Journal of Financial Markets xxx (xxxx) 100583

Contents lists available at ScienceDirect

Journal of Financial Markets

journal homepage: http://ees.elsevier.com

Speed and learning in high-frequency auctions☆

Marlene Haas[a], Mariana Khapko[b], Marius Zoican[c,*]

[a] *Cornerstone Research, 2001 K Street NW, North Tower, Suite 800, Washington, DC 20006, United States*
[b] *University of Toronto Scarborough, Rotman School of Management, Canada*
[c] *University of Toronto Mississauga, Rotman School of Management, Canada*

ARTICLE INFO

Article history:
Received 23 October 2018
Received in revised form 27 May 2020
Accepted 12 June 2020
Available online xxx

JEL classification:
D43
D47
G10
G14

Keywords:
High-frequency trading
Batch auction markets
Liquidity
Adverse selection

ABSTRACT

Faster trading improves liquidity in periodic call auction markets, in contrast to continuous-time markets. We build a model where high-frequency traders (HFTs) engage in duels to trade on stale quotes. More frequent periodic auctions increase the likelihood that a single HFT arrives in any given auction and subsequently acts as a monopolist on information. Higher trading speed increases the expected number of arbitrageurs participating in auctions, promoting competition between snipers and improving liquidity. We find that faster trading and longer auction intervals are substitute instruments to reduce bid-ask spreads. Relative to continuous-time trading, periodic batch auctions reduce HFT informational rents.

1. Introduction

Is continuous-time trading inherently flawed? Modern exchanges are largely organized as continuous-time limit order books: One can buy and sell assets at any given time. Traders who react first to profitable opportunities have a comparative advantage. Consequently, continuous-time trading generates incentives for each trader to become *marginally* faster than her competitors. As a result, an “arms' race” emerged between high-frequency traders: The round-trip trading times between New York City and Chicago dropped from 16 ms in 2010 to 8.02 ms in July 2015. A London-based trader can buy stocks in Frankfurt within just 2.21 ms. As a benchmark, light needs 2.12 ms to travel the same distance.[1]

☆ The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of Cornerstone Research. The paper received the 2016 Josseph de la Vega Prize from the Federation of European Securities Exchanges under the title “Discrete or continuous trading? HFT competition and liquidity on batch auction markets.” We have greatly benefited from discussion on this research with Matthieu Bouvard, Eric Budish, Jérôme Dugast, Carole Gresse, Björn Hagströmer, Mario Milone, Sophie Moinas, Klaus Ritzberger, and Bart Yueshen Zhou. We are grateful to two anonymous referees, seminar participants at the Université Paris-Dauphine and Bank of England for insightful comments, as well as conference participants at the 2016 New York Stern Microstructure meeting, the 2016 IFABS meeting, the 2016 Northern Finance Association conference, and the EUROFIDAI Paris December 2016 Finance Meeting. Mariana Khapko and Marius Zoican gratefully acknowledge Social Sciences and Humanities Research Council for an Insight Development Grant (430-2018-00125). Marius Zoican acknowledges financial support from Quantitative Management Research Initiative (QMI) under the aegis of the Fondation du Risque, a joint initiative by Université Paris-Dauphine, l'École Nationale de la Statistique et de l'Administration ParisTech, Addstones Groupe Français d'Informatique et La Française Investment Solutions.

* Corresponding author.
E-mail addresses: mhaas@cornerstone.com (M. Haas); mariana.khapko@rotman.utoronto.ca (M. Khapko); marius.zoican@rotman.utoronto.ca (M. Zoican)

[1] Sources for this paragraph: McKay Brothers Microwave Latencies Table and Bloomberg: “Wall Street Grabs NATO Towers in Traders Speed-of-Light Quest.”

https://doi.org/10.1016/j.finmar.2020.100583
1386-4181/© 2020.

Such arms' race is not necessarily benign: Ever higher trading speeds come at a non-trivial social cost. In 2010, Spread Networks spent USD 300 million building a straight-line fiber optic cable between New York City and Chicago for a 3 ms latency gain. Moreover, faster trading does not necessarily improve market quality (e.g., Ye et al., 2013; Shkilko and Sokolov, 2019).

Budish et al. (2015) argue that the "arms' race" for speed is a by-product of continuous-time trading, which is the prevailing market design for securities trading around the world. Their proposed alternative is a periodic batch auction market (a series of call auctions trading throughout the day). While traders can submit orders at any time, the batch auction market clears at discrete intervals, for example every second, through a uniform auction. Therefore, a batch auction market eliminates the "inherent flaw" of the continuous-time limit order book: the discrete advantage from being marginally (e.g., 1 ns) faster than competitors. Consequently, the scope for an HFT arms' race is limited.

Our paper "zooms in" on high-frequency batch auction markets, where the market clearing frequency is of the same order of magnitude as traders' speed to respond to arbitrage opportunities.[2] We complement Budish et al. (2015) and consider the role of learning frictions in batch auctions. Whereas in Budish et al. (2015) all high-frequency traders (HFTs) become informed simultaneously, we allow HFTs to learn the news at heterogeneous times. For very short auctions, there is a positive probability that a single HFT becomes informed and this trader can subsequently extract monopoly rents. We find that faster trading can be useful in high-frequency batch auction markets, as it reduces the probability that a single informed HFT participates in any given auction. Consequently, trading speed can improve liquidity as it stimulates competition between arbitrageurs. The economics of trading speed therefore crucially depends on the market design. Menkveld and Zoican (2017) show that in continuous-time markets, trading speed can actually harm liquidity as it triggers more zero-sum trading "duels" between high-frequency market-makers and arbitrageurs.

Our model builds on the general structure in Menkveld and Zoican (2017), where we allow for the market to clear at discrete time intervals. We model the discrete time market to be fully transparent before market clearing. This means that traders are able to observe order submissions and cancellations at all times. In addition, traders can submit and cancel orders at any time, but the market clears at equally-spaced discrete time intervals. Unlike in continuous-time markets, where a marketable order submission corresponds to the end of the trading round, the timing of market clearing is exogenous in periodic call auction markets, and thus unaffected by order arrival. High-frequency traders compete to provide liquidity at the beginning of the trading game. One of the HFTs, at random, is the first to post orders in the book and becomes a high-frequency market-maker (HFM). The other HFTs assume the role of high-frequency speculators (HFSs), or snipers, who try to profit from short-term stale quotes. Following an innovation in the asset value, all HFTs rush to the market: the HFM to cancel or update the now stale quotes, the HFSs to profit from a temporary arbitrage opportunity.

The main economic friction in the model is imperfect learning, which generates short-term asymmetric information. High-frequency traders still require time, albeit a short period, to process and react to trading signals (Brolley and Cimon, 2020). If the inter-auction interval is very short, on par with HFT speed, the number of HFTs in each auction becomes random.[3] This leads to three economically relevant auction types: (1) maker auctions, in which the market-maker removes stale quotes before any sniper can trade on them; (2) sniper auctions, where a single sniper can take advantage of the HFM's stale quotes; and (3) competitive auctions, where multiple snipers compete in the same auction until information rents are eliminated. We refer to these auction types as "economically relevant" because they are the only three types for which our model results in updates to the auction order book. We find that either a longer auction interval or a higher trading speed have the same effect: that is, they improve liquidity since they reduce the probability of sniper auctions and adverse selection costs.

In our model, continuous-time markets emerge naturally as a limiting case of periodic batch auction markets, as the duration of the auction interval becomes arbitrarily small. In the setup of Budish et al. (2015), HFT snipers engage in Bertrand competition in every auction, regardless of its duration. The outcome of this assumption is that zero-duration auction markets have better liquidity than continuous-time markets, although the market designs are otherwise equivalent. Introducing HFT speed as a parameter in the model allows us to eliminate the discontinuity: as the auction length decreases, the probability of two or more HFT arrivals in the same auction is lower. In the limit, at most one HFT can arrive in each auction and time-priority is de facto restored as in continuous-time markets.

