

Structural Credit Risk Models with Microstructure Noise: An Empirical Analysis for China

Tang Jia*

April 25, 2011

Abstract

In this paper a Bayesian Markov chain Monte Carlo (MCMC) method discussed in Huang and Yu (2010) is applied to estimate the credit risk models with microstructure noise, using the daily equity data from China. In literature, the observed equity prices are known to be influenced by market microstructure effects so that they deviate from the corresponding efficient prices. Credit risk models with microstructure noise is a way to depict this relationship. In the Bayesian framework, we employ Gibbs sampling, which is a Markov chain Monte Carlo (MCMC) technique, to analyze such models. We estimate the model with Gaussian iid microstructure term, using equity data of the firms in the Shanghai Stock Exchange 50 index constitutes. Estimates in the model converge well when we use the data of 6 firms out of 16 in our sample.

Keywords: Credit risk, Bayesian MCMC, Microstructure noise, Default probability

*School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Tel.: +65-96114751. E-mail: jia.tang.2009@me.smu.edu.sg. Supervised by Professor Yu Jun. Committee members: Professor Tse Yiu Kuen and An Sungbae.

Contents

1	Introduction	3
2	Literature review	4
2.1	Credit risk models	4
2.1.1	Structural models	4
2.1.2	Reduced-form models	5
2.1.3	Comparison of the two types of models	5
2.2	Microstructure noise analysis	5
2.3	MCMC methods	6
2.3.1	Metropolis-Hastings algorithm	6
2.3.2	Gibbs sampling	7
3	Model and Methodology	8
3.1	Merton's model with microstructure noise	8
3.2	Probability of default	9
3.3	Method of Bayesian estimation via MCMC	10
3.4	MCMC convergence diagnostics	11
4	Empirical Analysis	12
4.1	Data	12
4.2	Settings for the estimation via MCMC	13
4.3	Results	13
5	Conclusions	18

1 Introduction

Credit risk arises for an investor when a borrower fail to meet its obligations in accordance with agreed terms. In this paper, we mainly focus on the period of 2007-2008 when the global economy was affected by several critical events, such as the collapses of Bear Stearns, the government takeover of AIG, Fannie Mae and Freddie Mac and the bankruptcy of Lehman Brothers. Even though the two markets are far apart from each other, the turmoil in American financial market affects Chinese financial market in a large degree due to the global investment of many financial institutes in China. For example, after the bankruptcy of Lehman Brothers, its bond hold by Bank of China, Industrial and Commercial Bank of China and many other Chinese banks becomes totally worthless.

Huang and Yu (2010) explain how the Bayesian Markov chain Monte Carlo (MCMC) method can be used to estimate credit risk models with microstructure noise. Structural models which were first introduced by Merton (1974) have been used widely in credit risk analysis, such as credit spread and default probability. But the difficulty arises since the firm's asset value cannot be directly observed. One way to estimate parameters in this class of models is the transformed-data maximum likelihood estimation (MLE) method proposed by Duan (1994) which is based on the one-to-one relationship between equity and asset levels when there is no trading noise. Nevertheless, it has been shown in literature that observed equity prices may deviate from their true value if the effect of microstructure noise, such as infrequent trading, asymmetric information, is taken into account. In this case, the one-to-one relationship between equity and asset values is broken. In order to estimate the models with microstructure noises, Duan and Fulop (2009) suggests the simulation-base ML method while Huang and Yu (2010) discusses the Bayesian MCMC method. Giving similar results by the two approaches, which is shown in Huang and Yu (2010), the Bayesian MCMC method gives exact finite sample inference and is more computational efficient. Therefore, in this paper we apply the Bayesian MCMC method to investigate the effect of microstructure noise in Chinese stock market.

The MCMC technique we use is based on the general Bayesian approach with posterior computations performed by Gibbs sampling. The relationship between the contaminated traded equity value and firm's asset value is represented by the option pricing model with the perturbation of microstructure noises. Meanwhile, the firm's asset process follows Merton's model, that is, a geometric Brownian motion. Since both the efficient prices and microstructure noises are unobservable in such models, the asset's value equation and the equity equation together can be regarded as a nonlinear state space model, to which the MCMC is applied. Compared to filtering methods, Bayesian MCMC provides the exact solution to the smoothing problem of inferring about the unobserved equity value by including the latent variables in the parameter space. We consider three forms of model in this paper, i.e. the model with Gaussian iid microstructure term, the model with

student-t noise and the one with correlation between the noise and the innovation to the equity value. Sample used in the empirical analysis are data of Blue chips constituting Shanghai Stock Exchange (SSE) 50 Index over the two years 2007-2008.

2 Literature review

2.1 Credit risk models

In the literature that try to analysis default processes for debt or other defaultable financial instruments, there are generally two streams of models, i.e., structural and reduced-form (or intensity) models. Among the studies, some basic ones of structural models include Merton (1974) and Black and Cox (1976), etc. And some important analysis of reduced-form models are Jarrow and Turnbull (1995), Lando (1998), Duffie and Singleton (1999) and so on. In this section, we briefly discuss the key features of the two types of models.

2.1.1 Structural models

Structural models determine the default time using the evolution of firms' structural variables, such as asset and debt values. These models completely rely on the rule of a firm's asset value is divided between the shareholders and creditors, which means, the default time in these models depends on the capital structure of the firm.

