Appendix A. Proofs

Proof of Proposition 1

We verify that the strategies described in Proposition 1 for non-HFTs and HFTs are their best responses.

First, it is optimal for non-HFTs to submit market orders immediately upon arrival. Because the half-spread h_0 is stationary and v_t is a martingale, there is no benefit for anyone to wait.

Second, no HFT would deviate from the quoted bid-ask spread at $v_t \pm h_0$:

1. Any HFT who crosses the midpoint (sells below v_t or buys above v_t) loses money.
2. Liquidity provider(s) and snipers earn the same expected profits for each share in the LOB. Quoting a half-spread narrower than h_0 is not a profitable deviation because this strategy earns less than quoting h_0 or sniping at h_0.
3. Quoting at $v_t \pm h_0$ after an existing limit order is not a profitable deviation because a share at a later queue position is less likely to execute and more likely to be sniped. For example, there is a probability of $(1-\pi)^2$ that the second share at $v_t + h_0$ meets a non-HFT conditional on two buy events because it needs to wait for the first share to be executed, while the probability of being sniped is $1-(1-\pi)^2$.
4. The equilibrium half-spread cannot be wider than h_0. When the half-spread is wider than h_0, the expected profit of a sniper is strictly lower than that of a liquidity provider. Thus, a sniper would immediately quote a tighter spread. Therefore, a liquidity provider's quote in $(h_0, 1)$ can execute only during value jumps and loses money. A liquidity provider quoting a spread wider than one never sets the best bid-ask spread and will not affect the equilibrium outcome.

5. It is easier to show that a stale quote sniper has no incentive to deviate. During value jumps, all traders aim to snipe stale quotes, either because it is profitable to do so or because liquidity providers want to snipe their own quotes (equivalent to cancellation) to reduce their losses. Finally, steps 1-4 show that snipers have no incentive to deviate by providing liquidity. ■

Proof of Proposition 2

The difference between Proposition 2 and Proposition 1 is that a fraction β of non-HFTs, execution algorithms (EAs), can use limit orders to minimize their transaction costs. Thus, the equilibrium half-spread h_β is given in Eq. (4).

Second, submitting limit orders at v_t (stimulating orders) and facing zero transaction costs is the best outcome for EAs. All other execution strategies would lead to non-negative transaction costs (otherwise an HFT could deviate to such a strategy and earn a profit). For example, EAs who cross the midpoint always incur an instant positive transaction cost. For EAs who narrow the $v_t \pm h_\beta$ bid-ask spread by posting sell limit orders at $v_t + h$ or buy limit orders at $v_t - h$, where $h < h_\beta$, there is a probability of $\frac{(1-\beta)(1-\pi)}{(1-\beta)(1-\pi)+\pi}$ that these EAs meet an MOT and a probability of $\frac{\pi}{(1-\beta)(1-\pi)+\pi}$ that these EAs meet a sniper conditional on execution. Thus, the execution cost $C(h)$ is

$$-\left[\frac{(1-\beta)(1-\pi)}{(1-\beta)(1-\pi)+\pi} \cdot h - \frac{\pi}{(1-\beta)(1-\pi)+\pi} \cdot (1-h)\right] > -\left[\frac{(1-\beta)(1-\pi)}{(1-\beta)(1-\pi)+\pi} \cdot h_\beta - \frac{\pi}{(1-\beta)(1-\pi)+\pi} \cdot (1-h_\beta)\right] = 0. \quad \text{(A.1)}$$

Thus, we have $C(h) > 0$ when $h < h_\beta$. Note that in Eq. (A.1) we used the assumption that EAs are slower than HFTs. Thus, when a jump occurs, orders from EAs are always sniped. Moreover, $C(h_\beta) = 0$ means that EAs incur zero transaction costs if they provide the first unit at $v_t \pm h_\beta$.

Therefore, they will incur positive transaction costs if they queue after an existing limit order at $v_t \pm h_\beta$.[1] EAs who quote limit orders wider than $v_t \pm h_\beta$ can trade only with snipers because each non-HFT trades only one share and other HFTs will refill the liquidity-provision share after the share has been consumed by a non-HFT. A quote outside $v_t \pm 1$ is ruled out by assumption.[2]

Third, HFTs that accept an EA's order at v_t receive zero payoffs and no deviation can offer a positive profit. Also, for the same reason as with Proposition 1, no HFT can earn a greater payoff than other HFTs on the shares quoted by HFT(s) at $v_t \pm h_\beta$.

To summarize, no market participant can receive a higher payoff by deviating from the strategy defined in Proposition 2. Thus, Proposition 2 is an equilibrium. ■

Proof of Corollary 1

All the results follow directly by taking the derivative of h_β and $\bar{C}(\beta)$ with respect to β:

$$\frac{dh_\beta}{d\beta} = \frac{\pi(1-\pi)}{(1-\beta(1-\pi))^2} > 0, \tag{A.2}$$

$$\frac{d\bar{C}(\beta)}{d\beta} = \frac{-\pi^2}{(1-\beta(1-\pi))^2} < 0. \tag{A.3}$$

Thus, the half-spread h_β increases in β and the average transaction cost $\bar{C}(\beta)$ decreases in β.■

Proof of Proposition 3

When the bid-ask spread is binding at one tick, all non-HFTs take liquidity from HFTs. Thus, as with Proposition 1, and replacing N with ∞, HFTs provide liquidity at $v_t \pm \frac{\Delta}{2}$ if

---

[1] When pricing is discrete, our Assumption 1 requires that all limit orders be price-improving, which is not a binding constraint here.

