UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Algorithmic trading and market quality

Aitken, Michael; Cumming, Douglas; Zhan, Feng

DOI:
10.1016/j.jbankfin.2023.106843

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Aitken, M, Cumming, D & Zhan, F 2023, 'Algorithmic trading and market quality: International evidence of the impact of errors in colocation dates', *Journal of Banking & Finance*, vol. 151, 106843.
https://doi.org/10.1016/j.jbankfin.2023.106843

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 19. Jul. 2023

Journal of Banking and Finance 151 (2023) 106843

Contents lists available at ScienceDirect

Journal of Banking and Finance

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

Algorithmic trading and market quality: International evidence of the impact of errors in colocation dates ☆

Michael Aitken [a], Douglas Cumming [b,d,*], Feng Zhan [c]

[a] Macquarie Business School, Macquarie University, 4 Eastern Road, NSW 2109, Australia
[b] DeSantis Distinguished Professor, College of Business, Florida Atlantic University, 777 Glades Road, Boca Raton, Florida 33431
[c] DAN Department of Management & Organizational Studies, University of Western Ontario, London, Ontario N6A 5C2, Canada
[d] Visiting Professor, Birmingham Business School, University of Birmingham, University House, 116 Edgbaston Park Rd, Birmingham B15 2TY, United Kingdom

ARTICLE INFO

Article history:
Received 11 November 2021
Accepted 1 April 2023
Available online 2 April 2023

JEL Codes:
G12
G14
G18

Keywords:
High frequency trading
Algorithmic trading
Colocation

ABSTRACT

This paper examines evidence on colocation dates and their impact on market efficiency. International colocation dates can be sourced from a number of avenues including: [1] an 'exchange's news announcements and reports, [2] news media, and [3] by direct communication with the officers of an exchange. Boehmer et al. (2021) report colocation dates based on [1] and [2] and do not reference prior work that reports colocation dates that are primarily sourced from [3]. The consequence is that the discrepancies between prior studies and Boehmer et al. (2021) are significant and economically meaningful: the errors average 12.75 months with the largest being 46 months. This paper documents these discrepancies and provides evidence of how these differences in colocation dates matter for evidence of their impact on market efficiency.

© 2023 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

Non-standard errors, or errors arising from differences in human judgement, can give rise to mistaken inferences in empirical work in finance (Menkveld et al., 2022). In this paper, we show that non-standard errors are very pronounced in research pertaining to colocation. Also, we highlight a more general point: academics in financial economics are surprisingly hesitant to ask industry professionals directly for information. Other fields (e.g., management) do not suffer as much from this hesitancy. All too often, finance papers make inferences about certain facts rather than simply asking.[1]

Colocation refers to an exchange offering a dedicated space for high frequency traders (HFTs) and other parties to locate their servers in the same server room as the stock exchange servers. All parties are connected to stock exchange servers with each HFT box the exact distance from stock exchange servers as every other party. This innovation is potentially important for market efficiency because it enables HFTs to execute their trades more quickly (Aitken et al., 2015, 2017; Boehmer et al., 2021; Allen et al., 2021) but not more quickly than their direct competitors.[2]

Prior research indicates that colocation was introduced in response to HFTs entering the market (Aitken et al., 2015, 2017). Colocation was a service offered by the exchange to earn a fee from HFTs to gain a speed advantage. Instead of being located across the street, the HFT server is in the exchange computer room, and that distance of not having to go across the street saves HFTs a small but meaningful amount of time. This paper shows that an average of 44.9 months pass from the start of HFTs having a notable presence in the market to the exchange offering colocation services (and this fact is also reported in Aitken et al. (2015, 2017). Aitken et al. (2015, 2017) show that the introduction of colocation services does not seem to make a material difference for market efficiency and integrity in an international setting, that colocation is endogenous, and that the start of HFT seems to be a more pertinent economic and statistically significant factor in explaining changes in market efficiency and integrity.

A more recent paper by Boehmer et al. (2021) also examines whether colocation affects market efficiency – specif-

☆ We owe thanks to two anonymous reviewers, Carol Alexander, David Gallagher, Sofia Johan, April Knill, Evgeny Lyandres, Jay Ritter, and Tom Smith for helpful comments and suggestions.

* Corresponding author.
E-mail address: cummingd@fau.edu (D. Cumming).

[1] We owe thanks to an anonymous reviewer for this helpful comment and the language used in the latter part of this paragraph.

[2] Given there is a significant cost to implementing this service, the issue of fair dealing between HFT's versus the rest of the market remains.

https://doi.org/10.1016/j.jbankfin.2023.106843
0378-4266/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

ically liquidity, as measured by bid-ask spreads for example. Boehmer et al. (2021) do not refer to prior works on the topic. However, on page 16 of their working paper (as the time of writing this paper, their article was not yet in print), they state:

"As a potential caveat, the precision of reporting on colocation event dates could differ across countries. Yet, to the extent that the resulting errors are random, they should not affect the consistency of the IV estimator because such random errors would be captured by the regression error."

This statement seems to acknowledge that different sources provide different colocation dates. Specifically, the colocation dates reported in Aitken et al. (2015, 2017) differ from the colocation dates reported in Boehmer et al. (2021). The purpose of this paper is to inform readers of the discrepancies in colocation dates from primary and secondary sources and to explain why these differences may matter both statistically and economically. Also, we address the question as to whether these resulting errors are in fact "random". Finally, we point out that whether or not the errors in Boehmer et al. (2021) are systematic or random, the magnitude of the errors can and do impact the statistical inferences in their analyses such that they cannot be relied upon. With a panel dataset spanning approximately a decade, their average error is over a year, and their errors range up to four years; as such, it is reasonable to suspect that inferences drawn from the errors in Boehmer et al. (2021) are not going to be reliable, as we demonstrate in this paper.

