

Research Article

Modeling Financial Intraday Jump Tail Contagion with High Frequency Data Using Mutually Exciting Hawkes Process

Chao Yu^D,¹ Jianxin Bi,¹ and Xujie Zhao²

¹School of Statistics, University of International Business and Economics, Beijing, China ²School of International Trade and Economics, University of International Business and Economics, Beijing, China

Correspondence should be addressed to Chao Yu; chaoyu@uibe.edu.cn

Received 17 November 2019; Revised 23 January 2020; Accepted 4 February 2020; Published 20 May 2020

Academic Editor: Ricardo López-Ruiz

Copyright © 2020 Chao Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Financial extreme jumps in asset price may propagate across stock markets and lead to the market-wide crashes, which severely threatens the stability of the financial system. In order to analyzing the contagion features of jump tail risk, this paper proposes a mutually exciting contagion model based on Hawkes process with intraday high frequency data. We use a simple two-stage method that first extracts the jump component nonparametrically from the high frequency data and then models the intraday jump tail using mutually exciting Hawkes process. Moreover, we take both the occurrence time and magnitude of jump into account in modeling the conditional intensity of Hawkes process. The proposed method is applied to the five-minute high frequency data of the Chinese stock market. The empirical results show that, for the two main Chinese stock markets, only background intensity is significant in the Shanghai stock market, while mutually exciting effect is significant in the Shanghai stock market have significant stimulation to the next occurrences of jump in the Shenzhen stock market. Furthermore, the proposed model performs very well in predicting the future jump tail events.

1. Introduction

It is well recognized that the financial asset returns are not normally distributed, but instead exhibit more slowly decaying and asymmetric tails. Numerous studies have shown that these fatter tails may be attributable to stochastic volatility and/or occasionally large absolute price changes, called "jumps" in the underlying asset price process. With the availability of reliable financial high frequency data over the last two decades, many closer research studies on the dynamics of financial asset prices have documented the presence of jumps; see [1–6]. While both components can account for the extreme tail behavior, they have different mechanisms and further have very different implications on pricing and risk management, as explored by Bollerslev and Todorov [7].

In contrast to the numerous studies on tail risk resulting from stochastic volatility, there is less work to study the jump tail risk. However, the recent financial crisis has further spurred the interest of studying the jump tail events. Bollerslev et al. [8] first used the extreme value theory to study the tail distribution of jumps and the dependence of jumps with high frequency data. Their research reveals a strong degree of tail dependence between the market-wide jumps and the systematic jumps in individual stocks. Aït-Sahalia et al. [9] also pointed out that the jump occurred in one market may propagate over time and spread to other markets as well. They first used the Hawkes process to model the self-exciting and mutually exciting features of jumps and then established a class of jump contagion asset price model. The proposed jump contagion model is then applied to the problems of derivative pricing [10] and portfolio investment [11]. In addition, many studies also consider the self-exciting features of jump in the problems of options pricing [12] and volatility risk premium modeling [13]. However, these econometric analysis of jump transmission is often conducted using low sampling frequencies, such as daily or lower. As many researchers pointed out, the real price paths of many financial assets change fast at the microstructure level, especially in periods of financial crisis, and moreover, the tail-type jump events tend to occur more frequently at the intraday frequency, rather than daily or

lower frequencies [14–16]. Hence, combined with the methods on jumps with high frequency data, this paper proposes a simple two-stage method to model the jump contagion at the intraday frequency. We first use the non-parametric method to identify the intraday jump series with high frequency data and then combine the mutually exciting Hawkes process and peaks-over-threshold (POT) approach to construct a mutually exciting jump tail contagion model. By using the identified jump series at high frequency, we can focus on studying its self-exciting and mutually exciting features at the intraday frequency.

In comparison with the numerous research studies related to jump risk on developed financial markets, there is limited consideration on emerging markets. It is well known that stock returns in emerging markets usually exhibit different characteristics, such as higher volatility, fatter tails, and more sudden shifts. Therefore, jumps in emerging markets may be more frequent, and the ways in which they transmit among different markets may differ from developed markets. With the enormous growth in the past decades, the Chinese stock market has become a more and more influential emerging market among the world stock markets. Hence, the proposed method is applied to the Chinese stock market. We explore the contagion mechanism of the jump tail between the two stock markets in China's mainland: the Shanghai stock market and Shenzhen stock market based on the proposed bivariate mutually exciting jump contagion model.

Our research makes three differences from the existing literature. First, by employing Hawkes process, what we really care about is not the continuous risk caused by volatility, but the contagion features of jump tails or to say extreme jump risks. Second, instead of extending the existing asset price models with mutually exciting jump component, we use a simple nonparametric method to separate jump component first and then focus on modeling the jump tail contagion, which allows us to study the contagion feature of jump at the intraday frequency. Third, in the modeling of the conditional intensity of Hawkes process, we consider the stimulation effects of both the jump magnitude and occurrence time.

The rest of this paper is organized as follows. Section 2 presents the extraction method of jump in asset price with high frequency data and proposes the mutually exciting jump tail contagion model. The maximum likelihood estimation and goodness of fit of the model are further discussed. Section 3 presents the empirical analysis. Section 4 concludes the paper.

