Execution probabilities in a limit order book with stochastic order flows

Fenghui Yu

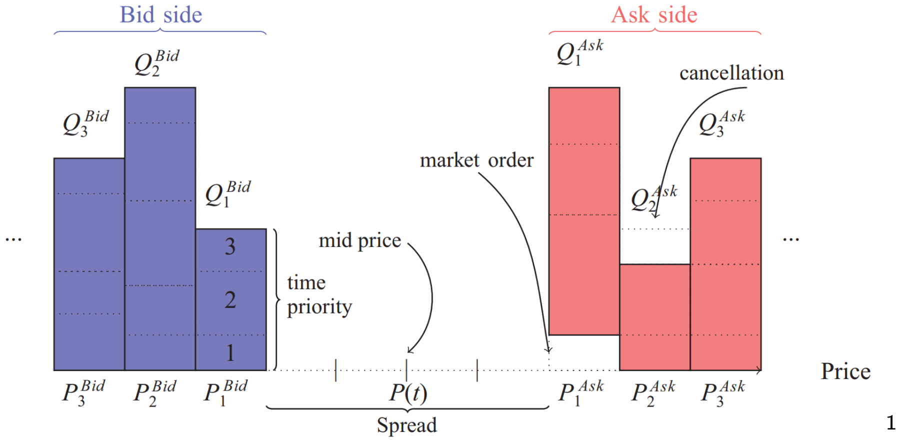
TU Delft

Joint work with Felix Lokin

International Conference on Computational Finance

Amsterdam, April 2nd, 2024

Limit order book



[1] C. Lehalle, O. Mounjid, and M. Rosenbaum. Optimal liquidity-based trading tactics. *Stochastic Systems*. 11(4), 2018.

Execution probabilities

- The execution probability, or fill probability, refers to the likelihood that a limit order is executed.
- This probability is affected by both intrinsic characteristics (price and quantity) of an order, and by external factors (market conditions).
- The order book changes at a very high frequency.
- Accurately predicting the fill probability is a key component of effective algorithmic trading.

What do we do

- We construct a **generic** stochastic order flow model that incorporates state-dependent arrival rates of limit and market orders, as well as cancellations.
- The state-dependent arrival and cancellation rates of orders are generally characterized as functions of stylized factors.
- Although the model and the derived analytical formulas for the execution probabilities are generic, we still provide explicit models that our model covers as examples.
- We conduct extensive numerical experiments using real order book data from the foreign exchange spot market.

Limit order book as a stochastic model

- The price grid is $\{1, ..., N\}$, where the upper boundary N is chosen to be sufficiently large.
- The volume state of the order book is monitored through a continuous-time process

$$\boldsymbol{Q}(t) \equiv (Q_1(t), ..., Q_N(t))_{t \geq 0},$$

where $|Q_i(t)|$ is the number of outstanding limit orders at time t at price level i, with $1 \leq i \leq N$.

Limit order book as a stochastic model

- The price grid is $\{1, \ldots, N\}$, where the upper boundary N is chosen to be sufficiently large.
- The volume state of the order book is monitored through a continuous-time process

$$\boldsymbol{Q}(t) \equiv (Q_1(t), \ldots, Q_N(t))_{t \geq 0},$$

  where $|Q_i(t)|$ is the number of outstanding limit orders at time t at price level i, with $1 \leq i \leq N$.
- To make a distinction between the price levels where bid orders are outstanding and price levels where sell orders are outstanding, the bid levels are denoted by negative quantities.

Dynamics of state-dependent order flows

- The dynamics of the model can be captured by queueing systems, and fully described by the following events:
  - **arrival of limit orders,**
  - **arrival of market orders,**
  - **cancellations of limit orders.**

Dynamics of state-dependent order flows

- The dynamics of the model can be captured by queueing systems, and fully described by the following events:
  - **arrival of limit orders,**
  - **arrival of market orders,**
  - **cancellations of limit orders.**

Dynamics of state-dependent order flows

- The dynamics of the model can be captured by queueing systems, and fully described by the following events:
  - **arrival of limit orders,**
  - **arrival of market orders,**
  - **cancellations of limit orders.**

- By constructing the dynamics of the order book as a sequence of state-dependent queueing systems, we can model the execution times as the first-passage times of the corresponding birth-death and pure-death processes.

