Probability in the Engineering and Informational Sciences, **33**, 2019, 591–617.
doi:10.1017/S0269964818000463

ON BOUNCING GEOMETRIC BROWNIAN MOTIONS

Xin Liu
School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA
E-mail: xliu9@clemson.edu.

Vidyadhar G. Kulkarni and Qi Gong
Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC 27599, USA E-mail: vkulkarn@email.unc.edu; qgong@email.unc.edu

A pair of bouncing geometric Brownian motions (GBMs) is studied. The bouncing GBMs behave like GBMs except that, when they meet, they bounce off away from each other. The object of interest is the position process, which is defined as the position of the latest meeting point at each time. We study the distributions of the time and position of their meeting points, and show that the suitably scaled logarithmic position process converges weakly to a standard Brownian motion as the bounce size $\delta \to 0$. We also establish the convergence of the bouncing GBMs to mutually reflected GBMs as $\delta \to 0$. Finally, applying our model to limit order books, we derive a simple and effective prediction formula for trading prices.

Keywords: diffusion approximations, geometric brownian motions, inverse gaussian distributions, limit order books, mutually reflected brownian motions, normal inverse gaussian distributions, reflected brownian motions, renewal reward processes, scaling limits

1. INTRODUCTION

Consider two particles, A and B, moving on the real line. Initially, particle A is to the right of B. The motion of each particle follows a geometric Brownian motion (GBM) until they meet. Once they meet, particle A instantaneously jumps to the right by an amount dictated by a bouncing size parameter $\delta > 0$, and particle B moves to the left by the same amount at the same time. The two particles then continue their motions according to their GBMs until they meet again, and this process continues for ever. We define a continuous time position process $\{P_\delta(t); t \geq 0\}$ as follows: the state of the position process at time t is the position of the last meeting of the particles at or before time t. We study the structure of the position process, and show that a suitably scaled and centered logarithmic position process converges to a standard Brownian motion (BM) as $\delta \to 0$. The trajectories of the particles are referred to as the bouncing GBMs, denoted by $\{A_\delta(t); t \geq 0\}$ and $\{B_\delta(t); t \geq 0\}$. It is shown that the time between consecutive meetings of the bouncing GBMs is inverse Gaussian (IG) distributed, and the change between positions of consecutive meeting points follows a normal inverse Gaussian (NIG) distribution (see Definition 3.1 for the definitions of IG and NIG distributions). Using these distributions, assuming we can observe position

 0269-9648/18 $25.00

process, but not the original bouncing GBMs, we use the method of moments to estimate the parameters of the model and show that the estimators are consistent and asymptotically normal.

We also study the convergence of the bouncing GBMs as $\delta \to 0$ and show that their limits are two mutually reflected GBMs. Mutually reflected GBMs can be constructed from mutually reflected BMs (the detailed construction is given in Section 2). Mutually reflected BMs have been studied by Burdzy and Nualart [12], and a related model of mutually reflected Brownian balls have been studied by Saisho and Tanaka [30]. Of these two papers, the one by Burdzy and Nualart is more relevant to our model. They study two Brownian motions in which the lower one is reflected downward from the upper one. Thus the upper process is unperturbed by the lower process, while the lower process is pushed downward (by an appropriate reflection map) when it hits the upper process. We use a similar construction in our mutually reflected GBMs, except that in our case both processes reflect off of each other in opposite directions whenever they meet. We assume that the reflection is symmetric, which will be made precise in Section 2.

One application for this model arises in the study of *limit order books (LOBs)*. A LOB is a system that a financial market uses to record all the orders sent to the market from sellers and buyers. Essentially, there are three types of orders, as listed below, can be submitted to the LOB.

- *Limit orders*: A limit order is an order which buys or sells an asset at a specified price or better.
- *Market orders*: A market order is an order which buys or sells an asset at the best available price in the market.
- *Order cancellation*: An order cancellation is used to cancel a limit order in the LOB.

The order book organizes the orders by their prices and by their arrival times within each price. The *market bid price* is the highest price of the limit buy orders, and the *market ask price* is the lowest price of the limit sell order. The market ask price cannot be less than the market bid price. The difference between the market ask and bid prices is called the *spread*. When a market order arrives, or a limit buy (resp. sell) order arrives with price at least equal to the market ask (resp. bid) price, the spread momentarily dips to zero, and a trade occurs. The two matched orders are removed from the LOB immediately, and the spread increases to a wider size. (We ignore the order sizes and assume that all orders are of size one in this simplified discussion.) Clearly, the market ask price is always above the market bid price, and when they momentarily become equal, a trade occurs instantaneously, and they separate again. Between two consecutive trading instances, the market ask and bid prices fluctuate due to new arrivals, cancellations, and so on.

There is an extensive literature in statistics and probability on the study of LOBs. In particular, Markov models have been developed in Abergel and Jedidi [1], Bayraktar *et al.* [5], Cont and De Larrard [14], Cont and Larrard [15], Cont *et al.* [16], Horst and Kreher [19], Horst and Paulsen [20], Kruk [26], to name a few. In such models, point processes are used to model arrival processes of limit and market orders, and the market bid and ask prices are formulated as complex jump processes. To simplify such complexity, one tries to develop suitable approximate models. Brownian motion type approximations are established, for example, in Abergel and Jedidi [1], Bayer *et al.* [4], Bayraktar *et al.* [5], Cont and Larrard [15], and the law of large numbers is recently studied in Horst and Kreher [19], Horst and Paulsen [20].

It is clear that the stochastic evolution of the market ask and bid prices is a result of complex dynamics of the trader behavior and the market mechanism. However, we focus on the following three important features of a LOB: (i) just before a trade occurs, the spread momentarily becomes zero; (ii) right after a trade, the spread becomes positive; (iii) between two consecutive trades, the market ask and bid prices change stochastically. To capture these features, we model the market ask and bid prices as two bouncing GBMs A_δ and B_δ such that (i) a trade occurs when the two prices become equal to each other; (ii) right after the trade, the two prices will jump away from each other with a jump size δ; (iii) between two consecutive trades, the two prices change according to independent GBMs. Thus we can think of the market ask price as particle A and market bid price as particle B. The meeting times of this particle correspond to the trading times, and the position process corresponds to the trading price process, which records the most recent trading price. To the best of our knowledge, this is the first work to model the market ask and bid prices using GBMs. Some relevant work on modeling market bid and ask prices are Bayer *et al.* [4], Cont and Larrard [15], Horst and Paulsen [20], all of which model the market ask and bid prices as jump processes, and then study their scaling limits. In Cont and Larrard [15], a random step function is derived as the scaling limit of the market bid price, where the function jumps when a volume process of orders reaches a certain boundary. In Horst and Paulsen [20], it is shown that the scaling limits of the market ask and bid prices are coupled ordinary differential equations, and later on, in Bayer *et al.* [4], under a suitable scaling, the price processes follow coupled stochastic differential equations.

Our main result for the bouncing GBM model says that under a suitable scaling, the logarithmic trading price process converges to a standard Brownian motion as $\delta \to 0$. Using these asymptotics, we derive a simple and effective prediction formula for trading prices in Section 5. It is interesting to see that we get an asymptotic GBM model for the trading prices in the limit. The GBM model captures the intuition that the rates of returns over non-overlapping intervals are independent of each other, and has been extensively used to model stock prices since the breakthrough made by Black and Scholes [8], Merton [27]. Thus our model of bouncing GBMs provides another justification for the GBM model of trading prices. Another interesting observation is the logarithmic returns between consecutive trading times are NIG distributed. In fact, empirical studies show that NIG distributions provide an excellent fit for the logarithmic returns of assets (see Brander–Nielsen [9,10], Rydberg [29]), and such a model is proposed in Barndorff–Nielsen [2].

Our bouncing GBM model is also related to asset-liability management (ALM). BM and GBM models have been used for stochastic controls in ALM (see Decamps *et al.* [17] and references therein). The asset and liability values are modeled as two GBMs constrained in an allowable set $\mathcal{A} = \{(x_1, x_2) \in \mathbb{R}_+^2 : \lambda_1 x_2 \le x_1 \le \lambda_2 x_2\}$, where λ_1 and λ_2 are constants satisfying $0 < \lambda_1 < \lambda_2$. Roughly speaking, inside the allowable set $\mathcal{A}$, the asset and liability value processes evolve according to two GBMs, and when they reaches the boundary of $\mathcal{A}$, an ALM strategy steps in to keep them in $\mathcal{A}$. Consider the special case when $\lambda_2 = \infty$. We model the asset and liability value processes as bouncing GBMs $(A_\delta, \lambda_1^{-1} B_\delta)$, and they will always lie in $\mathcal{A}$. Thus the bouncing GBM model provides a simple static control policy for ALM in this special case. Under such a control policy, whenever the liability value is about to exceed the asset value, a fixed amount (a function of δ) of funds is provided to prevent it happening. The exact analysis of bouncing GBMs in our work can be used to analyze business performance, for example, the next time that liability value threatens to exceed the asset value, and the value level when it happens.

It will be seen that the logarithmic position process of the bouncing GBMs is a renewal reward process (see Lemma 3.6). Renewal reward processes are also called continuous time random walks (CTRWs) in literature related to physics and finance. CTRWs were first introduced by Montroll and Weiss [28]. The process-level limit theorems for coupled CTRWs are studied in Becker–kern *et al.* [6], where they assume that the meeting times and meeting points are "coupled", that is they are not independent. The asymptotic distributions of CTRWs are studied in Kotulski [25]. We note that the logarithmic position process in our model is a coupled CTRW. However, this process is not covered by the limit theorems in Becker–kern *et al.* [6], although a similar process is studied in Kotulski [25] but for a fixed time t.

The rest of the paper is organized as follows. In Section 2, we introduce our bouncing GBM model in detail, where we also define a pair of mutually reflected GBMs that is shown to be the limit of the bouncing GBMs as the bouncing size parameter δ approaches 0. All the main results are summarized in Section 3, including the distributions of times and positions of meeting points, and the asymptotic behaviors of the logarithmic position process. In Section 4, the estimators of the model parameters are derived using the method of moments. Section 5 is then devoted to the application of bouncing GBMs in a limit order book. In particular, we use an asymptotic GBM model obtained in Section 3 for trading prices, from which we derive a simple and effective forecasting formula. We also apply the formula to real data, and show that the estimated δ parameter is indeed very small, and hence the asymptotic results are applicable, and work very well over short time horizons.

We use the following notation. Let $(\Omega, \mathcal{F}, P)$ denote a complete probability space satisfying the usual conditions. All the random variables and stochastic processes are assumed to be defined on this space. The expectation under P will be denoted by E. Denote by $\mathbb{R}^K$ the K-dimensional Euclidean space. For $y \in \mathbb{R}^K$, let $|y| = \sqrt{\sum_{k=1}^K y_k^2}$. For a real number x, define $x^+ = \max\{x, 0\}$ and $x^- = \max\{0, -x\}$. Similarly, for a real function f defined on $[0, \infty)$, define $f^+(t) = \max\{0, f(t)\}$ and $f^-(t) = \max\{0, -f(t)\}$, $t \geq 0$. Denote by $C([0,\infty); \mathbb{R}^K)$ the space of continuous functions defined from $[0, \infty)$ to $\mathbb{R}^K$ with the uniform topology, and $D([0,\infty); \mathbb{R}^K)$ the space of RCLL (right continuous with left limits) functions defined from $[0, \infty)$ to $\mathbb{R}^K$ with the Skorohod J_1 topology. A stochastic process X with values in $\mathbb{R}^K$ will be regarded as a random variable with values in $D([0,\infty); \mathbb{R}^K)$. All stochastic processes in this work will have RCLL sample paths. Convergence in distribution of random variables X^n to X will be denoted as $X^n \Rightarrow X$. For a $\mathbb{R}^K$-valued random variables X, denote by $f_X(\cdot)$ its density function. If Y is another $\mathbb{R}^L$-valued random variable, we use $f_X(\cdot|Y = y)$ to denote the conditional density function of X given that $Y = y$.