2. Mutually Exciting Jump Tail Contagion Model

2.1. Extraction of Jump with High Frequency Data. Assume that the efficient logarithmic price $p_t^{(j)}$ of the *j*th asset defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ evolves as

$$dp_t^{(j)} = b_t^{(j)}dt + \sigma_t^{(j)}dW_t + dJ_t^{(j)}, \quad j = 1, 2, \dots, M, \quad (1)$$

where $W = (W_t)$ is an \mathcal{F} -adapted standard Brownian motion. The drift $b^{(j)} = (b_t^{(j)})$ and the volatility $\sigma^{(j)} = (\sigma_t^{(j)})$

are progressively measurable processes which guarantee that (1) has a unique, strong solution, which are adapted and right continuous with left limits (càdlàg) processes. $J^{(j)} = (J_t^{(j)})$ is a pure jump process.

From the theory of high frequency data analysis, we can extract the jump component by consistently estimating the quadratic variation (QV) and the integrated volatility (IV) of price process. Considering that there are T trading days, for simplicity, assume that on the time horizon [0, T], the observations are equally spaced. On each trading day, there are totally n discrete observations of log-return. Hence, the realized volatility (RV) and the bipower variation (BV) on day t are given as follows:

$$RV_{t}^{(j)} = \sum_{i=(t-1)n+1}^{(t-1)n+n} \left| \Delta_{i}^{n} p^{(j)} \right|^{2},$$

$$BV_{t}^{(j)} = \frac{\pi}{2} \sum_{i=(t-1)n+2}^{(t-1)n+n} \left| \Delta_{i}^{n} p^{(j)} \right| \left| \Delta_{i-1}^{n} p^{(j)} \right|, \quad t = 1, 2, \dots, T,$$
(2)

where $\Delta_i^n p^{(j)}$ represents the log-return of the *j*th asset on the time interval [(i-1)/n, i/n], i.e.,

$$\Delta_i^n p^{(j)} = p_{(i/n)}^{(j)} - p_{((i-1)/n)}^{(j)}.$$
(3)

Barndorff-Neilsen and Shephard [1, 2] demonstrated that *RV* and *BV* converge to quadratic variation (QV) and integrated volatility (IV), respectively, when the time interval is small enough.

To eliminate the intraday effects of high frequency volatility, we adopt the *TOD* statistics proposed by Bollerslev et al. [8]:

$$TOD_{i}^{(j)} = \frac{n\sum_{t=1}^{T} \left| \Delta_{i_{t}}^{n} p^{(j)} \right|^{2} I\left(\left| \Delta_{i_{t}}^{n} p^{(j)} \right| \le \tau \sqrt{BV_{t}^{(j)} \wedge RV_{t}^{(j)}} n^{-\omega} \right)}{\sum_{s=1}^{nT} \left| \Delta_{s}^{n} p^{(j)} \right|^{2} I\left(\left| \Delta_{s}^{n} p^{(j)} \right| \le \tau \sqrt{BV_{[s/n]}^{(j)} \wedge RV_{[s/n]}^{(j)}} n^{-\omega} \right)},$$
$$i_{t} = (t-1)n + i, i = 1, 2, \dots, n,$$
(4)

where I(·) is an indicator function, and τ and ω are constants satisfying $\tau > 0, \omega \in (0, 0.5)$. $TOD_i^{(j)}$ actually reflects the intensity of the intraday effect of the *i*th interval. $TOD_i^{(j)}$ larger than 1 indicates that there is a strong intraday effect and vice versa.

Next, we can use the threshold method of Mancini [17] combined with the intraday effect of volatility to identify the intervals where jump occurred on each trading day. The locations of intervals containing the jump component on day *t* for asset *j*, denoted by $I_t^{(j)}$, is estimated by

$$\widehat{I}_{t}^{(j)} = \left\{ i \in \{1, 2, \dots, n\} \colon \left| \Delta_{i_{t}}^{n} p^{(j)} \right| > \alpha_{i_{t}}^{(j)} n^{-\omega}, \quad i_{t} = (t-1)n + i \right\},$$

$$t = 1, 2, \dots, T,$$
(5)

where

$$\alpha_{i_t}^{(j)} = \tau \sqrt{\left(BV_t^{(j)} \land RV_t^{(j)}\right) \times TOD_i^{(j)}}, \quad i = 1, 2, \dots, n.$$
(6)

In the empirical analysis in Section 3, we set $\tau = 2.5$ and $\omega = 0.49$. Then, the jump component can be estimated as the returns on the intervals where jumps occurred, because if there occurs a jump, the jump will dominate the return. Hence, the identified jump series on day *t* for asset *j* can be obtained by

$$\left\{\widehat{J}_{i_t/n}^{(j)} = \Delta_{i_t}^n p^{(j)} \colon i_t = (t-1)n + i, i \in \widehat{I}_t^{(j)}\right\}, \quad t = 1, 2, \dots, T,$$
(7)

where $\hat{I}_{i_t/n}^{(j)}$ denotes the estimated jump at occurrence time i_t/n . Then, $\{i_t/n: i_t = (t-1)n + i, i \in \hat{I}_t^{(j)}\}$ is the arrival times of jumps on day *t*.

With the identified jump series on each day, we can put the results of all the trading days over [0, T] together in the chronological order to obtain one series of jump. We denote it by $\{J_{t_k^{(j)}}^{(j)}\}$, where $t_k^{(j)}$ is the time of jump and $J_{t_k^{(j)}}^{(j)}$ is the jump size at time $t_k^{(j)}$ for asset *j*. In the following, we will focus on modeling the contagion behavior of these jump events based on the Hawkes process and extreme value theory.

