Journal of Financial Markets 66 (2023) 100853

Contents lists available at ScienceDirect

Journal of Financial Markets

journal homepage: www.elsevier.com/locate/finmar

The market quality implications of speed in cross-platform trading: Evidence from Frankfurt-London microwave☆

Khaladdin Rzayev [a,b,c,*], Gbenga Ibikunle [a,d], Tom Steffen [e]

[a] *University of Edinburgh, United Kingdom*
[b] *Koç University, Turkey*
[c] *Systemic Risk Centre; London School of Economics, United Kingdom*
[d] *European Capital Markets Cooperative Research Centre, Pescara, Italy*
[e] *Osmosis Investment Management, United Kingdom*

A R T I C L E I N F O

JEL classification:
G10
G14

Keywords:
Transmission latency
Microwave connection
High-frequency trading
Liquidity
Price discovery

A B S T R A C T

Exploiting information transmission latency between stock exchanges in Frankfurt and London, and speed-inducing technological upgrades, we show that when cross-market latency arbitrage opportunities are linked to the arrival of information, high-frequency traders' (HFTs') activities impair liquidity and enhance price discovery by facilitating the incorporation of public information into prices. Conversely, when cross-market latency arbitrage opportunities are driven by liquidity shocks, HFTs improve liquidity and reduce trading costs, thus incentivizing information acquisition and trading with private information. These findings underscore the complex nature of the association between trading speed and market quality and reconcile mixed evidence in the extant literature.

> The rise of high-frequency traders has opened up a debate among investors, brokers and exchanges. Critics have long claimed that speed-driven traders unfairly hurt traditional investors… Supporters argue that faster traders are now a vital element of modern markets….
>
> *Financial Times*, May 15, 2019

☆ For helpful comments, we are grateful to an anonymous referee, an anonymous advisory editor, the Editor (Gideon Saar), Bruno Biais, Maria Boutchkova, Marcos Carreira, Jo Danbolt, Jon Danielsson, Sean Foley, Angelica Gonzalez, Alexandre Laumonier, Ian Marsh, Han Ozsoylev, Talis Putnins, Mathieu Rosenbaum, Satchit Sagade, Avanidhar Subrahmanyam, Jean-Pierre Zigrand, as well as seminar/conference participants at the 2020 Annual Meeting of the American Finance Association, the 27th Conference on the Theories and Practices of Securities and Financial Markets, the Regulation and Operation of Modern Financial Markets Conference, The Future of Computer Trading in Financial Markets: has it happened? Conference, the 3rd European Capital Markets Workshop, Dublin City University, Koç University, London School of Economics, the University of Edinburgh, and Victoria University of Wellington. Khaladdin gratefully acknowledges the support of the Trans-Atlantic Platform by the Economic and Social Research Council (ESRC) [grant number ES/R004021/1].

* Corresponding author. University of Edinburgh, United Kingdom.
E-mail addresses: khaladdin.rzayev@ed.ac.uk (K. Rzayev), gbenga.ibikunle@ed.ac.uk (G. Ibikunle), tom.steffen@outlook.com (T. Steffen).

https://doi.org/10.1016/j.finmar.2023.100853
Received 1 January 2021; Received in revised form 7 June 2023; Accepted 1 July 2023
Available online 8 July 2023

1. Introduction

The speed of trading and, ultimately, of price adjustment is an important factor in the price discovery process. Today, this factor has a significance that transcends market quality implications. It is the driving force behind a recent upsurge in latency arbitrage trading strategies, defined as fast traders using their superior response speeds to exploit new symmetrically observable public information ahead of slower traders by executing against stale quotes (e.g., Budish et al., 2015). This suggests that the nature of the information that latency arbitrageurs exploit differs from the asymmetrically observable private information the informed trader exploits as modeled in classic market microstructure models (e.g., Glosten and Milgrom, 1985; Kyle, 1985). Aquilina et al. (2022) estimate that the deployment of latency arbitrage strategies results in losses of approximately $5 billion per year in global equity markets.

However, trading at high speeds can also be good for markets; the evidence on this has thus far been inconsistent.[1] On the one hand, the speed advantage of market-making high-frequency traders (HFTs) allows them to avoid adverse selection and inventory management risks and thus motivates them to provide more liquidity and increase price discovery (e.g., Hendershott et al., 2011; Brogaard et al., 2014b; Hoffmann, 2014); on the other hand, liquidity-consuming HFTs may use their speed advantage to pick off slow traders' orders, thereby imposing adverse selection costs on them and reducing their incentives for information acquisition (e.g., Biais et al., 2015; Foucault et al., 2016; Weller, 2018). As advanced by Foucault et al. (2017), the latency arbitrage opportunities that HFTs exploit can be toxic or non-toxic from the perspective of the slower traders. Conceptually, Foucault et al. (2017) define toxic latency arbitrage opportunities as those linked to asynchronous price adjustments that arise due to the arrival of new information and hence are a source of adverse selection risk for slow traders. Non-toxic latency arbitrage opportunities are in turn defined as those driven by temporary liquidity shocks; thus, HFTs may act as liquidity providers in exploiting them and reduce trading costs for slow traders. This implies that the effect of HFT activity or, more pointedly, the speed of trading, on market quality characteristics is not homogenous and dependent on whether the latency arbitrage opportunities that HFTs exploit are toxic or not. We therefore argue that the inconsistency observed in the literature is linked to the nature of the latency arbitrage opportunity HFTs exploit.

We explore this argument and attempt to reconcile the inconsistencies in the literature by investigating whether the effect of the latency, as an inverse proxy for the speed of trading, on market quality characteristics following the emergence of a toxic latency arbitrage opportunity differs from its impact following the onset of a non-toxic latency arbitrage opportunity. Our empirical approach centers on estimating transmission latency (*TL*), a variable that encapsulates the microwave/fiber optic connection latency between two trading venues/markets, traders' information execution time and exchange latency. This approach contrasts with previous studies where the focus has been on traders' execution speed alone. *TL* in this study is estimated between the Xetra Stock Exchange, Frankfurt (hereafter referred to as XSE), the home exchange/market, and Cboe Europe, London (hereafter referred to as Cboe), the satellite market, for a sample of the 100 most active XSE-listed German stocks that are also cross-listed on Cboe. When the prices of stocks traded on XSE are revised ahead of the prices of the same stocks traded on Cboe, because of their superior technological setup, fast traders can exploit the arbitrage opportunities these revisions present on Cboe. By using *TL* and deploying a two-stage least squares instrumental variable (2SLS IV) empirical framework that exploits three instruments, we characterize the effects of *TL* on the liquidity and price discovery dynamics of the cross-listed stocks in the satellite market. The three instruments used are: (i) two technological upgrades implemented by XSE; (ii) longitudinal data capturing the introduction of new microwave connections; and, in the spirit of Hasbrouck and Saar (2013), and (iii) the average *TL* in stock size groups.

For our sample period covering March 2017 to August 2018, we find that 49% (80%) of near-coincident price-changing trades on Cboe occur within 3 (5) milliseconds (ms) of similar and proportional price-changing trades on XSE. This means that the microwave and fiber optic connections enable price responses on Cboe within 3–5 ms of price changes on XSE. The estimates are also consistent with the latencies published by trading network infrastructure providers; for example, Perseus, one of the microwave connection providers between London and Frankfurt, states that the round-trip latencies via microwave and fiber optics between London and Frankfurt are 4.6 ms and 8.4 ms respectively.[2] We detect 1,927,000 latency arbitrage opportunities over the sample period; 677,328 (35.15%) of these opportunities are toxic, while the majority are classed as non-toxic. As a test of our classification algorithm, we compare the number of toxic and non-toxic latency arbitrage opportunities occurring on days with firm-relevant news and on days with no recorded firm-relevant news. To the extent that toxic arbitrage opportunities are driven by the arrival of new information, our expectation is that toxic arbitrage opportunities should be observed more frequently on days with firm news. Consistent with this, we find that 584,193 (86.25%) of the 677,328 toxic arbitrage opportunities occur during news days, while the corresponding percentage value for non-toxic arbitrage opportunities is 51.43%.

More importantly, our analysis shows that, as predicted by Foucault et al. (2017), the impact of HFTs on market quality characteristics is not homogeneous and depends on the type of cross-market latency arbitrage opportunities they are exploiting. HFT activity is linked to the deterioration of liquidity when HFTs exploit (toxic) latency arbitrage opportunities arising from asynchronous price adjustments to news. A 1 ms increase in latency reduces effective and quoted spreads by about 4.2% and 5.7%, respectively, following

[1] In this paper, we use speed and speed differentials interchangeably. This is because, as argued by Menkveld and Zoican (2017), any improvements in (exchange) speed will only directly impact a particular portion of traders (i.e., HFTs), although these improvements can be exploited by all traders. Brogaard et al. (2014a) also suggest that millisecond changes in speed will only have a direct impact on HFTs (fast traders). Thus, all fast traders will benefit from the latency improvements that we focus on in this study, which are measured in milliseconds, and they will increase speed differentials between fast and slow traders.

[2] https://www.reuters.com/article/us-highfrequency-microwave/lasers-microwave-deployed-in-high-speed-trading-arms-race-idUSBRE9400L920130501.

incidences of toxic latency arbitrage opportunities. However, one consequence of HFTs' exploitation of toxic latency arbitrage opportunities is that they help to convey information observed at the lead market (XSE) to the satellite market (Cboe) by trading on them as private information on the satellite market. We hypothesize that this process may cause market makers to update their quotes following toxic arbitrage opportunities to reflect the exploited private information and, hence, market maker action turns private information sourced from HFT activity into public information. Consistent with this argument, we find that a 1 ms increase in latency reduces the public information component of the price discovery by about 10.79% following incidences of toxic latency arbitrage opportunities.

Our results also show that HFTs' exploitation of latency arbitrage opportunities is beneficial for liquidity when these opportunities are driven by price pressures (i.e., when they are non-toxic); the economic magnitude of the impact of a 1 ms increase in latency yields an approximate increase of 18.5% for the effective spread and 17.0% for the quoted spread following episodes of non-toxic latency arbitrage opportunities. The linked reduction in transaction costs also incentivizes the acquisition of new information and thus increases the private information component of the price discovery process; a 1 ms increase in latency decreases the share of the private information of price discovery by 7.83% following episodes of non-toxic latency arbitrage opportunities.

The reported effects of latency on market quality characteristics, however, need to be set in the proper context. Significant reductions in latency of up to 1 ms are unlikely given that the current technological set ups in financial markets are close to the theoretical limit. For example, in the winter of 2012, Traderworx's microwave network linking New York City to Chicago was only able to shave an estimated 0.5 ms off Mackay Brothers' competing network, which opened months earlier. Factoring in that the straight line distance between Chicago and New York City is 711.99 miles underscores the challenge of achieving new meaningful reductions in transmission latencies in financial markets today. Hence, while, in line with standard practice, we report elasticities between a unit latency estimate (1 ms) and other variables of interest, it is crucial to recognize that these effects are unlikely to be realized at the magnitudes reported. Indeed, the largest impact of the two latency impacting technological upgrades we exploit as instruments corresponds to a 0.105 ms reduction in latency.

This study extends the literature on the impact of speed on market quality characteristics (e.g., Menkveld, 2016; Boehmer et al., 2018; Shkilko and Sokolov, 2020; Boehmer et al., 2021) by investigating the speed-market quality relationship in a cross-market setting and advancing a reconciliation of the mixed evidence in the extant literature. In particular, our study makes three main contributions to the literature.

First, we present significant new insights on the complexity of the relation between speed and market quality characteristics (liquidity and price discovery) by showing that the impact of HFT activity on market quality in the fragmented market environment that is the reality of modern trading depends on the type of latency arbitrage opportunities that HFTs exploit. Our analysis therefore complements Foucault et al. (2017), who propose a model in which liquidity depends on the latency arbitrage strategies deployed by fast traders in a cross-market setting. While Foucault et al. (2017, p. 1090) present some "suggestive evidence" on how the reaction speed of fast traders to "only one type of the … high-frequency arbitrage opportunities exploited by high-speed arbitrageurs" impacts liquidity, we exploit granular data and a set of novel liquidity measures to quantify market makers' spread-setting dynamics in response to both toxic and non-toxic latency arbitrage opportunities and their implications for price discovery. Thus, we empirically isolate the effects of latency on liquidity and price discovery in the aftermath of both toxic and non-toxic latency arbitrage opportunities, and respond to Foucault et al.'s (2017, p. 1090) call for "more research … to establish the robustness of our conclusions for other high-frequency arbitrage opportunities …".

Secondly, our latency measure, *TL*, and the empirical framework we use allow us to capture the cross-market latency arbitrage strategies of a broader spectrum of HFTs than in previous studies. The commonly used proxy for AT/HFT proxy in the literature is the normalized electronic message traffic (e.g., Hendershott et al., 2011; Malceniece et al., 2019; Boehmer et al., 2021). For instance, Boehmer et al. (2021) study the effects of trading speed on market quality using the ratio of the number of messages to the number of transactions as a measure of AT; they capture the exogenous variation in that measure using the introduction of colocation services as an instrument for such trading. The main issue with the normalized electronic message traffic measure is that it mainly captures AT-based market-making strategies (Hendershott et al., 2011). Brogaard et al. (2015) also note that colocation services are predominantly used by market-making HFTs. Boehmer et al. (2018) further show that these metrics are sensitive to the length of the interval they are computed at. Therefore, the measures often used in the literature are limited in terms of the scope of the insights they offer. This point is reinforced by Boehmer et al.'s (2018) classification of HFT strategies/product categories, which underscores the heterogeneity endemic in HFT. Thus, by proposing a new latency measure (*TL*) that simply captures HFTs' reaction speeds (see also Baron et al., 2019), and computing market quality proxies following incidences of toxic and non-toxic latency arbitrage opportunities, our empirical framework allows for a more comprehensive investigation of the association between the activities of HFTs and market quality characteristics.

Finally, we provide evidence on the origins of the speed-price discovery relation, an important insight that has been neglected in previous studies (e.g., Hasbrouck and Saar, 2013; Boehmer et al., 2018; Roşu, 2019; Boehmer et al., 2021). Indeed, while Boehmer et al. (2021) document higher AT-induced price volatility, their analysis is agnostic about the nature of such high volatility. However, consistent with our analysis of the contextual relation between speed and liquidity, understanding the drivers of the association between trading speed and volatility is crucial to considering the research-driven evidence in the right context. Therefore, we first estimate the private and public information components of the price discovery process separately, and then investigate the effects of HFT activity on them following incidences of toxic and non-toxic latency arbitrage. As such, this study also complements Brogaard et al. (2014b) and Boehmer et al. (2018), who mainly examine the impact of HFTs on permanent (aggregate information) and transitory price changes (pricing errors). We extend these studies by dividing information into its public and private components and are thus able to develop a more nuanced understanding of the main sources of information used by HFTs in a cross-market context. More importantly, we present novel evidence showing that the impact of HFTs on price discovery is not homogeneous but depends on the types of latency arbitrage opportunities that HFTs exploit.

The rest of the paper is organized as follows. Section 2 discusses the institutional and technical aspects of information transmission between trading venues. Section 3 describes the data used and the empirical properties of the latency arbitrage classification procedure. Results are presented in Section 4, and Section 5 concludes the paper.

