Liquidity and the rise of fast trading on Euronext

Christophe Desagre* Catherine D'Hondt† Mikael Petitjean‡

June 2, 2018

Abstract

Using full order book and trade data on the CAC40 stocks, we investigate how liquidity changes during the rise of fast trading on Euronext between 2002 and 2006. We find that market activity is more intense in 2006, with a much higher percentage of orders being cancelled. Although we observe disparities among stocks, we find that volatility goes down, spreads decrease, and depth rises from 2002 to 2006. Using market members' identification codes, we identify fast traders in 2006 based on message traffic and the lifetime of orders. We then separate the CAC40 stocks in two groups. By distinguishing stocks most exposed to fast trading in 2006 from those most immune from it, we show that the fast-traded stocks benefit the least from the liquidity increase. These fast-traded stocks could have maintained their liquidity 'edge' observed in 2002, had they been better immune from fast trading.

*Louvain School of Management (LFIN-IMMAQ, UCLouvain), 151 Chaussée de Binche, 7000 Mons, Belgium, Email : christophe.desagre@uclouvain.be

†Louvain School of Management (LFIN-IMMAQ, UCLouvain), 151 Chaussée de Binche, 7000 Mons, Belgium, Email : catherine.dhondt@uclouvain.be

‡Louvain School of Management (LFIN-IMMAQ, UCLouvain) and IESEG School of Management (LEM UMR 9221 CNRS). E-mail: m.petitjean@ieseg.fr.

We are grateful to Euronext for providing us with the data. We would like to thank Albert J. Menkveld, Sophie Moinas, Nicolas K. Scholtes, Matthias Saerens, Maurizio Montone and participants at the UCL-UNamur PhD Workshop (3^{rd} March 2016 - Namur, Belgium), 3L Workshop (29^{th} April 2016 - Brussels, Belgium), 33^{rd} AFFI Conference ($23^{rd}-25^{th}$ May 2016- Liège, Belgium) and other workshops and conferences for their useful comments. Any errors are the full responsibility of the authors.

1 Introduction

Financial markets have been considerably reshaped by technological innovation since the start of the new century. Trading in any asset now involves electronic markets on which some form of fast trading, algorithmic trading, or high frequency trading (HFT hereafter) takes place. Fast trading is a broad term that characterizes market participants who use speed to gain a short term advantage. In its origin, fast trading was not equivalent to algorithmic trading in the sense that gaining a speed advantage did not necessarily require the use of algos. Conversely, the first algos were not necessarily designed with the purpose of gaining a speed advantage. Today, this distinction tends to disappear as market participants who use algos to automatically trade and minimize the price impact of larger orders, do also consider speed in the execution of their consecutive smaller child orders. In fact, the most widespread form of algorithmic trading has become HFT, which can be described as the subset of algorithmic trading that most rapidly turns over its positions. This implies the use of algos at very high frequency and super-human speed to take profit of short-lived trading opportunities. It is typically characterized by a submission of a large number of orders that are cancelled shortly after submission, neutral positions at the end of a trading day, and the use of 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)).

This paper investigates the rise of fast trading (FT) on Euronext bewteen 2002 and 2006. As argued by Kirilenko and Lo (2013, p. 53), '*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*'. Chordia et al. (2013, p. 639) also recognizes that *'the question of whether financial markets before the advent of HFTs were better or worse than today's HFT-dominated markets remains unanswered'*. We shed additional light on this debate by using market members' identification codes (ID hereafter) to identify fast traders on Euronext in 2006, based on order submission, cancellation and execution ratios, as well as end-of-day net positions. Using a diff-in-diff approach, we estimate how liquidity evolved between 2002 and 2006 for a subset of stocks most exposed to fast trading in 2006, and compare them to a group of stocks less exposed to it.

The extant literature brings very strong evidence that both average implicit and explicit transaction costs have decreased with the rise of fast trading, for both retail and institutional investors.[1] A tentative explanation behind this trend is a fiercer competition among both exchanges and market participants, who increasingly rely on speed to sharpen their competitive edge. As illustrated by Menkveld (2016), the average effective spread and retail commission fees on the NYSE/Nasdaq decreased by more than 50% from 2001 to 2011. The trend was similar for institutional investors, with explicit costs going down from 15 bps to 6 bps. Their total transaction costs, as measured by the implementation shortfall, even felt down from 50 bps to 9 bps. As correlation is not causality, it is still possible that markets without fast trading could have offered even better services at lower cost, considering market stability as well. The benefits of higher competition between trading venues, extended trading opportunities, and market making by fast traders might still be outweighed by the costs related to large-order predation and run games by fast traders.

It is notoriously difficult to draw definitive conclusions about the consequences of fast trading on liquidity for the simple reason that counterfactual analysis of causation in market microstructure is extremely difficult to carry out. In the literature, we find studies of quasi-experiments around technological upgrades or regulatory changes. For example, Hendershott et al. (2011) use the decision to automate quote dissemination on the NYSE in 2003 as an exogenous instrument to measure the causal effect of fast trading on liquidity. This autoquote system was implemented to replace the manual and slower updates of the best bid and offer prices by specialist clerks. This change benefited fast traders the most by reflecting their orders more quickly in the NYSE disseminated quote. Autoquote was applied to stocks in several consecutive waves and could be used as an instrumental variable capturing the amount of fast trading in a diff-in-diff approach. Their findings indicate that, at least for large-cap stocks, price quotes experience lower adverse selection costs, which explains the lower spread. The authors also find that the use of algos enhances informativeness of quotes.

Riordan and Storkenmaier (2012) study the impact of an exchange speed upgrade on market quality for 98 stocks actively traded on the Deutsche Boerse. After a series of system upgrades, time between order entry and confirmation was reduced from about 50 milliseconds to about 10 milliseconds. Small- and medium-sized stocks benefited the most. The rate of order submission increased by more than 50%; effective spreads decreased by around 9%; average price impact

[1] To simplify the discussion, we refer to the broad term of FT in the paper, unless otherwise required. This choice is also motivated by the fact that no flag was available on Euronext in 2002 and 2006 to identify trades being generated by algorithms or by high frequency traders.

dropped by more than 60% ; and realized spreads tripled, suggesting that fast traders were able to earn greater revenues. The contribution of quotes to price discovery also doubled, indicating that prices became more efficient.

Jovanovic and Menkveld (2016) examine the effect on market quality of the entry of an HFT firm into the trading of Dutch stocks in July 2007. These stocks were mostly traded on Chi-X, which provided faster execution and rebates for liquidity provision. Using a diff-in-diff approach, the authors show that bid-ask spreads and other market quality measures were better on the 'treated' Dutch stocks than on a set of Belgian control stocks (that could not be traded on Chi-X and were therefore unaffected by the entry of the HFT firm). Effective bid-ask spreads on the Dutch stocks were about 15% narrower and adverse selection was 23% lower, while volatility remains unaffected.

Although the above-cited studies find rather positive effects associated with the advent of HFT, it is not always the case. Ye et al. (2013) investigate the effect of two Nasdaq technology upgrades in 2010 that cut the minimum time between messages from 950 nanoseconds to 200 nanoseconds. They find evidence consistent with quote stuffing,[2] with a significant increase in the number of cancelled orders but no real variation in overall trading volume. They report no change in the bid-ask spread, a decrease in market depth, and an increase in short-term volatility. The authors conclude that no social benefit is observed. 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. Regarding regulatory changes, Brogaard et al. (2017) study the SEC's short-sale ban in September 2008.[3] They find that HFT harms liquidity.

Co-location events offer another opportunity to study the consequences of fast trading on liquidity. Boehmer et al. (2014) find that co-location improves the bid-ask spread and the quality of price discovery, although volatility does not fall as much as expected based on the narrower spread. Brogaard et al. (2015) find that HFT firms use the fastest co-location service primarily for market making, in line with the traditional theoretical prediction that speed upgrade helps market makers maintain their quotes. Research on the impact of co-location on market liquidity is not limited to stock markets. Frino et al. (2014) study the Australian Securities Exchange's

[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)'.

[3]See also Boehmer et al. (2013) for an analysis of this event.

decision of allowing co-location in its futures market in 2012. They report that liquidity is improved through tighter spreads and higher depths, after the implementation of co-location. However, the use of co-location, proximity services or access to fast data feeds may overestimate the true level of fast trading, since traditional brokers trading exclusively on behalf of their clients (i.e. agency trading) may also use these tools to offer best execution to their clients, without any proprietary trading taking place.

It is also important to note that flags are frequently used to investigate the effect of fast trading on market quality. 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 is 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 have pros and cons, as indicated by ESMA (2014). On the one hand, flags apply to pure HFT firms only. On the other hand, it does not include HFT activity by other firms, such as investment and brokerage houses. It may also fail to include trading activity done by HFT firms going through another trading venue member with direct market or sponsored access, acting as a broker, unless the HFT firms are reported as its clients. The use of flags also implies that all trading by the flagged firms is HFT, although non-HFT strategies might also be implemented. All in all, the use of flags is likely to provide a lower bound estimate of fast trading, potentially failing to include the most cunning fast trading activity.

In this paper, we identify fast trading by considering message traffic and the lifetime of orders for *every* market member on Euronext using their ID codes.[4] In most studies, the method used to identify different groups of traders relies on the analysis of trading and quoting patterns at the exchange or the stock level, but rarely at the member level. It is also noteworthy that we

[4]Each market member is assigned a numerical code that allows us to identify their order submission and trades execution while preserving their anonymity.

can observe all the orders and trades on the CAC40 stocks in our analysis, since there was not yet any volume shift or market fragmentation due to the implementation of MiFID in 2002 and 2006.[5]

The identification of HFT is a task that Brogaard et al. (2017, p. 37) still denote as *'challenging, contentious, and difficult to enforce'.* This explains why various classification proxies are usually used in the existing literature: 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). In our case, ID codes enable us to trace back the activities of each market member and to fine-tune the identification of any change in trading style based on their message traffic and lifetime of orders. More specifically, we are able to compare the levels of fast trading for every market member between 2002 and 2006, and identify the fast traders active on Euronext in 2006.

Using full order book and transaction level data at the member level, we identify the stocks most exposed to fast trading in 2006 and separate them from the other stocks. We find that market activity is more intense in 2006, with a higher percentage of orders being cancelled. Although we observe disparities among stocks, we find that volatility goes down, spreads decrease, and depth rises from 2002 to 2006. Following a diff-in-diff approach, we show that the 'fast-traded' stocks exhibit the weakest decrease in spreads and the cost of round trip trade (CRT).