[2] Again, there is no way that EAs can enjoy negative transaction costs even without this assumption; otherwise an HFT can mimic the EA's strategy and earn a profit.

$$(1-\pi)\cdot\frac{\Delta}{2}-\pi\cdot\left(1-\frac{\Delta}{2}\right)\geq 0, \tag{A.4}$$

which is equivalent to $\pi \leq \frac{\Delta}{2}=\frac{1}{2L}$. No HFT wants to cancel her order and give up the above positive liquidity provision profits except when a fundamental value jump occurs. For the same reason, all HFTs compete to refill a liquidity provision order once it is consumed. No HFT can quote a spread narrower than Δ because of the tick-size constraint. A spread wider than Δ can never be the best bid-ask spread. All HFTs want to snipe stale quotes during value jumps; otherwise, the quote would be immediately sniped by other HFTs. Thus, Proposition 3 is an equilibrium. ■

Proof of Lemma 1

Denote the expected execution cost of undercutting h_β by k ticks at $v_t \pm \left(m+\frac{1}{2}-k\right)\Delta$ as $C(k)$. We show that $C(k)$ is larger than $\frac{\Delta}{2}$ for any $2 \leq k \leq m-1$ and $k \in \mathbb{N}^+$.

The limit order at $v_t \pm \left(m+\frac{1}{2}-k\right)\Delta$, where $k<m$, does not attract EAs.[3] Consider the best case where the order has never been further undercut, i.e. it can receive all future MOT order flows. Conditional on the order's execution, there is a probability of $\frac{(1-\beta)(1-\pi)}{(1-\beta)(1-\pi)+\pi}$ that the limit order trades with an MOT order and a probability of $\frac{\pi}{(1-\beta)(1-\pi)+\pi}$ that it trades with a sniper. Since quoted half-spread is $(m+\frac{1}{2}-k)\Delta$, the expected execution cost is

$$C(k)\geq-\frac{(1-\beta)(1-\pi)}{(1-\beta)(1-\pi)+\pi}\left(m+\frac{1}{2}-k\right)\Delta+\frac{\pi}{(1-\beta)(1-\pi)+\pi}\left(1-\left(m+\frac{1}{2}-k\right)\Delta\right)=\frac{\pi}{(1-\beta)(1-\pi)+\pi}-\left(m+\frac{1}{2}-k\right)\Delta. \tag{A.5}$$

Note that we have $\frac{\pi}{(1-\beta)(1-\pi)+\pi}=\frac{\pi}{1-\beta(1-\pi)}=h_\beta$ (Proposition 2). Therefore, $C(k)\geq h_\beta -$

[3] Again, this is because opposite-side EAs could always choose to trade immediately at $v_t \pm \frac{\Delta}{2}$, so EAs will not accept worse prices.

$\left(m+\frac{1}{2}-k\right)\Delta = h_\beta - \left(m-\frac{1}{2}\right)\Delta + (k-1)\Delta$. Since $h_\beta - \left(m-\frac{1}{2}\right)\Delta > 0$,[4] we have $C(k) > \Delta$ when $2 \leq k \leq m-1$ and $k \in \mathbb{N}^+$. ■

Proof of Propositions 4 and 5

The main task in constructing the proof for Proposition 4 and Proposition 5 is to determine the boundaries where EAs are indifferent between stimulating HFTs and undercutting HFTs by one tick. Therefore, we combine the proofs of both propositions.

HFTs' quoted spread

When HFTs' quotes are not binding at one tick (hence outside of $v_t \pm \frac{\Delta}{2}$), EAs never take liquidity from HFTs. When no EAs undercut HFTs, an HFT's breakeven half-spread is the $h_\beta = \frac{\pi}{1-\beta(1-\pi)}$ determined in Proposition 2, as only MOTs will take HFTs' orders. HFTs always post one unit to sell at the tick grid point immediately above $v_t + h_\beta$ and post one unit to buy at the tick grid point immediately below $v_t - h_\beta$:

Lemma 2: *When $\pi > \frac{\Delta}{2}$, HFTs almost always maintain one share at the ask price $v_t + (m+\frac{1}{2})\Delta$ and one share at the bid price $v_t - (m+\frac{1}{2})\Delta$ if they are not undercut by EAs, where m is the number of tick grid points strictly in the interval $(v_t, v_t + h_\beta)$.*

Note that the first, second, . . ., and m^{th} ticks above v_t is $v_t + \frac{\Delta}{2}$, $v_t + \frac{3\Delta}{2}$, $\cdots$, $v_t + \left(m-\frac{1}{2}\right)\Delta$, and hence the tick grid point immediately above $v_t + h_\beta$ is $v_t + \left(m+\frac{1}{2}\right)\Delta$. Eq. (7) in Section 5.2.2 characterizes HFT's optimal response if EAs choose to undercut HFTs.

[4] By definition, $v_t + \left(m-\frac{1}{2}\right)\Delta$ is the tick grid point immediately below $v_t + h_\beta$, so $h_\beta - \left(m-\frac{1}{2}\right)\Delta > 0$.

EAs' strategy: The boundary between stimulating and undercutting equilibria

We first define an EA's order-placement strategy.

