An Analysis of Simultaneous Company Defaults Using a Shot Noise Process

M. Egami[a], R. Kevkhishvili[b]

[a]Graduate School of Economics, Kyoto University, Sakyo-Ku, Kyoto 606-8501, Japan
[b]Graduate School of Economics, Kyoto University, Sakyo-Ku, Kyoto 606-8501, Japan

Abstract

During the subprime mortgage crisis, it became apparent that practical models, such as the one-factor Gaussian copula, had underestimated company default correlations. Complex models that attempt to incorporate default dependency are difficult to implement in practice. In this study, we develop a model for a company asset process, based on which we calculate simultaneous default probabilities using an option-theoretic approach. In our model, a shot noise process serves as the key element for controlling correlations among companies' assets. The risk factor driving the shot noise process is common to all companies in an industry but the shot noise parameters are assumed company-specific; therefore, every company responds differently to this common risk factor. Our model gives earlier warning of financial distress and predicts higher simultaneous default probabilities than commonly used geometric Brownian motion asset model. It is also computationally simple and can be extended to analyze any finite number of companies.

JEL classification: G01; G17; G21; G32

Keywords: Credit risk; Shot noise; Option-theoretic approach; Asset process; Simultaneous default probabilities; Risk management

1. Introduction

This study aims to calculate simultaneous default probabilities of multiple companies. Our research is motivated by the fact that the incumbent models did not predict the default corre-

☆This version: April 7, 2017.
Funding: This work was in part supported by Japan Society for the Promotion of Science [Grant-in-Aid for Scientific Research (B) No. 26285069].

Email addresses: egami@econ.kyoto-u.ac.jp (M. Egami), keheisshuiri.rusudan.73m@st.kyoto-u.ac.jp (R. Kevkhishvili)
URL: http://www.econ.kyoto-u.ac.jp/~~11~egami/ (M. Egami)

lations in the global financial crisis; simultaneous default probabilities were underestimated in the structuring and pricing of collateralized debt obligations (CDOs). We wish to compute joint default probabilities more accurately to enhance risk management quality for the portfolios of debt instruments.

In general, there are top-down and bottom-up approaches to default correlation analysis. Our model belongs to the latter one. Using structural approach proposed by Merton [22], we analyze company defaults in a particular industry through the behavior of unobservable asset process. For this, we define default as an event in which a company's asset value falls below a certain level. It is highly likely that company defaults in one industry are correlated. To incorporate this correlation into the calculation of simultaneous default probabilities, we include a common shot noise process in each company's asset model. In our model, each company's asset value is driven by a company-specific risk factor and by the shot noise process, the latter being common to all companies that belong to the industry and having negative effects on the asset value. The shot noise process can be seen as an aggregation of jumps up to each point in time. The effect of jumps does not disappear immediately but decreases gradually over time, and hence inclusion of the shot noise process may help to make default correlation among the companies more realistic. For computational simplicity, in this study we deal only with negative effect of jumps, and the shot noise process allows us to keep the negative effect of external shocks for a certain period of time. We assume that the parameters of this shot noise process are company-specific. This means that the sensitivity of each company to the jumps of the shot noise process is different.

The main contributions of our paper are the following. Even though we introduce an unobservable common jump process in the asset model, by using a justifiable approximation of this process, we are able to derive an equation linking observable equity and debt values to the unobservable asset value; thus, we are able to estimate the asset process that incorporates common shocks to the industry. Furthermore, our estimation procedure requires neither the assumption of numbers or distribution of jumps nor the observation of shock arrivals (i.e, specific arrival times of shocks). In addition, when simulating the asset process using the estimated parameters, there is no need to simulate the jump times or jumps directly. This makes the simulation simple, which in turn makes the calculation of simultaneous default probabilities computationally easy. The existing model that enables us to estimate the unobservable asset process in a similar way is the geometric Brownian motion model. Our model can be considered as an improvement over this model, since it incorporates common jump process and is able to capture the correlation among the asset processes of multiple companies in a particular industry.

Since we use a structural approach and include a shot noise process in the asset model, below we will provide an overview of the literature related to the structural approach and common jumps used in default modeling. Kunisch and Uhrig-Homburg [17] adopt the top-down approach which they base on the structural model of a firm. They use random thinning to decompose an economy's default intensity, which is driven by macroeconomic factors, into the intensities of defaultable company subsets. To this end, they define default of a company as an event in which company assets fall at the outstanding debt level. This study employs a structural framework similar to Merton [22], assuming that the asset process follows geometric Brownian motion and assets of different companies are correlated. Under these assumptions, they derive solutions to the default probabilities of the company subsets, and finally, they calculate thinning probabilities using these default probabilities. Asset model proposed by Ma and Xu [21] includes company-specific and also common self-exciting Hawkes process and is intended to model unexpected defaults and default clustering better than the geometric Brownian motion (GBM) model when used in a structural approach. Theoretically, their model is capable of reproducing jump clustering. Ma and Xu [21] derive closed-form formulas for the default correlation; however, this study does not show how to estimate asset model parameters using company data. Aït-Sahalia et al. [1] provides a financial asset model that includes mutually exciting jump component and a continuous Brownian component and aims to incorporate amplification of jumps. This study establishes an estimation method for this model and investigates the contagion patterns and jump excitation among five world stock markets. Since their model tries to capture the effect of jumps that remains over time and has both continuous Brownian and jump component, it is related to our model; however, in contrast to our paper, Aït-Sahalia et al. [1] models observable stock index returns and does not address the issue of modeling unobservable company asset process.

Common shocks are often included in intensity-based models that belong to the bottom-up approach. Mortensen [24] models default of a firm as the first jump of a Cox process, the intensity of which consists of an idiosyncratic and common (to all firms) elements. In the analysis, Mortensen [24] specifies that jump sizes are exponentially distributed, and checks the fit of the model to the market prices of synthetic CDOs. Giesecke et al. [12] develop a dynamic reduced-form model in which the default intensity of a firm in a portfolio is driven by idiosyncratic and systematic risk factors, as well as the past defaults in the portfolio. A default is assumed to cause a jump in the intensity processes of all surviving firms. Such model allows for self-exciting effects. Giesecke et al. [12] use this model to analyze the behavior of the default rate as the number of firms in the portfolio increases. Spiliopoulos et al. [26] and Giesecke et al. [13] study an approximation to a large portfolio's loss distribution based on this model. Other examples of intensity-based studies are Dong et al. [6], Herbertsson et al. [14], and Liang and Wang [20], which use shocks to model default intensity processes and to derive explicit formulas for the joint probability of default. However, the feasibility of the intensity-based common shock model approach in the analysis of simultaneous defaults of more than two companies is not sufficiently explored. The computational difficulty remains an issue.

We also want to review copula models since these models are often used in bottom-up default analysis. An example of the copula approach that is similar to our study, in a sense that it is based on the asset process, is Giesecke [11]. He describes the model in which the threshold at which a firm goes bankrupt is not publicly known because information about the firm's liabilities is not fully disclosed. Taking the connection of different firms into account, bond investors estimate threshold levels for firms using available information. Then, while observing asset dynamics and default events, the investors update their estimates. This study uses copula to model the dependence structure among thresholds of different companies. Copula of default times is modeled using asset and threshold dependence structures. Another example is Dalla Valle et al. [4]. They employ pair copula to model dependence structure among current and long-term portions of the asset and debt of the company, and express company equity as a function of the asset and debt using pair copula. They simulate the values of equity and define default as an event in which equity is at or below zero, but they do not describe the extension of the model to simultaneous defaults of multiple companies and its feasibility. Finally, the study by Elouerkhaoui [10] is an example of a copula approach that includes common jumps. Elouerkhaoui [10] adopts the Marshall-Olkin method and models the default times of the obligors using Cox processes with common trigger jumps. In this way, this study introduces dependency among default times and then employs time-dependent copula to model this interdependence. Elouerkhaoui [10] considers only those shocks that induce defaults, i.e fatal-shock model and assumes that conditional on each trigger event, the default of the firms are independent in time.

We want to emphasize where our model stands in relation to the abovementioned models.

Our model tries to achieve a balance between computational simplicity and the ability to predict realistic joint default probabilities. Theoretically, it is possible to model complicated default dependency structures using copula approach. Even though there is a wide variety of choices for copula models, to our knowledge, there does not exist a good copula model for multivariate analysis that is not too complex. To introduce default dependency, we use the shot noise process that keeps the effect of shocks over a certain period of time. Since we use this process in order to represent bad news affecting a particular industry, and it is natural to consider that the effect of bad news accumulates and gradually decreases, we think the shot noise process is a reasonable/natural choice for representing accumulated effects. In contrast to copula models, by adopting structural approach, we manage to keep calculations simple.

Another important feature of our model is that we can estimate the parameters of unobservable company asset processes. We explicitly derive the equation that links company assets, debt, and equity values (see equation (2.8)). This equation enables us to estimate the parameters. This is a contrasting point to Ma and Xu [21], who model asset value by similar model but do not provide insight into the parameter estimation. Structural approach assumes that the company equity is a European call option written on the company asset process with a strike price equal to the amount of debt at maturity. In order to estimate the parameters of the asset model, we need to link the unobservable asset process to observable equity and debt values. And for this, we need to derive an option pricing formula. When using GBM model for the asset process, this can be easily done. However, GBM model contains only one risk source and as we demonstrate in this study, it turns out insufficient compared to our shot noise model. Merton [23] and Kou [16] provide option pricing formulas when underlying asset model contains jumps. However, they need to specify jump size distributions in order to obtain closed-form solutions. Merton [23] assumes log-normal distribution, while Kou [16] does double exponential distribution. Moreover, the estimation procedure, provided by Duan [7] and Duan [8] that we will discuss later, requires the likelihood function for the asset process. As demonstrated in Consigli [3], when explicitly using jumps in the model, the likelihood function requires setting the maximum number of jumps in a time interval of interest. In our study, we bypass the necessity of specifying the jump size distribution, jump times or maximum number of jumps by employing an approximation of the shot noise process.

As for the analysis, our focus is on the subprime mortgage crisis. Therefore, we analyze daily data from 2005/12/30 to 2014/12/31. In our example, we present results for three companies. However, this model works for any number of companies and adding a new company to the model does not increase computational difficulty, since the shot noise process is assumed common for all companies that belong to that particular industry. We estimate simultaneous default probability matrices by our model (we call this shot noise model) and by the GBM model, that we will discuss later (Sections 2.5 and 4) and that is commonly used in a structural approach. The most important result is that our model reacts earlier to financial distress and predicts higher simultaneous default probabilities for 2008-2010. In addition, we test our model in the following two ways. First, we simulate asset values from our model and the GBM model and obtain implied equity values. Then, we compare these equity values to real equity data. The goodness of fit of our model turns out to be better or almost the same (in a few cases) as that of the GBM model. Second, we use simulated asset values and explore the relationship between assets and CDS spreads by using a linear regression analysis. We observe that all the coefficients are statistically significant at 1% level.

Finally, we would like to emphasize that we propose our model as a tool for risk management. Based on one year (or half a year) of data, one could estimate the parameters of our model and calculate joint default probabilities for the coming year. If those values are significantly different from the ones of the previous year, this could be considered as an alarm for investors or related parties. The behavior of the investors armed with this information obviously can affect the outcome of the coming year. Therefore, predicting a high simultaneous default probability for the next year does not mean that the default has to occur in order to justify the model. It only means that the model predicts high risk and the action of investors and the related parties has to change in a way that the default becomes avoidable.

The rest of the paper is organized as follows. In Section 2, we carefully construct the asset-value process based on the shot noise process, and derive the maximum likelihood function. After discussing the data used in Section 3, we present and compare the results from the GBM and shot noise model (see Sections 4.1, 4.2, and 4.3). All the mathematical proofs are presented in Appendix A and some statistical and graphical results are in Appendix B.

2. Methodology

2.1. Model

Throughout this study, we deal with probability space $(\Omega, \mathscr{F}_T, \mathbb{P})$. T can be viewed as any fixed point in time. Below we will specify the filtration $\mathscr{F}_T$. For an asset-value process, we aim to use a model that incorporates a shot noise process. To make the model simple, we use only one shot noise process for bad news, which means that the shot noise process will have negative effects on the asset value. We propose the following model for the company i's asset value and give the details behind this formulation below.

$$V_t^{(i)} = \exp\left(X_0^{(i)} + \left(\mu^{(i)} - \frac{1}{2}\left(\sigma^{(i)}\right)^2\right)t + \sigma^{(i)}B_t^{(i)} - \frac{\mu_1^{(i)}\rho}{\delta^{(i)}} - Z_t^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\right) \quad t \geq 0 \tag{2.1}$$

Here, $V_t^{(i)}$ denotes the asset value of company i at time t, and the superscript i denotes the company-specific parameters. $\mu^{(i)}$, $\sigma^{(i)} > 0$, $X_0^{(i)}$, $\mu_1^{(i)} > 0$, $\mu_2^{(i)} > 0$, $\delta^{(i)} > 0$, and $\rho > 0$ are constant parameters. ρ is a common parameter to all companies in a particular industry; therefore, it does not have the superscript i. $B_t^{(i)}$ is standard Brownian motion representing company-specific risk. $Z_t^{(i)}$ is an Ornstein-Uhlenbeck process and satisfies the differential equation

$$\mathrm{d}Z_t^{(i)} = -\delta^{(i)}Z_t^{(i)}\mathrm{d}t + \sqrt{2\delta^{(i)}}\mathrm{d}W_t,$$

where W_t is a standard Brownian motion. W_t is a risk factor common to all companies in the industry of interest. See (2.6) and the subsequent explanation of W. By the dynamics of $\mathrm{d}Z^{(i)}$, $Z_t^{(i)}$ can equivalently be written as

$$Z_t^{(i)} = Z_0^{(i)}e^{-\delta^{(i)}t} + \sqrt{2\delta^{(i)}}\int_0^t e^{-\delta^{(i)}(t-s)}\mathrm{d}W_s. \tag{2.2}$$

Before discussing further details of our model, we will explain the background behind this formulation.

Setting $\tilde{\lambda}_t^{(i)} = \frac{\mu_1^{(i)}\rho}{\delta^{(i)}} + Z_t^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}$, we observe that (2.1) can be written as

$$V_t^{(i)} = e^{X_0^{(i)} + \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)t + \sigma B_t^{(i)} - \tilde{\lambda}_t^{(i)}}, \tag{2.3}$$

where $\tilde{\lambda}_t^{(i)}$ serves as a decreasing factor of the asset process. It represents an approximation of a shot noise process $\lambda_t^{(i)}$ at time t based on Dassios and Jang [5]. We give a brief overview of this shot noise process. The shot noise process at time t, denoted by $\lambda_t^{(i)}$, is given by the following equation

$$\lambda_t^{(i)} = \lambda_0^{(i)}e^{-\delta^{(i)}t} + \sum_{j=1}^{M_t} Y_j^{(i)}e^{-\delta^{(i)}(t-S_j)}, \tag{2.4}$$

where $\lambda_0^{(i)}$ is the initial value of the shot noise process, $\delta^{(i)}$ is the exponential decay rate, $\{S_j\}_{j=1,2,\ldots}$ are event times of Poisson process M_t with constant rate ρ, and $\{Y_j^{(i)}\}_{j=1,2,\ldots}$ is a sequence of independent and identically distributed random variables with distribution function

$G^{(i)}(y), y > 0$. The $Y^{(i)}$'s represent the size of the shot noise jumps and are independent of M_t. In addition, we require that the first and second moments of the jumps be finite, following Dassios and Jang [5]:

$$\mathbb{E}\left(Y_j^{(i)}\right) = \mu_1^{(i)} < \infty \quad \text{and} \quad \mathbb{E}\left(\left(Y_j^{(i)}\right)^2\right) = \mu_2^{(i)} < \infty.$$

We use the superscipt i for the shot noise process because it includes company-specific parameters. The jump times of the shot noise process are common to all companies; however, we assume that companies respond differently to these jumps. For this reason, the Y's are used with the superscipt i, since the impacts of the jumps differ across companies. For the same reason, the first and the second moments of the jumps are company-specific, as well. Recall that ρ and the jump times are common to all companies. We want to point out that in this study $\tilde{\lambda}_t^{(i)}$ (or $\lambda_t^{(i)}$) is not used as an intensity of a stochastic process. It is simply used as a process that expresses accumulated effect of jumps. Since all jumps are assumed positive, we use $-\tilde{\lambda}_t^{(i)}$ to incorporate bad news in the asset model.

The shot noise process λ is widely used in the literature. As mentioned above, our main purpose of using this process is to represent jumps (= bad news), whose effects do not disappear immediately and remain for a certain period of time (see e.g., Fig. 1 below).

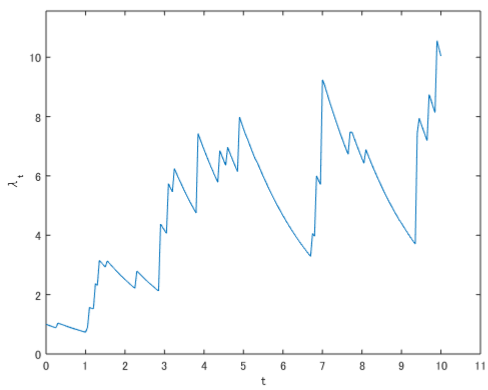


Fig. 1. An example of a simulated sample path of a shot noise process up to time $T = 10$. $\lambda_0 = 1$, $\delta = 0.5$, $\rho = 3$. Jumps follow exponential distribution with parameter 1.

2.2. *On Approximation of the Shot Noise Process*

As discussed in the introduction, the derivation of the equation that links an unobservable asset process to an observable equity and debt values is necessary for parameter estimation. This is the reason why we use the approximation of the shot noise process and not the process itself directly. This method allows us to estimate asset model parameters using observable debt and equity data. Below we will illustrate the idea behind this approximation and give its justification for our case.