2. BOUNCING GBMS

The bouncing GBMs are denoted by $\{(A_\delta(t), B_\delta(t)); t \geq 0\}$. Roughly speaking, A_δ and B_δ behave like independent GBMs except that when they are about to meet, they bounce off away from each other. More precisely, let $A_\delta(0)$ and $B_\delta(0)$ be two positive random variables satisfying $A_\delta(0) > B_\delta(0)$, and let δ be a strictly positive constant called the bouncing size parameter. For $t \geq 0$, we first define BMs X_a and X_b as follows. For $t \geq 0$,

$$X_a(t) = \mu_a t + \sigma_a W_a(t),$$
$$X_b(t) = \mu_b t + \sigma_b W_b(t),$$

where W_a, W_b are independent standard BMs independent of $A_\delta(0)$ and $B_\delta(0)$, and $\mu_a, \mu_b \in \mathbb{R}$ and $\sigma_a, \sigma_b \in (0, \infty)$ are the drift and volatility parameters. We assume that $\mu_a < \mu_b$. Define

$$D_\delta(t) = \ln(A_\delta(0)) + X_a(t) - \ln(B_\delta(0)) - X_b(t), \quad t \geq 0.$$

We note that $\{D_\delta(t); t \geq 0\}$ is a BM with drift $\mu_a - \mu_b < 0$, variance $\sigma_a^2 + \sigma_b^2$, and initial value $\ln(A_\delta(0)) - \ln(B_\delta(0)) > 0$. For $n \geq 1$, define the following stopping times: $T_{\delta,0} = 0$, and

$$T_{\delta,n} = \inf\{t \geq 0 : D_\delta(t) = -2(n-1)\delta\}. \quad \textbf{(1)}$$

It is clear that $E[T_{\delta,n}] < \infty$, $T_{\delta,n} \geq T_{\delta,n-1}$ and $T_{\delta,n} \to \infty$, almost surely, as $n \to \infty$.

For $t \geq 0$, define

$$A_\delta(t) = A_\delta(0)\exp\left\{X_a(t) + \sum_{n=1}^{\infty}(n-1)\delta 1_{\{t\in[T_{\delta,n-1},T_{\delta,n})\}}\right\}, \quad \textbf{(2)}$$

$$B_\delta(t) = B_\delta(0)\exp\left\{X_b(t) - \sum_{n=1}^{\infty}(n-1)\delta 1_{\{t\in[T_{\delta,n-1},T_{\delta,n})\}}\right\}. \quad \textbf{(3)}$$

We observe that the processes A_δ and B_δ have initial values $A_\delta(0)$ and $B_\delta(0)$, respectively, and for $t \in [0, T_{\delta,1})$,

$$A_\delta(t) = A_\delta(0)e^{X_a(t)} = e^{\ln(A_\delta(0))+X_a(t)}, \quad B_\delta(t) = B_\delta(0)e^{X_b(t)} = e^{\ln(B_\delta(0))+X_b(t)}.$$

From (1), the stopping time $T_{\delta,1}$ is the first time that $\ln(A_\delta(0)) + X_a$ and $\ln(B_\delta(0)) + X_b$ become equal, and so it is also the first time that A_δ and B_δ become equal. Define the position of the first meeting point to be

$$P_{\delta,1} = A_\delta(T_{\delta,1}-) = B_\delta(T_{\delta,1}-) = A_\delta(0)e^{X_a(T_{\delta,1})} = B_\delta(0)e^{X_b(T_{\delta,1})}.$$

According to the definitions in (2) and (3), right at the time $T_{\delta,1}$, A_δ and B_δ will separate in the following way.

$$A_\delta(T_{\delta,1}) = P_{\delta,1}e^{\delta} > P_{\delta,1}, \quad B_\delta(T_{\delta,1}) = P_{\delta,1}e^{-\delta} < P_{\delta,1}.$$

Starting from $T_{\delta,1}$, the processes A_δ and B_δ evolve again as two independent GBMs with initial values $P_{\delta,1}e^{\delta}$ and $P_{\delta,1}e^{-\delta}$. Their next meeting time is equal to the first time $A_\delta(0)e^{X_a(t)+\delta}$ and $B_\delta(0)e^{X_b(t)-\delta}$ meet, which by (1) is exactly $T_{\delta,2}$. Recursively, for $n \geq 1$, the stopping time $T_{\delta,n}$ will be the nth meeting time of A_δ and B_δ, and the position of the nth meeting point is defined as

$$P_{\delta,n} = A_\delta(T_{\delta,n}-) = B_\delta(T_{\delta,n}-), \quad \textbf{(4)}$$

and the bouncing GBMs at $T_{\delta,n}$ move to

$$A_\delta(T_{\delta,n}) = P_{\delta,n}e^{\delta} > P_{\delta,n}, \quad B_\delta(T_{\delta,n}) = P_{\delta,n}e^{-\delta} < P_{\delta,n}. \quad \textbf{(5)}$$

Right after $T_{\delta,n}$, the processes A_δ and B_δ evolve as two independent GBMs with initials $P_{\delta,n}e^{\delta}$ and $P_{\delta,n}e^{-\delta}$ until they meet again at $T_{\delta,n+1}$. The construction of $(T_{\delta,n}, P_{\delta,n})$ is illustrated in Figure 1, and the dynamics of a trajectory of (A_δ, B_δ) are shown in Figure 2.

In the following proposition, we identify the limit of (A_δ, B_δ) as $\delta \to 0$. Assume that $(A_\delta(0), B_\delta(0))$ converges to $(A(0), B(0))$, almost surely, as $\delta \to 0$, where $(A(0), B(0))$ is a positive bivariate random vector. We first define a pair of mutually reflected BMs (Y_a, Y_b) as follows. For $t \geq 0$, define

$$Y_a(t) = \ln(A(0)) + X_a(t) + \tfrac{1}{2}L(t), \quad \textbf{(6)}$$

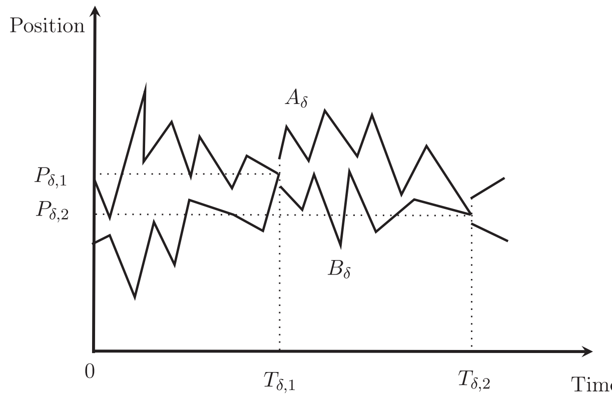

FIGURE 1. The construction of $T_{\delta,1}, P_{\delta,1}, T_{\delta,2}$, and $P_{\delta,2}$.

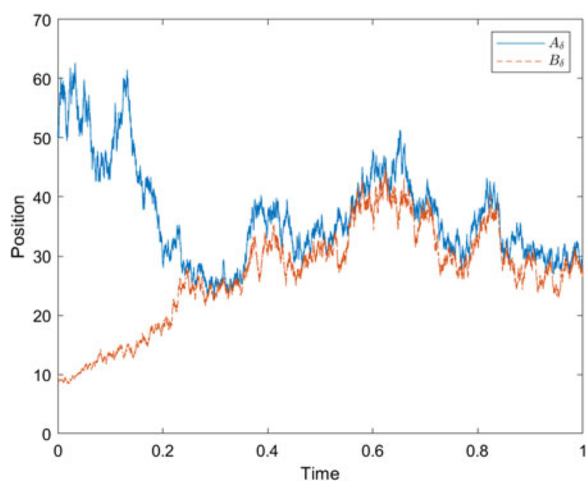

FIGURE 2. A sample path of the bouncing GBMs (A_δ, B_δ), where $A_\delta(0) = 50, B_\delta(0) = 10$, $\mu_a = -2, \mu_b = 2, \sigma_a = \sigma_b = 1$, and $\delta = 0.005$.

$$Y_b(t) = \ln(B(0)) + X_b(t) - \tfrac{1}{2}L(t), \tag{7}$$

where $\{L(t), t \geq 0\}$ is the unique continuous non-decreasing process such that

(i) $L(0) = 0$;

(ii) $L(t)$ can increase only when $Y_a(t) - Y_b(t) = 0$, that is,

$$\int_0^\infty 1_{\{Y_a(t)-Y_b(t)>0\}} \mathrm{d}L(t) = 0.$$

Noting that $\ln(A(0)) > \ln(B(0))$, with the process L satisfying (i) and (ii), it is clear that $Y_a(t) - Y_b(t) \geq 0$ for all $t \geq 0$. The existence and uniqueness of $\{L(t), t \geq 0\}$ follow from Skorohod lemma (see Karatzas and Shreve [23], Lemma 3.6.14). In fact, $L(t)$ has the following explicit formula

$$L(t) = \sup_{0 \leq s \leq t} D^-(s), \tag{8}$$

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

where

$$D(s) = \ln A(0) + X_a(s) - \ln B(0) - X_b(s),\ s \geq 0.$$

Roughly speaking, the processes $Y_a(t)$ and $Y_b(t)$ behave like two independent BMs when $Y_a(t) > Y_b(t)$, and whenever they meet, the process $Y_a(t)$ will be pushed up, while $Y_b(t)$ will be pushed down, so that $Y_a(t) \geq Y_b(t)$ for all $t \geq 0$. Here we assume the pushing effect for $Y_a(t)$ and $Y_b(t)$ are the same, and thus we have 1/2 before the regulator process $L(t)$ in both (6) and (7).

Finally, $A(t)$ and $B(t)$ are defined as

$$A(t) = e^{Y_a(t)}, \tag{9}$$

$$B(t) = e^{Y_b(t)}. \tag{10}$$

Thus $A(t)$ and $B(t)$ behave like two independent GBMs when $A(t) > B(t)$, and whenever they become equal, they will be pushed away from each other such that $A(t) \geq B(t)$ for all $t \geq 0$.

Proposition 2.1:

(i) *Assume that* $(A_\delta(0), B_\delta(0)) = (A(0), B(0))$. *Then for* $\delta > 0$ *and* $n \in \mathbb{N}$,

$$A(T_{\delta,n}) = B(T_{\delta,n}).$$

(ii) *Assume that* $(A_\delta(0), B_\delta(0)) = (A(0), B(0))$. *Then for* $\delta > 0$ *and* $t \geq 0$,

$$A_\delta(t) \geq A(t),\ \textit{and}\ B_\delta(t) \leq B(t).$$

(iii) *For* $t \geq 0$, *almost surely,*

$$\sup_{0\leq s\leq t} \frac{A_\delta(t)}{A(t)} \to 1, \quad \textit{and} \quad \sup_{0\leq s\leq t} \frac{B(t)}{B_\delta(t)} \to 1, \quad \textit{as}\ \delta \to 0.$$

Proof: In (i) and (ii), under the condition that $(A_\delta(0), B_\delta(0)) = (A(0), B(0))$, we have $D_\delta(t) = D(t)$ for $t \geq 0$. For (i), we note that

$$\ln(A(t)) - \ln(B(t)) = \ln(A(0)) + X_a(t) + \tfrac{1}{2}L(t) - \ln(B(0)) - X_b(t) + \tfrac{1}{2}L(t) = D(t) + L(t).$$

Thus it suffices to show that

$$D(T_{\delta,n}) = -L(T_{\delta,n}).$$

Now observe that $T_{\delta,0} = 0$ and $T_{\delta,n} = \inf\{t \geq 0 : D(t) = -2(n-1)\delta\}$, which yields that $D(T_{\delta,n}) = -2(n-1)\delta$, and $D(t) > -2(n-1)\delta$ for $t \in [0, T_{\delta,n})$. Thus from (8), $L(T_{n,\delta}) = 2(n-1)\delta$, and so (i) follows. To show (ii), we note that for $t \in [T_{\delta,n-1}, T_{\delta,n})$,

$D(t) > -2(n-1)\delta$, and at the time $T_{\delta,n-1}$, $D(T_{\delta,n-1}) = -2(n-2)\delta$. It follows that for $t \in [T_{\delta,n-1}, T_{\delta,n})$,

$$L(t) \equiv \sup_{0\le s\le t} D^-(s) \in [2(n-2)\delta,\ 2(n-1)\delta), \tag{11}$$

from which we have

$$\frac{A_\delta(t)}{A(t)} = \sum_{n=1}^{\infty} 1_{t\in[T_{\delta,n-1},T_{\delta,n})} \exp\left\{(n-1)\delta - \frac{1}{2}L(t)\right\} \in (1, e^\delta],$$

and

$$\frac{B_\delta(t)}{B(t)} = \sum_{n=1}^{\infty} 1_{t\in[T_{\delta,n-1},T_{\delta,n})} \exp\left\{-(n-1)\delta + \frac{1}{2}L(t)\right\} \in [e^\delta, 1).$$

For (iii), define a pair of mutually reflected GBMs $(\tilde{A}_\delta, \tilde{B}_\delta)$ with initial $(A_\delta(0), B_\delta(0))$ as follows.

