac{\gamma}{2})} \left\{ 2\xi\tau - \ln \left| \xi + \frac{\gamma}{2} + (\xi - \frac{\gamma}{2}) e^{2\xi\tau} \right| \right\}] + C_3 \\ &= \frac{2s}{\sigma^2} \ln \frac{\frac{1}{2} e^{-\frac{\gamma}{2}\tau}}{\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau} + C_3, \end{aligned} \tag{B.16}$$

where we use the following integral formula (B.17) for the integral of the second term in the second equal sign in the formula (B.16).

$$\int \frac{dx}{a + be^{cx}} = \frac{1}{ac}(cx - \ln |a + be^{cx}|) \tag{B.17}$$

Applying $A_0(0) = 0$ to (B.16), we have

$$A_0(\tau) = \frac{2s}{\sigma^2} \ln \frac{\xi e^{-\frac{\gamma}{2}\tau}}{\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau}.$$

In the protection leg, the ODEs are as follows:

$$\begin{cases} a_0'(\tau) = s a_1(\tau), \quad a_0(0) = 0 \\ a_1'(\tau) = \gamma a_1(\tau) - \sigma^2 A_1(\tau) a_1(\tau), \quad a_1(0) = 1 \end{cases}. \tag{B.18}$$

Since we already evaluate $\int A_1(\tau) d\tau$ in (B.16), $a_1(\tau)$ can be found by using the result in the second formula of (B.18):

$$\ln a_1(\tau) = \int \{\gamma - \sigma^2 A_1(\tau)\} d\tau = \ln \frac{1}{4(\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau)^2} + C_4. \tag{B.19}$$

Applying $a_1(0) = 1$ to (B.19), we have

$$a_1(\tau) = \frac{\xi^2}{(\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau)^2}. \tag{B.20}$$

In contrast, $a_0(\tau)$ can be found by integrating $a_1(\tau)$ in (B.20) with the current maturity τ:

$$\begin{aligned} a_0(\tau) &= s \int a_1(\tau) d\tau = s \int \frac{\xi^2 (\cosh^2 \xi\tau - \sinh^2 \xi\tau)}{(\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau)^2} d\tau \\ &= s \frac{\sinh \xi\tau}{\xi \cosh \xi\tau - \frac{\gamma}{2} \sinh \xi\tau} + C_5, \end{aligned} \tag{B.21}$$

where we use the following relational expression of hyperbolic function (B.22) in the second equal sign in the formula (B.21):

$$\cosh^2 x - \sinh^2 x = 1. \tag{B.22}$$

Applying $a_0(0) = 0$ to (B.21), we have

$$a_0(\tau) = s A_1(\tau).$$

The detailed derivation of the ODE analytical solutions can be found in Duffie and Singleton [22].

B.3 Specification of conventional standard CDS model

We assume the CIR model for a hazard rate (system equation) under the real probability P as follows:

$$d\lambda_t = (v + \alpha\lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t^P, \tag{B.23}$$

where $v > 0, \alpha < 0, \sigma > 0$. A Brownian motion is denoted by W_t^P under the real probability P. A CDS spread (observation equation) is given by

$$CDS_t^{mod,l} = \frac{(1-\theta)\Delta\sum_{k=1}^{\xi l} D(t,s_k)V_p(t,s_k)}{1/g\sum_{j=1}^{gl} D(t,t_j)V_f(t,t_j)}, \tag{B.24}$$

where

$$\begin{cases} V_f(t,s) = E^Q\left[e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right] & \text{(fixed leg)} \\ V_p(t,s) = E^Q\left[\lambda_s e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right] & \text{(protection leg)} \end{cases}. \tag{B.25}$$

The system equation (B.23) and observation equation (B.24) compose a state space model. Since we need to know the behavior of hazard rate (system equation) under the risk neutral probability Q to find the CDS spread, we change the probability measure. Let us define the market price of credit risk as

$$\psi_t = \phi_1\sqrt{\lambda_t} + \frac{\phi_2}{\sqrt{\lambda_t}}. \tag{B.26}$$