Dynamics of state-dependent order flows

- The dynamics of the model can be captured by queueing systems, and fully described by the following events:
  - **arrival of limit orders,**
  - **arrival of market orders,**
  - **cancellations of limit orders.**
- By constructing the dynamics of the order book as a sequence of state-dependent queueing systems, we can model the execution times as the first-passage times of the corresponding birth-death and pure-death processes.
- We assume that all of the events mentioned above are modelled by independent Poisson processes as often seen in the literature of modeling order flows.

- We then assume that
  - **limit orders arrive with rate $\lambda_{Q_i}(\boldsymbol{X}_i)$ at price level i,**

- We then assume that
    - **limit orders arrive with rate** $\lambda_{Q_i}(\boldsymbol{X}_i)$ **at price level** i**,**
    - **market orders arrive at the best bid and best ask with rate** $\mu_{Q_i}(\boldsymbol{X}_i)$ **with** $i = \{p_A, p_B\}$**,**

- We then assume that
  - **limit orders arrive with rate $\lambda_{Q_i}(\boldsymbol{X}_i)$ at price level i,**
  - **market orders arrive at the best bid and best ask with rate $\mu_{Q_i}(\boldsymbol{X}_i)$ with $i = \{p_A, p_B\}$,**
  - **and cancellation rate of limit orders at price level i to be $\phi_{Q_i}(\boldsymbol{X}_i)$,**

- We then assume that
  - **limit orders arrive with rate $\lambda_{Q_i}(\boldsymbol{X}_i)$ at price level i,**
  - **market orders arrive at the best bid and best ask with rate $\mu_{Q_i}(\boldsymbol{X}_i)$ with $i = \{p_A, p_B\}$,**
  - **and cancellation rate of limit orders at price level i to be $\phi_{Q_i}(\boldsymbol{X}_i)$,**

  where $\boldsymbol{X}_i$ is the vector of stylized factors that the arrival and cancellation rates of orders depend on, and $1 \leq i \leq N$.

- We then assume that
  - **limit orders arrive with rate $\lambda_{Q_i}(\boldsymbol{X}_i)$ at price level i,**
  - **market orders arrive at the best bid and best ask with rate $\mu_{Q_i}(\boldsymbol{X}_i)$ with $i = \{p_A, p_B\}$,**
  - **and cancellation rate of limit orders at price level i to be $\phi_{Q_i}(\boldsymbol{X}_i)$,**

  where $\boldsymbol{X}_i$ is the vector of stylized factors that the arrival and cancellation rates of orders depend on, and $1 \leq i \leq N$.

Schematic representation of the order book dynamics

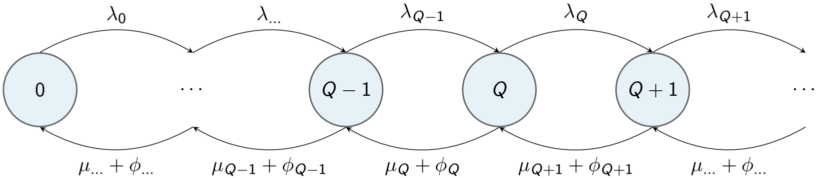


Figure: At the best ask $Q = Q_{p_A}$.

Schematic representation of the order book dynamics

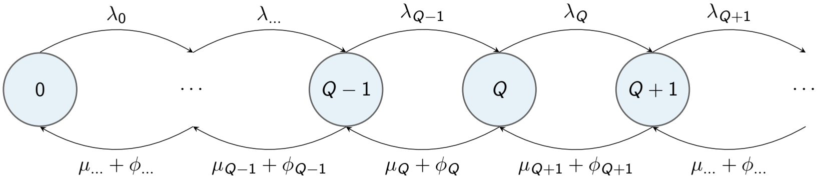


Figure: At the best ask $Q = Q_{p_A}$.

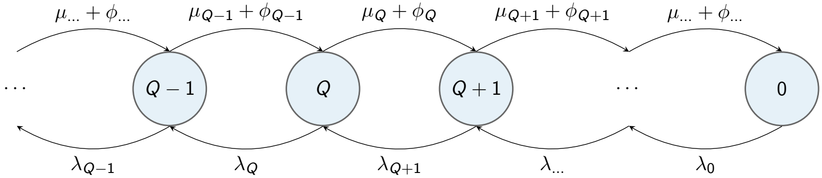


Figure: At the best bid $Q = Q_{p_B}$.