$$\tilde{A}_\delta(t) = A_\delta(0) \exp\left\{X_a(t) + \frac{1}{2}L_\delta(t)\right\},$$
$$\tilde{B}_\delta(t) = B_\delta(0) \exp\left\{Y_a(t) - \frac{1}{2}L_\delta(t)\right\},$$

where similar to (8),

$$L_\delta(t) = \sup_{0\le s\le t} D_\delta^-(s).$$

Similar to the proof of (ii), we have

$$\frac{A_\delta(t)}{\tilde{A}_\delta(t)} = \sum_{n=1}^{\infty} 1_{t\in[T_{\delta,n-1},T_{\delta,n})} \exp\left\{(n-1)\delta - \frac{1}{2}L_\delta(t)\right\} \in (1, e^\delta],$$

which yields

$$\sup_{0\le s\le t} \frac{A_\delta(s)}{\tilde{A}_\delta(s)} \to 1.$$

Now consider the functional, which defines L and L_δ. Denote it by Φ, and then for $t \ge 0$,

$$L(t) = \Phi(D(\cdot))(t),$$

and

$$L_\delta(t) = \Phi(D_\delta(\cdot))(t).$$

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

Next we note that $\Phi : C([0,\infty);\mathbb{R}) \to C([0,\infty);\mathbb{R})$ is Lipschitz continuous (cf. Dupuis and Ishii [18]), that is, there exists a constant $c \in (0,\infty)$ such that for $x_1, x_2 \in C([0,\infty);\mathbb{R})$ and $t \geq 0$,

$$\sup_{0\leq s\leq t} |\Phi(x_1(s)) - \Phi(x_2(s))| \leq c \sup_{0\leq s\leq t} |x_1(s) - x_2(s)|.$$

As a result, we have $\sup_{0\leq s\leq t} |\tilde{A}_\delta(s) - A(s)| \to 0$ almost surely. It follows that as $\delta \to 0$,

$$\begin{aligned}\sup_{0\leq s\leq t} \frac{A_\delta(s)}{A(s)} &\leq \sup_{0\leq s\leq t} \frac{A_\delta(s)}{\tilde{A}_\delta(s)} \cdot \sup_{0\leq s\leq t} \left(\frac{\tilde{A}_\delta(s) - A(s)}{A(s)} + 1 \right) \\ &\leq \sup_{0\leq s\leq t} \frac{A_\delta(s)}{\tilde{A}_\delta(s)} \cdot \left(\frac{\sup_{0\leq s\leq t} |\tilde{A}_\delta(s) - A(s)|}{\inf_{0\leq s\leq t} A(s)} + 1 \right) \to 1,\end{aligned}$$

and

$$\begin{aligned}\sup_{0\leq s\leq t} \frac{A_\delta(s)}{A(s)} &\geq \sup_{0\leq s\leq t} \frac{A_\delta(s)}{\tilde{A}_\delta(s)} \cdot \inf_{0\leq s\leq t} \left(\frac{\tilde{A}_\delta(s) - A(s)}{A(s)} + 1 \right) \\ &\geq \sup_{0\leq s\leq t} \frac{A_\delta(s)}{\tilde{A}_\delta(s)} \cdot \left(\frac{\inf_{0\leq s\leq t} |\tilde{A}_\delta(s) - A(s)|}{\sup_{0\leq s\leq t} A(s)} + 1 \right) \to 1,\end{aligned}$$

which yields that

$$\sup_{0\leq s\leq t} \frac{A_\delta(s)}{A(s)} \to 1, \text{ as } \delta \to 0.$$

Similarly, it can be shown that for $t \geq 0$,

$$\sup_{0\leq s\leq t} \frac{B(s)}{B_\delta(s)} \to 1, \quad \text{as } \delta \to 0.$$

■

3. MAIN RESULTS

The object of interest is the position of the latest meeting point. To formulate it, we let

$$\begin{aligned}&U_{\delta,n+1} = \ln(P_{\delta,n+1}/P_{\delta,n}), \quad V_{\delta,n+1} = T_{\delta,n+1} - T_{\delta,n}, \ n \geq 1, \\ &U_{\delta,1} = \ln P_{\delta,1}, \quad V_{\delta,1} = T_{\delta,1}.\end{aligned}$$

Define for $t \geq 0$,

$$N_\delta(t) = \max\{n \geq 0 : T_{\delta,n} \leq t\}, \tag{12}$$

which gives the number of times A_δ and B_δ meet each other up to time t. Now the position of the latest meeting point at time t can be formulated as

$$P_\delta(t) = P_{\delta,N_\delta(t)}, \quad \text{for } t \geq T_{\delta,1}. \tag{13}$$

For $t \geq T_{\delta,1}$, let $Z_\delta(t) = \ln(P_\delta(t))$, and so $Z_\delta(t) = \sum_{n=1}^{N_\delta(t)} U_{\delta,n}$. When $0 \leq t < T_{\delta,1}$, we simply let $Z_\delta(t) = 0$. Thus we have

$$Z_\delta(t) = \sum_{n=1}^{N_\delta(t)} U_{\delta,n}, \tag{14}$$

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

with the convention that $\sum_{n=1}^{0} U_{\delta,n} = 0$. Our goal is to establish a scaling limit theorem for Z as $\delta \to 0$, and develop an asymptotic model for real financial data.

We present the main results in the rest of this section. In particular, it is shown that $(U_{\delta,n}, V_{\delta,n}), n \geq 2$, are i.i.d. random variables (see Lemma 3.2), and $U_{\delta,n}$ follows a NIG distribution and $V_{\delta,n}$ is IG distributed (see Corollary 3.3). Using these results, it is clear that $\{Z_\delta(t), t \geq 0\}$ is a renewal reward process, and the scaling limit theorem is established in Theorem 3.8.

3.1. Distribution of $(U_{\delta,n}, V_{\delta,n})$

We first derive the joint distribution of $(U_{\delta,n}, V_{\delta,n})$ for each $n \geq 1$. To derive the marginal distributions of V_n and U_n, we introduce the following definitions of IG and NIG distributions (cf. Sephardi [31]).

Definition 3.1:

(i) *An IG distribution with parameters a_1 and a_2 has a density function*

$$f(x; a_1, a_2) = \frac{a_1}{\sqrt{2\pi x^3}} \exp\left\{ -\frac{(a_1 - a_2 x)^2}{2x} \right\}, \ x > 0,$$

which is usually denoted by $\mathrm{IG}(a_1, a_2)$.

(ii) *A random variable Y follows a NIG distribution with parameters $\bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta}$ with notation* $\mathrm{NIG}(\bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta})$ *if*

$$Y|X = x \sim N(\bar{\mu} + \bar{\beta}x, x), \ \text{and } X \sim \mathrm{IG}(\bar{\delta}, \sqrt{\bar{\alpha}^2 - \bar{\beta}^2}).$$

The density function of Y is given as

$$f(y; \bar{\alpha}, \bar{\beta}, \bar{\mu}, \bar{\delta}) = \frac{\bar{\alpha}}{\pi \bar{\delta}} \exp\left\{ \sqrt{\bar{\alpha}^2 - \bar{\beta}^2} + \frac{\bar{\beta}}{\bar{\delta}}(y - \bar{\mu}) \right\} \frac{K_1\left(\bar{\alpha}\sqrt{1 + (y - \bar{\mu}/\bar{\delta})^2}\right)}{\sqrt{1 + (y - \bar{\mu}/\bar{\delta})^2}},$$

where $K_1(z) = 1/2 \int_0^\infty \mathrm{e}^{-z(t+t^{-1})/2} \mathrm{d}t$ is the modified Bessel function of the third kind with index 1.

The joint distribution of $(U_{\delta,n}, V_{\delta,n})$ is given in the following lemma. Note that $(U_{\delta,1}, V_{\delta,1})$ will not depend on δ if the initial values $A_\delta(0)$ and $B_\delta(0)$ are independent of δ.

Lemma 3.2:

(i) *Assume $A_\delta(0) = e^\alpha, B_\delta(0) = e^\beta$, and $\alpha > \beta$. Then for $t \geq 0$ and $x \in \mathbb{R}$, the joint probability density function (PDF) of $(U_{\delta,1}, V_{\delta,1})$ is given by*

$$f_{(U_{\delta,1}, V_{\delta,1})}(x,t) = \frac{\alpha - \beta}{2\pi t^2 \sigma_a \sigma_b} \exp\left\{ -\frac{[\sigma_b/\sigma_a(x - \alpha - \mu_a t) + \sigma_a/\sigma_b(x - \beta - \mu_b t)]^2 + [\alpha - \beta - (\mu_b - \mu_a)t]^2}{2(\sigma_a^2 + \sigma_b^2)t} \right\}. \quad (15)$$

In particular, $V_{\delta,1}$ follows an IG distribution with the following density function

$$f_{V_{\delta,1}}(t) = \frac{\alpha-\beta}{\sqrt{2\pi(\sigma_a^2+\sigma_b^2)t^3}} \exp\left\{-\frac{[\alpha-\beta-(\mu_b-\mu_a)t]^2}{2(\sigma_a^2+\sigma_b^2)t}\right\}, \ t \geq 0, \tag{16}$$

and given $V_{\delta,1}=t$, $U_{\delta,1}$ is normal distributed with mean

$$\frac{(\sigma_b^2\alpha+\sigma_a^2\beta)+(\sigma_b^2\mu_a+\sigma_a^2\mu_b)t}{\sigma_a^2+\sigma_b^2}$$

and variance

$$\frac{\sigma_a^2\sigma_b^2 t}{\sigma_a^2+\sigma_b^2}.$$

(ii) *The sequence $(U_{\delta,n}, V_{\delta,n})_{n\geq 2}$ is an i.i.d. sequence, which is independent of $(U_{\delta,1}, V_{\delta,1})$ and has the same distribution as described in* (i) *with $\alpha=\delta$ and $\beta=-\delta$.*

PROOF: For $t \geq 0$, let

$$X(t) = \begin{pmatrix} X_a(t) \\ X_b(t) \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \mu_a t \\ \mu_b t \end{pmatrix} + \begin{pmatrix} \sigma_a W_a(t) \\ \sigma_b W_b(t) \end{pmatrix}.$$

Then $V_{\delta,1}$ and $U_{\delta,1}$ are the time and position of the first meeting point of X_a and X_b. Let $\theta = \arctan(\sigma_a\sigma_b^{-1})$, and define

$$M = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \sigma_a^{-1} & 0 \\ 0 & \sigma_b^{-1} \end{pmatrix} = \frac{1}{\sqrt{\sigma_a^2+\sigma_b^2}} \begin{pmatrix} \sigma_b/\sigma_a & \sigma_a/\sigma_b \\ -1 & 1 \end{pmatrix}, \tag{17}$$

where the last equality follows from the identities that

$$\cos\theta = \frac{\sigma_b}{\sqrt{\sigma_a^2+\sigma_b^2}}, \ \sin\theta = \frac{\sigma_a}{\sqrt{\sigma_a^2+\sigma_b^2}}. \tag{18}$$

Now for $t \geq 0$, define

$$\begin{aligned}
\check{X}(t) &= MX(t) \\
&= \begin{pmatrix} \alpha\sigma_a^{-1}\cos\theta + \beta\sigma_b^{-1}\sin\theta & -\alpha\sigma_a^{-1}\sin\theta + \beta\sigma_b^{-1}\cos\theta \end{pmatrix} \\
&+ \begin{pmatrix} \mu_a\sigma_a^{-1}\cos\theta + \mu_b\sigma_b^{-1}\sin\theta & -\mu_a\sigma_a^{-1}\sin\theta + \mu_b\sigma_b^{-1}\cos\theta \end{pmatrix} t \\
&+ \begin{pmatrix} W_a(t)\cos\theta + W_b(t)\sin\theta \\ -W_a(t)\sin\theta + W_b(t)\cos\theta \end{pmatrix}.
\end{aligned}$$

Let

$$\begin{aligned}
\check{a} &= \alpha\sigma_a^{-1}\cos\theta + \beta\sigma_b^{-1}\sin\theta, \\
\check{b} &= -\alpha\sigma_a^{-1}\sin\theta + \beta\sigma_b^{-1}\cos\theta, \\
\check{\mu}_a &= \mu_a\sigma_a^{-1}\cos\theta + \mu_b\sigma_b^{-1}\sin\theta, \\
\check{\mu}_b &= -\mu_a\sigma_a^{-1}\sin\theta + \mu_b\sigma_b^{-1}\cos\theta, \\
\check{B}_a(t) &= W_a(t)\cos\theta + W_b(t)\sin\theta, \\
\check{B}_b(t) &= -W_a(t)\sin\theta + W_b(t)\cos\theta.
\end{aligned}$$

