LUND UNIVERSITY

SCHOOL OF ECONOMICS AND MANAGEMENT

Research on Credit Risk Measurement of China's Listed Companies with KMV Model

by

Xiuyi Chen & Tong Liu

May 2020

Master Thesis in Finance

Supervisor: Jens Forssbaeck

Abstract

This thesis takes 200 Chinese listed companies as examples within ten years from 2009 to 2018, of which 100 are ST companies and the other 100 are non-ST companies. ST company is a company that has financial problems and then implemented with special treatment by the China Securities Regulatory Commission. The traditional credit risk measurement model, Z-Score model is chosen to be compared with the KMV model to test whether the KMV model is more suitable for the Chinese financial market. Through comparative analysis, we can judge whether KMV, which has been highly praised in recent years, has the apparent ability of identification and prediction for defaulting companies. Then, with the ROC curve, the ability to identify and predict of the two models can be further examined. A more accurate credit risk measurement model can help investors identify the company's financial status and avoid property losses caused by unnecessary risks as much as possible. The experimental results show that, for the Chinese stock market, the differentiation ability of the KMV model is more suitable than the Z-Score model, and it is also more suitable as a prediction model. Finally, if the KMV model is to be vigorously promoted in the Chinese market, it must overcome its shortcomings. Therefore, suggestions for improvement are listed in this paper.

Keywords: Credit risk measurement model, KMV model, Z-Score model, ST companies, ROC curve.

Acknowledgments

We would like to express our gratitude to our supervisor Jens Forssbaeck, for his great patience and valuable suggestions during the process of the thesis writing. Thank you, Jens, your help was invaluable. At the same time, we want to thank our family. We will never forget their continued support and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DPT	Default Point
DD	Distance to Default
EDF	Expected Default Frequency
CR	CreditRisk+
VaR	Value at Risk
CPV	CreditPortfolio View
ROC	Receiver Operating Characteristic
AUC	Area Under Curve

1 Introduction

Credit risk is the primary risk faced by commercial banks as credit intermediaries in their operation and management. In the increasingly fierce market competition and the increasingly turbulent macro-environment, only if the commercial banks grow to improve their ability to monitor credit risk, they will have a chance and ability to survive. As listed companies become more critical financing entities in the capital market, once their credit risk occurs, the scope of other industries they affect will be broader and deeper. Therefore, whether it is for commercial banks or investors in the capital market, the credit risk of listed companies has an important impact on banks or investors. The occurrence of credit risk does not happen overnight, but there is still a gradual change process. It is considered that strengthening the management of credit risk can control the occurrence of credit risk. However, the premise of credit risk management is the measurement and prediction of credit risk, which has become our first concern and needs to be solved.

1.1 Background

The World Bank (2020) shows that there were 3584 listed companies in China's capital market at the end of 2018. For many years, the futures trading volume of Chinese commodities has been one of the largest commodities in the world. However, since 2008, the financial crisis caused by the US subprime mortgage crisis has continued to affect the world. It has changed the real economy of the global market to varying degrees. In the fierce competition environment and the impact of various potential risks, the financial risks of listed companies in China are exposed more prominently.

The China Securities Regulatory Commission (CSRC) carried out special treatment for some listed companies (marked as ST companies) due to the failure of corporate financial risk management to warm them to improve. In other words, these companies have suffered many losses for many years, according to the annual reports of a part of listed companies. The reason why ST companies exist is that the delisting mechanism is not accomplished, and

China's delisting mechanism in the stock market is still in the process of establishment. Therefore, bankruptcy and liquidation of listed companies still account for only a small portion of the total number of stocks in the market. In the mature stock markets like western developed countries, the stocks' delisting has become a market mechanism, and it is an essential part of the "game rules" in the entire stock market. On the contrary, the companies which have severe losses still can continue to be listed or even reorganizing assets after listing in China.

On March 7, 2014, Shanghai Chaori Solar Technology Co., Ltd. (002506.SZ) defaulted on coupon payment. It is the first default on the onshore RMB corporate bond market since the establishment of the Chinese stock market (Cao, 2014). As a result of this default, Chinese retail investors gradually realized that they need to be more vigilant when they are investing. Well-functioning capital markets require investors not to blindly pursue high-return projects and ignore risks. When investors know how to allocate resources based on risk-yield ratios, Chinese capital markets will become more efficient, which will be a long process. According to Bloomberg data (2020), at the end of October 2019, 45 companies have defaulted their debts within ten months, and more than 40 companies in 2018, which was the highest number of defaults on Chinese debt in history. Therefore, in the gradually healthy and transparent Chinese stock market, investors and institutions need to find an effective method to measure and predict the company's credit risk. In other words, if Chinese companies want to gain a foothold in a complex operating environment and fierce competition, they must understand their financial characteristics, fully control financial risks, and establish a credit risk early warning system.

Basel III advocates that it is an appropriate way for the banking industry to measure credit risk through the internal rating-based approach (IRB). The IRB approach relies on a bank's assessment of its counterparties and exposures to calculate capital requirements for credit risk. Such an approach has two primary aspects:

Risk sensitivity - Capital requirements based on internal estimates are more sensitive to the credit risk in the bank's portfolio of assets. Incentive compatibility - Banks must adopt better risk management techniques to control the credit risk in their portfolio to minimize regulatory capital (Basel III, 2001).

However, it is apparent that it is difficult to implement IRB in China’s market. China's credit rating system is deficient, and the use of IRB is still at a low level. For example, the internal rating of many large commercial banks in China is still based on the lender's financial records and other available information in the past three to five years, which makes it extremely difficult to predict the future trend of the credit quality of the lender. At present, when China's financial institutions use the internal rating system, they mostly focus on the customer's credit quality screening and preliminary risk warning mechanism, which is still in the qualitative stage. Instead, the most common methods for China’s financial institutions are expert systems and credit scoring, which are easy to be affected subjectively and will make the establishment of the credit transfer matrix face enormous obstacles.

Furthermore, when the market environment changes, there is a significant lag in risk prediction and credit assessment. Compared with the risk measurement of developed countries, there is also still a significant gap in the use of IRB by Chinese commercial banks, which is far from the Basel committee’s basic standard. At the moment, there are only six commercial banks that use IRB to measure credit risk in China, namely: China Construction Bank, Bank of Communications, Bank of China, Industrial and Commercial Bank of China, Agricultural Bank of China and China’s Merchants Bank.

From the perspective of the current financial market, the application of IRB in China's banking industry is of great benefit to improve the credit risk measurement method in China. At the same time, it is necessary. The KMV model is more suitable for China's financial market. First, the input data requirements of the model are not as high as those of the other modern models such as CreditMetrics model, CreditPortfolio View model, and CreditRisk+ model. Secondly, the model combines financial data and equity value volatility, which can help financial institutions to make forward-looking predictions of credit risk. Finally, the theoretical basis of the KMV model is derived from the Black-Scholes option pricing formula, which has a solid academic foundation and overcomes the subjective defects in other models.

It is also the reason why this paper chooses the KMV model for empirical analysis.

1.2 Research purpose and thesis question

In view of the above background, based on the actual data of Chinese listed companies, this paper attempts to evaluate the prediction ability of the KMV model through a progressive test method suitable for China's financial market. There are several questions that need to be clarified:

How is the identification and prediction ability of the KMV model? Is it a reliable choice for Chinese listed companies? In this thesis, the KMV model will be compared with the Z-Score model to determine which model is more reliable based on the data obtained. The data are composed of 100 ST companies and corresponding non-ST companies for 10 years. The Z-Score model has been widely used by scholars in different industries and has become one of benchmarks for the comparison of financial distress prediction models. Both the KMV model and the Z-Score model are landmark achievements in the history of financial distress prediction technology. However, they rely on different information. Z-Score model filters information from publicly disclosed financial statements of listed companies. In contrast, the KMV model uses forward-looking stock price information, but both are a comprehensive measure that summarizes some variables related to the financial distress of an enterprise. From the perspective of application value, this paper constructs a set of credit risk measurement model for listed companies that conforms to the characteristics of China's financial market. For commercial Banks, they can better identify the operation status of listed companies to avoid credit risks. For the listed companies themselves, they can timely carry out an early financial warning, adjust business decisions, strengthen the management of credit risk, and avoid falling into a credit crisis. For related enterprises, they can change their strategies in time according to the changes in the credit status of the listed company to reduce losses as much as possible.

Given the points above, the main question is defined:

How is the ability to differentiate and predict of the KMV model compared with the Z-Score model?

1.3 Outline of the thesis

The structure of this thesis is as follows. Section 2 outlines the theoretical background and literature review of our analysis. The focus is set on the adaptability analysis and historical credit risk measurement models. Section 3 describes the methodology and models used in the thesis together, including the KMV model and ROC curve. Section 4 examines the data availability, implementation of two models and preliminary presentation of results. Finally, Section 5 states the results from the analyzed models and summarizes and makes conclusions of our research.

2 Literature review

2.1 Development of credit risk measurement

Credit risk measurements have developed dramatically in the past two decades. As people pay more and more attention to credit risk management, more and more models have been established, which measure credit risk from different perspectives. From an academic point of view, credit risk measurement methods can be divided into traditional methods and modern methods.

2.1.1 Traditional credit risk measurement model

Before the 1970s, people tended to use qualitative analysis methods to measure credit risk, which usually based on corporate financial data of corporate credit risk and relying on the subjective judgment of some credit experts. Since then, there have been some corporate credit trends, which can be measured with mathematical-statistical models, but qualitative analysis is still preferred. The traditional credit risk measurement model has gone through three stages: expert systems, credit scoring, and qualitative-response model.

2.1.1.1 Expert Systems (5C)

This method mainly relies on the subjective judgment of credit experts to analyze the basic information characteristics of borrowers and conduct credit risk assessment. This method evaluates the borrower's repayment ability at maturity from the following five aspects: Character, capacity, capital, collateral, and cycle condition.

Although the expert systems method has been widely used in the credit risk measurement of commercial banks in the early stage, it is difficult to popularize in banks because it requires a high level of personal experience ability of credit experts. Because the judgment ability based on the experience of a good credit expert cannot be replaced or duplicated, the method also faces significant challenges of subjectivity and consistency.

2.1.1.2 Credit Scoring

Altman (1968) established the Z-Score model, which based on multivariate mathematical statistics. The model selects key financial indicators as variables and uses large amounts of data from bankrupt American companies as samples for empirical research. Altman used 22 key financial indicators to analyze 66 manufacturing companies in the United States, trying to reflect the credit status of sample companies through these key financial indicators. Finally, he selected five key financial indicators. After weighing the industry heterogeneity, different key financial indicators were given different weights to construct a linear regression function, that is, the Z-Score model. The expression is as follows:

$$Z\ (\zeta) = 1.2A + 1.4B + 3.3C + 0.6D + 0.999E \tag{2-1}$$

Where:

$Z(\zeta)$ =The Altman Z-Score

A = Working capital/total assets

B = Retained earnings/total assets

C = Earnings before interest and taxes (EBIT)/total assets

D = Market value of equity/book value of total liabilities

E = Sales/total assets

The mechanism of this model is that a commercial bank can judge the credit status of the borrower through the Z-Score obtained by the above formula. The higher z-value it gets, the higher credit quality the enterprise is. On the contrary, the credit quality of enterprises is poor. Generally speaking, when Z-Score is less than 1.81, the credit quality of the enterprise is poor, and the commercial banks that lend to it will face higher credit risks. When Z-Score is greater than 2.99, the credit quality of the enterprise is high, the probability of default is low, and the enterprise has a high willingness to pay and solvency, which can guarantee the security of the loan. However, the model also pointed out that when Z-Score makes a judgment, there is an "unknown area" (1.81 < Z-Score < 2.99).

Altman, Haldeman and Narayanan (1977) expanded Z-Score model to make up for the deficiency of credit risk assessment caused by the grey proportion in the model and constructed the ZETA model with seven financial index variables, which can be able to detect the bankruptcy with an accuracy rate reaching 90 percent two years before its occurrence. Through the empirical study of 24 retail enterprises and 29 manufacturing enterprises from 1970 to 1977, it is found that the ZETA model has better explanatory ability compared with the Z-Score model. Due to the low cost, simple form and good prediction effect of the ZETA model, many countries have further developed other credit risk measurement models based on this model.

2.1.1.3 Qualitative-response model

The premise of the multivariate discrimination model is that the variables are normally distributed and the covariance matrices between the groups should be equal. These two assumptions are often violated. Press and Wilson (1978) and Ohlson (1980) proposed the multivariate credit score model to eliminate this strict assumption. If the probability of variable index occurrence obeys the cumulative standard normal distribution, it is called the Probit model. If the probability of variable index occurrence obeys the cumulative logistic distribution, it is called the Logit model. Due to the nonlinear property of the function, the maximum likelihood method is often used to estimate its parameters (Kumar & Tan, 2005). The score in logit analysis can be directly interpreted as the probability of failure (cumulative logistic function ranges from 0 to 1). The relationship between the Z-Score and logit model can be described as follows: $P(Z)=\frac{1}{1+e^{-Z}}$. The Probit and Logit models predict the default probability of an enterprise based on various financial ratio indicators, thereby setting a risk warning line according to the size of credit risk and provide decision suggestions for decision-makers. The basic form of the two models is consistent, the main difference is only in the cumulative probability distribution of the transformation is different. The Probit model is the cumulative standard normal probability function, and the Logit model is the cumulative logistic probability distribution function. Benos and Papanastasopoulos (2007) proposed that Discriminant analysis does very well provided that the variables in every group follow a multivariate normal distribution, and covariance matrices are equal for every a priori defined group. However, empirical studies have shown that especially defaulted firms violate the normality assumption.

Although the multi-credit score models have been widely promoted and applied in practice, the calculation basis of these models relies on the financial data of accounting books, and it is impossible to respond to the rapid and subtle changes in the credit status of evaluation objects promptly.

By the late 1980s, successive international debt and financial crises had led financial institutions and regulators to reconsider traditional risk measures. Odom and Sharda (1990) were the first to apply the back propagation (BP) neural network in the study of the credit risk of enterprises and established the new financial crisis warning model. Wilson and Sharda (1994) pointed out the default probability can be predicted by the neural network method, and the comparison and empirical study between the neural network method and the multi-factor analysis method could be carried out, the results showed that the prediction is effective.

