Competition for Order flow and Price Discovery: The Curious case of High-tech Entrants

GBENGA IBIKUNLE*
University of Edinburgh, Scotland
European Capital Markets Cooperative Research Centre, Pescara, Italy
Centre for Responsible Banking and Finance, University of St Andrews, Scotland

Abstract
I investigate how entrant markets attract order flow at the expense of established markets. High frequency analysis of BATS Chi-X Europe (Chi-X), a high-tech entrant, and an established national exchange (London Stock Exchange – LSE), suggests that Chi-X holds price leadership in the London market. This position proves critical to Chi-X's acquisition of market share at LSE's expense. Intraday variations in price leadership are driven by informed trading, liquidity constraints and institutional trading arrangements on both platforms, but are inconsistent with the theoretical liquidity-efficiency link. Asymmetric effects of dark and algorithmic trading across the London platforms further complicate this picture.

JEL Classification: G14; G15; G18

Keywords: High-tech entrant markets; Price discovery; Multilateral Trading Facilities; Regulated Markets; Dark trading; Algorithmic trading

---

* Contact information: University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh EH8 9JS, United Kingdom; e-mail: Gbenga.Ibikunle@ed.ac.uk; phone: +441316515186.
I am grateful to Yuxin Sun and Steven Kay for research assistance. I also thank the participants at the 27th Australasian Finance and Banking Conference in Sydney, the 18th Annual Conference of the Swiss Society for Financial Market Research in Zurich, Albert Menkveld, Tālis Putniņš, Petko Kalev, Richard Payne, Jo Danbolt, Øyvind Norli, Seth Armitage, Alex Frino, Ben Jacobsen, Chardin Wese-Simen, Tobias Dieler, Damien Wallace and colleagues at the University of Edinburgh for helpful comments and discussions; the usual disclaimer applies.
† An earlier version of this paper, using older datasets from a different (smaller) BATS Chi-X Europe's order book (BXE) was circulated under the titles "*Order flow, Liquidity and Price Leadership: The Curious Case of High-tech Entrant Markets*" and "*Competition for Order flow and Informativeness of Quotes*".

1. Introduction

New trading venues must compete with established national markets for market share in order to survive. Prior to the advent of large scale electronic trading, this task would have been quite impossible. Established exchanges hold the advantage of being able to draw on large levels of liquidity via their existing networks, thereby reducing search costs for counterparties trading on their own platforms. Given that search costs could be very expensive, this extent of power constituted a high entry barrier to entrants (see Pagano 1989). In this age of electronic trading however, search costs are rather insignificant because investors can survey a large cross section of trading venues from one location via a computer link up. Nevertheless, not all investors have access to the technology needed to negate the impact of the search costs at the speed required to remain competitive, thus entrant venues may still struggle to attract retail investor volumes (see Foucault & Menkveld 2008). Menkveld (2013) argues that high frequency traders (HFTs), who trade at very high speeds through the use of computer algorithms, could play an important role in ensuring that entrant markets compete favourably with the incumbents. Specifically, HFTs could generate competitive quotes, which new venues require in order to succeed. Accordingly HFTs are the dominant players in entrant high-tech markets and appear crucial for their survival as competition to established national exchanges.

BATS Chi-X is an entrant high-tech market that has successfully challenged established European national exchanges for order flow. Based on monthly trading estimates from Thomson Reuters over the last four years, BATS Chi-X Europe (hereinafter referred to as Chi-X) often alternates with the LSE for the title of Europe's largest share trading platform. This achievement is even more remarkable when one considers that previous entry attempts by entrant high-tech markets in Europe have largely failed; the failed EuroSETS challenge of NYSE-Euronext is an example. In this paper, I propose that new entrants must possess comparatively more informative quotes in order to successfully challenge established platforms for order flow. Therefore, I conduct a case study to examine the comparative informational quality of quotes from BATS Chi-X's largest order book, CXE and the London Stock Exchange (LSE)'s largest order book, the SETS. The influence of the quality of Chi-X quotes on its share in the London market for FTSE 100 stocks is then tested. Although the LSE has recently been replaced by Chi-X as the largest trading venue by value in Europe, it still holds the largest share of the market for trading its signature listings – FTSE 100 stocks, which are arguably the most liquid European stocks. I hypothesise that the level of information content in the orders and transactions on European entrant high-tech markets (e.g. Chi-X) is so high as to ensure that market makers protect themselves against informed trading by posting quotes with spreads that are generally wider than those of established national platforms trading in similar instruments (see Figure 1). This would imply that the entrant markets do not necessarily offer comparatively more competitive quotes; instead I postulate that they attract order flow through the provision of more efficient prices.[1] Consequently, I also advance a new argument that superior trading activity and liquidity (narrower spreads) in established markets do not necessarily translate into price leadership over high-tech entrant markets. However, given the demonstrated links between trading activity and price discovery in literature, it is also posited that since trading activity fluctuates across the trading day, price

[1] It is important to note that a venue's spread could be wider than its competitors' and its overall transaction costs could still be lower or approximate its competitors'. This might arise when the execution costs are lower for the venue with the wider spreads. This is the case with the two venues examined in this paper; while BATS Chi-X's spreads are generally wider on per-minute basis, when the execution costs are taken into consideration, the transaction costs for the platform is similar to the costs for the London Stock Exchange. On the LSE, standard execution charge for equities is 0.45bps for the first £2.5bn of an order executed. On BATS Chi-X it costs 0.15bps to add the first €8bn of an order; however, there is a charge of 0.30bps to remove also the first €8bn of an order.

discovery will also be time-varyingly connected with informed trading.[2] Generally, the results presented in this paper support my lines of argument.

I find that the competitiveness of the entrant market is strongly linked to the efficiency of the prices (information content of quotes) it generates. As hypothesised, the price discovery process for both platforms varies across the trading day. All the price discovery measures used are largely consistent, but the determination of which platform has price leadership at any one interval is dependent on whether one considers impounding fundamental information or a combination of fundamental information and the avoidance of noise. Price leadership, defined as the ability of a trading venue to attain efficient pricing of an instrument ahead of its competition, is an important pursuit for trading venues. This is because leading the price discovery process for a common instrument implies that the price leading venue acquires and efficiently incorporates information into instrument prices in a timely manner. It also speaks to the quality of the market's management and set up as well as to its liquidity,[3] these are significant considerations for attracting informed investors (Wang & Yang 2011). Therefore price leadership is important in attracting order flow and revenue that will ensure successful entry and survival for new venues. This study shows that entrant high-tech venues can achieve price leadership even with lower levels of trading activity simply by posting informative quotes at a faster pace than the incumbent, hence the importance of informed traders to entrant high-tech markets. The information going into the quotes indeed may come from keenly observing order flow from the incumbent exchange (see Chordia *et al.* 2008).

[2] Taylor (2011) and Frijns *et al.* (2015) also show that price discovery vary intraday in relation to macroeconomic announcements. However, this is the first paper to explicitly measure the share of price discovery (and in relation to informed trading) across the trading day.

[3] Liquidity is defined as the ability to trade large quantities of an instrument, relatively quickly, anonymously and with little or no price impact (see Campbell *et al.* 1997).

This is easily achievable if the entrant venue can develop the infrastructure to attract sophisticated traders with cross-market trading strategies and multi-venue trading operations.