This kind of models is first built by Merton (1974) through considering contingent claim approach (i.e. option pricing approach). Merton assumes a firm has two kinds of claims outstanding which are an equity and a zero-coupon debt with face value F and maturity T , and the asset value of the firm V_t is the sum of the equity value and the debt value. The firm's stock is regarded as an European call option with maturity T and strike price F . At the maturity date, if the firm's asset value is large than the face value of the debt ($V_T > F$), then the stockholders are able to afford the expenses to the creditors and thus to keep the firm. A popular commercial product, KMV model, is based on Merton's model. Otherwise, if the firm's asset value is less than the face value of the debt ($V_T < F$), the default occurs. Structural default models provide a link between the credit quality of a firm and the firm's economic and financial conditions. In this paper, the model based on the Merton's method is used.

Many extensions to the Merton's model relax one or more its assumptions, introducing assumptions that are more fit to the reality and easy to implement.

Black and Cox (1976) adjust Merton's model, bringing First Passage Models. In Merton's model, default is assumed to happen only at the maturity date. However, this not necessary true in the real world. To break this limit, Black and Cox (1976) relax this assumption, supposing defaults occur as soon as firm's asset value falls below a certain threshold, that is, default can occur at any time.

2.1.2 Reduced-form models

Reduced form approach, comparing with structural approach, does not consider that there is an explicit connection between default and firm value and regards default probability as an exogenous variable. Since this group of models reduced the economics background that leads to default, they are called reduced-form models. The time of default in intensity models is the first jump of an exogenously given jump process rather than being determined by the value of the firm. We get the inference of the parameters governing the default hazard rate from market data.

Jarrow and Turnbull (1995) gives a reduced form approach, in which companies are classified by their credit risk and default models which belong to Poisson process are established for the firms respectively. Later, Jarrow, Lando, and Turnbull (1997) extend the discrete time reduced-form model to continuous time, establishing a Markov model for the term structure of credit risk spreads. Then Lando (1998) allows for correlation between default intensity and interest rates.

Another kind of reduced-form models is provided by Duffie and Singleton (1999), which assume that the recovery rate is a fraction of the market value of the risky debt prior to default.

2.1.3 Comparison of the two types of models

Defaults are endogenously generated within the structural model instead of exogenously given as in the reduced approach. The treatment of recovery rates is another difference between the two approaches: whereas the value of the firm's assets and liabilities at default will determine recovery rates in structural models, in reduced models recovery rates are exogenously specified.

2.2 Microstructure noise analysis

Market microstructure investigates the process by which investors' latent demands are finally translated into transactions, influencing prices and volumes. It has been shown in the market microstructure literature, such as Harris (1990), Hasbrouck (1993) and Madhavan, Richardson, and Roomans (1997), that observed equity prices can deviate from their equilibrium values due to microstruc-

ture noise, for example, bid-ask spread, asymmetric information and illiquidity. Using structural credit risk models, if we ignore trading noise, the estimated asset volatility could be higher than the true value, and the estimates for credit spreads, default probabilities and other corporate contingent claims can be misleading. As Duan and Fulop (2009) and Huang and Yu (2010) show the microstructure condition using the US data, it is natural for one to consider applying the microstructure noise analysis to an emerging market. Such a market is usually more affected by microstructure noise due to its inefficiency and information asymmetry, etc. Therefore, in this paper we use Chinese data to estimate structural credit risk models with microstructure noise.

2.3 MCMC methods

Markov chain Monte Carlo (MCMC) methods aim to sample a target distribution $\pi(\cdot)$, which is usually multivariate, by constructing a Markov chain with invariant distribution $\pi(\cdot)$. The characteristics of $\pi(\cdot)$ (such as finite sample properties) are studied by investigating sample path averages of this Markov chain. MCMC methods can be used in both Bayesian and frequentist statistical inference. Bayesians employ the methods to integrate over the posterior distribution of model parameters given the data, and frequentists may use the technique to integrate over the distribution of observable given parameter values. Bayesian analysis is particularly much influenced by the fast development of MCMC technique because it usually investigates high dimensional integrals that can hardly be analyzed by more conventional numerical methods. Two widely used schemes for stochastic simulation using Markov chains, the Metropolis-Hastings sampling method and the Gibbs sampling algorithm, are summarized below.

2.3.1 Metropolis-Hastings algorithm

Metropolis-Hastings (M-H) algorithm was proposed and studied by Metropolis, Rosenbluth, Rosenbluth, Teller, Teller, et al. (1953) in physics and generalized to the field of statistics by Hastings (1970). Chib and Greenberg (1995) gives an intuitive explanation of this algorithm.

The goal is to sample from the target density $\pi(\cdot)$ of which the transition kernel is unknown. Therefore, we intend to find a transition kernel $P(x, dy)$ whose n th iteration converges to $\pi(\cdot)$ for n large enough. The transition kernel is supposed to contain a continuous and a discrete part, say,

$$P(x, dy) = p(x, y)dy + r(x)\delta_x(dy),$$

for some function $p(x, y)$, where $p(x, x) = 0$, $\delta_x(dy) = 1$ if $x \in dy$ and 0 otherwise, and $r(x) = 1 - \int p(x, y)dy$. Tierney (1994) shows that if the function $p(x, y)$ satisfies the reversibility condition

$$\pi(x)p(x, y) = \pi(y)p(y, x),$$

then $\pi(\cdot)$ is the invariant density of $P(x, \cdot)$. The M-H algorithm employs the candidate-generating density $q(\theta, \theta^*)$ which is used to get a candidate value θ^* given the current value θ . And we denote the function

$$\alpha(\theta, \theta^*) = \begin{cases} \min \left[\frac{\pi(\theta^*)q(\theta, \theta^*)}{\pi(\theta)q(\theta^*, \theta)}, 1 \right] & \text{if } \pi(\theta)q(\theta, \theta^*) > 0, \\ 1 & \text{otherwise.} \end{cases}$$

The M-H algorithm follows the steps below:

1. Choose an arbitrary initial value of $\theta^{(0)}$.
2. From $t = 0, 1, \dots, T$, generate a new sample θ^* from the proposal distribution $q(\theta^t, \cdot)$, and u from the uniform distribution $\mathcal{U}(0, 1)$. Then we draw

$$\theta^{(t+1)} = \begin{cases} \theta^* & \text{if } u \leq \alpha(\theta^{(t)}, \theta^*), \\ \theta^{(t)} & \text{otherwise.} \end{cases}$$

3. Return the values $\{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(T)}\}$

Here we present the case with single parameter θ , but the procedure of the M-H algorithm for higher-dimensional θ is similar to it. Relevant discussion can be found in Chib (2001).