Definition 1: *The order-placement strategy for an EA seller and an EA buyer are denoted as function* $\phi^s: (i_t^1, i_t^2, \cdots, i_t^m, j_t^m, \cdots, j_t^2, j_t^1) \rightarrow \{v_t - \frac{\Delta}{2}, v_t + \frac{\Delta}{2}, v_t + \frac{3\Delta}{2}, \cdots, v_t + \left(m - \frac{1}{2}\right)\Delta\}$ *and* $\phi^b: (i_t^1, i_t^2, \cdots, i_t^m, j_t^m, \cdots, j_t^2, j_t^1) \rightarrow \{v_t - \left(m - \frac{1}{2}\right)\Delta, v_t - \left(m - \frac{3}{2}\right)\Delta, \cdots, v_t - \frac{\Delta}{2}, v_t + \frac{\Delta}{2}\}$, *respectively, where* $i_t^k, j_t^k \in \{0,1\}$ *denote the depth on the LOB at prices* $v_t + \left(m + \frac{1}{2} - k\right)\Delta$ *and* $v_t - \left(m + \frac{1}{2} - k\right)\Delta$, *which are the prices* k *ticks below (above) HFT's ask (bid) in Lemma 2, where* $1 \leq k \leq m$ *and* $k \in \mathbb{N}^+$.

An EA's strategy is a mapping from each state of the LOB (the depth at each price level within HFTs' quotes) to a price level to place her order. Owing to Lemma 1, the EAs choose only between undercutting one tick, undercutting m ticks, and crossing the midpoint to sell at $v_t - \frac{\Delta}{2}$ or buy at $v_t + \frac{\Delta}{2}$ for immediate execution. Later, we show that EAs never choose to undercut m ticks unless $m = 1$, when the three possible strategies degenerate into two. The following proof determines the boundary when EAs are indifferent between undercutting one tick and stimulating HFTs. The formula for this boundary differs when $m = 1$ and $m > 1$, because EAs can attract other EAs when $m = 1$, but EAs do not execute with other EAs when $m > 1$.

Case $m = 1$. From Lemma 2, HFTs quote at $v_t \pm \frac{3\Delta}{2}$ if they are not undercut by EAs. EAs can choose to place orders only at $v_t + \frac{\Delta}{2}$ or $v_t - \frac{\Delta}{2}$. Thus, the order placement strategies for an EA seller and an EA buyer are $\phi^s: (i_t^1, j_t^1) \rightarrow \{v_t - \frac{\Delta}{2}, v_t + \frac{\Delta}{2}\}$ and $\phi^b: (i_t^1, j_t^1) \rightarrow \{v_t - \frac{\Delta}{2}, v_t + \frac{\Delta}{2}\}$, where $i_t^1, j_t^1 \in \{0,1\}$ denote the depth on the LOB at prices $v_t + \frac{\Delta}{2}$ and $v_t - \frac{\Delta}{2}$. Lemma 3 summarizes EAs' equilibrium order-placement strategy when $m = 1$.

Lemma 3: *When $m = 1$, an EA seller's equilibrium order-placement strategy is:*

$$\pi \in \begin{cases} \left(\frac{\Delta}{2}, \pi_2\right) & \phi^s(i_t^1, j_t^1) = v_t + \frac{\Delta}{2} \, if \; i_t^1 = 0; \; otherwise \; sell \; at \; v_t - \frac{\Delta}{2} \\ [\pi_2, \Delta] & \phi^s(i_t^1, j_t^1) = v_t + \frac{\Delta}{2} \, if \; i_t^1 = 0 \; and \; j_t^1 = 1; \; otherwise \; sell \; at \; v_t - \frac{\Delta}{2} \\ (\Delta, 1) & \phi^s(i_t^1, j_t^1) = v_t - \frac{\Delta}{2} \, for \; all \; (i_t^1, j_t^1). \end{cases}$$

Similarly, an EA buyer's equilibrium order-placement strategy is:

$$\pi \in \begin{cases} \left(\frac{\Delta}{2}, \pi_2\right) & \phi^b(i_t^1, j_t^1) = v_t - \frac{\Delta}{2} \, if \; j_t^1 = 0; \; otherwise \; buy \; at \; v_t + \frac{\Delta}{2} \\ [\pi_2, \Delta] & \phi^b(i_t^1, j_t^1) = v_t - \frac{\Delta}{2} \, if \; i_t^1 = 1 \; and \; j_t^1 = 0; \; otherwise \; buy \; at \; v_t + \frac{\Delta}{2} \\ (\Delta, 1) & \phi^b(i_t^1, j_t^1) = v_t + \frac{\Delta}{2} \, for \; all \; (i_t^1, j_t^1), \end{cases}$$

where $\pi_2 = \frac{\beta + 2 - \beta\Delta - \sqrt{(\Delta+1)^2\beta^2 + (4-12\Delta)\beta + 4}}{2\beta}$.

Intuitively, with low sniping risk $\pi \in (\frac{\Delta}{2}, \pi_2)$, the EAs would choose to undercut HFTs whenever possible. With high sniping risk $\pi > \Delta$, the EAs would choose to cross the midpoint for immediate execution regardless of the LOB state. In the interim region $\pi \in [\pi_2, \Delta]$, however, the EAs' action depends on the state of the LOB. We discuss these three cases below.

In the first case, an EA's equilibrium strategy is $\phi^s(0, j_t^1) = v_t + \frac{\Delta}{2}$ for all $j_t^1 \in \{0, 1\}$ and $\phi^b(i_t^1, 0) = v_t - \frac{\Delta}{2}$ for all $i_t^1 \in \{0,1\}$, because the cost of undercutting HFTs by one tick is lower than $\frac{\Delta}{2}$ even in the worst state of the book, i.e. (0,0). Denote C_0 as the cost an EA seller pays to sell at $v_t + \frac{\Delta}{2}$ when $j_t^1 = 0$ (no orders at $v_t - \frac{\Delta}{2}$ when the EA seller submits her order), and C_1 as the cost when $j_t^1 = 1$ or a buy order is present at $v_t - \frac{\Delta}{2}$. $C_1 < C_0$, because a buy order at $v_t - \frac{\Delta}{2}$ increases the probability with which future buyers will use market orders (future limit orders must be price-improving). Thus if $C_0 < \frac{\Delta}{2}$, C_1 must also cost less than stimulating HFTs, and EAs always choose to undercut in any book states as long as the price level does not contain another limit order.