Execution probability at the best quotes

- We now consider the fill probability of an order placed at the best ask or best bid before the mid-price moves, given that it is never cancelled.

Execution probability at the best quotes

- We now consider the fill probability of an order placed at the best ask or best bid before the mid-price moves, given that it is never cancelled.
- The conditional probability that an order placed at the best quote is executed before the mid-price moves is given by

$$\mathbb{P}[\epsilon_i < \tau \mid Q_A(0) = q_0^A, Q_B(0) = q_0^B, S(0) = s_0, NC_i], \tag{1}$$

where $i \in \{A, B\}$ and ϵ_i denotes the first-passage time at which a pure-death process reaches 0, given that it started in state q_0^i.

Execution probability at the best quotes

- We now consider the fill probability of an order placed at the best ask or best bid before the mid-price moves, given that it is never cancelled.
- The conditional probability that an order placed at the best quote is executed before the mid-price moves is given by

$$\mathbb{P}[\epsilon_i < \tau \mid Q_A(0) = q_0^A, Q_B(0) = q_0^B, S(0) = s_0, NC_i], \tag{1}$$

where $i \in \{A, B\}$ and ϵ_i denotes the first-passage time at which a pure-death process reaches 0, given that it started in state q_0^i.

- The execution probability of Equation (1) is then equivalent to

$$\mathbb{P}[\epsilon_i < \sigma_j \wedge \tau_A \wedge \tau_B] = \mathbb{P}[\epsilon_i - \sigma_j \wedge \tau_A \wedge \tau_B < 0]. \tag{2}$$

Computation formulas

Proposition *(Lokin and Y. '24)*

Let $\hat{f}^{s_0}_{\sigma_i}(s)$ denote the Laplace transform of the pdf of σ_i given s_0, we have that for the Laplace transform of the pdf of ϵ_i, denoted by $\hat{g}^{s_0}_{\epsilon_i}$, is

$$\hat{g}^{s_0}_{\epsilon_i}(s) = \prod_{k=1}^{q_0^i} \frac{\mu_k(\boldsymbol{X}_{p_i}) + \phi_k(\boldsymbol{X}_{p_i})}{\mu_k(\boldsymbol{X}_{p_i}) + \phi_k(\boldsymbol{X}_{p_i}) + s}, \tag{3}$$

where $q_i \geq 0$, $i \in \{A, B\}$, and $s_0 \geq 1$.
Let $\Lambda_{s_0} = \sum_{m=1}^{s_0-1} \lambda_0(\boldsymbol{X}_{p_A-m}) = \sum_{m=1}^{s_0-1} \lambda_0(\boldsymbol{X}_{p_B+m})$ again, then the fill probability given by Equation (1) can be calculated by the inverse Laplace transform of

$$\hat{F}^{s_0}_{\epsilon_i,\sigma_j}(s) = \frac{1}{s}\hat{g}^{s_0}_{\epsilon_i}(s)\left(\hat{f}^{s_0}_{\sigma_j}(2\Lambda_{s_0} - s) + \frac{2\Lambda_{s_0}}{2\Lambda_{s_0} - s}(1 - \hat{f}^{s_0}_{\sigma_j}(2\Lambda_{s_0} - s))\right), \tag{4}$$

evaluating at 0 for $i \neq j \in \{A, B\}$. Note that when $s_0 = 1$, we have

$$\hat{F}^{1}_{\epsilon_i,\sigma_j}(s) = \frac{1}{s}\hat{g}^{1}_{\epsilon_i}(s)\hat{f}^{1}_{\sigma_j}(-s). \tag{5}$$

Execution probability at a price level deeper than the best quotes

- We consider the case of a limit order posted at the price level deeper than the best quote, $p_A + 1$ and $p_B - 1$, before the opposite best quote price moves.
- The intuition can be extended to price levels even deeper in the order book.

Execution probability at a price level deeper than the best quotes

- We consider the case of a limit order posted at the price level deeper than the best quote, $p_A + 1$ and $p_B - 1$, before the opposite best quote price moves.
- The intuition can be extended to price levels even deeper in the order book.
- For an order to be executed when it is not submitted at the best quote price, the best quote price should move towards this order such that the order is at the new best quote.