A periodic batch auction market always improves liquidity and reduces HFT profits relative to the continuous-time limit order markets. We model HFT speed as an exogenous policy variable, rather than an endogenous investment in technology. In this sense, we interpret speed as being driven by common latency components, such as exchange architecture choices or co-location rules implemented by trading venues.[4]

There is an ongoing debate about the relative merits of continuous- and discrete-time trading mechanisms. The U.S. Securities and Exchanges Commission (SEC) chair indicated in June 2014 her interest in batch auction markets as a "more flexible, competitive" exchange design. In October 2015, Chicago Stock Exchange received approval from the SEC to launch a batch-auction platform, CHX SNAP. In March 2016, the London Stock Exchange launched a midday auction for the most liquid securities.[5] This paper offers an important policy-relevant insight: the HFT arms' race in a batch auction market intensifies competition between arbitrageurs and de-

[2] To understand the order of magnitude, we note that in 2011, the median arbitrage opportunity in U.S. index futures had a duration of 7 ms (Budish et al., 2015).

[3] In this sense, an "auction" stands for the time interval between two successive market clearings.

[4] An alternative interpretation, where HFT speed refers to information processing latency would yield identical results, as long as speed is common to all HFT, rather than an individual investment choice.

[5] Sources for this paragraph: SEC Chair Speech on June 5, 2014; the CHX SNAP Auction Market webpage; and the London Stock Exchange webpage.

creases the bid-ask spread. Stimulating HFT speed competition is effective in scenarios where longer auction intervals conflict with short-term trading needs. Discrete-time trading can thus align private and social incentives, since the arms' race is arguably privately optimal for HFTs.

Our paper contributes to a growing literature on HFT and market design. A closely related paper is Budish et al. (2015), who study a model of batch auction markets. The authors assume HFTs react to new information without a delay. As a consequence, arbitrageurs always compete à la Bertrand. We introduce a learning friction that generates adverse selection risk. In particular, we allow for auctions where a single HFT sniper participates and consequently acts as a monopolist on information. The richer model we propose features a positive bid-ask spread and unveils new economic channels, in particular the role of speed in discrete-time markets. Our model nests the Budish et al. (2015) setup for relatively infrequent auctions, as well as the continuous market outcome if the inter-auction delay is short enough.

In a policy paper, Farmer and Skouras (2012) estimate the worldwide benefits of the transition from limit order to periodic auction markets to be around USD 500 billion per year. Their study is similar to ours as they model the batch auction time as a Poisson process. However, the authors do not consider the effects of batch auctions on arbitrageur competition, nor the endogenous order choice for HFTs. Wah and Wellman (2013) and Wah et al. (2015) develop agent-based models to showcase the benefits of batch auction markets. Batch auctions improve welfare as they better aggregate supply and demand. If traders can choose between a limit order and a batch auction market, HFTs will always follow the choice of slow traders to increase arbitrage profits. Fricke and Gerig (2018) calibrate a batch auction trading model using U.S. data with risk-averse traders and find the optimal batch length to be between 0.2 and 0.9 s. However, the authors focus on liquidity risk rather than on an adverse selection channel. Aldrich and López Vargas (2019) build an experimental laboratory market to compare continuous- to discrete-time markets and find that the latter are less conducive to predatory trading, while having better liquidity and lower price volatility.

Madhavan (1992) and Economides and Schwartz (1995) argue that batch auction markets improve price efficiency as they aggregate disparate information from traders for a longer time interval. In a model of double-auction markets where "fast" traders act as intermediaries rather than arbitrageurs, Du and Zhu (2017) find that, if news is not scheduled, the socially optimal trading frequency exceeds information arrival frequency. However, trading speed does not offer an informational advantage. In our model, stochastic learning leads to a novel channel: infrequent trading improves the public-to-private news ratio, enhances learning, and reduces the adverse selection cost.

Other solutions to the HFT arms' race externality include exogenous order delays, or speed bumps (Baldauf and Mollner, 2020; Aoyagi, 2019; Brolley and Cimon, 2020). Kyle and Lee (2017) propose, at the opposite end of the spectrum, a fully-continuous exchange where traders submit buy or sell trade rates over time, rather than quantities. Brolley and Zoican (2020) argue for a FinTech-oriented solution where decentralized exchanges would incentivize HFTs to purchase low-latency infrastructure on an as-needed basis, reducing the overall cost of the arms' race.

Our paper is also related to the literature on auctions in financial markets. Janssen and Rasmusen (2002) and Jovanovic and Menkveld (2019) study auction mechanisms where the number of competing bids (in our case, the number of informed HFTs) is not common knowledge. The bidding equilibrium is always symmetric and in mixed strategies. We propose asymmetric information as a rationale for the uncertain number of auction participants. Kremer and Nyborg (2004) study various allocation rules in uniform price auctions and find that a discrete tick size or uniform rationing at infra-marginal prices eliminates arbitrarily large underpricing. Kandel et al. (2012), Pagano and Schwartz (2003), and Pagano et al. (2013) document that the introduction of opening and closing call auctions on the Paris Bourse in 1996 and Nasdaq in 2004, respectively, reduced spreads and volatility and improved price discovery.

Finally, our paper relates to a growing literature on high-frequency trading. Asymmetric information or asymmetric speed across traders may lead to "stale" quotes and adverse selection costs for market-makers, as first modeled in Copeland and Galai (1983) and Glosten and Milgrom (1985). As in Foucault et al. (2003) and Foucault et al. (2013), traders in our model engage in a speed race to react first to new arbitrage opportunities. Further, the benefits from the speed arms' race between HFTs are limited. High-frequency market markers have an option to cancel their quotes before being "sniped" by arbitrageurs, a mechanism explored in Baruch and Glosten (2013). Biais et al. (2015) find that investment in fast trading technology exceeds the social optimum. According to Kumar and Seppi (1994) and Menkveld (2014), the HFT arms' race can hurt market liquidity. In the same spirit, Menkveld and Zoican (2017) argue that ever faster exchanges promote a higher frequency of inter-HFT trades, increasing the adverse selection cost and consequently the spread. Empirical evidence suggests HFTs use strategies to snipe stale quotes. Hendershott and Moulton (2011) and Brogaard et al. (2014) find HFT market orders have a larger price impact. For more complete and in-depth analysis of the literature on the effects of HFT on market liquidity, we refer the interested reader to Menkveld (2016).

2. Model

In this Section, we develop the theoretical model and discuss, in turn, each assumption.

2.1. Trading environment

One risky asset is traded in a call auction market as in Budish et al. (2015). Traders can submit and cancel orders at any time, but the market clears at equally-spaced discrete time intervals of length τ, that is at $t \in \{\tau, 2\tau, \dots, k\tau \dots\}$. There is full pre-auction transparency, that is, traders observe new orders and cancellations at all times. Further, there is no time priority within the auction interval: all buy and sell orders entered in the kth auction between $t \in [(k-1)\tau, k\tau)$ are processed together and, if matched, executed at a unique clearing price. Traders can submit both limit orders and market orders. Limit orders are instructions to buy or sell a certain quantity of the asset at a given price or better. In contrast, market orders do not specify a limit price: Rather, they are instructions to execute at the best price that allows the order to be filled in full. Orders can be posted and cancelled at no cost before each market clearing.

Orders left outstanding after the market clears at the end of the kth auction are carried over to the $k+1$st auction, and have time priority over new orders at the same price. In each auction, if demand and supply intersect horizontally such that the price is uniquely determined and the quantity is not, then orders are executed with a non-unit probability such that the market clears. Conversely, if demand and supply intersect vertically, quantity is uniquely determined: all trades execute in full.[6] However, the market clears at a range of prices, e.g., between $[p_L, p_H]$. We follow Budish et al. (2015) and assume the market clearing price in this case is $\frac{p_L+p_H}{2}$.

2.2. Asset

The common value of the risky asset at the start of the trading game is v. Innovations about the value of the asset (*news*) arrive as a Poisson process with intensity η (as in Baruch and Glosten, 2013). The absolute size of value innovations is $\sigma > 0$. Conditional on news, the asset value either jumps to $v+\sigma$ for "good" news or $v-\sigma$ for "bad" news. The common value of the asset is a martingale; hence, good and bad news are equally likely.

2.3. Agents

The risky asset is traded by two types of agents: a single liquidity trader and $N > 2$ HFTs. There is no time discounting and all agents are risk neutral. The liquidity trader receives private value shocks according to a Poisson process with intensity μ. Conditional on the shock, they are equally likely to buy or sell exactly one unit of the asset.