2.2. Bivariate Mutually Exciting Jump Contagion Model. Here, we focus on modeling the contagion behavior of the left jump tail since we usually care about the events that lead to extreme losses. We combine the peaks-overthreshold (POT) approach in the extreme value theory (EVT) and the mutually exciting Hawkes process to construct the contagion model. The Hawkes process is a counting process that models a sequence of 'arrivals' of events over time, where each arrival excites the process in the sense that the chance of a subsequent arrival is increased for some period after the initial arrival. Hence, Hawkes process is often used to model the clustering of events. Ogata [18] first introduced self-exciting Hawkes process in studying the earthquake occurrence. There are increasing applications of Hawkes process in finance, for instance, modeling of the risk [19, 20], modeling of the duration between trades [21], or the arrival process of buy and sell orders [22]. Aït-Sahalia et al. [9] use Hawkes process to capture the contagion of jumps in different regions of the world. In their modeling of jump contagion, the factor of interest that affects the conditional intensity of jump is the occurrence times of earlier events. However, in the financial high-frequency context, it is natural to also let intensity depend on the magnitudes. Hence, in this paper, we consider that the conditional intensity could be affected by both the magnitudes and occurrence times of previous jumps.

Consider the bivariate situation. Assume that $\{J_{t_k}^{(j)}\}$, j = 1, 2, is the identified negative jump series of two assets for all the trading days over the fixed time interval [0, T], according to the approach proposed in Section 2.1. Set $u_1 > 0$

and $u_2 > 0$ which are the threshold values of these two jump series, respectively. If $|J_{t_k^{(j)}}^{(j)}| \ge u_j$, j = 1, 2, then extreme jump occurs. Define the excess of extreme jump occurring at $t_k^{(j)}$ over the threshold u_j as $X_k^{(j)} = |J_{t_k^{(j)}}^{(j)}| - u_j$. Following the notation from [19, 20], let $\{T_k^{(j)}\}_{k\in\mathbb{Z}}$ denote the series of occurrence times of extreme jumps, and $\{X_k^{(j)}\}_{k\in\mathbb{Z}}$ the corresponding series of magnitudes of excess jumps. Then, $(T_k^{(j)}, X_k^{(j)})_{k\in\mathbb{Z}}$ forms two marked point processes, where $T_k^{(j)}$ defines the arrival times and $X_k^{(j)}$ the corresponding marks for the *j*th asset. Let \mathscr{H}_t be the sigma algebra generated by the two processes, that is, the entire history of the marks and their occurrence up to but not including the time *t*. Assume that there are total $n^{(1)}$ and $n^{(2)}$ extreme jumps over interval [0, T] for two assets. Use $(t_k^{(1)}, x_k^{(1)}), k = 1, 2, \ldots, n^{(1)}$, to denote the observed sequence of the first marked point

process over the period [0,T] and $(t_k^{(2)}, x_k^{(2)})$, $k = 1, 2, ..., n^{(2)}$, the observed sequence of the second marked point process. Then, according to the theory of marked point process (see [23] for more details), assume that the conditional ground intensity for the *j*th marked point process, i.e., the conditional intensity of the *j*th marginal point process of locations of extreme jumps is given as

$$\lambda_{j}(t \mid \mathscr{H}_{t}) = \lambda_{j} + \sum_{m=1}^{2} \sum_{\{k:t_{k}^{(m)} < t\}} \theta_{jm} g_{jm}(t - t_{k}^{(m)}, x_{k}^{(m)}), \quad j = 1, 2,$$
(8)

where $\lambda_j > 0$ and $\theta_{jm} > 0$. λ_j is the background intensity, and $g_{jm}(t - t_k^{(m)}, x_k^{(m)})$ is the exciting function, which shows how the jump events before the time *t* affect the intensity of the jumps occurring at time *t*. Furthermore, we assume that

$$g_{jm}(t - t_k^{(m)}, x_k^{(m)}) = \exp(\delta_{jm} x_k^{(m)} - \eta_{jm}(t - t_k^{(m)})), \quad (9)$$

where $\delta_{jm} > 0$ and $\eta_{jm} > 0$. This exciting function indicates that the impact of extreme jump on the intensity of the process is driven exponentially by the magnitude of the excess jump and decreases exponentially in relation to the distance from the jump event. Then, the conditional ground intensity is given by

$$\lambda_{j}(t \mid \mathscr{H}_{t}) = \lambda_{j} + \theta_{j1} \sum_{\{k:t_{k}^{(1)} < t\}} \exp\left[\delta_{j1}x_{k}^{(1)} - \eta_{j1}(t - t_{k}^{(1)})\right] \\ + \theta_{j2} \sum_{\{k:t_{k}^{(2)} < t\}} \exp\left[\delta_{j2}x_{k}^{(2)} - \eta_{j2}(t - t_{k}^{(2)})\right].$$
(10)

In this model, we believe that the conditional intensity of extreme jump's arrival for asset *j* can be decomposed into two parts: the background intensity part λ_j , and the response part which is triggered by the self-exciting and cross-exciting effects of previous extreme jumps from the asset itself and the other asset, respectively. Specifically, if an extreme jump occurs in the first asset, its own intensity of extreme jumps would increase immediately by $\theta_{11} \exp(\delta_{11} x_k^{(1)})$ and then decay exponentially with the parameter η_{11} . And the intensity of extreme jumps in the second asset would increase

instantaneously by $\theta_{21} \exp(\delta_{21} x_k^{(1)})$ and then decay exponentially with the parameter η_{21} . Likewise, if one extreme jump occurs in the second asset, its own intensity of extreme jumps would rise by $\theta_{22} \exp(\delta_{22} x_k^{(2)})$ and decay exponentially with the parameter η_{22} , while the intensity in the first asset would increase by $\theta_{12} \exp(\delta_{12} x_k^{(2)})$ and then decay exponentially with the parameter η_{12} . Therefore, we can

analyze the contagion behavior of the extreme jump risks based on this bivariate mutually exciting Hawkes process.