2. Institutional and technical background

In today's increasingly fragmented markets, the speed of information transmission between trading venues plays an important role in facilitating price discovery. A decade ago, the most common way to transmit information from Frankfurt to London was via a fiber optic cable However, with fiber optic technology, "information" (photons) travels through cables and laying them in a straight line is difficult. According to Shkilko and Sokolov (2020), until 2010, the fiber optic cabling between Chicago and New York City exceeded the straight-line distance between the two cities by about 200 miles. With microwave technology, "information" travels through the air, offering transmission speeds that are 30%–50% faster than fiber optic technology, taking about 1.9 ms off the information transmission latency between Frankfurt and London, a reduction from 4.2 ms to 2.3 ms.[3] It is therefore not surprising that the past decade has seen the emerging operation of microwave networks between London and Frankfurt.[4] Some of these networks are operated by specialist network providers (e.g., McKay Brothers), while others are operated directly by HFTs (e.g., Jump Trading LLC, a Chicago-based company founded by former pit traders).

Microwave networks rely on the installation of dishes on towers that need to be positioned to establish a straight line — accounting for the Earth's curvature — between two points. In the case of HFTs, these two points are exchanges. This suggests that HFTs must compete over prime real estate, which requires the construction of towers in direct proximity to the exchanges. This competition for space also extends to competing interests involving mobile and radio operators. The feverish bidding among HFTs at the auction for an abandoned US/NATO military radio relay station in Houtem, Belgium aptly captures the frenzy around prime microwave infrastructure. Crucially, the ~1.31-ha former military post came with a 243.5 m communications tower that had at one time transmitted messages for the U.S. military. The tower offers an ideal foundation for establishing microwave connections between exchanges in mainland Europe and London, a possibility that was likely not apparent to the Belgian Ministry of Defense when it offered the station for €400,000 in December 2012.[5] The auction, which drew interest from established HFTs, eventually earned the ministry €5 million. The successful bidder was Jump Trading LLC. Intrigue surrounding the sale of old infrastructure is not unusual in HFT circles; in fact, the acquisition of former NATO towers in Europe by HFTs is an ongoing trend.[6]

3. Data, measures, and summary statistics

3.1. Data

Our primary data source is the Refinitiv Tick History (RTH) v2 from the DataScope database. The most important feature of this dataset is that it provides exact exchange timestamps, which are synchronized to UTC during the sample period, and are in milliseconds for exchange-traded transactions and order flow. It consists of ultra-high-frequency tick-by-tick data for the 100 most active German stocks that trade on both XSE in Frankfurt (home market) and Cboe in London (satellite market). The dataset includes transaction-level data for trading days between March 2017 and August 2018. We select this period for two reasons. First, Refinitiv/DataScope does not provide exchange timestamps for European markets before June 2015. Second, as noted above, to address potential endogeneity concerns, we employ a 2SLS IV approach using the two technological upgrades as instruments (discussed in Subsection 3.2). The upgrade dates are July 3, 2017 and April 9, 2018. We then select a data coverage period spanning a 120-day window before and after the upgrade dates for the 2SLS IV framework. The DataScope data contain standard transaction-level variables, such as date, time (both RTH and exchange timestamps), price, volume, prevailing bid price, prevailing ask price, prevailing bid volume, and prevailing ask volume.

To capture the evolution of HFT competition in terms of transmission speed, we also obtain data on the issuance of microwave connection licenses as granted by the United Kingdom's communications regulator, Ofcom, for the period from 2001 to 2018. The data obtained include the date of the license issue, license number, network status (whether active or not), network location, and licensee.[7] To obtain the data, we first identify the regions where microwave connections between XSE and Cboe potentially cross, namely Faversham (Dunkirk), Slough, Dover, Basildon, Ramsgate, Canterbury, and Sandwich (Richborough). Second, to ascertain that we are focusing on the microwave licenses used by HFTs, we obtain the list of licensee companies for these regions. From 2001 to 2018, a total of 346 licenses were granted to 80 licensee companies. We cross-check all 80 companies' public profiles with their customers, services, and related news from Factiva, leading to the identification of 12 companies involved in HFT activities. These companies are 12H AG, DRW NX UK Limited (Vigilant), Equinix (London) Limited, Equinix (UK) Limited, Flow Traders BV, GTT Communications BV, Gyron

[3] https://www.quincy-data.com/product-page/#latencies.
[4] https://www.bloomberg.com/news/articles/2014-07-15/wall-street-grabs-nato-towers-in-traders-speed-of-light-quest.
[5] https://sniperinmahwah.wordpress.com/.
[6] https://www.bloomberg.com/news/articles/2014-07-15/wall-street-grabs-nato-towers-in-traders-speed-of-light-quest.
[7] https://www.ofcom.org.uk/spectrum/information/spectrum-information-system-sis/spectrum-information-portal.

Internet Limited, McKay Brothers Communications Limited, McKay Brothers International SA, New Line Networks LLC (Jump), Optiver Holding BV, and Virtus (Data Centers) Limited.

Fig. 1 presents the number of microwave licenses obtained by the 12 HFT-linked companies from 2001 to 2018, which shows a sharp increase after 2013. While there are only nine microwave licenses in 2013, there are 150 by the end of 2018, highlighting that HFTs' participation in cross-platform trading between the two leading European financial centers intensifies substantially during our sample period. We use this evolution in HFTs' exploitation of microwave network connections in one of the approaches we adopt to instrument speed in our empirical framework (see Section 4).

3.2. Latency measurement

Latency can be considered as the delay between a signal and a response (e.g., Baron et al., 2019), and our latency (*TL*) estimation method encapsulates this view. In this method, we assume that information is transmitted from Frankfurt to London; therefore, the signal is the price-changing trade on XSE, and the response is the same-direction price-changing trade on Cboe. Laughlin et al. (2014) also define the signal as a price-changing trade in the home market, and the response as a near-coincident same-direction price-changing trade in the satellite market; they employ this method for futures-ETF pairs in the U.S. financial markets. We apply it to measure latency in the case of the 100 most active German stocks cross-listed on XSE and Cboe. The reasoning behind this method is that, according to the law of one price, the price of the cross-listed stocks should be the same regardless of location, and the difference between the prices of securities cross-listed in different exchanges should be instantaneously eliminated.

The above discussion implies that, by design, our *TL* measure is composed of the following elements: (i) the connection latency between XSE and Cboe; (ii) the exchange latency at Cboe; and (iii) the traders' execution latencies at Cboe. The connection latency is the time it takes for information to travel via microwave/fiber optic connection between XSE and Cboe. The exchange latency is the time it takes for the exchanges to process incoming and outgoing instructions. The transaction-level data from the RTH that we employ provide exact exchange timestamps for executed transactions. It therefore also accounts for the traders' execution latencies (i.e., their signal-processing and reaction times).

The expectation that information about German stocks largely flows from Germany is supported by the superior volume of transactions we find for XSE compared to Cboe, as reported in Table 1 and prior studies (e.g., Grammig et al., 2005); our approach to estimating *TL* reflects this expectation. However, as shown in Table 1, a non-negligible proportion of the averaged per stock trading activity in our sample also takes place in Cboe, which suggests that some relevant information may first be revealed on the London platform.

We explore this conjecture by estimating price leadership using Hasbrouck's (1995) information share metric (IS), Gonzalo and Granger's (1995) component share metric (CS), and the information leadership share metric (ILS) set out by Putniņš (2013).[8] Although the results (presented in Appendix A) are consistent with earlier studies in that price discovery mainly occurs on XSE for German stocks (0.69, 0.64, and 0.61 for IS, CS, and ILS estimations respectively), they also show that more than a third of the price discovery emanates from Cboe. Therefore, we extend our analysis to investigate the effects of latency when price change is driven by Cboe. We estimate *TL* from Cboe to XSE and replicate our full spectrum of analysis estimates on this basis. The results based on the assumption of XSE's price leadership are reported in subsequent sections as our main set of results, while the results based on Cboe having price leadership are presented in Appendix C; the latter set of results are also discussed in Subsection 4.3.

The latency measurement approach involves identifying the exact exchange timestamp for each price-changing trade on XSE or Cboe. We then look for a near-coincident same-direction price-changing trade on Cboe or XSE. To identify the near-coincident trade on Cboe, we examine trades occurring within 10 ms of each price-changing trade. We select the 10 ms interval given that the average information transmission latencies between Frankfurt and London are 2.3 ms and 4.2 ms for microwave and fiber optic connections respectively.[9] We follow Shkilko and Sokolov (2020) and only use the standalone signals (i.e., signals not preceded by another signal in the previous 10 ms) to reduce serial correlation effects.

Panel A in Table 2 reports the number of responses on Cboe to the signals on XSE for various latencies. We exclude the responses that fall into the 2 ms interval because the 2 ms interval is less than the theoretical minimum time it should take light to travel in a vacuum between the two locations. The number of responses in this interval only accounts for 2% of all responses; hence, the exclusion should not have any material impact on our analysis. Laughlin et al. (2014) argue that responses at less than the speed of light can be considered proof of the predictive capacity of HFTs. We do not examine this argument since it is outside the scope of this study.

Panel A in Table 2 shows that 48.61% (80.74%) of all responses (after excluding the 0–2 ms interval) fall within the 3 ms (5 ms) bin. These latencies are consistent with those provided by the microwave network and fiber optic connection providers and they corroborate the view that our latency measure indeed captures the transmission latency between the two trading venues. For example, McKay Brothers recently announced that their average microwave latency between the XSE (FR2) and Cboe (LD4) data centers is 2.3 ms (see Footnote 3). It is also generally accepted that the average latency via fiber optic connections is about 4.2 ms (see Footnote 2). These announced latencies — 2.3 ms and 4.2 ms — are only transmission latencies between exchanges and do not account for the exchange latency and traders' order execution latencies at Cboe. Therefore, we expect the actual trading latencies to be closer to our

[8] We obtain the SAS codes used in the computation of the price discovery metrics from the following website:http://pages.stern.nyu.edu/~jhasbrou/EMM%20Book/SAS%20Programs%20and%20Data/Description.html.

[9] https://www.reuters.com/article/us-highfrequency-microwave/lasers-microwave-deployed-in-high-speed-trading-arms-race-idUSBRE9400L920130501.

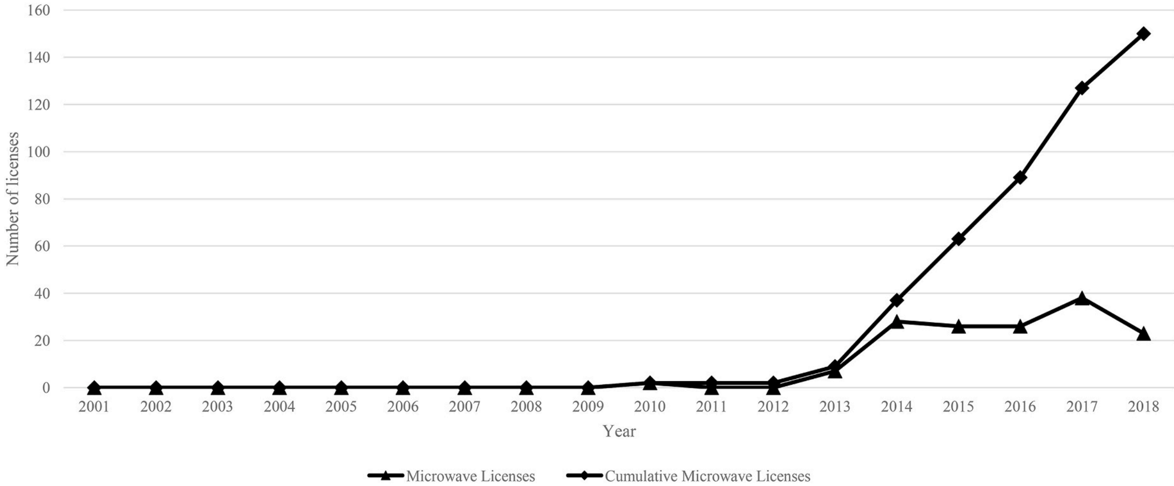

Fig. 1. The evolution of licenses obtained by HFT-linked companies.
This figure shows the time series of the number of active microwave licenses obtained by 12 companies linked to HFT activities in seven UK regions (Faversham (Dunkirk), Slough, Dover, Basildon, Ramsgate, Canterbury, and Sandwich (Richborough)). The 12 companies captured are: 12H AG, DRW NX UK Limited (Vigilant), Equinix (London) Limited, Equinix (UK) Limited, Flow Traders BV, GTT Communications BV, Gyron Internet Limited, McKay Brothers Communications Limited, McKay Brothers International SA, New Line Networks LLC (Jump), Optiver Holding BV, and Virtus (Data Centers) Limited. The sample period spans from 2001 to 2018, and the data are sourced from Ofcom.

Table 1
Transactions summary statistics and statistical tests.
This table presents trading summary statistics for XSE and Cboe and the statistical tests of the trading summary difference between XSE and Cboe. The statistical tests conducted are two-sample *t*-tests and pairwise Wilcoxon-Mann-Whitney tests. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe. The sample period covers March 2017 to August 2018. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

	Average trading volume per stock (€'000,000)	Average trading volume per stock (000,000s)	Average transactions per stock (000s)	Average trade size per stock (€'000)
XSE	16,263.46	428.56	984.02	14.94
Cboe	2739.96	64.09	356.29	6.87
XSE–Cboe	13,523.5***	364.47***	627.73***	8.07***
t-test *p*-value	<0.001	<0.001	<0.001	<0.001
W-M-W test *p*-value	<0.001	<0.001	<0.001	<0.001

estimated transmission latencies. Our estimates in Panel A's suggest that traders are more likely to employ microwave technology than fiber optic options to connect Frankfurt and London. Panel B in Table 2 presents the estimated mean, median, and standard deviation of the latencies; the average latency is 4.4 ms.

On July 3, 2017 and April 9, 2018, XSE implemented upgrades to increase the exchange speed. These technological upgrades include: (1) the "New T7 Trading Technology" upgrade first offered on July 3, 2017; and (2) the "Introduction of PS Gateways" upgrade first offered on April 9, 2018.[10] The Deutsche Börse T7 Trading Technology system significantly reduces the order processing time and should be captured by our *TL* measure as discussed above. The PS (partition-specific) gateways upgrade for all cash market instruments operates in parallel with the high-frequency gateways. Usually, latency jitters on parallel inbound paths encourage multiplicity to reduce latency. However, this leads to greater system loads and choking at busy times, thus less predictable latencies may arise. The PS gateways upgrade introduces a single low-latency point of entry, which addresses this issue and consequently reduces exchange latency at XSE. The upgrade therefore addresses the problem of less predictable latencies, which could discourage timely responses to XSE signals on the part of Cboe traders in our setting. If latency is stabilized, Cboe traders will likely pick up new information more reliably and act on it in a timely fashion. Hence, this reduction should also be captured by *TL*. This suggests that these two upgrades can be used to test the empirical relevance of *TL*. In particular, if *TL* indeed captures information transmission latency between London and Frankfurt, then the timing of the upgrades should align with a statistically significant reduction or downward adjustment in *TL*.

Fig. 2 illustrates the impact of both upgrades on *TL*. Panel A shows a sharp decrease in *TL* on the day of the New T7 Trading

[10] The details of the upgrades are at https://www.xetra.com/dbcm-en/newsroom/press-releases/New-T7-trading-technology-goes-live-on-Xetra-144756aand https://www.xetra.com/resource/blob/228942/0bbe6323aa5436a88648d298d9b41512/data/143_17e.pdf.