We solve C_0 and C_1 jointly using the following Markov process:

$$\begin{cases} C_0 = \frac{(1-\beta)(1-\pi)}{2}\cdot\left(-\frac{\Delta}{2}\right) + \frac{(1-\beta)(1-\pi)}{2}\cdot C_0 + \frac{\beta(1-\pi)}{2}\cdot C_1 + \frac{\beta(1-\pi)}{2}\cdot C_0 + \frac{\pi}{2}\cdot(1-\frac{\Delta}{2}) + \frac{\pi}{2}\cdot C_0 \\ C_1 = \frac{(1-\beta)(1-\pi)}{2}\cdot\left(-\frac{\Delta}{2}\right) + \frac{(1-\beta)(1-\pi)}{2}\cdot C_0 + \frac{\beta(1-\pi)}{2}\cdot(-\frac{\Delta}{2}) + \frac{\beta(1-\pi)}{2}\cdot C_0 + \frac{\pi}{2}\cdot(1-\frac{\Delta}{2}) + \frac{\pi}{2}\cdot C_0 \end{cases} \tag{A.6}$$

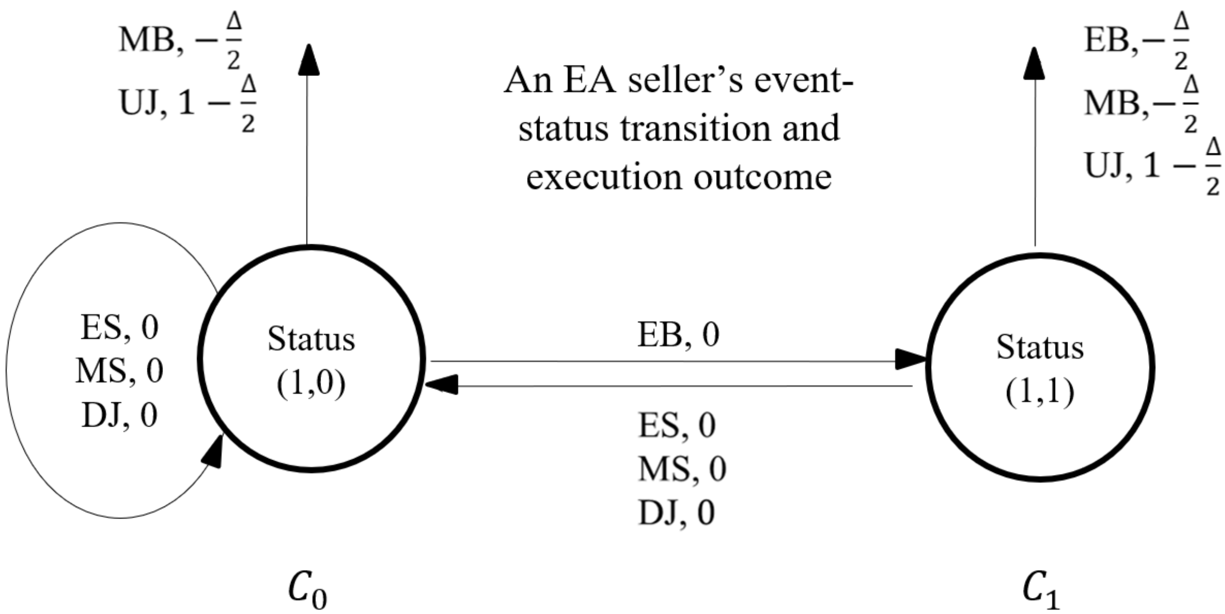

Figure A.1

This figure illustrates the Markov state-transition process for an EA seller who chooses to undercut HFTs. When she places a limit order in LOB state (0,0), the book state becomes (1,0). When she places a limit order in LOB state (0,1), the book status becomes (1,1). EB and ES represent the future arrival of EAs' buy and sell limit orders, MB and MS represent the arrival of MOTs' buy and sell market orders, and UJ and DJ denote upward and downward value jumps. The arrows between states represent state transitions, while arrows pointing toward the outside represent either order executions or cancellations. The number next to each event is the immediate payoff to EAs from the event. C_0 and C_1 are expected terminal payoffs of the EA.

Eq. (A.6) and Fig. A.1 show that six event types can change the state of the LOB. Still, consider C_0 on the ask side (the cost of selling at $v_t + \frac{\Delta}{2}$ while there is no buy order at $v_t - \frac{\Delta}{2}$). An MOT buyer and an MOT seller each arrive with probability $\frac{(1-\beta)(1-\pi)}{2}$. The EA seller enjoys a negative transaction cost of $-\frac{\Delta}{2}$ when the MOT buyer takes liquidity; the MOT seller hits an HFT's quote on the bid side and does not change the state of the LOB. An EA buyer and an EA seller arrive, each with probability $\frac{\beta(1-\pi)}{2}$. An EA buyer posts a buy limit order at $v_t - \frac{\Delta}{2}$ and changes the undercutting seller's cost to C_1; a new EA seller uses a stimulating limit order because she needs to improve the quotes because of Assumption 1, so the cost for the original undercutting EA seller remains C_0. Upward and downward value jumps occur with probability $\frac{\pi}{2}$. An upward jump leads to a sniping cost of $\left(1 - \frac{\Delta}{2}\right)$ whereas a downward jump does not change the state of the LOB because the undercutting EA seller updates her order accordingly.