High-frequency traders, by contrast to the liquidity trader, possess a monitoring technology that allows them to process changes to the security's fundamental value. Learning, however, takes time. The HFT market arrival times follow a Poisson process with rate $\psi > 0$, where ψ stands in for the traders' reaction speed. We interpret ψ as a technology-driven parameter: for example, the speed of incoming data feeds from trading platforms, the round-trip order latency between exchanges where the asset is cross-listed, or the performance of algorithms parsing news and social media feeds. Note that the traders' reaction speed ψ governs both the expected time from news to order execution (equal to ψ^{-1}), as well as the per-auction probability of HFT arrival. In continuous-time markets, only the former interpretation is relevant. In call auction markets, the latter interpretation becomes the most salient. Let t_{HFT} denote the time at which an HFT observes the value innovation. The probability that an HFT observes a value innovation in auction k for the first time, conditional on news, is:

$$\begin{aligned}\alpha \equiv \text{Prob}\left(t_{\text{HFT}} < k\tau \mid t_{\text{HFT}} \geq (k-1)\,\tau\right) \\ &= \frac{\exp(-\psi(k-1)\tau) - \exp(-\psi k\tau)}{\exp(-\psi(k-1)\tau)} \\ &= 1 - \exp(-\psi\tau),\end{aligned} \tag{1}$$

which is constant for each individual auction k.

2.4. Timing

The timing of our model mirrors Menkveld and Zoican (2017). At $t=0$, the N HFTs compete to provide price quotes. Let $v-s$ and $v+s$ denote the lowest buy and highest sell price quoted at $t=0$, respectively. Two high-frequency trading roles emerge as a result. The HFT that succeeds in posting their quotes becomes the high-frequency market-maker (HFM). The other $N-1$ HFTs act as quote snipers, poised to trade on stale quotes: we refer to such HFTs as high-frequency speculators (HFSs).

We denote the "trigger event" to be the first arrival of either news or a liquidity trader. Since both events are Poisson processes, the expected time until the trigger event is $\frac{1}{\eta+\mu}$. With probability $\frac{\mu}{\eta+\mu}$, the trigger event is liquidity-driven: an LT arrives at the market and consumes either the bid or the ask quote. In this case, since there is no change in the asset's value, the HFSs have no incentive to trade after the trigger event. Conversely, with the complementary probability $\frac{\eta}{\eta+\mu}$, news arrives instead. The $N-1$ HFSs and the HFM rush to submit orders as soon as they observe the value innovation: the HFSs send orders to trade against the stale quote, whereas the HFM rushes to cancel the order from $t=0$. The game ends upon the first market clearing with non-zero trading volume or a cancelled order.

Fig. 1 illustrates an example of the sequence of events in the model. The labels HFT_1, HFT_2, …, indicate HFT arrivals in chronological order. If the trigger event is an LT arrival, the liquidity trade is executed in the first auction at τ. Conversely, if the trig-

[6] In our model, trades always occur between limit orders on one side and market orders on the other side. In this case, the supply and demand always intersect horizontally, and the unique price is the one for which all market orders are executed in full.

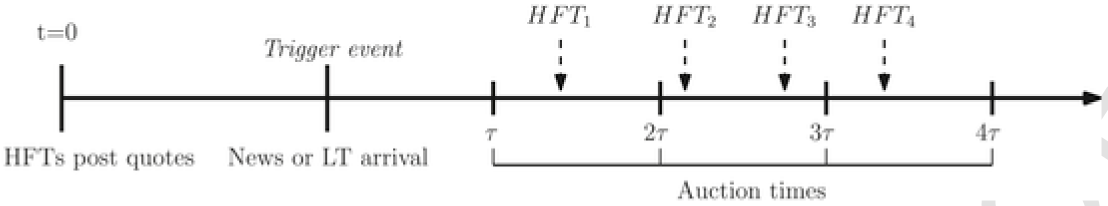

Fig. 1. Example of event sequence in the model.

ger event is a news arrival, the first auction elapses with no trades, since no HFTs arrive before τ. In the example in Fig. 1, trade only happens at $t = 2\tau$ (that is, when the second auction clears), upon the arrival of HFT_1.[7]

All model parameters are summarized in Appendix A.

3. Equilibrium

We search for symmetric Nash equilibria in the trading game. In particular, at any point in time $t \geq 0$ an equilibrium consists of (i) HFT strategies to submit limit or market orders to buy or to sell a given quantity of the asset (for limit orders, at a specified price), or a decision to cancel such orders, and (ii) market clearing prices at the batch auction times $t \in \{\tau, 2\tau, \dots, k\tau, \dots\}$.

3.1. Optimal sniping behavior

Let A denote the index of the first auction following news, where at least one HFT is present at the market (either an HFS or the HFM). Auction A therefore starts at $(A-1)\tau$ and ends at $t_A = A\tau$. Since HFTs do not participate in every auction, particularly for a very short τ, it could be that $t_A > t_{\text{news}} + \tau$ such that several auction intervals elapse between the news event and the first HFT order arrival. However, any auction interval that elapses at $t \in [t_{\text{news}}, t_A)$ is not economically relevant since it does not trigger any trades and does not reflect any order book updates.

We distinguish between three possible economically relevant scenarios. First, the HFM might return at the market during A. In this case, we refer to A as a *maker-auction* or *M-auction*. The probability of a maker auction, conditional on at least one arrival during A, is

$$\pi_M \equiv \mathbb{P}(\text{HFM order in } A \mid \text{at least one HFT arrival in } A) = \frac{\alpha}{1-(1-\alpha)^N}. \tag{2}$$

Second, if the HFM is not active in the market during A (either through cancelling or submitting orders), but the quotes submitted at $t = 0$ are still live, it must be that at least one sniper submits an order. We distinguish two possibilities: either (i) exactly one or (ii) at least two HFS submit orders in auction A. We denote A as a *sniper auction* or *S-auction* if exactly one HFS is submitting orders to the market (and the HFM is not). The probability of a sniper auction is:

$$\begin{aligned}\pi_S &\equiv \mathbb{P}(\text{Exactly one HFS order in } A \mid \text{at least one HFT arrival in } A)\\ &= \frac{(N-1)\alpha(1-\alpha)^{N-1}}{1-(1-\alpha)^N}.\end{aligned} \tag{3}$$

Finally, auction A is a *competitive-auction* or *C-auction* if at least two HFS submit orders during A, while the HFM does not act in A. The probability of a competitive auction is:

$$\begin{aligned}\pi_C &\equiv \mathbb{P}(\text{Two or more HFS orders in } A \mid \text{at least one HFT arrival in } A)\\ &= \frac{\sum_{k=2}^{N-1}\binom{N-1}{k}\alpha^k(1-\alpha)^{N-k}}{1-(1-\alpha)^N}.\end{aligned} \tag{4}$$

It is easy to verify that $\pi_M + \pi_S + \pi_C = 1$. Table 1 provides the three auction types and their respective probabilities.

Lemma 1 describes the HFT optimal order submission strategies for each of the three auction types in Table 1.

Lemma 1 *(Order submission upon a news trigger.) Following a news trigger event, HFTs optimally submit the following orders as a function of the auction type*:

[7] We also note that if, for example, HFT_2 would arrive before 2τ, both HFT_1 and HFT_2 orders are considered for market clearing in the second auction.

Table 1
Auction types.

Auction type	HFM active?	HFS active?	Probability
Maker (M) auction	Yes	Irrelevant	$\pi_M = \frac{\alpha}{1-(1-\alpha)^N}$
Sniper (S) auction	No	One	$\pi_S = \frac{(N-1)\alpha(1-\alpha)^{N-1}}{1-(1-\alpha)^N}$
Competitive (C) auction	No	Two or more	$\pi_C = \frac{\sum_{k=2}^{N-1}\left(\frac{N-1}{k}\right)\alpha^k(1-\alpha)^{N-k}}{1-(1-\alpha)^N}.$

(i) *In a maker-auction, the HFM cancels any quotes posted at $t = 0$. High-frequency snipers do not submit any new orders.*
(ii) *In sniper auctions and competitive auctions, each speculator present at the exchange sends a buy (sell) market order for one unit of the asset, if the trigger event was good (bad) news.*

Since the liquidity trader demand at most one unit of the security, the HFM infers that any directional order flow $|q| \geq 2$ indicates the presence of snipers. Consequently, competitive HFMs stand ready to buy (respectively, sell) any quantity of two units and above at a price of $v + \sigma$ (respectively, at $v - \sigma$).