Table 2
Information transmission latency between XSE and Cboe.
This table presents different statistics for the information transmission latency between XSE and Cboe. Panel A reports the number of responses on Cboe to price-changing trades on XSE for different time bins in milliseconds (ms). Panel B presents the mean, median, and standard deviation of the information transmission latency between XSE and Cboe. Panel C shows the average information transmission latencies for the 21 trading days before and 21 trading days after the technological upgrades on July 3, 2017 (T7) and April 9, 2018 (PS Gateway). The statistical tests conducted are two-sample t-tests and pairwise Wilcoxon-Mann-Whitney tests. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe. The sample period covers March 2017 to August 2018. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

Panel A: The distribution of the latencies of responses on Cboe to price-changing trades on XSE		
Speed (ms)	Frequency	Percentage
3	936,646	48.61
4	286,962	14.89
5	332,286	17.24
6	100,435	5.21
7	81,733	4.24
8	75,895	3.94
9	62,679	3.25
10	50,364	2.61
Panel B: Summary statistics of the information transmission latency		
Mean	Median	Standard deviation
4.399	4.371	0.498
Panel C: Information transmission latencies before and after the upgrades		
Period	Average latency for the full sample T7 upgrade	Average latency for the full sample PS Gateway upgrade
Before upgrade	4.40	4.27
After upgrade	4.30	4.20
Difference	0.10***	0.07***
t-test *p*-value	<0.001	<0.001
W-M-W test *p*-value	<0.001	<0.001

Technology upgrade, with the average latency falling by 0.105 ms–4.297 ms, a reduction of 2.4%. Panel B shows that the Introduction of PS Gateways leads to a significant latency reduction of 1.6%. Note that a reduction of 1.5%–3% is substantial given that latencies are already close to their limits due to the lower speed-of-light boundary. We also test the statistical significance of the difference between the latencies 21 trading days before and 21 trading days after the implementation of the upgrades. The estimates in Panel C in Table 2 show that the average latency reductions are statistically significant for both upgrades, and thus confirm the empirical validity of the *TL* metric.

3.3. Toxic and non-toxic latency arbitrage opportunities

TL, by design, inversely captures the latency-arbitrage-enabling speeds of HFTs actively engaged in cross-market trading (XSE-Cboe setting in our case). Hence, the speed of trading is informed by the inclination of HFTs to exploit latency arbitrage opportunities. However, according to Foucault et al. (2017), in a cross-market setting the effects of trading speed, as encapsulated by HFT activity, can have one of two contrasting impacts on market quality characteristics depending on the type of latency arbitrage opportunity being exploited. First, HFTs may increase liquidity in response to the emergence of (non-toxic) arbitrage opportunities due to price pressures; and second, HFTs may impair liquidity following the onset of (toxic) arbitrage opportunities due to lagged reactions to news. Hence, the effect of speed on market quality is a function of whether it evolves in response to toxic or non-toxic latency arbitrage opportunities. This implies that to disentangle and credibly investigate the effects of cross-market speed on market quality characteristics, we must begin by identifying the toxic and non-toxic latency arbitrage opportunities driving the pace of HFT activity in the cross-market setting.

We adapt the approach of Foucault et al. (2017) following the rationale that toxic (non-toxic) transactions arise due to information (transitory price pressures) and therefore generate permanent (temporary) price impacts. Thus, we first identify near-coincident same-direction price-changing trades on Cboe as in our latency measurement method (see Subsection 3.2). Thereafter, for each near-coincident transaction s at event time t, we compare the midpoint of the prevailing ask and bid quotes at time t (M_t) with the midpoint of the prevailing ask and bid quotes at time t+4 (M_{t+4}).[11] A price reversal of a minimum 50%[12] (i.e., M_{t+4} reversing toward

[11] For robustness, we compare the prevailing midpoint corresponding to the near-coincident trade with those corresponding to the trades at times t+6 and t+11. Naturally, the number of toxic and non-toxic arbitrage opportunities changes. However, our secondary regression results (in Section 4) are identical.

[12] For robustness, we identify toxic and non-toxic transactions using various reversal levels (from 10% to 100%) and employ them in the secondary regression models described in Section 4. The results are quantitatively and qualitatively similar.

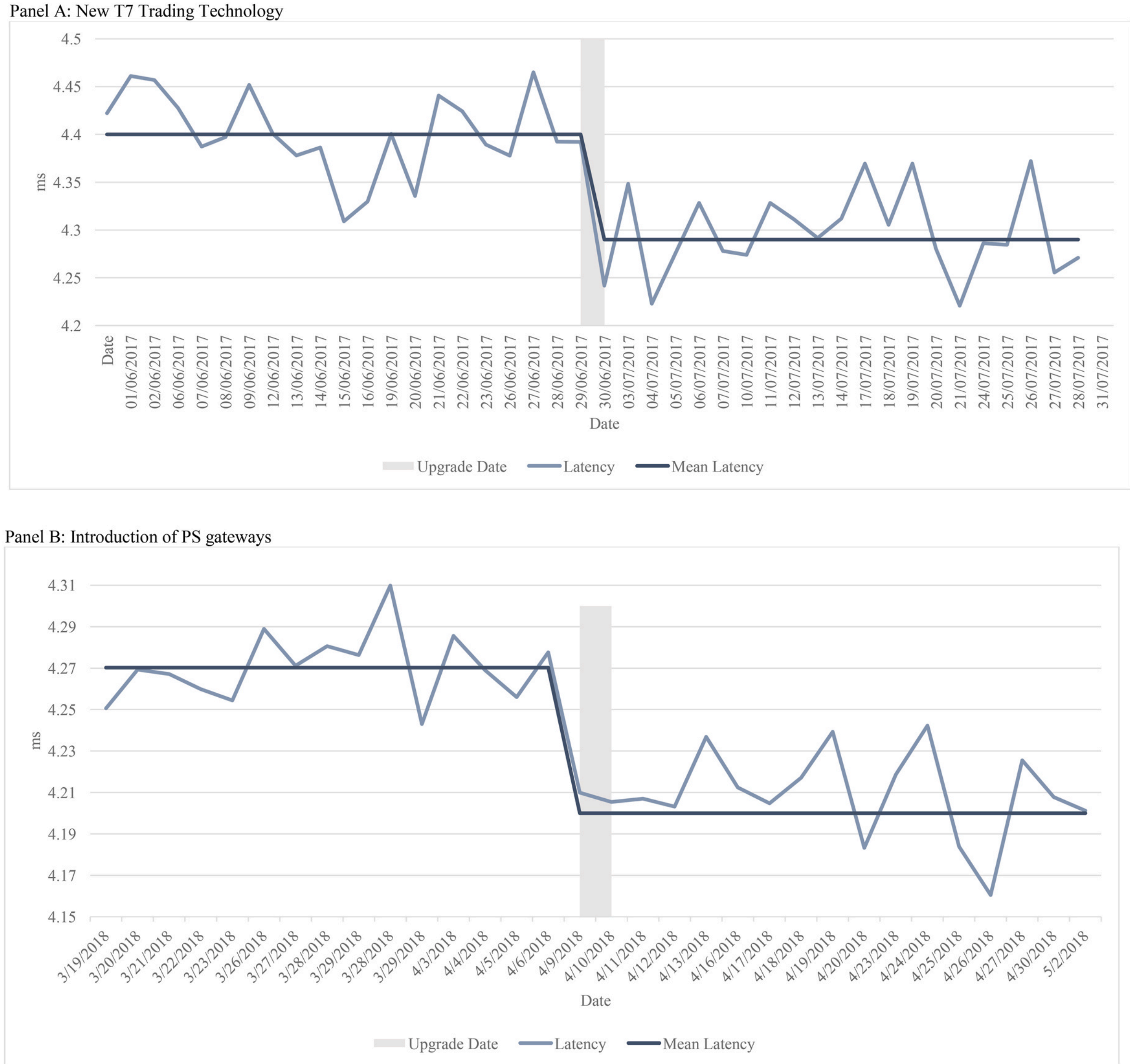

Fig. 2. Information transmission latency over time
Panel A: New T7 Trading Technology Panel B: Introduction of PS gateways.
This figure shows transmission latency (*TL*) around two technological upgrades. Panel A shows *TL* from June 1, 2017 to July 31, 2017, and Panel B shows *TL* from March 19, 2018 to May 2, 2018. The vertical bars indicate the technological upgrades, "New T7 Trading Technology," which took effect on July 3, 2017 (Panel A), and "Introduction of PS gateways," which took effect on April 9, 2018. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe.

M_t) implies that the near-coincident transaction is non-toxic. The remaining transactions are labeled as toxic. To clarify this mechanism, it is useful to consider the following illustration. Consider a price-changing transaction occurrence on XSE at time *t-k* and the same direction near-coincident transaction occurring on Cboe at time *t*. Suppose the prevailing midpoint on XSE and Cboe prior to these two transactions is €100 ($M_{t-k}^{XSE} = M_t^{Cboe} =$ €100), and a trader submits a market buy order on XSE at time *t-k*, which increases the prevailing midpoint on XSE to €100.1. This is followed by a near-coincident Cboe trade in the same direction, which increases the midpoint there to €100.1. We then observe the next four new transactions (i.e., executed at event times t+1, t+2, t+3, and t+4) executed on Cboe but only focus on the midpoint just prior to the fourth transaction (M_{t+4}^{Cboe}). For simplicity, we assume that no new information event (i.e., price-changing trade due to new information) occurs on XSE until the execution of a transaction at time t+4 on Cboe. To classify a transaction as toxic/non-toxic, we compare the midpoints prevailing at time t+4 and t (M_{t+4}^{Cboe} and M_t^{Cboe}). Thus, if M_{t+4}^{Cboe} is less than or equal to €100.05, implying a price reversal of at least 50%, the transaction is classified as non-toxic. If M_{t+4}^{Cboe} is greater than €100.05, the transaction is classified as toxic.

A caveat of our algorithm is that, as is the case with all (price impact) measures based on price comparison across time intervals in the market microstructure literature, is that we implicitly assume that no new information event occurs on Cboe between event times t and $t+4$. This assumption is highly plausible given that the average gap between event times t and $t+4$ in our sample is less than 1 s. For instance, O'Hara (2015) suggests that a 5- to 15-s window is a plausible horizon for computing realized spreads in a high-frequency environment. More importantly, our algorithm does not classify a latency arbitrage opportunity as toxic or non-toxic until the transaction execution at time $t+4$. If a new information event (i.e., price-changing trade due to new information) occurs on XSE before we observe the fourth transaction on Cboe, the algorithm will also not classify the arbitrage opportunity.

Nevertheless, to understand the potential impact of this narrow concern on our classification algorithm, we conduct a simple test. If intervening trades on Cboe between event times t and $t+4$ are belief-impacting (i.e., contain information) and cause a toxic arbitrage opportunity to be classified as non-toxic, we should expect these intervening trades to be contrarian (i.e., be in the opposite direction from the near-coincident price-changing trade that initiates the toxic latency arbitrage on Cboe at time t). We find that in 89% of the arbitrage opportunities that are classified as toxic, transactions between times t and $t+4$ consistently only occur in the same direction as the near-coincident price-changing transaction at time t. In only 3% of the toxic arbitrage opportunities, intervening trades are in the opposite direction from the near-coincident price-changing transaction at time t. There are mixed directions in the remaining toxic arbitrage opportunities (8%). The fact that the intervening trades between t and $t+4$ following 89% of toxic arbitrage opportunities have the same direction as the near-coincident price-changing trade at time t suggests that these intervening trades do not generally contain new information that causes price reversals. This is expected given that intervening trades are observed at sub-second intervals. Unsurprisingly, on the other hand, in 70% of the non-toxic arbitrage opportunities, intervening trades are in an opposite direction to the near-coincident price-changing trade at time t. Moreover, in 25% of the non-toxic arbitrage opportunities, intervening trade directions are mixed. In almost all of these mixed direction cases, the first and second intervening trades ($t+1$ and $t+2$) are in the same direction as the near-coincident trade at time t, while the intervening trade at $t+3$ is typically in the opposite direction. This suggests that, for non-toxic arbitrage opportunities, it takes a few trades for HFTs to conclude that an arbitrage opportunity is toxic or non-toxic. Only in 5% of the non-toxic latency arbitrage opportunities do we observe all intervening trades having the same trade direction as the near-coincident trade at time t.

The summary statistics for toxic and non-toxic transactions are presented in Table 3. The total number of latency arbitrage opportunities is 1,927,000 and only 35.15% (677,328) of them are toxic. This implies that the occurrence of non-toxic latency arbitrage opportunities is higher than toxic latency arbitrage opportunities. This is underscored by the distribution of the two types of opportunities along the time series. We find that the bulk (71%) of the non-toxic arbitrage opportunities in our sample occur following other non-toxic arbitrages; the remaining 21% of non-toxic latency arbitrage episodes are followed by toxic arbitrage opportunities. The distribution is more balanced in the cases of toxic latency arbitrage opportunities. Toxic latency arbitrage opportunities are only followed by other toxic latency arbitrage opportunities 46% of the time. Thus, non-toxic arbitrage opportunities are more likely to occur after non-toxic arbitrages, while we (almost) equally observe toxic and non-toxic arbitrage opportunities following the toxic arbitrage opportunities.

To test whether our classification method is empirically valid, we compare the number of toxic and non-toxic arbitrage opportunities on news and no-news days. Given that toxic latency arbitrage opportunities are driven by the arrival of information, we expect a higher incidence of such opportunities on days with firm-related news. We use two steps to identify the days with firm news (e.g., Hirschey, 2020). In the first step, we use Factiva, which contains news from over 35,000 sources, to identify news days. Then, to eliminate the possibility that traders respond to news not covered by Factiva, in the second step, we exclude days with absolute

Table 3

Number of toxic and non-toxic arbitrage opportunities.

This table reports the number of toxic and non-toxic arbitrage opportunities for the news and no-news days. To split the transactions into toxic and non-toxic groups, the approach of Foucault et al. (2017) is adapted, following the rationale that toxic (non-toxic) transactions arise due to information (transitory price pressures) and therefore generate permanent (temporary) price impacts. Firstly, near-coincident same-direction price-changing trades on Cboe are identified. Thereafter, for each near-coincident transaction s at event time t, the midpoint of the prevailing quote at time t (M_t) is compared with the midpoint of the prevailing quote at time $t+4$ (M_{t+4}). A price reversal, i.e., M_{t+4} reversing toward M_t, implies that the near-coincident transaction is non-toxic. The remaining transactions are labeled as toxic. Days with firm news are identified in two steps (see Hirschey 2020). In the first step, we use Factiva, which contains news from over 35,000 sources, to identify news days. Then, to eliminate the possibility that traders respond to news not covered by Factiva, in the second step, we exclude days with absolute market-adjusted returns greater than 1%. The statistical tests conducted are two-sample t-tests and pairwise Wilcoxon-Mann-Whitney tests. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe. The sample period covers March 2017 to August 2018. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

	Toxic arbitrage opportunities			Non-toxic arbitrage opportunities		
	Total number of arbitrage opportunities	Average number of arbitrage opportunities per stock	Fraction of arbitrage opportunities	Total number of arbitrage opportunities	Average number of arbitrage opportunities per stock	Fraction of arbitrage opportunities
News days	584,193	1604.93	86.25%	642,758	1765.82	51.43%
No-news days	93,135	255.87	13.75%	606,914	1667.35	48.57%
All days	677,328	1860.79		1,249,672	3433.16	
News–no-news days	491,058	1349.06***	72.50%	35,844	98.47	2.87%
t-test p-value		<0.001			0.19	
W-M-W test p-value		<0.001			0.20	

market-adjusted returns greater than 1%. As is evident in Table 3, the incidence of toxic latency arbitrage opportunities on news days is about 627% that of no-news days (584,193 toxic opportunities on news days to 93,135 toxic opportunities on no-news days). Overall, 86.25% of all toxic latency arbitrage opportunities are observed on news days. The difference in the incidence of toxic latency arbitrage opportunities between news and no-news days is also statistically significant, based on a null hypothesis of no difference in the occurrence of toxic latency arbitrage opportunities between news and no-news days. We do not detect the same order of difference in the incidence of non-toxic latency arbitrage opportunities between news and no-news days. The results of these two tests align with our expectations and strongly suggest that our classification algorithm is empirically relevant. Since toxic latency arbitrage opportunities imply the exploitation of information, observing toxic latency arbitrage opportunities more frequently on days when firm-related news is released underscores the relevance of the classification algorithm.