In failing to ask the exchanges when they first offered colocation, we believe Boehmer et al. (2021) have introduced errors into their analysis; for the exchanges in the Aitken et al. sample, their errors average 12.75 months with the largest being 46 months away from the date identified by the exchange itself. Evidence of those errors is provided here in this paper.

Further, in this paper we show that these errors matter for the evidence presented in Boehmer et al. (2021). For the exchanges in the Aitken et al. (2015, 2017) data, the incorrect colocation dates show a relation with market quality using raw comparison of means tests (not controlling for other things being equal), but no relation with market quality when controlling for other things being equal. Using corrected colocation dates can give rise to evidence that colocation affects market quality, but again only with incomplete controls. High-frequency traders pre-date the introduction of colocation in a market; that is, colocation is an endogenous service offered to serve high-frequency traders and the timing of the introduction of colocation can be predicted by the arrival of high-frequency traders. After controlling for the presence of high-frequency traders, there is no impact of colocation on market quality.

More generally, we believe the colocation dates reported in Boehmer et al. (2021) should not be relied on by others, as they could introduce mistakes that could materially affect statistical inferences in other research contexts. For some of the exchanges we have direct evidence reported here that the Boehmer et al. dates are incorrect, such as that for Canada and Switzerland. For other exchanges, we do not have direct evidence that the Boehmer et al. dates are necessarily incorrect, but merely evidence of conflicting reported dates. Hence, we encourage other scholars to look at and acknowledge different sources and check for robustness.

The paper proceeds as follows. In Section 2, we document different approaches to sourcing colocation dates. We also provide at the back of the paper selected examples of documentation from exchanges as to when colocation started. That is, we provide four Exhibits in the form of emails from exchanges identifying the reported start dates for colocation on their exchange. In Section 3 we introduce a simple dataset to address the question as to whether for selected data the issue of differences in colocation dates might matter for empirical evidence on the topic. Section 4 provides empirical tests, and these tests show that the differences and errors do in fact matter. Also, Section 4 provides evidence that suggests that the differences in colocation dates might not be random as claimed. The last section offers concluding remarks.

2. Direct evidence of errors in colocation dates

Table 1 summarizes the reported differences in colocation dates in Boehmer et al. (2021) and Aitken et al. (2015, 2017). Colocation dates may be found through alternative sources, including:

(1) an exchange's news announcements and/or reports.
(2) news media, and/or
(3) by directly asking the officers of an exchange.

Aitken et al. (2015, 2017) report colocation dates relying primarily on (3) while at the same time referring to (1) and (2). By contrast, Boehmer et al. (2021) report information for the same exchanges based on (1) and (2), but not (3), at least based on what we have learned from our correspondence with the JFQA handing editor that published Boehmer et al. (2021) (see the Online Appendix, Exhibit 1).[3]

Table 2 indicates some notable differences in the colocation dates reported by Aitken et al. (2015, 2017). These dates differ on average from the colocation dates reported in Boehmer et al. (2021) by an average of 12.75 months, ranging from 1 month (for Australia) to 46 months (for Switzerland).

What explains these differences? One explanation is that they are simply errors. We believe that explanation fits the Canadian exchange, as we document in full in the Online Appendix, Exhibit 1. In the case of Canada (where colocation started in April 2008, not November 2008, as Boehmer et al. 2021 report), Boehmer et al. (2021) referenced an incorrect source in the main body of their paper. Subsequently (after their paper was accepted at JFQA but before it was published in print), another source was added to their Online Appendix, but unfortunately, that additional source is incorrect as well. See Exhibit 1 in the Online Appendix to this paper.

We now take a closer look at some discrepancies for the dates other than the 7-month difference for Canada between Aitken et al. (2015, 2017) and Boehmer et al. (2021). Perhaps the second-best place to look is Switzerland, since the discrepancy is the largest at 46 months (June 2008 for Boehmer et al. and April 2012 for Aitken et al.). The Online Appendix Exhibit 2 of this paper presents our email evidence from the Swiss Exchange documenting the introduction of the Swiss colocation service, and an online news announcement that is still posted[4] on the Swiss 'Exchange's news feed. By contrast, Boehmer et al. provide a reference in their Online Appendix to the news source "Colt". The Colt news source states that the Swiss Exchange and Colt are teaming up to provide members of the Swiss Exchange' proximity 'services'. It does not say when they will provide these proximity services, or if they have reached a formal agreement, or if they have already started providing these proximity services. Irrespective, what we do know from our direct correspondence with the Swiss Exchange (Online Appendix Exhibit 2) is that the earliest colocation date on

[3] There can be a difference between the date from which colocation services are available (e.g., for testing) and the date from which colocation services are used for active trading (e.g., in production). Both Boehmer et al. (2021) and Aitken et al. (2015, 2017) attempt to focus on dates for colocation used for active trading. Also, the definition of 'colocation' might differ across trading venues, can change with technological upgrades, and can have some subjectivity. We thank an anonymous reviewer for these helpful comments.

[4] This information was still available online at the time of writing this paper, May 20, 2021, and hence available to Boehmer et al. (2021).

Table 1
The Conflicting Published Evidence on Colocation Dates in Academic Journals
This table summarizes colocation dates published in Aitken et al. (2015, 2017) and Boehmer et al. (2021) for the complete set of exchanges reported by these authors. Sources: Boehmer et al. (2021) Online Appendix; Aiken et al. (2015) Table A4, page 354.