The high-frequency snipers' rationale for trading is to profit from the existence of stale quotes as posted by the HFM at $t = 0$. If the HFM is actively present in the market at A, she optimally cancels the previously submitted quote on the news-relevant side of the market to avoid being adversely selected.[8] Since there are no stale quotes in the book, high-frequency snipers have no incentives to trade and optimally refrain from submitting orders to the matching engine. As a consequence, maker auctions have no trading volume and all HFTs earn zero profit.

The single HFS present at the market in a sniper auction is essentially a monopolist on information. Consider the case of good news (the case of bad news is symmetrical): the asset value becomes $v + \sigma$, and the market maker is willing to sell the security for $v + s$ as per the quote submitted in $t = 0$. A market buy order for one unit executes at $v + s$ (the lowest price that clears the market) and the unique sniper earns a monopoly profit equal to $\sigma - s$. Since the informed trade is a zero-sum game, the market maker loses $s - \sigma < 0$. As in Glosten and Milgrom (1985), a sniper auction essentially mimics the outcome obtained in a continuous-time market, where the first HFS at the market has a monopoly on information due to time priority rules.[9]

In competitive auctions, snipers submit two or more buy market orders. Consider again the case of good news: Since the auction is uniform price, the market clears at $v + \sigma$ and both the market maker and HFSs earn zero profit. Again, no HFS can profitably deviate from submitting a market order: a limit buy order with a price higher than $v + \sigma$ would yield a loss, whereas a limit buy order with a price lower than $v + \sigma$ would still get executed at $v + \sigma$.

Adverse selection on stale quotes, as in Menkveld and Zoican (2017), only emerges therefore in sniper auctions. In contrast, maker auctions correspond to zero trading volume, whereas competitive auction trades execute at the efficient price and generate no adverse selection.

Lemma 2 *(Sniping probability.) The probability of a sniping trade conditional on news, that is π_S as given in equation (3), decreases in the HFT speed ψ.*

The result in Lemma 2 is driven by two channels. First, since a higher ψ increases the per-auction probability of HFT arrival α, the HFM is more likely to arrive in auction A and cancel her quotes. Second, even if the HFM does not cancel her stale quotes in the auction, a higher ψ increases the likelihood of a competitive auction, where two or more snipers compete and drive the clearing price to the true asset value. Fig. 2 illustrates how the probabilities of each auction type change with HFT speed. As HFTs become faster, sniper-auctions where a single HFS can trade on stale quotes become less likely. At the same time, a faster market maker is more likely to update her quotes. The probability of an auction where snipers compete for stale quotes first increases, then decreases in HFT speed, as market makers are more likely to reach the market themselves and render sniper orders moot.

Equivalence with price competition. While the strategies above do not involve explicit competition on prices as in Budish et al. (2015), we note that a model where snipers can condition the order price on the presence or absence of competitors would yield identical results.

In a sniper auction, conditional on good news, the unique HFS can submit a buy order for any price $\mathbf{p}_{\text{HFS}} \in [v + s, v + \sigma]$. The market clears and one unit is traded at $\frac{\mathbf{p}_{\text{HFS}}+(v+s)}{2}$. As a result, the HFS earns a positive profit of $v + \sigma - \frac{\mathbf{p}_{\text{HFS}}+(v+s)}{2}$. The HFS optimally selects the lowest buy price that clears the market, $\mathbf{p}_{\text{HFS}} = v + s$.

Conversely, in a competitive auction HFSs engage in Bertrand competition on prices. Consider again the case of good news, for illustration purposes. Any market clearing price $\mathbf{p}_C < v + \sigma$ cannot be an equilibrium. If there are $m \geq 2$ HFS at the market posting an order to buy at $\mathbf{p}_C$, then each of them is randomly matched with probability $\frac{1}{m}$<1. It is optimal for any HFS to bid a slightly

[8] The quote on the no-news side is never executed in this case before the end of the game, and therefore the HFM is indifferent between cancelling it or not.
[9] Note that the HFS cannot profitably deviate by posting a limit order with price different from $v + s$: a buy order for $v + s - \varepsilon$ would not be executed, and a buy order for $v + s + \varepsilon$ would yield a lower profit, for any $\varepsilon > 0$.

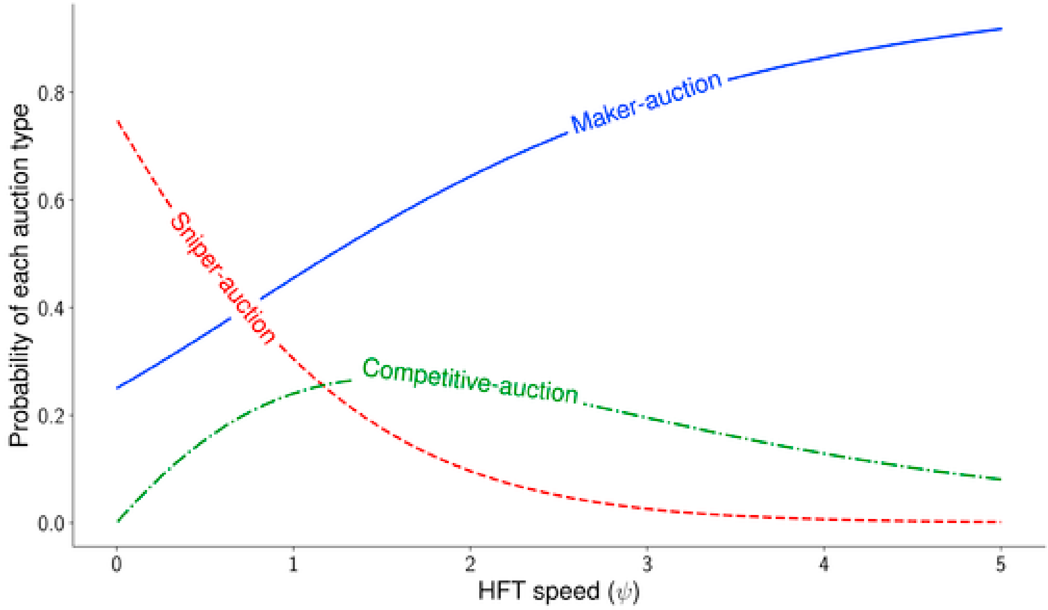

Fig. 2. Probabilities of each auction type. This figure illustrates the probabilities of each auction type (market, sniper, and competitive auction) as a function of the HFT speed parameter ψ. Parameter values: $\mu = \eta = \sigma = 1$, $\tau = 0.5$, and $N = 4$.

higher amount, $\mathbf{p}_C + \varepsilon$, and guarantee execution of his order, while still earning a positive profit. In equilibrium, the market clears at $v + \sigma$, and one HFS, randomly selected, is matched with the market maker.

A potential concern with the price competition is that snipers could wait to submit orders until an instant before market clearing, to avoid revealing themselves. The concern could be alleviated, for example, by allowing for a "soft close" auction, where market clearing is automatically delayed for a very brief interval upon the arrival of new orders.[10]

3.2. Liquidity in batch auction markets

In this subsection, we study the optimal HFT initial quotes (i.e., the ones submitted in $t = 0$). We focus on symmetric equilibria such that each HFT submits at $t = 0$ an order to buy the first unit of the asset for $v - s$ and an order to sell the first unit of the asset at $v + s$, where $s \geq 0$ denotes the outstanding half-spread on the limit orders around the common value v.

At $t = 0$, all HFTs submit quotes to the markets without being assigned a market marker or HFS role. The expected utility of a market maker is:

$$U_{HFM}(s) = \frac{\mu}{\mu + \eta} s + \frac{\eta}{\mu + \eta} \pi_S (s - \sigma). \tag{5}$$

With probability $\frac{\mu}{\mu+\eta}$, the trigger event is a liquidity trade. The liquidity trader submits an order at a price that clears the market and minimizes their trading cost: either buy one unit at $v + s$, the lowest price that the HFM accepts as a seller, or sell one unit at $v - s$, the highest price that the HFM requests as a buyer. In this case, the market maker earns s and the game ends.

With probability $\frac{\eta}{\mu+\eta}$, the trigger event is a news arrival. From Lemma 1 and the discussion in Subsection 3.1, the HFM earns zero profit in maker and competitive auctions, and makes a loss of $s - \sigma$ in sniper auctions, which occur with probability π_S.

