THE RISE OF FAST TRADING: CURSE OR BLESSING FOR LIQUIDITY?

Christophe Desagre, Catherine D'Hondt, Mikael Petitjean

Association Française de Finance | « Finance »

2022/3 Vol. 43 | pages 119 à 158

ISSN 0752-6180
DOI 10.3917/fina.pr.i

The Rise of Fast Trading: Curse or Blessing for Liquidity?

Christophe Desagre*, Catherine D'Hondt†, and Mikael Petitjean‡

Abstract

We study how market liquidity on Euronext has evolved with the rise of fast trading. We identify fast traders by directly measuring message traffic and the lifetime of orders for every individual market member on Euronext using their identification codes. We observe an overall improvement in terms of liquidity between 2002 and 2006. However, the most exposed stocks to fast trading exhibit the weakest increase in liquidity and lose the liquidity advantage they had before the rise of fast trading.

Keywords: Liquidity, Fast Tading, Euronext, Market Members

1. Introduction

This paper was motivated some years ago by Chordia et al. (2013)'s editorial in which they wrote that *'the question of whether financial markets before the advent of High-frequency trading (HFT) were better or worse than today's HFT-dominated markets remains unanswered'* (p. 639). Our motivation was further reinforced by Kirilenko and Lo (2013) who commented that '*a deeper understanding of the historical roots of algorithmic trading is especially important for predicting where it is headed and formulating policy and regulatory recommendations that affect it*' (p. 53).

* Louvain School of Management (UCLouvain, LFIN/LIDAM)
ICHEC Brussels Management School
Email: christophe.desagre@uclouvain.be

† Louvain School of Management (UCLouvain, LFIN/LIDAM)
Email: catherine.dhondt@uclouvain.be

‡ IESEG School of Management, Univ. Lille, CNRS, UMR 9221 – LEM – Lille Economie Management
Louvain School of Management (UCLouvain, LFIN/LIDAM)
E-mail: m.petitjean@ieseg.fr.

We are grateful to Euronext for providing us with the data. We would like to thank Jonathan Brogaard, Rudy De Winne, Corey Garriott, Carole Gresse, Paolo Mazza, Albert Menkveld, Sophie Moinas, Maurizio Montone, Matthias Saerens, Nicolas Scholtes, Gunther Wuyts, Yue Zhang, and the many participants at several conferences and seminars in business schools and universities. We apologize for not mentioning all the names of people who helped us improve the paper. The paper also benefited from the rigorous comments and suggestions from two reviewers, the editor and the coeditor. Christophe Desagre gratefully acknowledges financial support from the LFIN research center. Any errors are the full responsibility of the authors.

The historical perspective that we adopt in this paper is therefore timely since its goal is to study how the rise of fast trading (FT hereafter) has affected liquidity on Euronext in the 2000s.[1] We take advantage of a unique database to run a differences-in-difference (DID) study and compare the evolution of liquidity between a first group of CAC40 stocks most exposed to FT and a second group of all the remaining CAC40 stocks less exposed to it.

Although market liquidity on Euronext has very much improved in the 2000s, we show that the most exposed stocks to FT exhibit the weakest improvement in several liquidity metrics, including the relative spread, the cost of round trip trade, the effective spread, and the realized spread. We certainly do not pretend to draw definitive causal conclusions about the consequences of FT on liquidity, but our baseline regression findings are confirmed in a large number of robustness checks for endogeneity, non-linearity or the parallel trend hypothesis among many others.

The remainder of this paper is organized as follows. We briefly review the literature in Section 2 to stress the contribution of our work. We present our data and some descriptive statistics about market liquidity in Section 3. In Section 4, we first document how we identify fast traders based on their trading behavior and then analyze on which stocks they are particularly active. The empirical results are reported in Section 5 and robustness checks are available in Section 6. Section 7 concludes.

2. Brief Literature Review

In the existing literature, there is strong evidence that both average implicit and explicit transaction costs have decreased with the rise of FT, for both retail and institutional investors. However, as correlation is not causality, it is still possible that markets without FT could have offered even better services at lower cost. (H)FT can indeed deteriorate liquidity through higher information asymmetry, more aggressive orders widening the bid-ask spread, or a higher proportion of fleeting orders, discouraging patient traders to post limit orders and display depth.

1 FT is a broad term that characterizes market participants who use speed to gain a short term advantage. It is typically characterized by a submission of a large number of orders that are cancelled shortly after submission, and neutral positions at the end of the trading day. FT has preceded high frequency trading, which is a form of ultrafast trading that relies on colocation and proximity services to shorten latency (Kirilenko and Lo, 2013; Hendershott et al., 2011; Brogaard et al., 2014; Goldstein et al., 2014; Laughlin et al., 2014).

In the literature, we find studies of quasi-experiments around technological upgrades or regulatory changes. Although Hendershott et al. (2011), Riordan and Storkenmaier (2012), and Jovanovic and Menkveld (2016) find rather positive effects associated with the advent of HFT, Menkveld and Zoican (2017) study a decrease in latency from 350ms in 2007 to 5ms in 2009 on the NYSE platform and point to a detrimental effect on market quality because of reduced liquidity. Ye et al. (2013) also find evidence consistent with quote stuffing by HFT when investigating the effect of two Nasdaq technology upgrades in 2010 that cut the minimum time between messages from 950 nanoseconds to 200 nanoseconds.[2] They report a significant increase in the number of cancelled orders, no real variation in overall trading volume, no change in the bid-ask spread, a decrease in market depth, and an increase in short-term volatility. These authors conclude that no social benefit is observed.

Flags are also frequently used to investigate the effect of HFT on market quality. They correspond to binary variables taking the value of 1 each time an HFT firm active on the market enters an order and/or execute a trade. For example, Hendershott and Riordan (2013) and Brogaard et al. (2014) work with Nasdaq data wherein flags are available for trades involving HFTs. Hendershott and Riordan (2013) find that HFTs are associated with smaller trade sizes while humans are more related to block trading. They also suggest that HFTs consume (provide) liquidity when bid-ask spreads are relatively narrow (wide), thereby bringing down variation in market quality. In Brogaard et al. (2014), net buying by HFT liquidity takers move in line with future price changes, contributing to price discovery. Similarly, net buying by HFT liquidity suppliers and future price changes move in opposite direction, because of higher adverse selection from better-informed liquidity takers. HFTs are also found to initiate trades in the opposite direction to the transitory component of prices, contributing to price efficiency. This holds true during both volatile and quiet days.

The use of flags has pros and cons. As indicated by ESMA (2014), flags are typically based on self-disclosure by pure HFT firms. They do not include HFT activity by other firms, such as investment and brokerage houses. They may also fail to include trading activity done by HFT firms going through

2 Foucault et al. (2013, p. 41) define quote stuffing as market participants who ‘deliberately swamp platforms with messages (quotes and cancellations) solely in order to manipulate the tape (the quote and trade information reported to other participants)’.

another trading venue member with direct market or sponsored access, acting as a broker, unless the HFT firms are reported as its clients. Furthermore, the use of flags implies that all the trading activity of the flagged firms is considered as HFT, although part of it might be related to non-HFT strategies. All in all, the use of flags is likely to provide a lower bound estimate of HFT, potentially failing to include the most cunning HFT activity. Flags are also typically visible at the 'group' level only, i.e., it is impossible to know which HFT firm in particular is active at any point in time.

In this paper, we use a dataset that presents several key advantages over flags. We first have complete information on the full order book, including hidden depth and market members' ID codes. These ID codes enable us to study the trading behavior of *each* Euronext marker member in both late 2002 and early 2006, order by order and trade by trade. Reverse engineering is therefore possible as we can explain how we flag each member on an individual basis, without relying on self-disclosure by the exchange members or assuming that we live in a dichotomous trading landscape with pure fast traders on the hand and pure slow traders on the other hand. The ID codes enable us to fine-tune the identification of fast traders active on Euronext by comparing the trading activities for all market members with respect to the proportion of cancelled orders or the speed at which they cancel orders for example. In other words, we can estimate message traffic, the lifetime of orders and end-of-day inventories for *every* market member.

Such an analysis of trading and quoting patterns at the member level is important because the identification of (H)FT is a task that Brogaard et al. (2017, p. 37) denote as *'challenging, contentious, and difficult to enforce'*. Various classification proxies are usually used in past research: inventory management in Kirilenko et al. (2017); trading speed in Scholtus et al. (2014), Hasbrouck and Saar (2013), or Latza et al. (2014); message traffic and trading volume in Hendershott et al. (2011), Viljoen et al. (2014), or Harris and Saad (2014). We show in our analysis that fast traders do display characteristics close to those observed for HFTs nowadays, such as low execution-to-order ratios, high cancellation-to-order ratios, and high rapid cancellation-to-order ratios.

It is noteworthy that our empirical work covers a time period with no regulatory change on Euronext. We can observe close to 100% of all orders and trades on the CAC40 stocks in our analysis since there was no volume shift or market fragmentation before the implementation of MiFID in the late 2007.[3] As today's trading environment is much more fragmented than before because of the rise of multilateral trading facilities and dark pools, more recent datasets provide a less complete picture. In our case, all trades, quotes, and volumes are observable. Put differently, we rely on the full order book and use all the transaction level data at the member level.

Finally, most studies on FT use data from the Nasdaq or Deutsche Boerse. The use of Euronext data helps test the robustness of previous conclusions drawn mostly for the US and German equity markets, with the notable exceptions of Colliard and Hoffman (2017), Bellia (2018), Anagnostidis et al. (2020), and Bellia et al. (2020).

In the next Section, we give further details on the dataset at hand and provide descriptive statistics on trading activity and liquidity during the two periods of time under scrutiny.

3. Data

We use both order and trade data on Euronext, which cover two periods of time: 64 trading days over October 1-December 31, 2002, and 61 trading days over February 1-April 30, 2006. Our sample is made of all the 34 stocks included in the CAC40 index during both periods.[4] For each transaction, we have the ISIN code of the stock, the buyer's and the seller's ID, the time-stamp to the second, the number of shares traded and the execution price. For each order, we have the market member's ID, the order direction, both displayed and total quantities, order type (i.e., limit order, market order, and market-to-limit order), the limit price (if any), the order final state (i.e., executed, cancelled or expired), the time when the order enters the market and when it leaves. For both trades and orders, we also have the market member's account, which distinguishes proprietary trading

3 Volume shift and market fragmentation occurred later due to the implementation of MiFID starting from November 2007. MiFID stands for Markets in Financial Instruments Directive. This piece of European regulation is the second step in the harmonization of the capital markets industry across member states. MiFID swept away the very concept of central exchange and obligation of order concentration as it existed in several European countries. One of its main consequences was the opening of the execution landscape to full competition. In our sample, the cross-listing of some stocks on Brussels or Amsterdam, or in the US as ADRs was also too marginal to make any meaningful difference, as indicated later.

4 The correspondence between each stock ticker and the company name is available in the internet appendix Table A1.

from agency trading.[5] These rich data allow us to reconstruct the full order book for each stock using the Euronext market algorithm and order priority rules.[6]

Descriptive statistics are provided in Table 1. For each metric, we report the total in 2002 and 2006, as well as the corresponding daily averages and the change in percentage between these averages since the number of days differs in both periods. Panel A shows that there are respectively 11,279,320 and 14,395,179 submitted orders in 2002 and in 2006 for the stocks under scrutiny. The proportion of market and (market-to-) limit orders is relatively constant between the two periods. We nevertheless observe a significant decrease in the proportion of executed orders (from 62.16% to 57.19%), accompanied by a sharp increase in the proportion of cancelled orders (from 29.63% to 36.15%).

In Panel B of Table 1, the proportion of buy and sell orders is similar to what we observe for the full sample while the proportion of limit orders is slightly higher. We note however that there are much more (less) cancelled (executed) proprietary orders. The daily number of cancelled orders on average increases by 138.31% between 2002 and 2006.

Panel C of Table 1 provides the number of trades executed in both periods. The daily number of trades on average increases significantly by 31.83%.

3.1. Variables

Based on Harris (2003), we distinguish several dimensions of liquidity. While the choice of proxies is always disputable, we use a large set of measures, starting from the most widespread proxies, such as the quoted spread (QS), the relative spread (RS), the displayed and total depths at the Best Bid and Offer (BBO) as well as at the 5 best quotes (5BQ). We follow Hendershott et al. (2011) and Riordan and Storkenmaier (2012) in computing equally-weighted daily averages. We also consider two measures

5 The ranking of market members by proprietary orders submitted between February and April 2006 is available in the internet appendix (see Table A2). Proprietary orders are orders sent by market members for their own account, i.e., when they trade for their own inventory and not on behalf of their clients. In contrast to client orders, proprietary orders can reveal the fundamental strategy followed by a market member. There is no designated market maker, i.e., liquidity provider, for the CAC40 stocks.