2.1.2 Modern credit risk measurement model

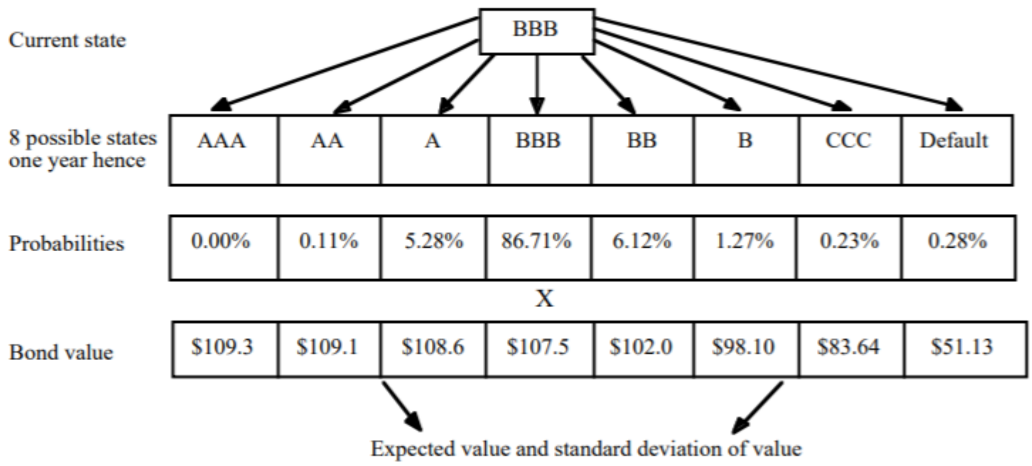
Although traditional credit risk measurement methods have obvious advantages in operation, however, as interest rate liberalization and financial derivatives become more and more complex, these methods have relatively long lags in data collection and data processing and cannot be based on economic conditions gradually. With the development of modern science and technology and the innovation of financial theory, scholars established a modern credit risk measurement framework based on Value-at-risk (VaR), with default probability and expected loss as core risk measurement indicators. VaR refers to the significant operating losses that may occur in the future for certain investment portfolios in a specific period under normal market conditions. Since then, changes that cannot be measured by traditional methods can also be adjusted and predicted. Crouhy, Galai and Mark (2000) made a comparative analysis of the theoretical basis, advantages, and disadvantages of these four models and found that KMV model and Credit Portfolio View are somewhat related since the market value of the firms’ assets depends on the form of the economy. Then it will be interesting to compare the transformation matrices by the two models. Kollár and Gondžárová (2015) examined the KMV model, CreditMetrics and CreditRisk+ model and found that the KMV model is most suitable for listed companies whose stock market determines asset value.

2.1.2.1 CreditMetrics Model

In 1997, J.P. Morgan and other cooperative banks (1997) developed a credit indicator model, the CreditMetrics model. This model is mainly applicable to bonds and bond derivatives with strong liquidity, and it can easily obtain price information and rating data every time. The core of the model is the historical data of the company's credit rating, which changes various industries in the evaluation system for each year.

The basic idea of the credit indicator model is that changes in asset value will be affected not only by debt default but also by changes in the debtor ’s credit rating (Zokirjonov, 2018). Therefore, investors first need a credit rating transfer matrix to calculate the possible asset value and loan recovery rate. Then they calculate the asset value level and rate of change based on the discount rate. Finally, they can get VaR. The process is shown in the figure 1.

Figure 1: The process of CreditMetrics model



$$\sigma_T = \sqrt{\sum_{i=1}^{s} p_i(\mu_i^2 + \sigma_i^2) - \mu_T^2} \qquad \text{where } \mu_T = \sum_{i=1}^{s} p_i \mu_i$$

Source from J.P. Morgan (1997)

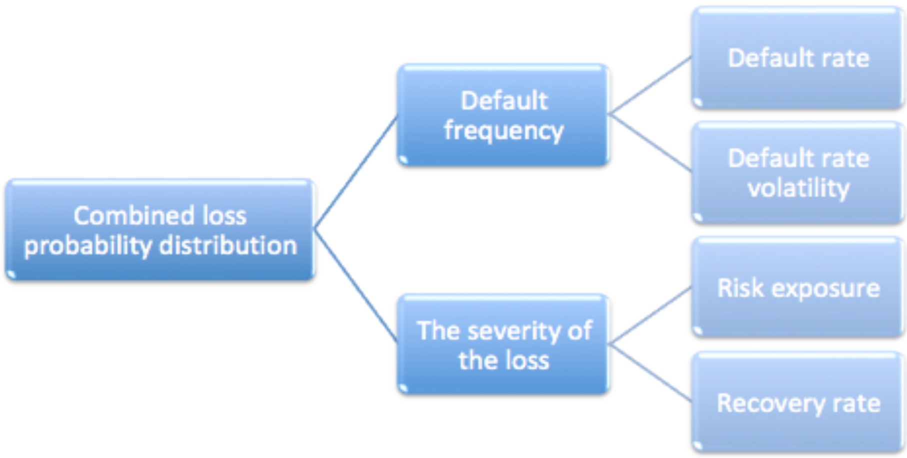
Zhang (n.d.) pointed out the advantages of this model. It has a wide range of applications, not only can measure personal loans or personal investment portfolio, but also can measure financial derivatives. The CreditMetrics model can identify credit risks uniformly in different markets by calculated the VaR index, thereby helping commercial banks weigh reserve and economic capital and effectively achieve the bank's income and security goals. This model is practical and can be used by companies in different industries.

Boris, Ivana and Anna (2015) presented the assumption of this model is that the term structure of interest rates is fixed and that market risks exist during a specific period, which is only applicable to short-term floating interest rate bills rather than credit derivatives and zero-coupon bonds. The premise of the model is to assume that the asset returns follow a normal distribution, but it is difficult to meet the conditions in reality.

2.1.2.2 CreditRisk+ Model

The CreditRisk+ (CR) model, the additional known credit risk issuance model, is a model used to measure credit risk developed by Credit Suisse Financial Products (CSFP) in 1997 (Gundlach & Lehrbass, 2013). The purpose is to determine the loss distribution of the bond or loan portfolio and the capital required to cover the risk. Diaz and Gemmill (2002) also pointed out that comparing with the CreditMetrics model, this model only considers the risk of default and does not be measured the reduction of credit rating. The calculation step of the CR model is that investors can improve the accuracy of risk measurement by dividing the risk exposure into different frequency intervals (Ong & International Monetary Fund, 2014). The model was constructed by Vandendorpe et al. (2008) through three steps as follows:

Figure 2: CreditRisk+ model frame diagram



This model has several advantages. The model is mainly used to analyze defaults. It only requires the distribution of defaults and risk exposures, so it just requires a few estimated variables and is easy to use. In addition, the model has good processing capacity and can handle different types of risk exposures in different periods, different regions, and different departments. Therefore, this model can also be used to handle large amounts of loans. Last but not least, the model can combine expected loss and unexpected loss through the combined value loss distribution function. (Kurth, Taylor & Wagner, 2002)

However, as Derbali (2008 cited in Hamisultane, 2018) presented, when grouping loan portfolios, since the model ignores changes in credit ratings, the credit risk exposure of each loan is fixed, which is inconsistent with the actual situation. The model is not considered a market risk, nor does it deal with non-linear financial products such as options and foreign currency swaps. The application range of this model is minimal because it is only valid for a few certain loan portfolios with specific characteristics.

2.1.2.3 CreditPortfolio View model

The CreditPortfolio View model (called the CPV model) is a model developed by McKinsey & Company in 1998 to quantify credit risk (Saunders & Allen, 2002). The model simulates the relationship between economic variables and grade conversion matrices and uses Monte

Carlo simulation to create macroeconomic “shocks.” It analyzes the migration probability and default situation of credit assets with different credit ratings in different industries. Combined with a conditional probability distribution, investors can analyze the credit risk of assets with different credit ratings in different industries.

Derbali and Hellara (2013) pointed out that the systemic risks of credit portfolios mainly depend on the health of the macroeconomics in the CPV model. Therefore, those risks cannot be resolved through effective decentralization. Since the model simulates the distribution of the joint conditions of credit rating transfer probability and default probability of different industries, the required variables are mainly macroeconomic variables, such as unemployment rate, GDP growth rate, and long-term interest rate.

Compared with other credit risk models, this model is considered the impact of macroeconomic variables, rather than blindly using historical credit rating data (Derviz & Kadl, 2001). The economic cycle, unemployment rate, interest rate, etc. are all systemic factors, so that they can be applied to companies in different countries and industries. In addition, non-dispersible risk portfolios are more suitable for management using this model because the model uses actual discrete distribution (Diez-Canedo, 2005).

However, using this model requires accurate macro data, it is difficult to obtain data from different countries and industries. When considering macro factors, the CPV model always ignores the company’s risk factors. The special process of the credit rating transfer matrix is established by the bank's credit department based on the subjective judgment of the target company's risk, which is what the model lacks (Kern & Rudolph, 2001).

2.1.2.4 KMV Model

Moody company developed the KMV model, and it is theoretically based on the Black-Scholes-Merton (BSM) stock option pricing model (Frenkel, Hommel & Rudolf, 2013). The expected default frequency (EDF) is a default forecast indicator used to reflect the enterprise's risk. The assumption of this model is that when the company's asset value is higher than the company's debt, the equity value is the difference between the two. On the contrary, when the company's asset value is lower than the company's debt, the company will sell all assets to repay all debts, even if the equity is not enough, and the company's equity value is equal to zero. For companies, this model construction is similar to buying European call options. Similarly, for creditors, this is equivalent to shorting put options.

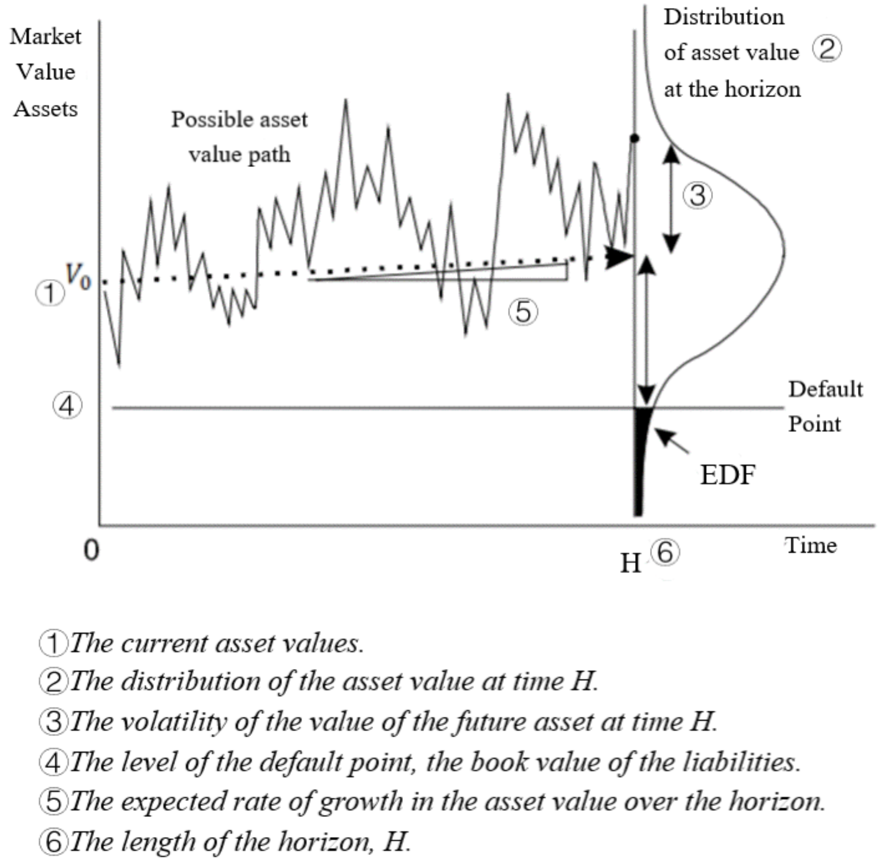
As mentioned above, the basis of the KMV model is the BSM model. The appearance of BSM provides an important reference for the pricing and theoretical basis of financial derivative products and investors' investment decisions. Before understanding the KMV model, understanding the assumptions of the BSM model is also one of the important conditions for exploring the KMV model. There are BSM Model assumptions (Merton, 1974):

- Stock returns are normally distributed.
- The market is frictionless, that is, there is no arbitrage.
- The risk-free interest rate is known and remains unchanged during the option period.
- The volatility of stocks is known and remains unchanged during the option period.
- The transaction costs are omitted from the model, and dividends are not included.
- Options cannot be exercised before they expire (spot options are European options).

On this basis, the pricing formula for non-income asset European call options is derived.

The investors estimate the market value and volatility of company assets based on the stock market value and volatility of option pricing companies. They can calculate default point (DPT), which is the deviation of the standard company value, based on the company’s liabilities. Finally, according to the database established by KMV, they can find the EDF, which shows the probability of default of the company. The greater the distance to default (DD) is, the less likely the company default is. The smaller the DD is, the greater the possibility of default risk is (Wang et al., 2010).

Figure 3: Principle of KMV model



Source: from Crosbie and Bohn (2003)

Kollár and Gondžárová (2015) cited several advantages of KMV model. The model has a strong theoretical basis, and it is based on the Merton option pricing theory. Because the input variable is capital market data, and real-time EDF can be obtained, the KMV model is forward-looking and can predict the future development prospects of the enterprise. Furthermore, the input data of the KMV model are financial statements and stock transaction data, which are relatively easy to obtain, so the estimation results of the model can be adjusted dynamically and in time.

There are disadvantages to the KMV model pointed out by Witzany (2017). The model is divided into debts into the short-term and long-term. It is assumed that the asset value of the borrowing company follows a normal distribution. However, it is difficult for all companies to meet this requirement. The market value and volatility of the company's assets are approximately equal to the value and volatility of the corresponding equity. Therefore, for long-term loans and interest rate sensitive products, the application of the KMV model will be limited. In addition, given the basic data requirements of the KMV model, it is difficult for unlisted companies to estimate the expected return and asset value fluctuations.

2.1.3 Comparison

At present, there are four main quantitative measurement models commonly used in credit risk assessment, which are the CreditMetrics model, CreditPortfolio View model, KMV model, and CreditRisk+ model. We measure from the five aspects, which are principle and analysis method, default state, default driving factor, measurement conditions, and discrete or continuous measurement, to show the differences among these four methods. As shown in the following table 1:

Table 1: Comparison of four models

Item	**CreditMetrics**	**Credit portfolio View**	**KMV**	**CreditRisk+**
Principle and analysis method	Historical data analysis of rating results	Macroeconomic Factor Adjustment and Simulation	Principles of Option Pricing	P & C insurance ideas
Default state	Mark-to-market model	Mark-to-market model or default model	Mark-to-market model or default model	Default model
Drivers of default	Asset value and its volatility	Macroeconomic factors	Asset value and its volatility	Expected default rate
Conditionality	Unconditional measure	Conditional measure	Conditional measure	Conditional measure
Discrete or continuous measure	Discrete measure	Discrete measure	Continuous measure	Continuous measure

From the analysis of Kern and Rudolph (2001), Kollár and Gondžárová (2015) and Derbali (2018), the migration matrix is used in the CreditMetrics model to measure the default rate of the research object. The migration matrix is based on the historical data of the company's credit rating, and these data from various industries change every year. Considering more macro data that affect the default rate and migration probability, it is the adjusted migration probability matrix and default rate matrix. KMV model is a quantitative tool used by Moody's KMV company to test the company's default rate through options pricing theory and is usually used to calculate the expected default rate of listed companies. The CreditRisk+ model is a type of credit risk management system launched by Credit Suisse, which is used as an actuarial insurance theory to calculate the probability distribution of bond loan portfolio losses.