$\tilde{\lambda}$ is the approximation of the shot noise process λ proposed in Theorem 2 in Dassios and Jang [5]. We will discuss a general case presented in Dassios and Jang [5], omitting the superscript i. Dassios and Jang [5] assume that the event arrival rate ρ tends to infinity and that λ_0 is a random variable, independent of everything else, satisfying $\frac{\lambda_0 - \frac{\mu_1\rho}{\delta}}{\sqrt{\frac{\mu_2\rho}{2\delta}}} \xrightarrow{d} Z_0$ as $\rho \to \infty$. Then, defining $Z_t^{(\rho)} := \frac{\lambda_t - \frac{\mu_1\rho}{\delta}}{\sqrt{\frac{\mu_2\rho}{2\delta}}}$, they prove that $Z_t^{(\rho)} \xrightarrow{d} Z_t$ as $\rho \to \infty$, where $\mathrm{d}Z_t = -\delta Z_t \mathrm{d}t + \sqrt{2\delta}\mathrm{d}B_t$ and B_t is a standard Brownian motion. For this approximation to hold, we need to verify that the event arrival rate ρ is large enough, that is, jumps are frequent. This means that events in this model are not catastrophes, but rather "common events of high frequency" (Dassios and Jang [5, p. 97]). We propose that the frequent jumps represent bad news about companies in a given industry, including deteriorating profit numbers, changing business environments, and sudden changes in management team as well as the jumps associated with systematic risk. If we use m companies of one industry in the analysis, these jumps would be related to 1) m idiosyncratic company risks, 2) risks associated with the remaining companies in the industry, and 3) systematic noise. We will be using daily data for the analysis, setting $\Delta_t = \frac{1}{360}$. The number of jumps in one year would be equal to ρ.

Since we have set $\tilde{\lambda}_t = \frac{\mu_1\rho}{\delta} + Z_t\sqrt{\frac{\mu_2\rho}{2\delta}}$ (by omitting superscript i), the approximation means

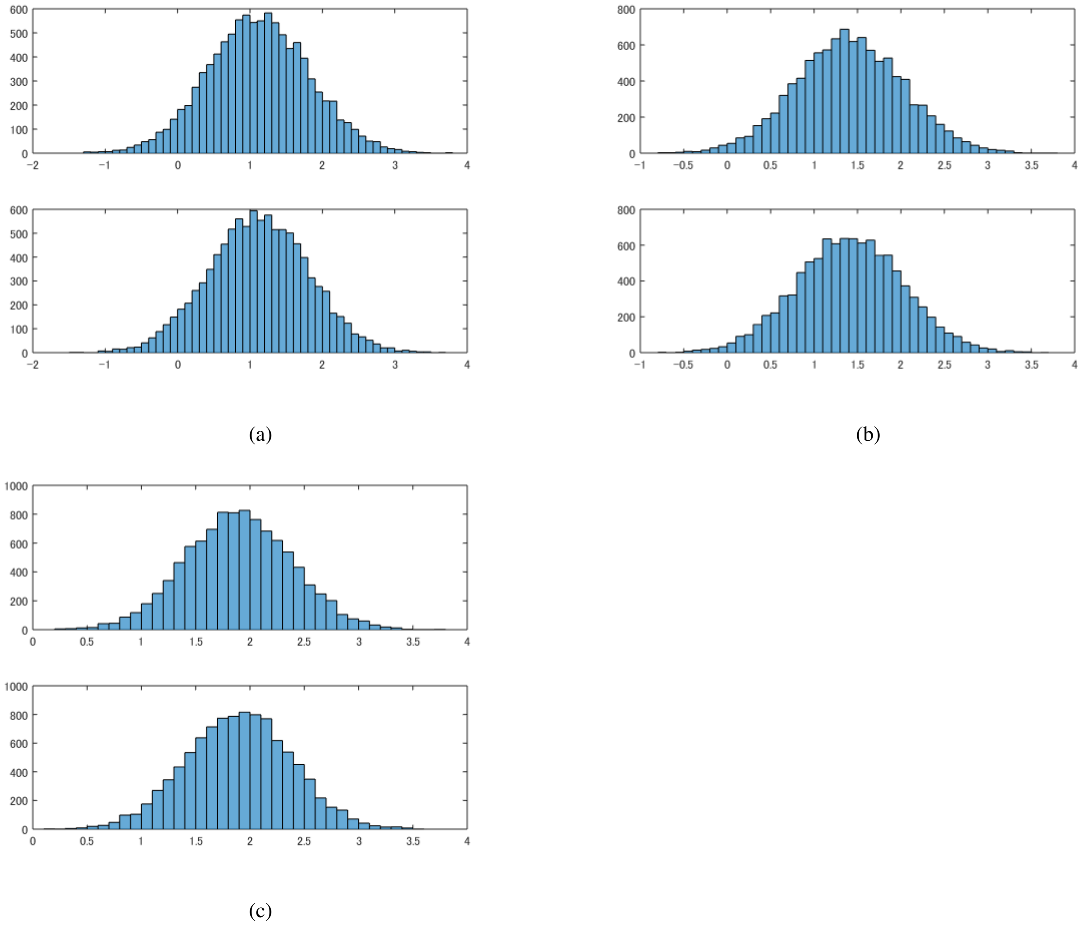
$$\lambda_t = \frac{\mu_1\rho}{\delta} + \sqrt{\frac{\mu_2\rho}{2\delta}} Z_t^{(\rho)} \stackrel{d}{=} \frac{\mu_1\rho}{\delta} + \sqrt{\frac{\mu_2\rho}{2\delta}} Z_t = \tilde{\lambda}_t \tag{2.5}$$

for each t when ρ is large enough. Note that λ_t is originally given by (2.4) (omitting the superscript i). Hence, checking the approximation of $Z_t^{(\rho)}$ by Z_t is equivalent to comparing λ_t and $\tilde{\lambda}_t$. We shall now make sure that this convergence in distribution holds with $\rho = 360$, one bad news per day in average in the industry. More specifically, we generate sample paths of λ_t based on (2.4) and generate $\tilde{\lambda}_t$ based on (2.2) with the right-hand side of (2.5). Again, we are discussing a general case and do not use i to indicate company-specific parameters. The precision of the approximation when $\rho = 360$ is demonstrated in Fig. 2, where we compared the distribution of λ_t and $\tilde{\lambda}_t$ at certain points of time t. We can see that the approximation of the distribution is quite accurate for such ρ. Since we are dealing with processes, we also display the trajectories of the shot noise and of the approximated process in Fig. 3. We simulated paths of the two processes λ_t and $\tilde{\lambda}_t$ including their initial values (a) 200, (b) 500, and (c) 1000 times and plotted the averaged paths. We show here the results that would most likely occur. During simulation, we only assumed that Z_0 and $\frac{\lambda_0-\frac{\mu_1\rho}{\delta}}{\sqrt{\frac{\mu_2\rho}{2\delta}}}$ have the same distribution, the sole assumption required in Dassios and Jang [5]. Table 1 reports the sums of squared differences between the two averaged paths. We also computed these differences for the case of $\rho = 1080$, that is, 3 jumps per day in average. This table shows that for each ρ, the differences diminish significantly as we increase the number of simulations. Note that since λ_t in (2.4) depends on ρ and hence its paths generated for the test of $\rho = 360$ and that of $\rho = 1080$ are different, a direct comparison of fit between $\rho = 360$ and $\rho = 1080$ is not relevant in this experiment. As we can see, the trajectories of the shot noise process are approximated rather well when there are $1 \sim 3$ jumps per day in average. We have tried other jump and decay parameters and obtained equally good results of fitness. $1 \sim 3$ jumps per day is not an unrealistic assumption. Thus, we proceed with this approximation $\tilde{\lambda}_t$ and hence with Z_t in (2.2).

Table 1

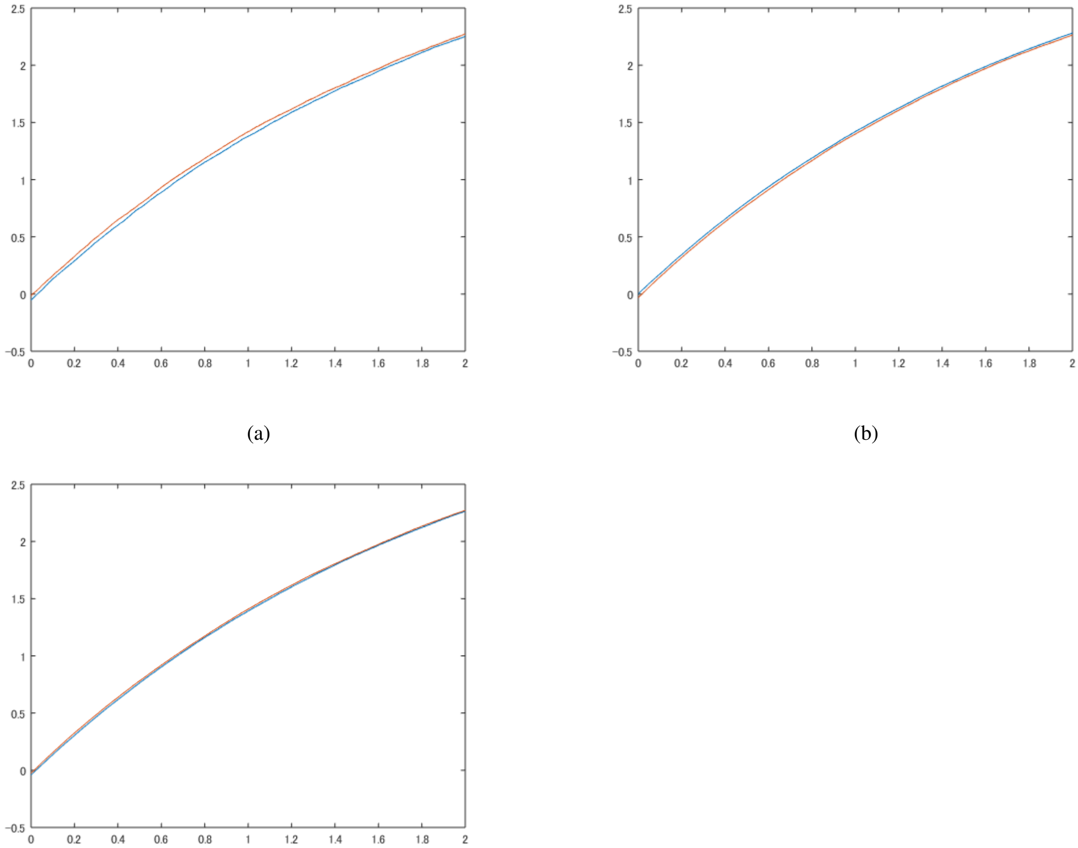
Sum of Squared Differences between the Shot Noise and the Approximated Process's paths.

	Simulation Number		
	200	500	1000
$\rho = 360$	0.8099	0.3377	0.1606
$\rho = 1080$	1.6368	0.6686	0.3030



(a) (b) (c)

Fig. 2. Histogram of a Shot Noise Process λ (Upper Panel) and Its Approximation $\tilde{\lambda}$ (Lower Panel). $T = 2$, $\delta = 0.5$, $\Delta_t = \frac{1}{360}$, $\rho = 360$. Jumps follow exponential distribution with parameter 200. λ_0 is assumed to be a standard normal random variable. The horizontal axis denotes the values of the process at $t = 0.75$ in (a), $t = 1$ in (b), and $t = 1.5$ in (c). The paths were simulated 10,000 times.



(a) (b)

(c)

Fig. 3. Average of the Simulated Trajectories of a Shot Noise Process λ (Blue Line) and Its Approximation $\tilde{\lambda}$ (Red Line). $T = 2$, $\delta = 0.5$, $\Delta_t = \frac{1}{360}$, $\rho = 360$. Jumps follow exponential distribution with parameter 200. λ_0 is assumed to be a standard normal random variable. Z_0 has the same distribution as $\frac{\lambda_0 - \frac{\mu_1 \rho}{\delta}}{\sqrt{\frac{\mu_2 \rho}{2\delta}}}$. The horizontal axis denotes a specific point in time. (a), (b), and (c) display the average of 200, 500, and 1000 simulated paths, respectively.

Now we return to our model (2.1) for any company i. As we mention, following Dassios and Jang [5], we have assumed that $\lambda_0^{(i)}$ is a random variable independent of everything else and satisfies $\frac{\lambda_0^{(i)} - \frac{\mu_1^{(i)}\rho}{\delta^{(i)}}}{\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}} \xrightarrow{d} Z_0^{(i)}$ when $\rho \to \infty$. Then, from Theorem 2 in Dassios and Jang [5],

$$\left(Z_t^{(i)}\right)^{(\rho)} \xrightarrow{d} Z_t^{(i)} \text{ as } \rho \to \infty.$$

Here, $\left(Z_t^{(i)}\right)^{(\rho)} = \frac{\lambda_t^{(i)} - \frac{\mu_1^{(i)}\rho}{\delta^{(i)}}}{\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}}$ and $Z_t^{(i)} = Z_0^{(i)} e^{-\delta^{(i)}t} + \sqrt{2\delta^{(i)}} \int_0^t e^{-\delta^{(i)}(t-s)} \mathrm{d}W_s$, which is (2.2). Note that W_t is a standard Brownian motion and since it represents shocks of the shot noise process, we assume that W_t is of the form

$$W_t = k^{(1)} B_t^{(1)} + k^{(2)} B_t^{(2)} + \cdots + k^{(m)} B_t^{(m)} + \tilde{k}\tilde{B}_t, \tag{2.6}$$

where $B^{(i)}$ is a standard Brownian motion that appears in (2.1), representing company i's idiosyncratic risk. After taking into consideration m companies' idiosyncratic risks, we are left with the systematic noise, which we assume is a standard Brownian motion, and $n_{\text{total}} - m$ number of companies' idiosyncratic risks (n_{total} denotes the total number of companies in the industry), i.e. $n_{\text{total}} - m$ number of standard Brownian motions. In our framework, company idiosyncratic risks are independent of each other and of the systematic noise. Then, the standard Brownian motion $\tilde{B}$ can be seen as a combination of the remaining independent Brownian motions (remaining idiosyncratic risks and the systematic noise) into one, since this is easily done mathematically. Hence, all the Brownian motions appearing in equation (2.6) are independent of one another. W has an instantaneous correlation coefficient $k^{(i)}$ with $B^{(i)}$. Note that we need the condition $\sum_{i=1}^{m} (k^{(i)})^2 < 1$, so that (W_t) is a standard Brownian motion. We shall explain this matter in Section 2.5, where we discuss parameter estimation. Finally, we assume that $m+1$-dimensional Brownian motion $\left((B_t^{(i)})_{1 \le i \le m}, \tilde{B}_t\right)$ is adapted to filtration $\mathscr{F}_t$.

Furthermore, we assume $Z_0^{(i)}$ is $\mathscr{F}_0$-measurable. There is no a priori information about the distribution of $Z_0^{(i)}$. We assume $Z_0^{(i)}$ is a bounded random variable. Given $Z_0^{(i)}$, $Z_t^{(i)}$ follows the normal distribution for t fixed. Hence, we can use $\tilde{\lambda}_t^{(i)} = \frac{\mu_1^{(i)}\rho}{\delta^{(i)}} + Z_t^{(i)} \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}$ as Gaussian approximation of $\lambda_t^{(i)}$. This is the process used in our model (2.1). Once again, we want to emphasize that the parameters (μ_1, μ_2, δ) and the random variable Z_0 are company-specific, since each company reacts differently to shocks (this means that the characteristics of jumps are different), while the driving force behind these jumps (W_t in our case) is assumed to be common to all companies in the industry.

2.3. Dynamics of $V_t^{(i)}$ under Approximation

Since $Z_t^{(i)}$ is a semimartingale, we can use Itô's formula to derive the dynamics of the asset process under the approximation (see Appendix A.1 for details):

$$
\begin{aligned}
\mathrm{d}V_t^{(i)} &= V_t^{(i)}\left(\mu^{(i)} - \frac{1}{2}\big(\sigma^{(i)}\big)^2\right)\mathrm{d}t + V_t^{(i)}\sigma^{(i)}\mathrm{d}B_t^{(i)} + \frac{1}{2}V_t^{(i)}\big(\sigma^{(i)}\big)^2\mathrm{d}t - V_t^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\mathrm{d}Z_t^{(i)} \\
&+ \frac{1}{2}V_t^{(i)}\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}\mathrm{d}\langle Z^{(i)}, Z^{(i)}\rangle_t - V_t^{(i)}\frac{1}{2}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\sigma^{(i)}\mathrm{d}\langle B^{(i)}, Z^{(i)}\rangle_t - V_t^{(i)}\frac{1}{2}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\sigma^{(i)}\mathrm{d}\langle B^{(i)}, Z^{(i)}\rangle_t \\
&= \Big(\mu^{(i)} + \frac{1}{2}\mu_2^{(i)}\rho + \delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_t^{(i)} - \sigma^{(i)}\sqrt{\mu_2^{(i)}\rho k^{(i)}}\Big)V_t^{(i)}\mathrm{d}t + \sigma^{(i)}V_t^{(i)}\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}V_t^{(i)}\mathrm{d}W_t \\
&= Q_t^{(i)}V_t^{(i)}\mathrm{d}t + \left(\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\sqrt{1-(k^{(i)})^2}\mathrm{d}W_t^{(-i)}\right)V_t^{(i)} \qquad (2.7)
\end{aligned}
$$

where $Q_t^{(i)} = \mu^{(i)} + \frac{1}{2}\mu_2^{(i)}\rho + \delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_t^{(i)} - \sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}k^{(i)}$ and $W_t^{(-i)} = \frac{\sum_{1\le j\le m, j\ne i} k^{(j)}\mathrm{d}B_t^{(j)} + \tilde{k}\mathrm{d}\tilde{B}_t}{\sqrt{1-(k^{(i)})^2}}$. Note that $B_t^{(i)}$ and $W_t^{(-i)}$ are independent standard Brownian motions. In the integral form, this is

$$
\begin{aligned}
&V_t^{(i)} = \\
&V_0^{(i)} + \int_0^t Q_s^{(i)}V_s^{(i)}\mathrm{d}s + \int_0^t \left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)V_s^{(i)}\mathrm{d}B_s^{(i)} - \int_0^t \sqrt{\mu_2^{(i)}\rho\big(1-(k^{(i)})^2\big)}V_s^{(i)}\mathrm{d}W_s^{(-i)}.
\end{aligned}
$$

The first integral is again a Lebesgue–Stieltjes integral and is a bounded variation process. The second and the third integrals are local martingales, since $\mathbb{P}\Big[\int_0^T \left(V_s^{(i)}\right)^2\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)^2 ds < \infty\Big] = 1$ and $\mathbb{P}\Big[\int_0^T \left(V_s^{(i)}\right)^2 \mu_2^{(i)}\rho\big(1-(k^{(i)})^2\big)ds < \infty\Big] = 1$, where T is our observation horizon. Their sum is also a local martingale. Therefore, $V_t^{(i)}$ is a semimartingale for $t \in [0, T]$.