6 When a stock is traded in several markets, we focus exclusively on the activity in the most liquid one for each stock. We provide further details in footnote 15.

of ex-post liquidity, i.e., the effective spread (*ES*) and the realized spread (*RealS*) to take implicit transaction costs into account. These two proxies are computed as follows:

$$ES = \begin{cases} Log(P_t) - Log(M_t) & \text{if buyer-initiated} \\ Log(M_t) - Log(P_t) & \text{if seller-initiated} \end{cases}$$

$$RealS = \begin{cases} Log(P_t) - Log(M_{t+1}) & \text{if buyer-initiated} \\ Log(M_{t+1}) - Log(P_t) & \text{if seller-initiated} \end{cases}$$

where P_t is the trade price at time *t*, M_t is the midpoint prevailing at time *t*, and M_{t+1} is the midpoint just after the trade occurring at time *t*.

Table 1: Descriptive statistics

Table 1 reports some descriptive statistics about our sample for both periods. For each order, we can identify the direction (buy/sell), the order type (limit, market-to-limit, and market order), the final state (executed, cancelled, partially filled, or expired), and the order account (proprietary vs non-proprietary orders). Column 2 (3) reports the total number of orders in 2002 (2006), respectively. Column 4 (5) reports the daily average number of orders in 2002 (2006), respectively. Column 6 indicates the variation (based on the daily averages) between both periods. Panel A reports descriptive statistics on the full sample of orders. Panel B provides these statistics on proprietary orders only, as defined in Footnote 5. Proprietary orders will serve to identify the fast traders active in 2006. Panel C reports the number of trades in the entire sample.

	2002 Total (64 days)	2006 Total (61 days)	2002 Daily avg.	2006 Daily avg.	Δ
Panel A: All orders	11,279,320	14,395,179	176,239.38	235,986.54	33.90%
Buy order	5,582,655	7,040,196	87,228.98	115,413.05	32.31%
	49.49%	48.91%			
Sell order	5,696,665	7,354,983	89,010.39	120,573.49	35.46%
	50.51%	51.09%			
Limit order	10,352,871	13,422,208	161,763.61	220,036.20	36.02%
	91.79%	93.24%			
Market-to-limit order	454,937	341,581	7,108.39	5,599.69	-21.22%
	4.03%	2.37%			
Market order	471,512	631,390	7,367.38	10,350.66	40.49%
	4.18%	4.39%			

Executed	7,010,831	8,233,070	109,544.23	134,968.36	23.21%
	62.16%	57.19%			
Cancelled	3,342,380	5,204,389	52,224.69	85,317.85	63.37%
	29.63%	36.15%			
Partially filled or expired	926,109	957,720	14,470.45	15,700.33	8.50%
	8.21%	6.65%			
Proprietary	5,372,706	8,224,750	83,948.53	134,831.97	60.61%
	47.63%	57.14%			
Non-proprietary	5,906,614	6,170,429	92,290.84	101,154.57	9.60%
	52.37%	42.86%			
Panel B: Proprietary orders	5,372,706	8,224,750	83,948.53	134,831.97	60.61%
Buy order	2,669,268	4,003,533	41,707.31	65,631.69	57.36%
	49.68%	48.68%			
Sell order	2,703,438	4,221,217	42,241.22	69,200.28	63.82%
	50.32%	51.32%			
Limit order	5,333,641	8,199,388	83,338.14	134,416.20	61.29%
	99.27%	99.69%			
Market-to-limit order	3,660	2,790	57.19	45.74	-20.02%
	0.07%	0.03%			
Market order	35,405	22,572	553.20	370.03	-33.11%
	0.66%	0.27%			
Executed	3,232,894	3,831,451	50,513.97	62,810.67	24.34%
	60.17%	46.58%			
Cancelled	1,746,141	3,966,236	27,283.45	65,020.26	138.31%
	32.50%	48.22%			
Partially filled or expired	393,671	427,063	6,151.11	7,001.03	13.82%
	7.33%	5.19%			
Panel C: Trades					
Number of trades	7,086,162	8,903,709	110,721.28	145,962.44	31.83%

In addition, we consider the cost of round trip trade (CRT), which has the advantage of depending on both tightness and depth (e.g., Irvine et al. (2000), Gomber et al. (2015)). It can be computed for several trade sizes and represents the cost associated with buying and selling a given number of shares, under the current market conditions.[7] Following Domowitz et al. (2005), the CRT for a size of q shares is computed as follows:[8]

$$CRT = [\sum_{\tau=1}^{k-1} q_\tau^A p_\tau^A + (q - \sum_{\tau=1}^{k-1} q_\tau^A) p_k^A] - [\sum_{\tau=1}^{k'-1} q_\tau^B p_\tau^B + (q - \sum_{\tau=1}^{k'-1} q_\tau^B) p_{k'}^B]$$
$$+\mathbf{1}(q > \sum_{\tau=1}^{k} q_\tau^A) \times [(q - \sum_{\tau=1}^{k} q_\tau^A)(p_k^A + x)] \quad (1)$$
$$+\mathbf{1}(q > \sum_{\tau=1}^{k'} q_\tau^B) \times [(q - \sum_{\tau=1}^{k'} q_\tau^B)(p_{k'}^B - x)]$$

where p_τ^A and p_τ^B are respectively the ask and bid prices at the τ^{th} limit, q_τ^A and q_τ^B are the corresponding number of shares, x is the minimum tick sixe, and k and k' are the indices of the last sell and buy quotes in the order book that are needed to entirely execute the round trip transaction, with $k \leq 5$ and $k' \leq 5$.

We use the number of submitted orders (NO) and the number of trades (NT) on stock i for a given day d as proxies for immediacy. To take intraday price variation into account, we compute the HighLow (HL) as the difference between the maximum and minimum transaction prices on stock i and day d. We also scale this HighLow by the daily VWAP to obtain a relative measure (RHL) as follows:

$$RHL_{d,i} = \frac{max\ price_{d,i} - min\ price_{d,i}}{VWAP_{d,i}} = \frac{HL_{d,i}}{VWAP_{d,i}} \quad (2)$$

where $VWAP_{d,i}$ is the daily volume-weighted average price for stock i on day d.

As for market activity proxies, we measure trade size as the number of shares traded. We also consider the execution-to-order ratio (EOR), the cancellation-to-order ratio (COR), and the rapid cancellation-to-order ratio (RCR). These ratios are computed as follows:

7 This measure assumes that one first buys all the quantities available, q^A_1 at price p^A_1, then the quantities available at q^A_2, and so on. Because of different minimum tick sizes, we form 3 price categories, i.e., price category 1 (stock 2 price between 0.01€ and 50€), price category 2 (stock price between 50.05€ and 100€), and price category 3 (stock price between 100.10€ and 500€). If the depth at the five best quotes is insufficient, we assume that one buys the remainder at a price p^A_5 + €0.01, €0.05, or €0.10, for the price category 1, 2, and 3, respectively. Conversely, one sells the shares, first at price p^B_1, and then at lower prices (p^B_2, p^B_3, etc.). If the depth at the five best quotes is insufficient, we assume that one sells the remainder at a price p^B_5 – €0.01, €0.05, or €0.10, for the price category 1, 2, and 3, respectively. We use the notation HCRT when we include the hidden quantities available in the order book in the computation.

8 For the sake of readability, we purposely omit the subscripts d and i while this measure is computed by day and stock, like the other proxies.

$$EOR_{d,i} = \frac{e_{d,i}}{n_{d,i}}, \quad COR_{d,i} = \frac{c_{d,i}}{n_{d,i}}, \text{ and } \quad RCR_{d,i} = \frac{r_{d,i}}{n_{d,i}} \tag{3}$$

where $e_{d,i}$ is the number of executed orders, $c_{d,i}$ is the number of cancelled orders, $r_{d,i}$ is the number of cancelled orders one second at the latest after its submission, $n_{d,i}$ the number of submitted orders, for each day d and each stock i by all the market members. We compute these ratios by relying on all orders but also on proprietary orders only. Such ratios based on cancellations or executions are complementary with the following caveat: in the database, an order that is partially filled can be classified either as 'executed' or 'cancelled'. Because of that shortcoming, we only use completely filled orders (i.e., 'executed' orders, $e_{d,i}$) and orders that lead to no transaction at all (i.e., 'cancelled' orders, $c_{d,i}$). In the subsequent analysis, we do not include agency orders to estimate these ratios because proprietary orders are the most likely to reveal the fundamental strategy followed by the market members.

Since HFTs are known for managing closely their net position (Benos and Sagade, 2016; Kirilenko et al., 2017), we also compute the daily net position for each market member j and each stock i. At the end of the day, we scale the net position (in absolute value) by the market member's trading volume as follows:

$$NP_{d,i,j} = \frac{|Net\ position_{d,i,j}|}{Volume_{d,i,j}} \tag{4}$$

where $Net\ position_{d,i,j}$ is the cumulative sum of shares weighted by either -1 for a sell or +1 for a buy, and $Volume_{d,i,j}$ is the cumulative sum of shares traded over a given day d, for stock i, and by market member j.[9]

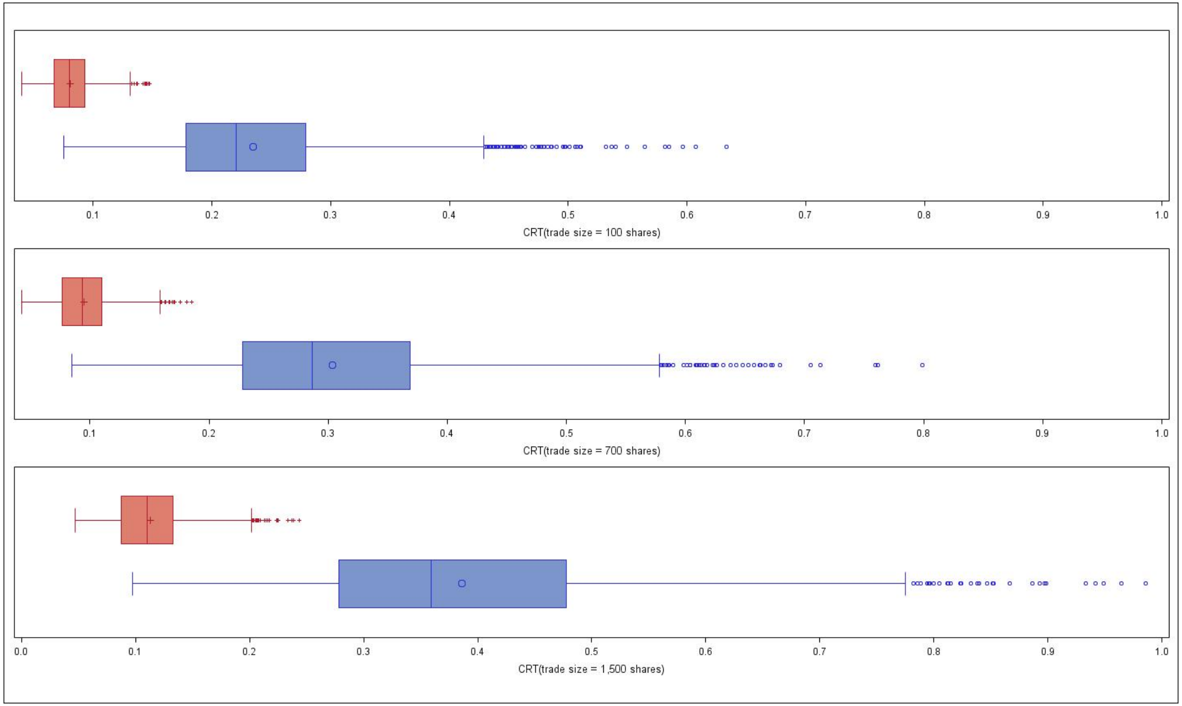
3.2. Descriptive statistics

The aforementioned proxies used to characterize liquidity and market activity are listed in Table 2. For each proxy, we compute an average for each stock and each day in each period, leading to 2,176 observations for 2002 and 2,074 observations for 2006. In Table 2, QS and RS decrease by more than 40% and 60%, respectively. By contrast, depth increases sharply in every case, whether displayed or total and whether computed at the BBO or at the five best limits. Taking into account both tightness and depth, the CRT goes down substantially, by around 70%. Figure 1 clearly displays the positive shift in terms of liquidity. Our volatility proxies show

9 We also compute $NP_{d,i,j}$ using monetary volumes. It gives almost identical results that are available upon request.