2.1.4 Applicability of credit risk model in China

For CreditRisk+ model, China ’s stock market has not yet developed, and a reasonable basic interest rate has not yet been formed, which does not meet the assumption of risk-free interest rate under the model premise. China's commercial banks tend to provide loans to customers with a long-term willingness to cooperate, and bank customers usually have mutual guarantees, which lead to the fact that each loan of the bank is not completely independent. Therefore, the existing lending situation in China does not meet the assumption of the model, Poisson distribution. At present, data on China's corporate default recovery rate and corporate credit rating conversion probability cannot be used to quantify and manage credit risk. Overall, the CreditRisk + model is not yet applicable under the current conditions in China.

The Chinese financial market lacks historical data and does not have a complete rating system. China does not have a professional and complete rating system and organization, so various rating standards cannot be unified. Since the CreditMetrics model is totally based on a rating system, it is restricted in China.

For the CreditPortfolio View model, the model's application in China has similar difficulties as the CreditMetrics model, such as the lack of historical data foundation and authenticity of the data. In addition, it is difficult to determine the economic meaning of the model in China, and the relationship of this model also lacks stability. Therefore, the application of this model in China is limited.

The KMV model has relatively large advantages in the application of the Chinese credit risk market. First, the KMV model predicts credit risk data more accurately and timely. The data of the KMV model are based on the daily updated data of listed companies on the stock market, which can reflect the status of corporate credit risk in time. In addition, the price of the stock market also implies future expectations. Using this parameter to estimate credit risk can also reflect people's expectations and forward-looking, and it can predict the status of credit risk more timely than other models. Third, the risk comparison is accurate, so each company can be calculated its own default probability, and investors can compare the risks of any two companies. Finally, the KMV model uses readily available stock prices of listed companies as data. Therefore, according to the actual situation in China, the KMV model is more applicable.

2.2 Previous empirical results of KMV model

Scholars have been exploring the usage conditions and background of the KMV model. In general, the KMV model is more suitable for listed companies because it contains more open and transparent data and companies' value. Furthermore, compared with other credit risk models, the KMV model has a better risk warning ability.

Compared with the CreditMetrics model, Crouhy, Galai and Mark (2000) emphasized that the KMV model had different default interest rates in a bond rating category. It shows that all companies with the same bond rating do not necessarily have the same default rate. Besides, since the KMV model is derived from Merton, the traditional credit migration table replaces the existence of EDF. In other words, each value of EDF is associated with credit ratings and spread curves. Generally, the KMV model applies to most listed companies because their stock prices can be directly observed. In order to explore a better quantitative risk model to accurately represent the degree of credit risk, Sobehart, Keenan, and Stein (2001) used four critical factors, accuracy rate, cumulative accuracy profile, conditional information entropy rate, and mutual information entropy, to measure credit risk models. The result shows that the KMV model has the highest risk prediction accuracy.

Crosbie and Bohn (2003) used financial companies as research objects to verify the effectiveness of the KMV model. The result proves that before the occurrence of a credit incident or bankruptcy, the French power company can indeed predict the credit status and changes of French and British companies more accurately and more sensitively. Takezawa and Takezawa (2003) used Vertical Keiretsu system samples to build different models to analyze the Japanese automotive industry. The result shows that the prediction result of KMV is more accurate than the Leland-Toft model and bivariate VAR model.

Crouhy, Galai and Mark (2000) compared the KMV model, CreditMetrics model, CreditRisk + model, and CreditPortfolio View model, and proposed a comprehensive framework that combines credit exposure and loss allocation to measure the value of derivative securities credit risk. They hoped to find a credit model that considers not only random interest rates but also the company's default probability and its probability of migration based on economic conditions such as interest rate levels and stock markets.

Crossen and Zhang (2011) analyzed the model performance of global financial companies before and after the US subprime mortgage crisis. They also distinguished the capabilities of defaulters and non-defaulters from different aspects. According to their results, the KMV model is superior to other risk measures, such as Agency Credit Ratings, because the KMV model can provide an early warning signal for 12 months before the default occurs. They also concluded that the EDF credit indicator is a handy forward-looking indicator of credit risk for analyzing global financial companies.

Kliestik et al. (2017) analyzed the bankruptcy of Slovakia in the financial derivatives index to improve the accuracy of the model by modifying it. Arora, Bohn and Zhu (2005) studied the extent to which companies disclose information when announcing information indicating difficulties. The existence of such "companies in distress" information stated that investors could infer from the abnormal returns of bonds or stocks issued by bankrupt companies or default companies. In order to make the manager's default decision may not completely surprise the market with the information of "troubled companies," the author chose a potential information method. Besides, this method can also help them analyze companies that have not yet disclosed information but conflicting information may have leaked to the market.

Blochwitz, Liebig and Nyberg (2000) compared the Gini curves and Gini coefficients determined on the same underlying data set by the Deutsche Bank default risk model, the KMV model, and the discriminatory German companies' common financial ratio. Because the KMV model is based on the market value of assets of the listed company and the asset value volatility is estimated based on the company's observation characteristics, they used the private enterprise model as a scale cutoff. Thus, scholars could use formal reports to update financial information used as model input and obtain updated EDF. Compared with traditional statistical methods, their modified model changed the company's capital structure while retaining the company's characteristics.

In order to explore the dependence between serious credit default time factors in the US national mortgage market, Gapko and ŠMíd (2012) established the modified VECM model. They changed the framework of the original model, which considers both the default rate (DR) and the default loss (LGD). DR and LGD are determined by a common and independent factor. Furthermore, they could implement the IRB method within the regulatory framework of Basel II.

2.3 Amendments of KMV model in China

Although previous papers provide theoretical and empirical proofs that the KMV model is a good reference for quantitative credit risk management in China, the theory cannot be applied directly to the Chinese market since considering that the capital market environment in China is different from that in developed countries, there will be some deviation in the empirical research of copying KMV model in China, so it is necessary to make the adjustment for the parameters.

In general, the parameters involved include the calculation method of equity value and asset value, the calculation method of equity volatility, and the setting of different default points.

Wang, Qu and Li (2015) measured the sample companies default distance based on the 2014 commercial bank loans to customers of financial data and stock trading data including 10 normal and 10 ST companies, then got the sample companies expected default rate, the empirical result showed that the EDF can well reflect the credit risk of listed companies. When calculating the equity value, they distinguished the tradable shares from the non-tradable shares. In addition, the credit risk of business cooperation between commercial banks and normal companies was low, and the credit risk of association with ST companies was high. In the aspect of adjustment for equity volatility, Xu and Zhang (2004) already used the GARCH model to predict the volatility of equity value. In addition, free cash flow was introduced to calculate the equity value instead of the stock market. Furthermore, they used Weibull distribution to describe the value of the company, which is different from the previous normal function mapping. The Weibull distribution derived the relationship function between the default distance and the default probability. It proved that the new function relationship is more conducive to the application of the KMV model in China. As for the adjustment for the DPT, Lee (2011) presented a new approach applying genetic algorithms (GA) to find an optimal DPT for the KMV model in Taiwan, and the GA-KMV model enhanced the accuracy of default forecasting. Yang et al. (2013) amended the parameters of the DPT through an appropriate clustering method and further also analyzed the applicability of the amended KMV model in different industries based on the data of Chinese listed companies. To solve the objective function as well as get the optimal DPT in China, Zhang and Shi (2016) obtained the price discount on non-tradable shares through this hybrid KMV model and shed light on how to price non-tradable shares by the hybrid KMV model.

3 Methodology

KMV model was proposed by the credit risk analysis professional company founded by Kealhofer, McQuown, and Vasicek. In order to maintain its core competitiveness, Moody's did not publish the details of the calculation of default distance and default probability in its credit risk evaluation. Therefore, the research and application of the model are mostly based on its framework.

3.1 Principle of KMV model

3.1.1 Asset value and volatility of asset returns

Based on the BSM model, the function between equity value and the asset value is constructed as follows:

$$V_E = f\ (V_A, K, \sigma_A, r, T) = V_A N(d_1) - Ke^{-rT} N(d_2) \tag{3-1}$$

Where:

$$d_1 = \frac{ln(\frac{V_A}{K})+(r+\frac{\sigma_A^2}{2})T}{\sigma_{A\sqrt{T}}} \tag{3-2}$$

$$d_2 = d_1 - \sigma_A\sqrt{T} = \frac{ln\ (\frac{V_A}{K})+(r-\frac{\sigma_A^2}{2})T}{\sigma_{A\sqrt{T}}} \tag{3-3}$$

Of which, V_E denotes equity value of the company, V_A denotes asset value of the company, *K* is the face value of debt, *DPT* is the default point, *r* is the risk-free rate, σ_E is the volatility of equity returns, σ_A is volatility of firm returns, and *T* is debt maturity. Besides, *N (*)* is a cumulative standard normal distribution function. Both V_A and σ_A are unknown and others in this paper are observed. In other words, these two unknown variables need to be solved finally.

According to Ito's formula, equity asset can be viewed as a function of asset value:

$$\sigma_E = N(d_1)\frac{V_A}{V_E}\sigma_A \tag{3-4}$$

3.1.2 Equity value and volatility of equity returns

The equity value V_E calculated by the traditional KMV model does not consider the special factor of non-tradable shares, so it cannot solve the pricing problem of non-tradable shares in China. In western mature capital markets, all the shares of listed companies are tradable. Since 2005, the reform of non-tradable shares in the Chinese capital market has theoretically achieved full circulation of stocks. However, there is still a transitional period for the real lifting of the ban on non-tradable shares, which leads to the fact that a large number of listed companies in China still have non-tradable shares.

Taking into account the particularity of China's securities market, it is inevitable to cause errors to directly estimate the equity value of enterprises by using ordinary tradable shares. Aiming at the calculation of the market value of non-tradable shares of listed companies, in order to avoid the overvaluation of the company's assets leading to the reduction of probability of default, this paper adopts the common correction method in the literature:

Equity value = the closing price of tradable shares × the number of tradable shares + net assets per share × the number of non-tradable shares

$$V_E = P_i \times N_i + M_j \times N_j \tag{3-5}$$

Of which listed tradable shares are only counted as tradable A-shares. A-shares are divided into two classes: one type being freely bought and sold by normal investors (tradable shares) and the other that cannot be freely traded (non-tradable shares). Meanwhile, in order to avoid the influence of speculation and other behaviors that lead to large stock fluctuations, the average daily closing price of the tradable shares is selected for the closing price P_i in this paper.

The method of sample variance, the historical volatility method, is used in the KMV model to predict the volatility of the stock market. It is assumed that the random process of the yield sequence of this method is independent and identically distributed. In the calculation process, the same weight is assigned to the data, and the volatility of the equity value is obtained by the standard deviation of the yield. The calculation formula is:

$$r_t = ln\frac{P_t}{P_{t-1}} \tag{3-6}$$

$$\sigma_{day} = \sqrt{\frac{1}{n-1}\sum_{i-1}^{n}(r-\bar{r})^2} \tag{3-7}$$

$$\sigma_{year} = \bar{\sigma}\times\sqrt{n} \tag{3-8}$$

Where P_t denotes the closing price of the stock on day t, r_t denotes daily return, $\bar{r}$ denotes the average value of r_t, σ_{day} is the daily volatility of a stock, and σ_{year} is the annual volatility of a stock. $\bar{\sigma}$ denotes the average of all the σ_{day} in this year. n denotes the trading dates.

After obtaining V_E and σ_E, V_A and σ_A can be solved through the formula of 3.1.1.

3.1.3 Calculation of DD and DPT

Crouhy, Galai and Mark (2000) said that *DD* is the number of standard deviations between the mean of the distribution of the asset value and a critical threshold, the “default point”. In other words, it refers to the distance between the expected value of the company's assets *E*(V_A) and the *DPT*. Among all structured models, the most influential is Moody's KMV model. *DD* is also the first concept proposed by Moody's KMV. The formula is as follows:

$$DD = \frac{E(V_A)-DPT}{E(V_A)\times\sigma_A} \tag{3-9}$$

Where: $E(V_A) = V_A \times (1+\varphi)\varphi$ denotes the expected growth rate of asset values. In the calculation of φ, this paper adopts the calculation method below:

$$\varphi = \frac{total\ assets\ at\ the\ end\ of\ that\ year - total\ assets\ at\ the\ beginning\ of\ that\ year}{total\ assets\ at\ the\ end\ of\ that\ year} \qquad (3\text{-}10)$$

It is usually assumed in empirical studies that φ is zero or risk-free interest rate, which is obviously inconsistent with the actual situation. The operating scale of high-quality listed companies continues to expand, and the value of assets increases year by year. In contrast, listed companies that fail to operate normally due to decision-making errors and other reasons may fall into the financial crisis, resulting in the decline of their asset values year by year. Therefore, simply setting the expected growth rate of asset value as zero or risk-free interest rate will lead to errors in the assessment results.

In theory, when a company's asset market value is less than its debt, it will default. However, in practice, the company's long-term debt can usually ease the company's short-term debt repayment pressure, so the company will not default due to financing. By analyzing a large amount of historical default data in the United States, KMV Company found that the default occurred most frequently at the company value = short-term debt + 0.5 long-term debt. Jessen and Lando (2015) specially studied the robustness of *DD* under different default points and concluded that the change of default point had little impact on the ability of *DD*. Thus 0.5 here is chosen to be the α value.

$$DPT = STD + 0.5LTD \qquad (3\text{-}11)$$

Via the formula (3-9), *DD* can be solved on the premise that $E(V_A)$, DPT and σ_A are already known. If the *DD* is small, it means that the company is more likely to default at maturity, less likely to repay debt, and has higher credit risk. On the contrary, the greater *DD* is, the less likely the company is to default at maturity, the more likely it is to repay the debt, and the lower the credit risk is. The stock prices of listed companies will be constantly updated on the trading day, and financial statements will be published regularly. Therefore, the default distance of the company can be calculated timely and the change of credit risk can be measured.