In order to change the probability measure using the Girsanov theorem, we define the Radon-Nikodym derivative as follows:

$$\begin{cases} Z_T = \dfrac{dQ}{dP} = \exp\left[-\displaystyle\int_0^T \psi_u dW_u^P - \frac{1}{2}\int_0^T \psi_u^2 du\right] \\ W_t^Q = W_t^P + \displaystyle\int_0^t \psi_u du \end{cases}, \tag{B.27}$$

where the maximum time horizon is theoretically denoted as T. In $0 \le t \le T$, the Radon-Nikodym derivative process is $Z_t = E^P[Z_T|\mathcal{G}_t]$. Then, $W_t^Q, 0 \le t \le T$ is a Brownian motion under the probability measure Q.

Accordingly, The SDE of hazard rate represented in Q is written as

$$\begin{aligned} d\lambda_t &= (v + \alpha\lambda_t)dt + \sigma\sqrt{\lambda_t}(dW_t^Q - \psi_t dt) \\ &= \{v - \sigma\phi_2 + (\alpha - \sigma\phi_1)\lambda_t\}dt + \sigma\sqrt{\lambda_t}dW_t^Q \\ &= (s + \gamma\lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t^Q \end{aligned} \tag{B.28}$$

by substituting (B.26) and (B.27) for (B.23), where

$$\begin{cases} s = v - \sigma\phi_2 \\ \gamma = \alpha - \sigma\phi_1 \end{cases}.$$

Incidentally, we set $\phi_2 = 0$.

The expectations of (B.25) in the CDS spread formula (B.24) are the solution $V(t,\lambda)$ of the PDE

$$V_t(t,\lambda) + (s + \gamma\lambda)V_\lambda(t,\lambda) + \frac{1}{2}\sigma^2\lambda V_{\lambda\lambda}(t,\lambda) = \lambda V(t,\lambda) \tag{B.29}$$

from (B.28), where the corresponding boundary conditions are provided as

$$V(T,\lambda) = \begin{cases} 1 & \text{(fixed leg)} \\ \lambda & \text{(protection leg)} \end{cases}.$$

We can obtain the ODEs with respect to the coefficients A_0, A_1, a_0, and a_1 of the exponential affine model

$$\begin{cases} V_f(t,s) = E^Q\left[e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right] = e^{A_0(\tau)-A_1(\tau)\lambda_t} \\ V_p(t,s) = E^Q\left[\lambda_s e^{-\int_t^s \lambda_u du}|\mathcal{G}_t\right] = e^{A_0(\tau)-A_1(\tau)\lambda_t}\{a_0(\tau)+a_1(\tau)\lambda_t\} \end{cases}, \quad \tau = s-t, \tag{B.30}$$

by substituting (B.30) for the PDE (B.29). In the fixed leg, the ODEs are as follows:

$$\begin{cases} A_0'(\tau) = -sA_1(\tau), \quad A_0(0) = 0 \\ A_1'(\tau) = \gamma A_1(\tau) - \frac{1}{2}\sigma^2 A_1^2(\tau) + 1, \quad A_1(0) = 0 \end{cases}. \tag{B.31}$$

In the protection leg, the ODEs are as follows:

$$\begin{cases} a_0'(\tau) = sa_1(\tau), \quad a_0(0) = 0 \\ a_1'(\tau) = \gamma a_1(\tau) - \sigma^2 A_1(\tau)a_1(\tau), \quad a_1(0) = 1 \end{cases}. \tag{B.32}$$

The analytical solutions of these ODEs (B.31) and (B.32) are as follows:

$$\begin{cases} A_0(\tau) = \dfrac{2s}{\sigma^2}\ln\dfrac{\xi e^{-\frac{\gamma}{2}\tau}}{\xi\cosh\xi\tau - \frac{\gamma}{2}\sinh\xi\tau} \\ A_1(\tau) = \dfrac{\sinh\xi\tau}{\xi\cosh\xi\tau - \frac{\gamma}{2}\sinh\xi\tau} \\ a_0(\tau) = sA_1(\tau) \\ a_1(\tau) = \dfrac{\xi^2}{(\xi\cosh\xi\tau - \frac{\gamma}{2}\sinh\xi\tau)^2} \end{cases},$$