Specifically, the results in this study suggest that on average Chi-X is faster at impounding fundamental information about the value of FTSE 100 stocks into their prices than the LSE, especially during early trading. This implies that the prices on Chi-X are comparatively more efficient than those on the LSE. Although LSE continues to enjoy dominance with respect to trading volumes across all trading periods, Chi-X appears to be the venue where most of the proportional informed trading occurs, hence Chi-X leads price discovery in more stocks and for most of the trading day. This implies that the larger trading volume of the LSE is not associated with price discovery during the continuous trading periods. This apparent violation of the price discovery-trading activity link (see for example Biais *et al.* 1999; Barclay & Hendershott 2003) is due to several factors. First, it is related to the differences in the structure of both markets; for example, the LSE is a hybrid trading venue with institutional trading-specific special arrangements on its upstairs market, while no such arrangements exist on Chi-X. Other complicit factors include informed trading activity, venue liquidity and Chi-X's share of trading volume in the London market. Results suggest that dark and algorithmic trading are inversely related to Chi-X's share of price discovery. However, for the highest volume stocks algorithmic trading helps improve Chi-X's share of price discovery. Generally, algorithmic trading on the LSE improves the venue's share of price discovery.

To my knowledge, only two studies have addressed questions related to the price leadership or informativeness of quotes involving a high-tech trading venue/mechanism and an established trading venue. Huang (2002) investigates the distribution of price discovery among a group of participants on NASDAQ, of which Electronic Communication Networks (ECNs), Instinet and

Island are a part. Huang (2002) compares the quality of 30 Dow stocks quotes submitted to the NASDAQ national bid and ask montage by traditional market makers and ECNs in July 1998 and November 1999. The study finds that the two most liquid ECNs in the United States at that time, Instinet and Island, post highly informative quotes. Baillie *et al.* (2002), adopting a similar framework in relation to Yahoo's stock from March 1999, arrive at the same conclusion.[4] These related studies provide important insights into how computer-driven trading networks contribute to the quality of NASDAQ quotes. This current paper makes new contributions to the literature by accounting for trading activity and liquidity dynamics as well as for the shift towards high frequency trading activity in financial markets. It also presents evidence on the drivers of the price discovery dynamics observed.

Recent evidence (see as examples, O'Hara & Ye 2011; Menkveld 2013) indicate that entrant high-tech markets are rapidly acquiring exchange market share at the expense of established exchanges.[5] One reason for this change is the lowering of the entry barrier to trading venues by both technological innovations and regulatory policy. Technological innovations especially are mostly responsible for the rapid lowering of the entry barrier to exchange trading. Hence, most entrant markets are high-tech enclaves where algorithmic trading (AT), the process of automated trading with computer algorithms, thrive. AT on these platforms is mainly based on high frequency trading (HFT) strategies, hence a substantial proportion of entrant high-tech markets' price discovery have been suggested to be linked to HFT activity (Menkveld, 2013). Most studies generally agree that HFT activity improve market quality. These studies are relevant to this

[4] Another stream of literature examines the impact of ECN quotes on market quality; for example, Barclay and Hendershott (2003) employ ECN transactions and quote data in their analysis of price discovery and trading after the market close.

[5] In a recent study, Kwan *et al.* (2015) also investigate competition between incumbent exchanges and dark pools. Their study is context-driven and based on the US regulatory environment; they find that dark pools hold a competitive advantage over exchanges when trading is spread-constrained. This current paper grapples with a different regulatory environment, based on the Markets in Financial Instruments Directive (MiFID) currently in force in Europe.

paper due to the significant HFT activity on high-tech entrants. Brogaard (2010) is one of the first studies to examine the contribution of HFTs to market quality. His study, which is based on an analysis of 120 NASDAQ stocks from 2008–2010, suggests that HFTs help to improve market quality. Four other studies employ the same dataset used by Brogaard (2010) in HFT and market quality activity-related studies. Firstly, Brogaard *et al.* (2014) report that HFTs are crucial contributors to price efficiency. Secondly, Carrion (2013) finds that HFTs are more likely to trade when there is reduced liquidity and when the market is relatively more efficient. Thirdly, Hirschey (2013) documents results consistent with HFTs being able to forecast short horizon returns. Finally, Zhang (2012) investigates the impact of information events on HFT and non-HFT activity.

Other papers have not employed the NASDAQ HFT data, but their findings are largely consistent with those who have. For example, Hasbrouck and Saar (2013) proxy HFT by proposing a novel measure of low latency activity, which correlates with NASDAQ-defined HFT trading. They find that HFT activity improves standard measures of market quality such as the bid-ask spread. Another stream of literature draws causal relationships based on events and HFT data. For example, Kirilenko *et al.* (2011) and Easley *et al.* (2011) examine the role of HFTs around the flash crash on 6th May 2010. Kirilenko *et al.* (2011) report that HFT could not have triggered the flash crash; however, HFTs aggravated the crash through their aggressive mopping up of liquidity in order to retain target inventory levels during the crash. Easley *et al.* (2011) also blame order flow toxicity for triggering the flash crash.

Hagströmer and Nordén (2013), using data from NASDAQ OMX Stockholm, suggest that HFT reduces intraday price volatility. Malinova and Park (2014) use HFT message traffic as a proxy for HFT activity, and also find that a reduction in HFT activity leads to the widening of trading spreads and an elevation of trading costs for all market participants. Apart from Hagströmer and Nordén (2013) and Menkveld (2013), there are few other studies that have employed data from European platforms in their analysis of HFT. For example, Jovanovic and Menkveld (2012) employ Chi-X trading data in examining the role of HFTs as middlemen in limit order markets. Jovanovic and Menkveld (2012) show that HFT increases welfare such that introduction of HFT coincides with a 23% drop in adverse selection and a 17% increase in trade latency. In their analysis of the impact of HFT on trade size across 24 trading venues across the world, Aitken *et al.* (2014) also employ data from Chi-X. They suggest that the reduction in average trade sizes may not be related to HFT after all, since average trade sizes have fallen even in countries with no HFT. Cumming *et al.* (2012) also use Chi-X data in their examination of HFT and end-of-day price dislocation. Their results suggest that HFT diminishes the prevalence and severity of price dislocation.

Based on the foregoing, the prevalence of HFT activity on a platform improves its pricing efficiency. This suggests that Chi-X, which is crucially reliant on HFTs (Menkveld 2013), may be more efficient in pricing instruments than the LSE. Specifically, since HFTs enhance informational efficiency by speeding up price discovery and eliminating arbitrage opportunities (see Brogaard *et al.* 2014; Chaboud *et al.* 2014), one would expect to see a significant proportion of FTSE 100 stocks' new prices being discovered first at Chi-X.

As all HFT activity is algorithm-based, this paper also relates to a wider stream of literature on AT. Studies conducted on AT are in general agreement with the findings in the studies reviewed above. For example, Hendershott *et al.* (2011) find that AT improves trading quality via increased informativeness of quotes and reduction in trading costs. Boehmer *et al.* (2013) report that algorithmic traders (ATs) improve liquidity; and Chaboud *et al.* (2014) show that ATs enhance informational efficiency by speeding up price discovery. The consistency in the findings for both

HFT and AT is unsurprising, because HFT is a sub-set of AT. Furthermore, results obtained by Brogaard *et al.* (2014) strongly imply that HFT is crucial to AT improving market quality.

This paper involves the direct examination of the distribution of price discovery between two trading venues; therefore, it is also related to yet another stream of literature, which examines the distribution of price discovery across different platforms. The earliest studies in this literature area examine the distribution of price discovery between the NYSE and regional stock exchanges (see as examples, Harris *et al.* 1995; Hasbrouck 1995; Harris *et al.* 2002). Although Bacidore and Sofianos (2002) argue that the home market should normally be the informationally dominant market, the results have not been generally consistent across different study settings. For example, Frijns *et al.* (2010) and Lok and Kalev (2006) study price discovery dynamics for cross listed Australian and New Zealand firms on the New Zealand Stock Exchange and the Australian Stock Exchange respectively (see also Roope & Zurbruegg 2002) and find that the home markets are the informationally dominant ones. Lieberman *et al.* (1999) also report similar results for Israeli firms cross-listed on NYSE. However, Kadapakkam *et al.* (2003) show that both the home and foreign markets contribute equally to the price discovery in the case of Indian firms listed on the London Stock Exchange. Eun and Sabherwal (2003) and Hupperets and Menkveld (2002) also show significant variations in the price discovery contributions across stocks cross-sections for the home and foreign markets. Most of the studies examining price discovery leadership between platforms have mainly resorted to examining the phenomenon by focusing on the time overlap periods, since those platforms are usually based in different time zones (see as examples Hupperets & Menkveld 2002; Pascual *et al.* 2006; Su & Chong 2007).[6]

[6] Wang and Yang (2011) propose alternative approaches to measuring price discovery in sequential markets. Those approaches are not relevant to this study since I focus on platforms operating in the same city and in the same time zone.