Figure 1: The cost of round trip trade (CRT) in 2006 and 2002

Figure 1 plots the cross-sectional daily average of the cost of round trip trade (CRT) in 2006 and 2002 for three trade sizes, namely, 100 shares, 700 shares, and 1,500 shares. The detailed description of how this variable is computed is available in Section 3.1. The red (blue) boxplot represents the CRT in 2006 (2002).



that transaction prices fluctuate within a smaller range in 2006 than in 2002. All these findings are significant at the 1% level and point to higher liquidity in 2006 than in 2002.

Although the number of orders increases, a larger proportion of them is cancelled in 2006, as indicated by the average COR. The average EOR is also lower in 2006, suggesting that orders lead to several and smaller executions more often in 2006 since the number of trades increases as well. Consistently, the trade size is found to decrease.

We replicate the above analysis on a stock-by-stock basis in Table 3. Our goal is to identify the number of stocks for which there is a positive or negative variation in each of our liquidity and market activity proxies. For each stock, we compare the averages using 64 and 61 daily observations in 2002 and 2006, respectively. While Table 3 confirms the previous findings at the stock level, it also shows that liquidity does not necessarily improve for all the stocks. For example, three stocks experience a significant rise in QS and two stocks exhibit a significant decrease in total depth at the BBO. However, in accordance with the RS, the CRT decreases significantly for all the stocks. We also observe a higher variability across stocks with respect to the EOR, COR, RCR, NO, NT, and trade size. As the variations in the QS and HL are likely to be impacted by the rise in stock prices between 2002 and 2006, we prefer relying on the RS or RHL, which both point to an improvement in liquidity.

Table 2: Liquidity and market activity in 2002 and 2006

Table 2 lists our different liquidity and market activity proxies, their cross-sectional daily average for each period, their evolution over time, and the result of the corresponding *t*-test. We perform this test for the quoted spread (QS), relative spread (RS), effective spread (ES), realized spread (RealS), cost of round trip trade (CRT), cost of round trip trade including hidden quantities (HCRT), displayed depth at the best bid and offer (DD_BBO), displayed depth at the five best quotes (DD_5BQ), total depth at the best bid and offer (TD_BBO), total depth at the five best quotes (TD_5BQ), number of orders (NO), number of Trades (NT), highlow (HL), relative highlow (RHL), execution-to-order ratio (EOR), cancellation-to-order ratio (COR), rapid cancellation-to-order ratio (RCR), with and without focusing on proprietary orders, and trade size. The detailed description of how these variables are computed is available in Section 3.1. ***, **, * indicate significance at 1%, 5%, and 10%, respectively.

Liquidity		2002	2006	% Change	*t*-value
Tightness	Quoted spread (QS)	0.0845	0.0492	-41.78%	24.20***
	Relative spread (RS)	0.2200%	0.0780%	-64.55%	82.97***
	Effective spread (ES)	0.0980%	0.0376%	-61.63%	73.78***
	Realized spread (RealS)	0.0515%	0.0223%	-56.70%	56.62***

	CRT (700 shares)	0.3038%	0.0948%	-68.80%	86.84***
	HCRT (700 shares)	0.2794%	0.0905%	-67.61%	86.28***
Depth	Displayed depth (DD_BBO)	4,050.6	8,322.0	+105.45%	-16.14***
	Displayed depth (DD_5BQ)	25,028.2	50,033.8	+99.91%	-14.32***
	Total depth (TD_BBO)	8450.7	14,783.5	+74.94%	-10.91***
	Total depth (TD_5BQ)	43,236.4	73,054.2	+68.96%	-12.87***
Immediacy	Number of order (NO)	5,183.5	6,940.8	+33.90%	-16.76***
	Number of Trades (NT)	3,256.5	4,293.0	+31.83%	-13.04***
Volatility	HighLow (HL)	1.8644	1.1385	-38.93%	21.06***
	Relative HighLow (RHL)	5.1993%	1.8998%	-63.46%	53.57***
Market activity					
COR	1. All orders	31.90%	37.67%	+18.09%	-18.39***
	2. Proprietary orders	33.59%	48.39%	+44.06%	-49.27***
EOR	1. All orders	59.93%	55.76%	-6.96%	14.49***
	2. Proprietary orders	58.97%	46.39%	-21.33%	44.30***
RCR	1. All orders	02.78%	04.80%	+72.66%	-12.59***
	2. Proprietary orders	02.58%	06.42%	+148.83%	-21.81***
	Trade size	717.6	626.5	-12.70%	5.89***

4. Methodology

Our goal in this paper is to study how the rise of FT has affected liquidity on Euronext using a DID analysis. For that purpose, we follow a two-step process. First, we identify the fast traders among all the market members active in 2006. Building on past research, these fast traders should display characteristics close to those observed for HFTs, i.e., high COR (or low EOR), high RCR, and/or small end-of-day net positions (NP). For that purpose, we follow a two-step process wherein we focus on proprietary orders only. This first step is described in Section 4.1. Once the fast traders are identified, we then identify the most exposed stocks to FT. This implies splitting our sample of 34 stocks into two groups: a first group including the stocks most exposed to the fast traders, and a second group including all the remaining stocks. This second step is explained in Section 4.2.

4.1. Identification of fast traders

Our identification method relies on the analysis of trading and order submission patterns at the member level. Thanks to the availability of the market members' ID codes, we do not have to assume perfect homogeneity in trading patterns among fast traders. As mentioned earlier, our data allow us to track order submission and trades for each market member in the two three-month periods under scrutiny. Since all the market members are not present in both periods, we refer to the "new market members" as those that were not registered in 2002. New market members are therefore only active in 2006. Our objective is to identify among them which ones behave like fast traders the most when trading on their own account (i.e., when submitting proprietary orders).

Table 3: Stock-by-stock variations in liquidity

Table 3 reports for each liquidity and market activity proxy the number of stocks for which we observe a decrease at 1%, 5%, and 10%, no variation, an increase at 1%, 5%, and 10%. We perform this test for the quoted spread (QS), relative spread (RS), effective spread (ES), realized spread (RealS), cost of round trip trade (CRT), cost of round trip trade including hidden quantities (HCRT), displayed depth at the best bid and offer (DD_BBO), displayed depth at the five best quotes (DD_5BQ), total depth at the best bid and offer (TD_BBO), total depth at the five best quotes (TD_5BQ), number of orders (NO), number of Trades (NT), highlow (HL), relative highlow (RHL), execution-to-order ratio (EOR), cancellation-to-order ratio (COR), rapid cancellation-to-order ratio (RCR), with and without focusing on proprietary orders, and trade size. The detailed description of how these variables are computed is available in Section 3.1. ***, **, * indicate significance at 1%, 5%, and 10%, respectively. N.S stands for 'Not significant'.

		Decrease				Increase		
Liquidity proxy		***	**	*	**N.S**	***	**	*
Tightness	Quoted spread (QS)	31				2	1	
	Relative spread (RS)	34						
	Effective spread (ES)	34						
	Realized spread (RealS)	33			1			
	CRT (700 shares)	34						
	HCRT (700 shares)	34						
Depth	Displayed depth (DD_BBO)		1			33		
	Displayed depth (DD_5BQ)	1				33		
	Total depth (TD_BBO)		1	1	4	26	1	1
	Total depth (TD_5BQ)		2		4	27	1	
Immediacy	Number of order (NO)	1		1	5	25	2	
	Number of Trades (NT)	2	1		6	21	2	2
Volatility	HighLow (HL)	29	2		3			
	Relative HighLow (RHL)	34						

Market activity								
COR	1. All orders	2	1		8	21	1	1
	2. Proprietary orders					34		
EOR	1. All orders	20			10	3		1
	2. Proprietary orders	33	1					
RCR	1. All orders	1	2		6	22	2	1
	2. Proprietary orders				1	32	1	
	Trade size	20	2	2	6	3		1

Table 4 provides the top 20 market members according to their respective number of proprietary orders (NO) between February and April 2006, wherein the ranking is used as ID code in this paper. When focusing on the most active market members, we distinguish a first cluster of 3 market members with more than 1,000,000 orders, a second cluster of 3 market members with more than 500,000 orders, and a third group of 4 market members with more than 200,000 orders. Taken together, these 10 market members generate more than 75% of the proprietary orders submitted on *each* stock. On average across stocks, they generate 85.89% of the orders. Among these 10 market members, 5 members are active in 2002, namely, those with ID n°1, 2, 4, 7, and 9. The new members in 2006 correspond to the following ID codes: n°3, 5, 6, 8, and 10. Among these new 5 market members, two of them (MM5 and MM8) are 'pure players', i.e., they only submit proprietary orders. The proportion of proprietary orders is also very high for the three other new members: about 80% for MM3, 97% for MM10, and 98% for MM6. However, unlike MM3 and MM10, MM6 can hardly be considered as a fast trader as we will see below.[10]

After the identification of the most active members in 2006, we now look at the usual suspects characterizing FT, i.e., the COR, EOR, and RCR. In Figures 2 to 4, we plot each of these variables (on the Y-axis) against the total number of proprietary orders (on the X-axis) for each market member over the period February-April 2006. The new market members in 2006 are indicated by a blue circle while the others (i.e., those already active in 2002) by a red cross.

In Figure 2 devoted to the average of $COR_{d,i}$ by member, the most crowded area (in the bottom-left corner) shows that most members submit

10 It is the reason why MM6 does not appear in bold in Table 4.

relatively few orders and do not cancel them often. Across all days, stocks, and members, the average of $COR_{d,i,j}$ is 25.47% in 2006. These market members are obviously not fast traders. A further visual inspection of Figure 2 enables us to identify four members newly active in 2006, with both high order submission and high COR. These members correspond to the following ID: n° 3, 5, 8, and 10. Of utmost interest is the level of their COR, higher than 50%. It is worth remembering that these four new members have been previously ranked among the top 10 members active in 2006 (see Table 4). By contrast, although MM6 is also included in the top 10 list, it exhibits a too low COR for being seriously considered as a fast trader. This is confirmed below.

Table 4: Top 20 Market members based on proprietary order submission

Table 4 reports the number of client orders and proprietary orders, as well as the proportion of proprietary orders for the 20 most active market members in terms of proprietary order submission. The group of fast traders is identified by their IDs in bold. Their identification is based on Figures 2 to 4.

MM ID	Client orders	Proprietary orders	Proportion of proprietary orders
1	528,296	1,423,257	72.93%
2	198,492	1,059,677	84.22%
3	**246,581**	**1,032,757**	**80.73%**
4	32,845	851,142	96.28%
5	–	**589,903**	**100.00%**
6	8,679	516,289	98.35%
7	96,232	495,688	83.74%
8	–	**420,172**	**100.00%**
9	342,834	367,880	51.76%
10	**6,438**	**225,929**	**97.23%**
11	21,954	162,010	88.07%
12	–	144,014	100.00%
13	10,216	138,774	93.14%
14	4,055	96,436	95.96%
15	97	70,959	99.86%
16	32	64,747	99.95%
17	2,746	60,623	95.67%
18	14	50,324	99.97%
19	399	41,776	99.05%
20	13,655	39,903	74.50%

Figure 2: Cross-sectional average cancellation-to-order ratio vs. number of orders

Figure 2 represents each market member *j* according to its number of proprietary orders submitted, n_j, and its cross-sectional average cancellation-to-order ratio (*COR*), with $COR_{(d,i)} = c_{d,i}/n_{d,i}$, where $c_{d,i}$ is the number of cancelled orders and $n_{d,i}$ is the number of submitted orders. 'New' market members, i.e., those only present in 2006, and 'Old' market members, i.e.' those present in both 2002 and 2006, are represented with a blue circle and a red cross, respectively.