The remainder of this paper is organized as follows. We present our data and some descriptive statistics about market liquidity in Section 2. In Section 3, 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 4 and robustness checks are available in Section 5. Section 6 concludes.

2 Data

We use both order and trade data on Euronext, which cover two periods: 64 trading days from October 1, 2002 to December 31, 2002, and 61 trading days from February 1, 2006 to April 30,

[5]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.

2006. These data (or part of them) have been used in De Winne and D'Hondt (2007), Mazza and Petitjean (2016), and Mazza and Petitjean (2018) among others. Our sample is made of all the 34 stocks included in the CAC40 index during both periods. 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. execution, cancellation or expiration), 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 from agency trading. These rich data allow us to reconstruct the full order book for each stock using Euronext's market algorithm and order priority rules.

The use of this dataset presents several key advantages. First, we have information on the full order book, including hidden depth and market members' ID codes. Second, we avoid the volume shift and market fragmentation that have been occurring since the implementation of MiFID in November 2007. 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 are often less representative and less reliable. Some recent studies lack information on the level of trading activity that prevails on all the competitive execution venues. By contrast, our Euronext data allow us to match for every order the quantity and the volume traded that resulted from that particular order. Market fragmentation makes nowadays such detailed and accurate information hard to obtain. Third, most studies on fast trading use data from the Nasdaq or Deutsche Boerse. The use of Euronext data brings some diversity and helps us test the robustness of previous conclusions drawn mostly for the US and German equity markets. Last but not least, we estimate message traffic, the lifetime of orders and end-of-day inventories for *every* market member on Euronext. We therefore observe disaggregated trading and quoting patterns for each member, without having to disentangle trading activity between fast and slow traders at the stock or, worse, the exchange level only. In Table 1, we list the different stock exchanges, securities, and time periods covered in several studies, including all the studies cited above. Interestingly, the use of data prior to 2006 in recent studies is no exception.

Table 1: Datasets in recent studies

Author(s)	Dataset and/or sample	Period
Hendershott et al. (2011)	NYSE	Feb. 2001 to Dec. 2005
Lee et al. (2013)	Korean Exchange	Nov. 2001 to Feb. 2002
Harris and Saad (2014)	Nasdaq-100	May 2003
D'Hondt et al. (2015)	Euronext	Feb. 2006 to Apr. 2006
De Winne and D'Hondt (2007)	Euronext	Feb. 2006 to Apr. 2006
Mazza and Petitjean (2016) and (2018)	Euronext	Feb. 2006 to Apr. 2006
Riordan and Storkenmaier (2012)	Deutsche Boerse	Feb. 2007 to Jun. 2007
	(110 stocks from HDAX)	
Jarnecic and Snape (2014)	London Stock Exchange	Jun. 2007, Jun. 2008,
		Apr. 2009 to Jun. 2009
Brogaard et al. (2014)	FTSE250	Nov. 2007 to Aug. 2011
Hendershott and Riordan (2013)	Deutsche Boerse	Jan. 2008
Brogaard et al. (2017)	Nasdaq	Aug. 2008 to Oct. 2008
Boehmer et al. (2013)	NYSE, NASDAQ and BATS	Aug. 2008 to Oct. 2008
Brogaard et al. (2014)	Nasdaq (120 random stocks)	2008 – 2009
Ye et al. (2013)	Nasdaq	2008 – 2009, Feb. 2010
	(120 random stocks)	and Oct. 2010
Scholtus et al. (2014)	Nasdaq (S&P500 ETF)	Jan. 2009 to Dec. 2011
Laughlin et al. (2014)	CME, NASDAQ, NYSE	Apr. 2010 to Aug. 2012
Kirilenko et al. (2017)	E-mini S&P500 futures	May 2010 (3 to 6)
Baron et al. (2012)	E-mini S&P500 futures	Aug. 2010 to Aug. 2012
Frino et al. (2014)	Australian Securities Exchange	Aug. 2011 to Aug. 2012
Friederich and Payne (2015)	Stoxx Europe 600 index	Feb. 2012 to May 2012
Benos and Sagade (2016)	LSE, Chi-X, BATS, Turquoise	Sep. 2012 to Dec. 2012

Table 1 reports the stock exchange and/or the sample of securities, as well as their corresponding sample period.

Descriptive statistics about the sample are provided in Table 2. Panel A shows that there were respectively 11,279,320 and 14,395,179 submitted orders in 2002 and in 2006 for the stocks under scrutiny. The number of orders has therefore increased by around 28%. The proportion of order types was relatively constant between the two periods. We nevertheless observe a significant decrease (increase) in the proportion of orders executed (cancelled). To identify fast traders, we will focus on market members' proprietary orders and trades. We therefore report statistics for this subsample of orders in Panel B. 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 also note that there are more (less) cancelled (executed) proprietary orders. Panel C provides the number of trades and the number of trading days in both sample periods. As expected, the number of trades is higher in the second period, despite the slightly lower number of trading days.

In Table 3 are listed the various proxies that we use to characterize liquidity and market activity. Based on Harris (2003), we distinguish several dimensions of liquidity, while keeping in mind that this classification is disputable since borders between proxies are not watertight. We compute equally-weighted quoted spreads (QS) and relative spreads (RS), displayed and total depths at the Best Bid and Offer (BBO) and at the 5 best quotes (5BQ).

In addition, we consider the cost of a round trip trade (CRT henceforth), the advantage of which is to depend on both tightness and depth. It can be computed for several trade sizes and represents the cost associated with buying and selling a given number of shares, under current market conditions.[6] The CRT for a given size is computed as follows:[7]

$$CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}} \tag{1}$$

where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, p_τ^A and p_τ^B are respectively the ask and bid prices at the τ^{th} limit, and q_τ^A and q_τ^B are the corresponding number of shares.

We use the number of submitted orders (NO) and the number of trades (NT) on a 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:

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

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

As for market activity proxies, we consider the execution-to-order ratio (EOR), the cancellation-to-order ratio (COR), and the rapid cancellation-to-order ratio (RCR). These are computed as

[6] A trader will first buy all the quantities available, q_1^A at price p_1^A, then the quantities available at q_2^A, and so on. If the depth at the five best quotes is insufficient, we assume that she can buy the remainder at a price $p_5^A + 0.01$€. Conversely, she sells her shares, first at price p_1^B, and then at lower prices (p_2^B, p_3^B, etc.). If the depth at the five best quotes is insufficient, we assume that she can sell the remainder at a price $p_5^B - 0.01$€.

[7] For reasons of readability, we omit the subscripts d and i but this measure is computed by day and stock, as for the other proxies.

Table 2: Descriptive statistics

	October-December 2002	February-April 2006
Panel A: All orders	11,279,320	14,395,179
Buy order	5,582,655	7,040,196
	(49.49%)	(48.91%)
Sell order	5,696,665	7,354,983
	(50.51%)	(51.09%)
Limit order	10,352,871	13,422,208
	(91.79%)	(93.24%)
Market-to-limit order	454,937	341,581
	(04.03%)	(02.37%)
Market order	471,512	631,390
	(04.18%)	(04.39%)
Executed	7,010,831	8,233,070
	(62.16%)	(57.19%)
Cancelled	3,342,380	5,204,389
	(29.63%)	(36.15%)
Partially filled or expired	926,109	957,720
	(08.21%)	(06.66%)
Proprietary orders	5,372,706	8,224,750
	(47.63%)	(57.14%)
Non-proprietary orders	5,906,614	6,170,429
	(52.37%)	(42.86%)
Panel B: Proprietary orders	5,372,706	8,224,750
Buy order	2,669,268	4,003,533
	(49.68%)	(48.68%)
Sell order	2,703,438	4,221,217
	(50.32%)	(51.32%)
Limit order	5,333,641	8,199,388
	(99.24%)	(99.69%)
Market-to-limit order	3,660	2,790
	(00.07%)	(00.03%)
Market order	35,405	22,572
	(00.66%)	(00.27%)
Executed	3,232,894	3,831,451
	(60.17%)	(46.58%)
Cancelled	1,746,141	3,966,236
	(32.50%)	(48.22%)
Partially filled and expired	393,671	427,063
	(07.33%)	(05.19%)
Panel C: Other information		
Number of trades	8,078,588	10,496,571
Number of trading days	64	61

Table 2 reports some descriptive statistics about our dataset 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). Panel A reports descriptive statistics on the full sample of orders. Panel B refines these statistics by considering only proprietary orders. Panel C indicates the number of trades.

Table 3: Liquidity proxies in 2002 and 2006

Liquidity		2002	2006	% Change	t-value
Tightness	QS	0.0845	0.0492	-41.78%	24.20***
	RS	0.2200%	0.0780%	-64.55%	82.97***
CRT	1. Trade size = 1000 shares	0.3349	0.1015	-71.11%	85.33***
	2. Trade size = 2000 shares	0.4313	0.1246	-71.11%	80.76***
	3. Trade size = 3000 shares	0.5128	0.1479	-71.16%	78.49***
Depth	Displayed depth (BBO)	4,050.6	8,322.0	+105.45%	−16.14***
	Displayed depth (5BQ)	25,028.2	50,033.8	+99.91%	−14.32***
	Total depth (BBO)	8450.7	14,783.5	+74.94%	−10.91***
	Total depth (5BQ)	43,236.4	73,054.2	+68.96%	−12.87***
Immediacy	NO	5,183.5	6,940.8	+33.90%	−16.76***
	NT	3,256.5	4,293.0	+31.83%	−13.04***
Volatility	HL	1.8644	1.1385	-38.93%	21.06***
	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***

Table 3 lists our different liquidity and market activity proxies, their cross-sectional average for each period, their evolution over time, and the result of the corresponding t-test. Quoted Spread (QS) and Relative Spread (RS) are respectively equal to $p_1^A - p_1^B$ and $2(p_1^A - p_1^B)/(p_1^A - p_1^B)$, with p_1^A and p_1^B the ask price and the bid price at the best quote, respectively. Cost of round trip trade (CRT) is computed as $CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}}$ where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, with p_τ^A and p_τ^B respectively the ask and bid prices at the τ^{th} limit, and q_τ^A and q_τ^B the corresponding number of shares. Displayed depth and total depth are measured both at the best quote (BBO) and cumulated at the five best quotes (5BQ). Number of orders (NO) and Number of Trades (NT) are the daily number of orders and trades. HighLow (HL) is the difference between the minimum and the maximum of transaction prices within a day for a given stock. 'Relative' HighLow (RHL) is equal to HighLow divided by the VWAP of that particular day and expressed in percentage. VWAP is the daily volume-weighted average price. Execution-to-order ratio (EOR), cancellation-to-order Ratio (COR), and rapid cancellation-to-order ratio (RCR) are respectively computed as the number of executed orders, cancelled orders, and cancelled orders within one second, divided by the total number of orders, with and without considering the 'proprietary order' criterion. Trade size is the number of shares traded. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%.

follows:

$$EOR_{d,i} = \frac{e_{d,i}}{n_{d,i}} \tag{3}$$

$$COR_{d,i} = \frac{c_{d,i}}{n_{d,i}} \tag{4}$$

$$RCR_{d,i} = \frac{r_{d,i}}{n_{d,i}} \tag{5}$$

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; and $n_{d,i}$ the number of submitted orders, for each day d and each stock i by all the market members, with and without distinguishing agency orders from proprietary orders.[8] Ratios based on cancellations or executions are complementary with the following caveat: in the database, an order which is partially filled can be classified either as 'executed' or 'cancelled'. In order to avoid that issue, 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, $c_{d,i}$). Finally, we measure trade size as the number of shares traded.