An oversight in all the aforementioned papers is the implicit assumption that the distribution of price discovery is constant across the trading day. Price discovery is a function of trading activity, which fluctuates across the trading day. Given the significant impacts trading volumes can have on instrument liquidity, informed trading content of a market may evolve in tandem with trading activity across the trading day. Friederich and Payne (2007) report that liquidity and informational risks influence choice of trading venue for investors. There is no reason why this decision on venue choice cannot be reviewed periodically on an intraday basis, and this will have significant consequences for price discovery. Kyle (1985) notes that even the most efficient markets still manage to reflect observable levels of private information. Hence, when markets become more liquid, they tend to more readily absorb information. This is because a reduction in trading costs due to higher volumes may incentivise informed trading (see Admati & Pfleiderer 1988; Chordia *et al.* 2008). Trading costs in this case may refer to microstructure impacts, the cost of searching for counterparties or costs due to adverse selection. Thus, in a trading environment where traders can select trading venues with ease, informed traders will likely choose a market where their trades will enjoy the best level of disguise at that point in time. It is therefore logical to expect that in a high-tech trading environment with sophisticated traders using cross-market strategies, the distribution of price discovery between two markets will evolve substantially across the trading day. The trading environment in London is such a high-tech trading environment, where sophisticated traders from all over the world survey stocks on multiple platforms within and outside the city.

The two largest equity trading venues in the city (and in Europe) are the LSE, an established national platform and Chi-X, a high-tech entrant. One of the hypotheses tested in this paper is that price discovery distribution between these two venues fluctuates across the trading day; I also propose that the price discovery distribution is related to the evolution of informed trading on both platforms. Therefore, this study extends the literature stream in this area by investigating price discovery distribution between two markets across intraday periods and linking those dynamics to informed trading activity and other price discovery determinants such as dark and algorithmic trading.

The remainder of this paper is structured as follows. Section II provides a background to the study by discussing the market structure and regulatory framework, as well as the relevant institutional background for the LSE and BATS Chi-X. Sections III and IV discuss the data and methodological framework respectively. Section V presents and discusses the empirical results, while Section VI concludes.

2. Background to the study

2.1. Market Structure and Regulatory Framework

Due to its competition-enhancing rules, the Markets in Financial Instruments Directive (MiFID) of the European Union in 2007 spurred the proliferation and growth of alternative high-tech trading venues to the established exchanges hitherto prevalent in the European markets, the so-called regulated markets (RMs). Such alternatives include multilateral trading platforms (MTFs), broker crossing networks (BCNs) and systematic internalisers (SIs). The market share held by MTFs especially has grown tremendously since MiFID's inception, and nothing exemplifies this more than the fact that in the first half of 2013, Chi-X, which then only had an MTF status, was the largest equity trading venue in Europe by market share.[7] As at 13th November

[7] Prior to 20th May 2013, BATS Chi-X Europe only had a licence to operate MTFs; however, since being granted a Recognised Investment Exchange (RIE) status, BATS Chi-X could now operate a listing exchange alongside its existing MTF operating business. The data employed in this paper is for a period after the transition of BATS Chi-X to RIE status. The price discovery dynamics of the BATS Chi-X order book employed in this analysis remains

2014 there were 151 MTFs listed on the CESR MiFID database managed by the European Securities and Markets Authority (ESMA).[8] Given that MTF operators such as Chi-X also operate prominent dark order books or dark pools[9] alongside their 'lit' ones, it is pertinent to examine how their trading volumes contribute to price discovery for instruments primarily listed on RMs. Such examination is of paramount interest to investors who may be apprehensive of the activities of HFTs on MTF dark pools (see Schacht *et al.* 2009).[10] For example, in April 2014, about £719.1 million worth of dark transactions were executed on Chi-X's BATS Europe (BXE) and Chi-X Europe (CXE) dark order books.[11]

EU market venues under MiFID are essentially defined on the basis of transaction execution; there are two main classes – bilateral or multilateral trading structures. MiFID generally defines the design, operation and interaction of financial markets in the EU. RMs and MTFs, which are the focus of this study, make up the multilateral class, and SIs are bilateral venues. RMs and MTFs are largely limit order markets, which operate by matching orders on established rules of price, time and visibility priorities. RMs are mainly the national stock exchanges such as the LSE. RMs and 'lit' MTFs regularly display and update market maker and limit order quotes on their order books; while 'dark' or dark pool MTFs do not display orders prior to execution, thus providing no pre-trade transparency (dark trades).[12] However, for all multilateral venues, all trades

---

essentially the same from before and after the transition. Enquiries made with BATS Chi-X confirm that their current order books are still the same as when BATS Chi-X could only operate MTFs; thus those books are still classic MTFs. Furthermore, achieving the RIE status was only expected to advance BATS Chi-X's fortunes with retail investors (see Financial-Times 2013). As at May 2014, BATS Trading Limited is still listed on the CESR MiFID database as an MTF. However, BATS Europe Regulated Market is also now listed as a regulated market.

[8] ESMA develops rule books for EU financial markets.

[9] Dark pools are order books where orders with no pre-trade transparency are executed. Orders submitted to dark pools are called dark orders and the trades arising on account of such orders are dark trades/transactions.

[10] A survey conducted by the CFA finds that 71% of respondents are of the view that dark pools constitute problems for price discovery.

[11] This is according to BATS Global Markets April 2014 Highlights, published by BATS Global Markets.

[12] It is also noteworthy that although RMs are not known to directly operate dark pools, some are associated with them. For example, the LSE Group is a majority shareholder in Turquoise, which operates one of the most liquid dark pools in Europe.

must be posted as soon as possible after execution; this directive is aimed at improving transparency.

Directive 2004/39/EC Article 4 (15) of the EU defines an MTF as '*a multilateral system operated by an investment firm or market operator, which brings together multiple third-party buying and selling interests in financial instruments – in the system and in accordance with non-discretionary rules - in a way that results in a contract in accordance to the provisions of Title II*'. This definition is not significantly different from that of RMs given in Article 4 (14) of the same directive. However, MTFs have evolved to provide pan-European market access, while RMs remain mainly domestic. Virtually all MTFs consist of integrated limit order books, which allow for non-discretionary and anonymous order matching among market participants. Although MiFID requires that RMs and MTFs to publish all current bid and ask prices as well as their corresponding bid and ask sizes, the regulations also provide for exemptions on four grounds.[13] MTFs rely on these exemptions to operate dark pools.