3.1.4 Calculation of expected default frequency (EDF)

The calculation of EDF is divided into theoretical *EDF* and empirical *EDF*.

The calculation of theoretical *EDF* is to select a certain time span and map the relation between *DD* and *EDF* by using a function. If the asset value of an enterprise follows a normal distribution, the theoretical calculation method is as follows:

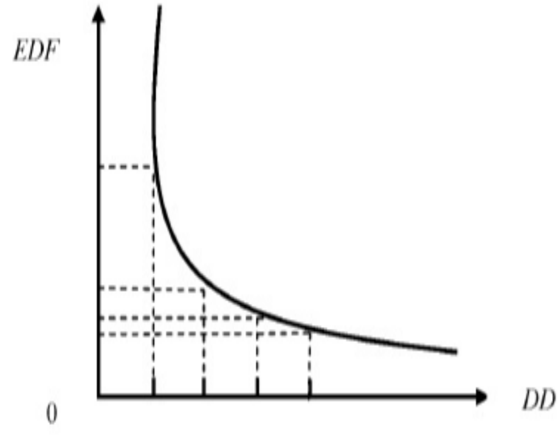
$$EDF = P(E(V_A) < DPT) = N(-DD) \tag{3-12}$$

However, it is not reasonable in practice to assume that V_A follows a normal distribution. And the *EDF* obtained by using the above formula is much smaller than the actual probability of default, which is inconsistent with the reality. Therefore, the KMV company does not advocate this algorithm.

Another method is empirical *EDF*, which was from KMV Company, aimed to help financial institutions represented by commercial banks to better measure credit risk. For an effective credit risk measurement model, as the time approaches the default date, the default risk of a company increases day by day, and the value of *EDF* will increase.

The graph about the relationship between *DD* and *EDF* is as follows:

Figure 4: Relationship between DD and EDF



As can be seen from the figure, the distance to default is inversely proportional to the probability of default. The smaller *DD* is, the larger *EDF* is. Conversely, the larger *DD* is, the smaller *EDF* is. In this thesis, we still use *DD* to measure the credit risk due to difficult measurement of *EDF*.

3.1.5 Risk-free rate

Since China's market lacks a developed interest rate market mechanism, there is no official risk-free rate in China. Many scholars simply assume a risk-free rate in their research, which is obviously inappropriate in practice. In this paper, the one-year time deposit interest rate announced by the people's Bank of China is selected as the risk-free interest rate in 2009-2018 in the KMV model. Interest rates from 2009 to 2018 are as follows:

Table 2: Interest rates from 2009 to 2018 (%)

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Interest rate	2.25	2.5	3.25	3	3	2.75	2.135	1.5	1.5	1.5

Source from: the People's Bank of China

3.1.6 Time horizon

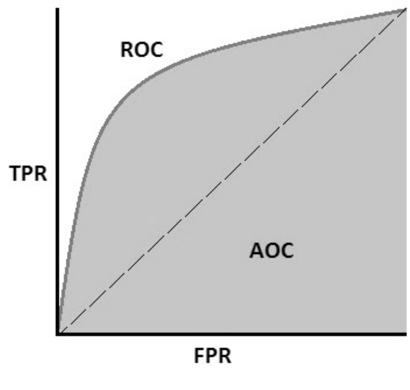
In a previous study of KMV model, scholars generally use one-year time span of the data, and the time span is less than one year may result in too much data to the data processing pressure, and time span more than one year may lead to data are insufficient to response the real situation, so this paper also uses one-year as time span.

The KMV model adjusted according to the characteristics of China's economic environment will be the KMV model used in the empirical study of this paper. The KMV model mentioned in the following studies, in the absence of special designation, refers to the KMV model adjusted by the parameters in this section.

3.2 ROC curve

The ROC (Receiver Operating Characteristics) curve is a graphical diagram that illustrates the diagnostic capabilities of the binary classifier system when the discrimination threshold changes. In other words, it is a performance measurement for the classification problem at various threshold settings. ROC curve is a technique for organizing and selecting classification models and visualizing their performance (Fawcett, 2006), and it is also the most popular technique for evaluating various rating methods and is widely used in the medical field. Hanley and McNeil (1982) described the statistical characteristics of the ROC curve in detail. Fawcett (2006) pointed out that the ROC curve has a particularly attractive property: it is not sensitive to changes in the distribution of categories. Even in a test, if the ratio of positive and negative instances changes, the ROC curve will not change. Sobehart and Keenan (2001) elaborated on how to test the internal credit rating model using the ROC curve and reached an important conclusion: the area under the ROC curve (AUC) is a decisive indicator of the model's ability to predict. The larger the AUC is, the better the prediction effect of the model has. The AUC cannot be directly observed, but statistical software such as SPSS and SAS can clearly draw the ROC curve and obtain the area under the curve.

Figure 5: AUC-ROC curve



The horizontal axis of the ROC curve: 1-specificity, false-positive rate (FPR), the proportion of all negative samples that are predicted to be positive but actually negative. The vertical axis of the ROC curve: sensitivity, true positive rate (TPR), the proportion of all positive samples that are predicted to be positive but actually positive.

In order to further explain more accurately and more intuitively observe the prediction effects of the KMV model and the Z-Score model, the ROC curve and its AUC are introduced to provide a judgment criterion for the model's predictive ability in table 3.

Table 3: Introduction of the ROC curve

Test	Present	Absent	Total		
Positive	True Positive (TP)	False Positive (FP)	Prediction Positive=(TP+FP)	PPV=TP/prediction positive (Precision)	FDR=FP/ prediction positive
Negative	False Negative (FN)	True Negative (TN)	Prediction Negative=(FN+TN)	FOR=FN/prediction positive	NPV=TN/ prediction positive
Total	Condition Positive =(TP+FN)	Condition Negative =(FP+TN)	N=TP+FN+FP+TN		
	TPR=TP/condition positive （Sensitivity）	FPR=FP/ condition negative (1-Specificity)			
	FNR=FN/condition positive (1- Sensitivity)	TNR=TN/condition negative (Specificity)			

When evaluating a credit risk model through a curve, four situations may occur. First, the ST company in the experimental group are correctly predicted. Second, the non-ST companies in the reference group are appropriately predicted. Third, the ST companies in the experimental group are misjudged as a non-ST company. Fourth, the non-ST companies in the reference group are mistaken for ST companies. Obviously, the better the model fitting result is, more frequently the first and the second occur, while less frequently the third and the fourth occur.

The AUC value, which is the area enclosed by the ROC curve with the horizontal axis and the vertical axis, indicates the prediction accuracy, and it is also an essential indicator in the ROC curve. The larger the AUC value is, the better the model explains. Hosmer, Lemeshow and Sturdivant (2013) provided an AUC identification standard. If the AUC is between 0.5 and 0.7, it means that the discriminative ability of the model is relatively weak, only higher than a coin toss. If the AUC is between 0.7 and 0.8, the test model is acceptable. If the value of the AUC is between 0.8 and 0.9, it means that this model has excellent discriminating ability.

4 Empirical study

4.1 Data description

ST companies are listed companies with special financial status who are engaged in stock trading. They are simply referred to as ST companies, which are special problem companies. There are two reasons why those companies are marked ST. The first is that during the audit process, the net profits of listed companies in the two fiscal years were negative. The second one is that the stock prices of listed companies are determined by the company's recent operating conditions. If a company’s net assets per share are lower than the par value of the stock, it proves that this listed company's financial situation is poor and is marked as an ST company. The main purpose of marking stocks is to provide shareholders with warning opinions that the stock may be in danger of delisting, and hope shareholders can remain cautious.

Therefore, in order to reduce the risk, the CSRC requires that the stock trading of these listed companies is restricted. Their stock transactions should be carried out in accordance with the regulations (Shenzhen Stock Exchange Trading Rules, 2016): The maximum daily upper and lower limit is 5%; The listed companies must pass accounting audits when performing interim report analysis.

In order to comprehensively analyze the Chinese economy, there are 100 ST companies and the corresponding 100 non-ST company sample companies, a total of 200 listed companies. They are chosen from Shenzhen Stock Exchange and Shanghai Stock Exchange and divided into two groups, default group and non-default group. Since most companies have data shortages in 2019, the data from 2009 to 2018 are chosen in this thesis. In 2018, there were 143 ST companies in the stock market. Since the requirements of the model on the length of the data interval in empirical analysis, 100 ST companies are chosen from 143 ST companies. Other ST companies are not selected because the time to market is shorter than ten years. These listed ST companies are involved in many industries in the stock market, including electricity, real estate, textiles, metals, internet services, chemical industry, computer software, trading, vehicles, etc. Furthermore, these industries develop independently and combine the major industry chains in the market, in other words, it can more fully represent the entire Chinese industry.

In order to better compare the difference of credit risk in companies, each selected ST company has a corresponding non-ST company. Their industries are the same, and the difference of circulation market value between the two group does not exceed 15%. The appendix includes detailed sample companies, and the data are from WIND[1]. According to "Shenzhen Stock Exchange Stock Listing Rules", if a model can measure the level of the credit risk effectively, in general, the DD of listed companies will be less than that of listed companies, and the model from the statistical test can distinguish different DD between the two sample groups significantly.

Data in this paper from WIND include working capital, total assets, retained earnings, earnings before interest and taxes (EBIT), the market value of equity, the book value of total liabilities, sales, the closing price of tradable shares, the number of tradable shares, net assets per share, the number of non-tradable shares, long term debt and short term debt.

It should be noted here that the ST company selected here was not always keeping the ST company during the selected period. Some of them had only changed once from non-ST companies to ST companies, and some companies were constantly changing between ST companies and non-ST companies. Therefore, the 4,000 data of the DD and Z-values used in the next applicability test, and 2000 for each, 361 of which are marked as the data of ST companies in the year, marked as 1, and the remaining 1639 data are from the ST companies when they were not marked and non-ST companies within 10 years, marked as 0.

[1] www.wind.com.cn Wind information is a leading financial data, information and software services company in mainland China, headquartered in Lujiazui financial center, Shanghai.

4.2 Limitation in data

KMV-Moody's business secrets are not publicly available, but KMV still provides an analytical framework. With respect to the analytical framework, there is no standard answer to this model, which means the model may be less prescriptive and precise even if amendments are done from the traditional model to adapt to China's financial markets in this paper. Furthermore, for the calculation formula of default distance, the assumptions of the model are quite strict and cannot be realized in reality, which will affect the accuracy of the results.

In this paper, ST companies and non-ST companies are treated as default and non-default groups. It has to be admitted that although sample results with research significance can be obtained according to this method, the difference between the ST group and the non-ST group is somewhat different from that between the default and non-default enterprises in terms of theoretical definition. Although a company is implemented special treatment, it is a better operating condition likely when it still does not take off "ST". However, the level of credit risk measured by the KMV model is found to be improved, which led to the inconsistency between the calculation results of the KMV model and the situation of the defaulted enterprise defined in this paper, affecting the judgment of the validity of KMV model. On the one hand, the method of adding "ST" before the name of a listed company is mainly based on the information in financial statements. However, the current stock market in China is not a strictly standardized market, and the information in financial statements is still more or less artificial, which will also affect the judgment ability when using the model to measure. On the other hand, although ST companies are specially treated companies, their operating conditions may become better due to the changes in the market environment and their internal operations. The company is marked an ST but still can continue to operate, unlike a default company in common sense. The above may cause some errors in the results we get.

The "ST" of a listed company is an objective event with high observability and easy to determine the research samples. Using ST companies is relatively precise because China has not set up the default company database that this paper chooses the best default company substitute, ST company. Moreover, since all the ST companies selected in this paper were marked or already had been marked to be ST in 2018, it does not mean that these companies had been in the state of ST in the sample period of 10 years. It is the first time for some companies to be marked ST these recent years, while some other companies changed from ST to non-ST and then changed back to ST later, which shows in our data sample, 10 years of data from ST companies are not necessarily all be regarded as defaulting data, in other words, when these companies were in the non-ST stage, their probability of default should be broadly similar with normal non-ST companies. As a result, non-default data are much more than default data in this paper. It has been 22 years since then, and there are about 100-200 ST companies every year. Due to the requirements of the sample range (10 years), only 100 ST companies that existed in 2018 are selected in this paper. As a result of sample availability, the total number of samples is not very large, which may affect the research results to some extent.

In terms of the risk-free interest rate, the current research in China generally adopts the one-year time deposit interest rate. Because of the change of national interest rate regulation, in addition to the risk-free rate in various industries and regions also can exist specific differences, it will lead to the interest rate of the year is different within one year. The interest rates of one-year time deposit after the last adjustment in the current year are uniformly adopted in this thesis. Whether the risk-free interest rates can adapt to the rapid changes in various industries and regional markets remains to be further studied.

In addition, the application of the ROC curve shows that the modified KMV model has practical significance in China but with such a small sample, ROC is not highly accurate.

4.3 Implementation of KMV model and Z-Score model

In chapter 3, the basic principles and framework of the KMV model have been explained and will not be elaborated again here. In this part, the Z-Score model from Altman as a comparison model will be used to measure the credit risk of the sample listed companies. The reason for choosing the Z-Score model is mainly based on the following three reasons: first, the model is mainly based on financial data to measure credit risk, which is easy to obtain in China's market. Secondly, the Z-Score model belongs to the multivariable credit scoring system, while the KMV model belongs to the structured credit risk model. The different properties of the two models make them more comparable. Thirdly, in the KMV model, DD is used instead of EDF since DD can be viewed as the same expression as the classic Z-Score model, which treats DD as a score and measures the likelihood of financial distress in the company by the score. Then in the next part, ROC curve analysis as one of the methods to verify the validity of the KMV model and Z-Score model will also be used.