For each proxy listed above, we build two samples of observations for 2002 and 2006 by computing an average for each stock and day, leading to 2176 observations for 2002 and 2074 observations for 2006. In Table 3, we observe that the QS and RS decrease by more than 40% and 60%, respectively. By contrast, depth increase 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 that transaction prices fluctuate within a smaller range in 2006 than in 2002. Although the number of orders increases, a larger proportion of them are 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. All these findings are significant at the 1% level and point to better liquidity in 2006 than in 2002.

We replicate the above analysis on a stock-by-stock basis in order to identify the number of stocks experiencing an increase, a decrease or no variation in each of our proxies. We use daily averages to obtain two samples with 64 and 61 observations in 2002 and 2006, respectively.[9] Table 4 reports the number of stocks experiencing either an increase or a decrease for each proxy. We confirm the previous findings at the stock level, but liquidity does not improve

[8] In the next section, we consider the $EOR_{d,i,j}$ and $COR_{d,i,j}$ for each market member j based on proprietary orders only.

[9] We follow Hendershott et al. (2011) and Riordan and Storkenmaier (2012) in computing daily averages.

Figure 1: Cost of round trip trade

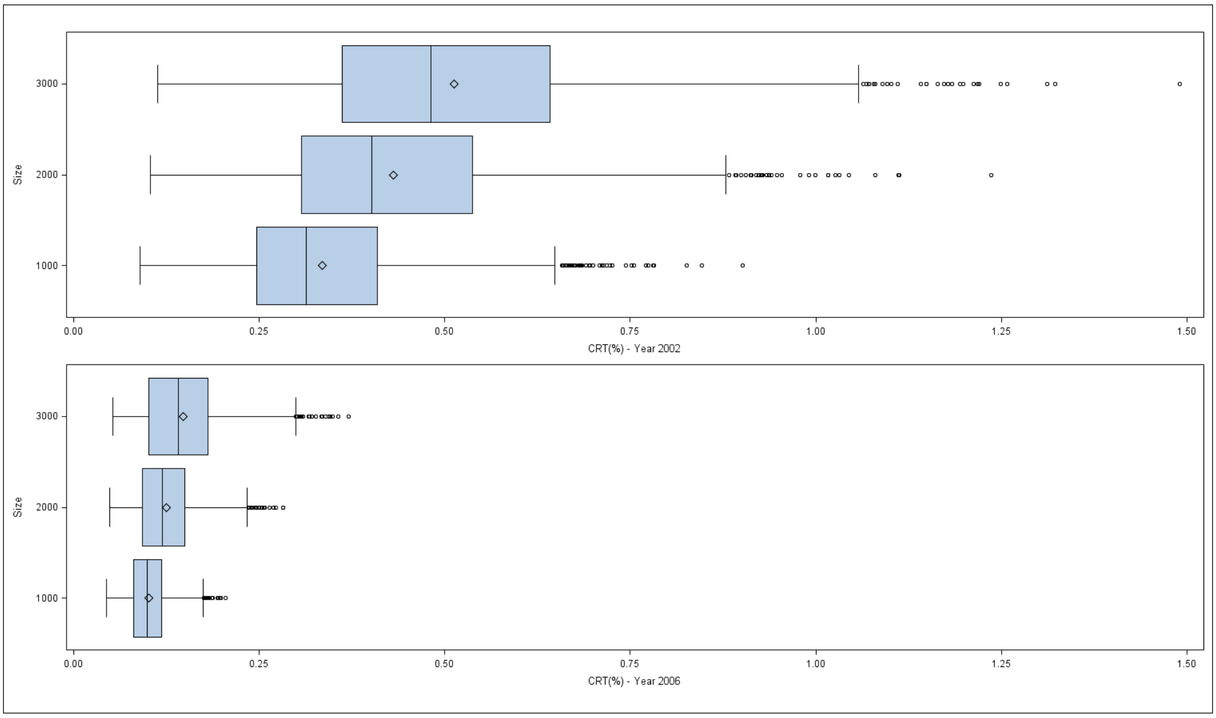

Figure 1 reports the cross-sectional average of the cost of round trip trade (CRT) in 2002 and 2006 for several trade sizes. It is computed as $CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}}$ where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, with p_τ^A and p_τ^B respectively the ask and bid prices at the τ^{th} limit, and q_τ^A and q_τ^B the corresponding number of shares.

necessarily 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, 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.

3 Methodology

The objective of our paper is to study how liquidity was impacted by the rise of fast trading on Euronext between 2002 and 2006. To that purpose, we employ the difference-in-differences method (DID method hereafter) whose first step requires the identification of fast traders active in 2006. These traders should display characteristics close to what can be expected from HFTs, i.e., they should show high COR (or low EOR), high RCR, and/or small end-of-day net positions (NP). In a second step, we separate stocks most exposed to fast trading from the other stocks.

3.1 Identification of fast traders

Our identification method relies on the analysis of trading patterns at the member level. Thanks to the ID code of each market member, we do not have to assume perfect homogeneity in trading patterns among fast traders. We are therefore able to test the sensitivity of our empirical results to the size and composition of the group of members identified as fast traders.

Table 5 lists all the market members according to their respective number of proprietary orders (NO) between February and April 2006. When focusing on the most active market members, we can 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. These members are ranked in the first ten positions in Table 5. Taken as a group and considering proprietary orders, these 10 market members generate more than 75% of the orders submitted on *each* stock. On average across stocks, they generate 85.89% of the orders. Among these 10 market members, 5 members were already active in 2002, i.e., those with ID n^o 504, 507, 517, 578, and 597. The new members correspond to the following ID codes: n^o 623, 643, 648, 4149, and 4460.

After identifying the most active members in 2006, we look at the usual suspects related to market activity in order to identify fast traders, i.e., the COR, EOR, and RCR. In Figures 2 to

Table 4: Stock-by-stock variations in liquidity proxies

	Decrease				Increase		
p-value	***	**	*	N.S	***	**	*
Liquidity proxy							
QS	31				2	1	
RS	34						
CRT (trade size = 1,000 shares)	34						
Displayed depth (BBO)		1			33		
Displayed depth (5BQ)		1			33		
Total depth (BBO)		1	1	4	26	1	1
Total depth (5BQ)		2		4	27	1	
Market activity proxy							
NO	1		1	5	25	2	
NT	2	1		6	21	2	2
HL	29	2		3			
RHL	34						
COR (All orders)	2	1		8	21	1	1
COR (Proprietary orders)					34		
EOR (All orders)	20			10	3		1
EOR (Proprietary orders)	33	1					
RCR (All orders)	1	2		6	22	2	1
RCR (Proprietary orders)				1	32	1	
Trade size	20	2	2	6	3		1

Table 4 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%. Quoted Spread (QS) and Relative Spread (RS) are respectively equal to $p_1^A - p_1^B$ and $2(p_1^A - p_1^B)/(p_1^A - p_1^B)$, with p_1^A and p_1^B the ask price and the bid price at the best quote, respectively. Cost of round trip trade (CRT) is computed as $CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}}$ where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, with p_τ^A and p_τ^B respectively the ask and bid prices at the τ^{th} limit, and q_τ^A and q_τ^B the corresponding number of shares. Displayed depth and total depth are measured both at the best quote (BBO) and cumulated at the five best quotes (5BQ). Number of orders (NO) and Number of Trades (NT) are the daily number of orders and trades. HighLow (HL) is the difference between maximum and transaction prices within a day for a given stock. 'Relative' HighLow (RHL) is equal to HighLow divided by the VWAP of that particular day and expressed in percentage. VWAP is the daily volume-weighted average price. Execution-to-order ratio (EOR), cancellation-to-order Ratio (COR), and rapid cancellation-to-order ratio (RCR) are respectively computed as the number of executed orders, cancelled orders, and cancelled orders within one second, divided by the total number of orders, with and without considering the 'proprietary order' criterion. Trade size is the number of shares traded. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%; N.S. stands for 'Not significant'.

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 and over February to April 2006 period. The new market members in 2006 are indicated by a circle while the 'old' market members (i.e. those already active in 2002) are indicated by a cross.

Table 5: Orders submitted bewteen February and April 2006

ID	Orders	ID	Orders	ID	Orders	ID	Orders	ID	Orders
504	1,423,257	29868	50,324	24984	12,882	1050	2,105	4125	72
597	1,059,677	4243	41,776	523	10,506	479	1,770	1451	43
648*	1,032,757	639	39,903	676	7,999	29645	1,498	583	42
578	851,142	521	32,518	645	7,206	29275	1,144	29101	41
624*	589,903	29005	29,367	468	6,573	29341	1,099	29183	37
4460*	516,289	1881	26,764	29028	6,351	4245	808	1946	36
517	495,688	559	21,355	1667	6,205	4066	801	577	31
643*	420,172	585	18,928	697	5,711	694	785	554	30
507	367,880	580	18,718	4239	4,730	460	602	4063	15
4149*	225,929	4120	16,437	360	4,474	1239	447	573	6
640	162,010	29561	15,667	685	3,467	482	424	1510	6
695	144,014	426	15,366	1417	3,465	4247	225	10007	6
505	138,774	536	15,135	4121	3,370	589	206	10004	5
29303	96,436	518	14,650	1207	3,176	1246	199	572	3
29002	70,959	560	14,088	4106	3,119	1669	103	499	1
29320	64,747	630	13,434	29159	2,511	452	92	1299	1
485	60,623	592	13,278	1770	2,283	683	73	1539	1

Table 5 lists all the market members' IDs with their respective number of proprietary orders, between February and April 2006. The top 10 market members are indicated in bold and $*$ indicates the new members among this top 10 list.