	Boehmer et al. (2021)		Aitken et al. (2015, 2017)	
Country	Co-location Date	Source	Co-location Date	Source (In Addition to Communications with Exchanges)
Australia	2008/11	2008 Annual Report	2008 Q4	ASX Group News Announcement
Belgium	2008/04	Info-Flash Euronext Cash Market		
Brazil	2009/06	Financial Industry Network Thinking		
Canada	2008/11	2008 Annual Report	2008/04	Information provided by TMX Datalinx [see the Online Appendix Exhibit 1 of this paper]
Denmark	2008/06	Financial Times		
Finland	2008/06	Financial Times		
France	2008/04	Info-Flash Euronext Cash Market		
Hong Kong			2012 Q4	Hong Kong Exchanges News
Germany	2006	Financial Times	2006/08	Information provided by XETRA Support [see the Online Appendix Exhibit 3 of this paper]
India BSE	2010/11	Business Standard	2010/02	The World Federation of Exchanges News
India NSE	2009/08	Business Standard	2010/01	NSE News [see the Online Appendix Exhibit 6 of this paper]
Italy	2009/09	Borsa Italiana		
Japan Osaka	2008/11	Osaka Securities Exchange		
Japan Tokyo	2009/05	Tokyo Stock Exchange (TSE) News	2010/01	TSE News
Norway			2010/04	OSLO News
Netherlands	2008/04	Info-Flash Euronext Cash Market		
Portugal	2008/04	Info-Flash Euronext Cash Market		
Singapore	2011/04	The Trade	2011/07	SGX News [see the Online Appendix 4 of this paper]
Sweden	2008/06	Financial Times	2011/03	AlipesNews
Switzerland	2008/06	Colt, provider of co-location service	2012/04	Six-Swiss Exchange News [see the Online Appendix Exhibit 2 of this paper]
Taiwan	2010/Q4	Asia ETrading	2010 Q4	The World Federation of Exchanges News
UK London	2008/09	Wall street tech	2009/09	LSE News
UK ChiX			2008/11	Information provided by ChiX Support
USA Nasdaq	2005/04	Wall street tech	2007/03	NASDAQ News
USA NYSE	2007/01	NYSE	2008/04	NYSE News

Table 2
Colocation and HFT Dates
This table summarizes colocation dates provided in Aitken et al. (2015, 2017) and Boehmer et al. (2021). Also, the table summarizes dates for proxies for high frequency trading (HFT) estimated by Aitken et al. (2015) based on structural breaks in trade size and order/cancellation ratios. The HFT evidence normally pre-dates colocation dates because colocation is an endogenous service offered by exchanges in response to market developments and market 'participants' interests in seeing those services, so that the exchange can make more money. The sample period is January 2003 to December 2011.

Exchange Name	Average Market Cap.Weighted Percentage Spread	HFT Trade Size (Aitken et al., 2015)	HFT Order Cancellation Ratio (Aitken et al., 2015)	Colocation (Aitken et al., 2015, 2017)	Colocation (Boehmer et al., 2021)	Number of Months from Colocation Boehmer et al. to Colocation Aitken et al.	Number of Months from HFT Trade Size to Colocation Aitken et al. (2015, 2017)
Australia	0.8377	2006/04	2006/06	2008/10	2008/11	-1	30
Bombay	0.3763	2009/05	2009/06	2010/02	2009/08	6	9
London	0.5510	2006/02	pre-2003	2009/09	2008/09	12	43
NASDAQ	0.3287	pre-2003	pre-2003	2007/03	2005/04	23	
NSE India	0.2004	2009/05	2009/05	2010/01	2009/08	5	8
NYSE	0.1118	2003/05	2003/07	2008/04	2007/01	15	59
Singapore	1.6359	N/A	N/A	2011/07	2011/04	3	
Stockholm	0.4443	2005/04	2005/02	2011/03	2008/06	33	71
Switzerland	1.6420	2004/01	pre-2003	2012/04	2008/06	46	99
Taiwan	0.5965	N/A	N/A	2010/10	2010/10	0	
Toronto (Canada)	0.6421	2005/05	2004/01	2008/04	2008/11	−7	35
XETRA (Germany)	0.2344	2003/01	2003/02	2006/08	2006/10	−2	43

the Swiss Exchange is April 2012 as reported in Aitken et al. (2015, 2017), and not June 2008 as reported by Boehmer et al. (2021).[5]

The 46-month discrepancy in Switzerland is, of course, non-trivial and the largest reported in Table 2 between Aitken et al. (2015, 2017) and Boehmer et al. (2021). There are other differences in the other exchanges, with the average/median difference being 12.75/6 months. In the Online Appendix Exhibits 3 and 4, respectively, we provide direct evidence for these exchanges of their colocation dates, at least insofar as the exchange officers sent it to us and with documentation. In Exhibit 3 in the Online Appendix, we provide evidence that shows Boehmer et al. (2021) provide a source for Germany that does not support the date that they use: the source states "mid-2006″ and Boehmer et al. (2021) use October 2006. In the Online Appendix in Exhibit 3, we show evidence from the exchange the colocation date for Germany is August 2006.[6] In the case of Singapore, it appears that the difference is likely due to using an unreliable news source or blog post, as verified by the exchange; please see Exhibit 4 in the Online Appendix.

While the "clear error" explanation appears to apply to Canada, Switzerland, Germany, and Singapore, of the remaining exchanges, other interpretations are possible. Aitken et al. asked the officers of the exchanges to identify the first date when they offered colocation services. See the replies in the Online Appendix Exhibits 1, 2, 3, and 4 for Canada, Switzerland, Germany, and Singapore as representative examples of replies from exchanges from different parts of the world. The exchanges replied either with a single date, with some referring to sources for us to link to. In Aitken et al., where an extra source was provided, we referenced that source (see, e.g., Aitken et al. 2015, Table A4, at page 354). If we only received a reply without an extra reference source, Aitken et al. (2015) reported that the exchange provided the information.