2.3. Maximum Likelihood Estimation. Assuming that the mark $\{X_k^{(j)}\}$ is independent of time $\{T_k^{(j)}\}$ and the past of the marked point process, the log-likelihood function of the bivariate marked point process $(T_k^{(j)}, X_k^{(j)})$ is given by

$$\log L = \sum_{j=1}^{2} \left(\sum_{\{k:t_{k}^{(j)} \le T\}} \log \lambda_{j}(t_{k}^{(j)}) - \int_{0}^{T} \lambda_{j}(t) dt + \sum_{\{k:t_{k}^{(j)} \le T\}} \log f_{j}(x_{k}^{(j)}) \right).$$
(11)

See the details in the study by Embrechts et al. [24] and Chavez-Demoulin et al. [20]. Let

$$\log L_{1} = \sum_{\{k:t_{k}^{(1)} \leq T\}} \log \lambda_{1}(t_{k}^{(1)}) - \int_{0}^{T} \lambda_{1}(t) dt,$$

$$\log L_{2} = \sum_{\{k:t_{k}^{(2)} \leq T\}} \log \lambda_{2}(t_{k}^{(2)}) - \int_{0}^{T} \lambda_{2}(t) dt,$$

$$\log L_{3} = \sum_{\{k:t_{k}^{(1)} \leq T\}} \log f_{1}(x_{k}^{(1)}),$$

$$\log L_{4} = \sum_{\{k:t_{k}^{(2)} \leq T\}} \log f_{2}(x_{k}^{(2)}),$$

(12)

where for j = 1, 2,

$$\begin{split} \int_{0}^{T} \lambda_{j}(t) dt &= \lambda_{j} T + \theta_{j1} \sum_{\{k:t_{k}^{(1)} \leq T\}} \int_{t_{k}^{(1)}}^{T} \exp\left(\delta_{j1} x_{k}^{(1)} - \eta_{j1} \left(t - t_{k}^{(1)}\right)\right) dt \\ &+ \theta_{j2} \sum_{\{k:t_{k}^{(2)} \leq T\}} \int_{t_{k}^{(2)}}^{T} \exp\left(\delta_{j2} x_{k}^{(2)} - \eta_{j2} \left(t - t_{k}^{(2)}\right)\right) dt \\ &= \lambda_{j} T + \frac{\theta_{j1}}{\eta_{j1}} \sum_{\{k:t_{k}^{(1)} \leq T\}} \exp\left(\delta_{j1} x_{k}^{(1)}\right) \\ &\times \left[1 - \exp\left(-\eta_{j1} \left(T - t_{k}^{(1)}\right)\right)\right] \\ &+ \frac{\theta_{j2}}{\eta_{j2}} \sum_{\{k:t_{k}^{(2)} \leq T\}} \exp\left(\delta_{j2} x_{k}^{(2)}\right) \\ &\times \left[1 - \exp\left(-\eta_{j2} \left(T - t_{k}^{(2)}\right)\right)\right]. \end{split}$$
(13)

 $f_j(x)$ is the probability density of $X_k^{(j)}$. According to the extreme value theory, if the threshold u_j is large enough, the distribution of $X_k^{(j)}$ can be approximated by generalized Pareto distribution (GPD). Hence, we can set

$$f_{j}\left(x;\xi_{j},\beta_{j}\right) = \begin{cases} \frac{1}{\beta_{j}}\left(1+\xi_{j}\frac{x}{\beta_{j}}\right)^{-1/\xi_{j}-1}, & \text{if } \xi_{j} \neq 0, \\ \frac{1}{\beta_{j}}\exp\left(-\frac{x}{\beta_{j}}\right), & \text{if } \xi_{j} = 0, \end{cases}$$
(14)

where $\beta_j > 0$ and ξ_j are the scale and shape parameters of the generalized Pareto distribution, respectively.

Then, the total log-likelihood function of the model can be written as a summation of $\log L_1$, $\log L_2$, $\log L_3$, and $\log L_4$. That is,

$$\log L = \log L_1 + \log L_2 + \log L_3 + \log L_4.$$
(15)

Hence, we can get the maximum likelihood estimation of the parameters for generalized Pareto distribution and the conditional ground intensity of Hawkes process separately. In particular, $\log L_1$ and $\log L_2$ can be used to estimate the parameters of the intensity, and $\log L_3$ and $\log L_4$ can be used to estimate the parameters of the generalized Pareto distribution.