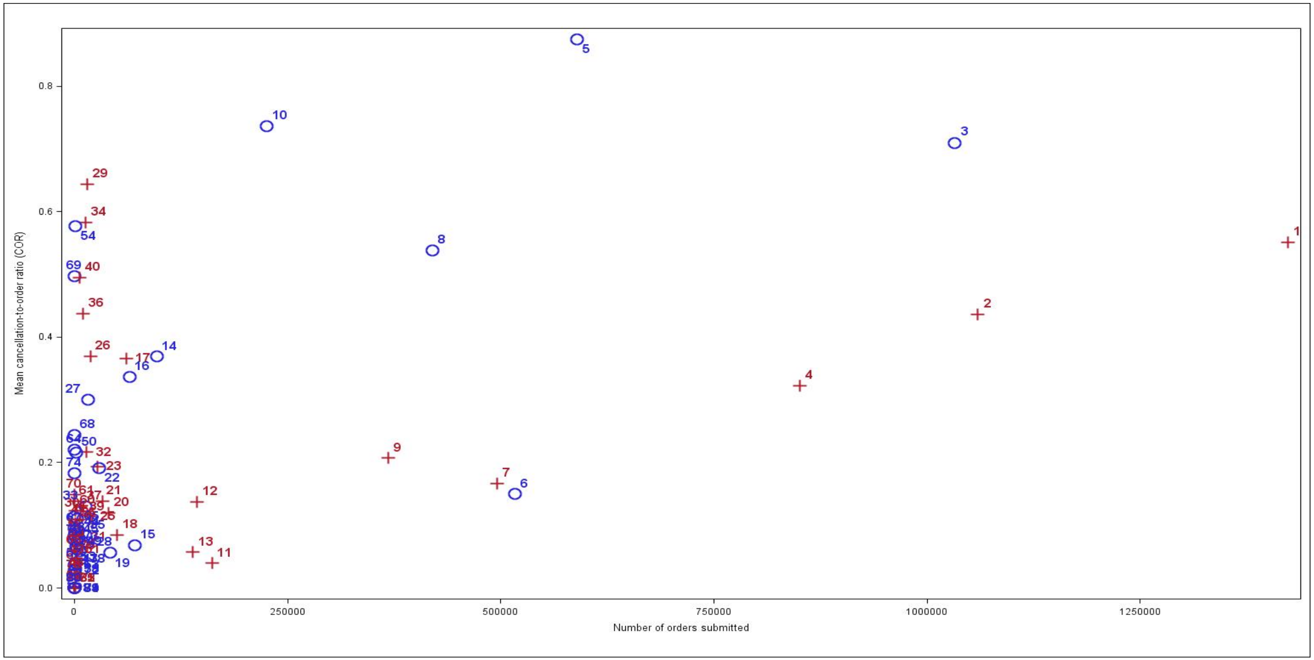


Figure 3: Cross-sectional average execution-to-order ratio vs. number of orders

Figure 3 represents each market member *j* according to its number of proprietary orders submitted, n_j, and its cross-sectional average execution-to-order ratio (*EOR*), with $EOR_{d,i} = e_{d,i}/n_{d,i}$, where $e_{d,i}$ is the number of executed orders and $n_{d,i}$ is the number of submitted orders. 'New' market members, i.e., those only present in 2006, and 'old' market members, i.e.' those present in both 2002 and 2006, are represented with a blue circle and a red cross, respectively.

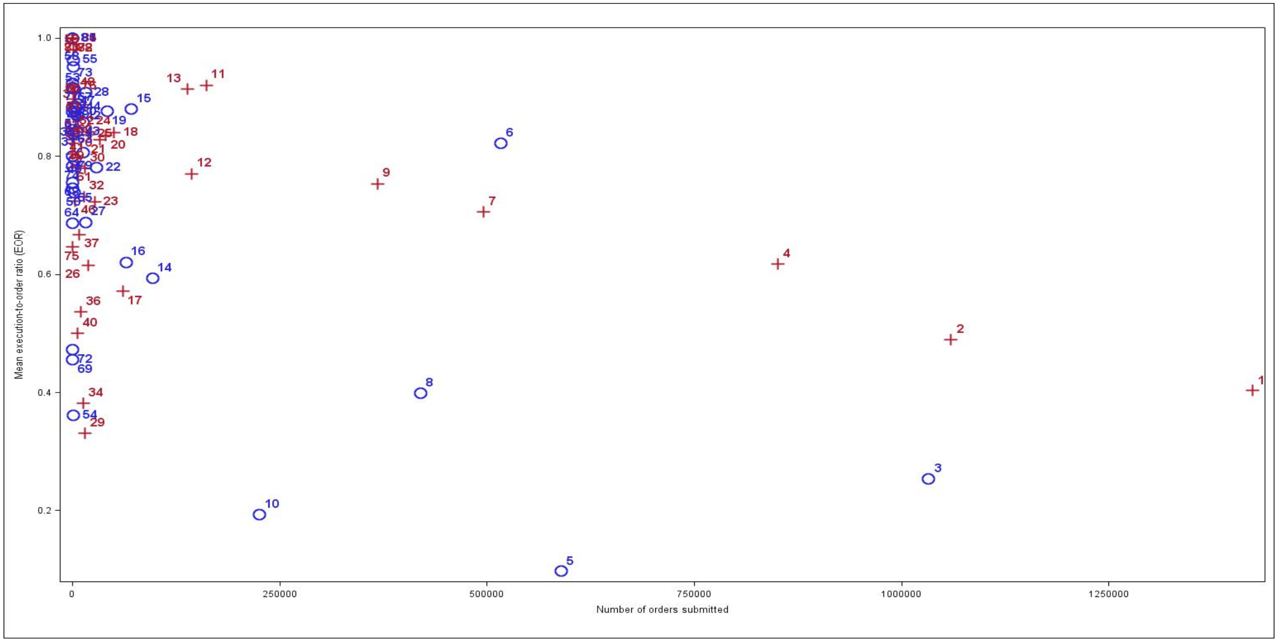
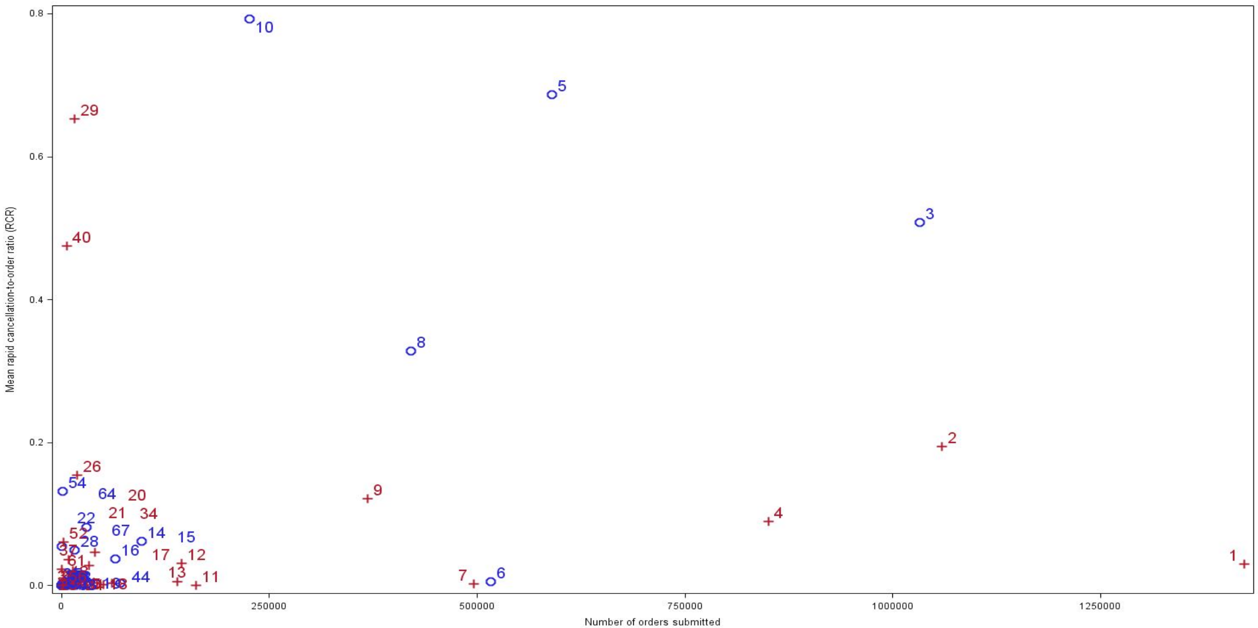


Figure 4: Cross-sectional average rapid cancellation-to-order ratio vs. number of orders

Figure 4 represents each market member *j* according to its number of proprietary orders submitted, n_j, and its cross-sectional average rapid cancellation-to-order ratio (*RCR*), with $RCR_{d,i} = r_{d,i}/n_{d,i}$, with $r_{d,i}$ the number of orders cancelled within one second at the latest, and $n_{d,i}$. 'New' market members, i.e., those only present in 2006, and 'old' market members, i.e.' those present in both 2002 and 2006, are represented with a blue circle and a red cross, respectively.



In Figure 3, we replace the COR by the EOR. As both measures are highly negatively correlated, we get almost identical results, except that the Y-axis is reversed.[11] The average of EOR (across members, stocks and days) is equal to 68.44%. Figure 3 leads to the same conclusion: MM3, MM5, MM8, and MM10 are the most likely fast traders.

Figure 4 displays the average of $RCR_{d,i}$ by member against the respective total number of submitted orders. The average of this ratio (across day, stocks, and members) is equal to 3.04%. Again, the same four new market members (ID: n° 3, 5, 8, and 10) stand out: they all exhibit a RCR higher than 5%. MM10 is clearly the fastest member, with 225,929 orders and a RCR of 35.63%. Therefore, the combination of the COR and RCR leads to the same group of fast traders.[12]

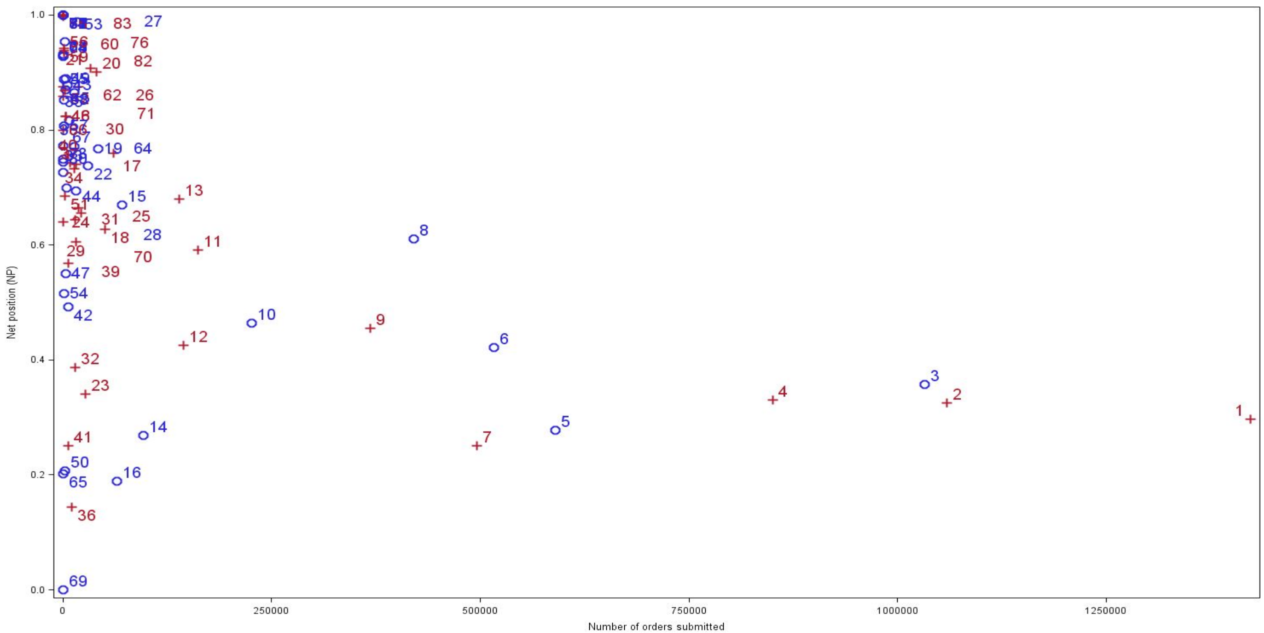
In Figure 5, we plot the total number of submitted orders against the average of NP across days and stocks for each market member. The 'average' market member closes the trading day with a net position representing 57.88% of their daily trading volume. In line with the other proxies, we observe heterogeneous behaviors across market members. Nevertheless, two clusters emerge. The first cluster includes the ten most active market members in terms of order submission; they are all holding smaller NPs than the average. The second cluster includes much smaller players whose NPs strongly differ. Regarding the selection of fast traders, these clusters are not really helpful. All the most active members display an NP at around 40% on average. Some members exhibit very low NPs, but they are tiny

11 The Pearson correlation coefficient is equal to -97.52% when we consider each market member's average COR and EOR, i.e., 85 observations. It decreases to -91.42% when we compute the Pearson correlation across the 65,285 observations (with one observation per market member, stock, and day). As expected, both coefficients are statistically different from zero with a p-value inferior to 1%. We also compute EOR_j as $\sum_{d=1}^{61}\sum_{i=1}^{34} e_{d,i,j} / \sum_{d=1}^{61}\sum_{i=1}^{34} n_{d,i,j}$, where $e_{d,i,j}$ and $n_{d,i,j}$ are the number of executed and submitted orders, respectively. These results are available upon request.