2.3.2 Gibbs sampling

Gibbs sampling is devised by Geman and Geman (1984) who use it to analyze Gibbs distributions in the context of image processing. However, this sampling method did not gain much attention from statisticians until Gelfand and Smith (1990) discuss its potential for more conventional statistical problems. A good introduction of this method is Casella and George (1992) in which simple examples are provided too.

Gibbs sampling is a special case of the M-H algorithm and is used to sample from each conditional distribution without calculating the density directly which may be very complicated. Usually there is a joint density and we are interested in getting characteristics (e.g. mean and variance) of the marginal density which is the integral of the joint density. However, in many cases, it is hard to evaluate this integral, either analytically or numerically. Gibbs sampling is a technique that solves this integral numerically. This method is especially useful in Bayesian procedures.

Suppose there is a parameter vector $\theta = (\theta_1, \theta_2, \dots, \theta_k)'$. $p(\mathbf{y}|\theta)$ is the likelihood, and $\pi(\theta)$ is the prior distribution. The full posterior conditional distribution of $\pi(\theta_i|\theta_j, i \neq j, \mathbf{y})$ is proportional to the joint posterior density, that is,

$$\pi(\theta_i|\theta_j, i \neq j, \mathbf{y}) \propto p(\mathbf{y}|\theta)\pi(\theta)$$

Gibbs algorithm works as follows:

1. Choose an arbitrary initial value of $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$.
2. From $t = 0, 1, \dots, T$, draw
 - $\theta_1^{(t+1)}$ from $\pi(\theta_1 | \theta_2^{(t)}, \theta_3^{(t)}, \dots, \theta_k^{(t)}, \mathbf{y})$
 - $\theta_2^{(t+1)}$ from $\pi(\theta_2 | \theta_1^{(t+1)}, \theta_3^{(t)}, \dots, \theta_k^{(t)}, \mathbf{y})$
 - ...
 - $\theta_k^{(t+1)}$ from $\pi(\theta_k | \theta_1^{(t+1)}, \theta_2^{(t+1)}, \dots, \theta_{k-1}^{(t+1)}, \mathbf{y})$
3. Return the values $\{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(T)}\}$

3 Model and Methodology

3.1 Merton's model with microstructure noise

Merton's model shows the evolution of a firm's structural variables, say asset and debt values. In the model, the unobservable variable V_t , the firm's asset value at time t , is supposed to follow a geometric Brownian motion, which is controlled by the annualized drift rate μ and the volatility of the stock's returns σ .

$$\frac{dV_t}{V_t} = \mu dt + \sigma dW_t \quad (1)$$

After applying the Ito's lemma we get

$$d \ln V_t = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t \quad (2)$$

The discrete time form of the model is

$$d \ln V_{t+1} = \left(\mu - \frac{\sigma^2}{2}\right) h + \ln V_t + \sigma \sqrt{h} \varepsilon_t \quad (3)$$

where $\varepsilon_t \stackrel{d}{\sim} i.i.d.N(0, 1)$, and h is the fixed sampling interval. F is the face value of a zero-coupon debt and T the maturity.

The observable equity prices are priced as a call option following the model in Black and Scholes (1973)

$$S_t \equiv S(V_t; \sigma) = V_t \Phi(d_t) - F e^{-r(T-t)} \Phi(d_t - \sigma \sqrt{T-t}) \quad (4)$$

where

$$d_t = \frac{\ln(\frac{V_t}{F}) + (r - \sigma^2)(T - t))}{\sigma \sqrt{T - t}}$$

Considering different forms of microstructure effects, Huang and Yu (2010) examines three types of the generalized Merton's models.

Model 1:

$$\begin{aligned} \ln S_t &= \ln S(V_t; \sigma) + \delta v_t, & v_t &\stackrel{d}{\sim} i.i.d.N(0, 1) \\ d\ln V_{t+1} &= (\mu - \sigma^2/2)h + \ln V_t + \sigma \sqrt{h}\varepsilon_t, & \varepsilon_t &\stackrel{d}{\sim} i.i.d.N(0, 1) \end{aligned} \quad (5)$$

where $\text{corr}(v_t, \varepsilon_t) = 0$.

Model 2:

$$\begin{aligned} \ln S_t &= \ln S(V_t; \sigma) + \delta v_t, & v_t &\stackrel{d}{\sim} t_k \\ d\ln V_{t+1} &= (\mu - \sigma^2/2)h + \ln V_t + \sigma \sqrt{h}\varepsilon_t, & \varepsilon_t &\stackrel{d}{\sim} i.i.d.N(0, 1) \end{aligned} \quad (6)$$

where $\text{corr}(v_t, \varepsilon_t) = 0$.

Model 3:

$$\begin{aligned} \ln S_t &= \ln S(V_t; \sigma) + \delta v_t, & v_t &\stackrel{d}{\sim} i.i.d.N(0, 1) \\ d\ln V_{t+1} &= (\mu - \sigma^2/2)h + \ln V_t + \sigma \sqrt{h}\varepsilon_t, & \varepsilon_t &\stackrel{d}{\sim} i.i.d.N(0, 1) \end{aligned} \quad (7)$$

where $\text{corr}(v_t, \varepsilon_t) = \rho$.