2.4. Goodness of Fit. With the observations $(t_k^{(j)}, x_k^{(j)})$, $k \in \{1, 2, ..., n^{(j)}\}$, j = 1, 2, of the two marked point processes, define

$$t_{k}^{(j)*} = \int_{0}^{t_{k}^{(j)}} \lambda_{j}(s) \mathrm{d}s.$$
 (16)

Based on model (10), formula (16) can be given as

$$t_{k}^{(j)*} = \lambda_{j} t_{k}^{(j)} + \frac{\theta_{j1}}{\eta_{j1}} \sum_{\{m:t_{m}^{(1)} \leq t_{k}^{(j)}\}} \exp(\delta_{j1} x_{m}^{(1)})$$

$$\times \left[1 - \exp(-\eta_{j1}(t_{k}^{(j)} - t_{m}^{(1)}))\right]$$

$$+ \frac{\theta_{j2}}{\eta_{j2}} \sum_{\{m:t_{m}^{(2)} \leq t_{k}^{(j)}\}} \exp(\delta_{j2} x_{m}^{(2)})$$

$$\times \left[1 - \exp(-\eta_{j2}(t_{k}^{(j)} - t_{m}^{(2)}))\right].$$
(17)

According to the residual analysis of point process, the sequence $\{t_1^{(j)*}, t_2^{(j)*}, \ldots, t_{n^{(j)}}^{(j)*}\}$ forms a Poisson process with unit rate. Thus, the interarrival times $\{\tau_1, \tau_2, \ldots, \tau_{n^{(j)}-1}\} = \{t_2^{(j)*} - t_1^{(j)*}, t_3^{(j)*} - t_2^{(j)*}, \ldots, t_{n^{(j)}}^{(j)*} - t_{n^{(j)}-1}^{(j)*}\}$ should satisfy $\tau_i^{i.i.d.}$ exp (1). Therefore, it is feasible to use QQ plot using the exponential distribution to see how well the Hawkes process fits the observations. Alternatively, we can use the Kolmogorov–Smirnov test to test whether τ_i follows exponential distribution.

3. Empirical Analysis

3.1. Estimation of Mutually Exciting Jump Tail Contagion Model. We collect the tick-by-tick transaction data of the Shanghai (SH) composite index in the Shanghai stock market and Shenzhen (SZ) component index in Shenzhen stock market from the Chinese RESSET high-frequency database. The sample starts on January 4, 2006, and ends on December 31, 2013. There are totally 1932 valid trading days after excluding the holidays and the trading days with successive missing values over intervals with length of ten minutes or above. Then, we sample the tick-by-tick data at five-minute frequency by assigning the last observation in each five-minute interval as the price at five-minute frequency. Finally, we calculate the five-minute log-returns with five-minute prices for these two indexes. The following analysis is carried out by using these five-minute return data.

Firstly, we extract the jump components from the fiveminute returns by using the threshold method proposed in Section 2.1. Also, we are interested in the negative jumps. Then, we use the mean excess function to find an appropriate threshold, which can make the jump tails subject to the generalized Pareto distribution. Figure 1 shows the mean excess functions of the SH composite index and SZ component index, respectively. From the figure, we can obtain that both mean excess functions tend to be positively linear when the threshold is over 0.3. Therefore, we select 0.3 as the threshold with which we can identify the extreme jumps and obtain the excesses of extreme jumps for the two indexes.

Table 1 reports some summary statistics about the extreme jumps over threshold identified in the SH composite index and SZ component index, respectively.

With the identified extreme jump times and their marks, i.e., magnitudes of the excesses of extreme jumps for these two indexes, we next estimate the bivariate mutually exciting jump contagion model by the maximum likelihood estimation method. Table 2 reports the estimation results of two generalized Pareto distributions.

Figure 2 gives the QQ-plots of two fitted generalized Pareto distributions for the SH composite index and SZ component index, respectively, which show that the negative jump tails of both indexes are well fitted by the generalized Pareto distribution.

Table 3 reports the estimation results of parameters in the conditional ground intensity of Hawkes process.

Moreover, we use the bootstrap method to obtain the standard errors of parameter estimation by Monte Carlo simulation. In particular, we first simulate 1000 paths of the bivariate Hawkes process with the estimated values of parameters and then run the parameter estimation procedure with these 1000 bootstrap samples to get 1000 estimates of the parameters. Then, we can compute the standard error of each parameter with these 1000 estimated values. From the results of intensity process in the Shanghai stock market, we can see that only background intensity is significant, which shows that there are no significant selfexciting and mutually exciting effects in the Shanghai stock market.

However, from the results of the Shenzhen stock market, we can see that λ_2 , δ_{21} , and η_{21} are significant at 5% level, and θ_{21} is significant at 10% level, which shows that the jump intensity in the Shenzhen stock market are mainly affected by the background factors and jumps occurred in the Shanghai stock market as well. Both the locations and sizes of jumps occurred in the Shanghai stock market have significant stimulations to the extreme jumps in the Shenzhen stock market. These results can be attributed to the different characteristics of the two stock markets. The Shanghai stock market mainly consists of large-cap stocks such as financial stocks or blue chips, and the companies listed in Shanghai stock exchange are often the leading enterprises in their own industry, while the main components of the Shenzhen stock market are medium and small-cap stocks and growth enterprises market board. Therefore, the performance of Shenzhen component index tends to follow the changes in the Shanghai composite index.

With the estimation results, the simulated paths of the conditional ground intensities for the two indexes are given in Figure 3. The figure shows that the intensity of the SZ component index is more volatile than the SH composite index since there exists cross-exciting effect in the Shenzhen stock market apart from the influence of common factors. Figure 4 presents the QQ-plots of two fitted Hawkes process for the SH composite index and SZ component index, respectively. From the figure, it can be easily seen that the sample quantiles and the theoretical quantiles of exponential distribution almost lie in a line, which indicates that the observations are well fitted by the model.