In Figure 2, we use the total number of orders submitted by each member on the X-axis and the average of $COR_{d,i}$ by member on the Y-axis. Across all days, stocks, and members, the average of $COR_{d,i,j}$ was 25.47% in 2006.[10] The most crowded area in Figure 2 (located in the

[10]The graph is similar when we plot the number of proprietary orders against the ratio of the total number of cancelled orders to the total number of submitted orders. In that case, we compute COR_j as $COR_j = \sum_{d=1}^{61}\sum_{i=1}^{34} c_{d,i,j} / \sum_{d=1}^{61}\sum_{i=1}^{34} n_{d,i,j}$, where $c_{d,i,j}$ and $n_{d,i,j}$ are the number of cancelled and submitted orders respectively. The figure is available upon request. As a reminder, the number of trading days in 2006 is equal to

bottom-left corner) indicates that most members submit relatively few orders and do not cancel them often. These market members are obviously not fast traders. A simple visual inspection of the graph enables us to identify four members newly active in 2006, with both high order submissions and high COR. These members correspond to the following ID: n^{o}624, 643, 648, and 4149. Most interesting is the level of their COR: they are all higher than 50%.[11] These four new members were ranked among the top 10 most active members in 2006 (Table 5). The fifth new member included in the top 10 ranking, i.e., n^{o} 4460, is unlikely to be a fast trader given its too low COR. This is confirmed below.

In Figure 3, we replace the COR by the EOR. As both measures are highly and negatively correlated, the graph is almost similar with the slight difference that the Y-axis is reversed.[12] The average of EOR (across members, stocks and days) is equal to 68.44%. Figure 3 leads to the same conclusion: the same four new market members are identified as potential fast traders.

Figure 4 displays the average of $RCR_{d,i}$ by member against their 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^{o}624, 643, 648, and 4149) can be clearly isolated: they all exhibit a RCR higher than 5%. Member n^{o}4149 is clearly the fastest member in the market, with 225,929 orders and a RCR of 35.63%.[13] Therefore, the combination of the COR and RCR leads to the same group of fast traders.

Fast traders are also known for managing closely their inventory (Benos and Sagade, 2016; Kirilenko et al., 2017). To take this trading behavior into account, we compute for each market

61 and the number of comparable stocks across both periods is equal to 34.

[11] The robustness of this visual classification is tested later. For example, we also include a new smaller exchange member in this group, i.e. member n^{o}29645. This market member also exhibits a COR higher than 50% but its number of orders submitted is slightly less than 30,000. The inclusion of this market member does not bring any change in our empirical results. We also disregard market members n^{o} 4125 and 29101, since they submit very few orders, respectively 72 and 41 orders. Needles to say, their inclusion does not make any difference.

[12] 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.

[13] If we zoom in on the very small members, we can identify two other fast traders: members n^{o}426 and n^{o}29028, with respectively 15,366 and 6,351 orders and a RCR of 29.93% and 20.20% respectively. In Section 5, we include these two small players in the group of fast traders as robustness checks. As expected, it makes no difference.

Figure 2: Cross-sectional average COR vs. number of orders

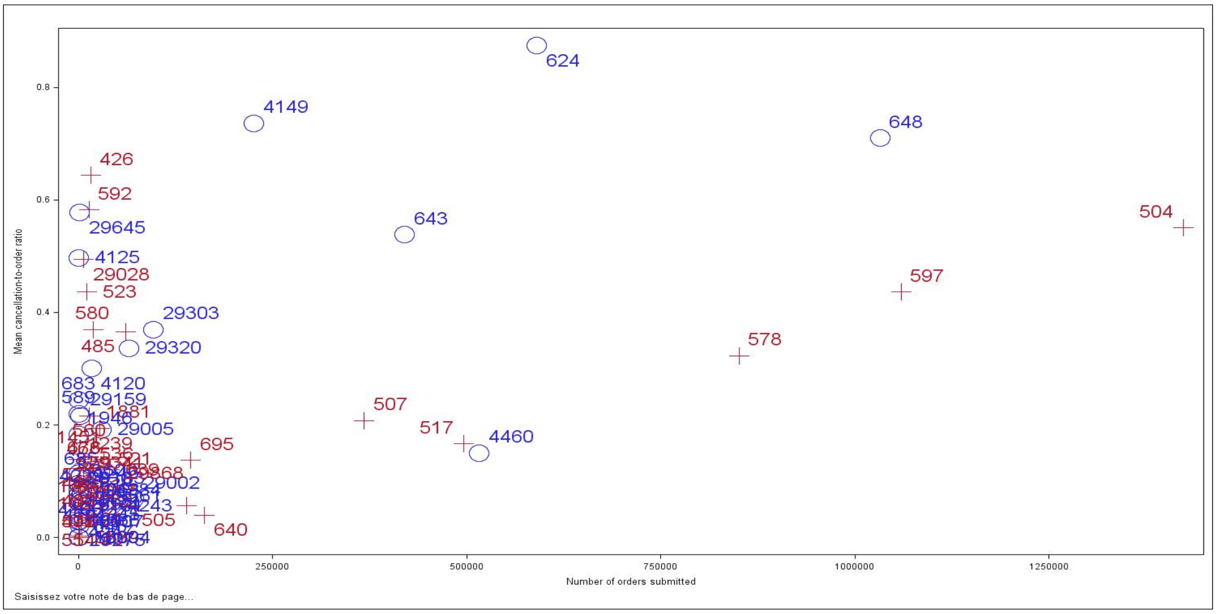

Figure 2 represents each market member j according to its number of proprietary orders submitted, n_j, and its cross-sectional average COR, with $COR_{d,i} = c_{d,i}/n_{d,i}$, where $c_{d,i}$ is the number of cancelled orders and $n_{d,i}$ is the number of submitted orders. 'New market members', i.e.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

Figure 3: Cross-sectional average EOR vs. number of orders

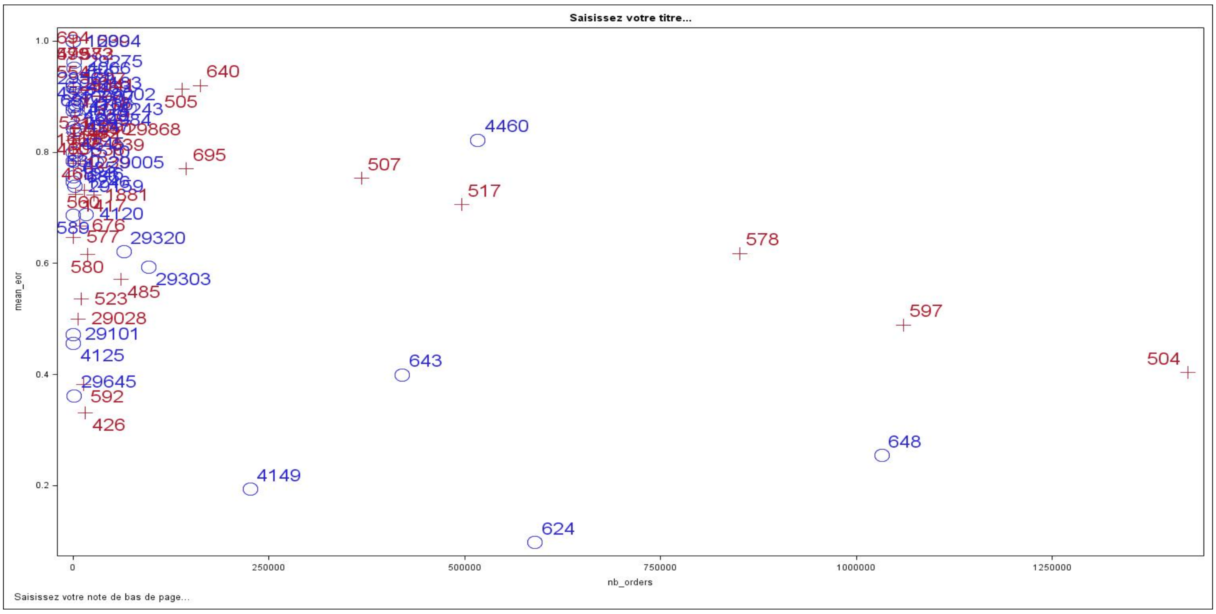

Figure 3 represents each market member j according to its number of proprietary orders submitted, n_j, and its cross-sectional average EOR, with $EOR_{d,i} = e_{d,i}/n_{d,i}$, with $e_{d,i}$, the number of executded orders, and $n_{d,i}$. 'New market members', i.e. those only present in 2006, and 'Old market members', i.e. those already present in 2002 are respectively represented with a circle and a cross.