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

In particular, $\check{B}_a$ and $\check{B}_b$ are independent standard BMs. Next from (17), we note that the second component of $\check{X}(t)$ is equal to $(-X_a(t)+X_b(t))/\sqrt{\sigma_a^2+\sigma_b^2}$, which yields that

$$\begin{aligned} V_{\delta,1} &= \inf\{t \geq 0 : X_a(t) = X_b(t)\} \\ &= \inf\left\{t \geq 0 : \check{X}(t) \in \{(x,y) : y = 0\}\right\} \\ &= \inf\{t \geq 0 : \check{B}_b(t) + \check{\mu}_b t = -\check{b}\}. \end{aligned}$$

Using Girsanov theorem, and from (5.12) in Karatzas and Shreve [23], Chapter 3.5.C, the density function of $V_{\delta,1}$ is given by

$$\begin{aligned} f_{V_{\delta,1}}(t) &= \frac{|\check{b}|}{\sqrt{2\pi t^3}} \exp\left\{-\frac{(-\check{b}-\check{\mu}_b t)^2}{2t}\right\} dt \\ &= \frac{\alpha-\beta}{\sqrt{2\pi(\sigma_a^2+\sigma_b^2)t^3}} \exp\left\{-\frac{[\alpha-\beta-(\mu_b-\mu_a)t]^2}{2(\sigma_a^2+\sigma_b^2)t}\right\} dt,\ t \geq 0. \end{aligned}$$

We next note that

$$U_{\delta,1} = X_a(V_{\delta,1}) = X_b(V_{\delta,1}) = \frac{\check{a} + \check{\mu}_a V_{\delta,1} + \check{B}_a(V_{\delta,1})}{\sigma_a^{-1}\cos\theta + \sigma_b^{-1}\sin\theta},$$

and $\check{B}_a$ and $V_{\delta,1}$ are independent. The joint density function of $(U_{\delta,1}, V_{\delta,1})$ is then given as follows: For $x \in \mathbb{R}$ and $t \geq 0$,

$$\begin{aligned} &f_{(U_{\delta,1},V_{\delta,1})}(x,t) \\ &= f_{U_{\delta,1}}(x|V_{\delta,1}=t) f_{V_{\delta,1}}(t) \\ &= f_{[\check{a}+\check{\mu}_a t+\check{B}_a(t)]/[\sigma_a^{-1}\cos\theta+\sigma_b^{-1}\sin\theta]}(x) f_{V_{\delta,1}}(t) \\ &= \frac{\sigma_a^{-1}\cos\theta+\sigma_b^{-1}\sin\theta}{\sqrt{2\pi t}} \exp\left\{-\frac{((\sigma_a^{-1}\cos\theta+\sigma_b^{-1}\sin\theta)x-\check{a}-\check{\mu}_a t)^2}{2t}\right\} f_{V_{\delta,1}}(t) \\ &= \frac{\alpha-\beta}{2\pi t^2\sigma_a\sigma_b} \exp\left\{-\frac{[\sigma_b/\sigma_a(x-\alpha-\mu_a t)+\sigma_a/\sigma_b(x-\beta-\mu_b t)]^2+[\alpha-\beta-(\mu_b-\mu_a)t]^2}{2(\sigma_a^2+\sigma_b^2)t}\right\}, \end{aligned}$$

where the last equality follows from (18). This proves (i). For (ii), we observe that for $t \in [0, T_{\delta,n+1} - T_{\delta,n}), n = 1, 2, \ldots,$

$$\begin{aligned} \tilde{X}_{a,n}(t) &\equiv \ln(A_\delta(t+T_{\delta,n})) - \ln(A_\delta(T_{\delta,n}-)) = X_a(t+T_{\delta,n}) - X_a(T_{\delta,n}) + \delta, \\ \tilde{X}_{b,n}(t) &\equiv \ln(B_\delta(t+T_{\delta,n})) - \ln(B_\delta(T_{\delta,n}-)) = X_b(t+T_{\delta,n}) - X_b(T_{\delta,n}) - \delta. \end{aligned}$$

From the strong Markov property of Brownian motions, $\{\tilde{X}_{a,n}(t); t \geq 0\}$ and $\{\tilde{X}_{b,n}(t); t \geq 0\}$ are independent Brownian motions with the initial values δ and $-\delta$, and the same drifts and volatility parameters as X_a and X_b. Furthermore, they are independent of $\mathcal{F}_{T_{\delta,n}}$, where

$$\mathcal{F}_t = \sigma\{(X_a(s), X_b(s)), 0 \leq s \leq t\}. \tag{19}$$

Thus if we let $\tilde{T}_n$ and $\tilde{L}_n$ denote the time and position of the first meeting point of $\tilde{X}_{a,n}$ and $\tilde{X}_{b,n}$, then $\{(\tilde{L}_n, \tilde{T}_n), n = 1, 2, \ldots\}$ is an i.i.d. sequence, which is independent of $(U_{\delta,1}, V_{\delta,1})$,

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

and has the same distribution as $(U_{\delta,1}, V_{\delta,1})$ with $\alpha = \delta$ and $\beta = -\delta$. Finally, noting that $X_a(T_{\delta,n}) - X_b(T_{\delta,n}) = -2(n-1)\delta$, we have that

$$\begin{aligned}\tilde{T}_n &= \inf\{t \geq 0 : \tilde{X}_{a,n}(t) = \tilde{X}_{b,n}(t)\} \\ &= \inf\{t \geq 0 : X_a(t + T_{\delta,n}) - X_a(T_{\delta,n}) + \delta = X_b(t + T_{\delta,n}) - X_b(T_{\delta,n}) - \delta\} \\ &= \inf\{t \geq 0 : X_a(t + T_{\delta,n}) - X_b(t + T_{\delta,n}) = -2n\delta\} \\ &= T_{\delta,n+1} - T_{\delta,n} \\ &= V_{\delta,n},\end{aligned}$$

and

$$\tilde{L}_n = \tilde{X}_{a,n}(\tilde{T}_n-) = \ln(A_\delta(T_{\delta,n+1}-)) - \ln(A_\delta(T_{\delta,n}-)) = \ln(P_{\delta,n+1}) - \ln(P_{\delta,n}) = U_{\delta,n}.$$

To summarize, we have shown that $\{(U_{n,\delta}, V_{n,\delta}), n = 2, 3, \ldots\}$ is an i.i.d. sequence, which is independent of $(V_{\delta,1}, U_{\delta,1})$, and has the same distribution as $(V_{\delta,1}, U_{\delta,1})$ with $\alpha = \delta$ and $\beta = -\delta$. ■

Using the above definitions, we have the following conclusion on the marginal distributions of $(U_n, V_n), n \geq 1$.

Corollary 3.3:

(i) *Assume $A_\delta(0) = e^\alpha$, $B_\delta(0) = e^\beta$, and $\alpha > \beta$. Then*

$$V_{\delta,1} \sim \text{IG}\left(\frac{\alpha - \beta}{\sqrt{\sigma_a^2 + \sigma_b^2}}, \frac{\mu_b - \mu_a}{\sqrt{\sigma_a^2 + \sigma_b^2}}\right),$$

and

$$U_{\delta,1} \sim \text{NIG}\left(\frac{\sqrt{(\sigma_a^2 + \sigma_b^2)(\mu_a^2\sigma_b^2 + \mu_b^2\sigma_a^2)}}{\sigma_a^2\sigma_b^2}, \frac{\mu_a\sigma_b^2 + \mu_b\sigma_a^2}{\sigma_a^2\sigma_b^2}, \frac{\alpha\sigma_b^2 + \beta\sigma_a^2}{\sigma_a^2 + \sigma_b^2}, \frac{(\alpha - \beta)\sigma_a\sigma_b}{\sigma_a^2 + \sigma_b^2}\right).$$

(ii) *For $n \geq 2$, $V_{\delta,n}$ and $U_{\delta,n}$ follow the same IG and NIG distributions as described in* (i) *with $\alpha = \delta$ and $\beta = -\delta$.*

Let (U_δ, V_δ) be a generic random variable with the same joint distribution as $(U_{\delta,n}, V_{\delta,n}), n \geq 2$. Next, we find the moment generating function of (U_δ, V_δ), which will be used in the proof of Theorem 3.8 and Section 4.

Lemma 3.4: *There exists $h > 0$ such that the moment generating function of (U_δ, V_δ) exists for $|(s,t)| \leq h$, and is given by*

$$\phi_\delta(s,t) = E\left[\exp\{sU_\delta + tV_\delta\}\right] = \exp\{[2\theta(s,t) - s]\delta\}, \tag{20}$$

where

$$\theta(s,t) = \frac{(\mu_b - \mu_a + s\sigma_b^2) - \sqrt{(\mu_b - \mu_a + s\sigma_b^2)^2 - (\sigma_a^2 + \sigma_b^2)(s^2\sigma_b^2 + 2t + 2s\mu_b)}}{\sigma_a^2 + \sigma_b^2}. \tag{21}$$

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

In particular, the first two moments of (U_δ, V_δ) *are given below:*

$$E(V_\delta) = \frac{2\delta}{\mu_b - \mu_a}, \quad E(U_\delta) = \frac{\delta(\mu_b + \mu_a)}{\mu_b - \mu_a},$$
$$Var(V_\delta) = \frac{2(\sigma_a^2 + \sigma_b^2)\delta}{(\mu_b - \mu_a)^3}, \quad Var(U_\delta) = \frac{2(\mu_b^2\sigma_a^2 + \mu_a^2\sigma_b^2)\delta}{(\mu_b - \mu_a)^3},$$
$$Cov(U_\delta, V_\delta) = \frac{2(\mu_b\sigma_a^2 + \mu_a\sigma_b^2)\delta}{(\mu_b - \mu_a)^3}.$$

Furthermore, for $k, l \in \mathbb{N} \cup \{0\}$ *and* $k + l \geq 1$*, there exists some constant* c_0 *such that*

$$\frac{E(U_\delta^k V_\delta^l)}{\delta} \to c_0, \quad as\ \delta \to 0. \tag{22}$$

Proof: Assume $A_\delta(0) = e^\delta$ and $B_\delta(0) = e^{-\delta}$. Then $(U_{\delta,1}, V_{\delta,1})$ has the same distribution as (U_δ, V_δ). Let

$$Y_a(t) = \exp\left\{\theta_1 X_a(t) - \left(\theta_1\mu_a + \frac{1}{2}\theta_1^2\sigma_a^2\right)t\right\},$$
$$Y_b(t) = \exp\left\{\theta_2 X_b(t) - \left(\theta_2\mu_b + \frac{1}{2}\theta_2^2\sigma_b^2\right)t\right\},$$

where θ_1 and θ_2 are arbitrary real numbers. Then $\{Y_a(t); t \geq 0\}$ and $\{Y_b(t); t \geq 0\}$ are independent, and $\{(Y_a(t), Y_b(t)); t \geq 0\}$ is an $\{\mathcal{F}_t\}_{t\geq 0}$ martingale (see the beginning of Section 5 of Chapter 7 in Karlin and Taylor [24]), where $\mathcal{F}_t$ is defined in (19). In fact, $\{Y_a(t)Y_b(t); t \geq 0\}$ is also an $\{\mathcal{F}_t\}_{t\geq 0}$ martingale. Indeed for $0 \leq s \leq t$, by the independence of Y_a and Y_b,

$$\begin{aligned} E(Y_a(t)Y_b(t)|\mathcal{F}_s) &= E([Y_a(t) - Y_a(s)][Y_b(t) - Y_b(s)]|\mathcal{F}_s) + E(Y_a(t)Y_b(s)|\mathcal{F}_s) \\ &\quad + E(Y_b(t)Y_a(s)|\mathcal{F}_s) - E(Y_a(s)Y_b(s)|\mathcal{F}_s) \\ &= 0 + Y_b(s)E(Y_a(t)|\mathcal{F}_s) + Y_a(s)E(Y_b(t)|\mathcal{F}_s) - Y_a(s)Y_b(s) \\ &= Y_a(s)Y_b(s). \end{aligned}$$

Also note that $V_{\delta,1}$ is an $\{\mathcal{F}_t\}_{t\geq 0}$ stopping time with finite mean and variance. The optional stopping theorem yields

$$E[Y_a(V_{\delta,1})Y_b(V_{\delta,1})] = E[Y_a(0)Y_b(0)].$$

Plugging the expressions of Y_a and Y_b into the above equation, and noting that $X_a(V_{\delta,1}) = X_b(V_{\delta_1}) = U_{\delta,1}$ yields that

$$E\left\{\exp\left\{(\theta_1 + \theta_2)U_{\delta,1} - \left(\theta_1\mu_a + \frac{1}{2}\theta_1^2\sigma_a^2 + \theta_2\mu_b + \frac{1}{2}\theta_2^2\sigma_b^2\right)V_{\delta,1}\right\}\right\} = \exp\left\{[\theta_1 - \theta_2]\delta\right\}.$$