3.4. Market quality measures: liquidity and price discovery proxies

We investigate and explain the interactions between trading speed and market quality by focusing on two market quality characteristics: liquidity and price discovery. We start with liquidity, where we use approximations of standard liquidity — the relative quoted and effective spreads. As discussed in Foucault et al. (2017) and in Subsection 3.3, market makers dynamically adjust their spread depending on their perception of whether latency arbitrage opportunities arise due to information or not. The implication of this for our study is that we need to compute spreads following toxic and non-toxic latency arbitrage opportunities separately to directly test the impact of the arbitrage strategies (liquidity-providing and liquidity-taking strategies) employed by HFTs. This is because a market maker does not know the source of a latency arbitrage opportunity before it ends; analogous to our algorithm for classifying latency arbitrage opportunities, they can only identify the nature of a latency arbitrage opportunity after it terminates. Thus, we compute two types of spreads for each day: toxic latency arbitrage and non-toxic latency arbitrage spreads.

We compute a toxic (non-toxic) latency arbitrage spread using the next 50 transactions following the conclusion of each toxic (non-toxic) latency arbitrage opportunity.[13] When a new latency arbitrage opportunity occurs before we observe the next 50 transactions that follow a latency arbitrage opportunity, the maximum number of transactions occurring prior to the new latency arbitrage opportunity is employed instead. For instance, when we observe a new cross-market latency arbitrage opportunity occurring 20 transactions after the previous one terminates, we instead estimate our spread measures using only the 19 transactions occurring prior to the latest latency arbitrage opportunity. On average, we use 13 transactions per latency arbitrage opportunity to compute our liquidity measures. Thus, the total number of transactions used to compute liquidity measures is 25,051,296, about 70.37% of the total transactions on Cboe. The selected maximum transactions threshold of 50 following the termination of a latency arbitrage opportunity is crucial to isolating the effect of high-frequency arbitrageurs on liquidity, given the many spread-impacting intraday events occurring in financial markets. The average time it takes to observe the 50 transactions is 4 s. By using only 50 transactions, we focus on the change in spread due to HFTs' cross-market latency arbitrage strategies.

We compute four spread measures based on the relative quoted and effective spreads. $Espread_{i,t}^{T}$ ($Espread_{i,t}^{NT}$) is the toxic (non-toxic) latency arbitrage effective spread computed as the daily volume-weighted average of twice the absolute value of the difference between the transaction prices that we observe following toxic (non-toxic) latency arbitrages and the midpoint of the bid-ask spread prevailing before these transactions, divided by the midpoint. $Qspread_{i,t}^{T}$ ($Qspread_{i,t}^{NT}$) is computed as the daily volume-weighted average of the difference between the ask and bid prices prevailing in transactions that we observe following toxic (non-toxic) latency arbitrages divided by the midpoint of these ask and bid prices.

The second market quality characteristic we examine is price discovery. Price drivers can be classified into two main components: an efficient/permanent component linked to the incorporation of information into price through, for example, informed traders trading with private information (e.g., Glosten and Milgrom, 1985) or information-driven quote updates, such as with no trading (e.g., Biais et al., 1999); and a noise/transitory component, linked to the activities of uninformed and noise traders, including microstructure impacts and reversible price effects (e.g., Hasbrouck, 1991). Information reflected in price may therefore be characterized as public [e.g., reflected in market maker quote updates as in Brogaard et al. (2019)] or private [e.g., through a privately informed trader as in Kyle (1985)]. Hence, we conduct a variance decomposition using a microstructure vector autoregression (VAR) model in the spirit of Hasbrouck (1991, 1995) and Barclay and Hendershott (2003) to disentangle the information-driven variance of stock returns into two components: private information and public information (see also Brogaard et al., 2022).

For parsimony, we provide the computational details of this decomposition method in Appendix B. We first model price as the sum of permanent (information) and transitory (noise) components. The information component is then further decomposed into two parts: (i) a private information component that is captured by signed natural logarithm of currency volume in the model, thus is reflected in trades; and (ii) a second component that is not attributable to the signed trading volume and captured using returns computed from prevailing midpoints corresponding to trades included in our sample; we call this "public information." We do so because previous studies (e.g., Biais et al., 1999; Brogaard et al., 2019) suggest that quote updates (and thus midpoint returns) not linked to individual trades reflect public information. However, as we are unable to observe quote updates without a trade in the sample, our ability to draw this conclusion is limited. Indeed, if a trade elicits a price change larger than those of trades with similar signed log volumes in the data, some of that change will be assigned to a different economic source, which is subsumed in the component we call public

[13] For robustness, we use various event windows after each opportunity, ranging from 10 to 200 transactions. The regression results (Section 4) are robust to alternative windows.

information. The crucial question then is: Why would the price impact of a trade be greater than the average for its amount of signed volume? We suggest this may either arise due to the price effects of quote changes (as in Biais et al., 1999; Brogaard et al., 2019) in between trades or as a result of the said trade reflecting a public information event that a fast trader exploits with the trade.[14] Both suggestions indicate that the second (non-private) information component estimated using the VAR essentially reflects what could plausibly be regarded as public information.

Consistent with our approach to estimating the spread measures, we compute the two components using toxic and non-toxic latency arbitrage transactions separately. Thus, we use four different price discovery measures divided into two groups: (1) price discovery due to private information, computed using transactions following the toxic and non-toxic latency arbitrage opportunities ($Prshare^{T}_{i,t}$ and $Prshare^{NT}_{i,t}$, respectively); and (2) price discovery driven by public information, computed using transactions following the toxic and non-toxic latency arbitrage opportunities ($Pushare^{T}_{i,t}$ and $Pushare^{NT}_{i,t}$, respectively).

Table 4 provides the definitions of the variables, along with their means, medians, and standard deviations. $Espread^{T}_{i,t}$ and $Qspread^{T}_{i,t}$ are 54.14% and 38.54% greater than their respective non-toxic latency arbitrage values ($Espread^{NT}_{i,t}$ and $Qspread^{NT}_{i,t}$). While this is not reported in the table, we also note that the differences between the toxic and non-toxic spread metrics are statistically significant at the 0.01 level. Given that toxic latency arbitrage opportunities are driven by information asymmetry, this pattern is consistent with the argument that market makers widen the spread in an informed trading environment (e.g., Copeland and Galai, 1983).

For toxic latency arbitrage transactions, 45.53% (54.47%) of the information-linked return variance is due to private information (public information). The respective percentages are 48.79% and 51.21% for the non-toxic latency arbitrage transactions. Thus, although there is no evidence of any significant difference in the private and public information components of transactions observed following the non-toxic latency arbitrage trades, the differences between the private and public information components of transactions observed following toxic latency arbitrage opportunities are economically and statistically significant.

At first glance, this finding seems to contradict the main reasoning behind toxic and non-toxic arbitrage opportunities classification. Given that toxic arbitrage opportunities are driven by information revealed through trading, we may expect that the private information component of transactions observed following toxic arbitrage opportunities would be higher than the public information component. However, in line with our focus on investigating the impact of observed cross-market latency arbitrage activity on market quality characteristics, our decomposition strategy employs transactions observed in the immediate aftermath of such latency arbitrage activities. This approach allows us to isolate the immediate effects of such activities on the market quality characteristics. Therefore, our explanation for the obtained estimates should focus on the drivers of the evolution of the price components following the termination of latency arbitrage opportunities, not the drivers linked to the evolution of the components during the latency arbitrage activities.

Consequently, the larger public information content of transactions following toxic arbitrage trades may be due to the mechanics of the market maker's reaction to latency arbitrage activity. When trading reveals private information on XSE, HFTs' cross-market arbitrage algorithms can bring this information onto Cboe by trading on them and thus generate toxic latency arbitrage. On the other hand, HFTs' market making algorithms update their quotes to reflect the occurrence of the recently-completed latency arbitrage opportunity. Our price decomposition approach captures the price impact generated by this quote update because we use the transactions following the termination of latency arbitrage opportunities to estimate the private and public components of price discovery. Given that market makers' quotes reveal public information (i.e., the midpoint can change without trading due to quote updates), the high levels of public information observed in price following the toxic arbitrage opportunities may suggest that quotes submitted by HFTs' market-making algorithms convert private information linked to HFTs' arbitrage strategies into public information. This is consistent with Boehmer et al. (2018) showing that competition among HFTs converts cross-stock private information into public information. Boehmer et al. (2018) also note that HFTs engaged in arbitrage activity may pursue both liquidity-providing (market--making) and liquidity-consuming (opportunistic) strategies. This suggests that the implied information-conversion activity and the subsequent market maker quote updates discussed above may be the work of the same HFTs. We are unable to ascertain whether this is the case, because testing it requires a dataset with trader identification, which is unavailable to us.

4. Empirical findings and discussion

4.1. Latency and liquidity

One of the main challenges in analyzing the impact of *TL* on market quality is endogeneity. In particular, an unobserved variable correlated with *TL* might be driving market quality, or there may be some reverse causality between the market quality characteristics we employ and *TL*. To address this endogenous determination, in addition to estimating fixed effects models, we employ a 2SLS IV approach. We estimate the following regression models:

$$TL_{i,t} = \alpha_i + \beta_t + \gamma Instrument_{i,t} + \sum_{k=1}^{7} \delta_k C_{k,i,t} + \varepsilon_{i,t} \tag{1}$$

[14] We are grateful to the editor for the ideas we use in enumerating these issues.

Table 4
Definition of variables and summary statistics.
This table provides the definition of the variables calculated for each stock *i* and day *t*, and reports the summary statistics for the main variables. To split transactions into toxic and non-toxic groups, we use the modified version of the classification algorithm described in Foucault et al. (2017). Further details of the classification algorithm are provided in Section 3.3 and Table 3. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe and the summary statistics are computed for Cboe. The sample period covers March 2017 to August 2018.

Variable	Definition	Mean	Median	SD
$Espread^{T}_{i,t}$ (bps)	Toxic latency arbitrage effective spread for stock *i* and day *t*, computed as the daily volume-weighted average of twice the absolute value of the difference between the transaction prices that are observed following toxic latency arbitrages and the midpoint of the bid-ask spread prevailing before these transactions, divided by the midpoint.	9.85	9.93	12.66
$Espread^{NT}_{i,t}$ (bps)	Non-toxic latency arbitrage effective spread for stock *i* and day *t*, computed as the daily volume-weighted average of twice the absolute value of the difference between the transaction prices that are observed following non-toxic latency arbitrages and the midpoint of the bid-ask spread prevailing before these transactions, divided by the midpoint.	6.39	6.44	8.69
$Qspread^{T}_{i,t}$ (bps)	Toxic latency arbitrage quoted spread for stock *i* and day *t*, computed as the daily volume-weighted average of the difference between the prevailing ask and bid prices corresponding to transactions that are observed following toxic latency arbitrages divided by the midpoint of these ask and bid prices.	10.82	9.75	9.83
$Qspread^{NT}_{i,t}$ (bps)	Non-toxic latency arbitrage quoted spread for stock *i* and day *t*, computed as the daily volume-weighted average of the difference between the prevailing ask and bid prices corresponding to transactions that are observed following non-toxic latency arbitrages divided by the midpoint of these ask and bid prices.	7.81	8.79	6.25
$Prshare^{T}_{i,t}$ (%)	The share of private information in the variance of returns for transactions that are observed following toxic latency arbitrages in stock *i* and day *t* as described in Appendix B.	45.53	47.61	23.22
$Prshare^{NT}_{i,t}$ (%)	The share of private information in the variance of returns for transactions that are observed following non-toxic latency arbitrages in stock *i* and day *t* as described in Appendix B.	48.79	45.62	20.39
$Pushare^{T}_{i,t}$ (%)	The share of public information in the variance of returns for transactions that are observed following toxic latency arbitrages in stock *i* and day *t* as described in Appendix B.	54.47	52.36	25.41
$Pushare^{NT}_{i,t}$ (%)	The share of public information in the variance of returns for transactions that are observed following non-toxic latency arbitrages in stock *i* and day *t* as described in Appendix B.	51.21	49.68	38.78
$lnTV_{i,t}$	Natural logarithm of trading volume for stock *i* and day *t*.	12.78	12.63	5.76
$lnDepth_{i,t}$	Natural logarithm of the sum of the ask and bid sizes for stock *i* and day *t*.	14.49	14.31	6.87
$InvPri_{i,t}$ (bps)	Inverse of last transaction price for stock *i* and day *t*.	302.16	261.27	340.52
$Momentum_{i,t}$ (bps)	The first lag of the daily return for stock *i* and day *t*.	0.611	0.547	25.712

$$DV_{i,t} = \alpha_i + \beta_t + \gamma \widehat{TL_{i,t}} + \sum_{k=1}^{7} \delta_k C_{k,i,t} + \varepsilon_{i,t}. \tag{2}$$

Equations (1) and (2) are the first- and second-stage regressions respectively. First, we compute the *TL* as the response times between pairs of linked trades, and then estimate $TL_{i,t}$ as the average *TL* for stock *i* and day *t*. $Instrument_{i,t}$ is the value of a given instrument for stock *i* and day *t*. For robustness, we use three instruments.

The first instrument is based on the XSE upgrades discussed in Subsection 3.2, which are used as shocks to *TL*. Therefore, $Instrument_{i,t}$ equals zero from March 1, 2017 to July 3, 2017, one from July 3, 2017 to April 9, 2018, and two from April 9, 2018 to August 31, 2018. As described in Hasbrouck and Saar (2013), HFT strategies are characterized by "strategic runs" (see also Menkveld, 2013), meaning that they rely in part on regularly observing or querying prices; thus, a fast trader on Cboe will need to continuously query prices on XSE in order to react to them in a timely fashion. An upgrade that impacts order processing at XSE will therefore reduce the response time for the Cboe trader observing price changes on XSE, and this is captured by *TL*. The execution latency of a Cboe trader not only directly depends on the speed with which she can act on information from linked exchanges, but also on the promptness with which she can receive information from these exchanges.

Using technological upgrade events at XSE as an instrument also allows us to investigate the upgrade spillover phenomenon described by Brogaard et al. (2015). Brogaard et al. (2015) show that the colocation upgrade at NASDAQ OMX improves liquidity in all linked markets through cross-listed stocks. However, the reason for this liquidity improvement across all exchanges has not been clearly analyzed in previous studies. In particular, our instrument allows us to test whether upgrades reducing latency on one exchange also reduce inter-exchange latency, i.e., *TL*, and thereby lessen the trading costs of competitive liquidity providers who operate in both markets. The likelihood of HFTs being engaged in cross-market strategies is established in the recent market microstructure studies and supported by evidence from the industry. For example, Menkveld (2013) shows that HFTs employ cross-market strategies in which they split their trades almost evenly across Euronext and Cboe. Boehmer et al. (2018) also show that HFTs extensively adopt cross-market arbitrate strategies.