At the time of preparing Aitken et al. (2015, 2017) we were not aware of any other evidence that the reported colocation dates were different than those used by others. Indeed, at the time of preparing the working papers of Aitken et al. (2015, 2017), we became aware that Boehmer et al. (2021) were using colocation dates in their working paper and asked the authors to compare dates. Boehmer et al. however, chose not to share their dates or post them publicly (they did so sometime after Aitken et al., 2015, 2017, were in print). On the other hand, Aitken et al. (2015, 2017) shared colocation dates and posted them publicly in working paper drafts a few years prior to the published versions. As such, if there is a mistake in Aitken et al. (2015, 2017), we acknowledge we are responsible, but we certainly did not know about it prior to publication and took all reasonable steps to ensure we were using the correct dates.

For other exchanges there is information online that conflicts with both the dates in Boehmer et al. (2021) and Aitken et al. (2015, 2017). For example, a London Stock Exchange service announcement provided by one of our reviewers suggests that colocation services became available on October 6, 2008, which corresponds to neither Aitken et al. (2015, 2017) or Boehmer et al. (2021) summarized in Table 1. This service announcement conflicts with the press release from the London Stock Exchange used in Aitken et al. (2015, 2017) (which we provide in Exhibit 5 of this paper in our Online Appendix), and the "Wall Street Tech" media news announcement used in Boehmer et al. (2021).[7] We believe the official news announcement from the London Stock Exchange that we reproduce in Exhibit 5 is the most reliable source, since it is an official media communication from the London Stock Exchange. There is similar conflicting evidence from India,[8] but given our source for India we believe our information from India is correct; see the Online Appendix Exhibit 6.

In 2012, Aitken et al. (2015, 2017) emailed the officers of all the exchanges reported in Aitken et al. (2015, 2017) to ask for colocation dates and did not receive any information that indicated Aitken et al.'s dates were incorrect. In 2022, we tried to repeat this exercise of emailing all of the exchanges, including the ones not in Aitken et al. (2015, 2017), but had scant responses and interest from the exchanges as those dates are significantly more remote.

Boehmer et al. consider more exchanges than those considered by Aitken et al. (2015, 2017). As such, Aitken et al. (2015, 2017) did not ask the other exchange officers to confirm the dates because they did not include those exchanges in their study (and likewise Boehmer et al. appear to have not done so either; see Online Appendix Exhibit 1). With differences in colocation dates like those reported in Switzerland and Canada, etc., in Boehmer et al. (2021), there is an opportunity to see if the extent of the magnitude of the errors are systematic or random. Boehmer et al. (2021, at p.16 of the working paper draft) claim the errors are likely random, implying that errors are not going to materially bias the findings of their analysis. So, the presentation of these things in Boehmer et al. (2021) offers an excellent opportunity to subject this information to empirical scrutiny. That is, the colocation differences in Boehmer et al. (2021) have generated a new research opportunity. The remainder of this paper takes a first look at the consequences of these differences.

3. Data

We report data here that was used in part in Aitken et al. (2015, 2017). The sample covers 12 exchanges (Australia, Bombay, London, NASDAQ, NSE India, NYSE, Singapore, Stockholm, Switzerland, Taiwan, Toronto, and XETRA (Germany)) for the time period January 2003 to December 2011. To keep the analysis here succinct, we examine one dependant variable: the market capitalization-weighted average percentage spread. We aim for succinctness here because our objective is to develop a single counterfactual to show that errors in colocation dates may matter, and they are not necessarily

[5] A helpful reviewer showed us that additional details of the Swiss Exchange colocation introduction are available here: https://www.six-group.com/dam/download/market-data/news/swiss-exchangemessages/2012/sse_message_201204031652_en.pdf.

[6] A helpful reviewer has pointed out that the Interim Report of the Deutsche Börse Group (Quarter 3/2006) states that colocation services ("Proximity Services") became available in September 2006, which is in-between the dates of Aitken, et al. (2015, 2017) and Boehmer, et al. (2021). September also appears in the Annual Report. https://www.deutscheboerse.com/resource/blob/34056/efd912d8c125190224e67cfc1da3f2e6/data/gdb-quartalsbericht-q3-2006_en.pdf. Our sources from Exhibit 3 are different. We cannot confirm date is more appropriate given these conflicting sources but we feel the preciseness of the information in Exhibit 3 from the officer of the exchange is sufficiently compelling to not ignore.

[7] We owe thanks to a helpful reviewer for pointing out these discrepancies and other sources. Also, in 2022 we contacted each of the exchanges to re-verify all colocation dates while working on the revise and resubmit. We received feedback about different colocation dates for NASDAQ and NYSE compared to what we received in 2012 (as reported in Aitken et al., 2015) and nothing that suggested anything different for the other exchanges. Due to the proximity of the timing of the feedback to the event, we left those dates as we reported Aitken et al. (2015). Regardless, our HFT start dates for NYSE and NASDAQ are both significantly prior to the colocation dates we report in Aitken et al. (2015) and prior to those reported in Boehmer et al. (2021); relatedly, these specific dates for these 2 exchanges do not impact the inferences drawn herein.

[8] A helpful reviewer has pointed out that in the case of the NSE in India, there was a SEBI case involving colocation connections. (See April 30, 2019 Orders of Chairman, https://www.sebi.gov.in/enforcement/orders/apr-2019/order-in-the-matter-of-nsecolocation_42880.html). The legal documents provide a lot of information, including: Paragraph 7.1.3 refers to August 2009, which corresponds to Boehmer et al. (2021). Paragraph 7.1.4 refers to December 2009, which is a month earlier than Aiken et al. (2015, 2017). But, in the paragraph 7.1.5, it appears that the TBT functionality for cash markets did not occur until July 2010. Given these technical details, it is unclear what date should be used.