3.2. Prediction. Here, we consider the problem of predicting the future jump events out of sample. Since our model is based on Hawkes point process, it is impossible to use it to predict the exact locations of future jumps. However, similar to the forecasting of earthquake, we can forecast the probability of the future jump's arrivals within a time period. Such kind of forecasting can be implemented by simulating the bivariate Hawkes process with the estimated intensity function repeatedly and then calculate the frequency of the arrivals within a given time interval. Then, we can calculate the frequencies of the arrivals within different time periods and finally get the empirical distribution of times for future jump arrivals. For instance, we can obtain the empirical distribution of days within which the next jump occurs or



FIGURE 1: Mean excess functions for the (a) SH composite index and (b) SZ component index.

| | Count | Mean | Std. dev. | Minimum | Maximum |
|----|-------|--------|-----------|---------|---------|
| SH | 946 | 0.8006 | 0.5947 | 0.3008 | 6.2527 |
| SZ | 948 | 0.8595 | 0.5801 | 0.3010 | 4.6786 |

TABLE 2: Estimation results for generalized Pareto distribution.

| | Scale <i>β</i> | Shape ξ |
|----|------------------------|-------------------|
| SH | 0.4257** (0.0210) | 0.1496** (0.0374) |
| SZ | 0.5434^{**} (0.0253) | 0.0286 (0.0334) |

Note. The values in parenthesis are the results of standard errors. ** The parameter is significant at 5% level.



FIGURE 2: QQ-plots of generalized Pareto distribution fitting for the (a) SH composite index and (b) SZ component index.

the next two jumps and other situations occur. Hence, with the estimated results in Section 3.1, we simulate 2000 paths of our model over the next month after December 31, 2013, i.e., January in 2014. We choose the following one month as the time interval of the simulation and prediction because this time period is long enough for the future occurrences of

TABLE 3: Estimation results for ground intensity process.

| Parameter | Value | Std. err. | Parameter | Value | Std. err. |
|---------------------|----------|-----------|---------------|-----------|-----------|
| λ_1 | 0.0102** | 0.0002 | λ_2 | 0.0101** | 0.0002 |
| θ_{11} | 0.0003 | 0.0012 | θ_{21} | 0.0025* | 0.0015 |
| δ_{11} | 0.1281** | 0.0387 | δ_{21} | 0.0962** | 0.0367 |
| η_{11} | 0.3631** | 0.0635 | η_{21} | 0.4728** | 0.0541 |
| $\dot{\theta}_{12}$ | 4.8e - 7 | 0.0015 | θ_{22} | 2.5e – 10 | 0.0012 |
| δ_{12} | 0.0579 | 0.0435 | δ_{22} | 0.2697** | 0.0544 |
| η_{12} | 0.6962** | 0.0496 | η_{22} | 0.5179** | 0.0567 |

Note. The results of standard errors are obtained by 1000 bootstrap simulations. * and ** denote that the parameters are significant at 10% and 5% level, respectively.



FIGURE 3: Paths of ground intensity processes of the (a) SH composite index and (b) SZ component index.



FIGURE 4: QQ-plots of Hawkes process fitting for the (a) SH composite index and (b) SZ component index.

jump events, which makes the prediction procedure described above and the checking of its performance feasible. Table 4 reports the frequencies of the first extreme jump occurring on the following days for the SH composite index and SZ component index based on simulation. Figure 5 shows the histograms.

TABLE 4: Frequencies of the first extreme jump occurrence on following days.

| | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 | Day 6 | Day 7 | Above 7 days |
|----|--------|--------|--------|--------|--------|--------|--------|--------------|
| SH | 0.3790 | 0.2405 | 0.1350 | 0.0865 | 0.0555 | 0.0365 | 0.0300 | 0.0370 |
| SZ | 0.3740 | 0.2320 | 0.1490 | 0.0935 | 0.0605 | 0.0410 | 0.0180 | 0.0320 |

Note. The values in the table are the frequencies of the fist extreme jump occurring on the following days. For instance, for the SH composite index, 0.3790 is the frequency of the first extreme jump occurring on the first day, and 0.2405 is the frequency of the first extreme jump occurring on the second day over the one-month prediction window.



FIGURE 5: Histogram of the next jump occurrence.

From Table 4, we can obtain that, for the SH composite index, the frequency that the next negative extreme jump occurs within one day after December 31, 2013, is 0.3790, within two days, 0.6195 (sum of the frequencies of the first two days) and within three days, 0.7545 (sum of the frequencies of the first three days). As for the SZ component index, the corresponding frequencies are 0.3740, 0.6060, and 0.7550, respectively. In summary, our prediction results show that in both stock markets the frequency that the next negative extreme jump occurs within three days is above 0.75, which implies that the next jump occurs within following three days with a relatively large probability.

To check the performance of our prediction, we use the intraday five-minute high frequency data from January 2, 2014, to January 31, 2014, and separate the negative extreme jumps for the two indexes to make a comparison with our prediction results. We find that in the Shanghai stock market, the first jump occurs at 10:10 a.m. on January 2, and in the Shenzhen stock market, the first jump occurs at 09:35 a.m. on January 3. Overall, the arrivals of the first jump are observed within two days for the two stock markets, which verifies our above prediction.

In order to make the idea of the prediction more clear, we proceed to report the frequency tables and histograms of the second extreme jump occurrence on following days for the SH composite index and SZ component index in Table 5 and Figure 6, respectively. From the results in Table 5, we can obtain that the frequencies that first two jumps occur in the SH composite index within 3 days and within 6 days are 0.4415 and 0.791, respectively. And the frequencies that first two jumps occur in the SZ component index within 3 days and within 6 days are 0.444 and 0.798, respectively. From the results of identified jumps with real data, we find that the second jump occurs at 9:35 a.m. on Januray 3 and 10:00 a.m. on January 3, respectively, for the SH composite index and SZ component index, which means that the arrivals of the next two jumps are observed within two days for the two indexes. The results are compatible with our prediction.