2.4. Risk-neutral Theory

The asset process of the company is not observable in the market. We know the book value of the assets only from the balance sheet. However, once we estimate the unknown parameters in (2.1), (2.2), and (2.6), we are able to simulate the asset process. For this purpose, first, we wish to connect company assets to its equity since the equity value is observable. We use the option-theoretic approach to equity value, as in Black and Scholes [2], Merton [22], and Lehar [19]. That is, we regard equity $E_t^{(i)}$ as a call option written on the company's assets with a strike price equal to the future value of the company's debt $D_{T_m}^{(i)} = D_t^{(i)} e^{r(T_m - t)}$ where $D_t^{(i)}$ is the debt of the company i at time t and T_m is the maturity of the option. In our case, we need to calculate the value of this call option when the underlying asset follows (2.1). Lehar [19] assumes that all debt is insured (risk-free) and grows at a risk-free rate r, which we adopt for simplicity. In addition, we assume that the time to maturity equals 1 year, as in Lehar [19]. For each t, we consider a new option maturing after 1 year, i.e. $T_m - t = 1$ for all t.

Proposition 1. *Let the company i's asset-value process $V_t^{(i)}$ follow the equation* (2.1). *Then, the value of equity $E_t^{(i)}$ for this company is given by the equation*

$$E_t^{(i)} = V_t^{(i)} \Phi\left(d_t^{(i)}\right) - D_t^{(i)} \Phi\left(d_t^{(i)} - M^{(i)} \sqrt{(T_m - t)}\right), \tag{2.8}$$

where $d_t^{(i)} := \frac{\ln\left(\frac{V_t^{(i)}}{D_t^{(i)}}\right) + \frac{\left(M^{(i)}\right)^2}{2}(T_m - t)}{M^{(i)}\sqrt{(T_m - t)}}$, $M^{(i)} := \sqrt{\left(\sigma^{(i)}\right)^2 + \mu_2^{(i)}\rho - 2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}\, k^{(i)}}$, *and* $\Phi(\cdot)$ *is the standard normal distribution function.*

Proof. Consider the discounted asset-value process $e^{-rt}V_t^{(i)}$, where r is the constant risk-free rate. Its dynamics are

$$\begin{aligned}
&\mathrm{d}\big(e^{-rt}V_t^{(i)}\big) \\
&= -re^{-rt}V_t^{(i)}\mathrm{d}t + e^{-rt}\mathrm{d}V_t^{(i)} \\
&= e^{-rt}V_t^{(i)}\big(-r + Q_t^{(i)}\big)\mathrm{d}t + e^{-rt}V_t^{(i)}\Big(\big(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}\, k^{(i)}\big)\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho\big(1 - (k^{(i)})^2\big)}\mathrm{d}W_t^{(-i)}\Big) \\
&= e^{-rt}V_t^{(i)}\left[\left(-r + Q_t^{(i)}\right)\mathrm{d}t + \left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}\, k^{(i)}\right)\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho(1 - (k^{(i)})^2)}\mathrm{d}W_t^{(-i)}\right].
\end{aligned}$$

The following argument is based on Shreve [25, pp. 226–228]. Let us define $\alpha_i := \sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}\, k^{(i)}$ and $\alpha_{-i} := -\sqrt{\mu_2^{(i)}\rho(1 - (k^{(i)})^2)}$. Then, we obtain

$$\mathrm{d}(e^{-rt}V_t^{(i)}) = e^{-rt}V_t^{(i)}\left[\left(-r + Q_t^{(i)}\right)\mathrm{d}t + \alpha_i \mathrm{d}B_t^{(i)} + \alpha_{-i}\mathrm{d}W_t^{(-i)}\right]. \tag{2.9}$$

In order to transform the discounted asset value into martingale, we rewrite the equation in the following way:

$$\mathrm{d}(e^{-rt}V_t^{(i)}) = e^{-rt}V_t^{(i)}\left(\alpha_i\left[-\theta_t^{(i)}\mathrm{d}t + \mathrm{d}B_t^{(i)}\right] + \alpha_{-i}\left[-\theta_t^{(-i)}\mathrm{d}t + \mathrm{d}W_t^{(-i)}\right]\right)$$

For an adapted process $\theta = \left(\theta_t^{(i)}, \theta_t^{(-i)}\right)_{t \geq 0}$ to satisfy the above equation, it is necessary that

$$\alpha_i\theta_t^{(i)} + \alpha_{-i}\theta_t^{(-i)} = \left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}\, k^{(i)}\right)\theta_t^{(i)} - \sqrt{\mu_2^{(i)}\rho(1 - (k^{(i)})^2)}\,\theta_t^{(-i)} = r - Q_t^{(i)}.$$

This is one equation in two unknowns; therefore, it has infinitely many solutions. We choose one of them (no matter what we choose, the resulting pricing formula is the same). We set $\theta_t^{(-i)} := 1$ and $\theta_t^{(i)} := \frac{r - Q_t^{(i)} + \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}}{\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}}$ and define $\tilde{B}_t^{(i)} := B_t^{(i)} - \int_0^t \theta_s^{(i)} ds$, $\tilde{W}_t^{(-i)} := W_t^{(-i)} - \int_0^t \theta_s^{(-i)} ds$. $\theta_t^{(i)}$ and $\theta_t^{(-i)}$ are measurable adapted processes. If we can show that

$$H_t(\theta) := e^{\int_0^t \theta_s^{(i)} dB_s^{(i)} + \int_0^t \theta_s^{(-i)} dW_s^{(-i)} - \frac{1}{2}\int_0^t ||\theta_s||^2 ds}$$

is a martingale, then by Theorem 5.1 (Karatzas and Shreve [15, p. 191]) $\left(\tilde{B}^{(i)}, \tilde{W}^{(-i)}\right)_t$ would be a two-dimensional standard Brownian motion for $0 \le t \le T$ on $(\Omega, \mathscr{F}_T, \tilde{\mathbb{P}}_T)$, where probability measure $\tilde{\mathbb{P}}_T$ is defined as $\tilde{\mathbb{P}}_T(A) = \mathbb{E}[\mathbb{1}_A H_T(\theta)]$ for $A \in \mathscr{F}_T$. Indeed, $H_T(\theta)$ is a martingale (see Appendix A.2) and we obtain

$$\begin{aligned}
\mathrm{d}\big(e^{-rt}V_t^{(i)}\big) &= e^{-rt}V_t^{(i)}\left[\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)\mathrm{d}\tilde{B}_t^{(i)} - \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}\mathrm{d}\tilde{W}_t^{(-i)}\right] \\
&= e^{-rt}V_t^{(i)}M^{(i)}\mathrm{d}\tilde{\tilde{B}}_t^{(i)}
\end{aligned}$$

as a martingale under $\tilde{\mathbb{P}}_T$, where

$$M^{(i)} := \sqrt{\left(\sigma^{(i)}\right)^2 + \mu_2^{(i)}\rho - 2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}k^{(i)}}$$

and

$$\tilde{\tilde{B}}_t^{(i)} := \frac{\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}}{M^{(i)}}\tilde{B}_t^{(i)} - \frac{\sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}}{M^{(i)}}\tilde{W}_t^{(-i)}$$

is a one-dimensional standard Brownian motion under $\tilde{\mathbb{P}}_T$ (see Appendix A.2). From (2.7), we obtain

$$\begin{aligned}
\mathrm{d}V_t^{(i)} &= V_t^{(i)}Q_t^{(i)}\mathrm{d}t + V_t^{(i)}\left(\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)\mathrm{d}\tilde{B}_t^{(i)} + \left(r - Q_t^{(i)} + \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}\right)\mathrm{d}t\right) \\
&\quad - V_t^{(i)}\left(\sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}\mathrm{d}\tilde{W}_t^{(-i)} + \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}\mathrm{d}t\right) \\
&= rV_t^{(i)}\mathrm{d}t + V_t^{(i)}\left(\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)\mathrm{d}\tilde{B}_t^{(i)} - \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}\mathrm{d}\tilde{W}_t^{(-i)}\right) \\
&= rV_t^{(i)}\mathrm{d}t + M^{(i)}V_t^{(i)}\mathrm{d}\tilde{\tilde{B}}_t^{(i)},
\end{aligned}$$

so that

$$V_t^{(i)} = V_0^{(i)}e^{\left(r - \frac{1}{2}\left(M^{(i)}\right)^2\right)t + M^{(i)}\tilde{\tilde{B}}_t^{(i)}} \quad \text{with} \quad V_0^{(i)} = e^{X_0^{(i)} - \frac{\mu_1^{(i)}\rho}{\delta^{(i)}} - Z_0^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}}. \tag{2.10}$$

By the risk-neutral pricing formula,

$$E_t^{(i)} = \tilde{\mathbb{E}}\left[e^{-r(T_m - t)}\left(V_{T_m}^{(i)} - D_{T_m}^{(i)}\right)^+ \middle| \mathscr{F}_t\right].$$

Recall that T_m denotes the maturity (a certain point in time) of the debt, $E_t^{(i)}$ is the equity (i.e., the price of the option written on the value of the firm at time t) and $D_{T_m}^{(i)}$ works as the strike price. This is the same as the Black–Scholes formula with the volatility parameter $M^{(i)}$ in view of (2.10). Then, we obtain the desired equation (2.8). $\square$

As we can see, the analysis for company i does not include parameters of the remaining $m-1$ companies on interest. The only common parameter is ρ and it appears always in the form of $\mu_2^{(i)}\rho$. The product $\mu_2^{(i)}\rho$ is company-specific. This means, that we can estimate the parameters for each company separately. This is demonstrated in the next subsection.

2.5. Estimation Procedure

Thanks to Proposition 1, we have established the relationship between the unobservable asset-value process and the equity values in our shot noise model. Using that, we estimate the parameters of $V_t^{(i)}$ by the maximum likelihood estimation technique introduced in Duan [7], Duan [8], and Duan et al. [9]. For the case of GBM model, we just follow the method in these articles. But for our model, we have to modify it as we shall explain here. Given the data of the equity process $E^{(i)} = (E_t^{(i)}, t = \Delta_t, 2\Delta_t, \ldots, n\Delta_t)$, we can estimate the company-specific parameters $\mu^{(i)}, \delta^{(i)}, \mu_2^{(i)}\rho, \sigma^{(i)}$, and the realized value of the random variable $Z_0^{(i)}$ by maximizing the following log-likelihood function:

$$
\begin{aligned}
&L(E_t^{(i)}, t = \Delta_t, 2\Delta_t, \ldots, n\Delta_t; \mu^{(i)}, \delta^{(i)}, \mu_2^{(i)}\rho, Z_0^{(i)}, \sigma^{(i)} \mid \mathscr{F}_0) = \\
&-\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{n} \ln \mathrm{Var}_{j\Delta_t}^{(i)} - \sum_{j=1}^{n} \frac{\left(\ln\left(\frac{\hat{V}_{j\Delta_t}^{(i)}}{\hat{V}_{(j-1)\Delta_t}^{(i)}}\right) - \mathrm{mean}_{j\Delta_t}^{(i)}\right)^2}{2\mathrm{Var}_{j\Delta_t}^{(i)}} - \sum_{j=1}^{n} \ln \hat{V}_{j\Delta_t}^{(i)} - \sum_{j=1}^{n} \ln \Phi(\hat{d}_{j\Delta_t}^{(i)}),
\end{aligned}
\tag{2.11}
$$

where $\hat{V}_{j\Delta_t}^{(i)}$ is the unique solution to (2.8) in Proposition 1, $\hat{d}_{j\Delta_t}^{(i)}$ is $d_{j\Delta_t}^{(i)}$ but with $V_{j\Delta_t}^{(i)}$ replaced by $\hat{V}_{j\Delta_t}^{(i)}$,

$$
\mathrm{mean}_{j\Delta_t}^{(i)} = \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)\Delta_t - \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}} Z_0^{(i)} e^{-\delta^{(i)} j\Delta_t}(1 - e^{\delta^{(i)}\Delta_t})
$$

and

$$
\mathrm{Var}_{j\Delta_t}^{(i)} = (\sigma^{(i)})^2\Delta_t + \frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}(1 - e^{-2\delta^{(i)}\Delta_t}) - \frac{2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}}{\delta^{(i)}}(1 - e^{-\delta^{(i)}\Delta_t})k^{(i)}.
$$

See Appendix A.3 for a derivation. Following Duan [7], we have dropped the first observation $V_0^{(i)}$ from the likelihood and used $\mathscr{F}_0$ to define the conditional distribution of the observations that follow; therefore, the above likelihood does not include the density of the first observation. As we have assumed in Section 2.2, $Z_0^{(i)}$ is $\mathscr{F}_0$ measurable. Therefore, the distribution of $Z_0^{(i)}$ does not come into play anymore and only the realized value of $Z_0^{(i)}$ will be estimated. Moreoever, since $V_0^{(i)} = e^{X_0^{(i)} - \frac{\mu_1^{(i)}\rho}{\delta^{(i)}} - Z_0^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}}$,

$$V_t^{(i)} = V_0^{(i)} e^{\left(\mu^{(i)} - \frac{1}{2}\left(\sigma^{(i)}\right)^2\right)t + \sigma^{(i)} B_t^{(i)} - Z_t^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}} + Z_0^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}} \tag{2.12}$$

and dropping $V_0^{(i)}$ from estimation means that we will not estimate $X_0^{(i)}$ and $\mu_1^{(i)}$. During maximization, when necessary, $V_0^{(i)}$ will be calculated by (2.8) in Proposition 1, using the observable values of $E_0^{(i)}$ and $D_0^{(i)}$.

As we have mentioned, we estimate the company-specific parameters separately for each company. However, we want to point out that from equation (2.6), the condition $\sum_{i=1}^{m} (k^{(i)})^2 + \tilde{k}^2 = 1$ must be satisfied when analyzing m companies. We estimate $k^{(1)}, \cdots, k^{(m)}, \tilde{k}$ separately, and then use these values in the calculation of the likelihood function. We illustrate the estimation procedure for $k^{(1)}, \cdots, k^{(m)}, \tilde{k}$ in the next paragraph.

We linearly regress (including an intercept) the time series of the industry price index on the time series of the individual company share prices (the share prices of all m companies at the same time) and apply ANOVA to the estimated linear model. For each company i, we calculate the *reduction* in residual sum of squares by adding company i's share price data to the model *that already contains all the other* $m-1$ *company share price data*.[1] We denote this reduction by (Sum of Squares)$_i$. Also, we calculate Total Sum of Squares (*TSS*), which is the sum of squared industry price index values after subtracting out the mean, and which in turn corresponds to the sum of squared residuals of the model that includes only the intercept. Then, we set

$$k^{(i)} = \sqrt{\frac{(\text{Sum of Squares})_i}{\text{Total Sum of Squares}}}.$$

Adding explanatory variables to the simple model with only intercept reduces the sum of squared residuals. The sum of the abovementioned individual contributions to this reduction

[1] We confirmed that when the companies' share prices are adjusted for the company size by a multiple of real numbers, the values of $k^{(i)}$ are the same. Hence our method here can be used whether or not the industry price index is adjusted for the company size.

(i.e. $\sum_{i=1}^{m}$ (Sum of Squares)$_i$) can never exceed ESS (which is the total contribution of the model that includes m companies' share price data). Hence, we have $\sum_{i=1}^{m}$ (Sum of Squares)$_i \leq ESS \leq TSS$. [2] Therefore, using this method, we can ensure that $\sum_{i=1}^{m} \left(k^{(i)}\right)^2 < 1$.

When implementing the likelihood maximization in MATLAB, the function that we maximize to obtain the parameters consists of two parts. Inside the function, $\hat{V}_{j\Delta_t}^{(i)}$ is estimated by a fixed-point iteration procedure from (2.8) using available equity and debt values. Thereafter, using these estimated asset values, the log-likelihood in (2.11) is calculated. The result is a function in the unknown five parameters and this function is maximized by the built-in interior-point method in MATLAB (precisely, we use the MATLAB function "fmincon" to minimize the negative log-likelihood).

2.6. Summary

Our model for company asset value $V^{(i)}$ is written by equation (2.1) with W as in (2.6) and $Z^{(i)}$ as in (2.2). The latter is part of our approximation of the shot noise process $\lambda^{(i)}$ in equation (2.4).

[1] We first estimate $k^{(i)}$, $i = 1, \cdots, m$ and $\tilde{k}$.

[2] We need to estimate parameters $\mu^{(i)}, \delta^{(i)}, \mu_2^{(i)} \rho, \sigma^{(i)}$ together with $Z_0^{(i)}$ by the log-likelihood function (2.11). This function is derived thanks to Proposition 1, in particular equation (2.8). These values are company-specific and hence we perform this estimation on a company by company basis.

[3] The required data for [2] are equity values $E^{(i)}$ and debt values $D^{(i)}$, which we discuss in details in Section 3.

[4] Once we obtain the parameters, we can simulate $V^{(i)}$ by (2.1), (2.2), and (2.6) and compute equity values implied by our model using (2.8). By comparing the implied equity values to the real equity data, we check whether the estimated parameters, including the

[2] The case that $\sum_{i=1}^{m}$ (Sum of Squares)$_i$ is equal to TSS would realize only if the following three events occur at the same time: (1) we have used all the companies in the industry (i.e. $n_{\text{total}} = m$), (2) there is no other source of variation, such as systemic risk (that is, $\tilde{k} = 0$ by definition), and (3) the sum of *marginal* contributions of each company is equal to the total reduction of variance (i.e., $\sum_{i=1}^{m}$ (Sum of Squares)$_i = ESS$). But this is highly unlikely to occur in practice.

realized value of $Z_0^{(i)}$, are accurate or not. With $V^{(i)}$, we can compute simultaneous default probabilities. For details of this part, see Section 4.