Figure 4: Cross-sectional average RCR vs. number of orders

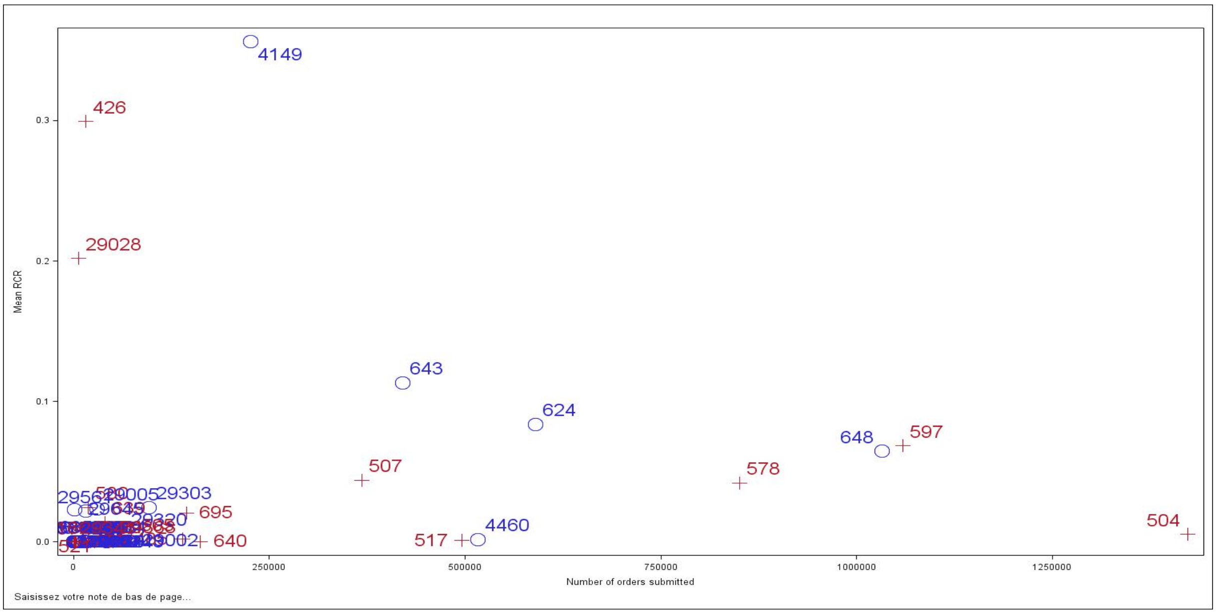

Figure 4 represents each market member j according to its number of proprietary orders submitted, n_j, and its cross-sectional average RCR, with $RCR_{d,i} = r_{d,i}/n_{d,i}$, with $r_{d,i}$, the number of orders cancelled within one second at the latest, and $n_{d,i}$. 'New market members', i.e. those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

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

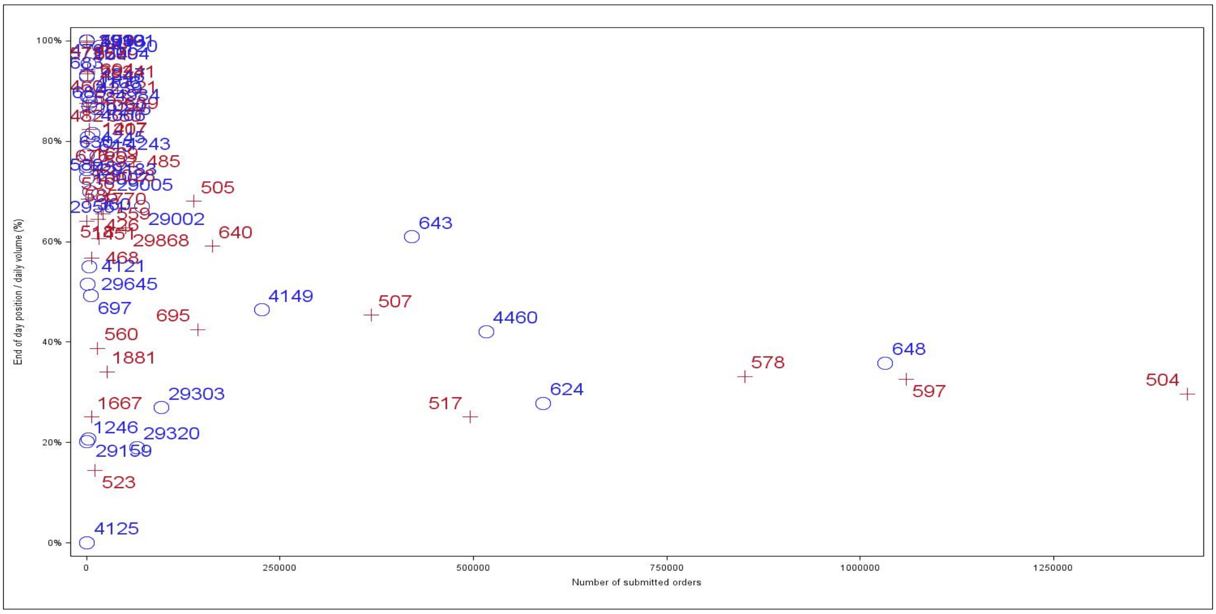

Figure 5 represents each market member j according to its number of proprietary orders submitted, n_j, and its cross-sectional average end-of-day NP, with $\frac{|Net\ position_{d,i}|}{Volume_{d,i}}$. $Net\ position_{d,i}$ is the cumulative sum of shares (weighted by a coefficient (-1 for a sell, +1 for a buy)) whereas $Volume_{d,i}$ 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 already present in 2002 are respectively represented with a circle and a cross.

member j her daily NP for each stock i. At the end of the day, we scale the net position (in absolute value) by the market member's trading volume:

$$NP_{d,i,j} = \frac{|Net\, position_{d,i,j}|}{Volume_{d,i,j}} \tag{6}$$

where $Net\ position_{d,i,j}$ is the cumulative sum of shares weighted by a coefficient (-1 for a sell, +1 for a buy), and $Volume_{d,i,j}$ is the cumulative sum of shares traded over a given day d, for a stock i, and by market member j.[14]

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 between members. We distinguish two clusters. 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 are very heterogeneous. The use of NPs does not help refine our selection of fast traders. All the big members seem to display NPs at around 40% on average. There exist members with very low NPs but they are tiny players. Most importantly, they fail to display both a low COR and a low RCR. Fast trading on Euronext in 2006 is unlikely to be characterized well by NPs.

When considering the fast trading identification criteria discussed above, we come to the conclusion that the members who behave like fast traders the most are n^{o}624, 643, 648, and 4149. All these 4 members were not active in 2002. Among them, the most serious new fast trader seems to be member n^{o}4149.

In Section 5, we carry out further robustness checks to validate the selection of these members. We show that our results remain robust to the choice of the identification criteria and the composition of the group of fast traders.

3.2 Identification of 'fast-traded' stocks

The objective of this section is to split in two groups the 34 stocks which belong to the CAC40 in both 2002 and 2006. In the first group, we include the stocks marginally affected by the fast

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

traders identified in the last section.[15] In the second group, we include the stocks which were exposed the most to fast trading in 2006.

To flag the stocks as 'fast-traded', we assume that orders are uniformly distributed across stocks. In such a case, the expected average proportion of orders for a stock is therefore equal to 1/34, which is roughly 2,95%. In Section 5, we show that the relaxation of this hypothesis does not affect our results.

Let us denote $n_{d,i,j}$, the number of submitted orders for day d, stock i, and market member j. Therefore, $n_{d,j} = \sum_{i=1}^{34} n_{d,i,j}$ 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. If $\beta_{i,j} > 2,95\%$ in the population, then stock i is actively traded by market member j.

For each of the four fast traders identified above (624, 643, 648, 4149), we check whether the null hypothesis of a uniform distribution of submitted orders is rejected at the 1% significance level. In this case, the stock is considered as being 'fast-traded'. As reported in Table 7, we obtain a matrix M_1 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.

As indicated in Table 7, we flag six stocks. Among them, two stocks (4161 and 4462) are actively traded by each of the four fast traders. The four other stocks (4157, 4187, 26990, and 29512) are actively traded by three of them.

Robustness checks about the selection of 'fast-traded' stocks are available in Section 5. In the next section, we answer the question of whether the liquidity of stocks most exposed to fast trading is better or worse than the liquidity of the other stocks.

[15] Although these stocks should be completely immune from fast trading in an ideal world, even purists in finance recognize the need for some tolerance since randomized control trials are arguably extremely difficult to conduct in finance.

4 Being a fast-traded stock: A curse of a blessing?

We use a DID specification to test the effect of fast trading on stock liquidity. DID estimation combines a control group with the treated sample to 'difference out' confounding factors and isolate the effect of an event. This methodology is applied in the HFT literature in Hendershott and Moulton (2011), Hendershott et al. (2011), or Kwan and Philip (2015), among others.

As explained in Section 2, we first construct a set of daily averages of liquidity metrics to obtain 4250 observations, i.e. one observation per stock and per trading days with 34 stocks, 64 days in 2002 and 61 days in 2006. We regress the RS and CRT on a set of treatment indicators that includes a dummy variable picking out the period (with $D1_d = 1$ if day d belongs to year 2006, and zero otherwise), a dummy variable picking out the treated group of 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' and day d is in 2006). After controlling for time series fluctuations and differences across the treated and control groups, we capture the difference in liquidity between the two groups of stocks due to the treatment of fast trading by the coefficient a_3 in the equation below:

$$L_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + \gamma X_{d,i} + \epsilon_{d,i} \tag{7}$$

where $L_{d,i}$ is a liquidity proxy (either the RS or the CRT), and $X_{d,i}$ is a matrix of control variables as defined in Table 6.

Table 6: Liquidity and control variables

Dimension	Liquidity Variables	Notation
Tightness	Relative spread	RS
	Cost of a round trip trade	CRT
Dimension	**Control Variables**	**Notation**
Size	Market capitalization	MC
Market Activity	Number of orders	NO
	Number of trades	NT
Volatility	HighLow	HL
	'Relative' HighLow	RHL

Table 6 reports the liquidity and control variables used in regressions 9 and 10. Relative Spread (RS) is equal to $2(p_1^A - p_1^B)/(p_1^A - p_1^B)$, with p_1^A and p_1^B the ask price and the bid price at the best quote, respectively. Cost of round trip trade (CRT) is computed as $CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}}$ where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, with p_τ^A and p_τ^B respectively the ask and bid prices at the τ^{th} limit, and q_τ^A and q_τ^B the corresponding number of shares. Number of orders (NO) and Number of Trades (NT) are the daily number of orders and trades. HighLow (HL) is the difference between maximum and transaction prices within a day for a given stock. 'Relative' HighLow (RHL) is equal to HighLow divided by the VWAP of that particular day. VWAP is the daily volume-weighted average price.

4.1 Results

In Table 8, we use the RS as the dependent variable and we disregard the control variables to more clearly show later that their inclusion does not modify our conclusions. The group of fast traders include four members (n^o 624, 643, 648, and 4149) and the group of fast-traded stocks include six stocks (4157, 4161, 4187, 4462, 26990, and 29512). We estimate the DID regression, but we also estimate the variation of liquidity across time for the two groups of stocks using time-fixed effects in the 'Dummy Time' (DT) regressions. Using group-fixed effects, we also estimate the variation in liquidity between the two groups of stocks in both 2002 and 2006.