An HFS has positive utility if both (i) the trigger event is a news arrival, with probability $\frac{\eta}{\mu+\eta}$, and (ii) she is the only trader to arrive in auction A. Since there are $N - 1$ potential snipers, each HFS is a monopolist with probability $\frac{\pi_S}{N-1}$. The expected utility of an HFS becomes:

$$U_{HFS}(s) = \frac{\eta}{\mu + \eta} \frac{1}{N - 1} \pi_S (\sigma - s). \tag{6}$$

In equilibrium, HFTs must be indifferent between the two roles. The equilibrium spread $s^\star$ is pinned down by equating the expected utilities in (5) and (6). Intuitively, if an HFT_i posts instead a larger half-spread $s^+ > s^\star$, then she does not become the HFM.

[10] This is similar to the auction mechanism implemented by Amazon, see Roth and Ockenfels (2002). Further, Cinnober, a Swedish financial infrastructure provider, patented adaptive micro-auctions, where market clearing is automatically delayed if new order arrivals change the clearing price. See patent US 8719146 B2, available at: https://bit.ly/2SKI6mb.

In this case, HFT_i would earn the sniper utility in (6) and is therefore indifferent between deviating or not.[11] Conversely, if HFT_i deviates and posts a tighter spread $s^- < s^\star$, she is sure to assume the HFM role. However, the deviation is not profitable since (i) the HFM and HFS utilities are equal in equilibrium and (ii) the HFM utility increases in the posted spread, that is:

$$U_{HFM}\left(s^-\right) < \frac{1}{N} U_{HFM}\left(s^\star\right) + \frac{N-1}{N} U_{HFS}\left(s^\star\right) = U_{HFM}\left(s^\star\right). \tag{7}$$

Lemma 3 describes the half-spread pinned down by this indifference condition.

Lemma 3 *[Half-spread] The unique half-spread $s^\star$ that ensures that $U_{HFM}(s^\star) = U_{HFS}(s^\star) > 0$ is given by:*

$$s^\star = \frac{\eta N \pi_S}{\eta N \pi_S + \mu(N-1)}\sigma = \frac{\eta N \alpha (1-\alpha)^N}{\mu(1-\alpha) + (1-\alpha)^N \left(\eta N \alpha - \mu(1-\alpha)\right)}\sigma. \tag{8}$$

Since from Lemma 3 the equilibrium spread $s^\star < \sigma$, from equation (6) it follows that HFTs earn positive utility in equilibrium. Therefore, all HFTs submit quotes at $t = 0$: with probability $\frac{1}{N}$, each of them is equally likely to arrive first and become HFM. The other $N - 1$ HFTs assume the role of HFS.

Proposition 1 characterizes the equilibrium of the trading game by merging the optimal HFT strategy at $t = 0$ described in Lemma 3 and the optimal HFT strategy for $t > 0$ in Lemma 1.

Proposition 1 *[Equilibrium] The following strategies form a Nash equilibrium in the trading game.*

(i) *At $t = 0$, each HFT submits a buy limit order for the first unit at price $v - s^\star$ and a sell limit order for the first unit at price $v + s^\star$, where $s^\star$ is given by (8). Further, HFTs submit buy and sell orders for any further incremental units at prices $v - \sigma$ and $v + \sigma$, respectively. The first HFT to arrive and post his quotes becomes the HFM.*
(ii) *In the event of LT arrival, no cancellations or new limit orders are submitted.*
(iii) *In the event of news arrival, the HFTs send orders as described in Lemma 1.*

4. HFT speed, auction duration, and liquidity

The equilibrium decribed in Proposition 1, as well as the results in Lemmas 1 and 3 allow us to study the impact of auction duration and HFT speed on liquidity, as measured by quoted spreads.

Corollary 1 *The equilibrium half-spread $s^\star$ increases in the size of value innovations, σ, news intensity, η, and decreases in the liquidity trader's arrival intensity, μ.*

Corollary 1 serves as a "reality check" and is consistent with results in the literature.[12] The positive spread emerges as a compensation for the adverse selection cost of being sniped by an HFS and is thus proportional to sniping profits. Sniping profits increase in the news intensity η and news size σ, and so does the equilibrium spread. A larger μ increases the HFM payoff from providing liquidity and therefore the spread decreases.

Proposition 2 states the main result of the paper. In periodic batch auction markets, a higher HFT speed improves liquidity. Proposition 2 describes the behavior of the equilibrium spread $s^\star$ with respect to the batch auction duration τ, the number of HFTs N, and the HFT speed ψ.

Proposition 2 *The equilibrium half-spread in the batch auction market, $s^\star$, decreases in (i) the HFT speed ψ, (ii) the duration of an auction τ, and (iii) the number of HFTs N.*

Faster HFTs are more likely to participate in any given auction. Keeping auction duration fixed, the probability that a *single* HFS arrives in a given auction interval decreases in HFT speed. If HFTs are faster, they are more likely to arrive in clusters and drive the clearing price to the fundamental value. At the same time, a faster HFM is also more likely to arrive and cancel quotes before being sniped.

Auction duration has an equivalent impact on equilibrium liquidity. Longer batch auctions, for a fixed HFT speed, increase the probability of either (i) the HFM cancels quotes before being sniped or (ii) two or more HFSs compete over the stale quote. Consequently, sniper auctions become relatively less likely, leading to lower adverse selection costs for the HFM and a lower spread.

[11] The intuition in this paragraph is identical to the Proof of Proposition 1 from Menkveld and Zoican (2017, p. 1217).
[12] For a detailed survey see, for example, Biais et al. (2005).

The number of competing HFTs and the HFT speed have a similar effect. Therefore, longer batch auctions and faster trading are substitutes for market liquidity.

Budish et al. (2015) argue that to encourage competition between HFTs, the auction length should be "long" enough. However, slowing down the market clearing frequency may create a negative externality for impatient traders, such as option traders who hedge in real time, or institutions with large positions to execute. We suggest that promoting low trading latencies, or at least not explicitly discouraging them, is a solution that can sustain both frequent market clearing and high liquidity. Fig. 3 illustrates the main result in Proposition 2: if HFTs are slow relative to the auction frequency, the equilibrium spread converges to the outcome on a continuous-time market. Conversely, the equilibrium spread decreases to zero if HFTs can react very fast relative to the duration of auction intervals, which is essentially the result in Budish et al. (2015).

Finally, the equilibrium spread decreases in the number of HFTs N. That is, stronger arbitrageur competition improves liquidity. As N becomes larger, two effects come into play. First, the probability of a sniping auction (π_S) is larger, leading to higher adverse selection costs for the market maker. However, it is the second effect that dominates in equilibrium. As more HFTs join the industry, the opportunity cost of the market maker is lower, which reduces the spread. A larger N reduces each sniper's individual probability of trading on a stale quote, leading to a lower sniper profit. Since in equilibrium HFTs are indifferent between the two roles, a lower HFS profit implies a lower HFM profit. Consequently, market makers require a lower spread to be compensated for assuming a liquidity provision role. The result is in line with Brogaard and Garriott (2019), who document that HFT arbitrageur competition improves liquidity.

In the limit, as τ decreases towards zero (very frequent auctions), the batch auction market converges to a continuous-time market. In this case, π_S converges to $\frac{N-1}{N}$ and we recover the equilibrium in Budish et al. (2015) stating that the equilibrium spread on a continuous-time market does not depend on the number of HFTs.

In Budish et al. (2015), competition between snipers emerges even as the duration of an auction becomes arbitrarily small. Therefore, in their model, the discrete auction market does not converge to a continuous-time market as the auction interval goes to zero. Lemma 4 shows that if we allow for a finite HFT learning speed, we obtain an equilibrium that bridges the continuous- and discrete-time markets.