where

$$\xi = \frac{1}{2}\sqrt{\gamma^2 + 2\sigma^2}.$$

B.4 Specification of interest rate model

We assume the CIR model for a risk-free interest rate (system equation) under the real probability P as follows:

$$dr_t = k_0(\theta_0 - r_t)dt + \sigma_0\sqrt{r_t}dW_{0t}^P, \tag{B.33}$$

where $k_0(>0)$ is a mean reversion speed, and $\theta_0(>0)$ is a mean reversion level under the real probability P. The volatility is denoted by $\sigma_0(>0)$. A Brownian motion is denoted by W_{0t}^P under the real probability P, which is independent of W_{1t}^P, W_{2t}^P. A default-free discount bond price (observation equation) is given by

$$D(t,s) = E^Q\left[e^{-\int_t^s r_u du}|\mathcal{F}_t\right]. \tag{B.34}$$

The system equation (B.33) and observation equation (B.34) compose a state space model. Since we need to know the behavior of risk-free rate (system equation) under the risk neutral

probability Q to find the discount bond price, we change the probability measure. Let us define the market price of risk-free interest rate risk as

$$\psi_{0t} = \phi_{01}\sqrt{r_t} + \frac{\phi_{02}}{\sqrt{r_t}}, \tag{B.35}$$

as with the case of credit risk. In order to change the probability measure using the Girsanov theorem, we define the Radon-Nikodym derivative as follows:

$$\begin{cases} Z_T = \dfrac{dQ}{dP} = \exp\left[-\displaystyle\int_0^T \psi_{0u} dW_{0u}^P - \frac{1}{2}\int_0^T \psi_{0u}^2 du\right], \\ W_{0t}^Q = W_{0t}^P + \displaystyle\int_0^t \psi_{0u} du \end{cases} \tag{B.36}$$

where the maximum time horizon is theoretically denoted as T. In $0 \le t \le T$, the Radon-Nikodym derivative process is $Z_t = E^P[Z_T|\mathcal{G}_t]$. Then, $W_{0t}^Q, 0 \le t \le T$ is a Brownian motion under the probability measure Q.

Accordingly, The SDE of risk-free interest rate represented in Q is written as

$$\begin{aligned} dr_t &= k_0(\theta_0 - r_t)dt + \sigma_0\sqrt{r_t}(dW_{0t}^Q - \psi_{0t}dt) \\ &= (k_0 + \sigma_0\phi_{01})\left(\frac{k_0\theta_0 - \sigma_0\phi_{02}}{k_0 + \sigma_0\phi_{01}} - r_t\right)dt + \sigma_0\sqrt{r_t}dW_{0t}^Q \\ &= \hat{k_0}(\hat{\theta_0} - r_t)dt + \sigma_0\sqrt{r_t}dW_{0t}^Q \end{aligned} \tag{B.37}$$

by substituting (B.35) and (B.36) for (B.33), where

$$\begin{cases} \hat{k_0} = k_0 + \sigma_0\phi_{01} \\ \hat{\theta_0} = \dfrac{k_0\theta_0 - \sigma_0\phi_{02}}{k_0 + \sigma_0\phi_{01}}. \end{cases}$$

Incidentally, we set $\phi_{02} = 0$ as with the case of credit risk.

The expectation of (B.34) is the solution $D(t,r)$ of the PDE

$$D_t(t,r) + \hat{k_0}(\hat{\theta_0} - r)D_r(t,r) + \frac{1}{2}\sigma_0^2 r D_{rr}(t,r) = rD(t,r) \tag{B.38}$$

from (B.37), where the corresponding boundary condition is provided as

$$D(T,r) = 1.$$

We can obtain the ODEs with respect to the coefficients B_0 and B_1 of the exponential affine model

$$D(t,s) = E^Q\left[e^{-\int_t^s r_u du}|\mathcal{F}_t\right] = e^{B_0(\tau) - B_1(\tau)r_t}, \quad \tau = s - t, \tag{B.39}$$

by substituting (B.39) for the PDE (B.38) as follows:

$$\begin{cases} B_0'(\tau) = -\hat{k_0}\hat{\theta_0}B_1(\tau), \quad B_0(0) = 0 \\ B_1'(\tau) = -\hat{k_0}B_1(\tau) - \dfrac{1}{2}\sigma_0^2 B_1^2(\tau) + 1, \quad B_1(0) = 0 \end{cases}. \tag{B.40}$$