4.3.1 The calculation process of KMV model

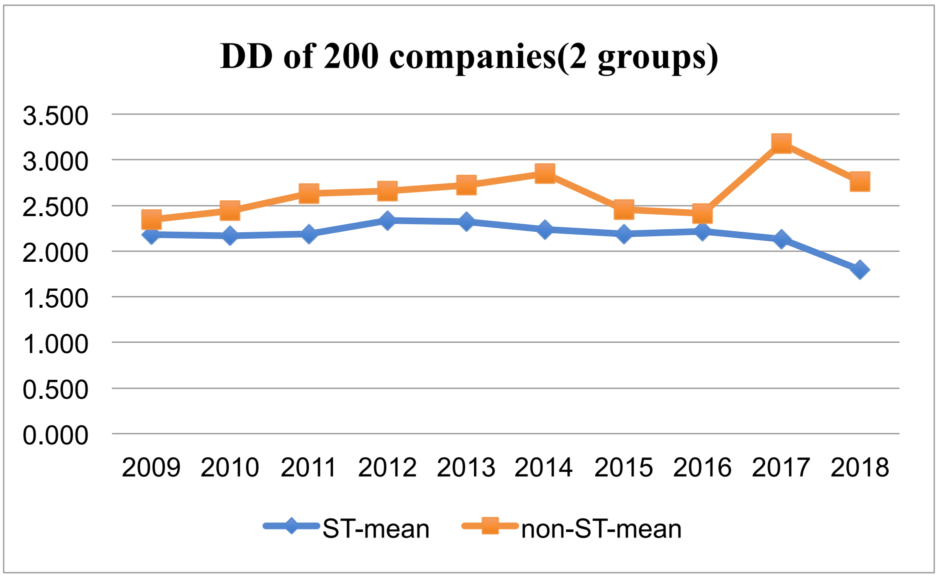
To solve DD, V_A and σ_A $these$ two key variables that cannot be directly observed. Before getting these two variables, asset value and asset value volatility need to be calculated, which can be obtained according to formula (3-5), (3-6), (3-7), (3-8). In addition, DPT can be obtained from formula (3-11). T=1 is known, and r varies from year to year (See table 2). Matlab2019 is used to solve nonlinear equations (3-1), (3-4), and (3-9), V_A, σ_A, and DD can finally be obtained (The specific Matlab2019 solver is described in detail in the appendix). Below is the descriptive statistics of DD from 2 groups.

Table 4: Descriptive statistics of DD

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
ST (Mean)	2.181	2.167	2.186	2.338	2.324	2.234	2.187	2.217	2.135	1.793
non-ST (Mean)	2.345	2.443	2.631	2.656	2.721	2.847	2.457	2.413	3.181	2.766
ST (Median)	2.113	2.293	2.237	2.436	2.428	2.373	2.243	2.227	2.113	1.767
non-ST (Median)	2.175	2.339	2.559	2.558	2.518	2.733	2.346	2.269	3.129	2.649
ST (Max)	3.847	3.553	3.846	5.007	3.262	3.771	3.785	3.975	4.391	3.242
non-ST (Max)	4.883	3.890	5.690	4.200	5.131	4.929	4.511	3.911	4.973	4.901
ST (Min)	0.397	0.432	-0.804	0.567	0.584	-0.071	0.819	0.019	0.292	0.160
non-ST (Min)	0.368	1.720	0.638	0.362	1.077	1.659	1.147	1.702	2.011	0.410
ST (std)	0.692	0.553	0.616	0.653	0.485	0.590	0.590	0.648	0.706	0.635
non-ST (std)	0.632	0.503	0.634	0.550	0.743	0.670	0.742	0.504	0.652	0.855
ST (obs)	100	100	100	100	100	100	100	100	100	100
non-ST (obs)	100	100	100	100	100	100	100	100	100	100

The table reports descriptive statistics for the annual observations of ST companies and non-ST companies. It should be noted that the ST companies in the data samples selected in this paper have not been marked by ST all the time in the past 10 years. It can be seen from the table that in 2009, the mean values and median values of ST companies are not significantly different from those of non-ST companies. This is because most of the ST companies in the sample were in normal operation, that is, non-ST status. However, from the overall point of view of 10 years, the mean and median values of ST companies are significantly smaller than those of non-ST companies. It can be seen that the default distance is different between the two types of companies.

Figue 6: Mean values of DD



In order to visualize the DD difference between the two groups, a line graph of the average DD values of the two groups is made. After calculating DD of 100 ST companies and 100 non-ST companies, the annual data are processed on average of these two groups, and it is found that DD value of ST companies is lower than that of non-ST companies in the same decade. Meanwhile, during this decade, the largest average DD in ST companies is also smaller than the smallest DD in non-ST companies. It can be noted that it is not difficult to find that the mean values of DD in the group of ST companies have been decreasing year by year since 2016. This is because many companies in the sample data selected started to be in the ST state in 2017 and 2018 especially in 2018. Most companies were in the non-ST state in the first few years of the sample range and were marked ST in the last two years.

4.3.2 The calculation process of Z-Score

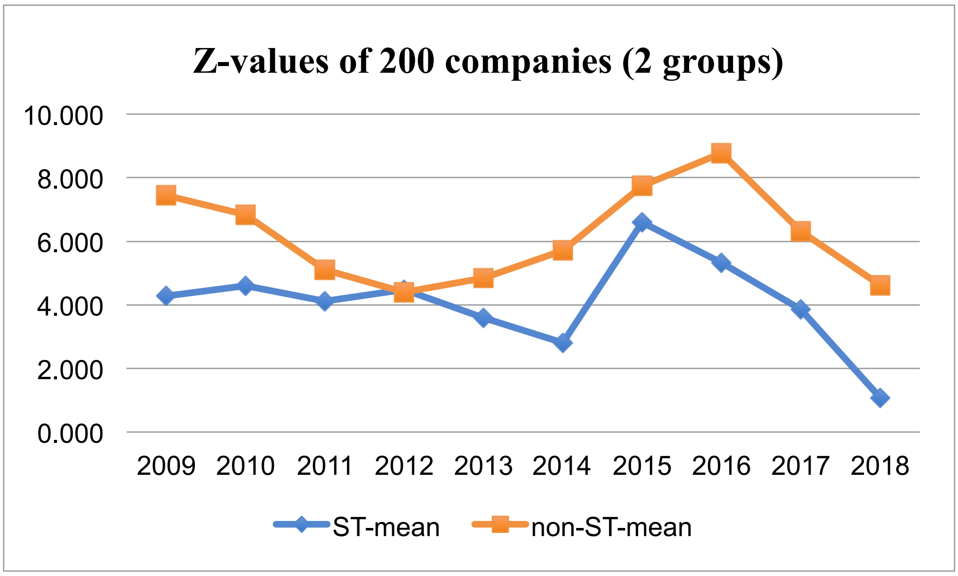
In order to evaluate the predictive performance of the default distance compared to the accounting model, Z-Score is used as the counterpart to compare with the KMV model. According to the formula (2-1), Z-values of 200 companies over 10 years are calculated. Through the financial statements of listed companies, the input variable data required by the model can be found. With the help of Excel calculation, the following descriptive analysis results can be obtained:

Table 5: Descriptive statistics of Z-values

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
ST (Mean)	4.278	4.601	4.107	4.477	3.591	2.813	6.591	5.323	3.854	1.083
non-ST (Mean)	7.447	6.839	5.108	4.398	4.850	5.701	7.737	8.764	6.309	4.627
ST (Median)	2.844	2.825	1.916	1.929	2.207	2.188	3.583	3.232	2.519	0.942
non-ST (Median)	5.321	7.285	6.921	6.659	6.401	5.988	8.264	6.674	7.387	6.660
ST (Max)	79.135	56.916	93.762	80.659	24.873	33.164	64.431	64.010	31.241	24.632
non-ST (Max)	127.518	40.625	24.811	16.633	56.007	62.752	94.075	199.014	68.333	50.086
ST (Min)	-70.466	-47.653	-48.543	-1.739	-1.325	-89.708	-34.449	-20.835	-22.389	-36.337
non-ST (Min)	-3.185	-8.272	-0.002	0.141	-0.039	-0.621	-0.538	-0.062	0.075	-0.316
ST (std)	11.638	9.395	13.694	9.466	4.360	11.211	10.870	8.841	6.395	6.301
non-ST (std)	14.037	8.734	5.346	3.990	6.897	8.486	11.064	20.868	10.124	6.185
ST (obs)	100	100	100	100	100	100	100	100	100	100
non-ST (obs)	100	100	100	100	100	100	100	100	100	100

From the table above, over a 10-year period, both the mean and median values of the non-ST group are larger than those of the ST group. The difference between the maximum value and the minimum value of Z-Score is more obvious than that of the KMV model, and the whole value distribution is more dispersed. In 2016, the maximum value of the non-ST group reached 199.014, and the minimum value was only -0.062, which makes the standard deviation of the Z-Score model larger.

Figure 7: Mean values of Z-value



To visualize the Z-value difference between the two groups, a line graph of the mean Z-values of the two groups is made. After calculating Z-values of 100 ST companies and 100 non-ST companies, the annual data of the two groups are averaged, and it is found that Z-value of ST companies is lower than that of non-ST companies in the same decade in general. Only in 2012 is the non-ST group even smaller than the ST group, 4.398 and 4.477. It can be seen that Z-value of ST group in 2015 is larger than Z-value of non-ST group in 2011, 2012, 2013 and 2014. There are three main reasons for this. First, the standard deviation of Z-values is much larger than that of DD, so the difference between Z-values is not easy to control. Second, in the sample of 10 years, there were not many companies that were marked ST in the middle of these years, thus the difference between the two groups is not obvious enough. For example, all the companies in the ST group were treated ST in 2018, so that the Z-value difference between the two groups is relatively obvious, 1.083 and 4.627, respectively. Thirdly, it also reflected that the prediction of Z-Score may not be very accurate.

4.4 Applicability test of modified KMV model and Z-Score model

The significance test is a method to infer the overall characteristics based on sample data. This method can distinguish whether data differences are due to unexpected fluctuations or significant differences. According to the above empirical results, the DD and Z-value of ST companies and non-ST companies are different. However, in order to prove that the KMV model and the Z-Score model can distinguish ST companies and non-ST companies, this thesis will still perform significance tests on the difference between the DD for the KMV model and the Z-value for Z-Score model to evaluate the reliability of these two models about the credit risk of listed companies in China. Here, t-test and Mann-Whitney U test will be implemented.

4.4.1 Normality test

In this thesis, a total of 2000 DD and 2000 Z-values are obtained in the above experiment. The significance test can distinguish whether these two models, the KMV model and the Z-Score model, have the ability to distinguish ST companies and non-ST companies. However, the significance test has strict requirements on the distribution of data. Therefore, before testing significance, one-sample non-parametric test methods should be used to test the data to determine whether they follow a normal distribution. Then through the parameter test method, it is inferred whether there is a significant difference between ST companies and non-ST companies. In this thesis, the Kolmogorov-Smirnov test method is used to test the data distribution.

In the one-sample test (See the appendix table A2) from SPSS, the values of DD and Z indicate that the significance level on 2-tailed is approximately 0, which is less than 0.05. Therefore either of them rejects the null hypothesis. It means that neither DD nor Z follows a standard normal distribution.

4.4.2 T-test

The T-test is performed by conducting independent sample tests on 100 ST companies and 100 paired non-ST companies' DD and Z-values according to their means and analyzing whether these two models can significantly distinguish these two types of companies.

Table 6: Group statistics for T-test

Group Statistics				
	ST or non-ST	N	Mean	Std. Deviation
DD	0	1639	2.5613240178501	0.6310001228625
	1	361	1.7292077597411	0.6203867031024
Z	0	1639	5.7310182119389	10.447659457382
	1	361	2.3733667239087	8,7865377313256

Table 7: T-test of DD and Z-value

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
DD	Equal variances assumed	0.001	0.981	0.000	0.8321163	0.036575	0.760385	0.903846
	Equal variances not assumed			0.000	0.8321163	0.036181	0.761042	0.903190
Z	Equal variances assumed	7.067	0.008	0.000	3.3576515	0.591187	2.198242	4.517060
	Equal variances not assumed			0.000	3.358E+15	0.529582	2.317613	4.397689

Calculated from SPSS

Although the DD and the Z do not follow the standard normal distribution as mentioned above, according to the central limit theorem, both of sample sizes are 2000, which is large enough. Thus the DD and Z-values can be regarded as the standard normal distribution. As shown in table 7, there are two independent variable sample tests, and the two-tailed significance level is around 0, which is less than 0.05. Therefore, in the T-test, both the KMV model and the Z-Score model can significantly distinguish the difference between ST companies and non-ST companies.

4.4.3 Mann-Whitney U test

The Mann-Whitney U test (also called Wilcoxon rank-sum test) is a non-parametric test that can be used to compare the distribution of two independent data. In this significance test, the Mann-Whitney test is used to compare the difference between two independent samples to determine whether the median values of the two independent samples is significantly different. Furthermore, it is further determined whether the KMV model and the Z-Score model can distinguish ST companies and non-ST companies.

Table 8: Descriptive statistics for the Mann-Whitney U test

Descriptive Statistics			
	Percentiles		
	25th	50th (Median)	75th
DD	2.02635807582	2.38300279500	2.7802811293
ST or not	0.00	0.00	0.00
Z	1.48665694374	3.02029544874	5.8657780931
ST or not	0.00	0.00	0.00

Table 9: Test statistics of DD and Z-value

Test Statistics[a]			
DD		Z	
Mann-Whitney U	90185.000	Mann-Whitney U	187442.000
Wilcoxon W	155526.000	Wilcoxon W	252783.000
Z	-20.704	Z	-10.913
Asymp. Sig. (2-tailed)	0.000	Asymp. Sig. (2-tailed)	0.000
a. Grouping Variable: ST or not		a. Grouping Variable: ST or not	

Calculated from SPSS

The table 9 reflects the test results of the KMV model and the Z-score model. Through the Mann-Whitney test, it can be noted that the significant 2-tailed values of the KMV model or the Z score model are approximately 0, which rejects the null hypothesis, indicating that these two models can distinguish the difference between ST companies and non-ST companies.

4.4.4 ROC curve

ROC curve is used to test the results of different categories (especially two categories) to determine the applicability. Therefore, the ROC curve can be used to test the judgment ability of the KMV model and the Z-Score model for ST companies in this thesis.

Here are steps of ROC curve: Firstly, import 2000 data from DD value obtained by KMV model and 2000 data from Z-values obtained by Z-Score into SPSS20.0 software. Secondly, mark non-ST data are number 0, ST data are number 1. It is important to note that the ST here is the data of the company that was marked ST that year, in other words, there are a total of three hundred and sixty-one 1. Thirdly, adjust the setting conditions of the ROC curve. The larger DD or Z-value is, the closer it is to the non-ST group, which is 0; On the contrary, the smaller DD or Z-value it is, the closer it is to the ST group, which is 1. Fourthly, perform ROC curve drawing.

Test result variable: Under the non-parametric assumption, DD and Z-values will be compared with the ROC curve indicator, diagonal. And the null hypothesis is the actual area = 0.5.

From the ROC graph in figure 8, the two curves are located on the upper left of the diagonal line, indicating that these two models have the obvious discriminative ability. In the ROC curve, it is not difficult to observe that the line of DD is significantly higher than that of Z, indicating that the discrimination ability of the KMV model is better than the Z-Score model in general.