Lemma 4 *For any $N > 2$, the equilibrium half-spread $s^\star$ on the batch auction market converges to zero as auctions are infinitely long or HFTs become infinitely fast, that is,*

$$\lim_{\tau\to\infty} s^\star = \lim_{\psi\to\infty} s^\star = 0. \tag{9}$$

As the auction length goes to zero, the equilibrium half-spread $s^\star$ converges to the continuous-time market spread

$$\lim_{\tau\to 0} s^\star = \sigma\frac{\eta}{\mu+\eta}. \tag{10}$$

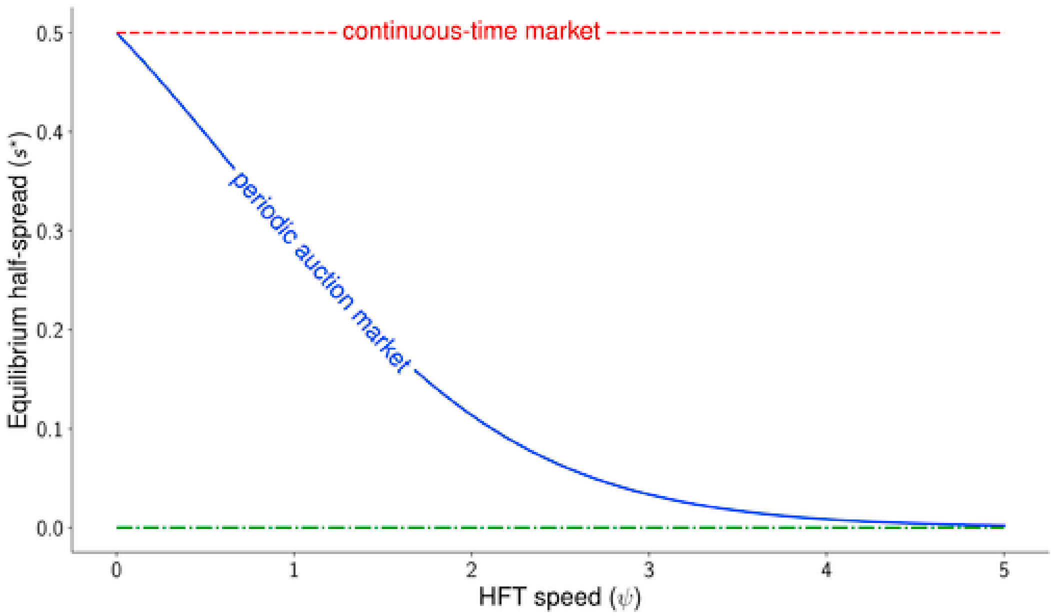

Fig. 3. Equilibrium half-spread. This figure illustrates the equilibrium half-spread ($s^\star$) on the periodic batch auction market as a function of the HFT speed parameter ψ. Parameter values: $\mu = \eta = \sigma = 1$, $\tau = 0.5$, and $N = 4$.

Lemma 4 implies that, while the periodic batch auction spread is positive, it is always lower than the spread on a continuous-time market, and converges to it as auctions become arbitrarily frequent. It follows that periodic batch auctions always improve liquidity.

The first limiting result in Lemma 4 is an exact counterpart to the batch auction market equilibrium in Budish et al. (2015). The authors assume "fast" HFTs act immediately on the information (infinite monitoring intensity generates a zero spread). Therefore, policies that intensify HFT monitoring (e.g., allowing for colocation) can reduce the equilibrium spread on a frequent call auction market. Fig. 4 illustrates that faster HFTs earn lower profits in equilibrium, as sniping opportunities become less frequent.

5. Discussion

In this section, we discuss the robustness of the model analysis. First, we consider the incentives of traders to invest in speed technology, when trading on a periodic call auction market. Second, we examine the effect of the tick size on the equilibrium spread. Finally, we discuss the possible implications of multiple "trigger" events in one auction.

5.1. Implications for the HFT arms race

In our model, the HFT speed parameter ψ is exogenous. In this subsection, we discuss the implications of allowing HFTs to invest in speed technology on a periodic auction market. Fig. 4 illustrates that the equilibrium HFT profit decreases in trading speed (ψ). A natural consequence of our model, in line with the intuition in Budish et al. (2015), would be that a frequent auction market reduces traders' incentives to invest in speed and alleviates the HFT arms race.

To better understand the economic mechanism, suppose that at the start of the game, each HFT i can choose individual speed level ψ_i at a cost $\mathscr{C}\left(\psi_i\right)$, where $\mathscr{C}'(\cdot)>0$ and $\mathscr{C}''(\cdot)>0$. Let $\Pi\left(\psi_i(\tau),\tau;\psi_{-i}\right)$ denote the ex ante expected profit of the HFT i, where ψ_{-i} is the vector of competitors' speed choices. When making the decision to invest in the reaction speed ψ_i the trader's objective is to maximize the net expected profit, for a given ψ_{-i},

$$\max_{\psi_i} \Pi\left(\psi_i(\tau),\tau;\psi_{-i}\right) - \mathscr{C}\left(\psi_i(\tau)\right). \tag{11}$$

The optimal speed investment $\psi_i^\star$ is determined by the first order condition of this problem:

$$\frac{\partial \Pi\left(\psi_i^\star(\tau),\tau;\psi_{-i}\right)}{\partial \psi_i} - \frac{\partial \mathscr{C}\left(\psi_i^\star(\tau)\right)}{\partial \psi_i} = 0, \tag{12}$$

where the competitors' speed levels, ψ_{-i}, are held constant from the perspective of HFT i. In any symmetric equilibrium, it needs to be that $\psi_i^\star = \psi_{-i}^\star$; that is, a fixed point of the HFTs' investment reaction function.

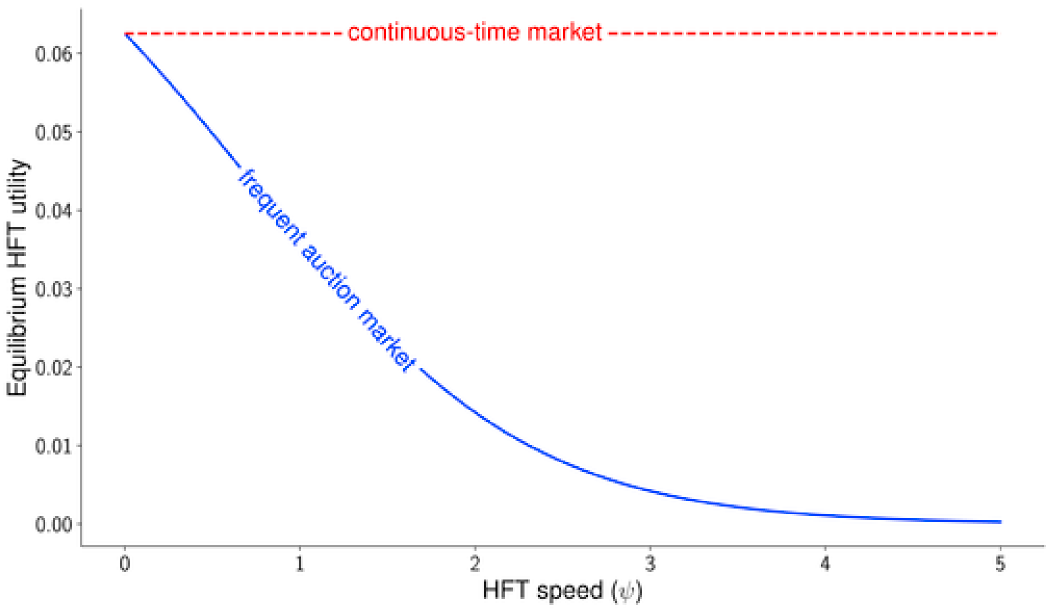

Fig. 4. Equilibrium HFT profit. This figure illustrates the equilibrium HFT profit on the periodic batch auction market as a function of the HFT speed parameter ψ. Parameter values: $\mu = \eta = \sigma = 1$, $\tau = 0.5$, and $N = 4$.

From the implicit function theorem, by taking the derivative of the optimality condition with respect to τ, we obtain:

$$\frac{\partial \psi_i^\star}{\partial \tau} = \frac{\frac{\partial^2 \Pi\left(\psi_i^\star(\tau), \tau; \psi_{-i}^\star\right)}{\partial \psi_i \partial \tau}}{\frac{\partial^2 \mathscr{C}\left(\psi_i^\star(\tau)\right)}{\partial \psi_i^2} - \frac{\partial^2 \Pi\left(\psi_i^\star(\tau), \tau; \psi_{-i}^\star\right)}{\partial \psi_i^2}}. \tag{13}$$

First, $\frac{\partial \mathscr{C}(\psi_i(\tau^\star))}{\partial \psi_i} > 0$ and $\frac{\partial^2 \mathscr{C}(\psi_i(\tau^\star))}{\partial \psi_i^2} > 0$ since faster trading requires additional investment and we assume that the cost increases more and more per incremental increase in reaction speed (the cost function is convex). Second, the term $\frac{\partial^2 \Pi\left(\psi_i^\star(\tau), \tau; \psi_{-i}^\star\right)}{\partial \psi_i^2}$ is negative, due to the decreasing returns to scale of the speed technology. Holding the competitors' speed constant at $\psi_{-i}^\star$, the probability of HFT i to be present in each given auction converges to a maximum of one. Finally, $\frac{\partial^2 \Pi\left(\psi_i^\star(\tau), \tau; \psi_{-i}^\star\right)}{\partial \psi_i \partial \tau} < 0$: intuitively, as the auction duration increases (higher τ), the marginal return from speed is lower as sniper auctions become less likely. It follows that $\frac{\partial \psi_i^\star}{\partial \tau} < 0$: a higher auction duration reduces investment in speed technology, therefore de-escalating the HFT arms' race.