12 To further indicate that fast trading is better identified by focusing on proprietary orders, we also compare the COR and the RCR in 2006 for MM3, which submits both agency and proprietary orders. With one COR (RCR) observation by day and by stock in 2006, we obtain 2,074 observations and perform a t-test for the difference between the COR (RCR) on proprietary orders versus the COR (RCR) on client orders. The average COR on proprietary orders is 70.97% while it is only 10.64% for the COR on client orders. This difference is highly statistically significant (t-stat = 128.77). Regarding the RCR on proprietary orders, it amounts to 1.59% while it is very close to zero for the RCR on client orders. Again, this difference is statistically significant (t-stat = 66.09). These results confirm that fast trading is done on proprietary orders. When fast market members execute client orders, their execution strategies are very different.

Figure 5: Cross-sectional average net position vs. number of orders

Figure 5 represents each market member *j* according to its number of proprietary orders submitted, *nj*, and its cross-sectional average end-of-day net position (*NP*), with $\frac{|Net\ position_{d,i}|}{Volume_{d,i}}$. *Net positiond,i* is the cumulative sum of shares (weighted by a coefficient (-1 for a sell, +1 for a buy)) whereas *Volumed,i* is the cumulative sum of shares traded over a given day *d* and for a stock *i*. 'New' market members, i.e., those only present in 2006, and 'old' market members, i.e.' those present in both 2002 and 2006, are represented with a blue circle and a red cross, respectively.



players without a high COR and a high RCR. FT on Euronext in 2006 is unlikely to be well-characterized by NPs.[13]

Taken all the above findings together, we conclude that the members that behave the most like fast traders are MM3, MM5, MM8, and MM10. As explained earlier, these 4 market members are only active in 2006 (i.e., not in 2002). Among them, MM10 looks like the most active fast trader. In Section 6.6, we carry out several robustness checks and show that our results remain robust to both the choice of the criteria and the composition of the group of fast traders.

4.2. Identification of the most exposed stocks to fast trading

Now that the fast traders are identified, we need to determine the most exposed stocks to FT. Our objective is basically to split our sample of stocks (i.e., the 34 stocks which belong to the CAC40 index in both 2002 and 2006) in two groups. In the first group, we include the most exposed stocks to FT in 2006. A given stock *i* considered as most exposed to FT when at least 3 out of the 4 fast traders (i.e., MM3, MM5, MM8, and MM10) submit a significantly higher proportion of orders for this stock than the proportion of orders submitted by all the members. In the second group, we include all the remaining stocks, that are marginally or not affected by FT in 2006.

For each stock *i*, we compare the order submission of market member *j* relative to the aggregate order submission of all market members. Let us denote $n_{d,i,j}$, the number of submitted orders for day *d*, stock *i*, by market member *j*. Therefore, $n_{d,j} = \sum_{i=1}^{34} n_{d,i,j}$, which represents the sum of submitted orders for day *d* and market member *j*. Dividing $n_{d,i,j}$ by $n_{d,j}$, we obtain $b_{d,i,j}$, that is, the realized proportion of submitted orders for day *d*, stock *i*, and market member *j*. We then compute $\beta_{i,j}$, i.e., the 2006 time-series average of $b_{d,i,j}$ representing the realized proportion of submitted orders by market member j for stock *i* in 2006. We also divide the number

13 Benos and Sagade (2016) report that HFTs have on average an end-of-day net position of 19% when analyzing U.K. stocks at the end of 2012. More interestingly, passive HFTs display an average end-of-day net position of 28%. With the benefit of hindsight, our estimation of 40% as end-of-day net position for the fast traders in 2006 seems therefore very reasonable, given that 2006 is really associated with the emergence of FT. End-of-day net position and mean reversion in inventory are also used by Kirilenko et al. (2017) who study E-mini S&P 500 futures contracts during the 2010 Flash Crash. They use a cutoff of 5% for end-of-day net position and 0.5% for the mean reversion proxy. However, as these authors note on page 977, *'these cutoff levels are specific to our sample and may need to be adjusted if applied to other markets.'* Needless to say, there are numerous differences between their study and ours (E-mini futures *vs* stocks, 2010 *vs* 2002 and 2006, U.S. market *vs* French market, etc.).

of orders submitted on day d for stock i ($n_{d,i}$) by the total number of orders submitted on day d (n_d) to obtain the proportion of orders submitted on day d for stock i by all market members, $x_{d,i}(= n_{d,i}/n_d)$. We subsequently compute $\tilde{X}_i$, i.e., the time-series average of $x_{d,i}$.

For each of the four fast traders (i.e., MM3, MM5, MM8, MM10), we check whether the null hypothesis $\beta_{i,j} <= \tilde{X}_i$ is rejected at the 1% significance level. In this case, stock i is considered as being 'fast-traded'.[14] Stock i is considered as most exposed to FT when at least 3 out of the 4 fast traders submit a higher proportion of orders for this stock than the proportion of orders submitted by all the members. We find that the four fast traders submit large proportions of orders on 5 stocks, namely, UG (Peugeot), SGO (Saint Gobain), GLE (Société Générale), BNP (BNP Paribas), and RNO (Renault).[15] In the next section, we investigate whether the liquidity of these 5 stocks is higher or lower than the liquidity of the other stocks included in the second group.

5. Being a fast-traded stock

As explained in Section 3, we construct a set of daily averages of liquidity metrics to obtain 4,250 observations, i.e., one observation per stock and per day, with 34 stocks, 64 days in 2002 and 61 days in 2006. For conducting

14 We obtain a matrix $M_{i,j}$ of binary values indicating whether stock i is actively traded by market member j, with $m_{i,j} = 1$ and zero otherwise. Next, we sum for each stock i the number of market members that trade the stock actively, i.e., $\sum_{j=1}^{N} m_{i,j}$ with N being the number of market members. We report this information in the internet appendix—Table A3.

15 We also check whether these five stocks are cross-listed as ADRs at the time of this study to make sure we have a reliable view on the way they are globally traded. Two of them are not listed as ADRs (Saint Gobain and Renault). For the remaining three stocks, we use Datastream to download daily prices and volumes (in number of shares) for their ADRs. We then compare the traded volume (in number of shares) on Euronext to the traded volume as ADRs (in number of shares). The median proportion, i.e., the volume traded as ADR relative to the volume traded on Euronext, is extremely low: 6%, 1%, and less than 1% for Societe Generale, BNP Paribas, and Peugeot, respectively. Not only the volume traded outside the domestic market is unlikely to change anything in our DID analysis, but we do not find any evidence that the fast traders identified in our sample are implementing in early 2006 some forms of multi-market trading on these 3 stocks. We have also identified 9 active cross-listings within Euronext in our sample. Only 2 of the 9 cross-listed stocks were "fast-traded" stocks, i.e., Saint-Gobain and Peugeot. In both cases, trading volume in euros on the secondary (and sometimes tertiary) exchange was too negligible to make any meaningful difference. It never accounted for more than 0.06% of the trading volume on the primary exchange for these two stocks in late 2002 or early 2006. For the non-fast traded stocks, trading volume in Brussels and/or Amsterdam never accounted for more than 0.33%, excluding Dexia which was quoted in Paris and Brussels, belonging to both the CAC40 and BEL20 indices. Trading volume in Brussels accounted for 61% and 73% of trading volume in Paris at the end of 2002 and early 2006, respectively. To make any meaningful difference in our analysis, we would have to find that fast trading was substantially higher on the shares of Dexia traded in Brussels than on those traded in Paris (since Dexia is identified as a 'non-fast-traded' in Paris). We find the opposite. The average number of orders by day is 5949 (3192), the average COR is 67.02% (36.01%), and the average RCR is 11.74% (2.5%) in the French (Belgian) order book, the differences in means being always statistically significant at 1%. In the "Belgian" order book, we also do not find any trading activity (in terms of order submission) from the two most important fast traders, i.e., MM1 and MM2. Cross-listing could not reasonably constitute an exogenous factor to explain why stocks are heavily traded by FT in our sample.

a DID analysis, we regress some of these liquidity measures on a set of both indicators and control variables. The indicators include a dummy variable picking out the period (with $D1_d = 1$ if day d belongs to year 2006, and zero otherwise), a dummy variable flagging the fast-traded stocks (with $D2_i = 1$ if stock i is fast-traded, and zero otherwise), and the interaction of those two dummies (with $D3_{d,i} = D2_i * D1_d = 1$ if stock i is fast-traded on day d in 2006). After controlling for time series fluctuations and differences across both groups of stocks, we capture the difference in liquidity due to FT by the coefficient a_3 in the following equation:

$$L_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + \gamma X_{d,i} + \sum_{p=1}^{2} \beta_p + \epsilon_{d,i} \qquad (5)$$

where $L_{d,i}$ is a liquidity proxy (either the RS, the ES, the RealS, or the CRT), and $X_{d,i}$ is a matrix of control variables listed in Table 5. It includes volatility (RHL), market cap (MC), market number of orders (NO) as a proxy for number of messages (Albuquerque et al., 2020; Eaton et al., 2020; Shkilko and Sokolov, 2020), and cancellation-to-order ratio (COR).[16] β_p is a fixed effect by price category as the tick size on Euronext depends on the stock price during the period under scrutiny. Descriptive statistics about the liquidity and control variables used in Equation 5 are available in Table 5.

Table 6 provides the results of the above DID regression. Our coefficient estimate of interest is related to the $D3_{d,i}$ dummy. For all the liquidity proxies, it is positive and statistically significant at 1%. The highest adjusted R^2 is obtained when the CRT is used as liquidity proxy, with a value equal to 77.03%. In that case, the coefficient estimate of $D3_{d,i}$ is equal to 5.58 basis points. This is economically very substantial since the average CRT for the fast-traded stocks in 2002 is 5.17 basis points lower in comparison to the stocks in the control group, as indicated by the coefficient estimate of $D2_{d,i}$. In other words, the most exposed stocks to FT in 2006 lost all the liquidity edge they experienced before the rise of FT. Another way to measure the economic impact of FT is to realize that the improvement in liquidity between 2002 and 2006 for the control stocks is 16.12 basis points, as measured by the CRT and indicated by the coefficient estimate of $D1_{d,i}$. It is still noteworthy that the two groups of stocks benefit overall from an

16 There is no noticeable difference when we use $NT_{d,i}$, i.e., the number of trades on day d and stock i, instead of $NO_{d,i}$. As expected, both variables are highly correlated with a Pearson correlation coefficient equal to 94.65%. To control for cancellation fees, we include the cancellation-to-order ratio (COR) as an additional control variable. There is no change when we use the number of cancelled orders (NC), instead of the COR. These unreported results are available upon request. In the internet appendix, we report the results for time-weighted relative spreads in Table A4.

improvement in liquidity between 2002 and 2006. The non-fast-traded stocks display a fall in their average CRT of around 16 basis points while the corresponding decrease is only 11 basis points for the fast-traded stocks. Again, this lower decrease in CRT is due to FT (5.58 basis points), whose magnitude is very close to the liquidity advantage of 5.17 basis points that these stocks exhibit in 2002.

Table 5: Liquidity and control variables—descriptive statistics

Table 5 reports descriptive statistics (minimum, mean, median, and maximum) for both dependent and control variables used in Equation 5. These descriptive statistics are provided for both 2002 and 2006 and computed across the control and the fast-traded stocks, respectively. The list of variables includes the relative spread (RS), cost of round trip trade (CRT), effective spread (ES), realized spread (RealS), relative highlow (RHL), market capitalization (MC), number of orders (NO), number of Trades (NT), cancellation-to-order ratio (COR), and volume average weighted price (VWAP). The detailed description of how these variables are calculated is available in Section 3.1. All these variables are computed by day and by stock.