Let

$$\theta_1 + \theta_2 = s,$$
$$\theta_1\mu_a + \tfrac{1}{2}\theta_1^2\sigma_a^2 + \theta_2\mu_b + \tfrac{1}{2}\theta_2^2\sigma_b^2 = -t.$$

Solving θ_1 and θ_2 in terms of s and t, we obtain

$$\theta_1(s,t) = \frac{(\mu_b - \mu_a + s\sigma_b^2) \pm \sqrt{(\mu_b - \mu_a + s\sigma_b^2)^2 - (\sigma_a^2 + \sigma_b^2)(s^2\sigma_b^2 + 2t + 2s\mu_b)}}{\sigma_a^2 + \sigma_b^2},$$
$$\theta_2(s,t) = s - \theta_1(s,t).$$

Letting $s = 0$, and noting that $V_{\delta,1}$ follows IG distribution (see (16)), the moment generating function of $V_{\delta,1}$ is known to be

$$E(\exp(tV_1)) = \frac{(\mu_b - \mu_a) - \sqrt{(\mu_b - \mu_a)^2 - 2t(\sigma_a^2 + \sigma_b^2)}}{\sigma_a^2 + \sigma_b^2}.$$

Thus the solutions of $\theta_1(s,t)$ should be $\theta(s,t)$ in (21), and the moment generating function $\phi(s,t)$ of (U_δ, V_δ) is given by (20). To compute the moments, we first need some simple results about $\theta(s,t)$ as follows.

$$\theta(0,0) = 0,$$
$$\left.\frac{\partial\theta(s,t)}{\partial t}\right|_{s=t=0} = \frac{1}{\mu_b - \mu_a}, \quad \left.\frac{\partial\theta(s,t)}{\partial s}\right|_{s=t=0} = \frac{\mu_b}{\mu_b - \mu_a},$$
$$\left.\frac{\partial^2\theta(s,t)}{\partial t^2}\right|_{s=t=0} = \frac{\sigma_a^2 + \sigma_b^2}{(\mu_b - \mu_a)^3}, \quad \left.\frac{\partial^2\theta(s,t)}{\partial s^2}\right|_{s=t=0} = \frac{\mu_b^2\sigma_a^2 + \mu_a^2\sigma_b^2}{(\mu_b - \mu_a)^3},$$
$$\left.\frac{\partial^2\theta(s,t)}{\partial s\partial t}\right|_{s=t=0} = \frac{\mu_b\sigma_a^2 + \mu_a\sigma_b^2}{(\mu_b - \mu_a)^3}, \quad \left.\frac{\partial^3\theta(s,t)}{\partial t^2\partial s}\right|_{s=t=0} = \frac{3(\sigma_a^2\mu_b + \sigma_b^2\mu_a)(\sigma_a^2 + \sigma_b^2)}{(\mu_b - \mu_a)^5}.$$

Therefore,

$$E(V_\delta) = \left.\frac{\partial\phi(s,t)}{\partial t}\right|_{s=t=0} = \frac{\partial}{\partial t}\exp\{[2\theta(s,t) - s]\delta\}|_{s=0,t=0}$$
$$= \frac{2\delta}{\mu_b - \mu_a}.$$

Similarly, we obtain

$$E(U_\delta) = \frac{\delta(\mu_b + \mu_a)}{\mu_b - \mu_a}$$
$$E(V_\delta^2) = \frac{4\delta^2}{(\mu_b - \mu_a)^2} + \frac{2(\sigma_a^2 + \sigma_b^2)\delta}{(\mu_b - \mu_a)^3}$$
$$E({U_\delta}^2) = \frac{\delta^2(\mu_a + \mu_b)^2}{(\mu_b - \mu_a)^2} + \frac{2(\mu_b^2\sigma_a^2 + \mu_a^2\sigma_b^2)\delta}{(\mu_b - \mu_a)^3}$$
$$E\left(U_\delta V_\delta\right) = \frac{2\delta^2(\mu_b + \mu_a)}{(\mu_b - \mu_a)^2} + \frac{2(\mu_b\sigma_a^2 + \mu_a\sigma_b^2)\delta}{(\mu_b - \mu_a)^3}.$$

Finally, for $k, l \in \mathbb{N} \cup \{0\}$ and $k + l \geq 1$, (22) follows by noting that

$$E(U_\delta^k V_\delta^l) = \left.\frac{\partial^{k+l}\phi(s,t)}{\partial s^k \partial t^l}\right|_{s=t=0} = \delta\left.\left(\phi(s,t)\ \frac{\partial^{k+l}}{\partial s^k\partial t^l}(2\theta(s,t) - s)\right)\right|_{s=t=0} + o(\delta),$$

where $o(\delta) \to 0$ as $\delta \to 0$. ■

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

Let $\alpha = \delta$ and $\beta = -\delta$, and define for $t \geq 0$,

$$T_\delta(t) = T_{\delta,\lfloor t/\delta \rfloor} = \sum_{k=1}^{\lfloor t/\delta \rfloor} V_{\delta,k},$$

and

$$S_\delta(t) = \frac{\ln P_{\delta,\lfloor t/\delta\rfloor} - \frac{(\sigma_b^2-\sigma_a^2)\delta\lfloor t/\delta\rfloor + (\sigma_b^2\mu_a+\sigma_a^2\mu_b)T_\delta(t)}{\sigma_a^2+\sigma_b^2}}{\sqrt{\frac{\sigma_a^2\sigma_b^2}{\sigma_a^2+\sigma_b^2}}} = \sum_{k=1}^{\lfloor t/\delta\rfloor} \frac{U_{\delta,k} - \frac{(\sigma_b^2-\sigma_a^2)\delta + (\sigma_b^2\mu_a+\sigma_a^2\mu_b)V_{\delta,k}}{\sigma_a^2+\sigma_b^2}}{\sqrt{\frac{\sigma_a^2\sigma_b^2}{\sigma_a^2+\sigma_b^2}}}.$$

The following lemma shows that as $\delta \to 0$, (T_δ, S_δ) converges weakly to an (IG, NIG) Lévy process. Following Example 4.3 in Barndorff–Nielsen *et al.* [3], the (IG, NIG) Lévy process can be defined by considering a two-dimensional Brownian motion (X, Y) with zero initial value, drift vector $(\gamma, 0)$, and identity covariance matrix. Let $\tau(t)$ be the first time X reaches the level ζt, and define $\xi(t) = Y(\tau(t))$. Then $\{(\tau(t), \xi(t)); t \geq 0\}$ is a bivariate Lévy process. Noting that $\tau(1) \sim \mathrm{IG}(\zeta, \gamma)$, and $\xi(1) \sim \mathrm{NIG}(\gamma, 0, 0, \zeta)$, the Lévy process $\{(\tau(t), \xi(t)); t \geq 0\}$ is thus called an (IG, NIG) Lévy process with parameters γ and ζ.

LEMMA 3.5: *As $\delta \to 0$, (S_δ, T_δ) converges weakly in $D([0,\infty);\mathbb{R}^2)$ to an (IG, NIG) Lévy process with parameters $2(\sigma_a^2 + \sigma_b^2)^{-1/2}$ and $(\mu_b - \mu_a)(\sigma_a^2 + \sigma_b^2)^{-1/2}$.*

PROOF: It is clear that (T_δ, S_δ) has independent increments. Note that both IG and NIG distributions are closed under convolutions. From Lemma 3.2 and Corollary 3.3, $T_\delta(t)$ follows $\mathrm{IG}(2\delta\lfloor t/\delta\rfloor(\sigma_a^2 + \sigma_b^2)^{-1/2}, (\mu_b - \mu_a)(\sigma_a^2 + \sigma_b^2)^{-1/2})$, and given that $T_\delta(t) = x$, $S_\delta(t)$ follows the normal distribution with mean 0 and variance x. For convenience, let

$$\bar{\zeta} = 2(\sigma_a^2 + \sigma_b^2)^{-1/2},\ \bar{\gamma} = (\mu_b - \mu_a)(\sigma_a^2 + \sigma_b^2)^{-1/2},\ \bar{\zeta}_\delta(t) = 2\delta\lfloor t/\delta\rfloor(\sigma_a^2 + \sigma_b^2)^{-1/2}.$$

From Example 4.3 in Barndorff–Nielsen *et al.* [3], the characteristic triplet of $(T_\delta(t), S_\delta(t))$ is $((\bar{\gamma}\bar{\zeta}_\delta(t), 0), 0, \Pi_\delta)$, where the Lévy measure Π_δ has the following density function

$$\pi_\delta(x, y) = \frac{\bar{\zeta}_\delta(t)}{2\pi x^2} \exp\left\{-\frac{1}{2}\left(\bar{\gamma}^2 x + \frac{y^2}{x}\right)\right\},\ x \geq 0, y \in \mathbb{R}.$$

Conversely, still from Example 4.3 in Barndorff–Nielsen *et al.* [3], the (IG, NIG) Lévy process with parameters $\bar{\zeta}$ and $\bar{\gamma}$ at time t has characteristic triplet $((\bar{\gamma}\bar{\zeta}t, 0), 0, \Pi)$, where the Lévy measure Π has the following density function

$$\pi(x, y) = \frac{\bar{\zeta}t}{2\pi x^2} \exp\left\{-\frac{1}{2}\left(\bar{\gamma}^2 x + \frac{y^2}{x}\right)\right\},\ x \geq 0, y \in \mathbb{R}.$$

It is clear that the δ-dependent characteristic triplet of (T_δ, S_δ) converges to the characteristic triplet of an (IG, NIG) Lévy process with parameters $\bar{\zeta}$ and $\bar{\gamma}$, as $\delta \to 0$. Noting that (T_δ, S_δ) is a semimartingale of independent increments, from Theorem VII.2.52 in Jacod and Shiryaev [22], the lemma follows. ■

3.2. Asymptotics of $\{Z_\delta(t); t \geq 0\}$

In this section, we study the behaviors of the $\{Z_\delta(t); t \geq 0\}$ process as either $t \to \infty$ or $\delta \to 0$. First from Corollary 3.3, it is clear that for each δ, $\{Z_\delta(t); t \geq 0\}$ is a renewal reward process, and we summarize it in the following lemma.

LEMMA 3.6: *For $\delta > 0$, $\{Z_\delta(t); t \geq 0\}$ is a renewal reward process and $\{P_\delta(t); t \geq 0\}$ is a semi-Markov process.*

The next result from Brown and Solomon Brown and Solomon [11] characterizes the asymptotic first and second moments of $Z_\delta(t)$ as $t \to \infty$, and is also helpful to identify the proper scaling in Theorem 3.8.

PROPOSITION 3.7 (Brown and Solomon Brown and Solomon [11]): *As $t \to \infty$, we have*

$$E(Z_\delta(t)) = mt + \mathcal{O}(1),$$
$$Var(Z_\delta(t)) = st + \mathcal{O}(1),$$

where

$$m = \tfrac{1}{2}(\mu_a + \mu_b), \quad s = \tfrac{1}{4}(\sigma_a^2 + \sigma_b^2), \tag{23}$$

and $O(1)$ represents a function of t, which is bounded as $t \to \infty$.

PROOF: From Brown and Solomon [11], we have the following results for a renewal reward process generated by $\{(U_{\delta,n}, V_{\delta,n}), n \geq 1\}$:

$$E(Z_\delta(t)) = mt + \mathcal{O}(1),$$

and

$$\mathrm{Var}(Z_\delta(t)) = st + \mathcal{O}(1),$$

where

$$m = \frac{E(U_\delta)}{E(V_\delta)},$$

and

$$s = \frac{E(V_\delta^2)E(U_\delta)^2}{E(V_\delta)^3} - \frac{2E(U_\delta V_\delta)E(U_\delta)}{E(V_\delta)^2} + \frac{E(U_\delta^2)}{E(V_\delta)},$$

Using the results of Lemma 3.4 in the above equations, we get (23). ■

The main result of this section is the following theorem. For $t \geq 0$, define

$$\hat{Z}_\delta(t) = \frac{\delta Z_\delta(t/\delta) - mt}{\sqrt{s\delta}},$$

where m and s are as given in (23).