2.2. Institutional Background

The London Stock Exchange's order book for its most liquid instruments is a hybrid market where trades are executed on its main order book, the stock exchange electronic trading system

[13] The first category of orders granted exemption from pre-trade transparency are large orders, which can have large market impacts if published pre-execution; this is referred to as the Large-in-Scale (LIS) waiver. To qualify for a LIS, trades must be of a minimum size, which is dependent on the average daily turnover for each instrument. The minimum order size ranges from 50,000 for the least active stocks to 500,000 for the most active ones. The second waiver is the Reference Price waiver, which is commonly used by MTFs to maintain dark pools of liquidity. Both RMs and MTFs may avoid abiding with pre-trade transparency requirements if they passively match orders to a widely published reference price obtained from another market. For example, BATS Chi-X's dark pools for FTSE index stocks commonly passively match orders to LSE's posted midpoints. The third waiver deals with transactions negotiated privately away from the exchanges by counterparties. These transactions are usually non-standard and must be conducted on the basis of prevailing volume weighted bid-ask spread or a reference price if the instrument is not subject to continuous trading. The final exemption relates to iceberg orders, and is known as the Order Management Facility waiver. RMs and MTFs can waive pre-trade transparency for orders subject to order management facility until such time when they will be disclosed to the market. In practice, only a fraction of submitted orders is displayed, and once filled, the portion is refreshed using part of the previously non-displayed order.

(SETS), and via broker-dealers. Broker-dealer trades must be reported within three minutes of execution. The FTSE 100 is the LSE's main index and it contains the largest 100 eligible UK firms listed on the platform;[14] those firms account for more than 81% of total LSE firms' market value. Despite the increased competition for market share faced as a result of MiFID's enactment, LSE has managed to retain the largest share of trading for its listed stocks. In October 2014, the platform accounted for 45.40% and 66% trading share of the FTSE 100 and of all its other listed instruments, respectively. The economic difference between the platform's shares in the largest volume stocks and other instruments is evidence of the increasing attractiveness of BATS Chi-X, to high volume stock traders.

Between its two order books, BXE and CXE, BATS Chi-X Europe held 35% of all FTSE 100 trades in October 2014. BATS Chi-X Europe is the largest equity trading venue by market share in Europe, with 21.40% of the European market in April 2014. It was formed in 2011 through the merging of two of the three largest continental MTFs, BATS Europe and Chi-X Europe, which was founded by Instinet. In May 2013, BATS Chi-X's owners, BATS Trading Limited, were granted a Recognised Investment Exchange (RIE) status by the Financial Conduct Authority in the United Kingdom. RIE status gives the firm the right to operate an RM for primary listings along with its current operation of an MTF. The BATS RM is for the trading of BATS listed securities and the BATS MTF trading platform consists of integrated order books (BXE and CXE)[15] as well as dark pools for trading of securities listed on other RMs. For the trading of stocks in my sample, both LSE and Chi-X commence continuous trading on the limit order book at 08:00:00hrs London local time and conclude the trading session at about 16:30:00hrs. Prior to the

[14] See http://www.ftse.com/Indices/UK_Indices/Downloads/FTSE_UK_Index_Series_Index_Rules.pdf for index rules

[15] There is also a trade reporting facility for OTC, BCN and SI trades; this is available to participants from 07:15:00hrs to 17:15:00hrs every trading day.

continuous trading session, the LSE opens with a 10-minute opening auction session to set the reference price for the trading day; it also closes with a five-minute closing auction session to set the closing price (see Ibikunle 2015 for a background on the LSE's auction sessions). Chi-X does not participate in these auctions; hence the dataset used consists of trading data for the hours between 08:00:00hrs and 16:30:00hrs. In any case, the price obtained for 08:00:00hrs should reflect the any information impounded into the price of instruments during the opening auction.

3. Data

3.1. Sample Selection

As earlier stated, Chi-X operates two integrated order books, BXE and CXE, which both trade all FTSE 100 stocks. Both order books have integrated lit and separate dark sections; both the lit and dark sections are normally allowed to interact throughout the trading day. However, whether an order hits both the lit and dark sections depends on the order type. It should be emphasised that the dark pools are separate from the integrated order books. I obtain two sets of data from the Thomson Reuters Tick History (TRTH) database. The first is for high frequency second by second quotes data for 47 of the highest volume FTSE 100 stocks trading on both the LSE's SETS and Chi-X's CXE order book between 1st July 2014 and 28th November 2014 (108 trading days). I also obtain intraday tick-by-tick trades and messages data, stamped to the nearest millisecond, for the same period and for both order books.

The BXE and CXE are integrated order books, the price innovation processes on both books are therefore inextricably linked. The CXE order book consistently accounts for about 75% of the entire Chi-X's FTSE 100 trading volume. The implication of this is that CXE adequately represents the trading environment at Chi-X just as much as SETS does for the LSE. The final sampling date in both datasets is 28th November 2014; on that date, the 47 stocks in my sample jointly account for 75.01% of the FTSE 100 index weight.

3.2. Sample Description

In order to better observe trading activity-related dynamics, the 47 stocks are exogenously split into pound volume quintiles by using average daily trading pound volume. The trading day is also exogenously divided into seven periods in order to further grasp the intraday dynamics of price discovery distribution between the two order books.[16] Panels A and B in Table 1 present the summary statistics for LSE's SETS and BATS Chi-X's CXE order books respectively. For SETS (CXE), the average daily trading value is over £2.05 billion (£736.33 million) from a daily average of about 75,111 (47,794) trades for all the 47 stocks in the sample. The total SETS (CXE) trading value for the entire period covered by this paper is about £221.36 billion (£79.52 billion). I observe that the 18 highest volume stocks on SETS account for more than 63.74% of that value; this large trading gap between highly active LSE stocks and the relatively less active ones is in line with observations from previous literature. It is also observed that the trading gap is most pronounced during the first trading hour, when the 9 highest volume stocks account for about 42.57% of the total transactions value.

The intraday trading dynamics observed for CXE are largely in line with those of SETS. One area of distinction is how average trade sizes vary from stock to stock across quintiles and between the two order books. Generally, the trade sizes on Chi-X are lower than those on the LSE. The average trade size on CXE is only about 56% of that on SETS. The values vary depending on trading interval, with the disparity greatest during the first half hour of trading for the most traded

[16] I experimented with three to ten trading windows/intervals before finally presenting results for seven due to the stability of the values for those intervals.

stocks. The typical CXE trade size is about 47% of SETS's during the early trading, and only about 42% for the 9 heaviest traded stocks for the same period. Also, there is less variation in the average trade sizes across stocks and time for CXE than there is for the LSE. Given that in the market microstructure literature, changes in trade sizes are thought to reflect changing composition of the traders/participants in a market, one may assume that the aggregate identity of CXE traders is more consistent than that of SETS.

INSERT TABLE 1 ABOUT HERE

4. Measures of Price Discovery

Measuring aggregate price discovery across parallel trading venues was first attempted in Hasbrouck's (1995) seminal study of NYSE-listed instruments. Hasbrouck's (1995) Information Share (IS) approach measures a market's contribution to the unobservable efficient price of an instrument, which is traded on more than one market. This approach's definition of price discovery involves capturing the variance of innovations to the common price factor across multiple venues. A separate approach is based on Gonzalo and Granger's (1995) work on cointegration econometrics; this is the Component Share (CS), which relates only to the error correction process. The process involves the decomposition of a cointegrated time series into transitory and permanent shock components by employing error correction coefficient components. The two methods are based on the vector error correction model (VECM). Baillie *et al.* (2002) show that both models are directly related and that the results obtained from both models mainly stem from the error correction vector in the VECM. The models usually provide qualitatively similar results if the VECM residuals are uncorrelated. However, if there is a significant level of serial correlation in the VECM residuals, the results obtained from the two methods may be different. This difference on account of residual autocorrelation is due to Hasbrouck's (1995) inclusion of contemporaneous correlation in his measure of price discovery contribution, and Gonzalo and Granger (1995) not following the same approach. In order to rectify this problem, Hasbrouck (1995) recommends the use of the Cholesky factorisation to eliminate the contemporaneous correlation.

However, both measures of price discovery potentially suffer from estimation bias if noise levels differ across trading venues/price series (see Yan & Zivot 2010; Putniņš 2013). Thus, in addition to the CS and IS price discovery measures, the information leadership share (ILS) prescribed by Putniņš (2013) on account of Yan and Zivot's (2010) information leadership (IL) metric is also computed. Abridged theoretical bases for all three approaches are presented below.