Because $C_0 - C_1 = \frac{\beta(1-\pi)}{2}\left(C_1 + \frac{\Delta}{2}\right) > 0$, the cost of undercutting is lower if the other side has another undercutting order. The solution for Eq. (A.6) is:

$$C_0 = \frac{2+\beta(1-\pi)}{2-\beta(1-\pi)}\pi - \frac{\Delta}{2}, \tag{A.7}$$

$$C_1 = \frac{2}{2-\beta(1-\pi)}\pi - \frac{\Delta}{2}. \tag{A.8}$$

$$C_0 < \frac{\Delta}{2} \Leftrightarrow \frac{2+\beta(1-\pi)}{2-\beta(1-\pi)}\pi < \Delta \Leftrightarrow \beta\pi^2 + (\beta\Delta - \beta - 2)\pi + (2-\beta)\Delta > 0.$$

Equation $\beta\pi^2 + (\beta\Delta - \beta - 2)\pi + (2-\beta)\Delta = 0$ has two roots:

$$\pi_{1,2} = \frac{\beta+2-\beta\Delta \pm \sqrt{(\Delta+1)^2\beta^2 + (4-12\Delta)\beta + 4}}{2\beta}, \tag{A.9}$$

$$\pi_1 > 1, \pi_2 = \frac{\beta+2-\beta\Delta - \sqrt{(\Delta+1)^2\beta^2 + (4-12\Delta)\beta + 4}}{2\beta}. \tag{A.10}$$

When $m = 1$, EAs choose to undercut when $\frac{\Delta}{2} < \pi < \pi_2$ because $C_1 < C_0 < \frac{\Delta}{2}$. This confirms the case of $\pi \in (\frac{\Delta}{2}, \pi_2)$ in Lemma 3. The intuition is that if jump probability π (the measure of sniping risk) is low, EAs prefer to undercut HFTs.

Next, we show that EAs always prefer stimulating orders (buy at $v_t + \frac{\Delta}{2}$ or sell at $v_t - \frac{\Delta}{2}$) over undercutting HFTs when $\pi > \Delta$. To see that, suppose that an EA deviates to undercut HFTs.

The best-case scenario for this EA is that she can attract both MOTs and other EAs. The cost in this best-case scenario is

$$(1-\pi)\cdot\left(-\frac{\Delta}{2}\right)+\pi\cdot\left(1-\frac{\Delta}{2}\right)=\pi-\frac{\Delta}{2}. \tag{A.11}$$

If $\pi>\Delta$, $\pi-\frac{\Delta}{2}>\frac{\Delta}{2}$. Therefore, when the sniping risk is high, undercutting costs more than stimulating even if the EA can attract both EAs and MOTs. EAs always then choose stimulating HFTs when $\pi\in(\Delta,1)$ in Lemma 3.

The subtle case arises when $\pi\in[\pi_2,\Delta]$, i.e. when the sniping risk is at a moderate level. The equilibrium strategy for EAs depends on the state of the LOB. An EA's equilibrium strategy is to undercut if there is an undercutting order on the other side of the LOB and use stimulating orders otherwise. That is, an EA seller sells at $v_t+\frac{\Delta}{2}$ only when there is a limit buy order at $v_t-\frac{\Delta}{2}$; otherwise, the EA seller will cross the midpoint to sell at $v_t-\frac{\Delta}{2}$. Hence, in this case $\phi^s(i_t^1,j_t^1)=v_t+\frac{\Delta}{2}$ only when $i_t^1=0$ and $j_t^1=1$. An EA buyer follows a similar strategy.

For the above strategy to be an equilibrium, we need to verify two scenarios: 1) If there are no undercutting orders on the other side of the LOB, is it optimal for the arriving EA to use stimulating order? 2) If the other side has an undercutting order, is it optimal for the arriving EA to undercut as well?

In scenario 1), if the arriving EA uses stimulating orders, her cost is $\frac{\Delta}{2}$. If she deviates to undercut, her cost will be C_0 in Eq. (A.7), because the other-side EA chooses to undercut according to the strategy outlined in Lemma 3. Since $C_0\geq\frac{\Delta}{2}$ when $\pi\geq\pi_2$, it is optimal for an EA to use a stimulating order if the other side does not have an undercutting order.

In scenario 2), the book state is $i_t^1=0$ and $j_t^1=1$ and the cost to undercut in this book state is

$$\frac{(1-\beta)(1-\pi)}{2} \cdot \left(-\frac{\Delta}{2}\right) + \frac{(1-\beta)(1-\pi)}{2} \cdot \frac{\Delta}{2} + \frac{\beta(1-\pi)}{2} \cdot \left(-\frac{\Delta}{2}\right) + \frac{\beta(1-\pi)}{2} \cdot \frac{\Delta}{2} + \frac{\pi}{2} \cdot \left(1-\frac{\Delta}{2}\right) + \frac{\pi}{2} \cdot \frac{\Delta}{2} = \frac{\pi}{2}. \quad \text{(A.12)}^{5}$$

Since $\frac{\pi}{2} \leq \frac{\Delta}{2}$ when $\pi \leq \Delta$, it is optimal for an EA to undercut if there is an undercutting order on the opposite side. Hence the strategy outlined in Lemma 3 is optimal when $\pi \in [\pi_2, \Delta]$. This completes the proof of Lemma 3.