Table 3
Comparison of Means Tests
This table provides a comparison of means of the average market capitalization weighted percentage spread on the exchanges listed in Table 2 for January 2003 to December 2011. The statistics are provided prior to, and after, HFT or colocation based on alternative definitions of the variables. *, **, *** significant at the 10%, 5%, and 1% levels, respectively. The data indicate that after the change, for any given variable documenting a change, there was a statistically significant reduction in the average market capitalization weighted percentage spread.

	HFT Trade Size (Aitken et al., 2015)	HFT Order Cancellation Ratio (Aitken et al., 2015)	Colocation (Aitken et al., 2015, 2017)	Colocation (Boehmer et al., 2021)
Market Cap Weighted Percentage Spread Prior to...	0.8270	0.7620	0.7588	0.7050
Market Cap Weighted Percentage Spread After...	0.5245	0.5944	0.2986	0.5398
Difference	0.3025	0.1676	0.4603	0.1651
Comparison of Means z-statistic	4.86***	2.31**	9.99***	2.72***

Table 4
Correlation Matrix
This table provides a correlation matrix for the variables in the data. All of the correlations are statistically significant at least the 1% level.

	[1]	[2]	[3]	[4]	[5]
[1] Market Capitalization Weighted Percentage Spread	1				
[2] Colocation (Boehmer et al., 2021)	-0.0777	1			
[3] Colocation (Aitken et al., 2015, 2017)	-0.195	0.7409	1		
[4] HFT Trade Size (Aitken et al., 2015)	-0.1479	0.5569	0.433	1	
[5] HFT Order Cancellation Ratio (Aitken et al., 2015)	-0.078	0.4646	0.3521	0.8097	1

random, so academics must always strive to obtain factual information from primary/multiple sources.

Table 2 highlights information for four possible explanatory variables: (1) colocation as defined by Boehmer et al. (2021) dates, colocation as defined by Aitken et al. (2015, 2017) dates, and two alternative dates for high frequency trading (HFT). HFT dates were proxied by Aitken et al. (2015) based on structural breaks in trade size and order/cancellation ratios. Structural breaks in trade size and order cancellation ratio serve as useful proxies for alternative dates for HFT because HFTs involve significantly lower trade sizes and significantly higher order cancellations than regular traders (Aitken et al., 2015, 2017). Those events are indicative of a notable increase in the presence of HFTs in the marketplace since HFT can be characterized by trade size and order cancellation ratios. Unfortunately, there are no known formal dates as to when HFT started in a particular market, and those dates are offered as proxies.

The summary of the data is presented in Table 2. Comparison tests are presented in Tables 3 and 4 provides a correlation matrix. The comparison tests and correlation matrix show exactly what we would expect: spreads have gone down over time, and for whichever date is chosen as having a possible material influence on spreads over 2003–2011, it is possible that the date shows up as statistically significant.[9]

Also, the correlations in Table 4 indicate that the explanatory variables are highly correlated; for example, the correlation between the Aitken et al. (2015, 2017) colocation variable and the Boehmer et al. (2021) colocation variable is 0.7401. So, even with mistakes in dates, it is possible that impacts are similar, and hence the mistakes in Boehmer et al. are not material for the purpose of examining the impact of colocation on liquidity. But it is of course worth examining further in a regression setting, as we do in the next section.

4. Regressions

4.1. Determinants of spreads

Table 5 presents panel data regressions for the determinants of the average monthly market capitalization-weighted spread across the 12 exchanges listed in Table 2 for the period January 2003 to December 2011. The regressions use exchange fixed effects and standard errors clustered at the exchange level. The data examined indicates that HFT proxied by trade size, order/cancellation, and colocation with Aitken et al. (2015, 2017) colocation dates all significantly negatively affect spreads.

The only robust variable that is statistically significant in all of the specifications is the HFT order/cancellation variable, which is significant at the 1% level in models [2] and [6], and significant at the 5% level in models [7] and [8]. The economic significance of the HFT order/cancellation variable is largest in Model [7] whereby HFT gives rise to a 47.8% drop in spreads relative to the average exchange-month spread in the data. By contrast, in Model [6], for example, colocation defined with Aitken et al. (2015, 2017) dates reduces spreads by 20.1% relative to the average exchange-month level in the data.

Notably, colocation with Boehmer et al. (2021) dates are insignificantly related to spreads in any of the models in Table 5. One explanation for the insignificance of the Boehmer et al. (2021) colocation dates variable is that those colocation dates have errors, as indicated above in Section 2 and the Online Appendix. Another explanation is that colocation is endogenous (Aitken et al., 2017). Either way, the results cast doubt on the idea that errors in colocation dates do not matter for liquidity, as claimed by Boehmer et al. (2021).[10]

[9] For other similar research on topic, see Cumming et al. (2011) and Cumming and Johan (2019) on the use of regulatory changes to explain liquidity changes in different countries.

[10] We further reran the regressions with the subset of data from the 4 exchanges reported in the Online Appendix. The data are consistent with that reported in Table 5. There is no specification for which the Boehmer et al. (2021) colocation dates matter for liquidity, but there is support for the relevance of the Aitken et al. (2015, 2017) colocation dates and the Aitken et al. (2015, 2017) HFT dates, similar to that reported in Table 5. Those results are available on request.