4.1. Component share

There is a natural expectation that the prices of LSE-listed stocks traded on Chi-X are cointegrated with the prices of those obtained on LSE, because the underlying instruments for the cross-listed Chi-X transactions are indeed those LSE stocks. Thus, if both price series are I(1) cointegrated, $P_t = (P_{1t}, P_{2t})'$, the following VECM can be estimated:

$$\Delta P_t = \alpha\beta' P_{t-1} + \sum_{j=1}^{k} A_j \Delta P_{t-j} + e_t, \qquad (1)$$

where α corresponds to the error correction vector, β is the cointegrating vector and e_t is a zero mean vector of serially uncorrelated innovations with covariance matrix Ω:

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \qquad (2)$$

$\sigma_1^2(\sigma_2^2)$ is equal to the variance of $e_1(e_2)$ and ρ is the correlation of e_{1t} and e_{2t}. The VECM comprises of two components. The first component, $\alpha\beta' P_{t-1}$, corresponds to the long-run equilibrium dynamic between the LSE and BATS Chi-X price series, while the second part,

$\sum_{j=1}^{k} A_j \Delta P_{t-j}$, describes the short-term dynamics caused by pricing flaws in the market. Such imperfections include noise induced by microstructure impacts of trading large sizes. The above specification is comparable to the common trend representation of Stock and Watson (1988):

$$Y_t = f_t + G_t, \tag{3}$$

where f_t represents the common factor component and G_t the transitory component, which leaves only a transient effect on Y_t. According to Gonzalo and Granger (1995), the common factor in this representation can be defined as a combination of the variables $Y_t = (y_{1t}, y_{2t})$, such that $f_t = \Gamma Y_t$, where $\Gamma = (\gamma_1, \gamma_2)$ is a 1×2 common factor coefficient vector. Gonzalo and Granger (1995) prove that Γ is statistically independent of the error correction vector α, which is denoted by $\Gamma = \alpha \perp'$. Γ is normalised such that $\sum \gamma_i = 1$. If f_t is taken as a portfolio of prices from the two markets, then Γ is the portfolio weights for the prices (see Harris *et al.* 2002). Thus, one can compute the price contribution of the first (second) market as $\gamma_1 (\gamma_2)$.

4.2. Information share

Hasbrouck's approach begins with the transformation of Equation (1) into a vector moving average (VMA):

$$\Delta P_t = \psi(L) e_t, \tag{4}$$

which in an integrated form can be expressed as follows:

$$P_t = \iota\psi\left(\sum_{s=1}^{t} e_s\right) + \psi^*(L) e_t, \tag{5}$$

where $\iota = (1,1)'$ is a column vector of ones and $\psi = (\psi_1, \psi_2)$ is a row vector. $\psi^*(L)$ is a matrix of polynomials in the lag operator, *L*. Equation (5) is analogous to Equation (3). The increment ψe_t in the first portion of Equation (5) is deemed by Hasbrouck (1995) as the permanent price innovation component due to new information; this component is the so-called common efficient price – the common factor. The decomposition of the variance of the common factor innovations, denoted by $\text{var}(\psi e_t) = \psi \Omega \psi'$, means that a market's information share corresponds to the part of $\text{var}(\psi e_t)$ that is due to innovations in that market. Suppose the covariance matrix *Ω* is diagonal, the information share of the *j-th* market will correspond to:

$$S_j = \frac{\psi_j^2 \sigma_j^2}{\psi \Omega \psi'}, \tag{6}$$

where ψ_j is the *j-th* element of ψ. Assuming that *Ω* is not diagonal, it will be impossible to systematically obtain the information share. As stated earlier, Baillie *et al.* (2002) show that $\frac{\psi_1}{\psi_2} = \frac{\gamma_1}{\gamma_2}$; thus, suppose the error terms are uncorrelated, the information share may be calculated using Equation 7:

$$S_j = \frac{\gamma_j^2 \sigma_j^2}{\gamma_1^2 \sigma_1^2 + \gamma_2^2 \sigma_2^2} \tag{7}$$

and Equation 8:

$$\frac{S_1}{S_2} = \frac{\gamma_1^2 \sigma_1^2}{\gamma_2^2 \sigma_2^2}. \tag{8}$$

In a situation where the price processes are significantly correlated across markets, Equation 7 will not hold. In order to eliminate the contemporaneous correlation, Hasbrouck (1995) suggests to use the Cholesky decomposition $\Omega = MM'$, where

$$M = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2(1-\rho^2)^{1/2} \end{pmatrix} \qquad (9)$$

is a lower triangular matrix, resulting in the information share computed as:

$$S_j = \frac{([\psi M])^2}{\psi\Omega\psi'}, \qquad (10)$$

where $[\psi M]_j$ is the *j-th* element of the row vector ψM . Based on Baillie et al.'s (2002) derivation, the conclusion is that:

$$\frac{S_1}{S_2} = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_2 m_{22})^2}. \qquad (11)$$

Since the information shares of both markets equal one, i.e. $S_1 + S_2 = 1$, then

$$S_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, \qquad (12)$$

and

$$S_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}. \qquad (13)$$

One advantage of these expressions is that one can easily compute the ISs based on the covariance matrix *Ω* of the residual vector and the common factor coefficient vector, $\Gamma = (\gamma_1, \gamma_2)$. When there is correlation between market innovations, the Cholesky factorisation is not invariant to the series ordering, and thereby levies a higher IS on the first price process. Hasbrouck (1995) suggests using different price orders and then averaging the upper and lower IS bounds to obtain a final result. Baillie *et al.* (2002) prove that the average of the upper and lower bounds, while re-ordering the price processes for the Cholesky factorisation, yields reasonable estimates of a market's price contribution. This approach is taken to compute the IS estimates in this paper.

4.3. Information leadership share

According to Putniņš (2013) the IS and CS measures of price discovery employed in this paper accurately allocate price discovery leadership in the sense that price discovery refers to the venue/price series to first reflect changes in the fundamental value of the underlying instrument. However, this holds only if the price series have similar levels of trading noise. If there are differences in the noise levels, the IS and CS measure to varying degrees a combination of speed of impounding information and relative avoidance of noise rather than price discovery. Although this still means that the price that incorporates new information first will have higher IS and CS; however, so is the price series with the lower level of noise. Hence when the levels of noise differ greatly relative to the variation in the speed of information incorporation, IS and CS estimates can lead to erroneous conclusions. This kind of econometric problem is more likely to arise when the price series being examined are for different asset classes, such as spot and futures. Hence such is highly unlikely in my analysis because the price series examined across the two markets relate to the same instruments.