3. Data

The data used in the estimations are obtained from Thomson Reuters Datastream and Worldscope. The debt and equity data are displayed in millions of US dollars. Our focus is on the banking industry; however, the abovementioned model can be used in the analysis of any industry. The choice of companies can be random. We analyze three companies: JP-Morgan Chase & Co. (JPM), Citigroup Inc. (Citi), and Bank of America Corporation (BAC) (see Tables B.1, B.2, and B.3 for the company profiles in Appendix B). For our equity data, we use "Market Value by Company" (MVC) from Datastream. "Market Value" (MV) from Datastream provides the market value of only one class of shares. MVC is the same as MV for companies with only one listed equity security. However, for companies with more than one type of listed or unlisted shares, MVC takes these other types of shares into consideration, too. For a detailed description, refer to Datastream Datatype Definition. For the estimation, we use daily data from 2005/12/30 to 2014/12/31. Fig. B.1a displays the daily MVC for the three companies.

From the company balance sheets, we calculate $D_t^{(i)}$ for each t. We define $D_t^{(i)}$ as the sum of Deposits-Total, Commercial Paper, Debt and equity instruments-Trading liabilities, Federal funds purchased under repurchase agreement, Other borrowed funds, and one half of Long-term debt. This addition of one-half of the long-term debt is a conventional method adopted by Moody's KMV, a unit that offers commercial packages of default probability. For JPM, all these categories are available (Federal funds purchased under repurchase agreement are displayed under the name of "under repurchase agreements"). However, for the other two companies, the data are not available under these names and we make small adjustments.

For Citi, we define D_t as the sum of Deposits-Total, Federal funds purchased and securities (sum of Federal Funds Purchased, Security Sold under Repurchase Agreement, and Federal funds purchased and securities), Commercial paper, Trading account liabilities, Short-term borrowings (sum of Short-term borrowings and Other borrowings - Balancing value), and one half of Long-term debt. For BAC, D_t is the sum of Deposits-Total, Federal funds purchased and securities (sum of Federal Funds Purchased, Security Sold under Repurchase Agreement, and Federal funds purchased and securities), Trading account liabilities, Short term borrowings, and one half of Long term debt. Commercial paper data are not available on the balance sheet for this company.

Deposits-Total for all companies and Long term debt for BAC are retrieved from Worldscope. The remaining data are from Reuters. The values are restated, updated, or reclassified values, that is, the latest values available. These debt data are quarterly. We use interpolation to obtain the daily values of the debt. Fig. B.1b displays the time-series (2609 observations) of the debt for all companies. Finally, the three companies' individual share prices and S&P500 Economic Sector Financials Price Index (S&P500 ES Financials) data was used for estimating parameters $k^{(i)}$.

4. Estimation, Results, and Discussion

Since there is a possibility that parameters of the model change over time, we split the estimation period into 1-year subperiods. The data in each 1-year subperiod are used to estimate the model parameters for the subsequent year's default probabilities. The first cohort is

$$\begin{aligned}
&2005/12/30 \sim 2006/12/29 \Rightarrow \text{for the year } 2007,\\
&2006/12/29 \sim 2007/12/31 \Rightarrow \text{for the year } 2008,\\
&2007/12/31 \sim 2008/12/31 \Rightarrow \text{for the year } 2009,\\
&2008/12/31 \sim 2009/12/31 \Rightarrow \text{for the year } 2010,
\end{aligned}$$

and the second cohort is

$$\begin{aligned}
&2009/12/31 \sim 2010/12/31 \Rightarrow \text{for the year } 2011,\\
&2010/12/31 \sim 2011/12/30 \Rightarrow \text{for the year } 2012,\\
&2011/12/30 \sim 2012/12/31 \Rightarrow \text{for the year } 2013,\\
&2012/12/31 \sim 2013/12/31 \Rightarrow \text{for the year } 2014.
\end{aligned}$$

For each 1-year interval, we estimate the parameters of our model (including $k^{(i)}$) and of the GBM model used in Lehar [19]. The difference between these two models is that the first incorporates the shot noise process and the second does not. We report the parameter values of both models for each period. We want to point out that the drift term of the asset process in our model (eq.(A.3) in Appendix A) is much more complicated than the drift term of the GBM model. For example, it is not possible to directly compare the parameter $\mu^{(i)}$ of our model and the counterpart of the GBM. [3] Also, each parameter of the shot noise model may change from one period to another, so that the final drift in eq.(A.3) can capture the overall trend of equity prices.

Using the estimated parameters of the 1-year subinterval, we generate the asset-value process for the next year. For this, we first retrieve the starting value of the asset process from the Black–Scholes equation using the estimated variance parameter and then simulate the next 360 observations by starting from that value. For example, if we use data from 2005/12/30 $\sim$ 2006/12/29 to estimate parameters, we compute the starting asset value of the next year (Year 2007) using the Black–Scholes formula (2.8), the estimated variance, and the observable data (equity and debt) of 2006/12/29. In order to simulate the asset path $V_t^{(i)}$, we need to simulate $B_t^{(i)}$ and $Z_t^{(i)}$, the latter depending on W_t. After estimating the coefficients $k^{(i)}, i = 1, \cdots, m$ and $\tilde{k} = \sqrt{(1 - \sum_{i=1}^{m} (k^{(i)})^2}$, we are able to simulate W_t using $B_t^{(i)}, i = 1, \cdots, m$ and $\tilde{B}_t$. We assume that default occurs only at the end of the simulation period (default model proposed by Merton [22]); therefore, we compare the final simulated asset value (i.e., the asset value for the end of the year 2007) to the debt amount of the company. We assume that debt grows at the rate of 1-year treasury bill r; that is, we use the year-end debt level of 2006 multiplied by e^r, where r is the last observed treasury rate in 2006. Table B.6 in Appendix B summarizes the interest rates used in the calculation of the debt. If the asset value falls at or below the level of debt D_{T_m}, we consider this as default. We simulate sample paths of the asset-value process 100,000 times and count the number of defaults. Then, dividing the number of defaults by 100,000 gives us default probabilities. An example demonstrating the precision of this simulation is displayed in Fig. B.5 in Appendix B.

Since our model has 5 unknown parameters, the values that we obtain from the minimization function depend on the initial values of the parameters. We mention our rules and procedures for choosing one parameter set among a few candidates. Let us take an example. Suppose we use the equity data of 2005/12/30$\sim$2006/12/29 to estimate parameters. Using these estimated parameter vectors, we simulate the asset-value process for the same 1-year period (for 2005/12/30$\sim$2006/12/29). The simulation procedure is the same as described in the previous paragraph. For each simulated asset path, we back out the equity values from

[3] We have observed during estimation that $Z_0^{(i)}$ and $\mu^{(i)}$ are closely related when determining the drift and most of the time take opposite signs.

the Black–Scholes equation (2.8). We do the asset simulation 50,000 times, so that we obtain 50,000 *implied* equity paths. Finally, we take the average of these paths. In this way, we are able to obtain the *implied* equity values from the model for 2005/12/30$\sim$2006/12/29. Then, we compare these equity values to the actual equity value time series of the company for the same period. We also calculate the sum of the squared differences (SSE) between the real and the *implied* equity values. In addition, we calculate standard errors (displayed in parentheses) by inverting the Hessian matrix of the constrained minimization problem and taking the square root of the diagonal elements.[4] The constraints are the following: $\mathrm{Var}^{(i)}_{j\Delta_t} > 0$, $\delta^{(i)} > 0$, $\mu_2^{(i)}\rho > 0$, $\sigma^{(i)} > 0$, $M^{(i)} > 0$. We choose the parameter vectors that provide the lowest likelihood (up to the first decimal point) and the best approximation of the equity process (the lowest SSE), together with realistic standard errors. We want to point out that even though multiple parameter vectors satisfy these criteria, the resulting SSE and the default probabilities are robust to the parameter estimates.

4.1. Results (1) 2007–2010

The estimated joint default probabilities of all three companies for 2011-2014 are close to 0 for both models . Therefore, we focus on 2007-2010 and report the results only for this period. The estimated model parameters are listed in Tables B.4 and B.5 in Appendix B. The graphs that compare the calculated (implied) equity values and real equity values are displayed below the tables of the estimated parameters. Refer to Table 2 for the estimated default probabilities. This table displays the probabilities of defaulting at the end of the indicated years. "All" denotes the probability that all three banks default at the end of the year. These default probabilities are predictions for the next 1 year but they also indicate how healthy the financial situation of the company is at the moment of estimation, based on 1-year data. If the parameters of the asset process remain the same for next year, what is the probability of defaulting at the end of the year? This is the question we try to answer here. For 2008, we notice that the default probability of Citi is quite high in the case of the shot noise model. The data of 2007 is used for estimating the parameters for 2008 and as we can see from Fig. B.3b, Citi's equity is decreasing in this period; therefore, this result is not surprising. Citi was bailed out in 2008 and we can say that our model predicted the deteriorating asset quality better than the GBM. For 2009, the

[4]However, from the description of the MATLAB manual of the function "fmincon", the reported standard errors may not be accurate. In addition, we observe that the standard errors depend on the initial values.

Table 2

Simultaneous Default Probability Matrices

GBM Model

2007	JPM	Citi	BAC
JPM	0	0	0
Citi	0	0	0
BAC	0	0	0
All		0	

Shot Noise Model

2007	JPM	Citi	BAC
JPM	0	0	0
Citi	0	0	0
BAC	0	0	0
All		0	

GBM Model

2008	JPM	Citi	BAC
JPM	0	0	0
Citi	0	0.02748	0
BAC	0	0	0
All		0	

Shot Noise Model

2008	JPM	Citi	BAC
JPM	0	0	0
Citi	0	0.99996	0
BAC	0	0	0
All		0	

GBM Model

2009	JPM	Citi	BAC
JPM	0.00566	0.00566	0.00486
Citi	0.00566	0.99961	0.85154
BAC	0.00486	0.85154	0.85187
All		0.00486	

Shot Noise Model

2009	JPM	Citi	BAC
JPM	0.00745	0.00745	0.00745
Citi	0.00745	0.99995	0.93235
BAC	0.00745	0.93235	0.93240
All		0.00745	

GBM Model

2010	JPM	Citi	BAC
JPM	0.09304	0.04043	0.00047
Citi	0.04043	0.42820	0.00180
BAC	0.00047	0.00180	0.00442
All		0.00019	

Shot Noise Model

2010	JPM	Citi	BAC
JPM	0.05618	0.01361	0.00299
Citi	0.01361	0.23726	0.01293
BAC	0.00299	0.01293	0.05545
All		0.00070	

probabilities predicted by the shot noise model are larger than those by the GBM (except for the case of JPM, where there is only slight difference). We confirm this feature by comparing the pair-wise default probabilities. For 2010, the joint default probabilities of the shot noise model are lower than those of the GBM for some pairs; however, the simultaneous default probability of all three companies is three times larger for our model than for the GBM. Even so, both estimates are close to zero.

As for goodness of fit, we review the graphs located below Tables B.4 and B.5 in Appendix B. In each graph, we plot, by the dashed lines, the actual equity value and by the solid lines, the implied equity value. The latter is implied from the asset value estimated by the shot noise model (upper panel) and the GBM (lower panel). Here, the indicated periods stand for the abovementioned 1-year subperiods. For example, 2006 denotes 2005/12/30∼2006/12/29. We briefly comment on these graphs. For JPM, see Fig. B.2. Except for Fig. B.2c, in which both models give similar results, the shot noise model has better performance, especially in 2006 and 2009. In the case of Citi, our model still outperforms the GBM, on average. Please refer to Fig. B.3. The performance of our model is especially good in 2008 and 2009. The results are similar for BAC (Fig. B.4). The two models give similar performance in 2007 and 2008 but the shot noise model provides much better fit of the equity values in 2006 and 2009.

We think 1-year period is too long to measure big fluctuations in market that occurred during short periods of time in subprime mortgage crisis. For this reason, we examine this period closely in the next subsection.

4.2. A Closer Look at the Subprime Mortgage Crisis

We know that during the subprime mortgage crisis, there was a big change in a short period of time. Therefore, we will take a closer look at the evolution of the simultaneous default probabilities during 2007–2009. We display the estimated probabilities of the three companies defaulting simultaneously in Fig. 4. In this analysis, we use a 6-month moving window estimation procedure. Using 6-month data, we estimate model parameters and with these estimates, we simulate asset paths for the next 6 months (by the same method as in a 1-year case) and calculate the probability of defaulting at the end of the sixth month. Then, we roll the estimation window forward by one month. For example, the default probability for the horizontal axis label 3 in Fig. 4 indicates the default at the end of December 2008, and the data used for the calculation is from the end of December 2007 to the end of June 2008. Table 3 summarizes the estimation periods and the 6-month risk-free rates used for the calculation of debt levels (the calculation method is the same as the one described on p.27 with 1 year changed to 6 months). The estimated model parameters are given in Tables B.8, B.9, B.10, B.11, and B.12 in Appendix B. The estimated values of default probabilities are displayed in Table B.7. These default probabilities are predictions for the next 6 months but they also indicate the financial situation of the company at the moment of estimation, just as in Section 4.1. If the parameters

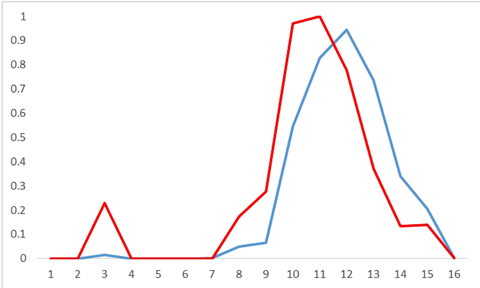


Fig. 4. The Simultaneous Default Probabilities of All Three Companies. The red line denotes the shot noise model and the blue line denotes the GBM model. The horizontal axis is explained in Table 3.

Table 3

This table explains the horizontal axis in Fig. 4 and lists 6-month treasury yields used as risk-free rates during default probability analysis of Section 4.2. The rates are observed at the end of each Period Used During Estimation. Source: US Department of the Treasury.

Horizontal Axis Label	Period Used During Estimation	Intereset Rate (%)	Horizontal Axis Label	Period Used During Estimation	Intereset Rate (%)
1	2007.10.31 - 2008.4.30	1.64	9	2008.6.30 - 2008.12.31	0.27
2	2007.11.30 - 2008.5.30	2.01	10	2008.7.31 - 2009.1.30	0.36
3	2007.12.31 - 2008.6.30	2.17	11	2008.8.29 - 2009.2.27	0.45
4	2008.1.31 - 2008.7.31	1.89	12	2008.9.30 - 2009.3.31	0.43
5	2008.2.29 - 2008.8.29	1.97	13	2008.10.31 - 2009.4.30	0.29
6	2008.3.31 - 2008.9.30	1.6	14	2008.11.28 - 2009.5.29	0.3
7	2008.4.30 - 2008.10.31	0.94	15	2008.12.31 - 2009.6.30	0.35
8	2008.5.30 - 2008.11.28	0.44	16	2009.1.30 - 2009.7.31	0.26

of the asset process remain the same for the next 6 months, what is the probability of defaulting at the end of the 6th month? Our goal is to answer this question. In case of the shot noise model, the data from the first 6 months of 2008 (horizontal axis label 3) already gives an alarm that there is a high probability (22.91%) of the simultaneous default, while the GBM model gives the probability of only 1.44%. After this alarm, the simultaneous default probabilities go to 0 from both models, and later start increasing. Furthermore, we observe that the shot noise model is faster than the GBM to capture the effect of negative shocks and give signals of high default risk for these subsequent periods.

4.3. Regression for CDS Spreads

In this study, our main focus is on the simultaneous default probabilities during 2007–2010, since this period encompasses the subprime mortgage crisis. After taking a closer look at this period in Section 4.2, we can say that our model captures the effect of negative shocks better and gives warning faster than the GBM model. In this subsection, we use the CDS spread data to test our model from a different perspective. We want to know if our model has the same power to explain the CDS spreads as the GBM model that is already widely used in practice.

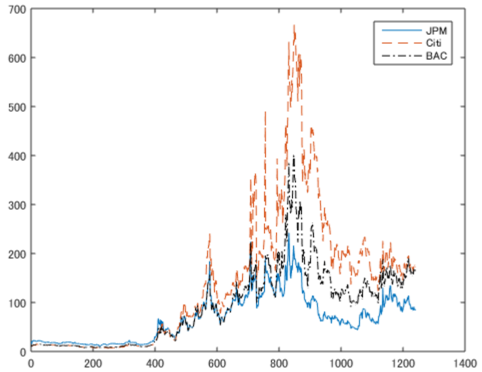


Fig. 5. CDS Spreads for All Three Companies. The vertical axis denotes CDS spread value in basis points and the horizontal axis denotes the corresponding $(\cdot)^{th}$ trading day starting from 2005/12/30.

We used Senior 5-year CDS Spread Mid data from Thomson Reuters (datasource-CMA) for JPM, Citi, and BAC. We call this data "CDS spread". Since we use daily data in our previous calculations, we also use daily CDS spreads. The data are available only until 2010/09/30. We focus on the 2-year period starting from the end of 2007 until the end of 2009 and use this data for our estimation. Fig. 5 illustrates the dynamics of CDS spreads. Observations 261–784 are from 2006/12/29 to 2008/12/31 and show that CDS spreads for Citi and BAC increased significantly in 2009. To check the connection between CDS spreads and asset values, we first simulate the asset processes for both models. We illustrate this by an example. Using the 2008 (again, this means 2007/12/31 $\sim$ 2008/12/31) data, we have already estimated parameters and used them to predict defaults at the end of 2009 in Section 4.1. Rather than predicting, we now use the 2008 parameters to simulate asset processes for the year of 2008. We retrieve initial asset values from the Black–Scholes equation (2.8) using the variance parameter and then, simulate asset paths 50,000 times and take the average. Using the 2009 data, we simulate asset paths for 2009 in a similar manner. We take the average of asset values for overlapping observations (e.g., the last value of 2008 and the first value of 2009). By putting these paths together, we obtain time-series of asset values from both models. Then, using R 3.1.3, we regress the CDS spreads of each company on all three asset time-series of the period 2007/12/31–2009/12/31. The ordinary least squares regression equations are as follows:

$$\begin{cases} JPMcds = c_1 + \alpha_1 JPMgbm + \beta_1 Citigbm + \gamma_1 BACgbm + \varepsilon_{JPM}, \\ Citicds = c_2 + \alpha_2 JPMgbm + \beta_2 Citigbm + \gamma_2 BACgbm + \varepsilon_{Citi}, \\ BACcds = c_3 + \alpha_3 JPMgbm + \beta_3 Citigbm + \gamma_3 BACgbm + \varepsilon_{BAC}, \end{cases}$$

and

$$\begin{cases} JPMcds = c_1^\star + \alpha_1^\star JPMshotnoise + \beta_1^\star Citishotnoise + \gamma_1^\star BACshotnoise + \varepsilon_{JPM}^\star, \\ Citicds = c_2^\star + \alpha_2^\star JPMshotnoise + \beta_2^\star Citishotnoise + \gamma_2^\star BACshotnoise + \varepsilon_{Citi}^\star, \\ BACcds = c_3^\star + \alpha_3^\star JPMshotnoise + \beta_3^\star Citishotnoise + \gamma_3^\star BACshotnoise + \varepsilon_{BAC}^\star. \end{cases}$$

JMPcds, Citicds, and BACcds are CDS spread data vectors. JPMgbm, Citigbm, and BACgbm are asset vectors simulated from the GBM model. JPMshotnoise, Citishotnoise, and BACshotnoise are asset vectors simulated from our model. Finally, c denotes constant terms. For each company's CDS spread, we perform two regressions: one based on assets simulated from the GBM model and another based on assets simulated from our model. See Tables 4, 5, and 6 below. The results from the GBM and shot noise model are headed by (1) and (2), respectively. All parameters of shot noise model are statistically significant at 1% level.