The second instrument is the number of microwave networks operated by HFTs during our sample period. At the start of our sample period, 94 microwave networks were in use by HFTs. This number increased to 148 by the end of our sample period, suggesting that 54 new microwave networks were launched by HFTs during the period of interest. Thus, for the second instrument, the variable $Instrument_{i,t}$ captures the launch of new microwave networks between Frankfurt and London. Our third instrument is designed in line with Hasbrouck and Saar (2013) and Degryse et al. (2015); we first split our sample into three groups based on their market capitalization. Then $Instrument_{i,t}$ equals the average *TL* across all stocks on day *t* in the same size group (calculated excluding stock *i*).

The dependent variable, $DV_{i,t}$, corresponds to one of the effective ($Espread^{T}_{i,t}$ and $Espread^{NT}_{i,t}$) or quoted ($Qspread^{T}_{i,t}$ and $Qspread^{NT}_{i,t}$) spreads for stock i and day t. $C_{k,i,t}$ is a set of k control variables, which includes the standard deviation of stock returns ($Stddev_{i,t}$), the inverse of price ($InvPri_{i,t}$), the natural logarithm of trading volume ($lnTV_{i,t}$), the natural logarithm of market depth ($ln\ Depth_{i,t}$), momentum ($Momentum_{i,t}$), $News_{i,t}$, and $Control_{i,t}$. $Stddev_{i,t}$ is computed as the standard deviation of the intraday hourly returns for stock i and day t; $InvPri_{i,t}$ is the inverse of the last transaction price for stock i and day t; $lnTV_{i,d}$ is the natural logarithm of the trading volume for stock i and day t; $ln\ Depth_{i,t}$ is the natural logarithm of the sum of the ask and bid volumes for stock i and day t; $Momentum_{i,t}$ is the first lag of daily returns ($Momentum_{i,t}$ is the return of stock i on day t-1); and $News_{i,t}$ is a dummy equaling 1 if there is firm-related news for stock i and day t.

$Control_{i,t}$ is the value of the corresponding market quality metric (the second-stage dependent variable) for the matched control stock. Our control group comprises 100 stocks that are listed on Cboe but not on XSE; they should therefore not be impacted by TL.[15] This variable allows us to remove variation that is common to Cboe stocks and not driven by TL. The main reason for using this approach over the standard difference-in-differences methodology is that we are unable to impose a one-to-one correspondence between changes in treatment stocks' dependent variable and changes in control stocks' dependent variable. This is because the dependent variable, $DV_{i,t}$, is separately computed for cross-market toxic and non-toxic latency arbitrage transactions. Given that our control stocks are only listed on Cboe, we are unable to compute toxic and non-toxic market quality variables for these stocks separately and, hence, we are unable to use the standard difference-in-differences approach in our empirical setting. As discussed in Foley and Putniņš (2016), by using the dependent variable of the control group as an independent variable, this approach avoids inflating the number of observations and provides more conservative standard errors.

The stock and fixed effects are α_i and β_t, respectively. However, for the 2SLS IV frameworks, time fixed effects are only included when $Instrument_{i,t}$ is the average level of the TL of the stocks in the same size group on day t. The reason for this is that our first (XSE upgrades) and second (number of microwaves) instruments only have time variation. Standard errors are double-clustered by stock and time (day).

Table 5 reports the results from the panel fixed effects ($TL_{i,t}$ is used as a key dependent variable) and the second stage of the 2SLS IV ($\widehat{TL_{i,t}}$ is used as a key dependent variable) estimations. For parsimony, the results of the first stage of the 2SLS estimations are not presented. However, there are two important points to note here. First, the relations between $TL_{i,t}$, on the one hand, and the three types of $Instrument_{i,t}$, on the other, are significant, and the signs of the relations are in line with expectations. The XSE upgrades and the number of microwave networks are both negatively correlated with $TL_{i,t}$, suggesting that the technological improvements decrease TL, which implies an increase in speed. The average TL in the same size group is also positively correlated with $TL_{i,t}$. Second, the F-statistics in the first-stage regressions are above the critical values specified in Stock and Yogo (2005); the minimum value across all the regressions is 39. These results indicate that the first-stage regression results support our instrument selection.

Panels A and B of Table 5 provide the results for equation (2) when $DV_{i,t}$ corresponds to $Espread^{T}_{i,t}$ ($Espread^{NT}_{i,t}$) and $Qspread^{T}_{i,t}$ ($Qspread^{NT}_{i,t}$), respectively. Three points stand out. First, the estimates across the various models are comparable and consistent. This suggests that our results are robust, allowing us to establish a causal relationship between trading speed and liquidity while effectively addressing endogeneity concerns. Second, there is a negative relationship between $TL_{i,t}$ on the one hand and $Espread^{T}_{i,t}$ and $Qspread^{T}_{i,t}$ on the other. This result suggests that the increases in TL are associated with improvements in liquidity following incidences of toxic latency arbitrage opportunities. By contrast, $TL_{i,t}$ is positively associated with $Espread^{NT}_{i,t}$ and $Qspread^{NT}_{i,t}$, implying that trading speed increases liquidity following non-toxic latency arbitrage opportunities. Both results are consistent with the predictions of Foucault et al. (2017) and show that the impact of HFTs on liquidity is not homogeneous and, in a cross-market setting, it is linked to the sources of the latency arbitrage opportunities HFTs seek to exploit.

Third, the magnitudes of the associations are also economically meaningful. For instance, on average, a 1 ms decrease in latency is expected to increase $Espread^{T}_{i,t}$ ($Qspread^{T}_{i,t}$) by about 0.41/9.85 = 4.16% (5.73%) and reduce $Espread^{NT}_{i,t}$ ($Qspread^{NT}_{i,t}$) by about 1.18/6.39 = 18.46% (17.03%). To contextualize the economic magnitude of these effects, we can look to the average latency changes resulting from the technological upgrades we exploit in our empirical framework. The adoption of the New T7 Trading Technology upgrade led to an average latency reduction of approximately 0.105 ms. This implies that implementing this rather (logistically and financially) significant upgrade results in a 0.42% (0.57%) increase in $Espread^{T}_{i,t}$ ($Qspread^{T}_{i,t}$) and a corresponding decrease of 1.85% (1.70%) in $Espread^{NT}_{i,t}$ ($Qspread^{NT}_{i,t}$).

Notwithstanding, the above estimates suggest that the economic impact of transmission latency on liquidity is about three to four times higher following incidences of non-toxic latency arbitrage opportunities when compared to its impact following incidences on toxic latency arbitrage opportunities. On balance, trading speed in financial markets likely enhances liquidity, and latency impairs it. To formally test this, we construct new spread measures ($Espread_{i,t}$ and $Qspread_{i,t}$) based on all transactions, without distinguishing between the influence of toxic and non-toxic latency arbitrage opportunities. We estimate equation (2) using the new spread measures as dependent variables. The results are reported in Panel C of Table 5. Consistent with our expectations, $TL_{i,t}$ is positively associated

[15] We employ the approach developed by Boulton and Braga-Alves (2010) to create the control group. We match each of the stocks to a corresponding control stock. The matching variables are market value, closing stock price, and trading activity. We select the control stocks by minimizing $\mathrm{Distance_f}$, where $\mathrm{Distance_f} = \sum_{\mathrm{f}} \left| \left(\mathrm{matchingvariable_f^{stock}} - \mathrm{matchingvariable_f^{control}}\right) / \left[\left(\mathrm{matchingvariable_f^{stock}} + \mathrm{matchingvariable_f^{control}}\right)/2\right] \right|$.

Table 5
Latency and liquidity.
This table reports the coefficient estimates from the following regression model:

$DV_{i,t} = \alpha_i + \beta_t + \gamma\widehat{TL_{i,t}} + \sum_{k=1}^{7}\delta_k C_{k,i,t} + \varepsilon_{i,t}$

where $DV_{i,t}$ corresponds to either the effective spread ($Espread^{T}_{i,t}$ and $Espread^{NT}_{i,t}$) (Panel A) or quoted spread ($Qspread^{T}_{i,t}$ and $Qspread^{NT}_{i,t}$) (Panel B) for stock *i* and day *t*. α_i and β_t are stock and day fixed effects. $TL_{i,t}$ is the average transmission latency between Frankfurt and London for stock *i* and day *t*. Four specifications of the model are estimated. In the regression for columns (1) and (5), the row level $TL_{i,t}$ is used as the key dependent variable. In the regression for columns (2) and (6), $TL_{i,t}$ is instrumented with XSE upgrades, and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. In the regression for columns (3) and (7), $TL_{i,t}$ is instrumented with the number of microwave networks between Frankfurt and London, and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. In the regression for columns (4) and (8), $TL_{i,t}$ is instrumented with the average $TL_{i,t}$ of all stocks on day *t* in the same size group (calculated by excluding stock *i*), and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. $C_{k,i,t}$ is a set of *k* control variables, which includes $Stddev_{i,t}$, $InvPri_{i,t}$, $lnTV_{i,t}$, $ln\ Depth_{i,t}$, $Momentum_{i,t}$, $Control_{i,t}$ and $News_{i,t}$. $Control_{i,t}$ is the value of the corresponding market quality metric ($DV_{i,t}$) for the matched stock, and $News_{i,t}$ is a dummy equal to 1 if there is news for firm *i*. All other variables and the procedure for identifying toxic transactions are defined in Tables 3 and 4
The sample consists of the 100 most active German stocks that are cross-listed on XSE and Cboe. All variables except $TL_{i,t}$ are computed for Cboe. The sample period covers March 2017 to August 2018. Standard errors are double clustered by stock and time, and *t*-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

Panel A: The impact of TL on private information following the toxic and non-toxic arbitrage opportunities

Independent variable	$Espread^{T}_{i,t}$				$Espread^{NT}_{i,t}$			
	Panel fixed effects (1)	2SLS IV: XSE upgrades (2)	2SLS IV: microwave (3)	2SLS IV: average $TL_{i,t}$ (4)	Panel fixed effects (5)	2SLS IV: XSE upgrades (6)	2SLS IV: microwave (7)	2SLS IV: average $TL_{i,t}$ (8)
$TL_{i,t}/\ \widehat{TL_{i,t}}$	−0.41*** (−3.66)	−0.39*** (−3.58)	−0.41*** (−2.83)	−0.47*** (−3.55)	1.18*** (6.25)	1.22*** (5.18)	1.24*** (3.88)	1.24*** (4.71)
$lnTV_{i,t}$	−0.28** (−2.18)	−0.31*** (−2.62)	−0.30*** (−2.74)	−0.33*** (−3.49)	−0.21** (−2.57)	−0.27*** (−3.10)	−0.26*** (−2.94)	−0.24*** (−2.81)
$ln\ Depth_{i,t}$	−0.17*** (−3.56)	−0.19*** (−3.67)	−0.16*** (−2.99)	−0.22*** (−3.28)	−0.14*** (−3.03)	−0.12*** (−3.68)	−0.09** (−2.29)	−0.15*** (−3.86)
$InvPri_{i,t}$	0.03*** (4.37)	0.02*** (3.68)	0.02*** (4.07)	0.03*** (4.54)	0.01** (2.56)	0.01** (2.45)	0.01*** (2.94)	0.01*** (2.75)
$Momentum_{i,t}$	−3.09** (−2.45)	−2.91** (−2.15)	−2.45** (−2.16)	−2.98*** (−3.03)	−1.34* (−1.76)	−1.10* (−1.94)	−0.90 (−1.09)	−1.24** (−2.15)
$Stddev_{i,t}$	0.68** (2.37)	0.73*** (2.71)	0.59** (2.58)	0.69** (2.16)	0.39* (1.73)	0.45** (2.03)	0.32* (1.94)	0.28 (1.63)
$Control_{i,t}$	0.27*** (3.49)	0.39*** (3.51)	0.36*** (3.82)	0.34*** (3.74)	0.19*** (2.78)	0.26*** (3.08)	0.28*** (2.90)	0.19** (2.15)
$News_{i,t}$	0.16*** (3.15)	0.20*** (2.90)	0.19*** (3.54)	0.14*** (2.74)	0.09** (2.31)	0.13*** (2.83)	0.14*** (2.87)	0.10** (2.33)
N	36,400	36,400	36,400	36,400	36,400	36,400	36,400	36,400
Time FE	Yes	No	No	Yes	Yes	No	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	27%	25%	23%	27%	26%	24%	24%	25%

Panel B: The impact of *TL* on the quoted spread following the toxic and non-toxic arbitrage opportunities

	$Qspread^{T}_{i,t}$				$Qspread^{NT}_{i,t}$			
$TL_{i,t}/\ \widehat{TL_{i,t}}$	−0.62*** (−3.54)	−0.64*** (−2.98)	−0.61*** (−3.08)	−0.66*** (−3.93)	1.33*** (5.17)	1.26*** (4.92)	1.27*** (4.30)	1.29*** (5.13)
$lnTV_{i,t}$	−0.47*** (−2.99)	−0.48*** (−3.61)	−0.51*** (−2.96)	−0.39*** (−2.71)	−0.21*** (−3.14)	−0.25*** (−3.37)	−0.28*** (−2.79)	−0.19** (−2.06)
$ln\ Depth_{i,t}$	−0.44*** (−4.27)	−0.41*** (−4.12)	−0.52*** (−4.82)	−0.47** (−4.04)	−0.19*** (−2.90)	−0.20*** (−3.07)	−0.18*** (−3.54)	−0.14** (−2.00)
$InvPri_{i,t}$	0.21*** (3.38)	0.24*** (3.05)	0.19*** (2.92)	0.18*** (3.59)	0.07** (2.03)	0.10** (2.11)	0.09** (2.15)	0.11** (2.41)
$Momentum_{i,t}$	−6.49*** (−4.17)	−4.45*** (−3.85)	−5.27*** (−2.83)	−4.91** (−2.39)	−3.68*** (−2.60)	−3.94*** (−2.79)	−3.79** (−2.07)	−3.38** (−2.00)
$Stddev_{i,t}$	0.78** (2.01)	0.82** (2.44)	0.70** (2.40)	0.77** (2.04)	0.46* (1.75)	0.38 (1.11)	0.45* (1.92)	0.51** (1.98)
$Control_{i,t}$	0.46*** (4.38)	0.42*** (3.79)	0.64*** (4.85)	0.45*** (4.36)	0.24*** (3.19)	0.18*** (2.76)	0.28*** (2.96)	0.21*** (2.70)
$News_{i,t}$	0.23*** (3.82)	0.25*** (4.29)	0.16*** (3.17)	0.25*** (3.80)	0.10** (2.19)	0.13*** (3.06)	0.12** (2.38)	0.10** (2.11)
N	36,400	36,400	36,400	36,400	36,400	36,400	36,400	36,400
Time FE	Yes	No	No	Yes	Yes	No	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	32%	34%	32%	32%	35%	35%	39%	34%

(*continued on next page*)

Table 5 (*continued*)

Panel A: The impact of TL on private information following the toxic and non-toxic arbitrage opportunities								
Independent variable	$Espread^{T}_{i,t}$				$Espread^{NT}_{i,t}$			
	Panel fixed effects (1)	2SLS IV: XSE upgrades (2)	2SLS IV: microwave (3)	2SLS IV: average $TL_{i,t}$ (4)	Panel fixed effects (5)	2SLS IV: XSE upgrades (6)	2SLS IV: microwave (7)	2SLS IV: average $TL_{i,t}$ (8)
Panel C: The effect of *TL* on total liquidity								
	$Espread_{i,t}$				$Qspread_{i,t}$			
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	0.48*** (3.26)	0.44*** (3.52)	0.44*** (3.24)	0.50*** (3.76)	0.77*** (3.63)	0.76*** (3.10)	0.70*** (3.04)	0.79*** (3.77)
$lnTV_{i,t}$	−0.40*** (−3.10)	−0.33*** (−2.74)	−0.38*** (−3.70)	−0.39*** (−3.03)	−0.35*** (−2.58)	−0.48*** (−2.92)	−0.47*** (−2.90)	−0.36*** (−2.56)
$ln\ Depth_{i,t}$	−0.21*** (−3.70)	−0.24*** (−3.58)	−0.25*** (−3.95)	−0.24*** (−3.72)	−0.27*** (−3.22)	−0.32*** (−3.84)	−0.31*** (−3.57)	−0.25*** (−3.24)
$InvPri_{i,t}$	0.03*** (3.93)	0.01*** (3.36)	0.01*** (3.11)	0.01*** (3.23)	0.09** (2.27)	0.11*** (2.80)	0.08** (2.14)	0.10*** (2.93)
$Momentum_{i,t}$	−3.05** (−2.45)	−2.76** (−2.26)	−2.48** (−2.09)	−2.11* (−1.91)	−3.64* (−1.78)	−3.71** (−2.05)	−3.95** (−2.02)	−4.05* (−1.94)
$Stddev_{i,t}$	0.23 (1.41)	0.26* (1.77)	0.32* (1.85)	0.24 (1.47)	0.38 (1.47)	0.37* (1.84)	0.40* (1.90)	0.31 (0.59)
$Control_{i,t}$	0.26*** (3.11)	0.35*** (3.84)	0.32*** (3.78)	0.42*** (3.92)	0.31*** (3.38)	0.39*** (3.66)	0.38*** (3.55)	0.33*** (3.13)
$News_{i,t}$	0.12*** (3.10)	0.16*** (2.68)	0.16*** (3.32)	0.14*** (3.13)	0.16*** (3.19)	0.11*** (3.84)	0.14*** (3.06)	0.20*** (3.44)
N	36,400	36,400	36,400	36,400	36,400	36,400	36,400	36,400
Time FE	Yes	No	No	Yes	Yes	No	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	23%	21%	22%	22%	31%	32%	30%	31%

with both $Espread_{i,t}$ and $Qspread_{i,t}$. This means that, on average, trading speed reduces trading costs and increases liquidity. This finding allows us to corroborate important earlier insights from the literature.