Under the strategies outlined in Lemma 3, LOB state $i_t^1 = 0$ and $j_t^1 = 1$ never occurs when $\pi \in [\pi_2, \Delta]$. Therefore, the equilibrium outcome is the same as in the stimulating equilibrium: all EAs choose to stimulate HFTs and the LOB state is always $i_t^1 = 0$ and $j_t^1 = 0$. To see this, note that the first EA chooses to simulate HFTs because the initial state of the book is $i_t^1 = 0$ and $j_t^1 = 0$. The second EA then faces the same state of the LOB and chooses to stimulate HFTs. Therefore, all future EAs choose to stimulate HFTs and the LOB state is always $i_t^1 = 0$ and $j_t^1 = 0$. For ease of exposition, we combine the case for $[\pi_2, \Delta]$ and $(\Delta, 1)$ in Proposition 4 because EAs use stimulating orders in the equilibrium state of the LOB. When $\pi \in \left(\frac{\Delta}{2}, \pi_2\right)$, Proposition 5 outlines the undercutting equilibrium.

Case $m \geq 2$. From Lemma 2, HFTs quote at $v_t \pm \left(m + \frac{1}{2}\right)\Delta$ if no EAs undercut them. When $m \geq 2$, Lemma 1 shows that if an EA chooses to undercut HFTs, she either undercuts by one tick at $v_t \pm (m - \frac{1}{2})\Delta$ or undercuts by m ticks at $v_t \pm \frac{\Delta}{2}$. Lemma 4 presents EA's strategy in each state of the LOB.

[5] Eq. (A.12) differs from C_1 in Eq. (A.6) along only one dimension. In Eq. (A.6), the EA chooses to undercut in any state of the book. Execution by an EA on the other side increases the cost of undercutting from C_1 to C_0, but the change is not dramatic enough for the EA to change her strategy, because $C_0 < \frac{\Delta}{2}$ when $\pi \in \left(\frac{\Delta}{2}, \pi_2\right)$. In (A.12), $\pi \in [\pi_2, \Delta]$, so the higher sniping risk induces the EA to switch from undercutting to simulating when the undercutting order on the opposite side executes. Therefore, in the second, fourth, and sixth terms the cost is $\frac{\Delta}{2}$ in Eq. (A.12) whereas the cost is C_0 in Eq. (A.6).

Lemma 4. *When $m \geq 2$, an EA seller's equilibrium order-placement strategy is:*

1) When $\Delta_\beta < \frac{\Delta}{2}$ (undercutting equilibrium):

$$\pi \in \begin{cases} \left(\frac{\Delta}{2}, \Delta\right) & \phi^s(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = \begin{cases} v_t + \left(m - \frac{1}{2}\right)\Delta, if\ i_t^1 = \cdots = i_t^m = 0\ and\ j_t^m = 0 \\ v_t + \frac{\Delta}{2}, if\ there\ exists\ k < m, i_t^k = 1;\ i_t^m = 0, and\ j_t^m = 1 \\ v_t + \left(m - \frac{1}{2}\right)\Delta\ or\ v_t + \frac{\Delta}{2}, if\ i_t^1 = \cdots = i_t^m = 0\ and\ j_t^m = 1^6 \\ v_t - \frac{\Delta}{2}, for\ all\ remaining\ cases \end{cases} \\ [\Delta, 1) & \phi^s(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = v_t + \left(m - \frac{1}{2}\right)\Delta\ if\ i_t^1 = \cdots = i_t^m = 0, otherwise\ sell\ at\ v_t - \frac{\Delta}{2}. \end{cases}$$

2) When $\frac{\Delta}{2} \leq \Delta_\beta < \Delta$ (stimulating equilibrium):

$$\pi \in \begin{cases} \left(\frac{\Delta}{2}, \Delta\right) & \phi^s(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = v_t + \frac{\Delta}{2}\ if\ i_t^m = 0\ and\ j_t^m = 1, otherwise\ sell\ at\ v_t - \frac{\Delta}{2} \\ [\Delta, 1) & \phi^s(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = v_t - \frac{\Delta}{2}\ for\ any\ \left(i_t^k, j_t^k\right). \end{cases}$$

Similarly, an EA buyer's equilibrium order-placement strategy is:

1) When $\Delta_\beta < \frac{\Delta}{2}$ (undercutting equilibrium):

$$\pi \in \begin{cases} \left(\frac{\Delta}{2}, \Delta\right) & \phi^b(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = \begin{cases} v_t - \left(m - \frac{1}{2}\right)\Delta, if\ j_t^1 = \cdots = j_t^m = 0\ and\ i_t^m = 0 \\ v_t - \frac{\Delta}{2}, if\ there\ exists\ k < m, j_t^k = 1;\ j_t^m = 0, and\ i_t^m = 1 \\ v_t - \left(m - \frac{1}{2}\right)\Delta\ or\ v_t - \frac{\Delta}{2}, if\ j_t^1 = \cdots = j_t^m = 0\ and\ i_t^m = 1 \\ v_t + \frac{\Delta}{2}, for\ all\ remaining\ cases \end{cases} \\ [\Delta, 1) & \phi^b(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = v_t - \left(m - \frac{1}{2}\right)\Delta\ if\ j_t^1 = \cdots = j_t^m = 0, otherwise\ buy\ at\ v_t + \frac{\Delta}{2}. \end{cases}$$

2) When $\frac{\Delta}{2} \leq \Delta_\beta < \Delta$ (stimulating equilibrium):

$$\pi \in \begin{cases} \left(\frac{\Delta}{2}, \Delta\right) & \phi^b(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = v_t - \frac{\Delta}{2}\ if\ i_t^m = 1\ and\ j_t^m = 0, otherwise\ buy\ at\ v_t + \frac{\Delta}{2} \\ [\Delta, 1) & \phi^b(i_t^1, \cdots, i_t^m, j_t^m, \cdots, j_t^1) = v_t + \frac{\Delta}{2}\ for\ any\ \left(i_t^k, j_t^k\right). \end{cases}$$

The intuition associated with the EA's equilibrium strategy is straightforward. As we