Execution probability at a price level deeper than the best quotes

- We consider the case of a limit order posted at the price level deeper than the best quote, $p_A + 1$ and $p_B - 1$, before the opposite best quote price moves.
- The intuition can be extended to price levels even deeper in the order book.
- For an order to be executed when it is not submitted at the best quote price, the best quote price should move towards this order such that the order is at the new best quote.
- We define τ_i^{quote} to be the first time of a change in mid-price as a result of the best quote price has moved to the price level of the concerned order,

$$\tau_i^{\text{quote}} \equiv \begin{cases} \inf\{t \geq 0 : p_A(t) > p_A(0)\}, & \text{for} \quad i = A, \\ \inf\{t \geq 0 : p_B(t) < p_B(0)\}, & \text{for} \quad i = B. \end{cases} \tag{6}$$

- Let τ_i^{other} be the first time of a change in mid-price as a result of a different event,

$$\tau_i^{\text{other}} \equiv \begin{cases} \inf\left\{t \geq 0 : \Big(p_A(t) < p_A(0)\Big) \wedge \Big(p_B(t) \neq p_B(0)\Big)\right\}, & \text{for} \quad i = A, \\ \inf\left\{t \geq 0 : \Big(p_B(t) > p_B(0)\Big) \wedge \Big(p_A(t) \neq p_A(0)\Big)\right\}, & \text{for} \quad i = B. \end{cases} \tag{7}$$

- Let τ_i^{other} be the first time of a change in mid-price as a result of a different event,

$$\tau_i^{\text{other}} \equiv \begin{cases} \inf\left\{t \geq 0 : \Big(p_A(t) < p_A(0)\Big) \wedge \Big(p_B(t) \neq p_B(0)\Big)\right\}, & \text{for} \quad i = A, \\ \inf\left\{t \geq 0 : \Big(p_B(t) > p_B(0)\Big) \wedge \Big(p_A(t) \neq p_A(0)\Big)\right\}, & \text{for} \quad i = B. \end{cases} \tag{7}$$

- The probability that best quote price moves towards the price level where the concerned order is at before the mid-price moves due to other events, is then given by

$$\mathbb{P}[\tau_i^{\text{quote}} < \tau_i^{\text{other}} \mid Q_A(0) = q_0^A, Q_B(0) = q_0^B, S(0) = s_0]. \tag{8}$$

- Let τ_i^{other} be the first time of a change in mid-price as a result of a different event,

$$\tau_i^{\text{other}} \equiv \begin{cases} \inf\left\{t \geq 0 : \Big(p_A(t) < p_A(0)\Big) \wedge \Big(p_B(t) \neq p_B(0)\Big)\right\}, & \text{for} \quad i = A, \\ \inf\left\{t \geq 0 : \Big(p_B(t) > p_B(0)\Big) \wedge \Big(p_A(t) \neq p_A(0)\Big)\right\}, & \text{for} \quad i = B. \end{cases} \tag{7}$$

- The probability that best quote price moves towards the price level where the concerned order is at before the mid-price moves due to other events, is then given by

$$\mathbb{P}[\tau_i^{\text{quote}} < \tau_i^{\text{other}} \mid Q_A(0) = q_0^A, Q_B(0) = q_0^B, S(0) = s_0]. \tag{8}$$

- Let τ^i be the time of the first change in mid-price after τ_i^{quote}, i.e.

$$\tau^i \equiv \inf\{t \geq \tau_i^{\text{quote}} : p_M(t) \neq p_M(\tau_i^{\text{quote}})\} - \tau_i^{\text{quote}}.$$

- Let τ_i^{other} be the first time of a change in mid-price as a result of a different event,

$$\tau_i^{\text{other}} \equiv \begin{cases} \inf\left\{t \geq 0 : \left(p_A(t) < p_A(0)\right) \wedge \left(p_B(t) \neq p_B(0)\right)\right\}, & \text{for} \quad i = A, \\ \inf\left\{t \geq 0 : \left(p_B(t) > p_B(0)\right) \wedge \left(p_A(t) \neq p_A(0)\right)\right\}, & \text{for} \quad i = B. \end{cases} \tag{7}$$

- The probability that best quote price moves towards the price level where the concerned order is at before the mid-price moves due to other events, is then given by

$$\mathbb{P}[\tau_i^{\text{quote}} < \tau_i^{\text{other}} \mid Q_A(0) = q_0^A, Q_B(0) = q_0^B, S(0) = s_0]. \tag{8}$$

- Let τ^i be the time of the first change in mid-price after τ_i^{quote}, i.e.