These models can be viewed in a state-space framework. To be specific, the first equation in each of the models is the observation equation and the second one the state equation.

3.2 Probability of default

One of the important credit applications based on credit risk models is probability of default (PD), which has been widely used in credit rating of firms.

It is the degree of likelihood that the borrower of a loan or debt will not be able to make the necessary scheduled repayments.

Following Duan and Fulop (2009), PD is described as a function of the unobserved asset value at the last time point of the sample.

$$p(V_n; \theta) = \Phi\left(\frac{\ln(\frac{F}{V_n}) - (\mu - \frac{\sigma^2}{2})(T - \tau_N)}{\sigma \sqrt{T - \tau_N}}\right) \quad (8)$$

where $t = \tau_i, i = 0, \dots, N$ are the time points that equity prices S_t are sampled. N is the sample size and the sampling interval h is $\tau_i - \tau_{i-1}$

3.3 Method of Bayesian estimation via MCMC

Let $\mathbf{X} = (\ln S_1, \dots, \ln S_n)'$, $\mathbf{h} = (\ln V_1, \dots, \ln V_n)'$ and $\theta = (\mu, \sigma, \delta)'$, then the likelihood function of model 1 is

$$p(\mathbf{X}; \theta) = \int p(\mathbf{X}, \mathbf{h}; \theta) d\mathbf{h} = \int p(\mathbf{X}|\mathbf{h}; \theta) p(\mathbf{h}; \theta) d\mathbf{h} \quad (9)$$

The nonlinear state space model can be Estimated with the maximum likelihood (ML) method developed by Duan and Fulop (2009) using a smoothed particle filter.

Alternatively, if Bayesian estimation is employed, as in Huang and Yu (2010), the joint posterior distribution is given by Bayes' theorem

$$p(\theta, \mathbf{h}|\mathbf{X}) \propto p(\mu, \sigma, \delta, \mathbf{h}) p(\mathbf{X}|\mu, \sigma, \delta, \mathbf{h}) \quad (10)$$

where the joint prior density is

$$p(\mu, \sigma, \delta, \mathbf{h}) = p(\mu) p(\sigma) p(\delta) p(\ln V_0) \prod_{t=1}^n p(\ln V_t | \ln V_{t-1}, \mu, \sigma) \quad (11)$$

as we assume the parameters μ , σ and δ are prior independent. And the likelihood function is

$$p(\mathbf{X}|\mu, \sigma, \delta, \mathbf{h}) = \prod_{t=1}^n p(\ln S_t | \ln V_t, \sigma) \quad (12)$$

In order to get the marginal distribution of the parameters from Equation (12), we need to deal with the integration of high-dimensional functions. One way to do it is Markov Chain Monte Carlo (MCMC) methods. We generate a random sample from a distribution by following a random walk for a long period (burn-in time) on a Markov chain. Finally, we will get a sample from its stationary distribution. If we repeat the process for many times and get a large amount of samples, by the ergodic theorem for Markov chains, we can calculate the characteristics of the marginal density by taking sample average to the desired degree of accuracy.

The Gibbs sampler samples one variate a time from the full conditional distributions. In our case, one iteration of all the univariate distributions is

1. Sample $\ln V_t$ from $\ln V_t | \ln V_{-t}, \mathbf{X}$
2. Sample $\sigma | \mathbf{X}, \mathbf{h}, \mu, \delta$
3. Sample $\delta | \mathbf{X}, \mathbf{h}, \mu, \sigma$
4. Sample $\mu | \mathbf{X}, \mathbf{h}, \sigma, \delta$

After initialize θ and \mathbf{h} , we implement the above iteration for several thousands of times, then the sample points will converge to a stationary distribution that is independent of the starting values. We discard the burn-in samples, for example, the first thousand sample points.

3.4 MCMC convergence diagnostics

The performance of MCMC methods is mainly evaluated by the convergence of the simulation which stands for the correct estimation of the posterior distribution of interest. Four diagnostic tests are often employed in practice and can be easily implemented using convergence diagnosis and output analysis (CODA) software in R. These tests presented here are Geweke diagnostics [Geweke (1992)], Gelman and Rubin diagnostics [Gelman and Rubin (1992) and Brooks and Gelman (1998)], Heidelberger and Welch diagnostics [Heidelberger and Welch (1981) and Heidelberger and Welch (1983)] and Raftery and Lewis diagnostics [Raftery and Lewis (1992) and Raftery and Lewis (1995)].

Geweke diagnostics checks whether the mean estimates have converged by comparing means from the early and latter part of the Markov chain. It is a two-sided test based on a Z-score statistic, where large absolute Z values implies rejection. Usually, if all Z values are within -2 and 2 , there is no differences in the means for the first and last sets of iterations.

Gelman and Rubin diagnostics tests whether parallel chains which have different initial values converge to the same target distribution. It is a one-sided test based on a variance ratio test statistic, where larger statistic values indicate rejection. One limitation of this test is that it is only applicable when there is more than one MCMC chain.

Raftery-Lewis test is designed to assess the accuracy of the estimated percentiles by investigating how many samples are needed to reach the desired accuracy of certain percentiles. If the Markov chain sample is more than the total samples needed, this means rejection. One limitation of this test is that it can only be used when chains have not been thinned.