[6] No matter whether the seller undercuts one or m ticks, the seller does not take liquidity from the buyer who undercuts m ticks. The subgame perfect strategy of the seller deters the buyer from being the first to undercut m ticks. We do not elaborate on the boundary between two off-equilibrium responses for the EA seller (one versus m ticks) because they lead to the same equilibrium outcome and convey no further economic intuition.

showed in Lemma 1, EAs never undercut HFTs by $2, 3, \cdots, m-1$ ticks. Thus, EAs choose to undercut HFTs only by one tick or m ticks, or they cross the midpoint to stimulate HFTs. The proof of Lemma 1 shows that the cost of undercutting HFTs by one tick is $\Delta_\beta = h_\beta - \left(m - \frac{1}{2}\right)\Delta$ as long as no EA further undercuts the order. The cost of stimulating HFTs is always $\frac{\Delta}{2}$. We will show that undercutting HFTs by m ticks would always cost more than $\frac{\Delta}{2}$ in equilibrium. Therefore, in equilibrium, i_t^m and j_t^m are always zero. As in Lemma 3, however, we also construct and verify our EAs' complete subgame perfect strategy when i_t^m or j_t^m is not zero.

According to Eq. (A.11), the lowest possible cost of undercutting HFTs by m ticks is $\pi - \frac{\Delta}{2}$, under the assumption that the EA can attract both other EAs and MOTs. Therefore, when $\pi > \Delta$, EAs never undercut by m ticks, because the cost of stimulating HFTs is lower.

When $\pi \leq \Delta$, as in the case of $\pi \in [\pi_2, \Delta]$ in Lemma 3: an EA may undercut m ticks only when the opposite side already has an EA who undercuts m ticks, but no EA chooses to be the first to undercut by m ticks. We use an EA buyer as an example to illustrate this result.

Suppose $i_t^m = j_t^m = 0$ (no orders at both $v_t + \frac{\Delta}{2}$ and $v_t - \frac{\Delta}{2}$), and an EA buyer decides to undercut m ticks on the bid side, hence buying at $v_t - \frac{\Delta}{2}$. We first show that the next EA seller would undercut on the ask side as well and would not take the EA buyer's order at $v_t - \frac{\Delta}{2}$ with cost $\frac{\Delta}{2}$. This is simply because the next EA seller can always choose to sell at $v_t + \frac{\Delta}{2}$ and implement the strategy in Eq. (A.12), which has cost $\frac{\pi}{2} < \frac{\Delta}{2}$.

Giving that the next EA seller always chooses to undercut on the ask side, now we calculate the EA buyer's cost of buying at $v_t - \frac{\Delta}{2}$ when $i_t^m = j_t^m = 0$. Denote her cost as D, we have

$$D = \frac{(1-\beta)(1-\pi)}{2} \cdot \left(-\frac{\Delta}{2}\right) + \frac{(1-\beta)(1-\pi)}{2} \cdot D + \frac{\beta(1-\pi)}{2} \cdot A + \frac{\beta(1-\pi)}{2} \cdot D + \frac{\pi}{2} \cdot \left(1-\frac{\Delta}{2}\right) + \frac{\pi}{2} \cdot D. \quad \text{(A.13)}$$

Here, A is the EA buyer's cost if the next event is an EA seller who will choose to undercut on the ask side as explained above. Eq. (A.11) is the best case for the EA who buys at $v_t - \frac{\Delta}{2}$, where the EA can attract all EAs and MOTs from the opposite side and costs $(1-\pi) \cdot \left(-\frac{\Delta}{2}\right) + \pi \cdot \left(1-\frac{\Delta}{2}\right) = \pi - \frac{\Delta}{2}$. Thus $A \geq \pi - \frac{\Delta}{2}$. Insert to Eq. (A.13), we have $D \geq \pi + \beta(1-\pi)\pi - \frac{\Delta}{2}$. It is easy to see that $D \geq \pi + \beta(1-\pi)\pi - \frac{\Delta}{2} > \frac{\Delta}{2}$ when $\pi \in (\frac{\Delta}{2}, \Delta]$ and $m \geq 2$.[7] Therefore, as in Lemma 3, no EAs will initiate trades to undercut HFTs by m ticks, and $i_t^m = j_t^m = 0$ always holds in equilibrium. The equilibrium outcome is that EAs undercut HFTs by only one tick when $\Delta_\beta < \frac{\Delta}{2}$ or by using stimulating orders to accomplish their trading needs when $\Delta_\beta \geq \frac{\Delta}{2}$. This completes the proof of Lemma 4.

Taking Lemmas 1 through 4 together, the equilibria in our model follow an intuitive structure despite the complex off-equilibrium path. HFTs widen their quoting spreads as jump probability π increases. When the tick size is not binding, HFTs widen the spread to $v_t \pm \left(m + \frac{1}{2}\right)\Delta$. EAs alternate between stimulating (Proposition 4) and undercutting (Proposition 5) HFTs, depending on how close HFTs' breakeven half-spread is to the nearest tick grid point.

A numerical example of a solution to Eq. (7).