Next, we adopt the procedure described above repeatedly and use the rolling window method to make prediction over the months from February to December in 2014. The performances of total 11 predictions with rolling window are reported in Table 6 for the SH composite index and in Table 7 for the SZ component index. In both tables, we report the frequencies of the first jump occurring within 3 days and 7 days over the forecasting interval and the identified dates and times of the first extreme jump occurrence by using the actual data for comparison. From Tables 6 and 7, we can see that, for all the rolling window predictions, the frequencies of the first extreme jump occurrence in 3 days are around 75%, a relatively large probability, and the frequencies in 7 days are around 97%, a very high probability. Furthermore, the jump arrivals with relatively large probability in every rolling window

TABLE 5: Frequencies of the second extreme jump occurrence on following days.

| | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 | Day 6 | Day 7 | Day 8 | Day 9 | Above 9 days |
|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------------|
| SH | 0.0865 | 0.1775 | 0.1775 | 0.1420 | 0.1245 | 0.0830 | 0.0535 | 0.0520 | 0.0365 | 0.0670 |
| SZ | 0.0990 | 0.1635 | 0.1815 | 0.1520 | 0.1120 | 0.0900 | 0.0620 | 0.0465 | 0.0245 | 0.0690 |

Note. The values in the table are frequencies of the second extreme jump occurring on the following days. For instance, for the SH composite index, 0.0865 is the frequency of the second extreme jump occurring on the first day, and 0.1775 is the frequency of the second extreme jump occurring on the second day over the one-month prediction window.



FIGURE 6: Histogram of the second jump occurrence.

TABLE 6: Performances of the rolling window prediction of the SH composite index.

| Prediction no. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Dates | Feb. 7 | Mar. 4 | Apr. 3 | May 7 | Jun. 5 | Jul. 3 | Aug. 6 | Sep. 9 | Oct. 9 | Nov. 3 | Dec. 3 |
| Times | 09:35 | 09:35 | 14:00 | 09:35 | 09:35 | 09:35 | 09:35 | 09:50 | 11:00 | 14:05 | 13:20 |
| In 3 days | 0.7630 | 0.7665 | 0.7715 | 0.7605 | 0.7855 | 0.7440 | 0.7635 | 0.7640 | 0.7415 | 0.7605 | 0.7575 |
| In 7 days | 0.9700 | 0.9730 | 0.9710 | 0.9645 | 0.9705 | 0.9620 | 0.9715 | 0.9640 | 0.9575 | 0.9645 | 0.9625 |

Note. The rows named "Dates" and "Times" in the table are the identified dates and times of the first jump on each prediction window, and the rows named "In 3 days" and "In 7 days" are the results of the cumulative frequencies of the first jump occurring within 3 days and 7 days over each prediction window.

TABLE 7: Performances of the rolling window prediction of the SZ component index.

| Prediction no. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Dates | Feb. 7 | Mar. 4 | Apr. 3 | May 5 | Jun. 4 | Jul. 9 | Aug. 5 | Sep. 9 | Oct. 9 | Nov. 3 | Dec. 3 |
| Times | 09:35 | 09:35 | 14:00 | 09:50 | 09:35 | 14:05 | 10:20 | 09:50 | 11:00 | 14:05 | 13:20 |
| In 3 days | 0.7630 | 0.7665 | 0.7550 | 0.7655 | 0.7500 | 0.7525 | 0.7635 | 0.7565 | 0.7675 | 0.7460 | 0.7840 |
| In 7 days | 0.9650 | 0.9660 | 0.9615 | 0.9630 | 0.9620 | 0.9590 | 0.9635 | 0.9600 | 0.9715 | 0.9595 | 0.9675 |

Note. The meaning of the results in the table is the same as the ones in Table 6.

prediction are almost observed. In particular, the next jump occurs within 3 days over prediction window for all the predictions except that in the sixth rolling prediction, the next jump for the SZ component index occurs in next 7 days, in the seventh rolling prediction, the next jump for the SH composite index occurs in next 4 days, and in the eighth rolling prediction, the next jumps for both indexes occurs in next 6 days. The results further demonstrate our model's power.

4. Conclusions

In this paper, we focus on modeling the contagion feature of the intraday jump tail with financial high frequency data. We use a two-stage method that first extracts the intraday jump nonparametrically with high frequency data and then constructs the mutually exciting jump tail contagion model based on Hawkes process and peaks-over-threshold approach. In the modeling of the conditional intensity of the Hawkes process, we take into account the effects of both occurrence time and excess jump magnitude. We then discuss the maximum likelihood estimation and the goodness of fit of the model. Finally, we apply the proposed method to the real high frequency data in the Chinese stock market. We first separate the negative jump series from the five-minute high frequency data of the Shanghai composite index and Shenzhen component index, which are the two main market indexes in China and then apply the proposed bivariate mutually exciting jump contagion model to the identified jump series. The empirical results show that mutually exciting feature is significant in the Shenzhen stock market, while both self-exciting and mutually exciting features are not significant in the Shanghai stock market. In particular, the intensity of extreme jump occurrence in the Shanghai stock market is only significantly affected by the background factors, while the intensity in the Shenzhen stock market is significantly affected by the background factors and the extreme jump events in the Shanghai stock market as well. Both the occurrence times and magnitudes of jumps occurred in the Shanghai stock market have significant stimulations to the extreme jumps in the Shenzhen stock market. Furthermore, the prediction results of the future jump events based on simulation demonstrate our model's power.