The Heidelberger-Welch stationary test examines whether the Markov chain comes from a covariance stationary process. It is a one-sided test based on a Cramer-von Mises statistic where small p -values indicate rejection. The Heidelberger-Welch half-width test checks the Markov chain sample size to see whether it is adequate to estimate the mean values accurately. If the calculated relative half-width of the confidence interval is greater than a predetermined tolerance value, then we can conclude that the data are not enough to estimate the mean accurately.

In this paper, we use the Geweke and Heidelberger-welch convergence diagnostics in the empirical analysis, since our chains are thinned and one chain is tested at a time.

4 Empirical Analysis

4.1 Data

The data is chosen from the stocks that constitute the SSE 50 Index. Because these stocks are with good liquidity and representativeness and are most influential in Shanghai security market, microstructure noise is supposed to be smaller for these stocks comparing to other stocks in SSE market. If microstructure noise is negligible, supporting evidence should be given by this data set. We analyze the data during the two year 2007-2008. As the financial market in China may somehow affected by the financial crisis spreading worldwide in this period, it would be interesting to investigate the default probability conditions of the firms in China. Noticing that the SSE 50 index makes constituent adjustment every 6 months following principle of Stability and Dynamic Tracking. Therefore, the stocks we are interested in are those remained as constituents in the index since the end of year 2006 to March 2011.¹ After this procedure, there are 17 firms left. Among them, we take China Yangtze Power out of the dataset due to its one-year trade suspension from May 8, 2008 to May 18, 2009. So in the end, our sample consists of daily equity values of 16 firms over year 2007-2008. The 16 firms are: shanghai Pudong Development Bank, Wuhan Iron and Steel Company (WISCO), Hua Xia Bank, China Minsheng Banking Corporation, Baoshan Iron and Steel, China Petroleum and Chemical Corporation (SINOPEC), Citic Securities (CITICS), China Merchants Bank, China United Network Communications, SAIC Motor Corporation (SAIC), Kweichow Moutai, Liaoning Cheng Da, GD Power Development, Daqin Railway, Industrial and Commercial Bank of China and Bank of China.

Unlike the US stocks, Chinese stocks are not fully circulating. For all the 16 firms, they have tradable shares and non-tradable shares. Accordingly, the way to calculate equity values are different. The equity value of a firm in a day is the sum of two parts. The first part is the daily closing price² times tradable A-shares³ outstanding. And the second part is the net assets per share times non-tradable shares. The net assets per share, which is the shareholders' equity divided by the total

¹The stocks are actually only required to be remained in the SSE 50 Index during the two year period 2007-2008.

²The price that is adjusted backward for dividends and stock splits is used.

³A shares on the SSE refers to those that are traded in Renminbi, the currency in mainland China.

number of shares, is our proxy for the value of the non-tradable shares. When the ownership of non-tradable shares are transferred, the transferring price is base on the net assets per share with some premium.

All data are acquired from WIND database. The initial maturity of debt is 10 years. To get the proxy of the face value of the debt, we take the book value of debt of the firms at the year end of 2006 and compound it for 10 years at the risk-free interest rate. We use the mean of the one-year term deposit rate during 2007-2008, which is reported by The People's Bank of China, as the proxy of the risk-free interest rate. There are 488 daily observations in our sample for each firm, so we set $h = 1/244$

4.2 Settings for the estimation via MCMC

We follow the choice of priors and initial values in Huang and Yu (2010). That is, the parameters μ , σ and δ are assumed to be prior independent. And the priors are set to be $\mu \sim N(0.3, 0.25)$, $\sigma \sim IG(3, 0.0001)$ and $\delta \sim IG(2.5, 0.025)$. The initial values of μ , σ^2 and δ^2 are 0.3, 0.0001 and 0.02, respectively. For WISCO, SINOPEC, CITICS and SAIC, we discard the first 10000 sample points (the burn-in samples), implement 110000 iterations and save the values every 20th iteration. For Kweichow Moutai and Liaoning Cheng Da, the burn-in periods are a little longer, which are 20000 and 60000, respectively.

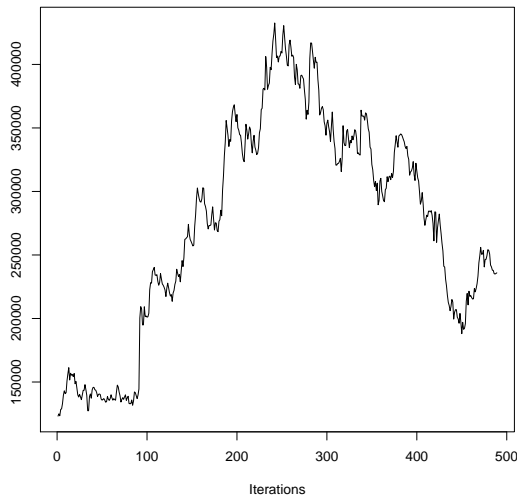
WinBUGS (Bayesian inference Using Gibbs Sampling) is a software for analyzing complex statistical models using MCMC methods. This software uses Gibbs sampling and the M-H algorithm to generate a Markov chain by sampling from full conditional distributions. We can also use R2WinBUGS package of R to call a BUGS model, summarize inferences and convergence in a table and graph, and save the simulations in arrays for later analysis.