	Panel A: **Control stocks**				Panel B: **Fasted-traded stocks**			
2002	**Minimum**	**Mean**	**Median**	**Maximum**	**Minimum**	**Mean**	**Median**	**Maximum**
RS(%)	0.0735	0.2262	0.2122	0.6139	0.0708	0.1845	0.1841	0.3853
CRT_700(%)	0.0851	0.3130	0.2963	0.7992	0.0924	0.2507	0.2500	0.4921
ES(%)	0.0366	0.1011	0.0931	0.4584	0.0375	0.0802	0.0790	0.3078
RealS(%)	0.0073	0.0534	0.0486	0.2208	0.0121	0.0406	0.0398	0.1153
RHL(%)	0.9398	5.2217	4.5892	28.4664	1.1822	5.0694	4.3348	13.4744
MC	0.0017	0.0171	0.0109	0.1026	0.0067	0.0179	0.0134	0.0406
NO	629	5120.91	3981.5	32681	769	5546.62	4697	21534
NT	235	3223.56	2370.5	27438	413	3447.6	2737	15995
COR	0.1081	0.3196	0.3045	0.7392	0.1402	0.3151	0.3035	0.5661
VWAP	2.3679	42.8701	29.2662	144.8043	19.3291	40.6408	41.9966	58.6017
2006	**Minimum**	**Mean**	**Median**	**Maximum**	**Minimum**	**Mean**	**Median**	**Maximum**
RS(%)	0.0392	0.0761	0.0754	0.1410	0.0466	0.0891	0.0889	0.1392
CRT_700(%)	0.0427	0.0934	0.0915	0.1851	0.0633	0.1027	0.1022	0.1812
ES(%)	0.0178	0.0366	0.0358	0.1119	0.0173	0.0432	0.0432	0.0936
RealS(%)	0.0073	0.0215	0.0209	0.0574	0.0063	0.0268	0.0254	0.0436
RHL(%)	0.3190	1.8959	1.6850	12.4527	0.6278	1.9222	1.7275	6.7192
MC	0.0039	0.0295	0.0189	0.1404	0.0111	0.0340	0.0235	0.0663
NO	1,351	6,909	6,224	30,495	2,495	7,123	6,598	17,557
NT	660	4,292	3,519	24,743	1,539	4,294	3,844	12,993
COR	0.1568	0.3744	0.3570	0.9166	0.2290	0.3905	0.3850	0.5508
VWAP	11.1279	60.8841	45.3169	228.7280	47.1223	76.8997	75.2289	125.4745

Table 6: Liquidity and fast trading

Table 6 reports the results for the following regression model: $L_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + b_1 * RHL_{d,i} + b_2 * MC_{d,i} + b_3 * NO_{d,i} + \sum_{p=1}^{2} \beta_p + \epsilon_{d,i}$. In column 2, $L_{d,i}$ is the relative spread ($RS_{d,i}$) on day d and stock *i*; in column 3, $L_{d,i}$ is the cost of round trip trade for a trade size of 700 shares ($CRT\_7000_{d,i}$); in column 4, $L_{d,i}$ is the effective spread ($ES_{d,i}$); in column 5, $L_{d,i}$ is the realized spread ($RealS_{d,i}$). *Constant* is the intercept. $D1_d$ is a dummy variable picking out the period (with $D1_d = 1$ in 2006, and zero otherwise). $D2_i$ is a dummy variable flagging the fast-traded stocks (with $D2_i = 1$ if the stock is fast-traded, and zero otherwise). $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$). The set of control variables include the relative highlow ($RHL_{d,i}$), the market capitalization ($MC_{d,i}$), the number of orders ($NO_{d,i}$), the cancellation-to-order ($COR_{d,i}$), and the price category fixed-effects (1 and 2). The detailed description of how these variables are calculated is available in Section 3.1. The number of observations (N) and the adjusted R-squared (R^2) are also given. ****, **, * indicate significance at 1%, 5%, and 10%, respectively. Standard errors are robust to heteroskedasticity.

	$RS_{d,i}$	$CRT\_700_{d,i}$	$ES_{d,i}$	$RealS_{d,i}$
Constant	0.1382***	0.2527***	0.0688***	0.0461***
$D1_d$	-0.1018***	-0.1612***	-0.0367***	-0.0192***
$D2_d$	-0.0379***	-0.0517***	-0.0190***	-0.0116***
$D3_{d,i}$	0.0546***	0.0558***	0.0263***	0.0167***
$RHL_{d,i}$	0.0127***	0.0143***	0.0072***	0.0030***
$MC_{d,i}$	0.0413	-0.0003***	-0.0001***	-0.0001***
$NO_{d,i}$	-0.0072***	-0.0105***	-0.0021***	-0.0009***
$COR_{d,i}$	0.1433***	0.2063***	0.0158***	-0.0035
β_1	0.0151***	-0.0251***	0.0022**	-0.0015**
β_2	0.0055***	-0.0187***	0.0018**	0.0007
N	4,250	4,250	4,250	4,250
R^2	75.70%	77.03%	69.38%	52.94%

We reach the same conclusion when considering the other three liquidity proxies used as dependent variables in Table 6. First, the liquidity level in 2002 is higher for the fast-traded stocks than for the control stocks ($a_2 < 0$), which is fully consistent with the literature since fast traders tend to primarily target more liquid stocks. Second, there is an improvement of liquidity over time ($a_1 < 0$). Third, the impact of FT on liquidity is detrimental ($a_3 > 0$).[17]

Although the a_3 coefficient estimates are smaller in magnitude in the *RS* and *RealS* regressions, they must be compared to the sum of their respective intercept and a_2 coefficient estimates, capturing what prevailed on average for the treatment group in 2002. In the *RS* regression, the ratio is 54%,

17 To have a proper interpretation of the constant, we also estimate the model without any fixed-effect. For example, we find that the average RS among the control stocks in 2002 is 14.72 basis points, which is quite close to the 13.82 basis points in Table 6. These unreported results are available upon request.

meaning that the average liquidity penalty for being most exposed to FT in 2006 represents around half of the relative spread (RS) estimated in 2002. This is 53% for the effective spread (ES) and 48% for the realized spread (RealS). Consequently, being most exposed to FT is detrimental in terms of ex-ante liquidity, but it also leads to a similar worsening of realized implicit trading costs.

In Table 6, all the coefficient estimates for the control variables display the expected signs, i.e., they reveal a positive relationship between liquidity and both RHL and COR, as well as a negative relationship with NO and MC (except in the RS regression where it is not significant).

6. Robustness checks

We run a large number of robustness checks to tackle methodological issues in our study, such as the risk of endogeneity, the risk of censorship in the RS, the imbalance between the number of treated and control stocks, the risk of a non-linear effect, the risk of violating the parallel trend assumption, and the risk of not identifying fast traders correctly.[18] None of them modifies the empirical results reported in Section 5.

6.1. Is there evidence of endogeneity?

Even though we do not observe the exact timing of FT entry, we do not identify violations of the exogeneity assumption and rule out the likely sources of endogeneity as explained below.

The most likely source of endogeneity would come from changes in the industrial organization of the exchange, as pointed out by Brogaard and Garriott (2019). This would affect both the timing of FT entry and liquidity. There was no specific organizational change on Euronext between the late 2002 and the early 2006. The most plausible explanation behind the rise of FT is that fast traders were essentially attracted to Euronext Paris because of structural wider spreads and lower competition, compared to what prevailed for example in the UK or the Netherlands at that time.

18 For the sake of completeness, we report another set of robustness checks in the internet appendix. Among these additional analyses, we investigate whether the effect of FT depends on trade size; whether the use of clustering in the standard errors makes any difference; whether there is evidence of FT in 2002; whether fast traders are more likely to employ spoofing-like strategies; and whether the correlation in order submission is higher for the fast-traded stocks.

We might still conjecture that FT depends on short-term variations in liquidity across large stocks on Euronext in early 2006. Although Brogaard and Garriott (2019) find that the exact time of entry by HFT firms on the Nasdaq is very unlikely to depend on unusual short-term variations in liquidity, we take this concern into account by looking at the situation that prevailed before February 2006 and after April 2006. Since it is impossible to get an extension of our dataset from Euronext, we use the TRTH database provided by Refinitiv to download hourly bid-ask prices for all the stocks in our sample from January 2002 to December 2006. We compute the daily averages of RS based on the best quotes and apply the usual filters to clean up the data (such as dropping negative spreads and excluding abnormally high levels of RS).

Table 7: Fast trading and short-term market conditions (Relative spread)

This table reports the regression results for Equation (5) without the control variables, using TRTH data. The model is specified as follows: $L_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + \epsilon_{d,i}$. $L_{d,i}$ is the relative spread. The TRTH data however do not enable us to include the usual control variables. The regression is estimated over five different time windows: (i) February 2006 – April 2006, (ii) January 2006 – March 2006, (iii) December 2005 – February 2006, (iv) March 2006 – May 2006, and (v) April 2006 – June 2006. For comparison purposes, the corresponding results based on the Euronext data are reported in the first row. As in Section 5, the group of fast traders includes MM3, MM5, MM8, and MM10 and the group of treated stocks includes 5 stocks (UG, SGO, GLE, BNP, and RNO).***, **, * indicate significance at 1%, 5%, and 10%, respectively. Standard errors are robust to heteroskedasticity.

	Constant	$D1_d$	$D2_i$	$D3_{di}$	*N*	*Adj.* R^2
Feb 2006–Apr 2006 (Euronext)	0.2262***	-0.1501***	-0.0416***	0.0547***	4,250	62.26%
Feb 2006–Apr 2006	0.1997***	-0.1300***	-0.0429***	0.0573***	3,717	49.41%
Jan 2006–Mar 2006	0.1997***	-0.1302***	-0.0429***	0.0569***	3,812	50.14%
Dec 2005–Feb 2006	0.1997***	-0.1292***	-0.0429***	0.0555***	3,782	46.54%
Mar 2006–May 2006	0.1997***	-0.1270***	-0.0429***	0.0586***	3,755	48.19%
Apr 2006–Jun 2006	0.1997***	-0.1217***	-0.0429***	0.0548***	3,703	45.62%

In Table 7, we provide the results for Equation 5 estimated over five different time windows: (i) February 2006 – April 2006 (base case); (ii) January 2006 – March 2006; (iii) December 2005 – February 2006; (iv) March 2006 – May 2006; and (v) April 2006 – June 2006. All the coefficient estimates and the adjusted R^2 are very stable and close to those we get when using the Euronext data (available in the top row of Table 7, for convenience). There is no evidence of endogenous entry by fast traders in late 2005 or early 2006.

To further alleviate the concern about reverse causality, we run a two-stage least squares regression analysis. In the first stage, we estimate the following LOGIT model:

$$D2_i = \alpha_0 + \alpha_1 LIQ_{d,i} + \epsilon_{d,i} \tag{6}$$

where $D2_i$ is a dummy variable set to one when stock i is fast-traded (and zero otherwise) and $LIQ_{d,i}$ is the relative spread on day d for stock i. The number of observations is 2,176, i.e., 34 stocks times 64 days in 2002. As expected, the liquidity of a stock in 2002 is positively related to its probability of being fast-traded in 2006. Next, we compute the fitted probabilities which are purged of their correlation with past liquidity. Accordingly, we obtain:

$$\hat{D}2_{d,i} = \frac{exp^{(\hat{\alpha}_0 + \hat{\alpha}_1 LIQ_{d,i})}}{1 + exp^{(\hat{\alpha}_0 + \hat{\alpha}_1 LIQ_{d,i})}} \tag{7}$$

We compute the averages of $\hat{D}2_{d,i}$ by stock and then flag the 5 stocks with the highest mean as fast-traded. In the second stage, we estimate Equation 5 with the adjusted fitted values. The results for the RS are reported in Table 8. They are consistent with our main findings provided in Section 5.

Our last ad hoc approach consists in estimating the following regression:

$$L_{d,i} = a_0 + a_1 D1_d + a_2 D2_i + a_3 D3_{d,i} + a_4 D4_{d,i} + \gamma X_{d,i} + \sum_{p=1}^{2} \beta_p + \epsilon_{d,i} \tag{8}$$

wherein we control for the level of liquidity in 2002 using the dummy $D4_{d,i}$. The latter is equal to 0 if $L_{d,i}$ is measured in 2002, and to $\overline{L}_i$ if $L_{d,i}$ is measured in 2006, with L_i equal to the mean of the liquidity proxy in 2002 for stock i. All the other variables were defined previously. The results for the RS are reported in Table 9. Again, when controlling for the past level of liquidity, our findings remain unaffected, i.e., $D1_d$ and $D2_i$ still have a negative coefficient estimate and $D3_{d,i}$ a positive coefficient estimate.