Figure 8: ROC curves of DD and Z-value

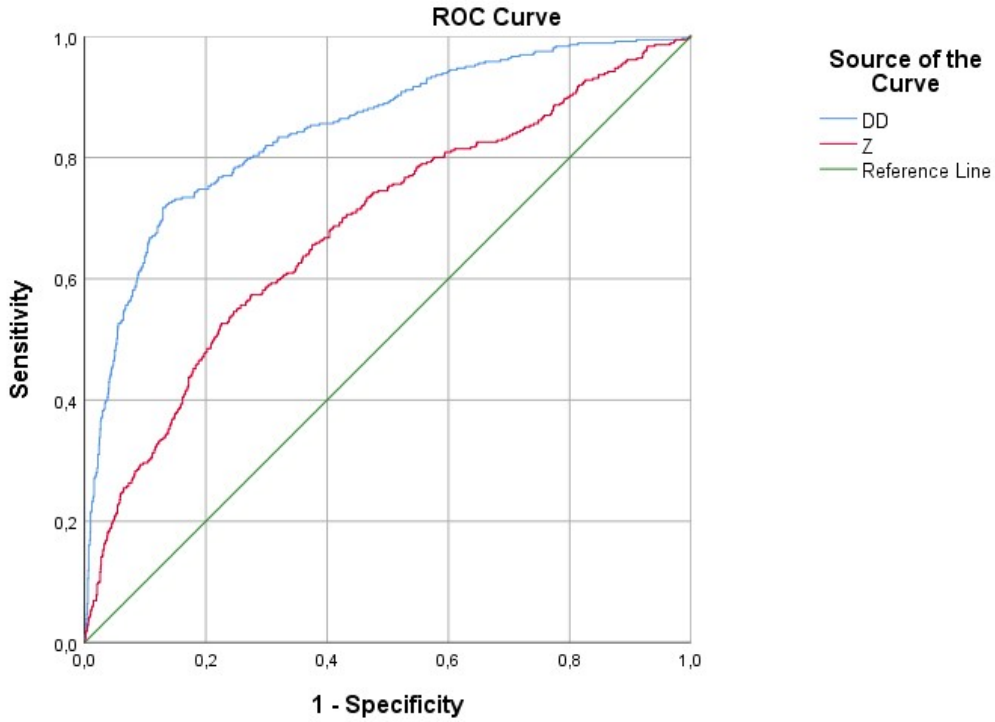


Table 10: AUC of DD and Z-value

Area Under the Curve					
				Asymptotic 95% Confidence Interval	
Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
DD	0.848	0.012	0.000	0.824	0.871
Z	0.683	0.016	0.000	0.651	0.715

a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Table 10 shows the area and related statistics. According to the theory of Hosmer, Lemeshow and Sturdivant (2013), the AUC of DD is 0.848 with a 95% confidence interval, which indicates that the KMV model has an excellent ability to distinguish ST companies and non-ST companies. In contrast, Z is relatively smaller and only 0.683. Compared with the KMV model, the discriminating ability of the Z-Score model is weaker.

4.5 Analysis of the prediction ability of KMV model and Z-Score model

After the above three adaptability tests, the KMV model and the Z-Score model can significantly identify ST companies and non-ST companies, and the KMV model has a better ability to identify. In this part of the analysis, the authors want to explore the predictive ability of the KMV model and the Z-Score model respectively, that is the sensitivity of these two models to the changes in credit risk data during the ST company's transition from the non-ST state to the ST state.

In this part, the prediction ability of the two models is judged by observing the data changes in the three years before ST company is judged as ST. There are 55 ST companies and 55 non-ST companies selected, which are from the 100 ST companies and 100 non-ST companies since during the 10-year sample range, only 55 companies have three consecutive years of data before being marked ST for the first time. We then selected these 55 ST companies and their corresponding 55 non-ST companies. These ST companies were marked ST at different times, which does not influence our research results. However, there is no denying that ST companies already had severe financial problems in the one or two years before they were marked as ST companies generally, and the deterioration of business conditions appeared gradually. Therefore, a good model should be able to accurately sense the trend changes one or two years before ST. Here, the year the ST company was marked as ST as time T. Correspondingly the year one year before T is as T-1, the year two years before T as T-2, and the year three years before T as T-3. In this way, the ROC curve can intuitively observe the identification ability of the KMV model and the Z-Score model at time T-1, T-2, and T-3, so that to compare the prediction accuracy of the KMV model and Z-Score model before companies were marked ST. If one model can accurately identify changes in ST company's credit risk data before the company was marked, investors can use this model to predict whether an ordinary company will become an ST company or not in the future. It also helps investors and commercial banks to avoid unnecessary risk. In order to obtain the predictive power of the model, the authors used the DD and Z values of ST and it's matching non-ST companies' at times T-1, T-2, and T-3. The results of the prediction ability test are shown in table 11 and figure 9. In this thesis, all the ROC and AUC diagrams are obtained from SPSS.

Figure 9: ROC curves of DD and Z-value for 3 years

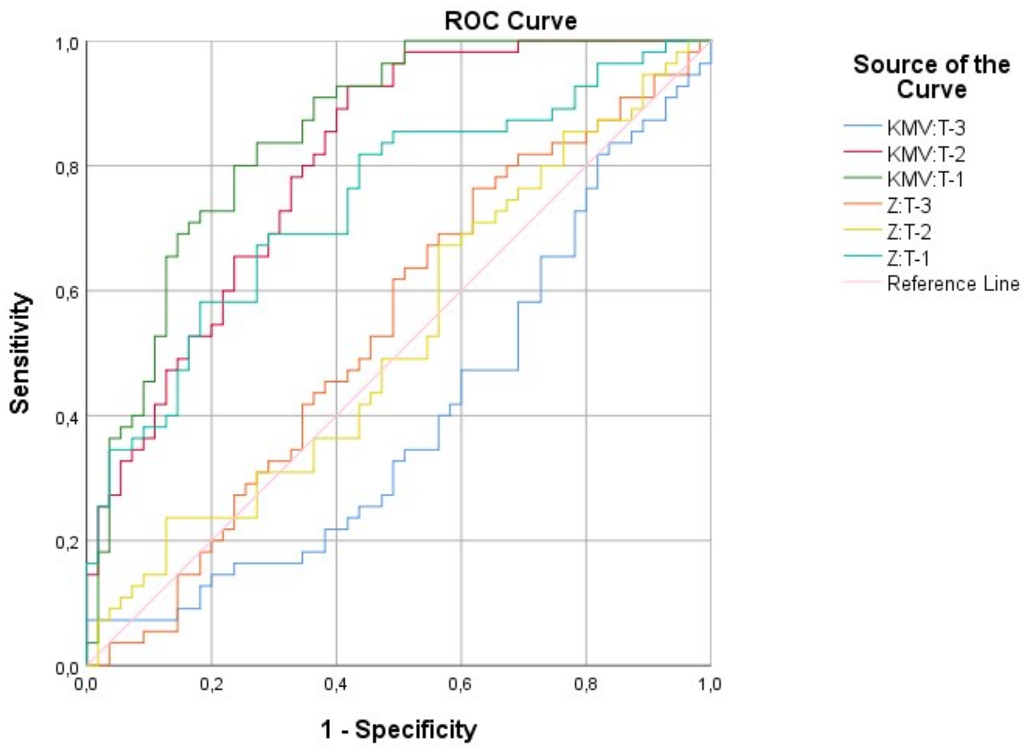


Table 11: AUC of DD and Z-value in prediction

Area Under the Curve					
				Asymptotic 95% Confidence Interval	
Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
KMV:T-3	0.400	0.055	0.070	0.293	0.507
KMV:T-2	0.806	0.041	0.000	0.725	0.886
KMV:T-1	0.852	0.036	0.000	0.781	0.923
Z:T-3	0.540	0.055	0.475	0.431	0.648
Z:T-2	0.525	0.055	0.647	0.417	0.634
Z:T-1	0.741	0.047	0.000	0.649	0.834

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Calculated from SPSS

Figure 9 and table 11 show that the ROC curves and AUC values of the DD and Z-values for the 55 ST companies and their corresponding non-ST companies at time T-1, T-2, and T-3.

Figure 9 depicts the ROC curve of KMV: T-1, KMV: T-2, Z: T-1, Z: T-3, Z: T-2 and KMV: T-3 (This is in order of occurrence from the top left corner to the bottom right corner in this graph). The figure shows that three of the curves are entirely on the upper left of the diagonal line, indicating that only lines of KMV: T-2, KMV: T-1, and Z: T-1 have the significant discriminant ability. Of which, KMV: T-2 is significantly higher than Z: T-1, indicating that identification ability of the KMV model at T-1 is better than that of the Z-Score model at T-1. That the position of the ROC curve of KMV: T-2 is lower than that of KMV: T-1 shows that the closer the time is to the company being ST, the better the prediction accuracy of the KMV model will be. The line of Z: T-1 is far above the other two Z-Score lines, which indicates that Z-Score has a relatively good perception ability for the year before the company is marked as ST.

For the other three curves Z: T-3, Z: T-2, and KMV: T-3, it is clear to see that the two lines of Z: T-2 and Z: T-3 are interlaced with each other. Normally, T-3 should be below T-2. It is enough to indicate that Z-Score does not have a good predictive ability to predict the changes from three years before the ST year to two years before the ST year. The curve of KMV: T-3 is below the diagonal, which means this model does not have any identification ability for two different types of companies back then, whereas it can be seen normal to some extent since in the year which is three years before the company was marked ST these two types of companies are not supposed to have much difference.

Table 11 shows 6 lines of AUC. According to the Hosmer, Lemeshow and Sturdivant (2013), the AUC of KMV: T-1 and KMV: T-2 are both higher than 0.8, which shows that the KMV model has an excellent capability and can distinguish between ST companies and non-ST companies at T-1 and T-2. Similarly, the AUC of Z: T-1 is 0.741, indicating that although its predictive ability is not as good as the KMV model at T-1 and T-2, it is also acceptable. The AUC of KMV: T-3, Z: T-3 and Z: T-2 are less than 0.7, indicating that their prediction ability for ST is not very good, and can almost be seen no prediction ability. To sum up, the KMV model has a better prediction ability than Z-Score.

The conclusion of this paper is similar to that of Hsiao and Gao's (2016) that the KMV model is more predictive than Altman's Z-Score model. At the same time, the result of this paper matches Liang's (2012) conclusion that KMV can better adapt to the Chinese stock market and give more accurate predictions. Miller (2009) found that Distance to Default has superior ordinal and cardinal bankruptcy prediction power within their universe. It also has a more durable bankruptcy signal, but it generates fewer stable ratings than the Z-Score. Also, there are some scholars concluded that Z-Score performed better than KMV model. However, it has to be admitted that due to different adjustments of the KMV model and Z-Score model used by each scholar, there are more or less differences in the results of different studies.

The performance of a model should be examined from two dimensions: its ability of discrimination and its reliability of prediction. The reliability of model prediction is also crucial. Although the reliability of prediction will increase with the improvement of model identification ability, there will always be calibration errors. The proper use of the method directly affects the judgment of the model. In combination with the previous relevant literature, some scholars only focused on the ability of model identification, and the sample size was very small, with a time span of only 1-3 years.

5 Conclusion

The credit risk measurement of listed companies in China as the research content are taken in this thesis, and through theoretical analysis, the theoretical support of the credit risk measurement model is discussed. Through the comparative analysis of the credit risk measurement models, the feasibility of KMV model in China's capital market is obtained. Then through the applicability of modified KMV model, the KMV model suitable for the credit risk measurement of listed companies in China is constructed. From the identification and prediction compared with Z-Score model, the study shows that both results of the modified KMV model and the Z-Score model calculated between ST companies and non-ST companies have obvious differences, and both are significantly different. According to the ROC curve analysis between KMV model and Z-Score model, it can be concluded that KMV model is superior to Z-Score model in judgment ability. In terms of model prediction ability, KMV is also better than Z-Score.

Although we conclude in this paper that the KMV model is more applicable in the Chinese market, it still has some shortcomings:

First of all, it is urgent to establish a default group and a non-default group in the research sample to verify the validity of the model. The important condition for the application of the credit risk measurement model in a new market is whether it is effective in this market, and to test whether the model is effective or not has become the key to the academic research on the applicability of the model. However, in the current Chinese market, there is no official default database, which seriously restricts the validation of the model in academic research. In order to solve this problem, China’s official research institutions should establish its enterprise's default database as soon as possible. The database should be constructed with the attention of the regulatory agencies, and with the combination of the credit data of Banks and financial intermediary rating agencies. At present, the main basis for the "special treatment" of listed companies in China's stock exchange listing rules is from financial statements, which mainly includes indicators such as the net interest rate of stocks in financial statements and the authenticity of financial statements. These indicators, especially the net interest rate of the stock to a certain extent, can reflect the operating performance of listed companies and the size of the credit risk, but the standard of implementing special treatment of the listed company only depends on these index would easily misjudge the validity of the model because of the error on the definition. To establish the prior sample data with strong credibility to verify the validity of the credit risk model, an authoritative credit risk experience database is indispensable. Moreover, the existence of empirical data is also the basis for KMV model to establish the corresponding relationship between default distance and default probability. In this paper, the default distance is used as the risk score to measure the credit risk, but there is no clear and accurate criterion for the default distance. If the corresponding relationship between the default distance and the default probability can be found, the KMV model will be more suitable for China.

Secondly, for the formula of default point in KMV model, no authoritative research in China has put forward a unified standard for it. In the future research, the empirical research on the specific formula of the default point should be strengthened, so as to better find the formula coefficient of the default point which is most suitable for Chinese enterprises. In the default point setting, the coefficient is related to the proportion of long-term debts. Especially for enterprises with more long-term debts, different coefficients have a great influence on the size of the default point. In KMV model, the default point plays an important role, which will directly affect the calculation of default distance. Therefore, finding the default point coefficient most suitable for Chinese enterprises is also an important aspect to enhance the applicability of KMV model in China.

References

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance*, vol. 23, no. 4, pp.589–609.

Altman, E. I., Haldeman, R. G. & Narayanan, P. (1977). ZETATM Analysis A New Model to Identify Bankruptcy Risk of Corporations, *Journal of banking & finance*, vol. 1, no. 1, pp.29–54.

Arora, N., Bohn, J. R. & Zhu, F. (2005). “Surprise” in Distress Announcements: Evidence from Equity and Bond Markets. Available online: https://pdfs.semanticscholar.org/9866/f2fd41954a2ac8664f33acc98b0537171f27.pdf?_ga=2.131696716.2143825280.1586864147-1210547374.1586378322 [Accessed 9 April 2020].

Basel Committee. (2001). The Internal Ratings-Based Approach, *Supporting Document the CP2*. Available online: https://www.bis.org/publ/bcbsca05.pdf [Accessed April 15 2020].