The analytical solutions of these ODEs (B.40) are as follows:

$$\begin{cases} B_0(\tau) = \dfrac{2\hat{k_0}\hat{\theta_0}}{\sigma_0^2}\ln\dfrac{\nu e^{\frac{\hat{k_0}}{2}\tau}}{\nu\cosh\nu\tau + \frac{\hat{k_0}}{2}\sinh\nu\tau} \\ B_1(\tau) = \dfrac{\sinh\nu\tau}{\nu\cosh\nu\tau + \frac{\hat{k_0}}{2}\sinh\nu\tau} \end{cases},$$

where

$$\nu = \frac{1}{2}\sqrt{\hat{k_0}^2 + 2\sigma_0^2}.$$

The other specifications of the interest rate model can be found in Brigo and Mercurio [8].

Bibliography

[1] Aboulaich, R., M. L. Hadji and A. Jraifi, 2013, Option Pricing with Constant Elasticity of Variance (CEV) Model, *Applied Mathematical Sciences*, **7**, 5443-5456.

[2] Adrian, T. and M. K. Brunnermeier, 2016, CoVaR, *The American Economic Review*, **106**, 1705-1741.

[3] Bali, T., M. Heidari and L. Wu, 2009, Predictability of Interest Rates and Interest-Rate Portfolios, *Journal of Business and Economic Statistics*, **27**, 517-527.

[4] Beckers, S., 1980, The Constant Elasticity of Variance Model and Its Implications for Option Pricing, *The Journal of Finance*, **35**, 661-673.

[5] Benmenzer, G., E. Gobet and C. Jerusalem, 2007, Arbitrage Free Cointegrated Models in Gas and Oil Future Markets, hal-00200422.

[6] Black, F. and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Politcal Economy*, **81**, 637-659.

[7] Blanco, R., S. Brennan and I. W. Marsh, 2005, An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps, *Journal of Finance*, **60**, 2255-2281.

[8] Brigo, D. and F. Mercurio, 2007, *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit, Second Edition*, Springer Finance.

[9] Brunel, V., 2004, Minimal Models for Credit Risk: an Information Theory Approach, Working paper.

[10] Chan, K. C., G. A. Karolyi, F. A. Longstaff, A. B. Sanders, 1992, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, **47**, 1209-1227.

[11] Cox, J. C., 1975, Notes on Option Pricing I: Constant Elasticity of Variance Diffusions, Working Paper Stanford University.

[12] Cox, J. C. and S. A. Ross, 1976, The Valuation of Options for Alternative Stochastic Processes, *Journal of Financial Economics*, **3**, 145-166.

[13] Cox, J. C., J. E. Ingersoll and S. A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica*, **53**, 385-407.

[14] Crank, J. and P. Nicolson, 1947, A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat Conduction Type, *Proc. Camb. Phil. Soc.*, **43**, 50-67.

[15] Credit Suisse Financial Products, 1997, *CreditRisk+: A Credit Risk Management Framework*.

[16] Crosbie, P. and J. Bohn, 2003, Modeling Default Risk, *Moody's KMV technical document.*

[17] Duan, J. C. and A. Fulop, 2009, Estimating the Structural Credit Risk Model when Equity Prices are Contaminated by Trading Noises, *Journal of Econometrics*, **150**, 288-296.

[18] Duane, S., A. D. Kennedy, B. J. Pendleton and D. Roweth, 1987, Hybrid Monte Carlo, *Physics Letters B*, **195**, 216-222.

[19] Duffie, D., 2005, Credit Risk Modeling with Affine Processes, *Journal of Banking and Finance*, **29**, 2751-2802.

