Market impact of trading strategies: An analysis using order book simulation

Tran Hoang Hai and Chua Wee Song

Department of Statistics and Applied Probability
National University of Singapore
The rise of algorithmic trading

Figure 1: Percentage of market volume traded using algo in the US

Market impact of trading strategies: An analysis using order book simulation —
Outline

1. Motivation ✓
2. Market microstructure
3. Market impact estimation

Market impact of trading strategies: An analysis using order book simulation —
A Limit Order Book

Figure 2: A limit order book

Market impact of trading strategies: An analysis using order book simulation —
Types of order

Two main orders:
- Limit orders: buy or sell a stock at a specific price or better.
- Market orders: executed immediately at current market prices.
Market participants

Two main categories:
- Liquidity takers: buy at the ask, sell at the bid.
- Liquidity providers: waits to buy at the bid, sell at the ask.
Market microstructure

Market actions

Spread = 0.02

Market impact of trading strategies: An analysis using order book simulation —
Market microstructure

Market actions

Market impact of trading strategies: An analysis using order book simulation —
Algorithmic trading architecture

Figure 3: Diagram of an algo system

Market impact of trading strategies: An analysis using order book simulation —
Algorithmic trading layers

An algo system normally has 2 layers
- Strategic Layer: Responsible for splitting a large order into blocks of certain intervals, such as every 5 minutes \(\Rightarrow\) Top-down models
- Tactical Layer: Responsible for interacting directly with exchanges and banks to liquidate each block \(\Rightarrow\) Bottom-up models
We focus on a class of strategic models, inspired by Robert Almgren and Neil Chriss in their 2002 paper

Market impact of trading strategies: An analysis using order book simulation —
The Almgren-Chriss framework

Assume a trader trade $q_t$ shares at time $t$ got the price

$$dY_t = k \times q_t \, dt$$

$$dS_t = \sigma \times dW_t + n \times q_t + Y_t$$

where $k$ is the magnitude of the permanent market impact
$n$ is the instant impact factor
The Horst et al., framework

Assume a trader trade $q_t$ shares at time $t$ got the price

$$dY_t = k \times q_t \, dt - b \times Y_t$$

$$dS_t = \sigma \times dW_t + n \times q_t + Y_t$$

where $k$ is the magnitude of the permanent market impact
$n$ is the instant impact factor
$b$ is a resilient factor that reduces the permanent impact

Market impact of trading strategies: An analysis using order book simulation —
Market impact

Both models have this similar market impact components
It’s desirable to be able to estimate those components reliably from data.
Types of market impact

In general, there are 2 types of market impact:
- Temporary impact: arising from the liquidity demands made by execution in a short time
- Permanent impact: Long term price deviation due to trade actions.
- (Almgren et al., 2005) proposed a method of estimation of market impacts using a set of proprietary trade data from Citigroup.

$\Rightarrow$ We aim to replicate the trade data using a simulation of trade execution in a replicated orderbook.
Orderbook modelling

We need to reconstruct three aspects of the orderbook messages:
- Order arrival rate.
- Order size.
- Order price.
In the following, we describe the modelling of each aspects.
Simple Hawkes process

(Hawkes, 1971) proposes an exponential kernel
\[ \nu(t) = \sum_{j=1}^{P} \alpha_j e^{-\beta_j t} 1_{t \in R^+} \]
so that the intensity of the model becomes

\[ \lambda(t) = \lambda_0(t) + \sum_{t_i < t} \sum_{j=1}^{P} \alpha_j e^{-\beta_j (t-t_i)} \]

A simplest version with \( P = 1 \) and \( \lambda_0(t) = \lambda_0 \) constant is defined as:

\[ \lambda(t) = \lambda_0 + \sum_{t_i < t} \alpha e^{-\beta (t-t_i)} \]
Simple Hawkes process with marks

Let $T_n$ be the time sequence of the simple Hawkes process. Let $Z_n$ be another i.i.d. sequence with distribution $Q$, and independent of $N_t$, then the double sequence $T_n, Z_n$ is a simple marked Hawkes process.

Similarly, let $Y_n$ be another i.i.d mark sequence with distribution $H$, then the triple sequence $(T_n, Z_n, Y_n)$ is a double marked Hawkes process.

In our case, $T_n$ will be the time of the limit order arrival, $Z_n$ is the lifespan of the limit order, and $Y_n$ is the associated size of the limit order.
Simple Hawkes process with marks and time-dependent base intensity

Due to the temporal pattern of the limit order liquidity, we can modify the intensity of Hawkes process as follows:

$$\lambda(t) = \lambda_0 \ast \eta(t) + \sum_{t_i < t} \alpha e^{-\beta(t-t_i)}$$

where $\eta(t)$ is a deterministic function represents the intraday variation of the order arrival intensity.
Order price levels

We assume the orderbook’s levels are constant, and each level forms its own population of "shares".
Simulated bid quantity

Figure 5: Simulation of bid quantity
Simulated ask quantity

![Figure 6: Simulation of ask quantity](image)

Market impact of trading strategies: An analysis using order book simulation —
Simulated bid-ask spread

Figure 7: Simulated bid-ask spread

Market impact of trading strategies: An analysis using order book simulation —
Daily adjustment factor