Table 5
Panel Regression for the Determinants of Average Exchange Market Capitalization Weighted Spread
This table presents panel model estimates of the determinants of the average exchange market capitalization weighted spread for the exchanges listed in Table 2. Fixed effects models are used in all specifications, and standard errors are clustered at the exchange level. *, **, *** significant at the 10%, 5%, and 1% levels, respectively.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
HFT Trade Size (Aitken et al., 2015)	-0.1511				-0.2314			-0.0523
	(-1.84)*				(-2.13)**			(-0.76)
HFT Order Cancellation Ratio (Aitken et al., 2015)		-0.2364				-0.1871	-0.301	-0.2523
		(-3.87)***				(-2.67)***	(-2.20)**	(-2.29)**
Colocation (Aitken et al., 2015, 2017)			-0.1776			-0.1275		-0.4657
			(-3.23)***			(-2.24)**		(-1.46)
Colocation (Boehmer et al., 2021)				0.0514	0.1354		0.1393	0.4453
				(0.24)	(0.56)		(0.058)	(1.11)
Constant	0.7379	0.8084	0.6713	0.6315	0.7392	0.8056	0.8108	0.81
	(15.28)***	(19.62)***	(52.59)***	(8.74)***	(15.62)***	(18.85)***	(0.82)***	(26.04)***
Number of Observations	1224	1224	1224	1224	1224	1224	1224	1224
R2-within	0.0034	0.0073	0.006	0.0006	0.0069	0.0101	0.0114	0.0303
R2-between	0.1005	0.0055	0.3907	0.2516	0.019	0.063	0.0036	0.0631
R2-overall	0.0219	0.0061	0.038	0.006	0.0157	0.0217	0.0023	0.0377
F	3.39*	14.99***	10.41***	0.06	4.30**	11.29***	8.64***	8.86***

Table 6
Panel Cox Duration Estimates of Time to Colocation
This table reports hazard ratios for the time to colocation using panel Cox duration models of time to Boehmer et al. (2021) colocation date, and time to Aitken et al. (2015, 2017) colocation dates. Hazard ratios are presented, meaning that for example in Model [1] the time to Boehmer et al.'s's colocation date is 7.75 times faster after HFT has significant presence in the market. z-statistics in parentheses. *, **, *** significant at the 10%, 5%, and 1% levels, respectively.

	Time to Boehmer et al. (2021) Co-location Date		Time to Aitken et al. (2015, 2017) Co-location Date			
	[1]	[2]	[3]	[4]	[5]	[6]
HFT Trade Size (Aitken et al., 2015)	7.7508		7.0166		5.7442	
	(6.99)***		(5.75)***		(4.54)***	
HFT Order Cancellation Ratio (Aitken et al., 2015)		5.4215		4.5265		2.3638
		(5.01)***		(4.62)***		(2.48)***
Colocation (Boehmer et al., 2021)					1.4665	3.5631
					(1.01)	(3.81)***
Number of Observations	1224	1224	1224	1224	1224	1224
LR Chi2	98.34***	45.05***	64.06***	33.73***	65.15***	41.70***

In Table 5 Model (8), we show that when you control for the presence of high-frequency traders, there is no impact of colocation on market quality regardless of the measure of colocation, which again is counter to the claims of Boehmer et al. (2021). High-frequency traders brought about improvements in market quality long before colocation services were endogenously offered, as indicated in Model (8). In the Section 4.2 below, we further show that colocation itself can be predicted by the presence of high-frequency traders. That is, colocation itself is an invalid endogenous instrument.

4.2. Determinants time to colocation dates

Boehmer et al. (2021) indicate on page 16 (working paper version) that "[*a*]s a potential caveat, the precision of reporting on colocation event dates could differ across countries. Yet, to the extent that the resulting errors are random,...". Here, we examine, indirectly, whether or not these errors are random or if the dates provided in Boehmer et al. (2021) are predictable. The regressions used for this purpose are reported in Table 6. Table 6 reports panel Cox regressions on the time to colocation as defined by Aitken et al. (2015, 2017) and time to colocation as defined by Boehmer et al. (2021).

The results in Table 6 models [1] and [2] show evidence that time to Boehmer et al. colocation is predictable. That is, when HFT (trade size definition) starts on an exchange, the hazard ratio is 7.75 and significant at the 1% level (model [1]), and when HFT (order/cancellation) starts on an exchange, the hazard ratio is 5.42 and significant at the 1% level (model [2]). So, despite the errors in Boehmer et al. (2021) colocation dates, we can still predict them, suggesting they are not perfectly random.

An equally interesting question is whether or not we can use Boehmer et al. (2021) colocation dates to predict actual colocation dates (that is, colocation dates verified by the exchanges as discussed above and in the Online Appendix). Models [5] and [6] show mixed support: Boehmer et al. (2021) colocation dates cannot predict Aitken et al. (2015, 2017) colocation dates in model [5], but can predict them in model [6]. By contrast, the HFT variables show much more robust evidence in models [5] and [6], meaning that once HFTs have a significant presence in the marketplace, the exchange endogenously offers colocation services in response to that new market development as a way to make money, as predicted by Aitken et al. (2017). The hazard ratios are 2.36–7.02 for HFT in models [3]–[6] and all significant at the 1% level, consistent with Aitken et al. (2017).

In sum, whether the errors in Boehmer et al. (2021) are random or not is a question that we cannot fully answer using the data in this paper. However, the data examined here points to a hypothesis that the errors are predictable, and hence they are unlikely to be random. But regardless, the question as to exactly how random Boehmer et al.'s errors is irrelevant, since the magnitude of the errors on colocation dates is likely to render any study based on them, unreliable.

5. Conclusions

This paper explained that colocation dates around the world are potentially important for explaining liquidity (Boehmer et al., 2021), or they are potentially endogenous responses to HFTs en-

tering a market and an exchange seeking a new revenue source (Aitken et al., 2017), or both (Aitken et al., 2015, 2017). This paper explained that there are differences in colocation dates reported in different studies due to the fact that colocation dates can be obtained from primary and secondary sources. A recent study by Boehmer et al. (2021) using secondary sources reports colocation dates that our primary sources suggest are materially incorrect, and likely have a material impact on inferences that can be drawn. We show direct evidence from the exchanges that the errors average 12.75 months, with the largest being 46 months. We also provide evidence that these differences in colocation dates with the exchange-provided dates almost always matter statistical inference in connection with liquidity.