While there are extensive studies on the effects of trading speed on liquidity, the results are not clear-cut. For example, while Hendershott et al. (2011), Brogaard et al. (2015), and Boehmer et al. (2021) show that HFTs improve liquidity by reducing market makers' adverse selection and inventory management costs, Biais et al. (2015) and Shkilko and Sokolov (2020) document that HFT activity leads to a deterioration in liquidity by exposing market makers to adverse selection risk. We reconcile these inconsistencies by adapting the arguments of Foucault et al. (2017) to show that, while an increase in trading speed improves liquidity when HFTs trade against price pressures and provide liquidity, it deteriorates liquidity when HFTs exploit arbitrage opportunities due to asynchronous price adjustments to information. Foucault et al. (2017) argue that, since HFTs commonly use latency arbitrage strategies to earn profits, their impact on liquidity should depend on the nature of these latency arbitrage opportunities. This implies that the effect of speed on liquidity is a function of the opportunities that speed is deployed to exploit and, ideally, it should be investigated in the context of the source of latency arbitrage opportunities.

4.2. Latency and price discovery

The inextricable link between the underlying drivers of liquidity and price discovery is well-established in the literature (e.g., O'Hara, 2003). Furthermore, in line with Foucault et al. (2017) and our analysis in the previous section, the effects of speed on the information-driven component of the price discovery process may also depend on the nature of the latency arbitrage opportunities HFTs aim to exploit with speed. We test this expectation in this section by first decomposing information-driven changes in the price discovery process into two components (private information and public information) as in Barclay and Hendershott (2003). Thereafter, we link the two components to latency, as captured by $TL_{i,t}$, in the multivariable framework presented in equations (1) and (2). Specifically, we re-estimate equations (1) and (2), with $DV_{i,t}$ corresponding to one of (i) private information ($Prshare^{T}_{i,t}$ and $Prshare^{NT}_{i,t}$) or (ii) public information ($Pushare^{T}_{i,t}$ and $Pushare^{NT}_{i,t}$) components. The details of the variance decomposition are in Appendix B. $C_{k,i,t}$ includes $Espread_{i,t}$, $InvPri_{i,t}$, $lnTV_{i,t}$, $lnDepth_{i,t}$, $Momentum_{i,t}$, $Control_{i,t}$, and $News_{i,t}$.

Panels A and B of Table 6 report the estimation results where $DV_{i,t}$ corresponds to the private information ($Prshare^{T}_{i,t}$ and $Prshare^{NT}_{i,t}$) and public information ($Pushare^{T}_{i,t}$ and $Pushare^{NT}_{i,t}$) components, respectively. The results in Panel A show a negative and statistically significant association between $TL_{i,t}$ and $Prshare^{NT}_{i,t}$. The estimate is also economically significant, with a 1 ms increase in $TL_{i,t}$ associated with a 7.83% decrease in $Prshare^{NT}_{i,t}$. This result may be explained by insights obtained from the preceding analysis on the liquidity impact of trading speed following non-toxic arbitrage opportunities. Given that, in aggregate, HFT activity improves liquidity (see also Hendershott et al., 2011) and reduces transaction costs following incidences of non-toxic latency arbitrage opportunities (see Table 5), higher trading speed may incentivize informed trading in the post-non-toxic latency arbitrage trading environment (see also Fang et al., 2009). This implies an increased use of private information, which is in line with our evidence of a negative relation between $TL_{i,t}$ and $Prshare^{NT}_{i,t}$.

There is also a negative, and statistically and economically significant association between $TL_{i,t}$ and $Pushare^{T}_{i,t}$. Economically, a 1 ms increase in $TL_{i,t}$ decreases $Pushare^{T}_{i,t}$ by 10.79%. This estimate shows that low latency increases the share of the public information component of price discovery following incidences of toxic arbitrage opportunities. As we discuss in Subsection 3.4, this appears consistent with the expectation that HFTs use information that they source from XSE on Cboe by trading on them, and thus exploit toxic arbitrage opportunities as they arise. The exploitation terminates the opportunity and prompts market making HFTs to update their quotes by widening the bid-ask spread, as also deduced from the results in Table 5. Intuitively, a market maker quote change generates a price impact and leads to a change in the midpoint. Given that midpoint changes are associated with market makers' quote updates, and not trading, this is captured by the public information component of the midpoint return variance decomposition approach we use (e.g., Brogaard et al., 2019).

The implication is therefore that, as shown by Boehmer et al. (2018), (high-frequency) market makers' quote updates are culpable in the conversion of private information sourced from HFTs' latency arbitrage strategies into public information following incidences of toxic latency arbitrage.[16] And, irrespective of the mechanics underpinning our results, ultimately, their economic implication is that HFT activity enhances the price discovery process on Cboe by facilitating the platform's incorporation of information first observed on XSE. Results presented in Subsection 4.3 show that the latency effects hold when latency arbitrage opportunities on XSE are considered as well.

Conversely, non-toxic arbitrage opportunities should not be associated with the conveying of information to Cboe from XSE (or vice versa) and, hence, there should be no need for market maker quote updates on the heels of the latency arbitrage opportunities. This also suggests that no (permanent) price impact is observed (e.g., Degryse et al., 2014) in this scenario. Our results are consistent with this expectation[17] as we observe no statistically significant relation between $TL_{i,t}$ and $Pushare^{NT}_{i,t}$.

[16] Consistent with Boehmer et al. (2018), the same HFTs may be responsible for both the implied information-conversion activity and the subsequent market maker quote updates.

[17] We thank the editor for suggesting this analysis.

Table 6
Latency and price discovery: variance decomposition.
This table reports the coefficient estimates from the following regression model:
$DV_{i,t} = \alpha_i + \beta_t + \gamma\widehat{TL_{i,t}} + \sum_{k=1}^{7}\delta_k C_{k,i,t} + \varepsilon_{i,t}$
where $DV_{i,t}$ corresponds to the private information share ($Prshare^{T}_{i,t}$ and $Prshare^{NT}_{i,t}$) and public information share ($Pushare^{T}_{i,t}$ and $Pushare^{NT}_{i,t}$) for stock i and day t, and α_i and β_t are stock and day fixed effects. $TL_{i,t}$ is the average transmission latency between Frankfurt and London for stock i and day t. Four specifications of the model are estimated. In the regression for columns (1) and (5), the row level $TL_{i,t}$ is used as the key dependent variable. In the regression for columns (2) and (6), $TL_{i,t}$ is instrumented with XSE upgrades, and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. In the regression for columns (3) and (7), $TL_{i,t}$ is instrumented with the number of microwave networks between Frankfurt and London, and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. In the regression for columns (4) and (8), $TL_{i,t}$ is instrumented with the average $TL_{i,t}$ of all stocks on day t in the same size group (calculated by excluding stock i), and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. $C_{k,i,t}$ is a set of k control variables, which includes $Espread_{i,t}$, $InvPri_{i,t}$, $lnTV_{i,t}$, $ln\ Depth_{i,t}$, $Momentum_{i,t}$, $Control_{i,t}$ and $News_{i,t}$. $Control_{i,t}$ is the value of the corresponding market quality metric ($DV_{i,t}$) for the matched stock, and $News_{i,t}$ is a dummy equal to 1 if there is news for firm i. All other variables and the procedure for identifying toxic transactions are defined in Tables 3 and 4 The sample consists of the 100 most active German stocks that are cross-listed on XSE and Cboe. All variables except $TL_{i,t}$ are computed for Cboe. The sample period covers March 2017 to August 2018. Standard errors are double clustered by stock and time, and *t*-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

Independent variable	$Prshare^{T}_{i,t}$				$Prshare^{NT}_{i,t}$			
	Panel fixed effects (1)	2SLS IV: XSE upgrades (2)	2SLS IV: microwave (3)	2SLS IV: average $TL_{i,t}$ (4)	Panel fixed effects (5)	2SLS IV: XSE upgrades (6)	2SLS IV: microwave (7)	2SLS IV: average $TL_{i,t}$ (8)
Panel A: The impact of *TL* on private information following the toxic and non-toxic arbitrage opportunities								
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	2.19* (1.72)	1.96 (1.49)	2.04	2.22*	−3.82***	−3.53***	−3.77***	−3.32***
			(1.61)	(1.82)	(−3.07)	(−2.84)	(−3.21)	(−2.74)
$lnTV_{i,t}$	2.38**	2.96**	2.88**	2.79**	2.12**	2.54**	2.62**	2.15*
	(2.06)	(2.28)	(2.35)	(2.01)	(1.99)	(2.03)	(2.18)	(1.93)
$ln\ Depth_{i,t}$	−2.04	−2.09	−2.41	−2.03	−1.97	−1.83	−1.33	−1.80
	(−0.75)	(−1.07)	(−1.42)	(−0.61)	(−0.69)	(−0.79)	(−1.21)	(−0.56)
$InvPri_{i,t}$	−0.03	−0.07	−0.04	−0.04	−0.03	−0.06	−0.04	−0.04
	(−0.05)	(−1.09)	(−0.14)	(−1.10)	(−0.04)	(−1.07)	(−0.14)	(−1.08)
$Momentum_{i,t}$	−1.70**	−2.02**	−1.89**	−1.31*	−1.47*	−1.69**	−1.51**	−1.18
	(−2.02)	(−2.23)	(−2.33)	(−1.95)	(−1.80)	(−2.00)	(−2.02)	(−1.64)
$Espread_{i,t}$	0.99***	1.06***	0.95***	0.94***	0.70***	0.89***	0.60***	0.75***
	(5.10)	(5.46)	(4.79)	(5.72)	(3.99)	(4.06)	(2.82)	(3.43)
$Control_{i,t}$	0.36***	0.37***	0.36***	0.41***	0.32***	0.35***	0.36***	0.40***
	(3.31)	(3.92)	(3.20)	(4.32)	(3.30)	(3.54)	(3.15)	(4.29)
$News_{i,t}$	6.42***	6.54***	5.60***	6.19***	4.74***	3.24***	3.06***	4.95***
	(9.35)	(8.15)	(8.01)	(7.49)	(6.52)	(5.97)	(6.91)	(5.86)
N	36,400	36,400	36,400	36,400	36,400	36,400	36,400	36,400
Time FE	Yes	No	No	Yes	Yes	No	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	18%	17%	18%	18%	20%	18%	19%	20%
Panel B: The impact of *TL* on public information following the toxic and non-toxic arbitrage opportunities								
	$Pushare^{T}_{i,t}$				$Pushare^{NT}_{i,t}$			
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	−5.88***	−5.94***	−5.99***	−5.76***	1.13	0.16	−0.08	1.18
	(−4.48)	(−4.79)	(−4.03)	(−4.46)	(1.58)	(0.04)	(−0.74)	(1.57)
$lnTV_{i,t}$	3.56***	4.05***	2.82**	3.68***	1.88**	2.71**	1.82*	2.05**
	(2.76)	(3.13)	(2.01)	(2.75)	(2.10)	(2.29)	(1.79)	(2.01)
$ln\ Depth_{i,t}$	−5.95***	−5.76***	−5.99***	−6.16***	−3.34**	−3.61**	−2.92*	−3.77**
	(−2.83)	(−2.81)	(−2.90)	(−2.88)	(−2.17)	(−2.00)	(−1.88)	(−2.03)
$InvPri_{i,t}$	0.10**	0.09**	0.06**	0.07**	0.06**	0.06**	0.04*	0.07**
	(2.44)	(2.51)	(2.30)	(2.21)	(2.03)	(2.14)	(1.71)	(2.10)
$Momentum_{i,t}$	6.50	5.87	5.61	6.23	5.30	3.18	4.74	4.93
	(0.78)	(0.26)	(0.83)	(0.55)	(0.55)	(0.02)	(0.57)	(0.23)
$Espread_{i,t}$	−1.26***	−1.21***	−1.28***	−1.25***	−0.98***	−0.82***	−0.96***	−0.97***
	(−7.48)	(−7.68)	(−7.47)	(−7.17)	(−5.46)	(−5.92)	(−6.07)	(−5.63)
$Control_{i,t}$	0.33***	0.37***	0.23***	0.29***	0.21***	0.17***	0.20***	0.20***
	(5.08)	(5.14)	(3.75)	(5.25)	(3.97)	(2.68)	(3.04)	(3.50)
$News_{i,t}$	7.77***	7.08***	8.32***	7.91***	4.32***	4.55***	4.94***	5.07***
	(10.99)	(9.42)	(8.83)	(9.22)	(5.49)	(4.39)	(3.37)	(6.74)
N	36,400	36,400	36,400	36,400	36,400	36,400	36,400	36,400
Time FE	Yes	No	No	Yes	Yes	No	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	26%	23%	22%	25%	25%	24%	24%	24%

4.3. Looking the other way: information transmission from London (Cboe) to Frankfurt (XSE)

As discussed in Subsection 3.2, in the main analysis, *TL* is computed based on the assumption that the information is transmitted from Frankfurt (XSE) to London (Cboe). However, given that Cboe has over a third of transactions and price discovery, it is important to examine how the latency of information transmission from Cboe to XSE impacts the market quality characteristics direction as well. For this purpose, we first estimate *TL* by using price-changing trade on Cboe as a signal and the same-direction near coincident price-changing trade on XSE as a response. Thereafter, we re-estimate all the models using *TL*, which is computed based on the Cboe-XSE information transmission direction. This inevitably requires modifications in the estimation of the models: all variables except *TL* are now also computed for XSE, and we use the selection criteria described above to identify a new control group of stocks only listed on XSE. We are, however, unable to estimate the 2SLS IV using the (XSE) upgrades as instruments.