4.3 Results

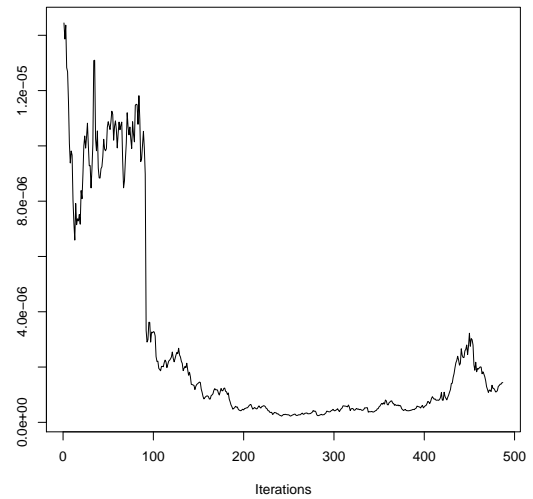
6 firms of the 16 in our sample have converging estimates using model 1 (equation 5) and the results are reported in Table 1. The first column is the abbreviation of names and their Shanghai Stock Exchange code, the second to the fourth columns contain the Bayesian MCMC estimates of posterior means and the estimates of posterior standard errors. As one nature of diffusion models, the drift term usually has large sampling errors, which is shown in the results.

Table 1: Bayesian estimation results for 7 companies in SSE 50 using model 1.

Company name (SSE code)	Estimates with trading noise (standard error)		
	μ	σ	$\delta \times 100$
WISCO (600005)	0.1073	0.5449	0.5385
	(0.3769)	(0.0196)	(0.1222)
SINOPEC (600028)	0.1924	0.2731	0.4662
	(0.1934)	(0.0103)	(0.0783)
CITICS (600030)	1.069	0.8868	0.5844
	(0.6034)	(0.0243)	(0.1467)
SAIC (600104)	0.1847	0.4452	0.5232
	(0.3156)	(0.0143)	(0.1084)
Kweichow Moutai (600519)	0.4397	0.4891	0.6042
	(0.3423)	(0.0181)	(0.1380)
Liaoning Cheng Da (600739)	0.9496	0.8892	0.7620
	(0.5963)	(0.0315)	(0.2266)



(a) Smoothed asset value



(b) Default probabilities

Figure 1: The smoothed firm asset values and default probabilities of SINOPEC.

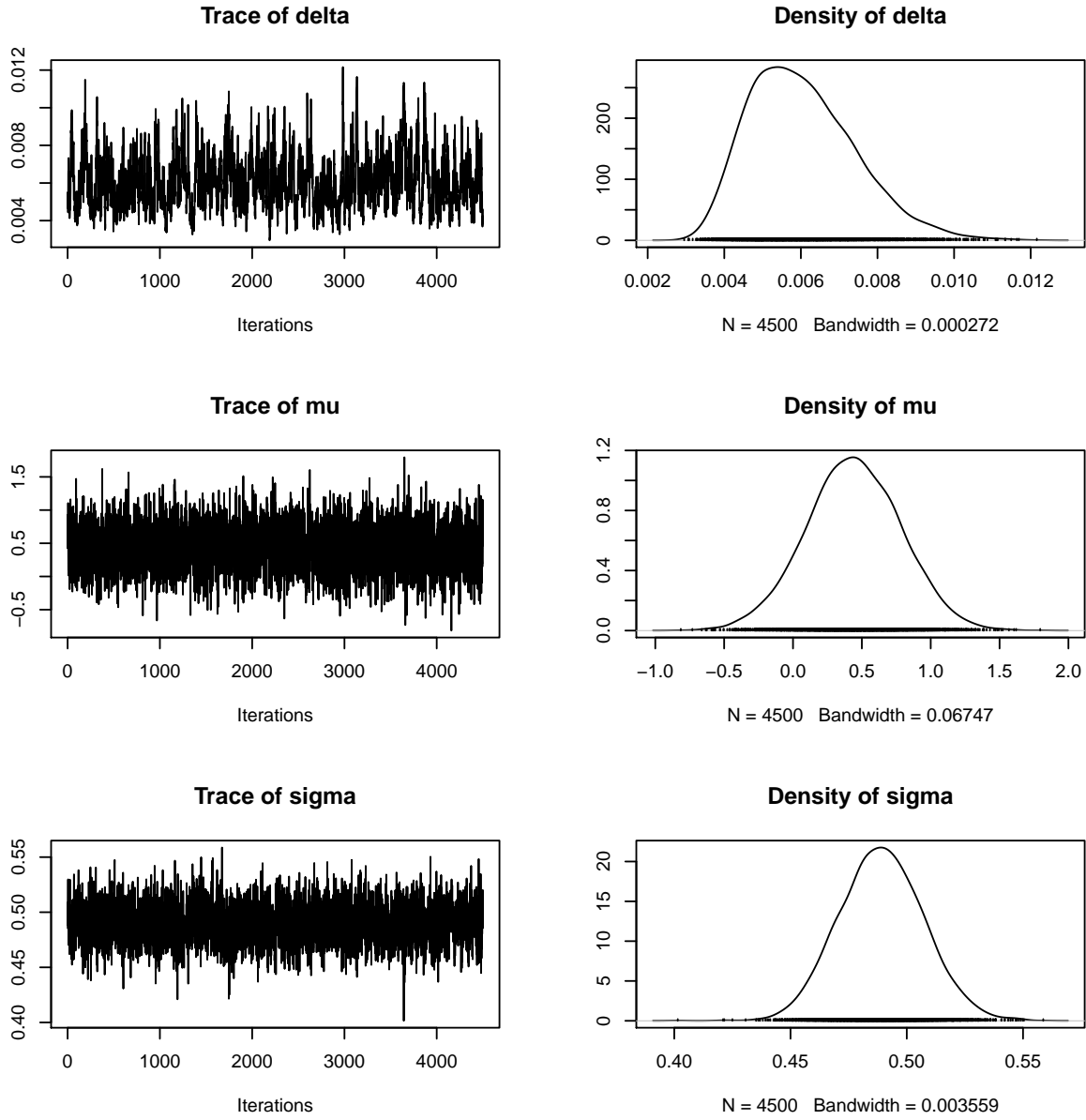


Figure 2: Trace and kernel density estimates of the marginal posterior distribution of parameters in model 1 for SINOPEC.