The evidence in this paper shows:

(1) The colocation dates presented by Boehmer et al. (2021) are incorrect, with the errors averaging 12.75 months and the largest error is 46 months from the correct colocation dates.
(2) Using corrected colocation dates appear to show that colocation can give rise to improvements in market quality, as shown in Boehmer et al. (2021).
(3) When you control for the presence of high-frequency traders prior to the introduction of colocation, there is no impact of colocation on market quality.
(4) Colocation itself is not a valid instrument for predicting market quality, because it is an endogenous service provided to market participants after high-frequency traders enter the market.
(5) The time to offering colocation is predictable based on the presence of high-frequency traders entering the market.

In writing this paper our hope is that future scholars will look more closely at alternative sources of information, in particular the use of primary and secondary sources. Further, in line with the doctrine of commensurability, authors would either take issue with factual differences in previously published works on the topic or at least identify that there are differences. Without doing so, readers could be misled, which could negatively affect subsequent scholarship that rely on those facts.

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jbankfin.2023.106843.

References

Aitken, M., Cumming, D.J., Zhan, F., 2015. High frequency trading and end-of-day price dislocation. J. Bank. Financ. 59, 330–349.
Aitken, M., Cumming, D., Zhan, F., 2017. Trade size, high frequency trading, and colocation around the world. Eur. J. Financ. 23 (7–9), 781–801.
Allen, F., Gu, X., Jagtiani, J., 2021. A survey of fintech research and policy discussion. Review of Corporate Finance 1 (3–4), 259–339.
Boehmer, E., Fong, K., Wu, J., 2021. Algorithmic trading and market quality: international evidence. J. Financ. Quant. Anal. forthcoming.
Cumming, D.J., Johan, S., 2019. Capital-market effects of securities regulation: prior conditions, implementation, and enforcement revisited. Financ. Res. Lett. 31, 425–435.
Cumming, D.J., Johan, S.A., Li, D., 2011. Exchange trading rules and stock market liquidity. J. Financ. Econ. 99 (3), 651–671.
Menkveld AJ. and Dreber A. and Holzmeister F. and Huber, Juergen and Johanneson, Magnus and Kirchler, Michael and Razen, Michael and Weitzel, Utz and Abad, David and Abudy, Menachem (Meni) and Adrian, Tobias and Ait-Sahalia, Yacine and Akmansoy, Olivier and Alcock, Jamie and Alexeev, Vitali and Aloosh, Arash and Amato, Livia and Amaya, Diego and Angel, James J. and Bach, Amadeus and Baidoo, Edwin and Bakalli, Gaetan and Barbon, Andrea and Bashchenko, Oksana and Bindra, Parampreet Christopher and Bjonnes, Geir Hoidal and Black, Jeffrey R. and Black, Bernard S. and Bohorquez, Santiago and Bondarenko, Oleg and Bos, Charles S. and Bosch-Rosa, Ciril and Bouri, Elie and Brownlees, Christian T. and Calamia, Anna and Cao, Viet Nga and Capelle-Blancard, Gunther and Capera, Laura and Caporin, Massimiliano and Carrion, Allen and Caskurlu, Tolga and Chakrabarty, Bidisha and Chernov, Mikhail and Cheung, William Ming Yan and Chincarini, Ludwig B. and Chordia, Tarun and Chow, Sheung Chi and Clapham, Benjamin and Colliard, Jean-Edouard and Comerton-Forde, Carole and Curran, Edward and Dao, Thong and Dare, Wale and Davies, Ryan J. and De Blasis, Riccardo and De Nard, Gianluca and Declerck, Fany and Deev, Oleg and Degryse, Hans and Deku, Solomon and Desagre, Christophe and Van Dijk, Mathijs A. and Dim, Chukwuma and Dimpfl, Thomas and Dong, Yun Jiang and Drummond, Philip and Dudda, Tom L. and Dumitrescu, Ariadna and Dyakov, Teodor and Dyhrberg, Anne Haubo and Dzieliński, Michał and Eksi, Asli and El Kalak, Izidin and ter Ellen, Saskia and Eugster, Nicolas and Evans, Martin D.D. and Farrell, Michael and Félez-Viñas, Ester and Ferrara, Gerardo and FERROUHI, El Mehdi and Flori, Andrea and Fluharty-Jaidee, Jonathan and Foley, Sean and Fong, Kingsley Y. L. and Foucault, Thierry and Franus, Tatiana and Franzoni, Francesco A. and Frijns, Bart and Frömmel, Michael and Fu, Servanna Mianjun and Füllbrunn, Sascha and Gan, Baoqing and Gehrig, Thomas and Gerritsen, Dirk and Gil-Bazo, Javier and Glosten, Lawrence R. and Gomez, Thomas and Gorbenko, Arseny and Güçbilmez, Ufuk and Grammig, Joachim and Gregoire, Vincent and Hagströmer, Björn and Hambuckers, Julien and Hapnes, Erik and Harris, Jeffrey H. and Harris, Lawrence and Hartmann, Simon and Hasse, Jean-Baptiste and Hautsch, Nikolaus and He, Xue-Zhong 'Tony' and Heath, Davidson and Hediger, Simon and Hendershott, Terrence J. and Hibbert, Ann Marie and Hjalmarsson, Erik