Nevertheless, for completeness, the Yan and Zivot (2010) information leadership (IL) metric, which according to simulation results by Putniņš (2013) correctly assign contributions to price discovery when noise levels differ in a bivariate system, is computed. For consistency the metric is expressed in share of price discovery terms as information Leadership share (ILS) (see Putniņš 2013). In Yan and Zivot's (2010) model CS quantifies the relative level of noise in the price series, while IS quantifies both relative noise levels and leadership in reflecting new information. Hence, the ILS is based on amalgamating CS and IS (as seen in Equations 14 and 15) in order to eliminate the relative noise levels and obtain a clear measure of relative information leadership. For each trading interval, the ILS for each of the LSE and BATS Chi-X is estimated as follows:

$$ILS_j^{LSE} = \frac{\left|\frac{IS_j^{LSE}}{IS_j^{BCE}}\frac{CS_j^{BCE}}{CS_j^{LSE}}\right|}{\left|\frac{IS_j^{LSE}}{IS_j^{BCE}}\frac{CS_j^{BCE}}{CS_j^{LSE}}\right| + \left|\frac{IS_j^{BCE}}{IS_j^{LSE}}\frac{CS_j^{LSE}}{CS_j^{BCE}}\right|} \tag{14}$$

$$ILS_j^{BCE} = \frac{\left|\frac{IS_j^{BCE}}{IS_j^{LSE}}\frac{CS_j^{LSE}}{CS_j^{BCE}}\right|}{\left|\frac{IS_j^{LSE}}{IS_j^{BCE}}\frac{CS_j^{BCE}}{CS_j^{LSE}}\right| + \left|\frac{IS_j^{BCE}}{IS_j^{LSE}}\frac{CS_j^{LSE}}{CS_j^{BCE}}\right|} \tag{15}$$

In the above expressions, ILS_j^{LSE} and ILS_j^{BCE} correspond to the information leadership share with respect to stock *j* for LSE and Chi-X respectively. IS_j^{LSE} and IS_j^{BCE} represent the IS with respect to stock *j* for SETS and CXE respectively, while CS_j^{LSE} and CS_j^{BCE} correspond to the CS with respect to stock *j* for SETS and CXE respectively.

5. Empirical Analysis, Results and Discussion

5.1. Price Discovery

5.1.1. Distribution of Price Discovery

In this section, I discuss the distribution of price discovery between the LSE order book (SETS) and the BATS Chi-X order book (BXE). To proceed, the complementarity of the CS and IS estimates is first considered. The CS estimates are not reported for two of the time periods examined as well as for two stock quintiles each in two other time periods. This is because of insufficient statistically significant α_1 values. As evident in Tables 2 and 3, there is a high degree of consistency for both sets of estimates. The inferences drawn regarding which platform leads the information incorporation process by looking at both sets of results are very similar. The only exception relates to the 12:00 – 13:00hrs trading interval for Quintile 3 stocks. According to the CS estimates, the LSE leads the aggregate price discovery for the period, whereas the IS estimates imply the opposite. The difference between the IS estimates is statistically significant, while that of CS is not.

INSERT TABLES 2 AND 3 ABOUT HERE

It is also important to note a high level of variation within the CS estimates when compared with the IS estimates. For example, the standard deviations for the cross-sectional CS estimates for the first 30 minutes of trading are 12.95%, 18.42% and 27.02% for Quintiles 4, 3 and 1, respectively. In contrast, the corresponding IS standard deviations are 1.34%, 1.48% and 0.82% respectively; and the highest IS standard deviation for a period is 8.57% compared with 27.02% for CS. These values imply that the CS results are being driven by a handful of stocks. However, the observed level of variation across the stocks' cross section is not unusual; Hupperets and Menkveld (2002) and Eun and Sabherwal (2003) find large variations in their analysis of Dutch and US listed stocks respectively. Furthermore, the differences between the CS and IS estimates for these stock classes are, on balance, in line with most of the previous literature (see as examples, Su & Chong 2007; Korczak & Phylaktis 2010). Indeed the variation reported in the higher level of variation in the CS estimates is only limited to about 17% of the trading periods examined. However, given the higher level of consistency in the IS estimates across stocks, I focus most of the initial price discovery discussion on the IS estimates as reported in Table 2.

The significant variation in the CS estimates notwithstanding, both the IS and CS measures consistently agree that most of the information incorporation into the sampled stocks occur at Chi-X ahead of the LSE. This is a very surprising result considering the huge differences in the share of information attributed to the two venues and the fact that most of the trading occurs on the LSE. The largest disparity in IS estimates is recorded for the first 30mins of trading, when BATS Chi-X accounts on average for 97.38% of the price innovation on the back of just about 23.80% of trading volume between the two venues for the period. It should be noted that there is very little variation in the estimates for all the stocks examined. While the gaps are less for the other periods, they are surprisingly large nonetheless. All Chi-X IS estimates are significantly different from the corresponding LSE estimates at the 0.001 level. It is important to understand why Chi-X holds such a commanding lead in the race for price leadership. Subsequent sections of this paper examine these issues in detail. However, first, I examine whether the results obtained for the CS and IS estimates could have been influenced by differing noise levels in both markets.

5.1.2. Price Leadership: eliminating the noise

Table 4 presents the average ILS estimates for stocks on per quintile basis. Since ILS is a derivate of CS, estimates could not be obtained for similar periods as the CS. Consistent with the IS and CS estimates, BATS Chi-X leads the information incorporation process for most stocks across all the five quintiles during most trading intervals of the day. However, there are notable exceptions where results suggest that either the LSE leads the information incorporation process at a conventional level of statistical significance or there is no clear price leader between the two markets. Thus the price leader between the two markets is not as clear cut once the differing levels of noise in the two markets is factored into the computation of the share of price discovery. Furthermore, if the length of the trading periods are considered it is likely that the LSE ends as many days as the price leader in London as Chi-X does; however, the stock cross section estimates for the intervals led by the LSE are not statistically significant. This suggests that clear price leadership could not be established by either venue during those trading intervals.

ILS estimates suggest that for Quintiles 2 and 1 stocks, fundamental information are mostly incorporated first on the LSE during the 09:00 – 10:00hrs interval. The same is the case for the highest trading stocks (Quintiles 5 and 4) during the longest trading interval in my analysis, 13:00 – 16:00hrs. The instances of deviation from the IS and CS appear to be related to noise differences at the two venues. This noise variations could have been generated on account of differences in market structure. For example, the LSE is essentially a hybrid market with both an upstairs and a downstairs market, while Chi-X has no formal upstairs market structure. Although it does have a reporting facility similar to the LSE's upstairs, the rules governing both structure are clearly different. Secondly, differences in the identities/classes of dominant traders on both markets could lead to substantial enough differences in both markets as to slightly bias IS and CS estimates. Menkveld (2013) identifies HFTs as the major drivers of trading activity on new entrant platforms like Chi-X, while large institutional traders are the usual trading activity drivers on established platforms like the LSE.

INSERT TABLE 4 ABOUT HERE

Generally, the ILS estimates are consistent with the IS and CS estimates and thus one can retain the suggestion that Chi-X leads the information incorporation process for a substantial proportion of the UK stocks examined despite executing significantly lower transaction volumes than the LSE. The strongest indication to that effect is recorded for the first half hour of trading as measured by the CS and IS estimates, and the same is also observed for the ILS estimates. The overall Chi-X ILS estimates for the first trading interval are 66.64%, 97.35%, 92.17%, 95.94% and 93.91% for Quintiles 5 to 1 stocks, which is consistent with corresponding IS estimates of 96.67%, 97.63%, 96.87%, 97.35% and 98.16%. Thus, the finding that Chi-X is the early front-runner of the regular trading day is unambiguous. The ILS estimates also show the same pattern of the LSE gradually increasing its share of price discovery as the progress, as observed in the IS and CS estimates. The extent of price leadership posted by Chi-X, as shown by the ILS estimates, is remarkable, especially given its inferior trading activity relative to the LSE's figures. The next three sub-sections motivate possible reasons for the price leadership competitiveness of Chi-X. The discussed variables are subsequently formally tested in Sections 5.2 and 5.3.