$$\tau^i \equiv \inf\{t \geq \tau_i^{\text{quote}} : p_M(t) \neq p_M(\tau_i^{\text{quote}})\} - \tau_i^{\text{quote}}.$$

- Given $\mathbb{P}[\tau_i^{\text{quote}} < \tau_i^{\text{other}}]$, the execution probability is

$$\mathbb{P}[\epsilon_{i-} < \tau^i \mid W_{i-}(\tau_i^{\text{quote}}) = q^{i-}_{\tau_i^{\text{quote}}}, Q_j(\tau_i^{\text{quote}}) = q^j_{\tau_i^{\text{quote}}}, S(\tau_i^{\text{quote}}) = s_0 + 1, NC_{i-}], \tag{9}$$

where $i \neq j \in \{A, B\}$.

Computation formulas

Proposition *(Lokin and Y. '24)*

The Laplace transform of the density function of the first-passage time σ_i of a birth-death process of $\tilde{Q}_i$ to 0 given s_0 for $i \in \{A, B\}$, denoted by $\hat{f}^{s_0}_{\sigma_i}(s)$, with $s_0 \geq 1$.
Again let $\Lambda_{s_0} = \sum_{m=1}^{s_0-1} \lambda_0(\boldsymbol{X}_{p_A - m}) = \sum_{m=1}^{s_0-1} \lambda_0(\boldsymbol{X}_{p_B + m})$, then Probability (8) can be calculated by the inverse Laplace transform of

$$\hat{G}^{s_0}_{\sigma_i,\sigma_j}(s) = \frac{1}{s}\hat{f}^{s_0}_{\sigma_i}(s)\left(\hat{f}^{s_0}_{\sigma_j}(2\Lambda_{s_0} - s) + \frac{2\Lambda_{s_0}}{2\Lambda_{s_0} - s}(1 - \hat{f}^{s_0}_{\sigma_j}(2\Lambda_{s_0} - s))\right), \tag{10}$$

evaluated at 0 for $i \neq j \in \{A, B\}$. In particular, if $s_0 = 1$, we have

$$\hat{G}^{1}_{\sigma_i,\sigma_j}(s) = \frac{1}{s}\hat{f}^{1}_{\sigma_i}(s)\hat{f}^{1}_{\sigma_j}(-s). \tag{11}$$

Computation formulas

Proposition *(Lokin and Y. '24)*

The Laplace transform of the density function of the first-passage time σ_i of a birth-death process of $\tilde{Q}_i$ to 0 given s_0 for $i \in \{A, B\}$, denoted by $\hat{f}^{s_0}_{\sigma_i}(s)$, with $s_0 \geq 1$.
Again let $\Lambda_{s_0} = \sum_{m=1}^{s_0-1} \lambda_0(\boldsymbol{X}_{p_A-m}) = \sum_{m=1}^{s_0-1} \lambda_0(\boldsymbol{X}_{p_B+m})$, then Probability (8) can be calculated by the inverse Laplace transform of

$$\hat{G}^{s_0}_{\sigma_i,\sigma_j}(s) = \frac{1}{s}\hat{f}^{s_0}_{\sigma_i}(s)\left(\hat{f}^{s_0}_{\sigma_j}(2\Lambda_{s_0} - s) + \frac{2\Lambda_{s_0}}{2\Lambda_{s_0} - s}(1 - \hat{f}^{s_0}_{\sigma_j}(2\Lambda_{s_0} - s))\right), \tag{10}$$

evaluated at 0 for $i \neq j \in \{A, B\}$. In particular, if $s_0 = 1$, we have

$$\hat{G}^{1}_{\sigma_i,\sigma_j}(s) = \frac{1}{s}\hat{f}^{1}_{\sigma_i}(s)\hat{f}^{1}_{\sigma_j}(-s). \tag{11}$$

Remark

Once we have the explicit expressions for the transition rates $\lambda_{Q_i}(\boldsymbol{X}_{p_i}), \mu_{Q_i}(\boldsymbol{X}_{p_i}), \mu_{Q_{i-}}(\boldsymbol{X}_{p_{i-}}), \phi_{Q_i}(\boldsymbol{X}_{p_i})$, and $\phi_{Q_{i-}}(\boldsymbol{X}_{p_{i-}})$ for $i = \{A, B\}$, we can then obtain the corresponding formulas to calculate all the execution probabilities according to the propositions.