Here we offer an example: $\pi = 0.3$, $\beta = 0.5$, and $\Delta = \frac{1}{2}$; therefore, $h_\beta = \frac{6}{13}$ and $m = 1$. We

---

[7] Observe that $\pi + \beta(1-\pi)\pi - \frac{\Delta}{2}$ increases in β, so it reaches its minimum when β is the smallest when $m = 2$ and $\Delta_\beta = 0^+$ (i.e. $h_\beta = \frac{\pi}{1-\beta(1-\pi)} = \frac{3\Delta}{2} + 0$). Insert $\frac{\pi}{1-\beta(1-\pi)} = \frac{3\Delta}{2}$ into $\pi + \beta(1-\pi)\pi - \frac{\Delta}{2}$, we have $D \geq \pi + \frac{3\Delta - 2\pi}{3\Delta}\pi - \frac{\Delta}{2}$. Observe that $\frac{\partial(\pi + \frac{3\Delta-2\pi}{3\Delta}\pi - \frac{\Delta}{2})}{\partial\pi} = 2 - \frac{4\pi}{3\Delta} < 0$, so $\pi + \frac{3\Delta-2\pi}{3\Delta}\pi - \frac{\Delta}{2}$ reaches its minimum when $\pi = \Delta$. Its minimum is $\frac{5\Delta}{6}$, still larger than $\frac{\Delta}{2}$.

analytically solve the four linear formulas assuming that $LP^{(i,j)}(1) > 0$ without truncation, and then we insert $\pi = 0.3$ and $\beta = 0.5$.[8] We obtain $LP^{(0,0)}(1) = 0.1406$, $LP^{(1,0)}(1) = -0.0180$, $LP^{(0,1)}(1) = 0.1449$, $LP^{(1,1)}(1) = 0.0069$.

As $LP^{(0,0)}(1) > 0$ and $LP^{(1,0)}(1) < 0$, the HFT supplying liquidity in state (0,0) will cancel her order when an EA undercuts her. As $LP^{(1,0)}(1) < 0$, we replace $LP^{(1,0)}(1)$ with $\overline{LP}^{(1,0)}(1) = 0$. Then,

$$\begin{cases} LP^{(0,0)}(1) = p_1 \cdot LP^{(0,1)}(1) + p_1 \cdot 0 + p_2 \cdot \frac{3}{2}\Delta + p_2 \cdot LP^{(0,0)}(1) + p_3 \cdot \left(-\frac{\Delta}{2}\right) + p_3 \cdot 0 \\ LP^{(0,1)}(1) = p_1 \cdot LP^{(0,1)}(1) + p_1 \cdot LP^{(1,1)}(1) + p_2 \cdot \frac{3}{2}\Delta + p_2 \cdot LP^{(0,0)}(1) + p_3 \cdot \left(-\frac{\Delta}{2}\right) + p_3 \cdot 0 \\ LP^{(1,1)}(1) = p_1 \cdot LP^{(0,1)}(1) + p_1 \cdot 0 + p_2 \cdot LP^{(0,1)}(1) + p_2 \cdot LP^{(1,0)}(1) + p_3 \cdot \left(-\frac{\Delta}{2}\right) + p_3 \cdot 0 \end{cases} \quad \text{(A.14)}$$

The solutions are $LP^{(0,0)}(1) = 0.1447$, $\overline{LP}^{(1,0)}(1) = 0$, $LP^{(0,1)}(1) = 0.1465$, $LP^{(1,1)}(1) = 0.0101$. Neither supplying liquidity when $LP^{(i,j)}(1) < 0$ nor canceling the limit order when $LP^{(i,j)}(1) > 0$ is a profitable deviation. ■

Proof of Corollary 2

By imposing a tick size of Δ, we discuss the three possible equilibrium outcomes and show that the average execution cost is always higher under a discrete tick size than under continuous pricing.

If a discrete tick size leads to the binding equilibrium, the average execution cost for all non-HFTs is

$$\bar{C}_{\Delta}(\beta) = \frac{\Delta}{2} \geq \pi > \pi \cdot \frac{1-\beta}{1-\beta+\beta\pi} = \bar{C}(\beta). \quad \text{(A.15)}$$

If a discrete tick size leads to the stimulating equilibrium, EAs pay more to stimulate HFTs and MOTs pay higher spreads than in the cases under continuous pricing.

[8] The solution includes 26 terms in its denominator because the denominator is the determinant of a 4-by-4 matrix with p_1, p_2, p_3 as its elements. Therefore, we do not report the analytical solution, but it is available upon request.

If a discrete tick size leads to the undercutting equilibrium, MOTs might be better off because they might trade with EAs at a half-spread of $h_\beta - \Delta_\beta$, which is Δ_β narrower than what HFTs quote under continuous pricing. On the other hand, EAs on average pay Δ_β to execute their orders under discrete pricing. We show that EAs' overpayments more than cancels out MOTs' savings.

Suppose that a mass μ of MOTs and EAs trade with each other, and $\beta - \mu$ of EAs have been sniped; in that case, $1 - \beta - \mu$ of MOTs take liquidity from HFTs. The total change in execution cost owing to the discrete pricing is

$$(1 - \beta - \mu)[\left(m + \tfrac{1}{2}\right)\Delta - h_\beta] + \mu(-\Delta_\beta) + \mu(\Delta_\beta - h_\beta) + (\beta - \mu)(1 - h_\beta + \Delta_\beta)$$

$$= (1 - \beta - \mu)[\left(m + \tfrac{1}{2}\right)\Delta - h_\beta] + (\beta - \beta h_\beta - \mu) + (\beta - \mu)\Delta_\beta. \tag{A.16}$$

The first term is nonnegative because $\left(m + \frac{1}{2}\right)\Delta \geq h_\beta$ and therefore HFTs quote wider spreads than continuous pricing in the undercutting equilibrium. The second term is zero because of the definition of h_β, i.e. $(\beta - \mu)(h_\beta - 1)\frac{N-1}{N} + \mu h_\beta = -(\beta - \mu)(h_\beta - 1)\frac{1}{N}$. The third term is positive because $\beta - \mu > 0$ and $\Delta_\beta \in (0, \Delta]$. Therefore, the average execution cost strictly increases under discrete pricing. ■