Table 7: Fast traders and actively traded stocks - Hypothesis of uniform distribution

Stock / ID	426	504	505	507	517	578	597	**624**	640	**643**	**648**	695	**4149**	4460	29002	29028	29303	29645
223								1										
4154	1	1			1		1				1	1		1				
4157	1	1	1	1	1		1	1	1	1		1	1				1	
4161		1	1	1	1	1		1	1	1	1	1	1	1			1	
4166	1									1			1					
4170							1											
4180		1	1		1	1					1	1				1		
4181		1		1	1						1		1	1				
4187	1	1	1	1	1	1	1	1	1	1	1	1		1	1		1	
4188					1			1				1		1			1	
4213													1					
4245		1	1	1	1		1		1	1	1	1		1	1			
4252	1							1										
4292													1					
4322	1										1							
4351							1									1		
4438			1	1			1		1									
4462	1				1			1		1	1	1	1	1	1		1	
26990	1	1	1		1		1	1		1		1	1		1	1	1	1
29512		1					1	1		1		1	1	1				
29636				1		1		1							1			
36064	1	1	1	1	1		1	1	1		1	1		1	1		1	
44540								1										
45057						1		1										
49388			1				1		1		1		1					
49471								1			1							
72275	1	1			1		1				1	1		1	1			

Table 7 indicates whether a market member j (identified with its ID in the colum) actively trade a stock i (identified with its ID in the rows) assuming that orders are uniformly distributed across stocks. In such a case, the expected average proportion of orders for a stock is therefore equal to 1/34, which is roughly 2,95%. We denote $n_{d,i,j}$, the number of submitted orders for day d, stock i, and by market member j. $n_{d,j} = \sum_{i=1}^{34} n_{d,i,j}$ is the total number of orders that market member j submit on day d. Dividing $n_{d,i,j}$ by $n_{d,j}$, we obtain $b_{d,i,j}$, i.e., 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 for stock i and market member j in 2006. If $\beta_{i,j} > 2,95\%$, then stock i is actively traded by market member j. A value of 1 indicates that the stock i is actively traded by market member j at a 1% significance level. See Section (3.2) for more details. The group of fast traders is identified by their IDs in bold.

If we first look at the DT regression in Panel 2A, we see that the average RS in 2002 for the stocks most exposed to fast trading is estimated at 16.3 basis points, i.e., around 7 basis points tighter than the RS for the other stocks, estimated at 23.2 basis points in Panel 2B. More interesting is the tightening of the spread which occurs between 2002 and 2006 for both groups of stocks. The variation is nevertheless less pronounced in the case of the fast-traded stocks with a decrease by 9.1 basis points only, versus 15.3 basis points for our control group. Still, the fast-traded stocks remain slightly more liquid than the control stocks, with the RS being approximately equal to 7 and 8 basis points respectively. This is confirmed if we look at the two DG regressions. In Panel 3B, we see that the average RS in 2002 was lower by 6.9 basis points in favor of the stocks fast-traded in 2006. Four years later, this liquidity 'edge' falls down to only 0.7 basis point in 2006, as indicated in Panel 3A.[16]

At this stage, we can conclude that the fast-traded stocks are more liquid in both 2002 and 2006. However, the DID regression shows that the treatment of fast trading in 2006 led to a widening of RS by 6.2 basis points on average by stock. In other words, the stocks traded by fast traders in 2006 could have maintained their liquidity differential of 6.9 basis points observed in 2002, had they been better immune from fast trading.

In the first row of Table 9 in bold, we replicate the results previously obtained for the DID equation while adding the control variables, i.e., volatility (RHL), market cap (MC), and market activity (NO).[17] The market members' ID of the fast traders are reported in column (2). As indicated in column (3), we use the hypothesis of a uniform distribution of submitted orders across stocks, i.e., 1/34 by stock. Column (4) indicates the required coverage ratio, i.e., the minimum number of fast traders who must trade the stock actively for the stock to be flagged as 'fast-traded'. In column (5), we display the IDs of the fast-traded stocks and their number in column (6).

[16]The coefficient of the interaction dummy $D3_{d,i}$ is 6.213 basis points. It is the 'difference-in-differences' coefficient because it is precisely equal to the difference between the two $D1_d$ coefficients in the DT regressions or to the difference between the two $D2_i$ coefficients in the DG regressions.

[17]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%. Results are available upon request.

Table 8: Difference-in-differences equation for the relative spread

	Constant	$D1_d$	$D2_i$	$D3_{d,i}$	N	*Adj.*R^2
Panel 1: Whole sample						
DID Regression	0,23224	-0,15299	-0,06920	0,06213	4250	65,14%
	***	***	***	***		
Panel 2A: Treated group of stocks						
DT Regression	0,16304	-0,09086			750	65,55%
	***	***				
Panel 2B: Control group of stocks						
DT Regression	0,23224	-0,15299			3500	64,10%
	***	***				
Panel 3A: 2006 Period						
DG Regression	0,07925		-0,00707		2074	2,14%
	***		***			
Panel 3B: 2002 Period						
DG Regression	0,23224		-0,06920		2176	11,51%
	***		***			

Table 8 reports the regression results of the following DID regression: $Liquidity_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + \epsilon_{d,i}$, which corresponds to Equation (7) without the control variables. The group of fast traders includes members 624, 643, 648, 4149) and the group of treated stocks includes 6 stocks (4161, 4462, 4157, 4187, 26990, and 29512). The minimum coverage ratio by fast traders is 3 (out of 4). Dummy Time (DT) regressions capture the variation of liquidity across time for each group of stocks, taken separately. They correspond to: $Liquidity_{d,i} = a_0 + a_1 * D1_d + \epsilon_{d,i}$. Dummy Group (DG) regressions capture the variation in liquidity between the two groups at each time period. They correspond to: $Liquidity_{d,i} = a_0 + a_2 * D2_i + \epsilon_{d,i}$. $Liquidity_{d,i}$ is captured by the relative spread $RS_{d,i}$ in this table. p-values are reported under the coefficient. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. Standard errors are robust to heteroskedasticity.

Table 9: Difference-in-differences equations: Effect of fast trading on relative spread

Market member Group	MM ID	Stocks Hyp.	Cov.	Stock ID	N	$Constant$	$D1_d$	$D2_i$	$D3_{d,i}$	$RHL_{d,i}$	$MC_{d,i}$	$NO_{d,i}$	Adj. R^2
1	**624, 643 648, 4149**	**1/34**	**3/4**	**4161, 4462, 4157, 4187, 26990, 29512**	**6**	**0,20743** ***	**-0,09848** ***	**-0,04172** ***	**0,07319** ***	**0,01217** ***	**-2,28E-13** ***	**-0,00000760** ***	**74,16%**
		1/34	4/4	4161, 4462	2	0,20122 ***	-0,08911 ***	-0,03564 ***	0,07315 ***	0,01243 ***	-1,9547E-13 ***	-0,00000778 ***	72,66%
		MB	3/4	4252, 4322, 4462 26990, 29512	5	0,20325 ***	-0,09259 ***	-0,03630 ***	0,05135 **	0,01252 ***	-1,58E-13 ***	-0,00000776 ***	72,97%
		MB	4/4	4252	1	0,19873 ***	-0,08467 ***	-0,02553 ***	0,01631 ***	0,01270 ***	-1,4552E-13 ***	-0,00000800 ***	71,91%
8	504, 578 597, 624 643, 648 4149	1/34	6/7	4161, 4187	2	0,19964 ***	-0,08823 ***	-0,03709 ***	0,06628 ***	0,01269 ***	-1,56E-13 ***	-0,00000786 ***	72,53%
		1/34	5/7	4161, 4187 4157, 26990, 29512	5	0,20577 ***	-0,09628 ***	-0,04281 ***	0,07067 ***	0,01221 ***	-1,83E-13 ***	-0,00000767 ***	73,76%
		1/34	4/7	4161, 4187, 4157, 26990, 29512, 4245, 4462, 36064	8	0,20611 ***	-0,09977 ***	-0,03900 ***	0,06923 ***	0,01276 ***	-2,47E-13 ***	-0,00000753 ***	74,38%
		MB	4/7	4166, 4252	2	0,19962 ***	-0,08615 ***	-0,02563 ***	0,02418 ***	0,01263 ***	-1,14E-13 ***	-0,00000806 ***	72.00%

Table 9 reports regression results of the following Equation: $L_{d,i} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{d,i} + b_1 * RHL_{d,i} + b_2 * MC_{d,i} + b_3 * NO_{d,i} + \epsilon_{i,d}$. In this table, $L_{d,i}$ is the relative spread ($RS_{d,i}$) on day d and stock i. Relative Spread (RS) is equal to $2(p_1^A - p_1^B)/(p_1^A - p_1^B)$, with p_1^A and p_1^B the ask price and the bid price at the best quote, respectively. In the first and second column, we respectively identify the group of fast traders and its composition. 'Hyp.' refers to the assumption used to identify actively traded stocks. In the uniform distribution (denoted as '1/34'), we compare the 2006 time-series average of $b_{d,i,j}$ representing the realized proportion of submitted orders for stock i and market member j in 2006 to 1/34, consistent with an equally-weighted proportion of orders submitted across stocks. In the 'market-based' distribution (denoted as 'MB'), we compare $b_{d,i,j}$ with the time-series average of $x_{d,i}$, $\tilde{X}_i$. See Section (3.2) for more details. 'Cov' refers to (at least) how many market members, among those identified in column 2, actively trade the stocks. Stock ID is the stock's identification number, and N is the number of stocks which are actively traded. a_0 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 therefore mentionned in column (5)), and zero otherwise); $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$); 'Relative' HighLow (RHL) is equal to HighLow divided by the VWAP of that particular day. VWAP is the daily volume-weighted average price.; $MC_{d,i}$ is the stock's daily market capitalization; $NO_{d,i}$ is the number of orders submitted on day d and stock i. The last column reports the adjusted R-squared. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. Standard errors are robust to heteroskedasticity.

Our coefficient of interest is related to the $D3_{d,i}$ dummy. It is estimated at 7.3 basis points while it was 6.9 basis points without the control variables. The level of significance is unchanged, with a p-value below 1%, but the adjusted R-squared is 74%, i.e., 9 p% higher than before. We also see that all the estimated coefficient signs for the control variables are consistent with expectations, i.e., a positive relation between liquidity and RHL, and a negative relation with MC and NO. In 2002, the average RS for the control group of stocks in 2006 is estimated at 20.7 basis points (versus 23.2 before). For those stocks, the decreases in the RS between 2002 and 2006 is estimated at 9.8 basis points (versus 15.3 before). Also, the average RS in 2002 is estimated to be 4.2 basis points wider for the control group of stocks than for the fast-traded stocks (versus 6.9 before). It is noteworthy that the liquidity 'edge' in 2002 in favor of fast-traded stocks is reversed in 2006: it goes down by 7.3 basis points (versus 6.2 before) such that the RS in 2006 become 3.1 basis points wider for fast-traded stocks (versus 0.7 tighter before).