5.1.3. Trading Activity and Price Discovery

The percentage differences between the IS estimates for the two venues are generally large and only relatively small during the last half-hour of trading. Although the picture is less clear for the CS and ILS, the trend is nevertheless consistent. Price leadership is very critical to the business model of trading venues since it implies that a venue is well managed and is liquid enough to attract informed investors. The shift in intraday price discovery leadership observed for the two venues suggests that both are able to attract informed investors trading FTSE 100 stocks. However, it is also possible that the investors generally are the same and they alternate trading venues in order to benefit from increased liquidity or stealth. The reason for stealth is the need of informed traders to hide their trading intentions (see as an example, Barclay & Warner 1993). Consider the case of CXE's trading volumes and average trade sizes, which are well below those of SETS at every point of the trading day. CXE's average trade size is at its lowest at the point when its share of price discovery is at the highest for both IS and ILS (and second highest for CS) – 08:00 to 08:30hrs. This is also the period that it enjoys its largest order submission rates as a proportion of SETS's order submission rate. All these factors suggest that CXE's price leadership is linked to its volume of small orders, which are preferred by informed traders looking to disguise their trading intentions. The increased aggression in order submission translates into higher liquidity and camouflage for the informed traders. The relative consistency in the average trade sizes on Chi-X also suggests that the identity of traders at the venue does not vary much across the day. Thus the attainment of price leadership appears less linked to the migration of informed traders from the LSE to Chi-X, but rather with increased market depth and liquidity afforded by a higher volume of orders and transactions.

It could also be argued that the SETS being a very liquid order book provides the right environment for camouflage of trading intentions hence should perhaps be preferred more frequently by informed traders. However, the superior trading activity on the LSE appears not to have resulted in price discovery leadership for the platform. This observation is consistent with the findings of Huang (2002), showing that NASDAQ market makers' contribution to price discovery is not systematically linked with market liquidity. Huang (2002) opined that this disconnect can be explained by the existence of institutional trading priority arrangements at NASDAQ. These arrangements invalidate the theoretical and demonstrated link between trading activity and price discovery (see as an example, Biais *et al.* 1999). It is believed that a similar phenomenon is at play at the LSE. This is because, as previously noted, the LSE operates as a hybrid market with a parallel upstairs market, from where executed institutional orders are reported to the SETS, which constitutes the downstairs market. The value of orders executed in the upstairs is non-negligible and has been shown by Armitage and Ibikunle (2015) to account for a substantial portion of information revealed at the LSE. The informativeness of the upstairs is linked to the activities of its dealers, who obtain information regarding unexpressed liquidity requirements of institutional traders and thus facilitate order execution with the accumulated information. Since upstairs trades on the LSE could be reported up to three minutes post-execution, the LSE ecosystem is likely to incorporate information regarding unexpressed liquidity into prices at a pace slower than the rival Chi-X's.

In a bid to avoid inducing unbeneficial price impacts due to their large order sizes, institutional traders on Chi-X are likely to trade using the dark order books available. Since high frequency traders are more dominant in these books (see Menkveld 2013), they can equally collect information on unexpressed liquidity demands of institutional traders and employ such information at a much faster pace than the dealers can reveal them through trades reported to the SETS on the LSE. The above offers an explanation on why increasing market depth on the LSE is disconnected from price discovery. Huang (2002) document, as I find in the case of Chi-X, that ECNs' (Instinet and Island) share of price discovery contribution is positively linked with trading volume. They associate this with informed traders trading on ECNs when there is sufficient liquidity to disguise their trading intentions (see also Kyle 1985).

I also note that the LSE competes most favourably for price leadership in the final half hour of trading, where the IS estimates for the quintiles range from 44.80% for Quintile 2 stocks to 46.40% for Quintile 5 stocks.[17] This strong showing by the LSE appears linked again to an institutional trading arrangement on the platform. Institutional investors are allowed to submit volume weighted average price (VWAP) orders to broker-dealers for execution at the close – 16:30hrs – just before the commencement of the closing batch auction. Such orders only specify buy or sell quantities and are executed at the end of the trading day at the VWAP observed for the trading day. What this implies is that no new prices are discovered at the close since the VWAP is dependent on earlier price discovery processes. Ibikunle (2015) also reports that the most active

[17] ILS and CS estimates are not available for this period. I extrapolate that had there been CS estimates, consistent with the IS trend, the ILS estimates would also be at their largest during this period.

trading period for institutional trades on the LSE is late in the continuous trading day and around the closing call auction, which is due to the arrival of VWAP orders. Since the VWAP orders cannot contribute to price discovery, owing to the fact that they are priced based on previous trades, it is likely that orders are placed in advance of the close in order to influence the VWAP for the day. This conjecture is strengthened by results obtained by Ibikunle (2015) suggesting that LSE trades become increasingly informative around this period. The trading summaries in Table 1 also imply that this is a tenable assertion. The largest value per unit time of trading on the LSE (£3.01 million/min) is recorded for the final half hour of the continuous trading day versus the middle of the trading day and the first half-hour on Chi-X.

Furthermore, the largest average trade sizes for the LSE are also recorded in the final half-hour of trading, suggesting a desperation to execute orders. Such implied desperation is usually linked to trading on some form of information. The average trade size on the platform rises by almost 13% in the final half hour from the previous hour to £34,873 and £34,943 for Quintiles 5 and 4 stocks respectively. This suggests that the trades recorded on the LSE in the final half hour are very informative and that the LSE is at its most informative during this period of the day, hence the improvement in the platform's contribution to price discovery. I examine this observation further in Section 5.2.

5.1.4. Dark orders, High frequency trading and Price discovery

The consistency in Chi-X's trade sizes across the day, as presented in Panel B of Table 1, suggests the presence of largely the same type of traders dominating trading across the different periods. These order sizes are always on average lower than those of the LSE as shown in Table 1. Given the evidence in Menkveld (2013) regarding trading on Chi-X and the elevated levels of messages to trades ratio (average of 27.51 across all quintiles and 30.61 for Quintile 5), I surmise that the dominant participants are likely to be HFTs. The increased frequency of order submission and execution during the periods when Chi-X holds comparatively highest levels of IS/ILS leadership is in line with the hypothesis that informed traders step up their activity with the appearance of higher transaction volumes. The higher trading volumes act as camouflage for informed trading (Admati & Pfleiderer 1988). Thus, during periods with high transaction levels per unit time such as the first half hour of trading, one should expect a correspondingly high level of informed trading activity. In Section 5.2, the question of intraday variation in informed trading activity is addressed.

Theoretically, sophisticated/informed traders crave the anonymity and volumes provided by entrant high-tech markets like BATS Chi-X (see Kyle 1985). Therefore, it is possible that the availability of dark pools, which allow for the interaction of dark orders with integrated lit order books, also serves as an added attraction. Most MTFs in Europe operate integrated order books, as is the case with the Chi-X order book I examine in this study, CXE. These books allow non-discretionary and anonymous matching of orders amongst traders, thus dark orders, which benefit from the MiFID pre-trade transparency waivers, may interact with a displayed order book like CXE. The priority rules will normally be in accordance with price, visibility and time; thus visible orders at the same price as dark orders will execute ahead of a dark order. However, the inability to observe the dark order can be problematic, for retail investors for example. Consider a limit order book where bid and ask orders are arranged in order of price. Suppose I have a best displayed buy order of 99, a displayed limit order sell of 99 and a hidden (dark) buy order of 99.10. Although it is invisible to other traders, the dark buy order has execution priority as a consequence of its better bid price. Thus a sell limit order of 99 posted into the integrated lit order book will first attempt to fill its stated quantity against the hidden buy quantity available at 99.10. Hence dark liquidity can significantly induce the innovation of price process for stocks listed on lit exchanges.

Considering that dark trading in Europe has increased sharply in the post-financial crisis period,[18] it is not surprising that an entrant high-tech market that allows the interaction of lit and dark order books has become a significant leader of price discovery for the instruments listed on a different regulated exchange. Normally the above scenario should be immaterial for Chi-X's price leadership prospects as examined in this paper. This is because the Chi-X dark order book matches orders using quote midpoints as displayed on lit exchanges. Therefore, the dark book's contribution to price discovery should be minimal if it does contribute at all. However, since in the case of institutional traders it is unlikely that full liquidity requirements can be satisfied in one trade, the execution of a trade could signal the liquidity needs of large traders in the market (cf. Hirschey 2013). HFTs could exploit this inadvertent revelation about the trading intentions of institutional traders and front-run them. Although the institutional trades could be largely liquidity driven, trading on the basis of anticipating them constitutes an informed trading activity. Hence, dark trades can lead to a significant level of informed trading activity.