Data Availability

The data that support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

Jianxin Bi's current address is as follows: School of Economics, Xiamen University, Xiamen, Fujian, P.R.China. The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported in part by the National Natural Science Foundation of China (NSFC) (71601048 and 11501104), National Social Science Found of China (NSSFC) (17CJY052), and the Fundamental Research Funds for the Central Universities in UIBE (CXTD9-07, CXTD10-10, and 13QD09).

References

- O. E. Barndorff-Nielsen and N. Shephard, "Power and bipower variation with stochastic volatility and jumps," *Journal of Financial Econometrics*, vol. 2, no. 1, pp. 1–37, 2004.
- [2] O. E. Barndorff-Nielsen and N. Shephard, "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, vol. 4, no. 1, pp. 1–30, 2006.
- [3] X. Huang and G. Tauchen, "The relative contribution of jumps to total price variance," *Journal of Financial Econometrics*, vol. 3, no. 4, pp. 456–499, 2005.
- [4] Y. Aït-Sahalia and J. Jacod, "Testing for jumps in a discretely observed process," *The Annals of Statistics*, vol. 37, no. 1, pp. 184–222, 2009.
- [5] S. S. Lee and J. Hannig, "Detecting jumps from Lévy jump diffusion processes," *Journal of Financial Economics*, vol. 96, no. 2, pp. 271–290, 2010.

- [6] S. S. Lee and P. A. Mykland, "Jumps in equilibrium prices and market microstructure noise," *Journal of Econometrics*, vol. 168, no. 2, pp. 396–406, 2012.
- [7] T. Bollerslev and V. Todorov, "Tails, fears, and risk premia," *The Journal of Finance*, vol. 66, no. 6, pp. 2165–2211, 2011.
- [8] T. Bollerslev, V. Todorov, and S. Z. Li, "Jump tails, extreme dependencies, and the distribution of stock returns," *Journal* of Econometrics, vol. 172, no. 2, pp. 307–324, 2013.
- [9] Y. Aït-Sahalia, J. Cacho-Diaz, and R. J. A. Laeven, "Modeling financial contagion using mutually exciting jump processes," *Journal of Financial Economics*, vol. 117, no. 3, pp. 585–606, 2015.
- [10] Y. Aït-Sahalia, R. J. A. Laeven, and L. Pelizzon, "Mutual excitation in eurozone sovereign CDS," *Journal of Econometrics*, vol. 183, no. 2, pp. 151–167, 2014.
- [11] Y. Aït-Sahalia and T. R. Hurd, "Portfolio choice in markets with contagion," *Journal of Financial Econometrics*, vol. 14, no. 1, pp. 1–28, 2015.
- [12] A. Fulop, J. Li, and J. Yu, "Self-exciting jumps, learning, and asset pricing implications," *Review of Financial Studies*, vol. 28, no. 3, pp. 876–912, 2015.
- [13] K. Chen and S. H. Poon, "Variance swap premium under stochastic volatility and self-exciting jumps," The University of Manchester, Manchester, England, Manchester Business School Working Paper, No. 634, 2013.
- [14] Y. Aït-Sahalia and J. Jacod, "Analyzing the spectrum of asset returns: jump and volatility components in high frequency data," *Journal of Economic Literature*, vol. 50, no. 4, pp. 1007–1050, 2012.
- [15] Y. Aït-Sahalia and D. Xiu, "Increased correlation among asset classes: are volatility or jumps to blame, or both?" *Journal of Econometrics*, vol. 194, no. 2, pp. 205–219, 2016.
- [16] M. Dungey, D. Erdemlioglu, M. Matei, and X. Yang, "Testing for mutually exciting jumps and financial flights in high frequency data," *Journal of Econometrics*, vol. 202, no. 1, pp. 18–44, 2018.
- [17] C. Mancini, "Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps," *Scandinavian Journal of Statistics*, vol. 36, no. 2, pp. 270–296, 2009.
- [18] Y. Ogata, "Statistical models for earthquake occurrences and residual analysis for point processes," *Journal of the American Statistical Association*, vol. 83, no. 401, pp. 9–27, 1988.
- [19] V. Chavez-Demoulin, A. C. Davison, and A. J. McNeil, "Estimating value-at-risk: a point process approach," *Quantitative Finance*, vol. 5, no. 2, pp. 227–234, 2005.
- [20] V. Chavez-Demoulin and J. A. McGill, "High-frequency financial data modeling using Hawkes processes," *Journal of Banking & Finance*, vol. 36, no. 12, pp. 3415–3426, 2012.
- [21] L. Bauwens and N. Hautsch, Handbook of Financial Time Series: Modelling Financial High Frequency Data Using Point Processes, Springer, Berlin, Germany, 2009.
- [22] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy, "Modelling microstructure noise with mutually exciting point processes," *Quantitative Finance*, vol. 13, no. 1, pp. 65–77, 2013.
- [23] D. Daley and D. Vere-Jones, An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods, Springer, New York, NY, USA, 2nd edition, 2003.
- [24] P. Embrechts, T. Liniger, and L. Lin, "Multivariate Hawkes processes: an application to financial data," *Journal of Applied Probability*, vol. 48, no. A, pp. 367–378, 2011.