THEOREM 3.8: *Assume that $E[\ln(A_\delta(0)) - \ln(B_\delta(0))]^2 < \infty$. Then the process $\hat{Z}_\delta$ converges weakly to a standard Brownian Motion as $\delta \to 0$ in $D([0, \infty); \mathbb{R})$.*

Proof: Consider an arbitrary nonnegative sequence $\{\delta_m\}_{m\geq 1}$ such that $\delta_m \to 0$ as $m \to \infty$. Define for $m, n \geq 1$,

$$\tilde{U}_{m,n} = \sqrt{\delta_m}(U_{\delta_m,n} - E(U_{\delta_m,n})),$$
$$\tilde{V}_{m,n} = \sqrt{\delta_m}(V_{\delta_m,n} - E(V_{\delta_m,n})).$$

We note that for each m, $\{(\tilde{U}_{m,n}, \tilde{V}_{m,n}), n \geq 2\}$ is an i.i.d. sequence. Furthermore,

$$\sum_{n=1}^{\lfloor \frac{t}{\delta_m^2} \rfloor} \mathrm{Var}(\tilde{U}_{m,n}) \to \frac{2(\mu_b^2\sigma_a^2 + \mu_a^2\sigma_b^2)t}{(\mu_b - \mu_a)^3}, \text{ and } \sum_{n=1}^{\lfloor \frac{t}{\delta_m^2} \rfloor} \mathrm{Var}(\tilde{V}_{m,n}) \to \frac{2(\sigma_a^2 + \sigma_b^2)t}{(\mu_b - \mu_a)^3}, \text{ as } m \to \infty.$$

We claim that $\{(\tilde{U}_{m,n}, \tilde{V}_{m,n}), m \geq 1, 1 \leq n \leq \lfloor t/\delta_m^2 \rfloor\}$ satisfies Lindeberg condition, that is, for any $\epsilon > 0$,

$$\sum_{n=1}^{\lfloor \frac{t}{\delta_m^2} \rfloor} E\left(\tilde{U}_{m,n}^2 1_{\{|\tilde{U}_{m,n}| \geq \epsilon\}}\right) \to 0, \text{ and } \sum_{n=1}^{\lfloor \frac{t}{\delta_m^2} \rfloor} E\left(\tilde{V}_{m,n}^2 1_{\{|\tilde{V}_{m,n}| \geq \epsilon\}}\right) \to 0, \text{ as } m \to \infty. \quad \mathbf{(24)}$$

We will prove (24) at the end of this proof. Thus from Billingsley [7], Theorem 18.2, letting

$$u_m(t) = \sum_{n=1}^{\lfloor \frac{t}{\delta_m^2} \rfloor} \tilde{U}_{m,n}, \quad \text{and} \quad v_m(t) = \sum_{n=1}^{\lfloor \frac{t}{\delta_m^2} \rfloor} \tilde{V}_{m,n},$$

then

$$(u_m, v_m) \Rightarrow W, \quad \text{as} \quad m \to \infty.$$

where W is a two-dimensional Brownian motion with drift 0 and covariance matrix

$$\frac{2}{(\mu_b - \mu_a)^3} \begin{pmatrix} \mu_b^2\sigma_a^2 + \mu_a^2\sigma_b^2 & \mu_b\sigma_a^2 + \mu_a\sigma_b^2 \\ \mu_b\sigma_a^2 + \mu_a\sigma_b^2 & \sigma_a^2 + \sigma_b^2 \end{pmatrix}.$$

Next from Iglehart and Whitt [21], Theorem 1 and Jacod and Shiryaev [22], Corollary 3.33, if

$$\tilde{N}_m(t) = (E(V_{\delta_m,1}))^{3/2}\left(N_{\delta_m}(t/\delta_m) - \frac{t}{\delta_m E(V_{\delta_m,1})}\right)$$
$$= \left(\frac{2\delta_m}{\mu_b - \mu_a}\right)^{3/2}\left(N_{\delta_m}(t/\delta_m) - \frac{(\mu_b - \mu_a)t}{2\delta_m^2}\right),$$

then $(u_m, v_m, \tilde{N}_m) \Rightarrow (W_1, W_2, -W_2)$ as $m \to \infty$, where W_1 and W_2 are the first and second components of the Brownian motion W. Finally, we note that

$$\hat{Z}_{\delta_m}(t) = \frac{1}{\sqrt{s}}\left[u_m\left(\delta_m^2 N_{\delta_m}(t/\delta_m)\right) + \frac{\mu_b + \mu_a}{\mu_b - \mu_a}\left(\frac{\mu_b - \mu_a}{2}\right)^{3/2}\tilde{N}_m(t)\right].$$

Furthermore, observing that

$$\delta_m^2 N_{\delta_m}(t/\delta_m) = \delta_m^2\left[\frac{\tilde{N}_m(t)}{(E(V_{\delta_m,1}))^{3/2}} + \frac{(\mu_b - \mu_a)t}{2\delta_m^2}\right] \to \frac{(\mu_b - \mu_a)t}{2}, \text{ as } m \to \infty,$$

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

we have that

$$\hat{Z}_{\delta_m}(\cdot) \Rightarrow \frac{W_1(((\mu_b-\mu_a)/(2))\ \cdot) + ((\mu_b+\mu_a)/(\mu_b-\mu_a))(((\mu_b-\mu_a)/(2)))^{3/2}W_2(\cdot)}{\sqrt{s}},$$

and it is easy to check that the weak limit on the right-hand side is a standard Brownian motion. Consequently, $\hat{Z}_\delta$ converges weakly to a standard Brownian motion as $\delta \to 0$. At last, we give the proof of the claim given in (24). The proofs for $\tilde{V}_{m,n}$ and $\tilde{U}_{m,n}$ are similar, and we only consider $\tilde{V}_{m,n}$. We first note that from Lemma 3.4,

$$E(V_{\delta_m,1}|A(0),B(0)) = \frac{\ln A(0) - \ln B(0)}{\mu_b-\mu_a},$$
$$\mathrm{Var}(V_{\delta_m,1}|A(0),B(0)) = \frac{(\ln A(0) - \ln B(0))(\sigma_a^2+\sigma_b^2)}{(\mu_b-\mu_a)^3},$$

and using conditional expectations, we have that for some $b_0 \in (0,\infty)$,

$$\begin{aligned}\mathrm{Var}(V_{\delta_m,1}) &= E(\mathrm{Var}(V_{\delta_m,1}|A(0),B(0))) + \mathrm{Var}(E(V_{\delta_m,1}|A(0),B(0)))\\ &\le b_0\left(E[\ln(A(0)/B(0))] + E[\ln^2(A(0)/B(0))]\right) < \infty.\end{aligned}$$

Next using Markov inequality, Holder's inequality and (22), we have for some $c_0 \in (0,\infty)$,

$$\begin{aligned}&\sum_{n=1}^{\lfloor \frac{t}{\delta_m^2}\rfloor} E\left(\tilde{V}_{m,n}^2 1_{\{|\tilde{V}_{m,n}|\ge\epsilon\}}\right)\\ &\le E(\tilde{V}_{m,1}^2) + \left\lceil \frac{t}{\delta_m^2}\right\rceil \sqrt{E(\tilde{V}_{m,2}^4)P(|\tilde{V}_{m,2}|\ge\epsilon)}\\ &\le \delta_m \mathrm{Var}(V_{\delta_m,1}) + \left\lceil \frac{t}{\delta_m^2}\right\rceil \sqrt{E(\tilde{V}_{m,2}^4)\epsilon^{-2}E(\tilde{V}_{m,2}^2)}\\ &\le \delta_m \mathrm{Var}(V_{\delta_m,1}) + \epsilon^{-1}\left\lceil \frac{t}{\delta_m^2}\right\rceil \delta_m^{3/2}\sqrt{E[(V_{\delta_m,2}-E(V_{\delta_m,2}))^4]\mathrm{Var}(V_{\delta_m,2})}\\ &\le \delta_m \mathrm{Var}(V_{\delta_m,1}) + \epsilon^{-1}\left\lceil \frac{t}{\delta_m^2}\right\rceil \delta_m^{3/2}\sqrt{c_0\delta_m^2}\\ &\to 0, \quad \text{as} \quad m\to\infty.\end{aligned}$$

■

Remark 3.9: We note that

$$Z_\delta(t) = \sqrt{\frac{s}{\delta}}\hat{Z}_\delta(\delta t) + mt,\ t\ge 0.$$

From Theorem 3.8, for small δ, we will use the following asymptotic model for logarithmic trading prices $Z_\delta(t)$ in Section 5:

$$\sqrt{\frac{s}{\delta}}W(\delta t) + mt, \tag{25}$$

where $\{W(t); t\ge 0\}$ is a standard Brownian motion. We note that (25) is normal distributed with mean mt and variance st.

4. PARAMETER ESTIMATIONS

We assume that the process $\{P_\delta(t); t \geq 0\}$ is observable, while $\{(A_\delta(t), B_\delta(t)); t \geq 0\}$ may not be publicly observable (e.g., the market ask and bid processes may be accessible to the brokers and dealers, but not to common traders). The question becomes how to find the parameters of $\{(A_\delta(t), B_\delta(t)); t \geq 0\}$ by observing $\{P_\delta(t); t \geq 0\}$. In this section, we will estimate the parameters μ_a, μ_b, σ_a, σ_b, and δ using the method of moments.

Suppose that the sample data for the time and position (t_i, p_i) of the ith meeting point are given for $i = 1, 2, \ldots, n$. Let

$$u_1 = \ln p_1, \quad v_1 = t_1, \quad \text{and} \quad u_{i+1} = \ln(p_{i+1}/p_i), \quad v_{i+1} = t_{i+1} - t_i, \ i \geq 1.$$

Then the sample data are given by $\{(u_i, v_i)\}_{i=1}^n$. Let

$$x_1 = \sum_{i=1}^n \frac{v_i}{n}, \quad x_2 = \sum_{i=1}^n \frac{u_i}{n}, \quad x_3 = \sum_{i=1}^n \frac{v_i^2}{n}, \quad x_4 = \sum_{i=1}^n \frac{u_i^2}{n}, \quad x_5 = \sum_{i=1}^n \frac{v_i u_i}{n}.$$

We aim to derive explicit estimators of the five parameters μ_a, μ_b, σ_a, σ_b, δ using moment estimations. Define the estimators of μ_a, μ_b, σ_a, σ_b, δ as follows.

$$\begin{aligned} &\hat{\mu}_a^n = \frac{y_1 - \sqrt{y_1^2 - (4(y_1 y_4 - y_3)/(y_2))}}{2}, \quad \hat{\mu}_b^n = \frac{y_1 + \sqrt{y_1^2 - (4(y_1 y_4 - y_3)/(y_2))}}{2}, \\ &\hat{\sigma}_a^n = \sqrt{(y_4 - \hat{\mu}_a^n y_2)(\hat{\mu}_b^n - \hat{\mu}_a^n)}, \quad \hat{\sigma}_b^n = \sqrt{(\hat{\mu}_b^n y_2 - y_4)(\hat{\mu}_b^n - \hat{\mu}_a^n)}, \\ &\hat{\delta}^n = (\hat{\mu}_b^n - \hat{\mu}_a^n) x_1, \end{aligned} \tag{26}$$

where

$$y_1 = \frac{2x_2}{x_1}, \quad y_2 = \frac{x_3 - x_1^2}{x_1}, \quad y_3 = \frac{x_4 - x_2^2}{x_1}, \quad y_4 = \frac{x_5 - x_1 x_2}{x_1}.$$

For convenience, denote $\Theta = (\mu_a, \mu_b, \sigma_a, \sigma_b, \delta)$ and $\hat{\Theta}^n = (\hat{\mu}_a^n, \hat{\mu}_b^n, \hat{\sigma}_a^n, \hat{\sigma}_b^n, \hat{\delta}^n)$. Let $\mathbf{g} : \mathbb{R}^5 \to \mathbb{R}^5$ be the differentiable function such that

$$\hat{\Theta}^n = \mathbf{g}(x_1, x_2, x_3, x_4, x_5).$$

Note that $\mathbf{g}$ can be uniquely determined by (26) and has an explicit expression.

LEMMA 4.1: *The estimators $\hat{\Theta}^n$ are well defined, that is,*

$$y_1^2 - \frac{4(y_1 y_4 - y_3)}{y_2} \geq 0, \quad (y_4 - \hat{\mu}_a^n y_2)(\hat{\mu}_b^n - \hat{\mu}_a^n) \geq 0, \quad (\hat{\mu}_b^n y_2 - y_4)(\hat{\mu}_b^n - \hat{\mu}_a^n) \geq 0, \tag{27}$$

and as $n \to \infty$, almost surely,

$$\hat{\Theta}^n \to \Theta. \tag{28}$$

Furthermore, $\sqrt{n}(\hat{\Theta}^n - \Theta)$ converges weakly to a five-dimensional normal distribution with zero mean and covariance matrix $\nabla \mathbf{g}(\Theta)\Sigma$, where Σ is the covariance matrix of $(V_\delta, U_\delta, V_\delta^2, U_\delta^2, U_\delta V_\delta)$, and $\nabla \mathbf{g}$ is the gradient of $\mathbf{g}$.