[20] Duffie, D., J. Ma and J. Yong, 1995, Black's Consol Rate Conjecture, *The Annals of Applied Probability*, **5**, 356-382.

[21] Duffie, D. and K. Singleton, 1999, Modeling Term Structures of Defaultable Bonds, *Review of Financial Studies*, **12**, 687-720.

[22] Duffie, D. and K. Singleton, 2003, *Credit Risk: Pricing, Measurement, and Management*, Princeton University Press.

[23] Engle, R. F. and C. W. J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, **55**, 251-276.

[24] Engle, R. F. and K. Sheppard, 2001, Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, NBER Working Paper.

[25] Filiz, I. O., X. Guo, J. Morton and B. Sturmfels, 2012, Graphical Models for Correlated Defaults, *Mathematical Finance*, **22**, 621-644.

[26] Gear, C. W., 1971, *Numerical Initial Value Problems in Ordinary Differential Equations* (Englewood Cliffs, N.J.: Prentice-Hall).

[27] Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin, 2013, *Bayesian Data Analysis, Third Edition*, Chapman and Hall/CRC.

[28] Gordon, N., D. Salmond and A. Smith, 1993, A Novel Approach to Non-Linear and Non-Gaussian Bayesian State Estimation, IEE Proceedings, F 140, 107-113.

[29] Guidolin, M., F. Melloni and M. Pedio, 2019, A Markov Switching Cointegration Analysis of the CDS-Bond Basis Puzzle, BAFFI CAREFIN Centre Research Paper No. 2019-121.

[30] Gupton, G. M., C. C. Finger and M. Bhatia, 1997, *CreditMetrics: Technical Document*, Morgan Guaranty Trust Company.

[31] Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press, Princeton.

[32] Heston, S. L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *The Review of Financial Studies*, **6**, 327-343.

[33] Higashide, T., K. Asai, J. Gotoh and T. Fujita, 2020, A New Formulation of Pair's Portfolio Selection with First Passage Time, *Transactions of the Japan Society for Industrial and Applied Mathematics*, **30**, 194-225.

[34] Hoffman, M. and A. Gelman, 2014, The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, *Journal of Machine Learning Research*, **15**, 1593-1623.

[35] Hogan, M., 1993, Problems in Certain Two-Factor Team Structure Models, *The Annals of Applied Probability*, **3**, 576-581.

[36] Hukushima, K. and K. Nemoto, 1996, Exchange Monte Carlo Method and Application to Spin Glass Simulations, *Journal of the Physical Society of Japan*, **65**, 1604-1608.

[37] Iba, Y., 2001, Extended Ensemble Monte Carlo, *International Journal of Modern Physics*, **C12**, 623-656.

[38] In't Hout, K., 2017, *Numerical Partial Differential Equations in Finance Explained: An Introduction to Computational Finance*, Macmillan Publishers Ltd.

[39] Jarrow, R., H. Li, X. Ye and M. Hu, 2019, Exploring Mispricing in the Term Structure of CDS Spreads, *Review of Finance*, **23**, 161-198.

[40] Jarrow, R. and S. Turnbull, 2000, The Intersection of Market and Credit Risk, *Journal of Banking and Finance*, **24**, 271-299.

[41] Jiao, Y., C. Ma and S. Scotti, 2017, Alpha-CIR Model with Branching Processes in Sovereign Interest Rate Modeling, *Finance and Stochastics*, **21**, 789-813.

[42] Johansen, S., 1988, Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, **12**, 231-254.

[43] Jones, C. S., 2003, Nonlinear Mean Reversion in the Short-Term Interest Rate, *The Review of Financial Studies*, **16**, 793-843.

[44] Kato, K., 2016, Long-Range Ising Model for Credit Portfolios with Heterogeneous Credit Exposures, *Physica A*, **462**, 1103-1119.

[45] Kato, K. and N. Nakamura, 2022, Cointegration Analysis of Hazard Rates and CDSs: Applications to Pairs Trading Strategy, Submitted.

[46] Kato, K. and N. Nakamura, 2022, PDE-based Bayesian Inference of CEV Dynamics for Credit Risk in Stock Prices, Submitted.