Should FT been found to improve liquidity, reverse causality would potentially be an issue in the baseline regressions of Table 6. When there is a positive feedback loop between liquidity and FT, there is the risk of attributing *too much* of the improvement in liquidity to FT. In such a case, since fast traders target the most liquid stocks in the first place,[19] it is hard to

19 The group of 'fast-traded' stocks is indeed on average more liquid than the control group in late 2002 since a_2 is negative in the regressions.

distinguish the positive effect of FT on liquidity from the positive effect of liquidity on FT. However, we do not document a positive feedback loop in this paper. We find the opposite since FT is detrimental to liquidity. Given this negative feedback loop, there is instead the risk of underestimating the negative effect of FT on liquidity.

Table 8. Two-stage least squares analysis—Relative spread

Table 8 reports the coefficient estimates and their level of significance for the first stage and second stage regressions. These equations are $D2_i = \alpha_0 + \alpha_1 LIQ_{d,i} + \epsilon_{d,i}$ and $L_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + b_1 * RHL_{d,i} + b_2 * MC_{d,i} + b_3 * NO_{d,i} + \sum_{p=1}^{2} \beta_p + \epsilon_{d,i}$, respectively. *Constant* is the intercept; $D1_d$ is a dummy variable picking out the period (with $D1_d = 1$ in 2006, and zero otherwise); $D2_i$ is a dummy variable picking out the treated group of stocks (with $D2_i = 1$ if the stock is fast-traded, and zero otherwise); $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$). The set of control variables include the relative highlow ($RHL_{d,i}$), the market capitalization ($MC_{d,i}$), the number of orders ($NO_{d,i}$), the cancellation-to-order ($COR_{d,i}$), and the price category fixed-effects (β_1 and β_2). A detailed description of these variables is available in Section 3.1. We report the number of observations (N) and the adjusted R-squared (R^2). ***, **, * indicate significance at 1%, 5%, and 10%, respectively. Standard errors are robust to heteroskedasticity.

First stage	**Coeff.**	***p*-value**
Constant	0.0885	
$LIQ_{d,i}$	-9.0843	***
N	2,176	
Pseudo R^2	4.19%	
Second stage	**Coeff.**	***p*-value**
Constant	0.1607	***
$D1_d$	-0.1117	***
$D2_d$	-0.0651	***
$D3_{d,i}$	0.0668	***
$RHL_{d,i}$	0.0112	***
$MC_{d,i}$	0.0000	
$NO_{d,i}$	-0.0058	***
$COR_{d,i}$	0.1444	***
β_1	-0.0050	***
β_2	-0.0125	***
N	4,250	
R^2	77.12%	

Table 9. Liquidity and fast trading—endogeneity (Relative spread)

Table 9 reports the results for Equation 8: $L_{d,i} = a_0 + a_1 D1_d + a_2 D2_i + a_3 D3_{d,i} + a_4 D4_{d,i} + \gamma X_{d,i} + \sum_{p=1}^{2} \beta_p + \epsilon_{d,i}$. *Constant* is the intercept; $D1_d$ is a dummy variable picking out the period (with $D1_d = 1$) in 2006, and zero otherwise); $D2_i$ is a dummy variable picking out the treated group of stocks (with $D2_i = 1$ if the stock is fast-traded, and zero otherwise); $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$); $D4_{d,i}$ which is equal to 0 if $L_{d,i}$ is measured in 2002, and to $\overline{L}_i$ if $L_{d,i}$ is measured in 2006, with $\overline{L}_i$ equal to the mean relative spread in 2002 for stock *i*. The set of control variables include the relative highlow ($RHL_{d,i}$), the market capitalization ($MC_{d,i}$), the number of orders ($NO_{d,i}$), the cancellation-to-order ($COR_{d,i}$), and the price category fixed-effects (β_1 and β_2). A detailed description of these variables is available in Section 3.1. We report the number of observations (N) and the adjusted R-squared (R^2). ***, **, * indicate significance at 1%, 5%, and 10%, respectively. Standard errors are robust to heteroskedasticity.

	Coeff.	p-value
Constant	0.1418	***
$D1_d$	-0.0700	***
$D2_d$	-0.0378	***
$D3_{d,i}$	0.0498	***
$D4_{d,i}$	-0.1328	***
$RHL_{d,i}$	0.0127	***
$MC_{d,i}$	0.0000	
$NO_{d,i}$	-0.0073	***
$COR_{d,i}$	0.1388	***
β_1	0.0155	***
β_2	0.0049	***
N	4,250	
R^2	75.96%	

There might still be a trend in a latent variable influencing FT and liquidity at the same time. The usual suspect is market capitalization, which is arguably positively correlated with both liquidity and FT. Nevertheless, because our sample includes only the largest blue-chip stocks in France, market capitalization disparities are very limited. The CAC40 index is quite narrow (with 40 stocks at best), in comparison with other broader stock indexes (such as the S&P500, for instance). We also find that the five most exposed stocks to FT in 2006 are not the largest stocks in our sample,[20] suggesting that there is no particular discrimination from fast traders based on market capitalization in our sample.

20 The five fast-traded stocks in our baseline regressions are ranked 3rd, 5th, 15th, 18th, and 28th.

6.2. Is the parallel trend hypothesis violated?

For the parallel trend assumption to hold in the DID analysis, liquidity in the control group must move parallel with liquidity in the treatment group until the event. This might not be the case if there are specific events or structural changes affecting this relationship, implying that the sign of the effect would not come from FT. As mentioned earlier, our empirical work covers a time period with no regulatory change on Euronext. On the contrary, it spans a period preceding the implementation of MiFID and the resulting market fragmentation.

To address the concern over the parallel trend hypothesis, Bilinski and Hatfield (2020) recently propose a non-inferiority approach. Building on the latter, we estimate the restricted model and its unrestricted version as follows:

$$L_{d,i} = a_0 + \sum_{k=D_0}^{D} \beta_k \mathbb{1}(k = d \cap D2_i = 1) + \alpha_i + \gamma_d + \epsilon_{d,i} \tag{9}$$

$$L_{d,i} = a'_0 + \sum_{k=D_0}^{D} \beta'_k \mathbb{1}(k = d \cap D2_i = 1) + \theta D2_i d + \alpha_i + \gamma_d + \epsilon'_{d,i} \tag{10}$$

where $L_{d,i}$ is a liquidity proxy for stock i on day d, β_0 is the intercept, D_0 is the time at which the treatment starts, D is the number of time periods, i.e., 61 days in 2006, d_i is the dummy related to the treatment, β_i is a stock-fixed effect, and β_d is a time-fixed effect. β and β' are then computed as $\beta = \frac{1}{k}\sum_{i=1}^{k} \beta_i$ and $\beta' = \frac{1}{k}\sum_{i=1}^{k} \beta'_i$, respectively.

With this specification, we get one estimate by day over the 2006 period, which leads to 61 estimates for β_k. Next, we compute the average of them to obtain the average treatment effect on the fast-traded stocks. When considering the RS in Equation 9, the average treatment effect is 0.0546, which replicates the coefficient estimate for a_3 in Table 6. In the unrestricted model, we get a highly significant estimate for β', equal to 0.0267 (t-stat = 35.97). When we include the control variables in the unrestricted regression as in our baseline regressions, the coefficient estimate is equal to 0.0478 (t-stat = 46.13). The parallel trend assumption is therefore not violated in our analysis.[21]

21 For the sake of brevity, the detailed results are unreported. They are however available upon request.

6.3. Is the effect of fast trading non-linear?

To consider a potentially non-linear dynamics between liquidity and FT, we adopt a quantile regression approach. The goal is to test whether the effect of FT (a_3) depends on the stock liquidity level. By contrast to an OLS regression that minimizes the sum of squared residuals, a quantile regression estimate coefficients by minimizing the following expression:[22]

$$Min \sum_{i:y_i \geq \beta_0+\beta_1 x_i} Q|y_i - (\beta_0 + \beta_1 x_i)| + \sum_{i:y_i < \beta_0+\beta_1 x_i} (1-Q)|y_i - (\beta_0 + \beta_1 x_i)| \quad (11)$$

where Q denotes the quantile of the dependent variable.[23] Based on this method, we estimate Equation 5 for each decile of the distribution for the RS. We report the results in Table 10. The increasing value of the coefficient estimates of the interaction variable indicates that the negative impact of FT on liquidity is confirmed and is stronger when the level of liquidity is lower, i.e., when spreads are wider.[24]

6.4. Is the imbalance between the number of fast-traded and control stocks an issue?

There are 5 fast-traded stocks and 29 control stocks in our analysis. Such an imbalance might be an issue for validating the DID methodology. The solution typically consists in performing a one-to-one matching using either the nearest neighbour, the Mahalanobis' distance, or propensity scores. With equity data, Davies and Kim (2009) recommend a matching on two criteria such as market capitalization and price. Applying this approach, we find no significant difference for both the coefficient estimates and the *p*-values related to the interaction dummy variable in the DID regression. For example, the results based on propensity score matching without replacement are available in Table 11. The coefficient estimate of the interaction dummy is equal to 5.8 (3.3) basis points when we match stocks on their characteristics in 2002 (2006), and the *p*-value for a comparison with the baseline coefficient estimate (5.46 basis points) is below 1% in each case.

22 We formalize the expression for a linear regression with a single independent variable. This setting can be easily extended to the case of a multiple linear regression.

23 When Q is equal to 0.5, it becomes equivalent as minimizing the least absolute deviation (LAD), i.e., minimizing the sum of absolute residuals.

24 Building on Athey and Imbens (2006), we also estimate a changes-in-changes model. This methodology allows time periods and groups to be treated asymmetrically. The unreported results are qualitatively similar and are available upon request.

Table 10: Quantile regression—Relative Spread

Table 10 reports the results for Equation (5) estimated through a quantile regression for each decile of the distribution of the relative spread. The group of fast traders includes MM3, MM5, MM8, and MM10, and the group of fast-traded stocks includes 5 stocks (UG, SGO, GLE, BNP, and RNO), as in Section 5. *Constant* is the intercept; $D1_d$ is a dummy variable picking out the period (with $D1_d$ = 1 in 2006, and zero otherwise); $D2_i$ is a dummy variable picking out the treated group of stocks (with $D2_i$ = 1 if the stock is fast-traded traded, and zero otherwise); $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$). The set of control variables include the relative highlow ($RHL_{d,i}$), the market capitalization ($MC_{d,i}$), the number of orders ($NO_{d,i}$), the cancellation-to-order ($COR_{d,i}$), and the price category fixed-effects (β_1 and β_2). A detailed description of these variables is available in Section 3.1.

	10th	20th	30th	40th	50th	60th	70th	80th	90th
Constant	0.1085	0.1183	0.1325	0.1383	0.1503	0.1516	0.1537	0.1708	0.1939
$D1_d$	-0.0667	-0.0768	-0.0833	-0.0896	-0.0990	-0.1049	-0.1137	-0.1311	-0.1515
$D2_d$	-0.0147	-0.0201	-0.0242	-0.0262	-0.0316	-0.0357	-0.0409	-0.0480	-0.0614
$D3_{d,i}$	0.0215	0.0292	0.0381	0.0426	0.0480	0.0519	0.0554	0.0634	0.0749
$RHL_{d,i}$	0.0095	0.0098	0.0103	0.0111	0.0113	0.0127	0.0136	0.0134	0.0135
$MC_{d,i}$	-0.0000	-0.0000	-0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
$NO_{d,i}$	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
$COR_{d,i}$	0.0754	0.0775	0.0836	0.0905	0.0932	0.1056	0.1291	0.1361	0.1549
β_1	-0.0033	0.0051	0.0001	0.0016	0.0025	0.0037	0.0062	0.0108	0.0173
β_2	0.0094	0.0129	0.0044	0.0030	0.0024	0.0016	0.0019	0.0028	0.0004

6.5. Does tick constrainedness lead to censoring in the relative spread?

Tick constrainedness implies that the minimum value of the quoted spread cannot go below 1 cent. Table 5 shows that the minimum value for the RS is 3.92 basis points in 2006. This value would be the minimum RS for a stock price of 25.13.[25] Hence, any stock with a price below €25.13 is subject to tick constrainedness. If there are more fast-traded stocks than control stocks trading at prices below €25.13, this could bias our empirical results. However, the Volume-Weighted Average Prices (or VWAP) for each stock in our sample are high enough to conclude that tick constrainedness does not lead to censoring in the RS. In 2006, most of the stocks display an average VWAP above €25.13. In addition, the stocks with VWAP lower than €25.13 are not identified as being exposed to FT.[26]

25 If *bid ask* = 0.01 and *RS* = 3.98 basis points, then the midpoint is equal to 25.13.

26 Table A1 in the internet appendix reports the minimum, average, and maximum prices for each stock in 2002 and in 2006.