The results are reported in Appendix C. As shown in Table C1, the average latency is similar to what we observe for the XSE-Cboe direction (see Table C1). Table C2 also shows that, consistent with the results in Table 3, the number of non-toxic arbitrage opportunities is higher than the number of toxic arbitrage opportunities. However, while for the XSE-Cboe information transmission direction analysis, the ratio of non-toxic arbitrage opportunities to toxic arbitrage opportunities is about 1.85, the corresponding ratio is 4.45 for the Cboe-XSE direction. This is not surprising as over two-thirds of price discovery occurs on XSE (see Appendix A).

More importantly, the key results for both sets of analyses are qualitatively similar: the impact of trading speed on market quality following incidences of toxic and non-toxic arbitrage opportunities are not homogenous. Consistent with the results in Tables 5 and 6, we find that *TL* is negatively (positively) associated with liquidity following non-toxic (toxic) latency arbitrage opportunities (see Table C4), and we find a positive relation between trading speed and the public (private) information component of price discovery following incidences of toxic (non-toxic) latency arbitrage opportunities (see Table C5). The only crucial difference between the results in this section and those in the preceding sections is that trading speed is not statistically significant when related to overall liquidity (i.e., when the construction of the liquidity proxies is not delineated by type of latency arbitrage opportunity). This suggests that the positive and negative effects of trading speed on liquidity following non-toxic and toxic arbitrage opportunities offset each other, mainly because the economic magnitude of the trading speed on liquidity following the incidences of non-toxic arbitrage opportunities is significantly lower in the Cboe-XSE direction. For instance, while a 1 ms increase in transmission latency increases effective spread by about 18% in the XSE-Cboe information direction, the respective effect is only 4% in the Cboe-XSE information transmission direction; thus, the impact of latency when considered from London to Frankfurt in the case of our German sample of stocks is about 4.5 times lower in magnitude.

This disparity may be explained by the magnitude of the temporary price impact that initiates the non-toxic arbitrage opportunities being lower on average when the opportunity starts on Cboe. To illustrate, consider that the prices for a given cross-listed stock on XSE and Cboe are the same: €100. Suppose in the first case that the price increases by 0.5% to €100.5 on XSE due to a (temporary) liquidity shock, and thus a non-toxic arbitrage opportunity emerges on Cboe and the Cboe price then increases to €100.5. Arbitrageurs will subsequently provide liquidity to push prices back to the previous level since the arbitrage opportunity is non-toxic. In a second case, the price increases by 0.1% to €100.1 on Cboe, also due to a liquidity shock, and a non-toxic arbitrage opportunity starts on XSE, followed by a price increase to €100.1 on XSE. Similarly, arbitrageurs provide liquidity to push the prices back to the previous level. Intuitively, arbitrageurs need to provide more liquidity to push prices back to the previous level (€100) in the first case as the magnitude of the price change is higher (0.5% vs. 0.1%). Given that market makers' quote updates reflect the amount of liquidity provision, bid-ask spread reduction following incidences of non-toxic arbitrage opportunity will be higher in the first case. To ascertain that our data confirms this interpretation, we compare the price impacts of trades that cause temporary liquidity shocks on XSE and Cboe. In particular, we compare the average percentage change in prices for price-changing trades that initiate non-toxic arbitrage opportunities on XSE and Cboe. We find that, as expected, the magnitude (absolute value) of the average percentage change in prices due to temporary liquidity shock is about 2.3 times lower on Cboe (0.12% on Cboe and 0.27% on XSE).

5. Conclusion

In this study, we examine the effect of latency, as an inverse proxy for the speed of trading, on market quality characteristics. By estimating the latency between Frankfurt (XSE) and London (Cboe) from transaction-level data, we show that prices in London largely respond to price changes in Frankfurt within 3–5 ms. This result is consistent with the latencies published by the providers of microwave and fiber optic connections between London and Frankfurt, and thus underscores the relevance of our information transmission latency estimation method. The evolution of the latency measure also closely aligns with technological upgrades that should enhance information transmission between the two venues.

By exploiting the technological upgrades, in a bid to address endogeneity concerns, we find that, consistent with the predictions of Foucault et al. (2017), the effects of speed on market quality characteristics in a cross-market setting are dependent on the underlying drivers of speed as deployed by HFTs. A decrease in the information transmission latency, between the home and satellite markets, amplifies liquidity in the satellite market when exploitable latency arbitrage opportunities arise due to temporary order imbalance (non-toxic price pressures). It impairs liquidity, however, when toxic latency arbitrage opportunities are driven by the arrival of new information.

Similarly, we show that HFTs' contribution to the price discovery process is also contingent on the nature of the latency arbitrage opportunities they pursue. Specifically, we find a positive link between HFT activity and the public information component of price changes following occurrences of toxic latency arbitrage opportunities. This finding can be attributed to HFTs' ability to convey private information from one market to the other through their trades when exploiting toxic latency arbitrage opportunities in a cross-market setting, information which is then reflected in high-frequency market maker quote updates (see Boehmer et al., 2018). We also find that HFT activity is linked to an increase in private informed trading following non-toxic latency arbitrage opportunities. We suggest that this is linked to a reduction in HFT activity facilitating a reduction in transaction costs (e.g., Hendershott et al., 2011), which encourages the use of private information.

Altogether, our findings help reconcile the inconsistencies in the extant literature on the relation between speed/AT/HFT and market quality by identifying the conditions under which the conflicting findings hold. Contradictory evidence on the effects of HFTs can complicate policymakers and exchange operators' efforts to craft policies/regulations and develop market structures aimed at effectively harnessing the potential benefits of HFTs, while limiting their possible negative effects. Thus, in highlighting the complexity of the link between trading speed and market quality characteristics, we address a significant gap in the literature and offer actionable evidence for policy and practice. Indeed, our findings have potentially wide-ranging implications for policymakers, market regulators, and exchange operators. Increasing competition between trading venues, and eliminating HFTs' speed advantages have been the subjects of heated debates among stakeholders, including politicians, who have advocated for reining in the activities of HFTs in recent years. For instance, after an initially lukewarm reception, exchanges are increasingly enthusiastic about the idea of implementing "speed bumps" to eliminate the speed advantages of HFTs.[18] Some exchanges have gone further and completely banned "aggressive" HFTs from their venues.[19] Our results, however, also show that HFTs can be valuable contributors to liquidity provision and price discovery. Importantly, we present new findings providing a more nuanced view on the crucial debate about the implications of HFTs on market quality.

Appendix

A. Price discovery analysis

This table presents the results for three different price discovery metrics estimating the share of price discovery for XSE and Cboe. IS is the information share metric as developed by Hasbrouck (1995), CS is the component share metric based on Gonzalo and Granger (1995), and ILS is the information leadership share as defined by Putniņš (2013). All estimates are computed based on price samples at the 1-s frequency. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe. The sample period covers March 2017 to August 2018.

Table A1

Metric	Value
IS	0.69
CS	0.64
ILS	0.61

B. Stock price return variance decomposition

Following Hasbrouck (1991) and Barclay and Hendershott (2003), we model the midpoint price as the sum of permanent (random-walk) and transitory (stationary) components:

$$p_t = m_t + s_t, \tag{B1}$$

where p_t is the midpoint price for trade t, m_t is the permanent (efficient price), and s_t is the transitory (pricing error) component. m_t follows a random walk with innovation v_t:

$$m_t = m_{t-1} + v_t, \tag{B2}$$

where $v_t \sim N(0, \sigma_v^2)$ and $Ev_t v_s = 0$ for $t \neq s$.

The difference between p_t and p_{t-1}, the stock midpoint return (r_t), is:

$$r_t = p_t - p_{t-1} = v_t + \Delta s_t. \tag{B3}$$

Consistent with Hasbrouck (1991) and Barclay and Hendershott (2003), we assume two sources of information: private information incorporated through trading and public information incorporated without trading (generally occurring through market makers' quote updates). This implies that the innovation in the permanent component (v_t) is either driven by innovation in r_t (public information, $\varepsilon_{r_t,t}$) or innovation in x_t (private information, $\varepsilon_{x_t,t}$):

$$v_t = a\varepsilon_{r_t,t} + b\varepsilon_{x_t,t}. \tag{B4}$$

[18] https://www.wsj.com/articles/more-exchanges-add-speed-bumps-defying-high-frequency-traders-11564401611.
[19] https://www.ft.com/content/04944696-c9d9-11e5-a8ef-ea66e967dd44.

We estimate both constituents of v_t, in line with prior studies (e.g., Hasbrouck 1991; Barclay and Hendershott 2003; Brogaard et al., 2022), by estimating the following vector autoregression (VAR) model:

$$\begin{aligned} r_t &= \sum_{i=1}^{p} \alpha_i r_{t-i} + \sum_{i=1}^{p} \beta_i x_{t-i} + \varepsilon_{r_t,t} \\ x_t &= \sum_{i=0}^{p} \gamma_i r_{t-i} + \sum_{i=1}^{p} \delta_i x_{t-i} + \varepsilon_{x_t,t,} \end{aligned} \tag{B5}$$

where x_t is the signed currency volume (the number of shares traded) calculated as $Ln(Volume_t)$ ($-Ln(Volume_t)$) for buy (sell) transactions[20], and r_t is the midpoint return for trade t. As in Brogaard et al. (2019), p equals 5.

Note the in-built assumptions regarding the contemporaneous effects in equation (B5). The summation i=0 in equation (B5) for lags of midpoint return, r_t, in the currency volume, x_t, equation, but i=1 for lags of signed currency volume in the midpoint return equation. This causality assumption arises from the timing convention inherent in our data, whereby an investor observes midpoint changes (r_t) during each interval t before executing the trade (x_t). Specifically, the data only captures quotes observed immediately prior to each transaction in the dataset. As such, the trade is the last event to occur in interval t, and its price impact is observable in period t+1, as captured by the midpoint return equation.[21] A similar timing convention is used in Brogaard et al. (2022), in which market return has a contemporaneous effect on trade.

Taking the variances of both sides of equation (B4) yields the following:

$$\sigma_v^2 = a^2 \sigma_{\varepsilon_r}^2 + b^2 \sigma_{\varepsilon_x}^2, \tag{B6}$$

where a and b are the impulse response functions obtained from the estimation of equation (B5). In equation (B6), σ_v^2 is the variance of stock return driven by information, $a^2\sigma_{\varepsilon_r}^2$ is the variance of stock return driven by public information, and $b^2\sigma_{\varepsilon_x}^2$ is the variance of stock return driven by private information. All the components of equation (B6) are computed using the VAR defined in equation (B5). We then normalize $a^2\sigma_{\varepsilon_r}^2$ and $b^2\sigma_{\varepsilon_x}^2$ by σ_v^2 to obtain the main variables of interest that we use in the regression models outlined in Subsection 4.2 [22]

$$\begin{aligned} PuShare_{i,t} &= \frac{a^2\sigma_{\varepsilon_r}^2}{\sigma_v^2} \\ PrShare_{i,t} &= \frac{b^2\sigma_{\varepsilon_x}^2}{\sigma_v^2} \end{aligned} \tag{B7}$$

C. Latency arbitrage opportunities in the direction of Cboe and XSE

In the main estimations, we assume that information is transmitted from Frankfurt (XSE) to London (Cboe). In the estimations for Tables C1 and C2 we relax this assumption and re-estimate all models by assuming that information is transmitted from Cboe to Frankfurt.

Table C1

Information transmission latency between Cboe and XSE. This table presents different statistics for the information transmission latency between Cboe and XSE. Panel A reports the number of responses on XSE to price-changing trades on Cboe for different time bins in milliseconds (ms). Panel B presents the mean, median and standard deviation of the information transmission latency between Cboe and XSE. The sample consists of the 100 most active German stocks cross-listed on Cboe and XSE. The sample period covers March 2017 to August 2018

Speed (ms)	Frequency	Percentage
Panel A: The distribution of the latencies of responses on Cboe to price-changing trades on XSE		
3	404,027	40.66
4	196,183	19.74
5	181,236	18.24
6	55,909	5.63

(*continued on next page*)

[20] We use the natural logarithm of trading volume to capture the convexity in the effect of trading size returns. We thank the editor for this suggestion.
[21] We thank an anonymous advisory editor for this suggestion.
[22] Another reasonable approach is normalizing the estimated private and public components by the total variance ($\sigma_v^2 + \sigma_s^2$). We normalize by σ_v^2 as it arguably captures long-term risk better. This is because we expect the noise disappears in the long-run.

Table C1 (*continued*)

Speed (ms)	Frequency	Percentage
7	49,052	4.94
8	35,734	3.60
9	38,859	3.91
10	32,697	3.29
Panel B: Summary statistics of the information transmission latency		
Mean	Median	Standard deviation
4.573	4	1.915

Table C2

Number of toxic and non-toxic arbitrage opportunities. This table reports the number of toxic and non-toxic arbitrage opportunities for the news and no-news days. To split the transactions into toxic and non-toxic groups, the approach of Foucault et al. (2017) is adapted, following the rationale that toxic (non-toxic) transactions arise due to information (transitory price pressures) and therefore generate permanent (temporary) price impacts. Firstly, near-coincident same-direction price-changing trades on XSE are identified. Thereafter, for each near-coincident transaction s at event time t, the midpoint of the prevailing quote at time t (M_t) is compared with the midpoint of the prevailing quote at time t+4 (M_{t+4}). A price reversal, i.e., M_{t+4} reversing toward M_t, implies that the near-coincident transaction is non-toxic. The remaining transactions are labeled as toxic. Days with firm news are identified in two steps (see Hirschey 2020). In the first step, we use Factiva, which contains news from over 35,000 sources, to identify news days. Then, to eliminate the possibility that traders respond to news not covered by Factiva, in the second step, we exclude days with absolute market-adjusted returns greater than 1%. The statistical tests conducted are two-sample t-tests and pairwise Wilcoxon-Mann-Whitney tests. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe. The sample period covers March 2017 to August 2018. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

	Toxic arbitrage opportunities			Non-toxic arbitrage opportunities		
	Total number of arbitrage opportunities	Average number of arbitrage opportunities per stock	Fraction of arbitrage opportunities	Total number of arbitrage opportunities	Average number of arbitrage opportunities per stock	Fraction of arbitrage opportunities
News days	126,085	346.39	69.10%	384,323	1055.83	47.38%
No-news days	56,377	154.889	30.90%	426,912	1172.84	52.62%
All days	182,462	501.27		811,235	2228.67	
News–no-news days	69,708	191.51 ***	38.20%	−42,589	−117.00	−5.25%
t-test p-value		<0.001			0.17	
W-M-W test p-value		<0.001			0.25	

Table C3

Definition of variables and summary statistics. This table defines the variables calculated for each stock i and day t, and reports the summary statistics for the main variables used in the study. To split transactions into toxic and non-toxic groups, we use the modified version of the classification algorithm described in Foucault et al. (2017). Further details of the classification algorithm are in Subsection 3.3 and Table 3. The sample consists of the 100 most active German stocks cross-listed on XSE and Cboe and the summary statistics are computed for XSE. The sample period covers March 2017 to August 2018.