5. Conclusion

We will summarize the main results of our analysis. The shot noise model predicted high probability of Citi's default at the end of 2008. A closer look at 2007–2009 demonstrated that the shot noise model is more reactive to the financial distress in the banking industry than the GBM, giving high simultaneous default probability predictions of all three companies earlier. This is because the shot noise model incorporates the dependencies of companies. In addition, the graphs displayed in Appendix B demonstrate how well our model fits the real equity data. Furthermore, the linear regression shows that the asset values generated from our model do not fall behind the GBM model's asset values when it comes to explaining the CDS spread data. Finally, we wish to emphasize that early alarming that the shot noise model can offer should be of critical importance for risk management.

To conclude, our model is promising and worth considering for further refinement. It is easy to expand this model and use it in an analysis of more than three companies. Adding more companies to the model would not increase computational burden since we use only one shot noise process. In addition, we wish to analyze simultaneous defaults in different industries in the future. For this, we consider using multiple shot noise processes to incorporate industry dependencies in the model.

Table 4

Regression Results for JPM CDS Spread

	Dependent variable:	
	JPMcds	
	(1)	(2)
JPMgbm	-0.0003^{***}	
	(0.00005)	
Citigbm	-0.0003^{***}	
	(0.00003)	
BACgbm	-0.0004^{***}	
	(0.00002)	
JPMshotnoise		-0.0002^{***}
		(0.00002)
Citishotnoise		-0.0002^{***}
		(0.00001)
BACshotnoise		-0.0004^{***}
		(0.00001)
Constant	$1,595.8240^{***}$	$1,290.5320^{***}$
	(121.2533)	(52.8553)
Observations	524	524
R^2	0.6003	0.6288
Adjusted R^2	0.5980	0.6267
Residual Std. Error (df = 520)	25.0123	24.1025
F Statistic (df = 3; 520)	260.2795^{***}	293.6332^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 5

Regression Results for Citi CDS Spread

	Dependent variable:	
	Citicds	
	(1)	(2)
JPMgbm	-0.0014^{***}	
	(0.0002)	
Citigbm	-0.0016^{***}	
	(0.0001)	
BACgbm	-0.0006^{***}	
	(0.0001)	
JPMshotnoise		-0.0008^{***}
		(0.0001)
Citishotnoise		-0.0008^{***}
		(0.00003)
BACshotnoise		-0.0005^{***}
		(0.00004)
Constant	$5{,}689.3910^{***}$	$3{,}389.5770^{***}$
	(418.2004)	(160.7776)
Observations	524	524
R^2	0.6122	0.7199
Adjusted R^2	0.6099	0.7183
Residual Std. Error (df = 520)	86.2669	73.3159
F Statistic (df = 3; 520)	273.6059^{***}	445.4525^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 6

Regression Results for BAC CDS Spread

	Dependent variable:	
	BACcds	
	(1)	(2)
JPMgbm	-0.0006^{***}	
	(0.0001)	
Citigbm	-0.0007^{***}	
	(0.0001)	
BACgbm	-0.0004^{***}	
	(0.00003)	
JPMshotnoise		-0.0003^{***}
		(0.00003)
Citishotnoise		-0.0004^{***}
		(0.00002)
BACshotnoise		-0.0003^{***}
		(0.00002)
Constant	$2{,}690.5180^{***}$	$1{,}721.1900^{***}$
	(211.5637)	(85.5596)
Observations	524	524
R^2	0.5846	0.6680
Adjusted R^2	0.5822	0.6661
Residual Std. Error (df = 520)	43.6416	39.0159
F Statistic (df = 3; 520)	243.9732^{***}	348.7913^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

Appendix A. Mathematical Proofs

Appendix A.1. Dynamics of $V^{(i)} = (V_t^{(i)})_{t\geq 0}$ under approximation (Section 2.3)

First, we prove that $Z_t^{(i)} = Z_0^{(i)} e^{-\delta^{(i)}t} + \sqrt{2\delta^{(i)}} \int_0^t e^{-\delta^{(i)}(t-s)} \mathrm{d}W_s$ is a semimartingale. Applying Itô's formula to $e^{\delta^{(i)}t} W_t$, we obtain

$$e^{\delta^{(i)}t} W_t = \int_0^t \delta^{(i)} e^{\delta^{(i)}s} W_s \mathrm{d}s + \int_0^t e^{\delta^{(i)}s} \mathrm{d}W_s.$$

Multiplying both sides by $e^{-\delta^{(i)}t}$, we get

$$W_t = \int_0^t \delta^{(i)} e^{-\delta^{(i)}(t-s)} W_s \mathrm{d}s + \int_0^t e^{-\delta^{(i)}(t-s)} \mathrm{d}W_s,$$

which is

$$\int_0^t e^{-\delta^{(i)}(t-s)} \mathrm{d}W_s = W_t - \int_0^t \delta^{(i)} e^{-\delta^{(i)}(t-s)} W_s \mathrm{d}s.$$

After substituting this into the equation of $Z_t^{(i)}$, we obtain

$$\begin{aligned} Z_t^{(i)} &= Z_0^{(i)} e^{-\delta^{(i)}t} + \sqrt{2\delta^{(i)}} W_t - \sqrt{2\delta^{(i)}} \int_0^t \delta^{(i)} e^{-\delta^{(i)}(t-s)} W_s \mathrm{d}s \\ &= Z_0^{(i)} + \left(Z_0^{(i)} e^{-\delta^{(i)}t} - Z_0^{(i)} - \sqrt{2\delta^{(i)}} \int_0^t \delta^{(i)} e^{-\delta^{(i)}(t-s)} W_s \mathrm{d}s \right) + \sqrt{2\delta^{(i)}} W_t \end{aligned} \tag{A.1}$$

The process inside the brackets starts at 0. For any fixed t, the integral inside the brackets is a well-defined Lebesgue–Stieltjes integral and as a function of t, it is of bounded variation (Karatzas and Shreve [15, Remark 4.6(i) p. 23, 150]). This means that $Z_t^{(i)}$ can be expressed as a sum of the initial value, the bounded variation process, and the local martingale. Hence, it is a semimartingale. In addition, we obtain

$$\langle Z^{(i)}, Z^{(i)} \rangle_t = 2\delta^{(i)} t. \tag{A.2}$$

Next, we show the dynamics of the asset value

$$V_t^{(i)} = e^{X_0^{(i)} + \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)t + \sigma^{(i)} B_t^{(i)} - \frac{\mu_1^{(i)}\rho}{\delta^{(i)}} - Z_t^{(i)} \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}}$$

by Itô's formula. Since $\mathrm{d}Z_t^{(i)} = -\delta^{(i)}Z_t^{(i)}\mathrm{d}t + \sqrt{2\delta^{(i)}}\mathrm{d}W_t$ and $\langle B^{(i)}, Z^{(i)}\rangle_t = \langle B^{(i)}, \sqrt{2\delta^{(i)}}W\rangle_t = \sqrt{2\delta^{(i)}}k^{(i)}t$, we obtain

$$\begin{aligned}
\mathrm{d}V_t^{(i)} &= \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)V_t^{(i)}\mathrm{d}t + \sigma^{(i)}V_t^{(i)}\mathrm{d}B_t^{(i)} + \frac{1}{2}(\sigma^{(i)})^2V_t^{(i)}\mathrm{d}t - \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}V_t^{(i)}\mathrm{d}Z_t^{(i)} \\
&+ \frac{1}{2}\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}V_t^{(i)}\mathrm{d}\langle Z^{(i)}, Z^{(i)}\rangle_t - V_t^{(i)}\frac{1}{2}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\sigma^{(i)}\mathrm{d}\langle B^{(i)}, Z^{(i)}\rangle_t - V_t^{(i)}\frac{1}{2}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\sigma^{(i)}\mathrm{d}\langle B^{(i)}, Z^{(i)}\rangle_t \\
&= \mu^{(i)}V_t^{(i)}\mathrm{d}t + \sigma^{(i)}V_t^{(i)}\mathrm{d}B_t^{(i)} - V_t^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}(-\delta^{(i)}Z_t^{(i)}\mathrm{d}t + \sqrt{2\delta^{(i)}}\mathrm{d}W_t) + \frac{1}{2}\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}(2\delta^{(i)})V_t^{(i)}\mathrm{d}t \\
&- V_t^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\sigma^{(i)}\sqrt{2\delta^{(i)}}k^{(i)}\mathrm{d}t \\
&= V_t^{(i)}\left(\left(\mu^{(i)} + \frac{1}{2}\mu_2^{(i)}\rho + \delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\sigma^{(i)}k^{(i)}\right)\mathrm{d}t + \sigma^{(i)}dB_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\mathrm{d}W_t\right).
\end{aligned} \tag{A.3}$$

Furthermore, since W_t is a Brownian motion, $\tilde{k}^2 + \sum_{j=1}^{m}(k^{(j)})^2 = 1$ and we have

$$\begin{aligned}
&\sigma^{(i)}dB_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\mathrm{d}W_t \\
&= \sigma^{(i)}dB_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\left(\sum_{1\le j\le m, j\ne i}k^{(j)}\mathrm{d}B_t^{(j)} + \tilde{k}\mathrm{d}\tilde{B}_t\right)\frac{\sqrt{1-(k^{(i)})^2}}{\sqrt{1-(k^{(i)})^2}} \\
&= (\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)})\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\sqrt{1-(k^{(i)})^2}\mathrm{d}W_t^{(-i)} \\
&= (\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)})\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho}\sqrt{1-(k^{(i)})^2}\mathrm{d}W_t^{(-i)}
\end{aligned}$$

where $B_t^{(i)}$ and $W_t^{(-i)} = \left(\sum_{1\le j\le m, j\ne i}k^{(j)}\mathrm{d}B_t^{(j)} + \tilde{k}\mathrm{d}\tilde{B}_t\right)\frac{1}{\sqrt{1-(k^{(i)})^2}}$ are independent standard Brownian motions. Then,

$$\begin{aligned}
\mathrm{d}V_t^{(i)} = V_t^{(i)}\Bigg(&\Bigg(\mu^{(i)} + \frac{1}{2}\mu_2^{(i)}\rho + \delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_t^{(i)} - \\
&- \sqrt{\mu_2^{(i)}\rho}\sigma^{(i)}k^{(i)}\Bigg)\mathrm{d}t + (\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)})\mathrm{d}B_t^{(i)} - \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}\mathrm{d}W_t^{(-i)}\Bigg)
\end{aligned}$$

which is (2.7).

Appendix A.2. Proof of the Novikov condition being satisfied (Section 2.4)

We define

$$H_t(\theta) = e^{\int_0^t \theta_s^{(i)}dB_s^{(i)} + \int_0^t \theta_s^{(-i)}dW_s^{(-i)} - \frac{1}{2}\int_0^t ||\theta_s||^2 ds},$$

where

$$\theta_t^{(i)} = \frac{r - Q_t^{(i)} + \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}}{\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}} \quad \text{and} \quad \theta_t^{(-i)} = 1.$$

We have to show that Novikov's condition $E\left[e^{\frac{1}{2}\int_0^T ||\theta_t||^2 dt}\right] < \infty$ is satisfied. Let $K = r - \mu^{(i)} - \frac{\mu_2^{(i)}\rho}{2} + \sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}k^{(i)} + \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)}$.

$$\begin{aligned}
\frac{1}{2}||\theta_t||^2 &= \frac{\left(r - \mu^{(i)} - \frac{\mu_2^{(i)}\rho}{2} + \sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}k^{(i)} + \sqrt{\mu_2^{(i)}\rho(1-(k^{(i)})^2)} - \delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_t^{(i)}\right)^2}{2\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)^2} + \frac{1}{2} \\
&= \frac{\left(K - \delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_t^{(i)}\right)^2}{2\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)^2} + \frac{1}{2}.
\end{aligned}$$

Recalling (2.2), we have

$$\frac{1}{2}||\theta_t||^2 \le \frac{1}{2} + \frac{\left(|K| + \left|\delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_0^{(i)}e^{-\delta^{(i)}t}\right| + \delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left|\int_0^t e^{-\delta^{(i)}(t-s)}dW_s\right|\right)^2}{2\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)^2}$$

Since $e^{-\delta^{(i)}t} \le 1$, $|K| + \left|\delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_0^{(i)}e^{-\delta^{(i)}t}\right| \le |K| + \left|\delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_0^{(i)}\right|$. Since we have assumed that $Z_0^{(i)}$ is bounded, $|K| + \left|\delta^{(i)}\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_0^{(i)}\right| \le C$, where C is some constant. Let $W_T^* := \sup_{0\le t\le T}|W_t|$. Then,

$$\left|\int_0^t e^{-\delta^{(i)}(t-s)}dW_s\right| \le |W_t| + \delta^{(i)}\int_0^t |W_s|\,ds \le W_T^* + \delta^{(i)}W_T^*T \tag{A.4}$$

and we obtain

$$\frac{1}{2}||\theta_t||^2 \le \frac{1}{2} + \frac{\left(C + \delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(W_T^* + \delta^{(i)}W_T^*T\right)\right)^2}{2\left(\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho}k^{(i)}\right)^2}. \tag{A.5}$$

We take any sequence that satisfies $0 = t_0 < t_1 < \ldots < t_n \to \infty$ and consider

$$\mathbb{E}\left[e^{\frac{1}{2}\int_{t_{n-1}}^{t_n}||\theta_t||^2 dt}\right] \le e^{\frac{t_n - t_{n-1}}{2}}\mathbb{E}\left[e^{\frac{1}{2}\int_{t_{n-1}}^{t_n}\frac{\left(C+\delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(1+\delta^{(i)}T\right)W_T^*\right)^2}{\left(\sigma^{(i)}-\sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)^2}dt}\right]$$

$$= e^{\frac{t_n - t_{n-1}}{2}}\mathbb{E}\left[e^{\frac{1}{2}(t_n - t_{n-1})\frac{\left(C+\delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(1+\delta^{(i)}T\right)W_T^*\right)^2}{\left(\sigma^{(i)}-\sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)^2}}\right],$$

where $\frac{C+\delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(1+\delta^{(i)}T\right)W_t}{\sqrt{2}\left(\sigma^{(i)}-\sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)}$ is a Gaussian process. For each ω, Brownian motion is continuous in t; therefore, its maximum on $[0,T]$ exists. From the Burkholder–Davis–Gundy inequality (Karatzas and Shreve [15, p. 166]), there is some constant $K_{\frac{1}{2}}$ and $\mathbb{E}[W_T^*] \le K_{\frac{1}{2}}\sqrt{T}$. Now, we have established

$$\sup_{0\le t\le T}\left|\frac{C+\delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(1+\delta^{(i)}T\right)W_t}{\sqrt{2}\left(\sigma^{(i)}-\sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)}\right| < \infty \quad a.s. \tag{A.6}$$

Then, by Landau and Shepp [18], some $\varepsilon > 0$ exists, such that

$$\mathbb{E}\left[\exp\left(\frac{\varepsilon\left(C+\delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(1+\delta^{(i)}T\right)W_T^*\right)^2}{2\left(\sigma^{(i)}-\sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)^2}\right)\right] < \infty.$$

We make the interval $t_n - t_{n-1}, \ \forall n \ge 1$ equal to ε. Then,

$$\mathbb{E}\left[\exp\left(\left(t_n - t_{n-1}\right)\frac{\left(C+\delta^{(i)}\sqrt{\mu_2^{(i)}\rho}\left(1+\delta^{(i)}T\right)W_T^*\right)^2}{2\left(\sigma^{(i)}-\sqrt{\mu_2^{(i)}\rho k^{(i)}}\right)^2}\right)\right] < \infty, \quad \forall n \ge 1.$$

This means that

$$\mathbb{E}\left[e^{\frac{1}{2}\int_{t_{n-1}}^{t_n}||\theta_t||^2 dt}\right] < \infty, \quad \forall n \ge 1 \tag{A.7}$$

and from Corollary 5.14 from Karatzas and Shreve [15, p. 199], we can say that $H_t(\theta)$ is a martingale.[5] Because $H_t(\theta)$ is a martingale, we can change the measure and $\left(\tilde{B}^{(i)}, \tilde{W}^{(-i)}\right)_t$ becomes two-dimensional Brownian motion under $\tilde{P}_T$.