PROOF: For convenience, we omit the superscript n for the estimators of $\mu_a, \mu_b, \sigma_a, \sigma_b$, and δ. Using the moments in Lemma 3.4, we consider the following equations.

$$x_1 = \frac{2\hat{\delta}}{\hat{\mu}_b - \hat{\mu}_a} \tag{29}$$

$$x_2 = \frac{\hat{\delta}(\hat{\mu}_b + \hat{\mu}_a)}{\hat{\mu}_b - \hat{\mu}_a} \tag{30}$$

$$x_3 = \frac{4\hat{\delta}^2}{(\hat{\mu}_b - \hat{\mu}_a)^2} + \frac{2(\hat{\sigma}_a^2 + \hat{\sigma}_b^2)\hat{\delta}}{(\hat{\mu}_b - \hat{\mu}_a)^3} \tag{31}$$

$$x_4 = \frac{\hat{\delta}^2(\hat{\mu}_a + \hat{\mu}_b)^2}{(\hat{\mu}_b - \hat{\mu}_a)^2} + \frac{2(\hat{\mu}_b^2\hat{\sigma}_a^2 + \hat{\mu}_a^2\hat{\sigma}_b^2)\hat{\delta}}{(\hat{\mu}_b - \hat{\mu}_a)^3} \tag{32}$$

$$x_5 = \frac{2\hat{\delta}^2(\hat{\mu}_b + \hat{\mu}_a)}{2(\hat{\mu}_b - \hat{\mu}_a)^2} + \frac{2(\hat{\mu}_b\hat{\sigma}_a^2 + \hat{\mu}_a\hat{\sigma}_b^2)\hat{\delta}}{(\hat{\mu}_b - \hat{\mu}_a)^3}. \tag{33}$$

Next we solve the above equations for $\hat{\mu}_a$, $\hat{\mu}_b$, $\hat{\sigma}_a$, $\hat{\sigma}_b$, $\hat{\delta}$ in terms of $x_k, k = 1, 2, \ldots, 5$. Let

$$\begin{aligned}
y_1 &= \frac{2x_2}{x_1} = \hat{\mu}_b + \hat{\mu}_a \\
y_2 &= \frac{x_3 - x_1^2}{x_1} = \frac{\hat{\sigma}_a^2 + \hat{\sigma}_b^2}{(\hat{\mu}_b - \hat{\mu}_a)^2} \\
y_3 &= \frac{x_4 - x_2^2}{x_1} = \frac{\hat{\mu}_b^2\hat{\sigma}_a^2 + \hat{\mu}_a^2\hat{\sigma}_b^2}{(\hat{\mu}_b - \hat{\mu}_a)^2} \\
y_4 &= \frac{x_5 - x_1x_2}{x_1} = \frac{\hat{\mu}_b\hat{\sigma}_a^2 + \hat{\mu}_a\hat{\sigma}_b^2}{(\hat{\mu}_b - \hat{\mu}_a)^2}.
\end{aligned}$$

We then note that

$$y_1^2 - 4\frac{y_1y_4 - y_3}{y_2} = (\hat{\mu}_b - \hat{\mu}_a)^2.$$

Letting $\hat{\mu}_b > \hat{\mu}_a$, we obtain

$$\begin{aligned}
\hat{\mu}_a &= \frac{y_1 - \sqrt{y_1^2 - 4((y_1y_4 - y_3)/(y_2))}}{2} \\
\hat{\mu}_b &= \frac{y_1 + \sqrt{y_1^2 - 4((y_1y_4 - y_3)/(y_2))}}{2}
\end{aligned}$$

and

$$\begin{aligned}
\hat{\sigma}_a &= \sqrt{(y_4 - \hat{\mu}_a y_2)(\hat{\mu}_b - \hat{\mu}_a)}, \\
\hat{\sigma}_b &= \sqrt{(\hat{\mu}_b y_2 - y_4)(\hat{\mu}_b - \hat{\mu}_a)}, \\
\hat{\delta} &= (\hat{\mu}_b - \hat{\mu}_a)x_1.
\end{aligned}$$

To see the above estimators are well-defined, we only need to show (27). We first note that

$$y_1^2 - \frac{4(y_1y_4 - y_3)}{y_2} = \frac{4}{x_1^2(x_3 - x_1^2)}\left[x_2^2(x_3 - x_1^2) - 2x_1x_2(x_5 - x_1x_2) + x_1^2(x_4 - x_2^2)\right].$$

It is clear that

$$x_1^2 = \left(\sum_{i=1}^{n} \frac{v_i}{n}\right)^2 > 0,$$

$$x_3 - x_1^2 = \frac{n\sum_{i=1}^{n} v_i^2 - \sum_{i=1}^{n} v_i}{n^2} > 0.$$

We next note that

$$\begin{aligned}
&x_2^2(x_3 - x_1^2) - 2x_1x_2(x_5 - x_1x_2) + x_1^2(x_4 - x_2^2)\\
&\geq 2x_1x_2\sqrt{(x_3 - x_1^2)(x_4 - x_2^2)} - 2x_1x_2(x_5 - x_1x_2)\\
&= 2x_1x_2(\sqrt{(x_3 - x_1^2)(x_4 - x_2^2)} - (x_5 - x_1x_2))\\
&= 2\frac{\sum v_i}{n}\frac{\sum u_i}{n}\left(\sqrt{\left(\frac{\sum v_i^2}{n} - \left(\frac{\sum v_i}{n}\right)^2\right)\left(\frac{\sum u_i^2}{n} - \left(\frac{\sum u_i}{n}\right)^2\right)}\right.\\
&\quad \left. - \left(\frac{\sum v_i \Delta p_i}{n} - \frac{\sum v_i}{n}\frac{\sum u_i}{n}\right)\right)\\
&= 2\frac{\sum v_i}{n}\frac{\sum u_i}{n}\left(\sqrt{\frac{\sum (v_i - \sum v_i/n)^2}{n}\frac{\sum (u_i - \sum u_i/n)^2}{n}} - \left(\frac{\sum v_iu_i}{n} - \frac{\sum v_i}{n}\frac{\sum u_i}{n}\right)\right)\\
&\geq 2\frac{\sum v_i}{n}\frac{\sum u_i}{n}\left(\frac{\sum (v_i - \sum v_i/n)(u_i - \sum u_i/n)}{n} - \left(\frac{\sum v_iu_i}{n} - \frac{\sum v_i}{n}\frac{\sum u_i}{n}\right)\right)\\
&= 0.
\end{aligned}$$

This shows the first inequality in (27). To show the last two inequalities in (27), we observe that

$$\begin{aligned}
y_4 - \hat{\mu}_a y_2 &= \frac{y_2\sqrt{y_1^2 - ((4(y_1y_4 - y_3))/(y_2))}}{2} + \left(y_4 - \frac{y_1y_2}{2}\right),\\
\hat{\mu}_b y_2 - y_4 &= \frac{y_2\sqrt{y_1^2 - ((4(y_1y_4 - y_3))/(y_2))}}{2} - \left(y_4 - \frac{y_1y_2}{2}\right).
\end{aligned}$$

Hence it suffices to show

$$\frac{y_2^2\left(y_1^2 - ((4(y_1y_4 - y_3))/(y_2))\right)}{4} \geq \left(y_4 - \frac{y_1y_2}{2}\right)^2.$$

After simplifying above inequality, it suffices to show that $y_2y_3 \geq y_4^2$. Note that

$$y_2y_3 \geq y_4^2$$

is equivalent to

$$(x_3 - x_1^2)(x_4 - x_2^2) \geq (x_5 - x_1x_2)^2,$$

and the latter one is proved above. This completes the proof of (27). Next from the construction of the estimators, we see that they are the unique solutions of (29)–(33). Using the strong law of large numbers and the continuous mapping theorem, we have (28). Finally, the central limit theorem for $\hat{\Theta}$ follows immediately from Delta method (cf. Casella and Berger [13]) and the central limit theorem for $(x_1, x_2, \ldots, x_5)$, that is,

$$\sqrt{n}[(x_1, x_2, x_3, x_4, x_5) - E(x_1, x_2, x_3, x_4, x_5)] \Rightarrow \mathcal{N}_5(0, \Sigma),$$

where Σ is the covariance matrix of $(V_\delta, U_\delta, V_\delta^2, U_\delta^2, U_\delta V_\delta)$. ∎

5. APPLICATION IN LOBS

In this section, we apply our model to a LOB as described in the introduction. We aim to forecast the trading price movement over a short period. We develop an asymptotic GBM model for trading prices as follows. As in Section 4, suppose that the sample data for the time and price (t_i, p_i) of the ith trade are given for $i = 1, 2, \ldots, n$. Let

$$u_1 = \ln p_1, \quad v_1 = t_1, \quad \text{and} \quad u_{i+1} = \ln(p_{i+1}/p_i), \quad v_{i+1} = t_{i+1} - t_i, \ i \geq 1.$$

We first estimate the parameters $\mu_a, \mu_b, \sigma_a, \sigma_b$, and δ as in (26), and use the estimators $\hat{\mu}_a^n, \hat{\mu}_b^n, \hat{\sigma}_a^n$, and $\hat{\sigma}_b^n$ to compute m and s by substituting $\mu_a, \mu_b, \sigma_a, \sigma_b$ with $\hat{\mu}_a^n, \hat{\mu}_b^n, \hat{\sigma}_a^n, \hat{\sigma}_b^n$, respectively, in (23). Typically, the estimator $\hat{\delta}^n$ is small (see Figures 6–9) and so from Theorem 3.8, we approximate $Z(t)$ by a $N(mt, st)$ random variable. Hence the prediction formula for $\ln P(t) - \ln P(0)$ is

$$\frac{(\hat{\mu}_a^n + \hat{\mu}_b^n)t}{2}, \tag{34}$$

and the upper and lower bounds are chosen to be

$$\frac{(\hat{\mu}_a^n + \hat{\mu}_b^n)t}{2} + \frac{3\sqrt{[(\hat{\sigma}_a^n)^2 + (\hat{\sigma}_b^n)^2]t}}{2}, \quad \frac{(\hat{\mu}_a^n + \hat{\mu}_b^n)t}{2} - \frac{3\sqrt{[(\hat{\sigma}_a^n)^2 + (\hat{\sigma}_b^n)^2]t}}{2}. \tag{35}$$

We next apply the above formulas to real data. Here we select the stock SUSQ (Susquehanna Bancshares Inc). The data are chosen from 01/04/2010 9:30AM to 01/04/2010 4:00PM, including the trading prices and trading times. The unit of trading prices is dollars and the unit of the difference of consecutive trading times is seconds. We perform the back test to evaluate the performance of the prediction. To be precise, we predict the logarithmic trading price at each trading time using the 10-minute data 1-minute before the trading time. For example, observing that there is a trade at 10:34:56, we then use the data from 10:23:56 to 10:33:56 to estimate the parameters and predict the logarithmic trading price at 10:34:56, and the last trading price during the time interval from 10:23:56 to 10:33:56 is regarded as $P(0)$. At the same time, we calculate the upper and lower bounds of the prediction at that trading time. We note that even though the drift and volatility parameters in the asymptotic model (25) is constant, the estimated parameters for predictions are actually time-varying. We compare this predicted logarithmic trading prices with the real trading prices in Figure 3. We do a similar prediction for each trading time but using the 10-minute data 2-, 5-, 10-minute before the trading time respectively. The comparisons are shown in Figures 4–6. In each of these figures, there are four curves. The curves at the top and bottom are the upper and lower bounds given by (35). The two curves in the middle are the predicted logarithmic prices, which are computed by (34), and the real logarithmic

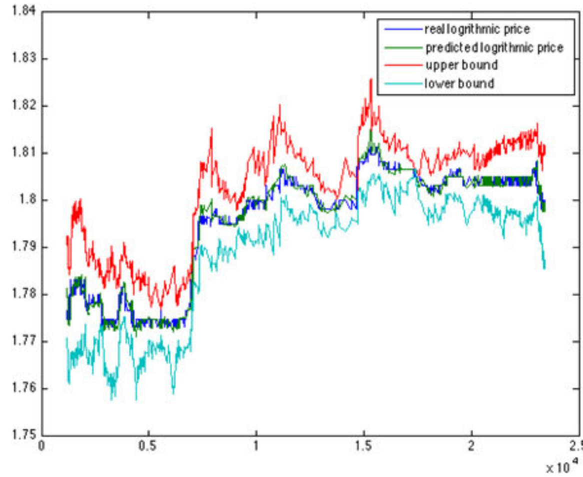

FIGURE 3. Predictions of trading prices using 10-minute data 1-minute before each trading time.