Variable	Definition	Mean	Median	SD
$Espread^{T}_{i,t}$ (bps)	Toxic latency arbitrage effective spread for stock i and day t, computed as the daily volume-weighted average of twice the absolute value of the difference between the transaction prices that are observed following toxic latency arbitrages and the midpoint of the bid-ask spread prevailing before these transactions, divided by the midpoint.	7.22	6.58	14.33
$Espread^{NT}_{i,t}$ (bps)	Non-toxic latency arbitrage effective spread for stock i and day t, computed as the daily volume-weighted average of twice the absolute value of the difference between the transaction prices that are observed following non-toxic latency arbitrages and the midpoint of the bid-ask spread prevailing before these transactions, divided by the midpoint.	5.35	5.11	8.95
$Qspread^{T}_{i,t}$ (bps)	Toxic latency arbitrage quoted spread for stock i and day t, computed as the daily volume-weighted average of the difference between the prevailing ask and bid prices corresponding to transactions that are observed following toxic latency arbitrages divided by the midpoint of these ask and bid prices.	7.91	7.43	7.36
$Qspread^{NT}_{i,t}$ (bps)	Non-toxic latency arbitrage quoted spread for stock i and day t, computed as the daily volume-weighted average of the difference between the prevailing ask and bid prices corresponding to transactions that are observed following non-toxic latency arbitrages divided by the midpoint of these ask and bid prices.	4.82	4.59	5.22
$Prshare^{T}_{i,t}$ (%)	The share of private information in the variance of returns for transactions that are observed following toxic latency arbitrages in stock i and day t as described in Appendix B.	49.51	40.28	22.60
$Prshare^{NT}_{i,t}$ (%)	The share of private information in the variance of returns for transactions that are observed following non-toxic latency arbitrages in stock i and day t as described in Appendix B.	51.17	41.45	23.15
$Pushare^{T}_{i,t}$ (%)	The share of public information in the variance of returns for transactions that are observed following toxic latency arbitrages in stock i and day t as described in Appendix B.	50.49	50.03	29.43

(continued on next page)

Table C3 (*continued*)

Variable	Definition	Mean	Median	SD
$Pushare^{NT}_{i,t}$ (%)	The share of public information in the variance of returns for transactions that are observed following non-toxic latency arbitrages in stock *i* and day *t* as described in Appendix B.	48.83	45.49	36.25
$lnTV_{i,t}$	Natural logarithm of trading volume for stock *i* and day *t*.	14.24	15.11	7.48
$lnDepth_{i,t}$	Natural logarithm of the sum of the ask and bid sizes for stock *i* and day *t*.	16.05	15.83	13.77
$InvPri_{i,t}$ (bps)	Inverse of last transaction price for stock *i* and day *t*.	303.02	268.91	365.69
$Momentum_{i,t}$ (bps)	The first lag of the daily return for stock *i* and day *t*.	0.581	0.526	17.02

Table C4
Latency and liquidity, This table reports the coefficient estimates from the following regression model:
where $DV_{i,t}$ corresponds to either the effective spread ($Espread^{T}_{i,t}$ and $Espread^{NT}_{i,t}$) (Panel A) or quoted spread ($Qspread^{T}_{i,t}$ and $Qspread^{NT}_{i,t}$) (Panel B) for stock *i* and day *t*. α_i and β_t are stock and day fixed effects. $TL_{i,t}$ is the average transmission latency between London and Frankfurt for stock *i* and day *t*. Three specifications of the model are estimated. In the regression for columns (1) and (4), the row level $TL_{i,t}$ is used as the key dependent variable. In the regression for columns (2) and (5), $TL_{i,t}$ is instrumented with the number of microwave networks between Frankfurt and London, and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. In the regression for columns (3) and (6), $TL_{i,t}$ is instrumented with the average $TL_{i,t}$ of all stocks on day *t* in the same size group (calculated by excluding stock *i*), and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. $C_{k,i,t}$ is a set of *k* control variables, which includes $Stddev_{i,t}$, $InvPri_{i,t}$, $lnTV_{i,t}$, $ln\,Depth_{i,t}$, $Momentum_{i,t}$, $Control_{i,t}$ and $News_{i,t}$. $Control_{i,t}$ is the value of the corresponding market quality metric ($DV_{i,t}$) for the matched stock, and $News_{i,t}$ is a dummy equal to 1 if there is news for firm *i*. All other variables and the procedure for identifying toxic transactions are defined in Tables 3 and 4 The sample consists of the 100 most active German stocks that are cross-listed on XSE and Cboe. All variables except $TL_{i,t}$ are computed for XSE. The sample period covers March 2017 to August 2018. Standard errors are double clustered by stock and time, and *t*-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively.

Panel A: The impact of TL on the effective spread following the toxic and non-toxic arbitrage opportunities						
Independent variable	$Espread^{T}_{i,t}$			$Espread^{NT}_{i,t}$		
	Panel fixed effects (1)	2SLS IV: microwave (2)	2SLS IV: average $TL_{i,t}$ (3)	Panel fixed effects (4)	2SLS IV: microwave (5)	2SLS IV: average $TL_{i,t}$ (6)
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	−0.18** (−2.31)	−0.19** (−2.07)	−0.21** (−2.52)	0.23** (2.42)	0.26** (2.14)	0.27** (2.35)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,400	36,400	36,400	36,400	36,400	36,400
R^2	30%	28%	30%	31%	28%	29%
Panel B: The impact of *TL* on the quoted spread following the toxic and non-toxic arbitrage opportunities.						
$Qspread^{T}_{i,t}$				$Qspread^{NT}_{i,t}$		
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	−0.09* (−1.83)	−0.10** (−2.01)	−0.10** (−2.08)	0.13** (2.17)	0.14** (2.01)	0.13** (2.06)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,400	36,400	36,400	36,400	36,400	36,400
R^2	37%	25%	37%	34%	34%	33%
Panel C: The effect of *TL* on total liquidity.						
$Espread_{i,t}$				$Qspread_{i,t}$		
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	0.04 (1.27)	0.05 (1.12)	0.04 (0.86)	0.02 (1.25)	0.02 (0.66)	0.01 (0.51)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,400	36,400	36,400	36,400	36,400	36,400
R^2	22%	19%	20%	28%	27%	25%

$$DV_{i,t} = \alpha_i + \beta_t + \gamma \widehat{TL_{i,t}} + \sum_{k=1}^{7} \delta_k C_{k,i,t} + \varepsilon_{i,t}$$

Table C5. Latency and price discovery: variance decomposition

Panel A: The impact of *TL* on private information following the toxic and non-toxic arbitrage opportunities						
Independent variable	$Prshare^{T}_{i,t}$			$Prshare^{NT}_{i,t}$		
	Panel fixed effects (1)	2SLS IV: microwave (2)	2SLS IV: average $TL_{i,t}$ (3)	Panel fixed effects (4)	2SLS IV: microwave (5)	2SLS IV: average $TL_{i,t}$ (6)
$TL_{i,t}$/ $\widehat{TL_{i,t}}$	2.43** (2.03)	2.36** (2.21)	2.28** (2.02)	−3.05*** (−3.09)	−2.98*** (−2.79)	−2.49** (−2.35)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

(*continued on next page*)

(*continued*)

Panel A: The impact of *TL* on private information following the toxic and non-toxic arbitrage opportunities						
Independent variable	$Prshare^{T}_{i,t}$			$Prshare^{NT}_{i,t}$		
	Panel fixed effects (1)	2SLS IV: microwave (2)	2SLS IV: average $TL_{i,t}$ (3)	Panel fixed effects (4)	2SLS IV: microwave (5)	2SLS IV: average $TL_{i,t}$ (6)
Time FE	Yes	No	Yes	Yes	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,400	36,400	36,400	36,400	36,400	36,400
R^2	20%	19%	19%	21%	20%	23%
Panel B: The impact of *TL* on public information following the toxic and non-toxic arbitrage opportunities						
	$Pushare^{T}_{i,t}$			$Pushare^{NT}_{i,t}$		
$TL_{i,t}/\ \widehat{TL_{i,t}}$	−3.92*** (−3.53)	−4.05*** (−3.36)	−2.99** (−2.45)	1.67* (1.72)	1.92* (1.81)	1.84* (1.86)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	Yes	No	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
N	36,400	36,400	36,400	36,400	36,400	36,400
R^2	29%	30%	25%	26%	27%	26%

This table reports the coefficient estimates from the following regression model:

$DV_{i,t} = \alpha_i + \beta_t + \gamma \widehat{TL_{i,t}} + \sum_{k=1}^{7} \delta_k C_{k,i,t} + \varepsilon_{i,t},$

where $DV_{i,t}$ corresponds to the private information share ($Prshare^{T}_{i,t}$ and $Prshare^{NT}_{i,t}$) and public information share ($Pushare^{T}_{i,t}$ and $Pushare^{NT}_{i,t}$) for stock i and day t, and α_i and β_t are stock and day fixed effects. $TL_{i,t}$ is the average transmission latency between London and Frankfurt for stock i and day t. Three specifications of the model are estimated. In the regression for columns (1) and (4), the row level $TL_{i,t}$ is used as the key dependent variable. In the regression for columns (2) and (5), $TL_{i,t}$ is instrumented with the number of microwave networks between Frankfurt and London, and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. In the regression for columns (3) and (6), $TL_{i,t}$ is instrumented with the average $TL_{i,t}$ of all stocks on day t in the same size group (calculated by excluding stock i), and the fitted value of $TL_{i,t}$ ($\widehat{TL_{i,t}}$) is used as the key dependent variable. $C_{k,i,t}$ is a set of k control variables, which includes $Espread_{i,t}$, $InvPri_{i,t}$, $lnTV_{i,t}$, $ln\,Depth_{i,t}$, $Momentum_{i,t}$, $Control_{i,t}$ and $News_{i,t}$. $Control_{i,t}$ is the value of the corresponding market quality metric ($DV_{i,t}$) for the matched stock, and $News_{i,t}$ is a dummy equal to 1 if there is news for firm i. All other variables and the procedure for identifying toxic transactions are defined in Tables 3 and 4 The sample consists of the 100 most active German stocks that are cross-listed on XSE and Cboe. All variables except $TL_{i,t}$ are computed for XSE. The sample period covers March 2017 to August 2018. Standard errors are double clustered by stock and time, and t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 levels respectively

References

Aquilina, M., Budish, E., O'Neill, P., 2022. Quantifying the high-frequency trading "arms race". Q. J. Econ. 137, 493–564.
Barclay, M.J., Hendershott, T., 2003. Price discovery and trading after hours. Rev. Financ. Stud. 16, 1041–1073.
Baron, M., Brogaard, J., Hagströmer, B., Kirilenko, A., 2019. Risk and return in high-frequency trading. J. Financ. Quant. Anal. 54, 993–1024.
Biais, B., Foucault, T., Moinas, S., 2015. Equilibrium fast trading. J. Financ. Econ. 116, 292–313.
Biais, B., Hillion, P., Spatt, C., 1999. Price discovery and learning during the preopening period in the Paris Bourse. J. Polit. Econ. 107, 1218–1248.
Boehmer, E., Fong, K.Y., Wu, J., 2021. Algorithmic trading and market quality: International evidence. J. Financ. Quant. Anal. 56, 2659–2688.
Boehmer, E., Li, D., Saar, G., 2018. The competitive landscape of high-frequency trading firms. Rev. Financ. Stud. 31, 2227–2276.
Boulton, T.J., Braga-Alves, M.V., 2010. The skinny on the 2008 naked short-sale restrictions. J. Financ. Mark. 13, 397–421.
Brogaard, J., Hagströmer, B., Nordén, L., Riordan, R., 2015. Trading fast and slow: colocation and liquidity. Rev. Financ. Stud. 28, 3407–3443.
Brogaard, J., Hendershott, T., Hunt, S., Ysusi, C., 2014a. High-frequency trading and the execution costs of institutional investors. Financ. Rev. 49, 345–369.
Brogaard, J., Hendershott, T., Riordan, R., 2014b. High-frequency trading and price discovery. Rev. Financ. Stud. 27, 2267–2306.
Brogaard, J., Hendershott, T., Riordan, R., 2019. Price discovery without trading: evidence from limit orders. J. Finance 74, 1621–1658.
Brogaard, J., Nguyen, T.H., Putnins, T.J., Wu, E., 2022. What moves stock prices? The roles of news, noise, and information. Rev. Financ. Stud. 35, 4341–4386.
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, 1547–1621.
Copeland, T.E., Galai, D., 1983. Information effects on the bid-ask spread. J. Finance 38, 1457–1469.
Degryse, H., de Jong, F., Kervel, V.v., 2015. The impact of dark trading and visible fragmentation on market quality. Rev. Finance 19, 1587–1622.
Degryse, H., de Jong, F., van Kervel, V., 2014. Does Order Splitting Signal Uninformed Order Flow. VU University, Amsterdam. Unpublished working paper.
Fang, V.W., Noe, T.H., Tice, S., 2009. Stock market liquidity and firm value. J. Financ. Econ. 94, 150–169.
Foley, S., Putniņš, T.J., 2016. Should we be afraid of the dark? Dark trading and market quality. J. Financ. Econ. 122, 456–481.
Foucault, T., Hombert, J., Roşu, I., 2016. News trading and speed. J. Finance 71, 335–382.
Foucault, T., Kozhan, R., Tham, W.W., 2017. Toxic arbitrage. Rev. Financ. Stud. 30, 1053–1094.
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.
Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. J. Bus. Econ. Stat. 13, 27–35.
Grammig, J., Melvin, M., Schlag, C., 2005. Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects. J. Empir. Finance 12, 139–164.
Hasbrouck, J., 1991. Measuring the information content of stock trades. J. Finance 46, 179–207.
Hasbrouck, J., 1995. One security, many markets: determining the contributions to price discovery. J. Finance 50, 1175–1199.
Hasbrouck, J., Saar, G., 2013. Low-latency trading. J. Financ. Mark. 16, 646–679.
Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does algorithmic trading improve liquidity? J. Finance 66, 1–33.
Hirschey, N., 2020. Do high-frequency traders anticipate buying and selling pressure? Manag. Sci. 67, 3321–3345.
Hoffmann, P., 2014. A dynamic limit order market with fast and slow traders. J. Financ. Econ. 113, 156–169.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
Laughlin, G., Aguirre, A., Grundfest, J., 2014. Information transmission between financial markets in Chicago and New York. Financ. Rev. 49, 283–312.
Malceniece, L., Malcenieks, K., Putniņš, T.J., 2019. High frequency trading and comovement in financial markets. J. Financ. Econ. 134, 381–399.
Menkveld, A.J., 2013. High frequency trading and the new market makers. J. Financ. Mark. 16, 712–740.
Menkveld, A.J., 2016. The economics of high-frequency trading: taking stock. Annual Review of Financial Economics 8, 1–24.
Menkveld, A.J., Zoican, M.A., 2017. Need for speed? Exchange latency and liquidity. Rev. Financ. Stud. 30, 1188–1228.
O'Hara, M., 2003. Presidential address: liquidity and price discovery. J. Finance 58, 1335–1354.
O'Hara, M., 2015. High frequency market microstructure. J. Financ. Econ. 116, 257–270.
Putniņš, T.J., 2013. What do price discovery metrics really measure? J. Empir. Finance 23, 68–83.
Roşu, I., 2019. Fast and slow informed trading. J. Financ. Mark. 43, 1–30.
Shkilko, A., Sokolov, K., 2020. Every cloud has a silver lining: fast trading, microwave connectivity and trading costs. J. Finance 75, 2899–2927.
Stock, J.H., Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press.
Weller, B.M., 2018. Does algorithmic trading reduce information acquisition? Rev. Financ. Stud. 31, 2184–2226.