and Hoelscher, Seth and Hoffmann, Peter and Holden, Craig W. and Horenstein, Alex R. and Huang, Wenqian and Huang, Da and Hurlin, Christophe and Ivashchenko, Alexey and Iyer, Subramanian R. and Jahanshahloo, Hossein and Jalkh, Naji and Jones, Charles M. and Jurkatis, Simon and Jylha, Petri and Kaeck, Andreas and Kaiser, Gabriel and Karam, Arzé and Karmaziene, Egle and Kassner, Bernhard and Kaustia, Markku and Kazak, Ekaterina and Kearney, Fearghal and van Kervel, Vincent and Khan, Saad and Khomyn, Marta and Klein, Tony and Klein, Olga and Klos, Alexander and Koetter, Michael and Krahnen, Jan Pieter and Kolokolov, Aleksey and Korajczyk, Robert A. and Kozhan, Roman and Kwan, Amy and Lajaunie, Quentin and Lam, Full Yet Eric Campbell and Lambert, Marie and Langlois, Hugues and Lausen, Jens and Lauter, Tobias and Leippold, Markus and Levin, Vladimir and Li, Yijie and Li, (Michael) Hui and Liew, Chee Yoong and Lindner, Thomas and Linton, Oliver B. and Liu, Jiacheng and Liu, Anqi and Llorente-Alvarez, Jesus-Guillermo and Lof, Matthijs and Lohr, Ariel and Longstaff, Francis A. and Lopez-Lira, Alejandro and Mankad, Shawn and Mano, Nicola and Marchal, Alexis and Martineau, Charles and Mazzola, Francesco and Meloso, Debrah C and Mihet, Roxana and Mohan, Vijay and Moinas, Sophie and Moore, David and Mu, Liangyi and Muravyev, Dmitriy and Murphy, Dermot and Neszveda, Gabor and Neumeier, Christian and Nielsson, Ulf and Nimalendran, Mahendrarajah and Nolte, Sven and Nordén, Lars L. and O'Neill, Peter and Obaid, Khaled and Ødegaard, Bernt Arne and Östberg, Per and Painter, Marcus and Palan, Stefan and Palit, Imon and Park, Andreas and Pascual Gascó, Roberto and Pasquariello, Paolo and Pastor, Lubos and Patel, Vinay and Patton, Andrew J. and Pearson, Neil D. and Pelizzon, Loriana and Pelster, Matthias and Pérignon, Christophe and Pfiffer, Cameron and Philip, Richard and Plíhal, Tomáš and Prakash, Puneet and Press, Oliver-Alexander and Prodromou, Tina and Putnins, Talis J. and Raizada, Gaurav and Rakowski, David A. and Ranaldo, Angelo and Regis, Luca and Reitz, Stefan and Renault, Thomas and Wang, Renjie and Renò, Roberto and Riddiough, Steven and Rinne, Kalle and Rintamäki, Paul and Riordan, Ryan and Rittmannsberger, Thomas and Rodríguez Longarela, Iñaki and Rösch, Dominik and Rognone, Lavinia and Roseman, Brian and Rosu, Ioanid and Roy, Saurabh and Rudolf, Nicolas and Rush, Stephen and Rzayev, Khaladdin and Rzeźnik, Aleksandra and Sanford, Anthony and Sankaran, Harikumar and Sarkar, Asani and Sarno, Lucio and Scaillet, Olivier and Scharnowski, Stefan and Schenk-Hoppé, Klaus Reiner and Schertler, Andrea and Schneider, Michael and Schroeder, Florian and Schuerhoff, Norman and Schuster, Philipp and Schwarz, Marco A. and Seasholes, Mark S. and Seeger, Norman and Shachar, Or and Shkilko, Andriy and Shui, Jessica and Sikic, Mario and Simion, Giorgia and Smales, Lee A. and Söderlind, Paul and Sojli, Elvira and Sokolov, Konstantin and Spokeviciute, Laima and Stefanova, Denitsa and Subrahmanyam, Marti G. and Neusüss, Sebastian and Szaszi, Barnabas and Talavera, Oleksandr and Tang, Yuehua and Taylor, Nicholas and Tham, Wing Wah and Theissen, Erik and Thimme, Julian and Tonks, Ian and Tran, Hai and Trapin, Luca and Trolle, Anders B. and Valente, Giorgio and Van Ness, Robert A. and Vasquez, Aurelio and Verousis, Thanos and Verwijmeren, Patrick and Vilhelmsson, Anders and Vilkov, Grigory and Vladimirov, Vladimir and Vogel, Sebastian and Voigt, Stefan and Wagner, Wolf and Walther, Thomas and Weiss, Patrick and van der Wel, Michel and Werner, Ingrid M. and Westerholm, P. Joakim and Westheide, Christian and Wipplinger, Evert and Wolf, Michael and Wolff, Christian C. P. and Wolk, Leonard and Wong, Wing-Keung and Wrampelmeyer, Jan and Xia, Shuo and Xiu, Dacheng and Xu, Ke and Xu, Caihong and Yadav, Pradeep K. and Yagüe, José and Yan, Cheng and Yang, Antti and Yoo, Woongsun and Yu, Wenjia and Yu, Shihao and Yueshen, Bart Zhou and Yuferova, Darya and Zamojski, Marcin and Zareei, Abalfazl and Zeisberger, Stefan and Zhang, Sarah and Zhang, Xiaoyu and Zhong, Zhuo and Zhou, Z. Ivy and Zhou, Chen and Zhu, Xingyu and Zoican, Marius and Zwinkels, Remco C.J. and Chen, Jian and Duevski, Teodor and Gao, Ge and Gemayel, Roland and Gilder, Dudley and Kuhle, Paul and Pagnotta, Emiliano and Pelli, Michele and Sönksen, Jantje and Zhang, Lu and Ilczuk, Konrad and Bogoev, Dimitar and Qian, Ya and Wika, Hans C. and Yu, Yihe and Zhao, Lu and Mi, Michael and Bao, Li and Vaduva, Andreea and Prokopczuk, Marcel and Avetikian, Alejandro and Wu, Zhen-Xing, Non-standard errors (November 16, 2022). Available at SSRN: https://ssrn.com/abstract=3961574 or 10.2139/ssrn.3961574