Explicit stochastic intensity model

Based on the assumptions in *Cont et al. '10* and our empirical analysis, we have the following explicit arrival, cancellation rates of limit and market orders at time t_j in the numerical experiments:

$$\begin{cases} \lambda_{Q_i}(\boldsymbol{X}_i) = \lambda(p_A(t_j) - i, S(t_j)) & \text{for} \quad i < p_A(t_j), \\ \lambda_{Q_i}(\boldsymbol{X}_i) = \lambda(i - p_B(t_j), S(t_j)) & \text{for} \quad i > p_B(t_j), \\ \mu_{Q_i}(\boldsymbol{X}_i) = \mu(S(t_j)) & \text{for} \quad i = p_A(t_j) \text{ or } i = p_B(t_j), \\ \phi_{Q_i}(\boldsymbol{X}_i) = \theta(i - p_B(t_j), S(t_j))|Q_i(t_j)| & \text{for} \quad i \geq p_A(t_j), \\ \phi_{Q_i}(\boldsymbol{X}_i) = \theta(p_A(t_j) - i, S(t_j))|Q_i(t_j)| & \text{for} \quad i \leq p_B(t_j), \end{cases}$$

where $\lambda, \theta \colon \mathbb{N} \times \mathbb{N} \to \mathbb{R}_+$ are functions of the price and spread, and $\mu \colon \mathbb{N} \to \mathbb{R}_+$ is a function of the spread.

Empirical limit order executions

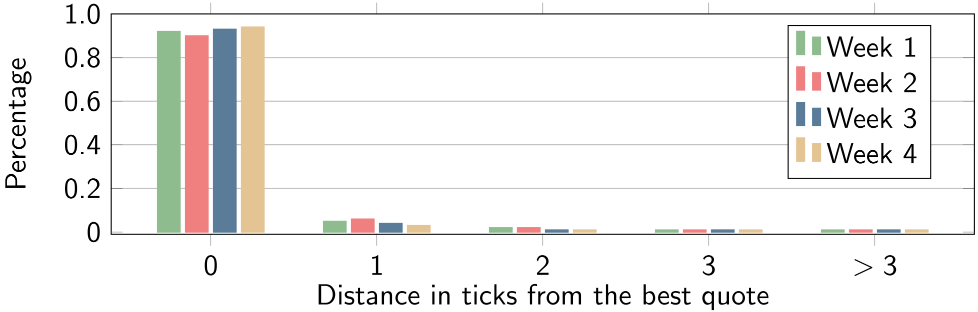


Figure: Distribution of executed limit orders based on their distance in ticks from the best quote at the time of execution.

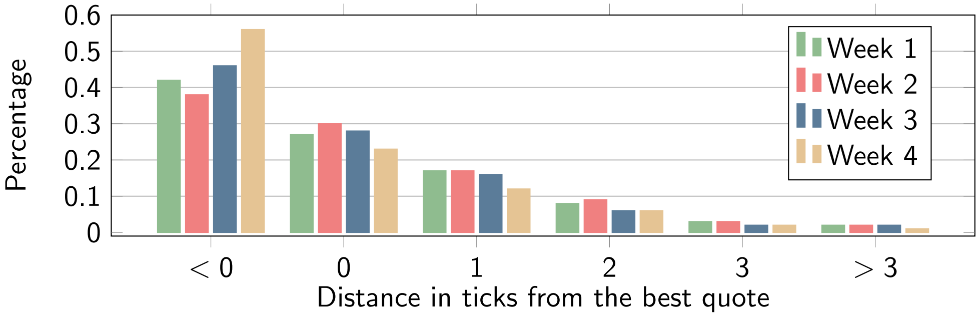


Figure: Distribution of executed limit orders based on their distance in ticks from the best quote at the time of submission.

Empirical intensity rates

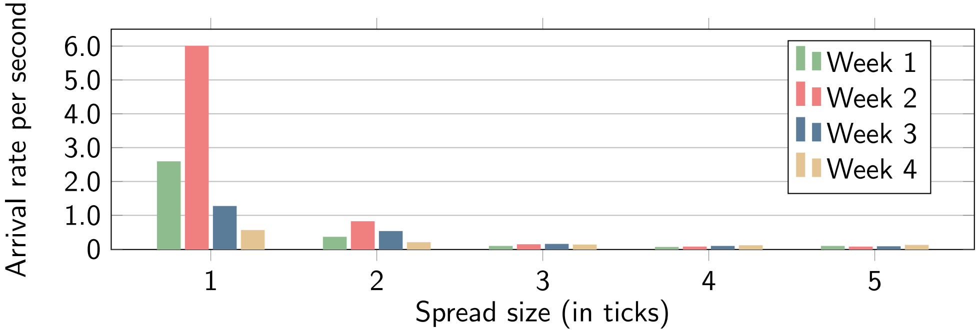


Figure: Arrival rates per second for market orders for each spread size between one and five ticks.