[47] Kitabayashi, T. and N. Nakamura, 2021, IPDE-Based Bayesian Statistical Inference for CIR Interest Rate Model with Poisson Jump, *Proceedings of the 54-th JAFEE meeting*, **54**, 24-35.

[48] Kitsukawa, K., S. Mori and M. Hisakado, 2006, Evaluation of Tranche in Securitization and Long-Range Ising Model, *Physica A*, **368**, 191-206.

[49] Koenker, R. and G. Bassett, Jr., 1978, Regression Quantiles, *Econometrica*, **46**, 33-50.

[50] Lando, D., 1998, On Cox Processes and Credit Risky Securities, *Review of Derivatives Research*, **2**, 99-120.

[51] Li, D. X., 2000, On Default Correlation: a Copula Function Approach, *Journal of Fixed Income*, **9**, 43-54.

[52] Marsaglia, G., 2003, Xorshift RNGs, *Journal of Statistical Software*, **8**, 1-6.

[53] Merton, R., 1974, On the Pricing of Corporate Debt: the Risk Structure of Interest Rates, *Journal of Finance*, **29**, 449-470.

[54] Molins, J. and E. Vives, 2005, Long Range Ising Model for Credit Risk Modeling in Homogeneous Portfolios, *AIP Conference Proceedings*, **779**, 156-161.

[55] Molins, J. and E. Vives, 2015/2016, Model Risk on Credit Risk, *Risk and Decision Analysis*, **6**, 65-78.

[56] Nakajima, K. and K. Ohashi, 2012, A Cointegrated Commodity Pricing Model, *The Journal of Futures Markets*, **32**, 995-1033.

[57] Neal, R., 2011, *MCMC Using Hamiltonian Dynamics, In Handbook of Markov Chain Monte Carlo*, edited by S. Brooks, A. Gelman, G. L. Jones, and X. Meng, 116-62. Chapman; Hall/CRC.

[58] Nishimori, H., 2001, *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press.

[59] Pan, J. and K. Singleton, 2008, Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads, *The Journal of Finance*, **63**, 2345-2384.

[60] Pardoux, E. and S. G. Peng, 1990, Adapted Solution of a Backward Stochastic Differential Equation, *Systems and Control Letters*, **14**, 55-61.

[61] Phillips, P. C. B. and S. Ouliaris, 1990, Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, **58**, 165-193.

[62] Pitt, M., 2002, Smooth Particle Filters Likelihood Evaluation and Maximisation, Working Paper, University of Warwick.

[63] Roncalli, T., 2013, *Introduction to Risk Parity and Budgeting*, Chapman and Hall/CRC.

[64] Sankaran, M., 1959, On the Non-Central Chi-Square Distribution, *Biometrika*, **46**, 235-237.

[65] Sankaran, M., 1963, Approximations to the Non-Central Chi-Square Distribution, *Biometrika*, **50**, 199-204.

[66] Schneider, P., L. Sogner and T. Veza, 2010, The Economic Role of Jumps and Recovery Rates in the Market for Corporate Default Risk, *Journal of Financial and Quantitative Analysis*, **45**, 1517-1547.

[67] Shreve, S., 2004, *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer Finance.

[68] Stan Development Team, 2021, *Stan Modeling Language Users Guide and Reference Manual, Version 2.28*. `https://mc-stan.org`

[69] Tsay, R. S., 2010, *Analysis of Financial Time Series (3rd Edition)*, John Wiley & Sons, New York.

[70] Tourin, A. and R. Yan, 2013, Dynamic Pairs Trading using the Stochastic Control Approach, *Journal of Economic Dynamics and Control*, **37**, 1972-1981.

[71] Vasicek, O. A., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, **5**, 177-188.

[72] Vasicek, O. A., 1991, Limiting Loan Loss Probability Distribution, KMV Corporation.

[73] Vidyamurthy, G., 2004, *Pairs Trading: Quantitative Methods and Analysis*, Wiley Finance, Vol.217, John Wiley & Sons, New York.

[74] Watanabe, S., 2010, Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory, *Journal of Machine Learning Research*, **11**, 3571-3594.