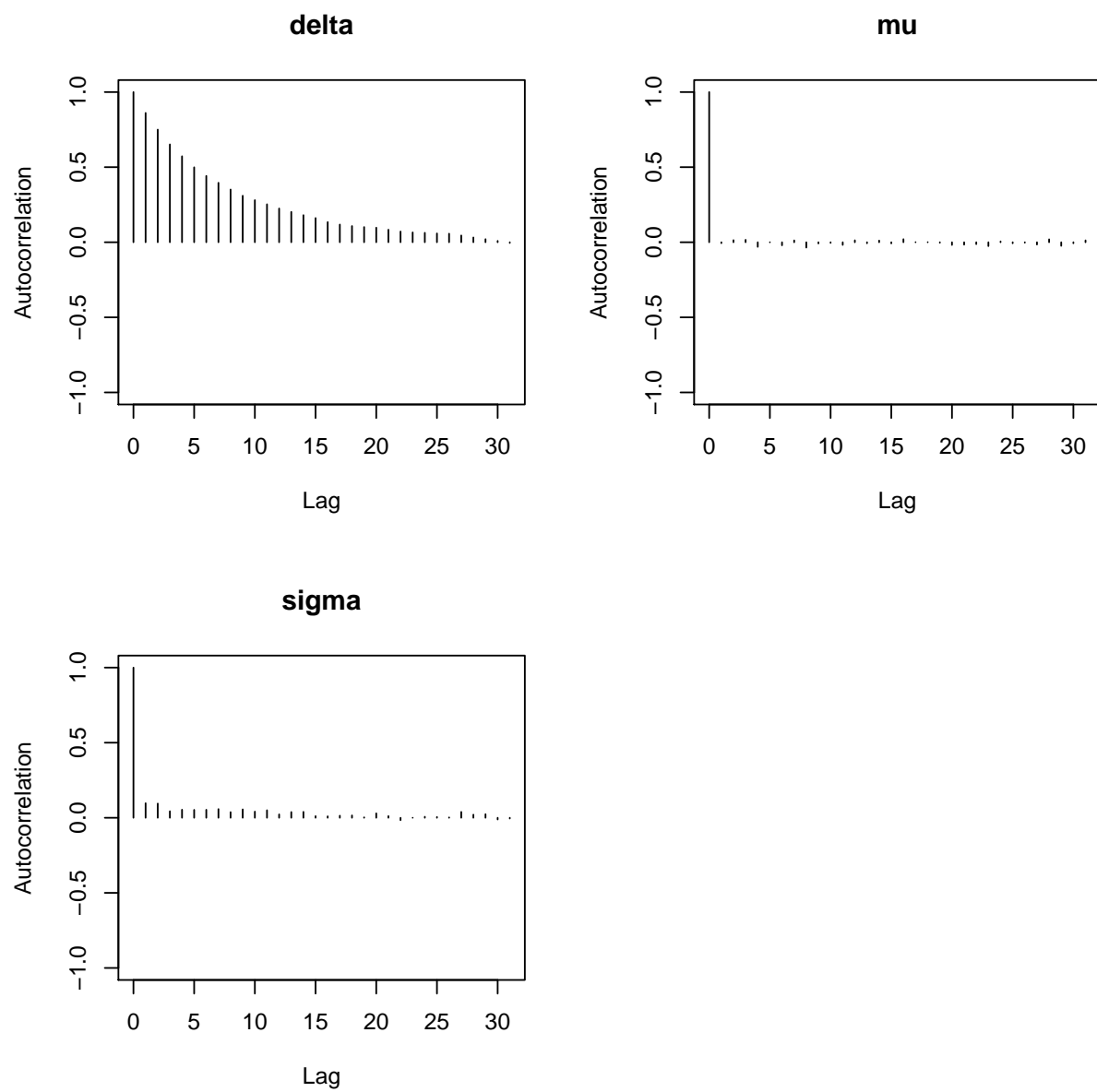


Figure 3: Autocorrelation functions of parameters in model 1 for SINOPEC.

We present the smoothed firm asset values and default probabilities of SINOPEC in Figure 1.⁴ We can see the default probabilities have different values according to the variation of asset values. But basically the smoothed estimates for the default probabilities are very small, which is quite reasonable for SINOPEC.

Figure 2 presents the trace and density estimates of marginal posterior distribution of model parameters δ , μ , σ for the 6 firms reported above. All chains mix well and the marginal posterior distribution is very symmetric for μ and σ but with some skewness for δ .

Figure 3 shows the autocorrelation function for each chain. For each parameter, the autocorrelation goes to zero after some lags, suggesting the convergence is fast, especially for μ and σ .

Refereing to convergence test, for the 6 firms all parameters pass the Heidelberger-Welch stationary and half-width test and their Geweke's Z-scores are reported in Table 2. In the table, we can see that the Z-scores are all between -2 and 2 , which suggests the chains' convergence. However, for the 10 firms left, one or more chains do not mix well, in most cases, σ has a convergence issue. This may be caused by the oversimplification of Merton's model. A little more complicated but realistic model may be considered to solve this problem, for example, assume sigma is stochastic rather than constant.

Table 2: Geweke's Z-score for the estimates in model 1.

Company name (SSE code)	Z-scores		
	μ	σ	δ
WISCO (600005)	0.7427	-0.8269	-0.2498
SINOPEC (600028)	0.1479	-0.8332	1.3878
CITICS (600030)	-0.9536	-0.6109	-1.1123
SAIC (600104)	0.9298	-1.6417	0.2397
Kweichow Moutai (600519)	0.8457	0.1326	0.1083
Liaoning Cheng Da (600739)	0.2836	0.9554	0.9343

When we try to estimate parameters in model 2 (equation 6) using data of the 10 firms that do not perform well in model 1, the convergence situation is not improved. When we only use one year data of 2007 for these 10 firms, we don't get much improvement as well.