[5] The idea of this proof was borrowed from Nate Eldredge http://math.stackexchange.com/questions/133691/can-i-apply-the-girsanov-theorem-to-an-ornstein-uhlenbeck-process Accessed: 2015-05-08

Now, $\tilde{\tilde{B}}_t^{(i)} := \frac{\sigma^{(i)} - \sqrt{\mu_2^{(i)}\rho} k^{(i)}}{M^{(i)}} \tilde{B}_t^{(i)} - \frac{\sqrt{\mu_2^{(i)}\rho\left(1-(k^{(i)})^2\right)}}{M^{(i)}} \tilde{W}_t^{(-i)}$ is a martingale and its quadratic variation is t (by the definition of $M^{(i)}$). Therefore, it is a standard Brownian motion under $\tilde{P}_T$. Finally, $\mathrm{d}(e^{-rt}V_t^{(i)}) = e^{-rt}V_t^{(i)} M^{(i)} \mathrm{d}\tilde{\tilde{B}}_t^{(i)}$ and $e^{-rt}V_t^{(i)}$ is a local martingale under $\tilde{\tilde{\mathbb{P}}}_T$. However, we also obtain

$$\tilde{\tilde{\mathbb{E}}}\left[e^{\frac{1}{2}\int_0^T (M^{(i)})^2 \mathrm{d}t}\right] < \infty$$

and from Proposition 14.2 in Steele [27, p.241], we conclude that $e^{-rt}V_t^{(i)}$ is a martingale under $\tilde{\tilde{\mathbb{P}}}_T$.

Appendix A.3. Derivation of the Log-likelihood Function (Section 2.5)

First, we calculate $\mathbb{E}\left[\ln V_{j\Delta_t}^{(i)} \mid V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right]$ and $\mathrm{Var}\left[\ln V_{j\Delta_t}^{(i)} \mid V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right]$. From (2.1), we have

$$\ln\left(\frac{V_{j\Delta_t}^{(i)}}{V_{(j-1)\Delta_t}^{(i)}}\right) = \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)\Delta_t + \sigma^{(i)}\left(B_{j\Delta_t}^{(i)} - B_{(j-1)\Delta_t}^{(i)}\right) - \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}\left(Z_{j\Delta_t}^{(i)} - Z_{(j-1)\Delta_t}^{(i)}\right),$$

where we compute

$$\begin{aligned}
&Z_{j\Delta_t}^{(i)} - Z_{(j-1)\Delta_t}^{(i)} \\
&= Z_0^{(i)} e^{-\delta^{(i)} j\Delta_t} + \sqrt{2\delta^{(i)}} e^{-\delta^{(i)} j\Delta_t} \int_0^{j\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s - Z_0^{(i)} e^{-\delta^{(i)}(j-1)\Delta_t} \\
&- \sqrt{2\delta^{(i)}} e^{-\delta^{(i)}(j-1)\Delta_t} \int_0^{(j-1)\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s \\
&= Z_0^{(i)} e^{-\delta^{(i)} j\Delta_t}(1 - e^{\delta^{(i)}\Delta_t}) + \sqrt{2\delta^{(i)}}\left[e^{-\delta^{(i)} j\Delta_t} \int_0^{j\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s - e^{-\delta^{(i)}(j-1)\Delta_t} \int_0^{(j-1)\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s\right]
\end{aligned}$$

and

$$\begin{aligned}
&e^{-\delta^{(i)} j\Delta_t} \int_0^{j\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s - e^{-\delta^{(i)}(j-1)\Delta_t} \int_0^{(j-1)\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s = \\
&\qquad e^{-\delta^{(i)} j\Delta_t} \int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s + \left(\frac{1}{e^{\delta^{(i)}\Delta_t}} - 1\right) e^{-\delta^{(i)}(j-1)\Delta_t} \int_0^{(j-1)\Delta_t} e^{\delta^{(i)} s} \mathrm{d}W_s,
\end{aligned}$$

and it follows that

$$\begin{aligned}
&\ln\left(\frac{V_{j\Delta_t}^{(i)}}{V_{(j-1)\Delta_t}^{(i)}}\right) \\
&= \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)\Delta_t + \sigma^{(i)}(B_{j\Delta_t}^{(i)} - B_{(j-1)\Delta_t}^{(i)}) - \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}} Z_0^{(i)} e^{-\delta^{(i)} j\Delta_t}(1 - e^{\delta^{(i)}\Delta_t}) \\
&- \sqrt{\mu_2^{(i)}\rho} \cdot e^{-\delta^{(i)} j\Delta_t} \int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)} s} dW_s - \sqrt{\mu_2^{(i)}\rho}\left(\frac{1}{e^{\delta^{(i)}\Delta_t}} - 1\right) e^{-\delta^{(i)}(j-1)\Delta_t} \int_0^{(j-1)\Delta_t} e^{\delta^{(i)} s} dW_s.
\end{aligned}$$

Since we use daily data for the analysis, $\Delta_t = \frac{1}{360} \approx 0.0028$. We assume $\delta^{(i)}$ is small enough and satisfies $e^{\delta^{(i)}\Delta_t} \approx 1$. Note that we can check our assumption on this point by comparing the equity values implied by our model (with the estimated parameters) to the real data. Note also that our experiments without removing this term produces $\delta^{(i)}$ small enough to justify this assumption. $\mathbb{E}\left[e^{-\delta^{(i)}(j-1)\Delta_t}\int_0^{(j-1)\Delta_t} e^{\delta^{(i)}s}dW_s\right] = 0$. Also, from equation (A.4) we have

$$\left|e^{-\delta^{(i)}(j-1)\Delta_t}\int_0^{(j-1)\Delta_t} e^{\delta^{(i)}s}dW_s\right| \leq W_T^* + \delta^{(i)}W_T^*T < \infty.$$

Therefore, we may assume that

$$\sqrt{\mu_2^{(i)}\rho}\left(\frac{1}{e^{\delta^{(i)}\Delta_t}} - 1\right)e^{-\delta^{(i)}(j-1)\Delta_t}\int_0^{(j-1)\Delta_t} e^{\delta^{(i)}s}dW_s \approx 0.$$

We do not have any information about $Z_0^{(i)}$, expect that it takes bounded values according to our assumption. To check how big the actual value of $Z_0^{(i)}$ is, we estimate the realized value of $Z_0^{(i)}$; therefore, we do not ignore the term $\sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_0^{(i)}e^{-\delta^{(i)}j\Delta_t}\left(1-e^{\delta^{(i)}\Delta_t}\right)$. Following the above arguments, conditioning on $V_{(j-1)\Delta_t}^{(i)}$ and $\mathscr{F}_0$, $\ln V_{j\Delta_t}^{(i)}$ follows the normal distribution with the (conditional) mean

$$\begin{aligned}
&\mathbb{E}\left[\ln V_{j\Delta_t}^{(i)} \,\middle|\, V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right] \\
&= \ln V_{(j-1)\Delta_t}^{(i)} + \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)\Delta_t - \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}}Z_0^{(i)}e^{-\delta^{(i)}j\Delta_t}\left(1-e^{\delta^{(i)}\Delta_t}\right)
\end{aligned}$$

since $\mathbb{E}\left[\int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)}s}\mathrm{d}W_s \,\middle|\, V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right] = 0$, and the (conditional) variance

$$\begin{aligned}
\mathrm{Var}\left[\ln V_{j\Delta_t}^{(i)} \,\middle|\, V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right] &= \mathbb{E}\left[\left(\ln V_{j\Delta_t}^{(i)} - \mathbb{E}\left[\ln V_{j\Delta_t}^{(i)} \,\middle|\, V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right]\right)^2 \,\middle|\, V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right] \\
&= \mathbb{E}\left[\sigma^{(i)}\left(B_{j\Delta_t}^{(i)} - B_{(j-1)\Delta_t}^{(i)}\right) - \sqrt{\mu_2^{(i)}\rho}e^{-\delta^{(i)}j\Delta_t}\int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)}s}\mathrm{d}W_s\right]^2
\end{aligned}$$

which in evaluated as

$$\begin{aligned}
&(\sigma^{(i)})^2\Delta_t + \mu_2^{(i)}\rho e^{-2\delta^{(i)}j\Delta_t}\mathbb{E}\left(\int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)}s}\mathrm{d}W_s\right)^2 \\
&- 2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}e^{-\delta^{(i)}j\Delta_t}\mathbb{E}\left[\int_{(j-1)\Delta_t}^{j\Delta_t}\mathrm{d}B_s^{(i)} \cdot \int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)}s}\mathrm{d}W_s\right] \\
&= (\sigma^{(i)})^2\Delta_t + \mu_2^{(i)}\rho e^{-2\delta^{(i)}j\Delta_t}\int_{(j-1)\Delta_t}^{j\Delta_t} e^{2\delta^{(i)}s}ds - 2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}e^{-\delta^{(i)}j\Delta_t}\int_{(j-1)\Delta_t}^{j\Delta_t} e^{\delta^{(i)}s}k^{(i)}ds \\
&= (\sigma^{(i)})^2\Delta_t + \mu_2^{(i)}\rho e^{-2\delta^{(i)}j\Delta_t}\frac{1}{2\delta^{(i)}}\left(e^{2\delta^{(i)}j\Delta_t} - e^{2\delta^{(i)}(j-1)\Delta_t}\right) - \frac{2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}}{\delta^{(i)}}\left(1-e^{-\delta^{(i)}\Delta_t}\right)k^{(i)} \\
&= (\sigma^{(i)})^2\Delta_t + \frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}\left(1-e^{-2\delta^{(i)}\Delta_t}\right) - \frac{2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}}{\delta^{(i)}}\left(1-e^{-\delta^{(i)}\Delta_t}\right)k^{(i)}.
\end{aligned}$$

The log likelihood function is written as

$$L\left(V_t^{(i)}, t = \Delta_t, 2\Delta_t, \ldots, n\Delta_t \mid \mathscr{F}_0\right) =$$

$$\ln\left(f\left(V_{n\Delta_t}^{(i)} | V_{(n-1)\Delta_t}^{(i)}, \mathscr{F}_0\right) f\left(V_{(n-1)\Delta_t}^{(i)} | V_{(n-2)\Delta_t}^{(i)}, \mathscr{F}_0\right) \ldots f\left(V_{\Delta_t}^{(i)} | \mathscr{F}_0\right)\right),$$

where f denotes the density function. Let us now define

$$\begin{cases} \text{mean}_{j\Delta_t}^{(i)} = \left(\mu^{(i)} - \frac{1}{2}(\sigma^{(i)})^2\right)\Delta_t - \sqrt{\frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}} Z_0^{(i)} e^{-\delta^{(i)} j\Delta_t}\left(1 - e^{\delta^{(i)}\Delta_t}\right) \\ \text{Var}_{j\Delta_t}^{(i)} = (\sigma^{(i)})^2\Delta_t + \frac{\mu_2^{(i)}\rho}{2\delta^{(i)}}\left(1 - e^{-2\delta^{(i)}\Delta_t}\right) - \dfrac{2\sigma^{(i)}\sqrt{\mu_2^{(i)}\rho}}{\delta^{(i)}}\left(1 - e^{-\delta^{(i)}\Delta_t}\right)k^{(i)} \end{cases}$$

Then, using the density of log-normal distribution,

$$f\left(V_{j\Delta_t}^{(i)} \,\middle|\, V_{(j-1)\Delta_t}^{(i)}, \mathscr{F}_0\right) = \frac{1}{\sqrt{2\pi \text{Var}_{j\Delta_t}^{(i)}} V_{j\Delta_t}^{(i)}} \exp\left(-\frac{\left(\ln\left(\frac{V_{j\Delta_t}^{(i)}}{V_{(j-1)\Delta_t}^{(i)}}\right) - \text{mean}_{j\Delta_t}^{(i)}\right)^2}{2\text{Var}_{j\Delta_t}^{(i)}}\right)$$

The log-likelihood function of unobserved $V_{j\Delta_t}^{(i)}$ becomes

$$\begin{aligned} & L\left(V_t^{(i)}, t = \Delta_t, 2\Delta_t, \ldots, n\Delta_t \,\middle|\, \mathscr{F}_0\right) \\ & = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{n} \ln \text{Var}_{j\Delta_t}^{(i)} - \sum_{j=1}^{n} \frac{\left(\ln\left(\frac{V_{j\Delta_t}^{(i)}}{V_{(j-1)\Delta_t}^{(i)}}\right) - \text{mean}_{j\Delta_t}^{(i)}\right)^2}{2\text{Var}_{j\Delta_t}^{(i)}} - \sum_{j=1}^{n} \ln V_{j\Delta_t}^{(i)} \end{aligned}$$

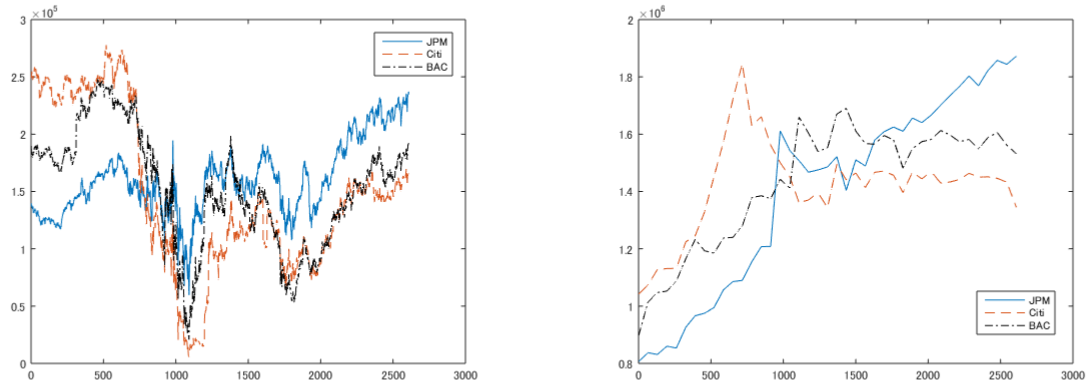
We have an element-by-element transformation from an unobserved asset sample to an observed equity sample through equation (2.8) and $\frac{\partial E_t^{(i)}}{\partial V_t^{(i)}} = \Phi(d_t^{(i)})$. Then, we can write the log-likelihood function of equity as

$$\begin{aligned} & L\left(E_t^{(i)}, t = \Delta_t, 2\Delta_t, \ldots, n\Delta_t); \mu^{(i)}, \delta^{(i)}, \mu_2^{(i)}\rho, Z_0^{(i)}, \sigma^{(i)} \,\middle|\, \mathscr{F}_0\right) = \\ & -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{n} \ln \text{Var}_{j\Delta_t}^{(i)} - \sum_{j=1}^{n} \frac{\left(\ln\left(\frac{\hat{V}_{j\Delta_t}^{(i)}}{\hat{V}_{(j-1)\Delta_t}^{(i)}}\right) - \text{mean}_{j\Delta_t}^{(i)}\right)^2}{2\text{Var}_{j\Delta_t}^{(i)}} - \sum_{j=1}^{n} \ln \hat{V}_{j\Delta_t}^{(i)} - \sum_{j=1}^{n} \ln \Phi(\hat{d}_{j\Delta_t}^{(i)}), \end{aligned}$$

where $\hat{V}_{j\Delta_t}^{(i)}$ is the unique solution to (2.8) and $\hat{d}_{j\Delta_t}^{(i)}$ is $d_{j\Delta_t}^{(i)}$ with $V_{j\Delta_t}^{(i)}$ replaced by $\hat{V}_{j\Delta_t}^{(i)}$. The derivation of (2.11) is now complete.

Appendix B. Statistical and Graphical Results

We collect some statistical results here. Matlab R2015a is used for all the calculations and to produce graphs.



(a) (b)

Fig. B.1. Equities (a) and Debts (b) for All Three Companies. The vertical axis denotes values in millions of US dollars and the horizontal axis denotes the corresponding $(\cdot)^{th}$ trading day starting from 2004/12/31.

Table B.1

JPM Company Profile

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Total Assets	1157248	1198942	1351520	1562147	2175052	2031989	2117605	2265792	2359141	2415689	2572773
Total Liabilities	1051595	1091731	1235730	1438926	2008168	1866624	1941499	2082219	2155072	2204511	2341046
ROE	5.87%	7.98%	12.96%	12.86%	3.82%	6.01%	9.69%	10.21%	10.72%	8.40%	9.75%
ROA	0.63%	0.95%	1.67%	1.35%	0.59%	0.75%	0.96%	1.01%	1.06%	0.92%	1.00%
Net Income	-33.53%	89.95%	70.27%	6.38%	-63.52%	109.24%	39.89%	10.92%	12.82%	-15.26%	21.96%

Table B.2

Citi Company Profile

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Total Assets	1484101	1494037	1884318	2187631	1938470	1856646	1913902	1818846	1864660	1880382	1842530
Total Liabilities	1374810	1381500	1764535	2074033	1794448	1701673	1748113	1639273	1673663	1674249	1630485
ROE	16.56%	22.33%	18.66%	3.08%	-31.88%	-8.28%	6.66%	6.37%	4.04%	6.88%	3.37%
ROA	1.49%	2.00%	1.73%	0.72%	-0.84%	0.10%	0.98%	-	0.72%	-	-
Net Income	-4.52%	44.25%	-12.41%	-83.21%	-	-	-	3.51%	-32.22%	81.83%	-46.29%

Table B.3

BAC Company Profile

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Total Assets	1110432	1291803	1459737	1715746	1817943	2230232	2264909	2097047	2176936	2102273	2104534
Total Liabilities	1010197	1190270	1324465	1568943	1640891	1998788	2036661	1866946	1939980	1869588	1861063
ROE	19.18%	16.35%	18.07%	10.77%	1.82%	-1.33%	-1.77%	0.04%	1.29%	4.61%	1.71%
ROA	1.70%	1.61%	1.87%	1.33%	0.59%	0.81%	0.30%	0.42%	0.48%	-	-
Net Income	30.83%	18.05%	28.35%	-29.11%	-73.25%	59.18%	-	-	189.69%	173.03%	-57.71%

Table B.4

Shot Noise Model Parameters. "Year" indicates the year from which the data was used during estimation. Standard errors are displayed in parentheses.