In the first row of Table 10, we replicate the same analysis but we replace the RS by the CRT with a trade size equal to 1,000 shares. Again, the coefficient of $D3_{d,i}$ is positive and statistically significant at 1%. Being fast traded in 2006 is a curse since it increases the CRT by 12.9 basis points on average. We also see that the liquidity 'edge' that those stocks benefited from in 2002, equal to 8 basis points, is reversed. In 2006, the CRT turns out to be 4.9 basis points *wider* for the fast-traded stocks (than for the other stocks). This is economically very meaningful since the average CRT for the fast-traded stocks was 22.4 basis points in 2006. More than half of the CRT in 2006 (i.e. 12.9 divided by 22.4) is explained by the fact that these stocks became fast-traded.

Table 10: Difference-in-differences equations: Effect of fast trading on CRT

Market member Group	MM ID	Hyp.	Cov.	Stocks Stock ID	N	$Constant$	$D1_d$	$D2_i$	$D3_{d,i}$	RHL	MC	NO	Adj. R^2
1	**624, 643 648, 4149**	**1/34**	**3/4**	**4161, 4462, 4157, 4187, 26990, 29512**	**6**	**0,35834** ***	**-0,18312** ***	**-0,08030** ***	**0,12869** ***	**0,01326** ***	**-3,83E-13** ***	**-0,00001383** ***	**76,16%**
		1/34	4/4	4161, 4462	2	0,34734 ***	-0,16521 ***	-0,05876 ***	0,12166 ***	0,01373 ***	-3,9623E-13 ***	-0,000001419 ***	74,14%
		MB	3/4	4252, 4322, 4462 26990, 29512	5	0,35068 ***	-0,16925 ***	-0,06155 ***	0,07326 ***	0,0139 ***	-3,22E-13 ***	-0,00001418 ***	74,37%
		MB	4/4	4252	1	0,34354 ***	-0,15767 ***	-0,04749 ***	0,02339 ***	0,01417 ***	-3,1409E-13 ***	-0,00001457 ***	73,41%
8	504, 578 597, 624 643, 648 4149	1/34	6/7	4161, 4187	2	0,34566 ***	-0,16595 ***	-0,08221 ***	0,14241 ***	0,01413 ***	-3,21E-13 ***	-0,00001426 ***	74,55%
		1/34	5/7	4161, 4187 4157, 26990, 29512	5	0,35646 ***	-0,18017 ***	0,08388 ***	0,13207 ***	0,01325 ***	-3,44E-13 ***	-0,00001393 ***	75,92%
		1/34	4/7	4161, 4187, 4157, 26990, 29512, 4245, 4462, 36064	8	0,35625 ***	-0,18754 ***	-0,08199 ***	0,12768 ***	0,01428 ***	-4,07E-13 ***	-0,00001338 ***	76,84%
		MB	4/7	4166, 4252	2	0,34459 ***	-0,15962 ***	-0,042 ***	0,03169 ***	0,0141 ***	-2,54E-13 ***	-0,0000147 ***	73,46%

Table 10 reports regression results of the following Equation: $L_{i,d} = a_0 + a_1 * D1_d + a_2 * D2_i + a_3 * D3_{i,d} + b_1 * RHL_{d,i} + b_2 * MC_i + b_3 * NO_{d,i} + \epsilon_{i,d}$. In this table, $L_{d,i}$ is the cost of round trip trade ($CRT_{d,i}$) on day d and stock i. Cost of round trip trade (CRT) is computed as $CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}}$ where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, with p_τ^A and p_τ^B respectively the ask and bid prices at the τ^{th} limit and q_τ^A and q_τ^B the corresponding number of shares. In the first and second column, we respectively identify the group of fast traders and its composition. 'Hyp.' refers to the assumption used to identify actively traded stocks. In the uniform distribution (denoted as '1/34'), we compare the 2006 time-series average of $b_{d,i,j}$ representing the realized proportion of submitted orders for stock i and market member j in 2006 to 1/34, consistent with an equally-weighted proportion of orders submitted across stocks. In the 'market-based' distribution (denoted as 'MB'), we compare $b_{d,i,j}$ with the time-series average of $x_{d,i}$, $\tilde{X}_i$ (see Section (3.2) for more details). 'Cov' refers to (at least) how many market members, among those identified in column 2, actively trade the stocks. Stock ID is the stock's identification number, and N is the number of stocks which are actively traded. a_0 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 therefore mentionned in column (5)), and zero otherwise); $D3_{d,i}$ is the interaction dummy (with $D3_{d,i} = D2_i * D1_d$); 'Relative' HighLow (RHL) is equal to HighLow divided by the VWAP of that particular day. VWAP is the daily volume-weighted average price.; $MC_{d,i}$ is the stock's daily market capitalization; $NO_{d,i}$ is the number of orders submitted on day d and stock i. The last column reports the adjusted R-squared. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. Standard errors are robust to heteroskedasticity.

In Section 5, we show that our results remain robust to changes in the methods used to identify fast traders and fast-traded stocks.

5 Robustness checks

We divide our robustness checks in two main parts. In the first part, we measure the sensitivity of our empirical results to the method used to identify fast traders. We show that there was no evidence of fast trading in 2002; market members already active in 2002 behave very differently from the fast traders identified previously. Our empirical results are unchanged when different combinations of fast traders are adopted. In the second part, we show that the empirical results are not sensitive to the distributional hypothesis regarding the submission of orders between stocks. Changes in the coverage ratio do not modify the empirical results. Our findings are even stronger for the CRT which increases due to fast trading irrespective of the trade size.

Is there evidence of fast trading in 2002?

Our empirical results may be invalidated by the early existence of fast trading in 2002. To test whether some fast traders could have been active in 2002, we have to study the trading behavior of the 44 market members which were already active during that year. We call them the 'old' market members. Using our proxies of market activity (EOR, COR, and RCR), we compute one ratio per day, per stock and per market member, leading to 41,255 and 40,998 cross-sectional observations in 2002 and 2006, respectively. After verifying the equality of variances, we perform three tests of means between both periods. Results are presented in Table 11. As expected, EOR decreases from 74.18% to 72.18% while COR and RCR increase from 18.14% to 21.02% and from 1.10% to 1.65% respectively (Panel A). These variations are not economically very significant, although they are all highly statistically significant with p-values lower than 1% in each case. As expected, these 'old' market members do not behave like fast traders. On average, they have too high EORs and too low CORs and RCRs in 2002. This holds true in 2006 as well.

Do the 'old' market members trade differently from the 'new' members and fast traders?

In comparison to the 'new' market members, i.e. those only active in 2006, the 'old' market members exhibit significantly higher EORs, lower CORs and RCRs on average, as indicated in Panel B of Table 11. In particular, the 'new' market members exhibit on average an EOR of

Table 11: Market members' trading behaviour

Panel A: Old market members in 2002 vs. Old market members in 2006

	Old in 2002	Old in 2006	t-value
EOR	74.18%	72.18%	10.50***
COR	18.14%	21.02%	−16.65***
RCR	1.10%	1.65%	−14.45***
N	41,255	40,998	

Panel B: New market members in 2006 vs Old market members in 2006

	New in 2006	Old in 2006	t-value
EOR	62.13%	72.18%	−36.98***
COR	32.98%	21.02%	45.85***
RCR	5.39%	1.65%	45.91***
N	24,287	40,998	

Panel C: Fast traders in 2006 vs. Old market members in 2006

	Fast traders in 2006	Old in 2006	t-value
EOR	23.64%	72.18%	−177.86***
COR	71.46%	21.02%	180.92***
RCR	15.13%	1.65%	76.54***
N	8,148	40,998	

Table 11 reports t-test results on the comparison of three market activity proxies, i.e. the cancellation-to-order ratio (COR), the execution-to-order ratio (EOR), and the rapid cancellation-to-order ratio (RCR). We test whether the trading behaviour of 'old market members' in 2006 is different from (a) the trading behaviour of 'old market members' in 2002, (b) 'new market members' in 2006, (c) the group of fast traders (Market members n^o 624, 643, 648, and 4149 - as identified in Table 12). We compute for each day d, for each stock i and for each market member j its $COR_{d,i,j}(= c_{d,i,j}/n_{d,i,j})$, $EOR_{d,i,j}(= e_{d,i,j}/n_{d,i,j})$, and $RCR_{d,i,j}(= r_{d,i,j}/n_{d,i,j})$, with $e_{d,i,j}$, the number of executed orders, $c_{d,i,j}$, the number of cancelled orders, $r_{d,i,j}$ the number of orders cancelled one second at the latest after their submission, and $n_{d,i,j}$ the number of proprietary orders. N is the number of observations of the sample. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%.

62.13% versus 72.18% for the 'old' market members. The average COR is 32.98% for the 'new' members and 21.02% for the 'old' members. Most importantly, the RCR for 'new' members is more than triple the RCR of 'old' members, i.e., 5.39% against 1.65% respectively. For the three tests of means, we can reject the null hypotheses that both groups exhibit the same market activity proxies, with a p-value lower than 1% in each case. In Panel C, the difference between the 'old' market members and the group of fast traders is even stronger, as expected.

Do different combinations of fast traders make any difference?

Table 12 indicates the different alternative groups of potential fast traders that we consider. Group 1 includes the four members initially identified as fast traders. We then vary the selection criteria. For example in Group 2, we only include new members on condition that they belong to the top-10 list ranked by order submission in 2006. This leads to the inclusion of member n^{o}4460. In Group 8, we combine Group 1 with all the members included in the top-5 list ranked by order submission in 2006. This leads to the inclusion of three members already active in 2002, i.e., members n^{o}504, 578, 497. It is noteworthy that the four members identified as fast traders initially (Group 1, in bold), are always included in the alternative groups.

For the sake of brevity, we only report the results for two groups in Tables 9 and 10, i.e. Group 1 (already studied in Section 4) and Group 8. Our results are robust to misclassification risk. The estimated coefficient for $D3_{d,i}$ remains always positive and significant. For example in Table 9 when the hypothesis of uniform distribution is kept unchanged, the estimations for Group 8 are between 0.06628 and 0.07067. Across all groups (1 to 14), the minimum estimation for this coefficient is 0.04542 and its maximum estimation is 0.08271, all estimations being statistically significant at 1%. Regarding the CRT in Table 10, the estimations for Group 8 are between 0.12768 and 0.14241. Across all groups (1 to 14), the minimum estimation for this coefficient is 0.09061 and its maximum estimation is 0.14560, all estimations being statistically significant at 1%.[18]

Is the uniform distribution hypothesis too simplistic?