Table 11: Relative spread and fast trading using propensity score matching

Table 11 replicates the DID regression stated in Equation (5) using propensity score matching without replacement in either 2002 or 2006. The baseline regression is estimated using the 5 fast-traded and 29 control stocks, for which the results are also reported in Table 6. *Constant* is the intercept; $D1_d$ is a dummy variable picking out the period (with $D1_d = 1$ in 2006, and zero otherwise); $D2_i$ is a dummy variable picking out the treated group of stocks (with $D2_i = 1$ if the stock is actively traded, and zero otherwise); $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$). The set of control variables include the relative highlow ($RHL_{d,i}$), the market capitalization ($MC_{d,i}$), the number of orders ($NO_{d,i}$), the cancellation-to-order ($COR_{d,i}$), and the price category fixed-effects (β_1 and β_2). A detailed description of these variables is available in Section 3.1. We report the number of observations (N) and the adjusted R-squared (R^2). ***, **, * indicate significance at 1%, 5%, and 10%, respectively. Standard errors are robust to heteroskedasticity.

	Coeff.	*p*-value	Coeff.	*p*-value	Coeff.	*p*-value
	Baseline		Matching in 2002		Matching in 2006	
Constant	0.1382	***	0.08386	***	0.15325	***
$D1_d$	-0.1018	***	-0.1257	***	-0.1211	***
$D2_d$	-0.0379	***	-0.0421	***	-0.0173	***
$D3_{d,i}$	0.0546	***	0.05804	***	0.03253	***
$RHL_{d,i}$	0.0127	***	0.01319	***	0.00911	***
$MC_{d,i}$	0.0413		0.00038	***	-0.0002	**
$NO_{d,i}$	-0.0072	***	-0.008	***	-0.0065	***
$COR_{d,i}$	0.1433	***	0.29409	***	0.21887	***
β_1	0.0151	***	0.02028	***	-0.0306	***
β_2	0.0055	***	0.0241	***	-0.0091	***
N	4,250		1,250		1,250	
R^2	75.70%		79.99%		79.79%	

It is therefore practically impossible that tick constrainedness in the quoted spread leads to censorship in the RS. Furthermore, the other liquidity proxies, which are not affected by any constrainedness, point to the same conclusions.[27]

6.6. Do different combinations of fast traders make any difference?

Table 12 provides alternative groups of fast traders, depending on various selection criteria. The first group (G1) refers to the benchmark selection established in Section 4.1 and used in Section 5. For the sake of brevity, we

27 The CRT is not affected by any constrainedness since it is always feasible to trade 1,500 shares at any point in time during any trading day for the CAC40 stocks in our sample.

do not review each alternative.[28] The crucial point is that all the alternative groups either just fit with the benchmark selection of fast traders or extend it without affecting our findings.[29]

Table 12: Composition of alternative groups of fast traders

Table 12 lists alternative groups of fast traders, depending on various selection criteria. The first group (G1) corresponds to the benchmark selection presented in Section 4.1 and used in Section 5.

Groups	Market Members	Criteria
G1	**3, 5, 8, 10**	Only in 2006, RCR >5%, and COR >50%
G2	**3, 5, 8, 10**	Only in 2006 and RCR >5%
G3	**3, 5, 8, 10, 54**	Only in 2006 and COR >50%
G4	**3, 5, 8, 10**	Only in 2006 and top-ten volume and RCR >5%
G5	**3, 5, 8, 10**	Only in 2006 and top-ten volume and COR >50%
G6	**2, 3, 5, 8, 10, 29, 40**	RCR >5%
G7	**2, 3, 5, 8, 10, 29, 40**	COR >50%
G8	**2, 3, 5, 8, 10**	Top-ten volume and RCR >5%
G9	**2, 3, 5, 8, 10**	Top-ten volume and COR >50%

7. Conclusion

Using two three-month periods at the end of 2002 and the start of 2006, we study the rise of fast trading (FT) on Euronext. Based on descriptive statistics, we first observe an overall increase in liquidity between 2002 and 2006. For example, the relative and effective spreads become tighter for all the CAC40 stocks included in our sample; this is also the case for the cost of round trip trade. Nevertheless, market members cancel a higher percentage of orders in 2006. They also trade smaller quantities, compared to 2002.

Next, we identify fast traders in 2006 by directly measuring the cancellation-to-order ratio (COR), rapid cancellation-to-order ratio (RCR), and end-of-day net positions (NP) for every Euronext market member, using their ID codes. Building on past research, fast traders are expected to display characteristics close to those observed for HFTs, i.e., high COR, high RCR,

28 In the G6 to G9 groups for example, we include both ‘old’ and ‘new’ market members. In G8 and G9, they must also belong to the top-10 members ranked by decreasing order submission in 2006. This leads to the inclusion of three fast-traders at best, that is, MM2, MM29, and MM40.

29 For the sake of brevity, the results are unreported but are available upon request.

and/or small NP. After the identification of four fast traders, we determine the most exposed stocks to FT. Practically, we split our sample of 34 stocks into two groups: a first group including the most exposed stocks to these fast traders, and a second group including all the remaining stocks.

Although liquidity improves for both groups of stocks between 2002 and 2006, we provide empirical evidence that the most exposed stocks to FT benefit the least. Being significantly fast-traded in 2006 is economically quite detrimental since these stocks lost in 2006 the liquidity 'edge' they had in 2002, before the rise of FT. All the different robustness checks lead to the same conclusion. The most exposed stocks to FT over time could have maintained their liquidity edge observed in 2002 had they been less subject to FT. As in Ye et al. (2013), Menkveld and Zoican (2017), or Brogaard et al. (2017), we question the positive effects associated with FT across all stocks since our findings show that FT has adverse effects on liquidity beyond a significant dose.

What remains an open question is whether the most exposed stocks to FT are at the same time less exposed to noise trading in such a way that the smaller improvement in market liquidity might be due to higher information content in trading (i.e., lower noise trading) *and/or* to FT.[30] This joint hypothesis remains to be tested empirically and is left for future research, on top of the causality issue since more FT may be the cause or the consequence of less noise trading. This is the reason why we have resisted the temptation to recommend stricter, across-the-board regulation of FT.

30 Bloomfield et al. (2009) show that noise traders are *uninformed* contrarian traders preventing prices from converging to fundamental asset values but increasing market volume and depth, while *reducing* bid-ask spreads and the temporary price impact of trades, i.e., the CRT in our study.

References

Albuquerque, R., S. Song, and C. Yao (2020). The price effects of liquidity shocks: A study of the sec's tick size experiment. *Journal of Financial Economics 138* (3), 700–724.

Anagnostidis, P., P. Fontaine, and C. Varsakelis (2020). Are high–frequency traders informed? *Economic Modelling 93*, 365–383.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74* (2), 431–497.

Bellia, M. (2018). High-frequency market making: liquidity provision, adverse selection, and competition. *Working Paper*.

Bellia, M., L. Pelizzon, M. G. Subrahmanyam, and D. Yuferova (2020). Designated market makers: Competition and incentives. Technical report, SAFE Working Paper.

Benos, E. and S. Sagade (2016). Price discovery and the cross-section of high-frequency trading. *Journal of Financial Markets 30*, 54–77.

Bilinski, A. and L. A. Hatfield (2020). Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions. *Working paper*.

Bloomfield, R., M. O'Hara, and G. Saar (2009, 01). How Noise Trading Affects Markets: An Experimental Analysis. *The Review of Financial Studies 22* (6), 2275–2302.

Brogaard, J. and C. Garriott (2019). High-frequency trading competition. *Journal of Financial and Quantitative Analysis 54* (4), 1469–1497.

Brogaard, J., T. Hendershott, S. Hunt, and C. Ysusi (2014). High-frequency trading and the execution costs of institutional investors. *Financial Review 49* (2), 345–369.

Brogaard, J., T. Hendershott, and R. Riordan (2014). High-frequency trading and price discovery. *Review of Financial Studies 27* (8), 2267–2306.

Brogaard, J., T. Hendershott, and R. Riordan (2017). High frequency trading and the 2008 short-sale ban. *Journal of Financial Economics 124* (1), 22–42.

Chordia, T., A. Goyal, B. N. Lehmann, and G. Saar (2013). High frequency trading. *Journal of Financial Markets 16* (4), 637–645.

Colliard, J.-E. and P. Hoffman (2017). Financial transaction taxes, market composition, and liquidity. *Journal of Finance 72* (6), 2685–2716.

Davies, R. and S. Kim (2009). Using matched samples to test for differences in trade execution costs. *Journal of Financial Markets 12* (2), 173–202.

Domowitz, I., O. Hansch, and X. Wang (2005). Liquidity commonality and return comovement. *Journal of Financial Markets 8* (4), 351–376.

Eaton, G. W., P. J. Irvine, and T. Liu (2020). Measuring institutional trading costs and the implications for finance research: The case of tick size reductions. *Journal of Financial Economics*.

ESMA (2014). High-frequency trading activity in EU equity markets. *Economic Report Number 1*.

Foucault, T., M. Pagano, and A. Röell (2013). *Market liquidity: theory, evidence, and policy*. Oxford University Press.

Goldstein, M. A., P. Kumar, and F. C. Graves (2014). Computerized and high-frequency trading. *Financial Review 49* (2), 177–202.

Gomber, P., U. Schweickert, and E. Theissen (2015). Liquidity dynamics in an electronic open limit order book: An event study approach. *European Financial Management 21* (1), 52–78.

Harris, J. H. and M. Saad (2014). The sound of silence. *Financial Review 49* (2), 203–230.

Harris, L. (2003). *Trading and exchanges: market microstructure for practitioners*. Oxford University Press.

Hasbrouck, J. and G. Saar (2013). Low-latency trading. *Journal of Financial Markets 16* (4), 646–679.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does algorithmic trading improve liquidity? *Journal of Finance 66* (1), 1–33.

Hendershott, T. and R. Riordan (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis 48* (4), 1001–1024.

Irvine, P. J., G. J. Benston, and E. Kandel (2000). Liquidity beyond the inside spread: Measuring and using information in the limit order book. *Available at SSRN 229959*.

Jovanovic, B. and A. J. Menkveld (2016). Middlemen in limit order markets. *Available at SSRN 1624329*.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun (2017). The flash crash: high-frequency trading in an electronic market. *Journal of Finance 72* (3), 967–998.

Kirilenko, A. A. and A. W. Lo (2013). Moore's law versus murphy's law: Algorithmic trading and its discontents. *Journal of Economic Perspectives 27* (2), 51–72.

Latza, T., I. W. Marsh, and R. Payne (2014). Fast aggressive trading. *Available at SSRN 2542184*.

Laughlin, G., A. Aguirre, and J. Grundfest (2014). Information transmission between financial markets in Chicago and New York. *Financial Review 49* (2), 283–312.

Menkveld, A. J. and M. A. Zoican (2017). Need for speed? Exchange latency and liquidity. *Review of Financial Studies 30* (4), 1188–1228.

Riordan, R. and A. Storkenmaier (2012). Latency, liquidity and price discovery. *Journal of Financial Markets 15* (4), 416–437.

Scholtus, M., D. van Dijk, and B. Frijns (2014). Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking & Finance 38*, 89–105.

Shkilko, A. and K. Sokolov (2020). Every cloud has a silver lining: Fast trading, microwave connectivity, and trading costs. *The Journal of Finance 75* (6), 2899–2927.

Viljoen, T., P. J. Westerholm, and H. Zheng (2014). Algorithmic trading, liquidity, and price discovery: An intraday analysis of the SPI 200 futures. *Financial Review 49* (2), 245–270.

Ye, M., C. Yao, and J. Gai (2013). The externalities of high frequency trading. *Available at SSRN 2066839*, 1–56.