JPM

Year	2006		2007		2008		2009	
μ	0.04140	(0.02013)	0.10105	(0.52028)	-3.22915	(1.02244)	0.06508	(0.15777)
δ	3.36241	(0.80350)	8.87134	(4.73650)	0.03510	(0.02235)	21.52812	(2.38332)
$\mu_2\rho$	0.00036	(0.00042)	0.00006	(0.00008)	0.01821	(0.00673)	0.00001	(0.00000)
Z_0	19.34641	(5.02009)	21.98025	(3.94756)	200.44205	(11.45851)	-132.08146	(19.11733)
σ	0.02484	(0.00530)	0.04024	(0.00199)	0.02051	(0.00520)	0.07110	(0.02344)
k^2	0.05886		0.03889		0.00393		0.02768	
M	0.02731		0.03945		0.13522		0.07075	
Likelihood	-2283.98727		-2420.59085		-2705.45474		-2622.28717	

Citi

Year	2006		2007		2008		2009	
μ	5.46832	(0.91507)	-11.39959	(0.54377)	-0.86360	(0.93395)	0.49460	(0.16508)
δ	0.08332	(0.02373)	0.16515	(0.01983)	1.31767	(0.45313)	9.15615	(0.34043)
$\mu_2\rho$	0.00081	(0.00034)	0.00182	(0.00020)	0.00027	(0.00120)	0.00011	(0.00009)
Z_0	-913.16063	(12.48526)	990.97999	(13.09358)	65.86908	(21.49292)	-119.24921	(7.37256)
σ	0.00107	(0.03044)	0.00930	(0.30066)	0.10217	(0.01584)	0.21295	(0.01736)
k^2	0.05428		0.04538		0.04519		0.00577	
M	0.02828		0.04164		0.09998		0.21242	
Likelihood	-2376.54011		-2546.32342		-2641.99974		-2473.97455	

BAC

Year	2006		2007		2008		2009	
μ	-1.92586	(0.14746)	1.68165	(0.65830)	-7.55677	(2.54445)	-0.63579	(0.30954)
δ	0.44366	(0.06549)	0.10704	(0.08674)	0.05976	(0.05595)	1.30789	(0.93302)
$\mu_2\rho$	0.00063	(0.00009)	0.00073	(0.00009)	0.01147	(0.00509)	0.00032	(0.00280)
Z_0	205.30077	(11.71072)	-258.45660	(6.83843)	410.12122	(14.65558)	88.98744	(8.78695)
σ	0.02665	(0.00241)	0.02550	(0.00286)	0.10346	(0.06139)	0.09340	(0.01545)
k^2	0.00497		0.00191		0.00936		0.00926	
M	0.03534		0.03633		0.14153		0.09341	
Likelihood	-2414.38806		-2445.22217		-2721.79010		-2621.63717	

Table B.5

GBM Model Parameters. "Year" indicates the year from which the data was used during estimation. Standard errors are displayed in parentheses.

JPM

Year	2006		2007		2008		2009	
μ	0.21919	(0.03292)	0.15618	(0.04840)	0.31776	(0.06802)	-0.00104	(0.03397)
σ	0.02798	(0.00203)	0.03969	(0.00602)	0.13522	(0.01437)	0.07167	(0.01510)
Likelihood	-2290.69922		-2422.15893		-2705.46914		-2624.44213	

Citi

Year	2006		2007		2008		2009	
μ	0.31259	(0.03517)	0.03907	(0.01463)	-0.29999	(0.09195)	0.10987	(0.03191)
σ	0.02867	(0.00284)	0.04680	(0.00397)	0.10117	(0.00404)	0.22568	(0.04988)
Likelihood	-2380.15378		-2575.81747		-2642.87053		-2476.83399	

BAC

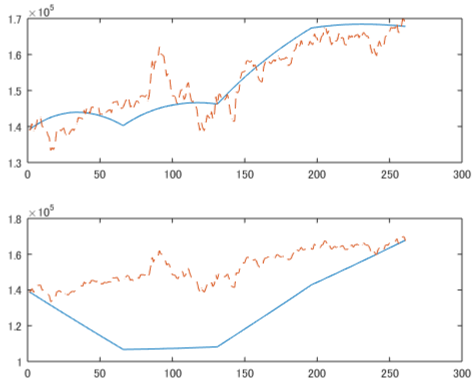
Year	2006		2007		2008		2009	
μ	0.15641	(0.02101)	0.12783	(0.04977)	-0.12223	(0.04184)	0.20402	(0.09560)
σ	0.03676	(0.00719)	0.03638	(0.00990)	0.14169	(0.01362)	0.09836	(0.01537)
Likelihood	-2424.69174		-2445.56137		-2721.94870		-2623.65780	

Table B.6

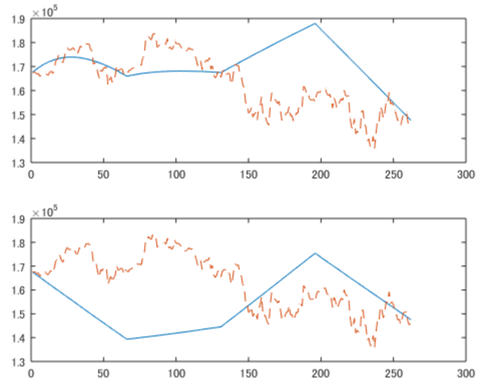
1-Year Risk-free Interest Rates

Simulation Year	Observed Date	1-Yr Treasury Yield (%)
2007	12/29/06	5
2008	12/31/07	3.34
2009	12/31/08	0.37
2010	12/31/09	0.47
2011	12/31/10	0.29
2012	12/30/11	0.12
2013	12/31/12	0.16
2014	12/31/13	0.13
2015	12/31/14	0.25

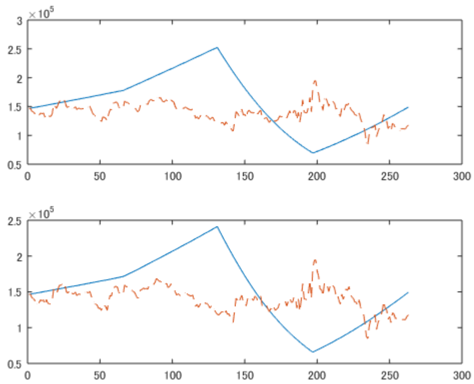
Source: US Department of the Treasury



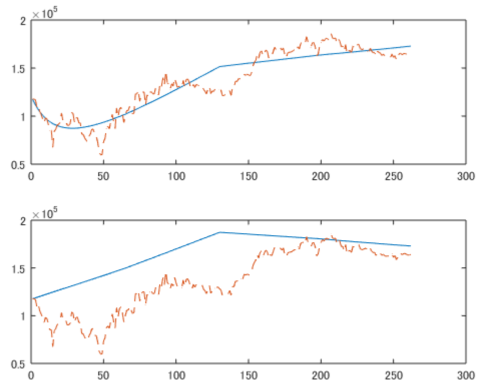
(a) 2006



(b) 2007



(c) 2008



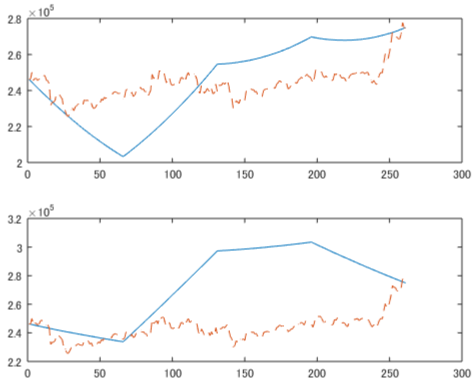
(d) 2009

Fig. B.2. Implied (Solid Line) and Realized (Dashed Line) Equities for JPM for the Corresponding Subperiods. The shot noise model is in the top panel, and the GBM model is in the bottom panel. The indicated year is the year from which the data was used during the estimation. The vertical axis denotes equity values in millions of US dollars and the horizontal axis denotes the corresponding $(\cdot)^{th}$ trading day starting from the last day of the previous year for each subperiod.

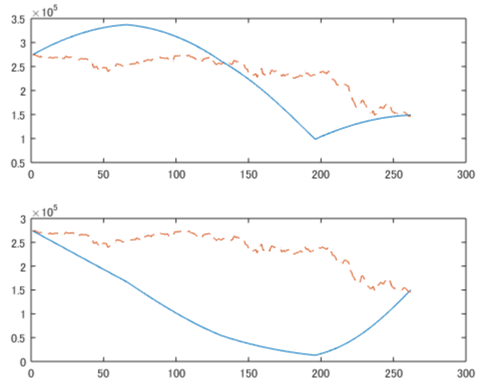
Table B.7

Simultaneous Default Probabilities Displayed in Fig. 4

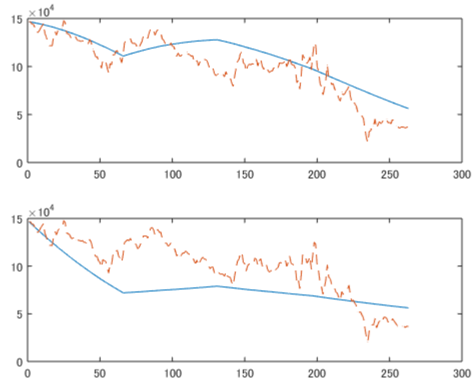
Horizontal Axis Label	GBM	Shot Noise
1	0	0
2	0	0.0003
3	0.0144	0.2291
4	0	0
5	0	0
6	0	0
7	0.0014	0
8	0.0487	0.1733
9	0.0662	0.2756
10	0.5464	0.9709
11	0.8275	0.9991
12	0.9433	0.7783
13	0.7346	0.3720
14	0.3397	0.1332
15	0.2056	0.1389
16	0.0014	0.0013



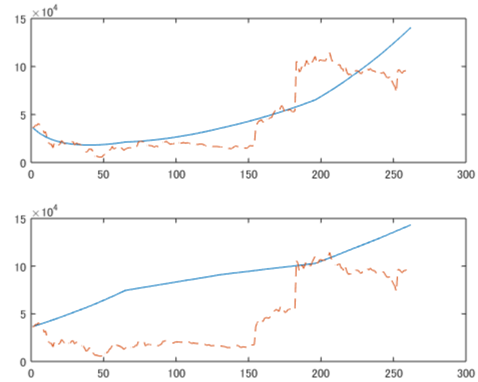
(a) 2006



(b) 2007

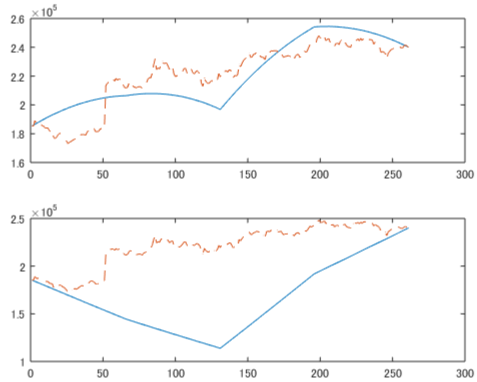


(c) 2008

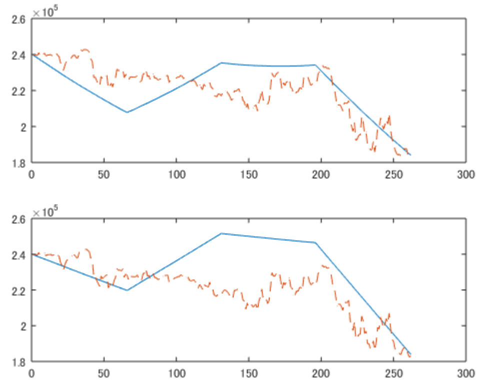


(d) 2009

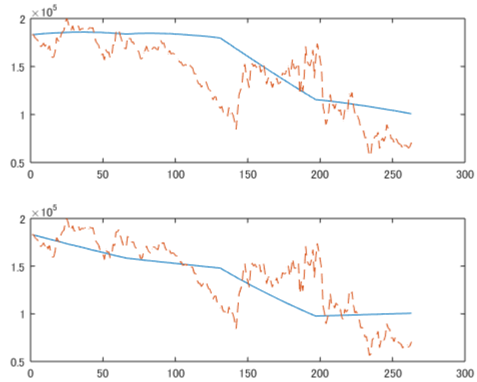
Fig. B.3. Implied (Solid Line) and Realized (Dashed Line) Equities for Citi for the Corresponding Subperiods. The shot noise model is in the top panel, and the GBM model is in the bottom panel. The indicated year is the year from which the data was used during the estimation. The vertical axis denotes equity values in millions of US dollars and the horizontal axis denotes the corresponding $(\cdot)^{th}$ trading day starting from the last day of the previous year for each subperiod.



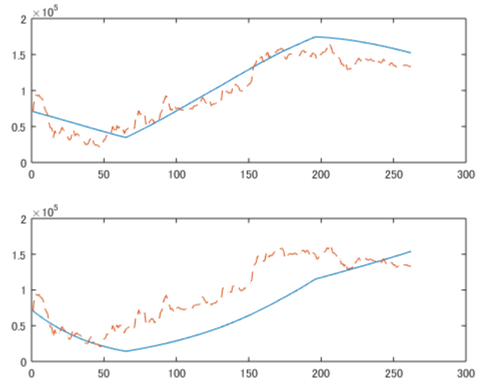
(a) 2006



(b) 2007



(c) 2008



(d) 2009

Fig. B.4. Implied (Solid Line) and Realized (Dashed Line) Equities for BAC for the Corresponding Subperiods. The shot noise model is in the top panel, and the GBM model is in the bottom panel. The indicated year is the year from which the data was used during the estimation. The vertical axis denotes equity values in millions of US dollars and the horizontal axis denotes the corresponding $(\cdot)^{th}$ trading day starting from the last day of the previous year for each subperiod.

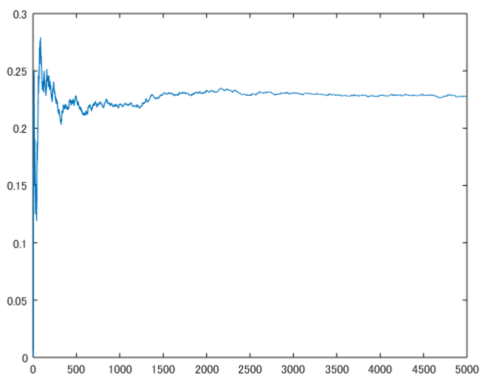


Fig. B.5. Convergence of the Simultaneous Default Probability (Table B.7, Horizontal Axis Label 3, Shot Noise) Displayed up to 5000 trials. The value in this example is 0.22882 and 95% confidence interval around the mean of the (100000,1)-dimensional binomial vector (1 for default, 0 for non-default) with 100,000 simulations is $[0.226216, 0.231424]$.

Table B.8

Shot Noise Model Parameters for JPM for each 6-month Subperiod. "Period" indicates the period from which the data was used during estimation. For example, 2007.10-2008.4 denotes the period from the end of October, 2007 to the end of April, 2008. The rest should be understood in a similar way. Standard errors are displayed in parentheses.

Period	2007.10-2008.4		2007.11-2008.5		2007.12-2008.6		2008.1-2008.7	
μ	0.16935	(0.43178)	-5.15671	(4.92242)	-10.16382	(2.40578)	19.19598	(2.80149)
δ	0.00019	(0.00581)	0.14788	(0.06974)	0.17928	(0.09812)	0.17991	(0.04333)
$\mu_2\rho$	0.00126	(0.00076)	0.00264	(0.00149)	0.00271	(0.00178)	0.00658	(0.00180)
Z_0	126.35071	(5.31810)	389.37951	(13.25713)	677.06440	(21.69214)	-802.62820	(16.25556)
σ	0.05829	(0.00339)	0.04070	(0.02248)	0.04422	(0.02041)	0.00870	(0.46101)
k^2	0.00590		0.01323		0.01611		0.00803	
M	0.06588		0.06179		0.06387		0.08078	
Likelihood	-1275.47934		-1270.02487		-1273.64867		-1296.63644	

Period	2008.2-2008.8		2008.3-2008.9		2008.4-2008.10		2008.5-2008.11	
μ	18.70039	(2.51287)	25.29645	(2.87620)	0.96619	(0.43719)	-25.73898	(9.77697)
δ	0.24412	(0.11168)	0.28635	(0.11546)	14.79082	(3.04631)	0.11864	(0.11682)
$\mu_2\rho$	0.00695	(0.00158)	0.01513	(0.00606)	0.02939	(0.00850)	0.04364	(0.01861)
Z_0	-652.22744	(10.79559)	-554.95077	(6.32650)	-4.18307	(1.14583)	527.02632	(15.22436)
σ	0.00353	(0.34587)	0.00625	(0.07229)	0.01769	(0.01325)	0.04131	(0.66705)
k^2	0.00022		0.00031		0.04281		0.04710	
M	0.08337		0.12304		0.16865		0.20397	
Likelihood	-1300.28382		-1348.44774		-1375.78055		-1366.66309	

Period	2008.6-2008.12		2008.7-2009.1		2008.8-2009.2		2008.9-2009.3	
μ	-2.76816	(0.75406)	-1.02960	(0.29805)	-0.57461	(0.28468)	17.97071	(7.94993)
δ	2.45224	(0.51889)	6.10800	(1.31877)	20.05494	(4.33431)	0.17358	(0.10876)
$\mu_2\rho$	0.04488	(0.01246)	0.05336	(0.00351)	0.06589	(0.02506)	0.02668	(0.01198)
Z_0	21.51110	(5.25524)	7.11308	(2.32964)	3.92360	(1.01425)	-393.73101	(8.07602)
σ	0.01920	(0.01638)	0.01508	(0.00412)	0.00851	(0.00955)	0.10442	(0.04185)
k^2	0.03189		0.01671		0.00440		0.00900	
M	0.20927		0.22954		0.25626		0.18533	
Likelihood	-1391.06585		-1381.67195		-1367.81741		-1351.70137	

Period	2008.10-2009.4		2008.11-2009.5		2008.12-2009.6		2009.1-2009.7	
μ	1.37390	(2.80019)	0.96900	(0.99689)	0.05868	(0.13764)	0.11554	(0.07194)
δ	2.54631	(0.97409)	2.45263	(0.81466)	25.74678	(3.47023)	24.12009	(5.03220)
$\mu_2\rho$	0.00129	(0.00100)	0.00081	(0.00047)	0.00025	(0.00041)	0.00006	(0.00013)
Z_0	-59.68576	(7.84254)	-49.38486	(5.58259)	-24.02740	(5.16684)	-24.94549	(2.11805)
σ	0.18338	(0.07628)	0.16536	(0.11461)	0.13094	(0.01280)	0.10181	(0.01837)
k^2	0.03641		0.06573		0.05383		0.02393	
M	0.18001		0.16044		0.12820		0.10088	
Likelihood	-1335.12304		-1340.53716		-1320.65833		-1320.03700	

Table B.9

Shot Noise Model Parameters for Citi for each 6-month Subperiod. "Period" indicates the period from which the data was used during estimation. For example, 2007.10-2008.4 denotes the period from the end of October, 2007 to the end of April, 2008. The rest should be understood in a similar way. Standard errors are displayed in parentheses.