Benos, A. & Papanastasopoulos, G. (2007). Extending the Merton Model: A Hybrid Approach to Assessing Credit Quality, *Mathematical and Computer Modelling*, vol. 46, no. 1, pp.47–68.

Blochwitz, S., Liebig, T. & Nyberg, M. (2000). Benchmarking Deutsche Bundesbank's Default Risk Model, the KMV Private Firm Model and common financial ratios for German corporations. In *Workshop on Applied Banking Research, Basel Committee on banking supervision*. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.424.1397&rep=rep1&type=pdf [Accessed 9 April 2020].

Boris, K., Ivana, W. & Anna, S. (2015). Quantification of Credit Risk with the Use of CreditMetrics, *Procedia Economics and Finance*, vol. 26, pp.311–316.

Cao, L., (2014). What Does The Chaori Default Mean For The Chinese Financial Markets?. CFA Institute Enterprising Investor. Available Online: https://blogs.cfainstitute.org/investor/2014/03/13/what-does-the-chaori-default-mean-for-the-chinese-financial-markets/ [Accessed 15 April 2020].

Crosbie, P. and Bohn, J., (2003). Modeling default risk. Moody’s KMV White Paper. Available online: https://www.moodysanalytics.com/-/media/whitepaper/before-2011/12-18-03-modeling-default-risk.pdf [Accessed 11 May 2020].

Crossen, C. and Zhang, X., (2012). Validating the public EDF model for global financial firms. *Moody’s Analytics*. Available online: https://www.moodysanalytics.com/-/media/whitepaper/2011/11-01-10-Financials-Validating-Public-EDF-Model.pdf [Accessed 9 April 2020].

Crouhy, M., Galai, D. & Mark, R. (2000). A comparative analysis of current credit risk models, *Journal of Banking & Finance*, vol. 24 (1-2), pp.59-117.

Derbali, A. (2018). How the Default Probability Is Defined by the CreditRisk+ Model?, Available online: https://hal.archives-ouvertes.fr/hal-01696011/document [Accessed 6 May 2020].

Derbali, A. & Hellara, S. (2013). Analysis of default probability: a comparative theoretical approach between the credit portfolio view model and the creditrisk+ model, *International Journal of Business Management & Research* (IJBMR), vol. 3, pp.157–170.

Derviz, A. and Kadlčáková, N. (2001). *Methodological problems of quantitative credit risk modeling in the Czech economy*, Česká národní banka, Sekce měnová. Available online: https://www.researchgate.net/profile/Alexis_Derviz/publication/237782802_METHODOLOGICAL_PROBLEMS_OF_QUANTITATIVE_CREDIT_RISK_MODELING_IN_THE_CZECH_ECONOMY/links/0c960532170512eafa000000.pdf [Accessed 11 May 2020].

Diaz, D. & Gemmill, G. (2002). A systematic comparison of two approaches to measuring credit risk: CreditMetrics versus CreditRisk+, *Trans. 27th ICA*. Available online: http://actuaires.org/EVENTS/Congresses/Cancun/ica2002_subject/credit_risk/credit_x_diazledezma.pdf [Accessed 11 May 2020]

Diez-Canedo, J. M. (2005). A simplified credit risk model for supervisory purposes in emerging markets, *Press & Communications CH-4002 Basel, Switzerland*, pp.328-360.

Fawcett T. (2006). An introduction to ROC analysis, *Pattern recognition letters*, vol. 27, no. 8, pp.861-874.

Frenkel, M., Hommel, U. & Rudolf, M. (2013). *Risk Management: Challenge and Opportunity*, Springer Science & Business Media.

Gapko, P. & ŠMíd, M. (2012). Dynamic Multi-Factor Credit Risk Model with Fat-Tailed Factors, *Finance a Uver: Czech Journal of Economics & Finance*, vol. 62, no. 2, pp.125–140.

Gundlach, M. & Lehrbass, F. (2013). *CreditRisk+ in the Banking Industry*, Springer Science & Business Media.

Hanley J A, McNeil B J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, 143(1), pp.29-36.

Jessen, C. & Lando, D. (2015). Robustness of Distance-to-Default, *Journal of Banking & Finance*, vol. 50, pp.493–505.

Hosmer, D. W., Sturdivant, R. X. and Lemeshow, S. (2013) *Applied logistic regression*. [Elektronisk resurs], 3rd ed. Wiley (Wiley Series in Probability and Statistics). Available online: https://search.ebscohost.com/login.aspx?direct=true&db=cat07147a&AN=lub.5907326&site=eds-live&scope=site [Accessed: 16 May 2020]

Hsiao, I.T. and Gao, L. (2016). Models of Bankruptcy Prediction Since the Recent Financial Crisis: KMV, Naïve, and Altman's Z-score. Available online: http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8877522&fileOId=8877523 [Accessed: 16 May 2020]

Kern, M. & Rudolph, B. (2001). Comparative Analysis of Alternative Credit Risk Models: An Application on German Middle Market Loan Portfolios, *CFS Working Paper*, 2001/03. Available online: https://www.econstor.eu/bitstream/10419/78084/1/755300432.pdf [Accessed 6 May 2020]

Kliestik, T., Misankova, M., Valaskova, K. & Svabova, L. (2017). Bankruptcy prevention: new effort to reflect on legal and social changes. *Science and Engineering Ethics*, Vol. 24, no.2, pp.791-803. Available Online: http://link.springer.com/10.1007/s11948-017-9912-4 [Accessed 9 April 2020].

Kollár, B. & Gondžárová, B. (2015). Comparison of Current Credit Risk Models, *Procedia Economics and Finance*, vol. 23, pp.341–347.

Kumar K, Tan C. (2005). Some recent developments in financial distress prediction, *Bulletin of the International Statistical Institute*.

Kurth, A., Taylor, H. & Wagner, A. (2002). An Extended Analytical Approach to Credit Risk Management, *Economic Notes*, vol. 31, no. 2, pp.237–253.

Lee, W.-C. (2011). Redefinition of the KMV Model's Optimal Default Point Based on Genetic Algorithms – Evidence from Taiwan, *Expert Systems with Applications,* vol. 38, no. 8, pp.10107–10113.

Liang, X., 2012. An empirical estimation of the default risk of Chinese listed company based on the Merton-KMV Model (Master's thesis). Available online: https://openaccess.nhh.no/nhh-xmlui/bitstream/handle/11250/169767/Liang_2012.pdf?sequence=1&isAllowed=y [Accessed 25 May 2020].

Listed Domestic Companies, Total - China | Data. (2020). Available Online: https://data.worldbank.org/indicator/CM.MKT.LDOM.NO?locations=CN [Accessed 11 May 2020].

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *The Journal of Finance*, vol. 29, no. 2, pp.449–470.

Miller, W., (2009). Comparing models of corporate bankruptcy prediction: Distance to default vs. Z-Score. (July1, 2009). Available online: https://pdfs.semanticscholar.org/5295/fafcc8f81455b18f13043f0632a72e99096a.pdf [Accessed: 26 May 2020]

Morgan, J.P., 1997. Creditmetrics-technical document, *JP Morgan, New York.* Available online: http://homepages.rpi.edu/~guptaa/MGMT4370.10/Data/CreditMetrics.pdf [Accessed: 10 April 2020]

Number of China Bond Defaults Touches Last Year's Record - Bloomberg. (2020). Available. online: https://www.bloomberg.com/news/articles/2019-10-18/number-of-china-bond-defaults-just-touched-last-year-s-record [Accessed 15 April 2020].

Odom M D, Sharda R. (1990). A neural network model for bankruptcy prediction, *1990 IJCNN International Joint Conference on neural networks*, pp.163-168. IEEE.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, vol. 18, no. 1, pp.109–131.

Ong, L. L. & International Monetary Fund (eds). (2014). *Guide to IMF Stress Testing: Methods and Models*, Washington, DC: International Monetary Fund.

Press, S. J. & Wilson, S. (1978). Choosing Between Logistic Regression and Discriminant Analysis, *Journal of the American Statistical Association*, vol. 73, no. 364, pp.699–705.

Saunders, A. & Allen, L. (2002). Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms, John Wiley & Sons, Incorporated, Hoboken. pp.107-120. Available from: ProQuest Ebook Central. [15 April 2020].

Szse.cn. (2016). Trading rules of Shenzhen stock exchange. Available online: https://www.szse.cn/English/rules/siteRule/P020181124401737559498.pdf [Accessed 26 April

2020].

Sobehart, J.R., Keenan, S.C. and Stein, R., (2000). Benchmarking quantitative default risk models: a validation methodology, *Moody's Investors Service*. Available Online: http://www.rogermstein.com/wp-content/uploads/53621.pdf [Accessed 9 April 2020]

Takezawa, N. & Takezawa, N. (2003). A Note on Credit Risk of Vertical Keiretsu Firms: Preliminary Evidence from the Japanese Automobile Industry, *Asia-Pacific Financial Markets*, vol. 10, no. 4, pp.377–398.

Vandendorpe, A., Ho, N.-D., Vanduffel, S. & Van Dooren, P. (2008). On the Parameterization of the CreditRisk + Model for Estimating Credit Portfolio Risk, *Insurance: Mathematics and Economics*, vol. 42, no. 2, pp.736–745.

Wang, C.-P., Qu, L.-J. & Li, J.-W. (2015). Commercial bank credit risk measurement based on KMV model studies, *In 2015 International conference on Engineering Management, Engineering Education and Information Technology*. Atlantis Press. Available Online: http://www.atlantis-press.com/php/paper-details.php?id=25841947 [Accessed 9 April 2020].

Wang, Y., Lin, L., Kuo, H.-C. & Piesse, J. (2010). *Handbook of quantitative finance and risk management*, Springer Science & Business Media. Available Online: http://link.springer.com/10.1007/978-0-387-77117-5_107 [Accessed 15 April 2020].

Wilson, R. L. & Sharda, R. (1994). Bankruptcy Prediction Using Neural Networks, *Decision Support Systems*, vol. 11, no. 5, pp.545–557.

Witzany, J. (2017). *In Credit Risk Management: Pricing, Measurement, and Modeling*. Springer, Cham.

Xu Z, Zhang S. (2004). The comparative research on volatility prediction ability of adjusted realized volatility, GARCH model and SV model, *Systems Engineering*, 8, pp.6-63.

Yang, Y., Li, L., Zhou, Z. & Fei, W. (2013). The Research on Applicability of Amended KMV Model with Different Industries, *Journal of Risk Analysis and Crisis Response*, vol. 3, no. 1, pp.27.

Zhang, Y. & Shi, B. (2016). Non-Tradable Shares Pricing and Optimal Default Point Based on Hybrid KMV Models: Evidence from China, *Knowledge-Based Systems*, vol. 110, pp.202–209.

Zhang, Y.Y. (n.d.), On Credit Risk Management Models: CreditMetrics vs. KMV. Available online: https://cpb-us-w2.wpmucdn.com/sites.wustl.edu/dist/0/590/files/2016/12/Yvonne_CreditMetrics-and-KMV-1nffz84.pdf [Accessed 11 May 2020].

Zokirjonov, M. (2017). Methodology of creditmetrics for credit risk assessment. *International Finance and Accounting*, 2018(2), pp.135. Available online: https://www.researchgate.net/publication/332781824_Methodology_of_creditmetrics_for_credit_risk_assessment [Accessed 11 May 2020]

Appendix

Table A1: Data samples; Left are ST companies and right are non-ST companies

Stock	Name	Industry	Stock	Name
000670.SZ	Infotmic Co., Ltd.	Semiconductor	002119.SZ	Ningbo Kangqiang Electronics Co., Ltd.
600275.SS	Hubei Wuchangyu Co.,Ltd.	Animal husbandry	000702.SZ	Hunan Zhenghong Science and Technology Develop Co.,Ltd.
000939.SS	KAIDI ECOLOGICAL	Electricity	600163.SS	Zhongmin Energy Co., Ltd.
600726.SS	Huadian Energy Company Limited	Electricity	600310.SS	Guangxi Guidong Electric Power Co., Ltd.
002289.SZ	Shenzhen Success Electronics Co., Ltd.	Electronic device	300032.SZ	Jinlong Machinery & Electronic Co.,Ltd
600462.SS	Shenzhen Geoway Co., Ltd.	Electronic device	600353.SS	Chengdu Xuguang Electronics Co., Ltd.
002188.SZ	Bus Online Co., Ltd.	Electronic component	002141.SZ	Infund Holding Co., Ltd.
600234.SS	Guanghe Landscape Culture Communication Co., Ltd.	Real estate services	000560.SZ	5i5j Holding Group Co., Ltd.
000897.SZ	Tianjin Jinbin Development Co., Ltd.	Real estate development	600533.SS	Nanjing Chixia Development Co., Ltd.
002147.SZ	Neoglory prosperity Inc.	Real estate development	000502.SZ	Lvjing Holding Co., Ltd.
000982.SZ	Ningxia Zhongyin Cashmere Co., Ltd.	Textile	002083.SZ	Sunvim Group Co.,Ltd.
000611.SZ	Inner Mongolia TianShou Technology&development, CO., LTD.	Textile	600689.SS	Shanghai Sanmao Enterprise (Group) Co., Ltd.
600696.SS	Shanghai Guijiu Co.,Ltd.	Non-bank finance	600318.SS	Anhui Xinli Finance Co., Ltd.
600399.SS	Fushun Special Steel Co.,Ltd.	Steel	000717.SZ	SGIS Songshan Co., Ltd.
600608.SS	Shanghai Broadband Technology Co.,Ltd.	Steel	600117.SS	XiNing Special Steel Co., Ltd.
002005.SZ	Elec-Tech International Co., Ltd.	Optoelectronics	002189.SZ	Costar Group Co., Ltd.
600807.SS	Jinan High-tech	Precious metals	000506.SZ	Zhongrun Resources