5.2. Discrete price grid

An implicit assumption of the model is that HFTs can submit orders on a continuous price grid: That is, the equilibrium spread $s^\star$ defined in equation (8) is precisely pinned down by the HFT indifference condition between the two roles of market maker and speculator. From equations (5) and (6), the HFS (HFM) utility is linearly decreasing (increasing) in the half-spread s.

In reality, however, prices are likely to be constrained to a discrete grid. We normalize $v = 0$ and denote the tick size as $\Delta > 0$, such that the feasible prices for buy and sell orders are $\{0, \Delta, 2\Delta, \dots, \kappa\Delta, \dots\}$. We focus on the ask quote, that is the quote at which the HFM stands ready to sell the asset (the case of the bid quote is symmetric). If $v = 0$, then from Proposition 1 the equilibrium ask price is equal to the half-spread $s^\star$. Assume $s^\star$ falls in between two consecutive feasible prices $\kappa\Delta$ and $(\kappa + 1)\Delta$, as illustrated in Fig. 5.

We next analyze the two equilibrium candidate ask quotes, $s' = \kappa\Delta$ and $s'' = (\kappa + 1)\Delta$, following the same process as in Subsection 3.2. First, we note that s' cannot be an equilibrium: if $N - 1$ HFTs quote $s' = \kappa\Delta$, the Nth HFT can profitably deviate by choosing not to submit a quote as she earns a higher expected utility as a speculator. However, no HFT can profitably deviate from posting an ask quote with half-spread s''. If an HFT submits a larger half-spread, for example, $s^+ = s'' + \Delta$, then she does not become the HFM and will earn the strictly lower HFS expected utility. Conversely, if an HFT submits a smaller half-spread (e.g., s'), then she is sure to assume the HFM role, but will earn lower utility.

The equilibrium ask quote therefore is the lowest price on the grid that exceeds $s^\star$: that is, there is a minimum spread equal to the tick size – even if the unconstrained spread $s^\star$ becomes arbitrarily low due to high trading speed or longer auctions.

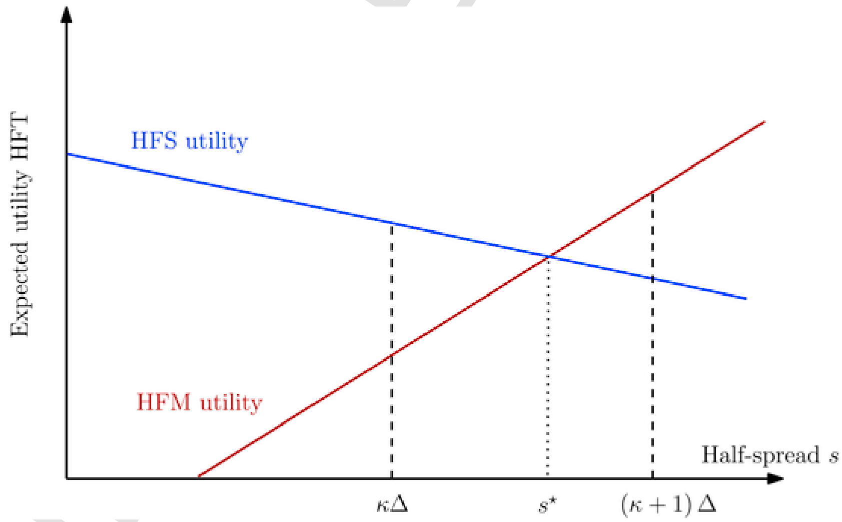

Fig. 5. Equilibrium spread on a price grid.

5.3. Multiple events in one auction

In our model, a single news event or a single liquidity trader may arrive (i.e., the trigger event). A richer setting could allow for a second event in between the "trigger" and market clearing. Menkveld and Zoican (2017, Lemma 1 on p. 1203) provide useful insights into the impact of a second news or LT event on adverse selection costs and the equilibrium spread.

First, since HFTs are risk-neutral and the asset value is a martingale, a second news event does not impact the equilibrium spread. For example, conditional on a "bad" news trigger, it is equally likely that the second news is also "bad" (doubling the potential HFM loss) or "good" (that is, a value reversal that eliminates adverse selection risk altogether).

Further, allowing for a subsequent LT arrival would lead to even lower spreads in markets with lower auction intervals. The rationale is that, since LT arrivals are uncorrelated with news, an LT trade on the no-news side compensates the market maker for a potential sniping loss.

6. Conclusion

We find that in periodic batch auctions in contrast to a continuous-time limit order market, faster trading improves liquidity by stimulating competition across high-frequency snipers. A lower HFT speed, or shorter auction interval, increases the probability that each individual HFT has a monopoly on information in any given auction. This allows a single sniper to take advantage of the stale HFM quotes and consequently to earn rents. Conversely, faster HFTs are more likely to become informed in the same auction and compete to trade on arbitrage opportunities. Such competition reduces arbitrage rents and adverse selection costs: as a consequence, the equilibrium spread decreases. Longer batch auctions, for a fixed HFT speed, have the same effect.

Importantly, the equilibrium in our model bridges two alternative market designs: continuous and discrete markets. As the auctions become more frequent and the auction length decreases to zero, the equilibrium spread converges to the continuous-time market counterpart. Conversely, as the auction length increases or HFTs become infinitely fast, the spread on the batch auction market converges to zero as in Budish et al. (2015).

Our findings contribute to the public debate on high-frequency trading and alternative market designs. It generates an important implication: Policies that aim at intensifying HFT monitoring (e.g., allowing for colocation) can reduce the equilibrium spread and make batch call auction markets more attractive for investors. This in turn implies that the "arms race" implication strikingly differs on batch auction markets when compared to limit order markets. Whereas for limit order markets speed does not impact liquidity, it turns out to be beneficial in the case of periodic batch auction markets.

Appendix.

A. Notation summary

Parameter	Definition
v	Common value of the risky asset.
τ	Length of auction interval (auction duration).
η	Poisson intensity of news arrival.
μ	Poisson intensity of LT arrival.
ψ	Poisson intensity of HFT news monitoring (HFT speed).
σ	Size of news, i.e., common value innovations.
N	Number of high-frequency traders.

B. Proofs

Lemma 1

Proof The proof is presented in the main body of the paper. □

Lemma 2

Proof To sign the derivative of $\pi_S = \frac{(N-1)\alpha(1-\alpha)^{N-1}}{1-(1-\alpha)^N}$ with respect to ψ, we first compute the first order derivative of π^S with respect to α:

$$\frac{\partial \pi_S}{\partial \alpha} = -\frac{(N-1)(1-\alpha)^{N-2}\left((1-\alpha)^N + N\alpha - 1\right)}{\left((1-\alpha)^N - 1\right)^2}. \tag{B.1}$$

It is enough to show that $f(\alpha) \equiv (1-\alpha)^N + N\alpha - 1 \geq 0$. This is true since $f(0) = 0$ and $\frac{\partial f}{\partial \alpha} = N\left[1-(1-\alpha)^{N-1}\right] > 0$. Therefore, π_S decreases in α. Since $\frac{\partial \alpha}{\partial \psi} > 0$, we conclude that π_S decreases in ψ. □

Lemma 3

Proof The equilibrium spread obtains immediately from setting equal the market-maker utility in (5) and the sniper utility in (6). □

Proposition 1

Proof The trading game equilibrium follows from Lemmas 1 and 3. □

Corollary 1

Proof We compute the first order derivatives of $s^{\star}$ with respect to σ, η, and μ.