Period	2007.10-2008.4		2007.11-2008.5		2007.12-2008.6		2008.1-2008.7	
μ	-0.01481	(0.50430)	-0.12663	(0.25197)	-9.34337	(5.88169)	-0.41996	(0.10097)
δ	15.88077	(8.23615)	23.42112	(10.20578)	0.21140	(0.21454)	5.11785	(1.67285)
$\mu_2\rho$	0.00006	(0.00003)	0.00020	(0.00019)	0.00225	(0.00165)	0.00007	(0.00009)
Z_0	-80.49119	(3.03780)	-20.13731	(4.77693)	615.87192	(390.85580)	32.07261	(6.15784)
σ	0.05560	(0.00514)	0.05283	(0.00406)	0.02853	(0.08514)	0.05304	(0.01442)
k^2	0.20751		0.24140		0.11439		0.02126	
M	0.05247		0.04749		0.04633		0.05247	
Likelihood	-1281.55245		-1268.90658		-1263.24572		-1272.26501	

Period	2008.2-2008.8		2008.3-2008.9		2008.4-2008.10		2008.5-2008.11	
μ	-0.31684	(0.15745)	5.62609	(0.88356)	-0.29450	(0.15878)	-9.28147	(2.40997)
δ	12.71235	(1.23708)	0.12135	(0.02038)	14.38931	(6.71972)	0.17434	(0.24781)
$\mu_2\rho$	0.00004	(0.00008)	0.00659	(0.00151)	0.00113	(0.00339)	0.01505	(0.02226)
Z_0	23.08451	(7.48412)	-300.75735	(3.02128)	-4.72012	(1.95102)	250.30261	(9.04846)
σ	0.05506	(0.00531)	0.05208	(0.00506)	0.12582	(0.01985)	0.20110	(0.07373)
k^2	0.05239		0.17733		0.24171		0.16302	
M	0.05395		0.07576		0.11316		0.18860	
Likelihood	-1271.68354		-1308.30624		-1337.57901		-1334.62633	

Period	2008.6-2008.12		2008.7-2009.1		2008.8-2009.2		2008.9-2009.3	
μ	-0.91023	(0.67059)	-1.47071	(0.61863)	-1.47157	(0.85461)	53.89664	(22.20254)
δ	7.27340	(1.26724)	6.37209	(2.71264)	0.00014	(0.00045)	0.09717	(0.03635)
$\mu_2\rho$	0.00093	(0.00119)	0.00024	(0.00064)	0.13651	(0.08986)	0.15840	(0.09172)
Z_0	21.54567	(4.90313)	58.82569	(5.38875)	-19.38073	(9.07170)	-641.73389	(26.07933)
σ	0.22587	(0.05294)	0.28858	(0.06108)	0.23506	(0.03056)	0.24874	(0.16337)
k^2	0.06970		0.00382		0.00192		0.00055	
M	0.21979		0.28804		0.42913		0.46438	
Likelihood	-1342.95369		-1319.15642		-1290.43938		-1249.73621	

Period	2008.10-2009.4		2008.11-2009.5		2008.12-2009.6		2009.1-2009.7	
μ	46.63025	16.32559	21.32822	(6.74110)	0.23543	(0.38680)	0.12873	(0.17642)
δ	0.16543	(0.06531)	0.20249	(0.07077)	8.75045	(3.64768)	13.34913	(8.96509)
$\mu_2\rho$	0.15256	(0.05021)	0.02986	(0.01800)	0.00007	(0.00008)	0.00003	(0.00003)
Z_0	-436.39042	(15.63711)	-409.58304	(4.38073)	-99.49593	(5.42703)	-76.08406	(3.35488)
σ	0.11484	(0.02696)	0.04149	(0.01018)	0.13182	(0.03666)	0.06554	(0.01549)
k^2	0.00283		0.01257		0.01889		0.00745	
M	0.40121		0.17313		0.13092		0.06530	
Likelihood	-1205.51610		-1178.36287		-1146.79620		-1126.73272	

Table B.10

Shot Noise Model Parameters for BAC for each 6-month Subperiod. "Period" indicates the period from which the data was used during estimation. For example, 2007.10-2008.4 denotes the period from the end of October, 2007 to the end of April, 2008. The rest should be understood in a similar way. Standard errors are displayed in parentheses.

Period	2007.10-2008.4		2007.11-2008.5		2007.12-2008.6		2008.1-2008.7	
μ	-3.37983	(0.47615)	-3.55630	(1.33077)	-7.39615	(1.99630)	14.71927	(6.22488)
δ	0.23609	(0.13508)	0.24033	(0.17118)	0.22304	(0.13708)	0.10526	(0.02281)
$\mu_2\rho$	0.00329	(0.00108)	0.00226	(0.00115)	0.00318	(0.00159)	0.00539	(0.00492)
Z_0	180.72506	(5.00186)	222.83691	(6.93309)	400.40584	(31.58148)	-893.79524	(121.98668)
σ	0.02506	(0.02073)	0.03295	(0.02014)	0.01297	(0.60251)	0.03960	(0.08950)
k^2	0.01522		0.01719		0.04484		0.03650	
M	0.05972		0.05415		0.05511		0.07648	
Likelihood	-1286.15317		-1274.38596		-1274.23330		-1310.98785	

Period	2008.2-2008.8		2008.3-2008.9		2008.4-2008.10		2008.5-2008.11	
μ	11.60681	(2.95546)	16.43999	(9.47100)	0.14265	(0.10871)	-27.85228	(10.52505)
δ	0.14211	(0.11217)	0.16729	(0.33359)	10.60600	(2.28166)	0.06383	(0.05119)
$\mu_2\rho$	0.00458	(0.00385)	0.01933	(0.01709)	0.00010	(0.00031)	0.07456	(0.04061)
Z_0	-659.84760	(21.27672)	-419.49585	(17.34311)	-31.11603	(6.29688)	572.69955	(23.31740)
σ	0.06043	(0.03142)	0.03139	(0.00660)	0.21584	(0.10925)	0.03752	(0.01485)
k^2	0.02817		0.00031		0.00735		0.01016	
M	0.08284		0.14198		0.21520		0.27186	
Likelihood	-1315.24563		-1364.76216		-1393.52336		-1373.59448	

Period	2008.6-2008.12		2008.7-2009.1		2008.8-2009.2		2008.9-2009.3	
μ	-1.03986	(0.70420)	-0.75168	(0.40036)	-0.52689	(0.19876)	69.73410	(39.28902)
δ	6.84679	(2.65173)	10.50677	(2.16160)	0.00014	(0.00128)	0.14858	(0.06703)
$\mu_2\rho$	0.10298	(0.02470)	0.00073	(0.00356)	0.15073	(0.02633)	0.19479	(0.08298)
Z_0	4.09323	(2.28340)	17.31140	(6.76433)	-95.34258	(7.22529)	-599.81814	(20.77161)
σ	0.01857	(0.00795)	0.37177	(0.03363)	0.36854	(0.03308)	0.12892	(0.63276)
k^2	0.01336		0.01992		0.02136		0.00884	
M	0.31929		0.36893		0.49470		0.44801	
Likelihood	-1390.56558		-1377.49228		-1354.06623		-1326.79338	

Period	2008.10-2009.4		2008.11-2009.5		2008.12-2009.6		2009.1-2009.7	
μ	1.30978	(0.58653)	1.25057	(0.70362)	0.74751	(0.35939)	-1.18236	(0.42266)
δ	7.71079	(1.71445)	6.38245	(1.68945)	11.11703	(2.97884)	0.54900	(0.22048)
$\mu_2\rho$	0.00050	(0.00027)	0.00020	(0.00045)	0.00068	(0.00192)	0.00081	(0.00079)
Z_0	-92.04925	(5.83663)	-94.89346	(39.83479)	-20.05959	(8.81854)	124.88194	(8.73892)
σ	0.40124	(0.09052)	0.37155	(0.14852)	0.31051	(0.03315)	0.17637	(0.04986)
k^2	0.01104		0.00457		0.02673		0.01908	
M	0.39951		0.37086		0.30733		0.17473	
Likelihood	-1301.42445		-1307.62951		-1294.66381		-1296.57820	

Table B.11

GBM Model Parameters for JPM and Citi for each 6-month Subperiod."Period" indicates the period from which the data was used during estimation. For example, 2007.10-2008.4 denotes the period from the end of October, 2007 to the end of April, 2008. The rest should be understood in a similar way. Standard errors are displayed in parentheses.

JPM

Period	2007.10-2008.4		2007.11-2008.5		2007.12-2008.6		2008.1-2008.7	
μ	0.21319	(0.07546)	0.14392	(0.10206)	0.05234	(0.03572)	0.29755	(0.14505)
σ	0.06588	(0.01400)	0.06198	(0.00449)	0.06488	(0.00723)	0.08479	(0.02948)
Likelihood	-1275.47934		-1270.34259		-1275.25656		-1299.98972	
Period	2008.2-2008.8		2008.3-2008.9		2008.4-2008.10		2008.5-2008.11	
μ	0.52313	(0.15131)	0.77804	(0.24394)	0.62224	(0.15034)	0.52853	(0.04598)
σ	0.09099	(0.01889)	0.14147	(0.02683)	0.16879	(0.02603)	0.20780	(0.00315)
Likelihood	-1305.41168		-1354.22146		-1377.42600		-1367.12046	
Period	2008.6-2008.12		2008.7-2009.1		2008.8-2009.2		2008.9-2009.3	
μ	0.57975	(0.12866)	0.14285	(0.06586)	-0.10157	(0.07189)	-0.38367	(0.09895)
σ	0.23471	(0.05275)	0.24464	(0.04475)	0.25563	(0.03797)	0.18816	(0.03521)
Likelihood	-1393.84983		-1383.36441		-1369.39328		-1352.27919	
Period	2008.10-2009.4		2008.11-2009.5		2008.12-2009.6		2009.1-2009.7	
μ	-0.21601	(0.12280)	-0.06371	(0.06939)	-0.08846	(0.05967)	0.03647	(0.07179)
σ	0.18446	(0.03695)	0.16224	(0.03116)	0.13028	(0.05879)	0.10152	(0.05995)
Likelihood	-1336.06880		-1341.01779		-1321.51138		-1320.41296	

Citi

Period	2007.10-2008.4		2007.11-2008.5		2007.12-2008.6		2008.1-2008.7	
μ	-0.32850	(0.10201)	-0.24301	(0.04263)	-0.20397	(0.11480)	-0.22784	(0.02103)
σ	0.06029	(0.01176)	0.04964	(0.01066)	0.04828	(0.01968)	0.05308	(0.01252)
Likelihood	-1292.07017		-1272.20962		-1266.54887		-1272.91080	
Period	2008.2-2008.8		2008.3-2008.9		2008.4-2008.10		2008.5-2008.11	
μ	-0.23564	(0.12060)	-0.25415	(0.08126)	-0.37658	(0.07756)	-0.49848	(0.16085)
σ	0.05449	(0.02084)	0.07602	(0.01226)	0.11405	(0.01420)	0.18875	(0.01749)
Likelihood	-1272.22788		-1308.48679		-1337.73282		-1334.75718	
Period	2008.6-2008.12		2008.7-2009.1		2008.8-2009.2		2008.9-2009.3	
μ	-0.47486	(0.03022)	-0.84348	(0.26948)	-1.46722	(0.23355)	-1.35366	(0.40535)
σ	0.22184	(0.08199)	0.29100	(0.06274)	0.42913	(0.03707)	0.46784	(0.12217)
Likelihood	-1343.34248		-1319.51646		-1290.43938		-1249.99944	
Period	2008.10-2009.4		2008.11-2009.5		2008.12-2009.6		2009.1-2009.7	
μ	-0.90736	(0.23758)	-0.38332	(0.16801)	-0.32403	(0.13933)	-0.08896	(0.04624)
σ	0.41582	(0.16110)	0.18736	(0.07493)	0.15287	(0.03573)	0.07953	(0.01767)
Likelihood	-1206.24911		-1179.59436		-1148.82478		-1129.24235	

Table B.12

GBM Model Parameters for BAC for each 6-month Subperiod. “Period” indicates the period from which the data was used during estimation. For example, 2007.10-2008.4 denotes the period from the end of October, 2007 to the end of April, 2008. The rest should be understood in a similar way. Standard errors are displayed in parentheses.

BAC

Period	2007.10-2008.4		2007.11-2008.5		2007.12-2008.6		2008.1-2008.7	
μ	0.03705	(0.01377)	-0.03992	(0.01533)	-0.15159	(0.04318)	-0.05345	(0.03154)
σ	0.05991	(0.00833)	0.05437	(0.00204)	0.05584	(0.00498)	0.07723	(0.00804)
Likelihood	-1286.51132		-1274.86403		-1275.91633		-1311.79893	

Period	2008.2-2008.8		2008.3-2008.9		2008.4-2008.10		2008.5-2008.11	
μ	-0.00017	(0.06862)	0.08554	(0.05414)	-0.04136	(0.15008)	-0.19891	(0.01676)
σ	0.08378	(0.00562)	0.14429	(0.05251)	0.21607	(0.02688)	0.27256	(0.09420)
Likelihood	-1316.02280		-1365.50322		-1393.66006		-1373.67766	

Period	2008.6-2008.12		2008.7-2009.1		2008.8-2009.2		2008.9-2009.3	
μ	-0.09935	(0.04838)	-0.47891	(0.09599)	-0.78673	(0.23733)	-0.45968	(0.23744)
σ	0.31777	(0.01987)	0.36967	(0.11545)	0.49469	(0.06288)	0.47975	(0.09863)
Likelihood	-1391.26790		-1377.59291		-1354.06623		-1327.81849	

Period	2008.10-2009.4		2008.11-2009.5		2008.12-2009.6		2009.1-2009.7	
μ	-0.07077	(0.14157)	0.31080	(0.10390)	0.44122	(0.23059)	0.51233	(0.15189)
σ	0.42651	(0.09214)	0.37516	(0.02479)	0.30205	(0.08278)	0.17732	(0.03982)
Likelihood	-1302.57972		-1308.10172		-1294.84802		-1296.63112	

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R. J., 2015. Modeling financial contagion using mutually exciting jump processes. Journal of Financial Economics 117 (3), 585–606.

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

[3] Consigli, G., 2002. Tail estimation and mean-VaR portfolio selection in markets subject to financial instability. Journal of Banking and Finance 26 (7), 1355–1382.

[4] Dalla Valle, L., De Giuli, M. E., Tarantola, C., Manelli, C., 2016. Default probability estimation via pair copula constructions. European Journal of Operational Research 249 (1), 298–311.

[5] Dassios, A., Jang, J.-W., 2005. Kalman–bucy filtering for linear systems driven by the cox process with shot noise intensity and its application to the pricing of reinsurance contracts. Journal of Applied Probability 42 (1), 93–107.

[6] Dong, Y., Wang, G., Yuen, K. C., 2014. Bilateral counterparty risk valuation on a cds with a common shock model. Methodology and Computing in Applied Probability 16 (3), 643–673.

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

[8] Duan, J.-C., 2000. Correction: Maximum likelihood estimation using price data of the derivative contract. Mathematical Finance 10 (4), 461–462.

[9] Duan, J.-C., Gauthier, G., Simonato, J.-G., 2004. On the equivalence of the kmv and maximum likelihood methods for structural credit risk models. http://www.rmi.nus.edu.sg/duanjc/index_files/files/KMV-MLE_Nov%2030%202004.pdf, accessed: 2015-12-19.

[10] Elouerkhaoui, Y., 2007. Pricing and hedging in a dynamic credit model. International Journal of Theoretical and Applied Finance 10 (4), 703–731.

[11] Giesecke, K., 2004. Correlated default with incomplete information. Journal of Banking and Finance 28 (7), 1521–1545.

[12] Giesecke, K., Spiliopoulos, K., Sowers, R. B., 2013. Default clustering in large portfolios: Typical events. The Annals of Applied Probability 23 (1), 348–385.

[13] Giesecke, K., Spiliopoulos, K., Sowers, R. B., Sirignano, J. A., 2015. Large portfolio asymptotics for loss from default. Mathematical Finance 25 (1), 77–114.

[14] Herbertsson, A., Jang, J., Schmidt, T., 2011. Pricing basket default swaps in a tractable shot noise model. Statistics and Probability Letters 81 (8), 1196–1207.

[15] Karatzas, I., Shreve, S. E., 1998. Brownian Motion and Stochastic Calculus, 2nd Edition. Springer Science+Business Media New York.

[16] Kou, S., 2002. A jump-diffusion model for option pricing. Management Science 48 (8), 1086–1101.

[17] Kunisch, M., Uhrig-Homburg, M., 2008. Modeling simultaneous defaults: A top-down approach. Journal of Fixed Income 18 (1), 25–36.

[18] Landau, H. J., Shepp, L. A., 1970. On the supremum of a gaussian process. The Indian Journal of Statistics 32 (4), 369–378.

[19] Lehar, A., 2005. Measuring systemic risk: A risk management approach. Journal of Banking and Finance 29 (10), 2577–2603.

[20] Liang, X., Wang, G., 2012. On a reduced form credit risk model with common shock and regime switching. Insurance: Mathematics and Economics 51 (3), 567–575.

[21] Ma, Y., Xu, W., 2016. Structural credit risk modelling with hawkes jump diffusion processes. Journal of Computational and Applied Mathematics 303, 69–80.

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

[23] Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3 (1), 125–144.

[24] Mortensen, A., 2006. Semi-analytical valuation of basket credit derivatives in intensity-based models. Journal of Derivatives 13 (4), 8–26.

[25] Shreve, S. E., 2004. Stochastic Calculus for Finance II Continuous-Time Models. Springer-Verlag New York.

[26] Spiliopoulos, K., Sirignano, J. A., Giesecke, K., 2014. Fluctuation analysis for the loss from default. Stochastic Processes and their Applications 124 (7), 2322–2362.

[27] Steele, J. M., 2001. Stochastic Calculus and Financial Applications. Springer-Verlag New York.