Figure 8: Daily adjustment factor

Market impact of trading strategies: An analysis using order book simulation —
Modified Hawkes process

We propose the following modified version of the Hawkes process to model the market impact following a trade:

$$\lambda(t) = \mu \eta(t) + \sum_{t_i < t} \alpha \exp(-\beta(t-t_i)) + \sum_{t_j^* < t} \alpha \exp(-\beta(t-t_j^*)) \tag{1}$$

where $t_j$ are exogenous trade times generated by our algo.
We modify the rejection sampling procedure in (Ogata, 1981) to accommodate the exogeneous trades as follows:

**Algorithm - Initialization**
- Set $\lambda = \mu$, $n = 1$
- Generate $U \sim U[0, 1]$ and set $s = -\frac{1}{\lambda} \ln U$
- If $s \leq T$ then $t_1 = s$ else exit
Event sampling

- **Algorithm - General**
  - Set $n = n + 1$
  - **Update maximum intensity** Set $\lambda(n)$ according to equation (1)
  - **Generate new event** Generate $U \sim U[0, 1]$ and set $s = -\frac{1}{\lambda} \ln U$
  - If $s \geq T$ then exit
  - **Rejection test** Generate $D \sim U[0, 1]$ then let $t(n) = t_{n-1} + s$
    - If $t(n) > T$ then exit
    - else if $D \leq t(n)/\lambda(n)$ then
      - If there is $t^*_j$ between $t(n)$ and $t(n - 1)$ then
        - $t(n) = \min(t^*_j; t(n - 1) \leq t^*_j \leq t(n))$
  - Repeat

Market impact of trading strategies: An analysis using order book simulation —
A more general Hawkes process

Hawkes process in its general form has the following intensity:

\[ \lambda(t) = \beta(t) + \sum_{n \in \mathbb{Z}} h(t - T_n, Z_n)1_{R}(T_n) \]

\( \beta(t) \) is the external infection rate \( \sum_{n \in \mathbb{Z}} h(t - T_n, Z_n)1_{R}(T_n) \) is the internal contagion rate
**Stability condition**

Hawkes process is stationary, in the sense that there exists a unique process $N^+$ such that

$$P(N(t, \infty) = N + \forall t \geq T) = 1$$

if the following conditions are satisfied:

$$\int_0^\infty E[h(t, Z_t)]dt < 1$$

$$\int_0^\infty \beta(t)dt < \infty$$

$$\int_0^\infty tE[h(t, Z_t)]dt < \infty$$

Market impact of trading strategies: An analysis using order book simulation —
Orderbook model

We propose a two-layer, hybrid class of orderbook model. The macro layer is responsible for simulating the price formation process, including the changing spreads among bid-ask levels. It is extracted from historical prices. The micro layer is responsible for the queuing of different orders into the orderbook and their interactions.
The micro layer

Simulating the interaction of different order types in the orderbook: Levels are constant, and each level forms its own population of "shares. Evolution only depends on the inter- and intra-interation of orders among levels.
Limit orders in the micro layer

A submission of a limit order is considered a birth event. A cancellation of a limit order is considered a death event. Number of limit orders is a birth-death marked Hawkes process, with the marks representing the sizes and the duration of the orders and the intensity function:

$$\lambda_t = \mu_t + \sum_{T_n < t} h(t - T_n, Z_t, V_t)$$
Market orders in the micro layer

Simulate as a marked Hawkes process with general immigrants with intensity functions:

\[ \lambda_t = m(t) + \sum_{S_n < t} \Phi_t(t - S_n, X_n) + \sum_{Q_k < t} \Psi_t(t - Q_k, Y_k) \]

where:

- \( S_n \) are the occurrence times of market orders \( \hat{N}_t \), along with \( X_n \) as their sizes. Following the jump at time \( S_n \), the intensity of the process grows by an amount \( \Phi_t(t - S_n, X_n) \).

- \( Q_k \) are the occurrence times of orders from the trading strategy \( \hat{N}_t \), along with \( Y_k \) as their sizes. Following the jump at time \( Q_k \), the intensity of the process grows by an amount \( \Psi_t(t - Q_k, Y_k) \).

Market impact of trading strategies: An analysis using order book simulation —
Market orders in the micro layer

Figure 9: Market orders

Market impact estimation
Estimation of market impact

- Simulate the microlayer of the orderbook without the strategy trades.
- Generate a random sequence of trades from $T_0$ to $T_1$, where the sizes of all the orders sum up to $K$.
- Simulate the microlayer of the orderbook with the strategy trades.
- Bootstrap the microlayer, both with and without the trading strategy, with the macro layer.
- Calculate the relative spread differences between two simulated orderbook.
Estimation of market impact

Figure 10: Estimation of market impact
Based on the procedure describe above, we simulate the orderbook to execute a random strategy of 1000 trades with total shares of 854,900 (20% of total daily volume of 4,274,461 shares).
Market impact estimation

Figure 11: Market impact estimation
Market impact of trading strategies: An analysis using order book simulation
References

Cartea, A et al
Algorithmic and high-frequency trading

Olivier Gueant

Marco Avellaneda and Sasha Stoikov
High-frequency trading in a limit order book.