Probability of an increase in mid-price

		Empirical Probability					Model Probability				
		q_A					q_A				
	q_B	1	2	3	4	5	1	2	3	4	5
	1	50.3%	33.0%	22.9%	27.1%	22.4%	50.0%	34.7%	27.1%	22.6%	19.6%
	2	70.5%	56.6%	-	-	-	65.3%	50.0%	41.1%	35.3%	31.1%
$S = 1$	3	78.7%	-	-	-	-	72.9%	58.9%	50.0%	43.8%	39.2%
	4	78.2%	-	-	-	-	77.4%	64.7%	56.2%	50.0%	45.3%
	5	81.4%	-	-	-	-	80.5%	68.9%	60.8%	54.8%	50.0%
	1	49.6%	38.5%	32.2%	25.2%	22.0%	50.0%	36.8%	30.9%	27.6%	25.6%
	2	58.1%	48.5%	52.3%	45.5%	27.5%	63.3%	50.0%	43.2%	39.1%	36.4%
$S = 2$	3	70.1%	49.9%	-	-	-	69.2%	56.8%	50.0%	45.7%	42.7%
	4	75.1%	43.9%	-	-	-	72.4%	60.9%	54.4%	50.0%	47.0%
	5	81.6%	-	-	-	-	74.5%	63.7%	57.3%	53.1%	50.0%

Execution probability at the best quotes

	q_B	Empirical Probability q_A 1	2	3	4	5	Model Probability q_A 1	2	3	4	5
$S=1$	1	2.0%	5.3%	7.1%	13.4%	-	3.0%	3.9%	4.6%	5.1%	5.5%
	2	0.6%	0.0%	-	-	-	1.9%	2.7%	3.2%	3.7%	4.0%
	3	0.3%	-	-	-	-	1.5%	2.2%	2.6%	3.0%	3.3%
	4	0.0%	-	-	-	-	1.2%	1.8%	2.2%	2.6%	2.9%
	5	0.0%	-	-	-	-	0.9%	1.5%	1.9%	2.2%	2.5%
$S=2$	1	1.3%	2.5%	4.1%	5.9%	5.9%	1.5%	1.8%	2.0%	2.1%	2.2%
	2	0.3%	0.1%	0.0%	-	-	1.0%	1.2%	1.3%	1.4%	1.5%
	3	0.3%	1.6%	-	-	-	0.7%	0.9%	1.0%	1.1%	1.2%
	4	0.2%	-	-	-	-	0.6%	0.8%	0.9%	1.0%	1.0%
	5	0.0%	0.0%	0.0%	-	-	0.5%	0.7%	0.8%	0.8%	0.9%

Execution probability at a price deeper

	q_B	Empirical Probability q_A 1	2	3	4	5	Model Probability q_A 1	2	3	4	5
$q_{B-} = 1$	1	0.19%	0.27%	0.68%	2.16%	-	0.75%	0.97%	1.09%	1.16%	1.21%
	2	0.23%	0.25%	0.39%	0.67%	-	0.53%	0.75%	0.88%	0.97%	1.03%
	3	0.11%	0.69%	-	-	-	0.42%	0.62%	0.75%	0.84%	0.91%
	4	0.12%	-	-	-	-	0.35%	0.54%	0.66%	0.75%	0.82%
$q_{B-} = 2$	1	0.06%	0.44%	-	-	-	0.63%	0.81%	0.91%	0.96%	1.00%
	2	0.12%	0.37%	0.00%	-	-	0.47%	0.67%	0.78%	0.86%	0.91%
	3	0.12%	0.00%	-	-	-	0.38%	0.57%	0.69%	0.77%	0.83%
	4	0.00%	-	-	-	-	0.32%	0.50%	0.61%	0.70%	0.76%

Thank you for listening!

- F. Lokin and F. Yu. Fill probabilities in a limit order book with state-dependent stochastic order flows. *https://arxiv.org/abs/2403.02572*, 2024.