	Development Co., Ltd.			Investment Corporation
600687.SS	Gansu Gangtai Holding (Group) Co., Ltd.	Precious metals	600766.SS	YanTai Yuancheng Gold Co., Ltd.
600652.SS	Shanghai U9 Game Co.,Ltd.	Internet service	600242.SS	Zhongchang Big Data Corporation Limited
002113.SZ	Hunan Tianrun Digital Entertainment &Cultural Media Co.,Ltd.	Internet service	000038.SZ	Shenzhen Capstone Industrial Co.,Ltd.
002072.SZ	Kairuide Holding Co., Ltd.	Internet service	002247.SZ	Zhejiang Juli Culture Development Co.,Ltd.
600634.SS	Shanghai Fukong Interactive Entertainment Co.,Ltd.	Internet service	000835.SZ	Great Wall International ACG Co., Ltd.
000971.SZ	Gosun Holding Co., Ltd.	Internet service	300017.SZ	Wangsu Science & Technology Co.,Ltd.
000792.SZ	Qinghai Salt Lake Industry Co.,Ltd.	Fertilizer	600426.SS	Shandong Hualu-Hengsheng Chemical Co., Ltd.
600423.SS	Liuzhou Chemical Industry Co., Ltd.	Fertilizer	600731.SS	Hunan Haili Chemical Industry Co., Ltd.
000953.SZ	Guangxi Hechi Chemical Co., Ltd.	Fertilizer	600796.SS	Zhejiang Qianjiang Biochemical Co., Ltd.
600228.SS	JiangXi ChangJiu Biochemical Industry Co., Ltd.	Chemical materials	002165.SZ	Hongbaoli Group Corporation, Ltd.
000422.SZ	Hubei Yihua Chemical Industry Co., Ltd.	Chemical materials	000545.SZ	Gpro Titanium Industry Co., Ltd.
600091.SS	Baotou Tomorrow Technology Co., Ltd.	Chemical materials	600367.SS	Guizhou RedStar Developing Co.,Ltd.
000707.SZ	Hubei Shuanghuan Science and Technology Stock Co.,Ltd.	Chemical materials	600714.SS	Qinghai Jinrui Mineral Development Co., Ltd
002263.SZ	Zhejiang Great Southeast Corp.Ltd.	Chemicals	000973.SZ	FSPG Hi-Tech CO., Ltd.
600319.SS	Weifang Yaxing Chemical Co., Ltd.	Chemicals	002211.SZ	Shanghai Hongda New Material Co., Ltd.
000737.SZ	Nafine Chemical Industry Group Co.,Ltd.	Chemicals	000985.SZ	Daqing Huake Company Limited
000504.SZ	NanHua Bio-medicine Co., Ltd.	Environmental protection	000040.SZ	Tunghsu Azure Renewable Energy Co.,Ltd.
000820.SZ	Shenwu Energy Saving Co.,	Environmental	000005.SZ	Shenzhen Fountain

	Ltd.	protection		Corporation
600595.SS	Henan Zhongfu Industrial Co.,Ltd.	Basic metal	600961.SS	Zhuzhou Smelter Group Co.,Ltd.
600654.SS	China Security Co., Ltd.	Computer software	600355.SS	Routon Electronic Co., Ltd.
600701.SS	Harbin Gong Da High-Tech Enterprise Development Co.,Ltd.	Computer software	000638.SZ	Vanfund Urban Investment and Development Co., Ltd.
000010.SZ	Shenzhen Ecobeauty Co., Ltd.	Building Construction	600248.SS	Shaanxi Yanchang Petroleum Chemical Engineering Co., Ltd.
002200.SZ	Yunnan Yuntou Ecology and Environment Technology Co., Ltd.	Building Construction	000628.SZ	ChengDu Hi-Tech Development Co., Ltd.
600539.SS	Taiyuan Lionhead Cement Co.,Ltd.	Energy saving and environmental protection	300023.SZ	Bode Energy Equipment Co., Ltd.
600877.SS	CETC Energy Joint-Stock Co., Ltd.	New metal and non-metal materials	600172.SS	Henan Huanghe Whirlwind Co., Ltd.
600614.SS	Pengqi Technology Development Co., Ltd.	New metal and non-metal materials	002297.SZ	Hunan Boyun New Materials Co.,Ltd.
600265.SS	Yunnan Jinggu Forestry Co.,Ltd.	Forestry	600076.SS	Kangxin New Materials Co.,Ltd.
002102.SZ	Guanfu Holding Co., Ltd.	Trading	600278.SS	Orient International Enterprise, Ltd.
600145.SH	Xinjiang Yilu Wanyuan Industrial Investment Holding Co. Ltd.	Trading	000626.SZ	Grand Industrial Holding Co.,Ltd.
600149.SS	Langfang Development Co., Ltd.	Trading	000151.SZ	China National Complete Plant Import & Export Corporation Limited
600301.SS	Nanning Chemical Industry Co., Ltd.	Trading	000701.SZ	Xiamen Xindeco Ltd.
600870.SS	Xiamen Overseas Chinese Electronic Co.,Ltd.	Trading	600605.SS	Shanghai Huitong Energy Co.,Ltd.
600247.SS	Jilin Chengcheng Group Co.,Ltd.	Trading	600250.SS	Nanjing Textiles Import & Export Corp., Ltd.
600725.SS	Yunnan Yunwei Company Limited	Coal	600740.SS	Shanxi Coking Co., Ltd.

600408.SS	Shanxi Antai Group Co., Ltd.	Coal	000780.SZ	Inner Mongolia PingZhuang Energy Resources Co.,Ltd.
000911.SZ	Nanning Sugar Industry Co., Ltd.	Agriculture	600359.SS	Xinjiang Talimu Agriculture Development Co., Ltd.
000972.SZ	Chalkis Health Industry Co.,Ltd.	Agriculture	600540.SS	Xinjiang Sayram Modern Agriculture Co., Ltd.
002220.SZ	Dalian Tianbao Green Foods Co., Ltd.	Agriculture	600506.SS	XinJiang Korla Pear Co.,Ltd.
002190.SZ	Sichuan Chengfei Integration Technology Corp.Ltd.	Other electrical equipment	600379.SS	Shaanxi Baoguang Vacuum Electric Device Co., Ltd.
000981.SZ	Yinyi Co.,Ltd.	Vehicle	600686.SS	Xiamen King Long Motor Group Co., Ltd.
000572.SZ	Haima Automobile Co.,Ltd.	Vehicle	600081.SS	Dongfeng Electronic Technology Co.,Ltd.
000868.SZ	Anhui Ankai Automobile Co., Ltd.	Vehicle	000980.SZ	Zotye Automobile Co., Ltd.
600698.SS	Hunan Tyen Machinery Co., Ltd.	Vehicle	000757.SZ	Sichuan Haowu Electromechanical Co., Ltd.
000760.SZ	Steyr Motors Corp.	Vehicle	600213.SS	Yangzhou Yaxing Motor Coach Co., Ltd.
002259.SZ	Sichuan Shengda Forestry Industry Co., Ltd.	Gas	600333.SS	Changchun Gas Co., Ltd.
600856.SS	Changchun Sinoenergy Corporation	Gas	000669.SZ	Jinhong Holding Group Co., Ltd.
600817.SS	Zhengzhou Deheng Hongsheng Technology Co., Ltd.	Commercial property management	000007.SZ	Shenzhen Quanxinhao Co., Ltd.
002207.SZ	Xinjiang Zhundong Petroleum Technology Co., Ltd.	Petroleum gas	000637.SZ	Maoming Petro-Chemical Shihua Co., Ltd.
600290.SS	Huayi Electric Company Limited	Power transmission and transformation equipment	600192.SS	Lanzhou GreatWall Electrical Co., Ltd.
000806.SZ	Galaxy Biomedical Investment Co., Ltd.	Power transmission and	000533.SZ	Guangdong Shunna Electric Co., Ltd.

		transformation equipment		
000585.SZ	Northeast Electric Development Company Limited	Power transmission and transformation equipment	002184.SZ	Shanghai Hi-Tech Control System Co., Ltd.
600112.SS	Guizhou Changzheng Tiancheng Holding Co.,Ltd.	Power transmission and transformation equipment	002112.SZ	SAN BIAN SCIENCE& TECHNOLOGY Co., LTD.
002089.SZ	New Sea Union Technology Group Co.,Ltd.	Communication device	600293.SS	Hubei Sanxia New Building Materials Co., Ltd.
600289.SS	Bright Oceans Inter-Telecom Corporation	Communication device	002231.SZ	Allwin Telecommunication Co., Ltd.
000410.SZ	Shenyang Machine Tool Co., Ltd.	General Equipment	600765.SS	AVIC Heavy Machinery Co., Ltd.
002122.SZ	Tianma Bearing Group Co.,Ltd.	General Equipment	000530.SZ	Bingshan Refrigeration and Heat Transfer Technologies Co., Ltd.
000816.SZ	Jiangsu Nonghua Intelligent Agriculture Technology Co.ltd	General Equipment	000570.SZ	Changchai Company, Limited
002175.SZ	Oriental Times Media Corporation	General Equipment	600202.SS	Harbin Air Conditioning Co.,Ltd.
600179.SS	Antong Holdings Co., Ltd.	Logistics	600794.SS	Zhangjiagang Freetrade Science & Technology Group Co., Ltd.
600119.SS	Y.U.D.Yangtze River Investment Industry Co.,Ltd.	Logistics	002245.SZ	Jiangsu Aucksun Co., Ltd.
002210.SZ	Shenzhen Feima International Supply Chain Co., Ltd.	Logistics	300013.SZ	Jiangsu Xinning Modern Logistics Co.,Ltd.
002260.SZ	DEA General Aviation Holding Co. Ltd.	Home appliances	002035.SZ	Vatti Corporation Limited
002306.SZ	Cloud Live Technology Group Co.,Ltd.	Leisure services	002186.SZ	China Quanjude(Group) Co.,Ltd.
600209.SS	Lawton Development Co., Ltd.	Leisure services	000430.SZ	Zhang Jia Jie Tourism Group Co., Ltd.

600485.SS	Beijing Xinwei Technology Group Co., Ltd.	Medical service	002219.SZ	Hengkang Medical Group Co.,Ltd.
600721.SS	Xinjiang Baihuacun Co., Ltd.	Medical service	000150.SZ	Yihua Healthcare Co., Ltd.
600767.SS	Winsan (Shanghai) Medical Science and Technology Co., Ltd.	Medical service	002173.SZ	Innovation Medical Management Co., Ltd.
600084.SS	Citic Guoan Wine CO.,LTD	Beverage	600199.SS	Anhui Golden Seed Winery Co., Ltd.
600238.SS	HaiNan Yedao (Group) Co., Ltd.	Beverage	600962.SS	SDIC Zhonglu Fruit Juice Co.,Ltd.
000995.SS	Huangtai Wine-Marketing	Beverage	600573.SS	FuJian YanJing HuiQuan Brewery Co.,Ltd
000752.SZ	Tibet Galaxy Science & Technology Development Co., Ltd.	Beverage	000929.SZ	Lanzhou Huanghe Enterprise Co., Ltd
002086.SZ	Shandong Oriental Ocean Sci-Tech Co., Ltd.	Fishery	600257.SS	Dahu Aquaculture Co.,Ltd.
600518.SS	Kangmei Pharmaceutical Co., Ltd.	Chinese medicine production	600252.SS	Guangxi Wuzhou Zhongheng Group Co.,Ltd
600781.SS	FUREN Group Pharmaceutical Co., Ltd.	Chinese medicine production	000590.SZ	Tus-Guhan Group Corp.,Ltd.
600385.SS	Shandong Jintai Group Co., Ltd.	Jewelry	600086.SS	Eastern Gold Jade Co.,Ltd.
600891.SS	Harbin Churin Group Jointstock Co., Ltd.	Jewelry	000587.SZ	Jinzhou Cihang Group Co., Ltd.
600421.SS	Hubei Yangfan Holding Co., Ltd.	professional service	000056.SZ	Shenzhen Wongtee International Enterprise Co., Ltd.
600815.SS	Xiamen XGMA Machinery Company Limited	Professional setting	000680.SZ	Shantui Construction Machinery Co., Ltd.
600666.SS	Aurora Optoelectronics Co.,Ltd.	Professional setting	002278.SZ	Shanghai SK Petroleum & Chemical Equipment Corporation Ltd.
002021.SZ	Zoje Resources Investment Co., Ltd.	Professional setting	300022.SZ	Gifore Agricultural Science & Technology Service Co., Ltd.
000571.SZ	Sundiro Holding Co., Ltd.	Comprehensive	600784.SS	Luyin Investment Group Co., Ltd.

000409.SZ	ShanDong Geo-Mineral Co.,Ltd.	Comprehensive	000833.SZ	Guangxi Yuegui Guangye Holdings Co., Ltd.
600193.SS	Shanghai Prosolar Resources Development Co., Ltd.	Comprehensive	600128.SS	Jiangsu Holly Corporation

Source from WIND

Matlab2019 for DD, asset value and asset value volatility

```
folder='/Users/Jennifer/Desktop/data/data 10';
cd(folder)
%myFiles=dir([folder,'/*.xlsx'])
Initialdata=xlsread([folder,'/2011.xlsx']);
% Data selection
fai=Initialdata(:,8)';
DPT=Initialdata(:,11)';
DPT=DPT./1e+9;%
Sigma_e=Initialdata(:,13)';
v_e=Initialdata(:,4)';
v_e=v_e./1e+9;
r=0.0325;
%T=1;
Va=zeros(1,length(fai));
Sigma_a=zeros(1,length(fai));
DD=zeros(1,length(fai));


function F=fun1(x,DPT,r,Sigma_e,v_e)

F=[x(1)*normcdf((log(x(1)/DPT)+(r+0.5*x(2)^2))/x(2),0,1)-DPT*exp(-r)*normcdf((log(x(1)/DPT)+(r+0.5*x(2)
^2))/x(2)-x(2),0,1)-v_e;normcdf((log(x(1)/DPT)+(r+0.5*x(2)^2))/x(2),0,1)*x(1)*x(2)/v_e-Sigma_e];
end


for i=1:length(fai)
    %for a=1:9
        %for b=0:0.01:1
         x0=[1,0.5];
           VathetaX=fsolve(@(x)fun1(x,DPT(i),r,Sigma_e(i),v_e(i)),x0);
           Va(i)=VathetaX(1);
           Sigma_a(i)=VathetaX(2);
        %[Va(i),Sigma_a(i)]=fun2(DPT(i),r,T,Sigma_e(i),v_e(i),x0);
             DD(i)=(Va(i)*(1+fai(i))-DPT(i))/(Va(i)*(1+fai(i))*Sigma_a(i));
        %end
    %end
end

Va=Va'*1e+9;
Sigma_a=Sigma_a';
DD=DD';
```

Table A2: Normal distribution test of DD and Z-value

One-Sample Kolmogorov-Smirnov Test			
		DD	Z
N		2000	2000
Normal Parameters[a,b]	Mean	2.411127	5.124962
	Std. Deviation	0.705722	10.24761
Most Extreme Differences	Absolute	0.066	0.261
	Positive	0.066	0.232
	Negative	-0.048	-0.261
Test Statistic		0.066	0.261
Asymp. Sig. (2-tailed)		0.000[c]	0.000[c]

a. Test distribution is Normal.
b. Calculated from data.
c. Lilliefors Significance Correction.

Calculated from SPSS