We also relax the assumption of a uniform distribution of submitted orders across stocks. For each stock i, we instead compare the submission of market member j relative to the aggregate

[18]Results are available upon request.

submission of all market members. To do so, we divide the number of orders submitted on day d for stock i ($n_{d,i}$) by the total number of orders submitted on day d (n_d). We obtain a ratio $x_{d,i}(= n_{d,i}/n_d)$, which is the proportion of orders submitted on day d for stock i by all market members. We compute $\tilde{X}_i$, i.e. the time-series average of $x_{d,i}$. We replace the 1/34 by $\tilde{X}_i$. We refer to this hypothesis as the 'market-based' (MB) distribution hypothesis.[19]

Considering Group 1 in Tables 9 and 10, we find that the list of flagged stocks remains relatively stable. Let us take an example in Table 9 for Group 1 when the coverage ratio is 3/4. In comparison with the list of stocks flagged by the members of Group 1 under the null hypothesis of a uniform distribution (1/34 versus MB), we observe that stocks 4462, 26990, and 29512 are fast-traded under both distributional assumptions. Stocks 4157, 4161, and 4187 are nevertheless replaced by stocks 4252 and 4322. The estimated coefficient for $D3_{d,i}$ remains always positive and significant. Results are even stronger for the CRT in Table 10. We proceed in the same way for the other groups by considering the two distributional hypotheses.

Do changes in the coverage ratio matter?

In Tables 9 and 10, we also vary the coverage ratio for each group of fast traders, taking the two distributional hypotheses of submitted orders into account. No significant variation in the estimated coefficient for $D3_{d,i}$ is observed. Again, empirical findings for the CRT are even stronger. Being most exposed to fast trading leads to a increase in the CRT, for a trade size of 1,000 shares, by a minimum value of 2.3 basis points and a maximum value of 14.2 basis points, as reported in Table 10.

Does the effect of fast trading depend on trade size?

Table 13 provides an even more comprehensive analysis of the variations in CRT between 2002 and 2006. We consider six different trade sizes and replicate Table 10 by considering the same two groups of fast traders (Groups 1 and 8), the two distributional hypotheses, and the same changes in the coverage ratio. The estimated coefficient for $D3_{d,i}$ remains always positive and significant, between a minimum of 2 basis points for a trade size of 500 shares and a maximum of 25 basis points for a trade size of 3000 shares.

[19]The matrix M_2 of binary values with the null of a 'market-based' distribution, equivalent to Table 7, is available upon request.

Table 12: Group compositions

Groups	Market Members	Criteria
1	**624, 643, 648, 4149**	Only in 2006, RCR >5%, and COR >50%
2	**624, 643, 648, 4149**, 4460	Only in 2006 and top-ten volume
3	**624, 643, 648, 4149**, 4460, 29002, 29303	Only in 2006 and top-fifteen volume
4	**624, 643, 648, 4149**	Only in 2006 and RCR >5%
5	**624, 643, 648, 4149**, 29645	Only in 2006 and COR >50%
6	**624, 643, 648, 4149**	Only in 2006 and top-ten volume and RCR >5%
7	**624, 643, 648, 4149**	Only in 2006 and top-ten volume and COR >50%
8	504, 578, 597, **624, 643, 648, 4149**	Group 1 and Top-five volume (ex-group 5)
9	504, 507, 517, 578, 597, **624, 643, 648, 4149**, 4460	Top-ten volume
10	504, 505, 507, 517, 578, 597, **624**, 640, **643, 648**, 695, **4149**, 4460, 29002, 29303	Top-fifteen volume
11	426, 597, **624, 643, 648, 4149**, 29028	RCR >5%
12	426, 597, **624, 643, 648, 4149**, 29028	COR >50%
13	597, **624, 643, 648, 4149**	Top-ten volume and RCR >5%
14	597, **624, 643, 648, 4149**	Top-ten volume and COR >50%

Table 12 lists all the combinations of criteria used to identify the group of fast traders.

Table 13: Robustness: CRT size

	$D1_d$			$D2_i$			$D3_{d,i}$		
Size	Min.	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.
500	-0,1470	-0,1346	-0,1241	-0,0645	-0,0524	-0,0345	0,0200	0,0782	0,1070
1000	-0,1875	-0,1711	-0,1577	-0,0839	-0,0673	-0,0420	0,0234	0,0976	0,1424
1500	-0,2241	-0,2041	-0,1881	-0,1011	-0,0798	-0,0472	0,0242	0,1142	0,1743
2000	-0,2559	-0,2327	-0,2146	-0,1178	-0,0904	-0,0503	0,0229	0,1282	0,2027
2500	-0,2833	-0,2574	-0,2373	-0,1321	-0,0990	-0,0523	0,0210	0,1397	0,2274
3000	-0,3062	-0,2781	-0,2565	-0,1440	-0,1060	-0,0534	0,0192	0,1491	0,2482

Table 13 reports summary statistics about the regression coefficient of $D1_d$, $D2_i$, and $D3_{d,i}$. In accordance with Tables 9 and 10, we vary the composition of fast traders (Group 1 and Group 8), the underlying assumption about the distribution of submitted stocks (1/34 and 'market-based' (MB)), and the coverage ratio. For each size, we obtain 8 regressions and we report the minimum, mean, and maximum value of each coefficient. Cost of round trip trade (CRT) is computed as $CRT(size) = \frac{\sum_{\tau=1}^{5}(q_\tau^A * p_\tau^A) - \sum_{\tau=1}^{5}(q_\tau^B * p_\tau^B)}{size * \frac{p_1^A + p_1^B}{2}}$ where $\sum_{\tau=1}^{5}(q_\tau^A) = \sum_{\tau=1}^{5}(q_\tau^B) = size$, with p_τ^A and p_τ^B respectively the ask and bid prices at the τ^{th} limit and q_τ^A and q_τ^B the corresponding number of shares.

6 Conclusion

Using two three-month periods at the end of 2002 and the start of 2006, we study the rise of fast trading on Euronext. We observe an overall improvement in terms of liquidity between 2002 and 2006. On average, spreads are tighter, depths in the order book are higher, and there are lower intraday price fluctuations. However, market members cancel a higher percentage of orders, especially when they trade on a proprietary basis. They also trade smaller quantities.

We study the effect of fast trading on the liquidity of the CAC40 stocks by trying to identify new market members in 2006 with higher speed of execution, lower lifetime of orders, and tighter inventories. Thanks to the ID codes of market members, we compute their cancellation-to-order ratio, rapid cancellation-to-order ratio, and end-of-day net positions. Stocks are then separated in two groups. We identify the stocks most exposed to fast trading and use a diff-in-diff approach. We show that the 'fast-traded' stocks exhibit the weakest decrease in both the relative spread and the cost of round trip trade.

Although the market environment on Euronext between 2002 and 2006 benefited both groups of stocks, the 'slow-traded' stocks took the most out of it. Being a fast-traded stock in 2006 was economically quite detrimental. In one of our base case regressions, the CRT in 2006 turns out to be on average 4.9 basis points *wider* for the fast-traded stocks than for the other stocks. This is in spite of the fact that the fast-traded stocks enjoyed a liquidity 'edge' in 2002, before the rise of fast trading. Considering an average CRT for the fast-traded stocks equal to 22.4 basis points in 2006, we estimate that 12.9 basis points, i.e., more than half if the CRT, can be explained by the fact that these stocks became fast-traded.

All the different robustness checks lead to the same conclusion. The stocks which became most exposed to fast trading over time could have maintained their liquidity 'edge' observed in 2002, had they been better immune from fast trading. As in Ye et al. (2013), Menkveld and Zoican (2017), or Brogaard et al. (2017), our findings question the positive effects typically associated with fast trading.

References

Baron, M., J. Brogaard, and A. Kirilenko (2012). The trading profits of high frequency traders. *Unpublished Manuscript*.

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

Boehmer, E., K. Fong, and J. Wu (2014). International evidence on algorithmic trading, singapore management university. Technical report, unpublished working paper.

Boehmer, E., C. M. Jones, and X. Zhang (2013). Shackling short sellers: The 2008 shorting ban. *The Review of Financial Studies 26*(6), 1363–1400.

Brogaard, J., B. Hagströmer, L. Nordén, and R. Riordan (2015). Trading fast and slow: Colocation and liquidity. *The Review of Financial Studies 28*(12), 3407–3443.

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. *The 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.

De Winne, R. and C. D'Hondt (2007). Hide-and-seek in the market: placing and detecting hidden orders. *Review of Finance 11*(4), 663–692.

D'Hondt, C., C. Majois, and P. Mazza (2015). Commonality on euronext: Do location and account type matter? *International Review of Financial Analysis 42*, 183–198.

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

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

Friederich, S. and R. Payne (2015). Order-to-trade ratios and market liquidity. *Journal of Banking & Finance 50*, 214–223.

Frino, A., V. Mollica, and R. I. Webb (2014). The impact of co-location of securities exchanges' and traders' computer servers on market liquidity. *Journal of Futures Markets 34*(1), 20–33.

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

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*, Volume Financial Management Association survey and synthesis series. 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? *The Journal of Finance 66*(1), 1–33.

Hendershott, T. and P. C. Moulton (2011). Automation, speed, and stock market quality: The nyse's hybrid. *Journal of Financial Markets 14*(4), 568–604.

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

Jarnecic, E. and M. Snape (2014). The provision of liquidity by high-frequency participants. *Financial Review 49*(2), 371–394.

Jovanovic, B. and A. J. Menkveld (2016). Middlemen in limit order markets.

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

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

Kwan, A. and R. Philip (2015). High-frequency trading and execution costs. In *EFMA Annual Meetings-2015, Amsterdam.*

Latza, T., I. W. Marsh, and R. Payne (2014). Fast aggressive trading.

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

Lee, E. J., K. S. Eom, and K. S. Park (2013). Microstructure-based manipulation: Strategic behavior and performance of spoofing traders. *Journal of Financial Markets 16*(2), 227–252.

Mazza, P. and M. Petitjean (2016). On the usefulness of intraday price ranges to gauge liquidity in cap-based portfolios. *Economic Modelling 54*, 67–81.

Mazza, P. and M. Petitjean (2018). Implicit transaction cost management using intraday price dynamics. *Forthcoming in Applied Economics*, 1–11.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics 8*, 1–24.

Menkveld, A. J. and M. A. Zoican (2017). Need for speed? exchange latency and liquidity. *The 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.

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.