$$\frac{\partial s^{\star}}{\partial \sigma} = \frac{\eta N \pi_S}{\eta N \pi_S + \mu(N-1)} > 0 \tag{B.2}$$

$$\frac{\partial s^{\star}}{\partial \eta} = \frac{\mu(N-1) N \pi_S \sigma}{(\eta N \pi_S + \mu(N-1))^2} > 0 \tag{B.3}$$

$$\frac{\partial s^{\star}}{\partial \mu} = -\frac{\eta(N-1) N \pi S \sigma}{(\eta N \pi S + \mu(N-1))^2} < 0. \tag{B.4}$$

It follows that $s^{\star}$ increases in σ and η and decreases in μ, which concludes the Proof. □

Proposition 2

Proof We first show that the equilibrium spread increases in the sniping probability π_S. Indeed,

$$\frac{\partial s^{\star}}{\partial \pi_S} = \frac{\eta\mu(N-1)N\sigma}{(\eta N \pi S + \mu(N-1))^2} > 0. \tag{B.5}$$

From Lemma 2, the sniping probability π_S decreases in the per-auction arrival probability of an HFT, α. Further, from equation (1), α increases in ψ and τ. Therefore,

$$\frac{\partial s^{\star}}{\partial \psi} = \underbrace{\frac{\partial s^{\star}}{\partial \pi_S}}_{>0} \underbrace{\frac{\partial \pi_S}{\partial \alpha}}_{<0} \underbrace{\frac{\partial \alpha}{\partial \psi}}_{>0} < 0 \text{ and} \tag{B.6}$$

$$\frac{\partial s^{\star}}{\partial \tau} = \underbrace{\frac{\partial s^{\star}}{\partial \pi_S}}_{>0} \underbrace{\frac{\partial \pi_S}{\partial \alpha}}_{<0} \underbrace{\frac{\partial \alpha}{\partial \tau}}_{>0} < 0 \tag{B.7}$$

Therefore, the equilibrium spread $s^{\star}$ decreases in the HFT speed ψ and the duration of an auction τ.Next, we compute the partial derivative of $s^{\star}$ with respect to N:

$$\frac{\partial s^{\star}}{\partial N} = \frac{\eta\mu\alpha\sigma(1-\alpha)^{N+1}\left(1-(1-\alpha)^N + N\log(1-\alpha)\right)}{\left(\mu + (1-\alpha)^N(\eta N\alpha + \mu(\alpha-1)) - \mu\alpha\right)^2}. \tag{B.8}$$

The sign of the derivative in (B.8) is given by the sign of:

$$g \equiv 1-(1-\alpha)^N + N\log(1-\alpha).$$

We note that $g\,(\alpha = 0) = 0$ and that g decreases in α since:

$$\frac{\partial g}{\partial \alpha} = -\frac{N\left(1-(1-\alpha)^N\right)}{1-\alpha} < 0. \tag{B.9}$$

It follows that $g < 0$ and consequently $\frac{\partial s^{\star}}{\partial N} < 0$ which implies that the equilibrium spread decreases in the number of HFTs. □

References

Aldrich, E M, López Vargas, K, 2019. Experiments in high-frequency trading: comparing two market institutions. Exp. Econ. 23, 322–352.
Aoyagi, J, 2019. Speed Choice by High-Frequency Traders with Speed Bumps. Manuscript.
Baldauf, M, Mollner, J, 2020. High-frequency trading and market performance. J. Finance 75, 1495–1526.
Baruch, S, Glosten, L R, 2013. Fleeting Orders. Columbia Business School Research Paper 13-43.

Biais, B, Foucault, T, Moinas, S, 2015. Equilibrium fast trading. J. Financ. Econ. 116, 292–313.
Biais, B, Glosten, L, Spatt, C, 2005. Market microstructure: a survey of microfoundations, empirical results, and policy implications. J. Financ. Mark. 8, 217–264.
Brogaard, J, Garriott, C, 2019. High-frequency trading competition. J. Financ. Quant. Anal. 54, 1469–1497.
Brogaard, J, Hendershott, T, Riordan, R, 2014. High frequency trading and price discovery. Rev. Financ. Stud. 27, 2267–2306.
Brolley, M, Cimon, D, 2020. Order flow segmentation, liquidity and price discovery: the role of latency delays. J. Financ. Quant. Anal. (forthcoming).
Brolley, M, Zoican, M, 2020. Liquid Speed: A Congestion Fee for Low-Latency Exchanges. Rotman School of Management Working Paper No. 3377346.
Budish, E, Cramton, P, Shim, J, 2015. The high-frequency trading arms race: frequent batch auctions as a market design response. Q. J. Econ. 130, 1547–1600.
Copeland, T E, Galai, D, 1983. Information effects on the bid-ask spread. J. Finance 38, 1457–1469.
Du, S, Zhu, H, 2017. What is the optimal trading frequency in financial markets? Rev. Econ. Stud. 84, 1606–1651.
Economides, N, Schwartz, R A, 1995. Electronic call market trading. J. Portfolio Manag. 21, 10–18.
Farmer, J D, Skouras, S, 2012. Review of the Benefits of a Continuous Market vs. Randomised Stop Auctions and of Alternative Priority Rules Tech. rep.. United Kindgom's Government Office for Science.
Foucault, T, Kadan, O, Kandel, E, 2013. Liquidity cycles and make/take fees in electronic markets. J. Finance 68, 299–341.
Foucault, T, Roell, A, Sandas, P, 2003. Market making with costly monitoring: an analysis of the soes controversy. Rev. Financ. Stud. 16, 345–384.
Fricke, D, Gerig, A, 2018. Too fast or too slow? Determining the optimal speed of financial markets. Quant. Finance 18, 519–532.
Glosten, L R, Milgrom, P R, 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. J. Financ. Econ. 14, 71–100.
Hendershott, T, Moulton, P C, 2011. Automation, speed, and stock market quality: the NYSE's Hybrid. J. Financ. Mark. 14, 568–604.
Janssen, M, Rasmusen, E, 2002. Bertrand competition under uncertainty. J. Ind. Econ. 50, 11–21.
Jovanovic, B, Menkveld, A J, 2019. Equilibrium Bid-Price Dispersion. Manuscript.
Kandel, E, Rindi, B, Bosetti, L, 2012. The effect of a closing call auction on market quality and trading strategies. J. Financ. Intermediation 21, 23–49.
Kremer, I, Nyborg, K G, 2004. Underpricing and market power in uniform price auctions. Rev. Financ. Stud. 17, 849–877.
Kumar, P, Seppi, D J, 1994. Information and index arbitrage. J. Bus. 67, 481–509.
Kyle, A P S, Lee, J, 2017. Toward a fully continuous exchange. Oxf. Rev. Econ. Pol. 33, 650–675.
Madhavan, A, 1992. Trading mechanisms in securities markets. J. Finance 47, 607–641.
Menkveld, A, Zoican, M, 2017. Need for speed? Exchange latency and liquidity. Rev. Financ. Stud. 30, 1188–1228.
Menkveld, A J, 2014. High frequency traders and market structure. Financ. Rev. (Special Issue: Computerized and High-Frequency Trading) 49, 333–344.
Menkveld, A J, 2016. The economics of high-frequency trading: taking stock. Ann. Rev. Financial Econ. 8, 1–24.
Pagano, M S, Peng, L, Schwartz, R A, 2013. A call auction's impact on price formation and order routing: evidence from the NASDAQ stock market. J. Financ. Mark. 16, 331–361.
Pagano, M S, Schwartz, R A, 2003. A closing call's impact on market quality at Euronext Paris. J. Financ. Econ. 68, 439–484.
Roth, A E, Ockenfels, A, 2002. Last-minute bidding and the rules for ending second-price auctions: evidence from eBay and Amazon auctions on the internet. Am. Econ. Rev. 92, 1093–1103.
Shkilko, A, Sokolov, K, 2019. Every cloud has a silver lining: fast trading, Microwave connectivity and trading costs. J. Finance (forthcoming).
Wah, E, Hurd, D, Wellman, M, 2015. Strategic market choice: frequent call markets vs. continuous double auctions for fast and slow traders. In: Proceedings of 3rd EAI Conference.
Wah, E, Wellman, M, 2013. Latency arbitrage, market fragmentation, and efficiency: a two-market model. In: Proceedings of the 14th ACM Conference.
Ye, M, Yao, C, Gai, J, 2013. The Externalities of High Frequency Trading. WBS Finance Group Research Paper 180.