⁴For other 5 firms, they have similar graphs of asset values and default probabilities.

5 Conclusions

We examine the structural credit risk models with microstructure noise using Chinese capital market data. In order to estimate the parameters in the model, which can actually be regarded as a state space model, we employ Bayesian MCMC method via Gibbs sampling. This method is easy to implement, and provides smoothed estimates of latent variables, which in our case are the firm's value. More importantly, we can get the exact posterior distribution by Bayesian MCMC method.

We perform empirical analysis using a sample of the data of 16 firms among the constituents of Shanghai Stock Exchange 50 index. We get the posterior mean and standard errors for the estimates using data of 6 firms in our sample. Chains mix very well in the cases of all the 6 firms. The reason that the performance of Bayesian MCMC in the cases using other firms' data may be the mis-specification of Merton's model, which can be solved by considering more complicated models.

References

- BLACK, F., AND J. COX (1976): "Valuing corporate securities: Some effects of bond indenture provisions," *The Journal of Finance*, 31(2), 351–367.
- BLACK, F., AND M. SCHOLES (1973): "The pricing of options and corporate liabilities," *The journal of political economy*, 81(3), 637–654.
- BROOKS, S., AND A. GELMAN (1998): "General methods for monitoring convergence of iterative simulations," *Journal of Computational and Graphical Statistics*, 7(4), 434–455.
- CASELLA, G., AND E. GEORGE (1992): "Explaining the Gibbs sampler," *The American Statistician*, 46(3), 167–174.
- CHIB, S. (2001): "Markov chain Monte Carlo methods: computation and inference," *Handbook of Econometrics*, 5, 3569–3649.
- CHIB, S., AND E. GREENBERG (1995): "Understanding the metropolis-hastings algorithm," *The American Statistician*, 49(4), 327–335.
- DUAN, J. (1994): "Maximum likelihood estimation using price data of the derivative contract," *Mathematical Finance*, 4(2), 155–167.

- DUAN, J., AND A. FULOP (2009): “Estimating the structural credit risk model when equity prices are contaminated by trading noises,” *Journal of Econometrics*, 150(2), 288–296.
- DUFFIE, D., AND K. SINGLETON (1999): “Modeling term structures of defaultable bonds,” *Review of Financial Studies*, 12(4), 687.
- GELFAND, A., AND A. SMITH (1990): “Sampling-Based Approaches to Calculating Marginal Densities,” *Journal of the American Statistical Association*, 85(410), 398–409.
- GELMAN, A., AND D. RUBIN (1992): “Inference from iterative simulation using multiple sequences,” *Statistical science*, 7(4), 457–472.
- GEMAN, S., AND D. GEMAN (1984): “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 6, 721–741.
- GEWEKE, J. (1992): “Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (Disc: P189-193),” *Bayesian Statistics*, 4, 169–188.
- HARRIS, L. (1990): “Estimation of stock price variances and serial covariances from discrete observations,” *Journal of Financial and Quantitative Analysis*, 25(03), 291–306.
- HASBROUCK, J. (1993): “Assessing the quality of a security market: A new approach to transaction-cost measurement,” *Review of Financial Studies*, 6(1), 191.
- HASTINGS, W. (1970): “Monte Carlo sampling methods using Markov chains and their applications,” *Biometrika*, 57(1), 97.
- HEIDELBERGER, P., AND P. WELCH (1981): “A spectral method for confidence interval generation and run length control in simulations,” *Communications of the ACM*, 24(4), 233–245.
- (1983): “Simulation run length control in the presence of an initial transient,” *Operations Research*, 31(6), 1109–1144.
- HUANG, S., AND J. YU (2010): “Bayesian analysis of structural credit risk models with microstructure noises,” *Journal of Economic Dynamics and Control*.
- JARROW, R., D. LANDO, AND S. TURNBULL (1997): “A Markov model for the term structure of credit risk spreads,” *Review of financial studies*, 10(2), 481.

- JARROW, R., AND S. TURNBULL (1995): “Pricing derivatives on financial securities subject to credit risk,” *The Journal of Finance*, 50(1), 53–85.
- LANDO, D. (1998): “On Cox processes and credit risky securities,” *Review of Derivatives research*, 2(2), 99–120.
- MADHAVAN, A., M. RICHARDSON, AND M. ROOMANS (1997): “Why do security prices change? A transaction-level analysis of NYSE stocks,” *Review of Financial Studies*, 10(4), 1035.
- MERTON, R. (1974): “On the pricing of corporate debt: The risk structure of interest rates,” *Journal of finance*, 29(2), 449–470.
- METROPOLIS, N., A. ROSENBLUTH, M. ROSENBLUTH, A. TELLER, E. TELLER, ET AL. (1953): “Equation of state calculations by fast computing machines,” *The journal of chemical physics*, 21(6), 1087.
- RAFTERY, A., AND S. LEWIS (1992): “How many iterations in the Gibbs sampler,” *Bayesian statistics*, 4(2), 763–773.
- RAFTERY, A., AND S. LEWIS (1995): “The number of iterations, convergence diagnostics and generic Metropolis algorithms,” *forthcoming in Practical Markov Chain Monte Carlo*, WR Gilks, DJ Spiegelhalter and S. Richardson, eds.(Chapman and Hall: London).
- TIERNEY, L. (1994): “Markov chains for exploring posterior distributions,” *the Annals of Statistics*, 22(4), 1701–1728.