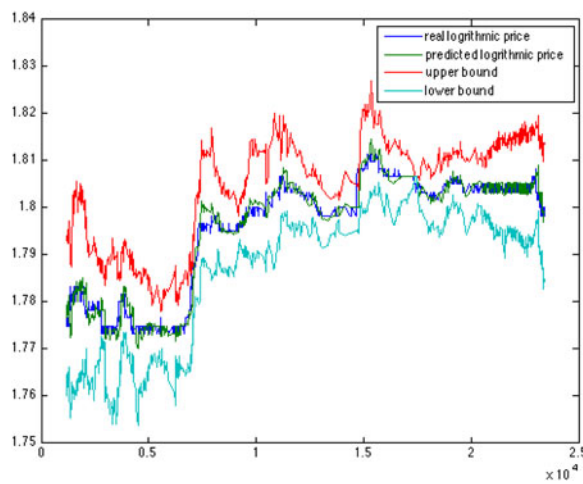

FIGURE 4. Predictions of trading prices using 10-minute data 2-minute before each trading time.

prices. We note that these two curves stay very close, especially in Figure 3. Define

$$\text{Relative error (RE)} = \frac{\text{Real price - Predicted price}}{\text{Real price}}.$$

For the predictions 1-, 2-, 5-, 10-minute into the future, the maximum absolute REs are 0.0055, 0.0058, 0.0080, 0.0152, respectively. We see that the prediction 1-minute into the future provides very good forecasting, and the accuracy of the prediction deteriorates as we try to predict farther into the future, which is to be expected. We note that our asymptotic model is obtained when δ is small. We present the values of $\hat{\delta}^n$ for all four predictions in Figures 6–9, and observe that all values are $\mathcal{O}(10^{-3})$. Thus it is reasonable to use the asymptotic results in the regime $\delta \to 0$.

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

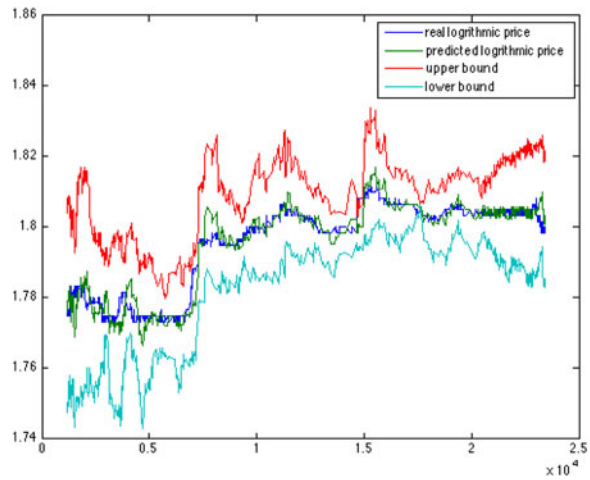

FIGURE **5.** Predictions of trading prices using 10-minute data 5-minute before each trading time.

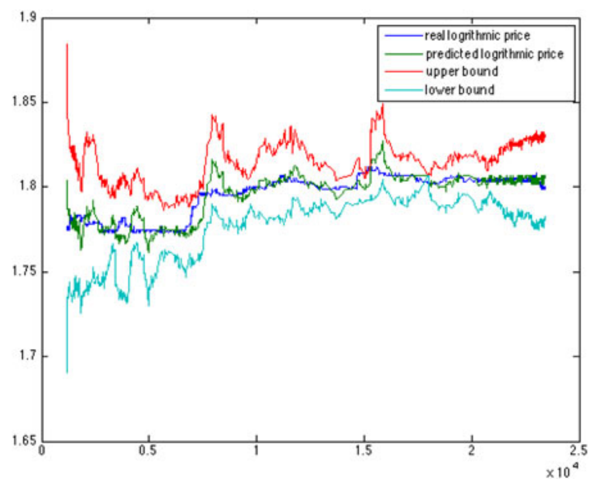

FIGURE **6.** Predictions of trading prices using 10-minute data 10-minute before each trading time.

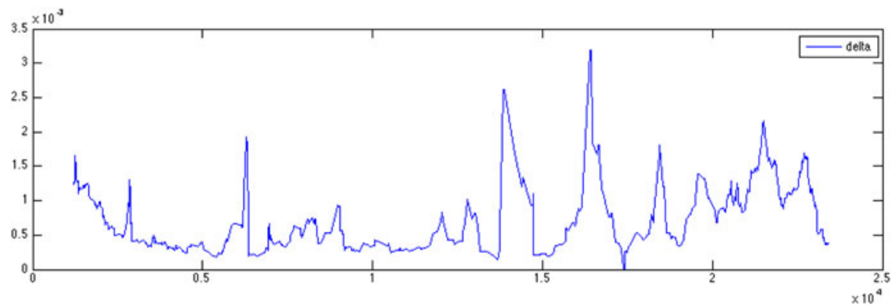

FIGURE **7.** Values of $\hat{\delta}^n$ when using 10-minute data 1-minute before each trading time.

https://doi.org/10.1017/S0269964818000463 Published online by Cambridge University Press

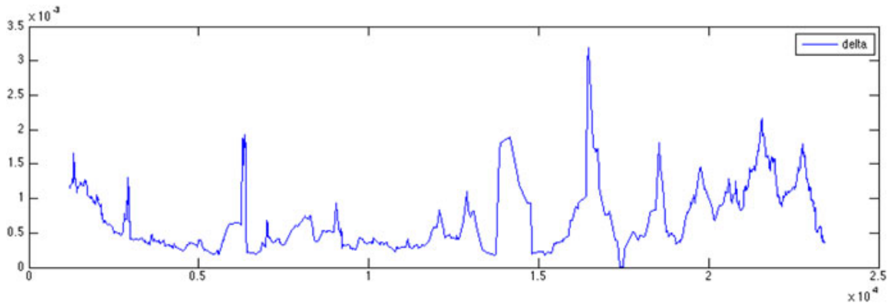

FIGURE 8. Values of $\hat{\delta}^n$ when using 10-minute data 2-minute before each trading time.

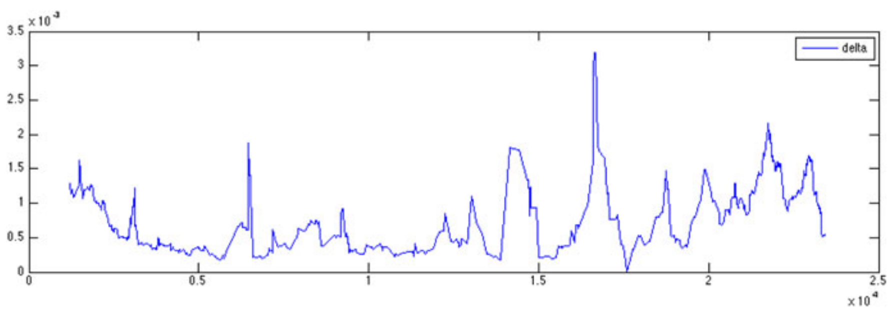

FIGURE 9. Values of $\hat{\delta}^n$ when using 10-minute data 5-minute before each trading time.

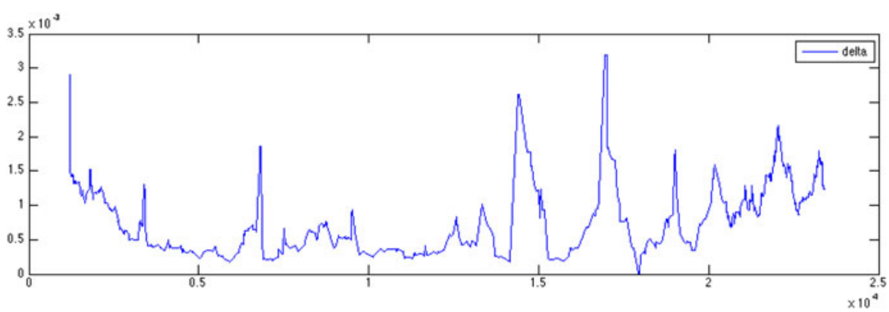

FIGURE 10. Values of $\hat{\delta}^n$ when using 10-minute data 10-minute before each trading time.

References

1. Abergel, F. & Jedidi, A. (2013). A mathematical approach to order book modeling. *International Journal of Theoretical and Applied Finance* 16(05): 1350025.
2. Barndorff–Nielsen, O.E. (1997). Processes of normal inverse Gaussian type. *Finance and Stochastics* 2(1): 41–68.
3. Barndorff–Nielsen, O.E., Mikosch, T., & Resnick, S. I. (2001). *Lévy processes: theory and applications*. Basel, Switzerland: Birkhäuser Basel.
4. Bayer, C., Horst, U., & Qiu, J. (2017). A functional limit theorem for limit order books with state dependent price dynamics. *The Annals of Applied Probability* 27(5): 2753–2806.
5. Bayraktar, E., Horst, U., & Sircar, R. (2007). Queueing theoretic approaches to financial price fluctuations. In Birge, J.R., & Linetsky, V. (eds.), *Handbooks in Operations Research and Management Science* 15: 637–677.
6. Becker–kern, P., Meerschaert, M.M., & Scheffler, H. (2004). Limit theorems for coupled continuous time random walks. *The Annals of Probability* 32(1B): 730–756.

7. Billingsley, P. (1999). *Convergence of probability measures*, 2nd ed. New York: Wiley.
8. Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy* 81(3): 637–654.
9. Brander–Nielsen, O.E. (1994). *Normal inverse Gaussian processes and the modelling of stock returns*. Research report. Department of Mathematical Sciences, Aarhus University.
10. Brander–Nielsen, O.E. (1996). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* 24: 1–13.
11. Brown, M. & Solomon, H. (1975). A second-order approximation for the variance of a renewal reward process. *Stochastic Processes and their Applications* 3(3): 301–314.
12. Burdzy, K. & Nualart, D. (2002). Brownian motion reflected on Brownian motion. *Probability Theory and Related Fields* 122(4): 471–493.
13. Casella, G. & Berger, R. (2002). *Statistical inference*. 2nd ed. Pacific Grove, CA: Duxbury Press.
14. Cont, R. & De Larrard, A. (2013). Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics* 4(1): 1–25.
15. Cont, R. & Larrard, A.D. (2012). Order book dynamics in liquid markets: Limit theorems and diffusion approximations. Preprint, SSRN 1757861.
16. Cont, R., Stoikov, S., & Talreja, R. (2010). A stochastic model for order book dynamics. *Operations Research* 58(3): 549–563.
17. Decamps, M., De Schepper, A., & Goovaerts, M. (2006). A path integral approach to asset-liability management. *Physica A: Statistical Mechanics and its Applications* 363(2): 404–416.
18. Dupuis, P. & Ishii, H. (1991). On lipschitz continuity of the solution mapping to the skorokhod problem, with applications. *Stochastics: An International Journal of Probability and Stochastic Processes* 35(1): 31–62.
19. Horst, U. & Kreher, D. (2017). A weak law of large numbers for a limit order book model with fully state dependent order dynamics. *SIAM Journal on Financial Mathematics* 8(1): 314–343.
20. Horst, U. & Paulsen, M. (2017). A law of large numbers for limit order books. *Mathematics of Operations Research* 42(4): 1280–1312.
21. Iglehart, D.L. & Whitt, W. (1971). The equivalence of functional central limit theorems for counting processes and associated partial sums. *The Annals of Mathematical Statistics* 42(4): 1372–1378.
22. Jacod, J. & Shiryaev, A.N. (2003). *Limit theorem for stochastic processes*. 2nd ed. Berlin: Springer-Verlag.
23. Karatzas, I. & Shreve, S. (1991). *Brownian motion and stochastic calculus* 2nd ed. New York: Springer.
24. Karlin, S. & Taylor, H.M. (1975). *A first course in stochastic processes* 2nd ed. New York: Academic press.
25. Kotulski, M. (1995). Asymptotic distributions of continuous-time random walks: a probabilistic approach. *Journal of Statistical Physics* 81(3–4): 777–792.
26. Kruk, L. (2003). Functional limit theorems for a simple auction. *Mathematics of Operations Research* 28(4): 716–751.
27. Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3(1–2): 125–144.
28. Montroll, E.W. & Weiss, G.H. (1965). Random walks on lattices. II. *Journal of Mathematical Physics* 6(2): 167–181.
29. Rydberg, T.H. (1997). The normal inverse Gaussian Lévy process: simulation and approximation. *Stochastic Models* 13(4): 887–910.
30. Saisho, Y. & Tanaka, H. (1986). Stochastic differential equations for mutually reflecting Brownian balls. *Osaka Journal of Mathematics* 23(3): 725–740.
31. Seshadri, V. (1994). *The inverse Gaussian distribution: a case study in exponential families*. Clarendon Press.