Consistent with the view that dark trades are validly informative, Comerton-Forde and Putniņš (2015) find that although dark orders are generally less informative than lit trades, they are not totally without information. As a result, the fact that dark trades have no pre-trade transparency implies that the market is not able to incorporate the information they harbour in a timely fashion, thereby impairing price discovery. Furthermore, Zhu (2014) holds that under natural conditions, uninformed traders gravitate towards dark pools, while informed traders largely

[18] According to the March 2014 edition of the Thomson Reuters Equity Market Share Reporter, dark pool and broker cross activity for all European equities accounted for more than €80.23 billion in March 2014, which is about 8.50% of all traded equities on the continent. For the 12-month period ending March 2014, the total value of dark trades stood at more than €898.22, which represents more than 9.55% of all equity trades in Europe.

trade on the lit markets. This manner of self-selection should result in Chi-X's dark order book contributing less to price discovery than its corresponding lit order book. Hence, one would expect to see a reduction in the proportion of price discovery share of Chi-X as its share of dark trading relative to its lit trading increases. Commensurately, with the dark end of the Chi-X venue siphoning away uninformed traders, one could expect to observe an increase in adverse selection costs in the lit market. It is therefore unsurprising that Chi-X's spreads are consistently wider than those of the LSE throughout the trading day as shown in Figure 1. The expectation that dark trading activity is inversely related to price discovery is formally tested/examined in Section 5.3.

As implied above, the price leadership potentials of Chi-X could also be linked to HFT activity Menkveld (2013) already suggests that the drivers of the success of markets like Chi-X are HFTs. And Carrion (2013) reports that HFTs actually do actively supply liquidity, which leads to widening spreads. The larger spreads (than LSE's) observed for BATS Chi-X in Figure 1 could therefore be as a result of HFTs supplying liquidity for most of the day. And the narrowing of the effective spread differences between BATS Chi-X and the LSE late in the day (see Figure 1) could be because of HFTs abandoning their positions by taking liquidity in order to achieve a neutral position by the end of the trading day. There is evidently an element of timing to this action, an issue on which Carrion (2013) has also presented evidence. Carrion (2013) suggests that HFTs possess intraday market timing ability, which may be related to Chi-X maintaining price leadership across the day. The timing ability of HFTs appears to be related with their ability to forecast non-HFTs' liquidity demands (see Hirschey 2013). Thus when HFTs increase their activity, especially when they mop up liquidity, they eliminate order imbalance and therefore enhance the information incorporation process (see as an example, Chaboud *et al.* 2014). Such heightened level of information processing may have been responsible for Chi-X maintaining price leadership throughout the day despite posting substantially inferior trading volumes. If this is the case, one would expect informed trading on Chi-X to be at generally higher levels than on the LSE. I examine these questions in the next section (5.2); in Section 5.3 I also formally link HFT activity to price discovery.

5.2. Informed Trading

The results so far suggest that informed traders, who are active on Chi-X throughout the trading day, dominate the release of trading-related information in London. However, activities of informed traders are also stepped up on the LSE as the trading day progresses. In what appears to be linked to institutional trading arrangements on the LSE, the activities of arbitrageurs/informed traders is most evident in prices towards the end of the day. As average trade sizes increase in later trading hours, the increased trading activity on the LSE allow informed participants the opportunity to better disguise their trading intentions (see Admati & Pfleiderer 1988). The idea that Chi-X leads price discovery is due to higher levels of informed trading activities is also underscored by the fact that throughout the trading day spreads on the platform are wider than those of the LSE. This is despite both markets trading the same instruments as contained in the sample employed in this study.

INSERT FIGURE 1 ABOUT HERE

I now turn my attention to substantiating the hypothesis that BATS Chi-X has higher levels of informed trading activity than the LSE. It is also expected that the first period of trading, which encompasses the first hour, will be highly informative, especially for lower volume quintile stocks.[19] In order to investigate the dynamics of informed trading across both the LSE and BATS Chi-X, I estimate the fraction of the aggregate trades during each trading period, which are based on information events. Hence, I investigate the relative participation rates of liquidity and informed traders during the various trading periods by employing the probability of informed trading (PIN) model of Easley *et al.* (1996) and Easley *et al.* (1997).[20] This model assumes that trading between informed traders, liquidity traders and market makers ensues repeatedly over numerous trading intervals. As presented in Figure 2, every trading interval commences with the informed traders obtaining a private signal on the value of the instrument with a probability of α. Subject to the arrival of a private signal, bad news will arrive with a probability of δ, while good news arrives with the probability of $(1 - \delta)$. The market maker in the mix determines bid and ask prices for his inventory, with orders arriving from liquidity traders at the arrival rate ε. If there is new information to act on, informed traders also trade and their orders arrive at the rate μ. Thus informed traders will buy if they receive a good news signal and sell if the signal is bad news. It is important to note that setting different arrival rates for uninformed buyers and sellers does not qualitatively alter estimations of the probability that an informed trade has occurred (see Easley *et al.* 2002).

By using the PIN model, I am able to infer the unobservable distribution of trades between informed and uninformed traders from buys and sells volume data.[21] Thus, the typical quantity of

[19] This is in line with recent evidence from the LSE as presented in Ibikunle (2015). Ibikunle (2015) reports that more than 50% of close-to-close price discovery occurs for all FTSE 100 stocks prior to 09:00:01hrs. While most of the price discovery (about 30% of the day's total) for the highest trading stocks occurs during the opening call auction (07:50 – 08:00hrs), a similar proportion of price discovery is recorded for lower volume stocks within the first 10 minutes of continuous trading (08:00 – 08:10hrs). Ibikunle (2015) argues that the results show the rapid pace of the incorporation of accumulated overnight information into the stock prices. The delay in the incorporation of information for the lower volume stocks is linked to their routine failure to open via the opening call auction (see also Friederich & Payne 2007). This argument is also consistent with the results obtained by Barclay and Hendershott (2003) and Jiang *et al.* (2012) in their analyses of price discovery after hours.

[20] For robustness, I also estimate adverse selection costs as proxy for informed trading by using the Huang and Stoll (1997) model; the inferences obtained from the PIN analysis are unchanged.

[21] Buy and sell trades are inferred by using the Lee and Ready (1991) trade classification algorithm.

buys and sells in an instrument is interpreted as uninformed trade by the model; this information is employed in estimating ε. An unusual level of buys or sells is interpreted as information trade and used in determining μ. Further, the frequency of intervals during which ‘abnormal’ levels of buys and sells are recorded is used to compute the values of α and δ. These computations are done simultaneously by using maximum likelihood estimation. If I assume that the uninformed and informed trades arrive via a Poisson distribution, the likelihood function for the PIN model for each interval estimated can be expressed as:

$$L((B,S)\mid\theta)=(1-\alpha)e^{-\varepsilon T}\frac{(\varepsilon T)^{B}}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^{S}}{S!}+\alpha\delta e^{-\varepsilon T}\frac{(\varepsilon T)^{B}}{B!}e^{-(\mu+\varepsilon)T}\frac{((\mu+\varepsilon)T)^{S}}{S!}$$
$$+\alpha(1-\delta)e^{-\varepsilon T}\frac{(\varepsilon T)^{S}}{S!}e^{-(\mu+\varepsilon)T}\frac{((\mu+\varepsilon)T)^{B}}{B!}, \qquad (16)$$

where B and S respectively correspond to the total number of buys and sells for the day within a trading interval; $\theta = (\alpha, \delta, \mu, \varepsilon)$ is the parameter vector for the structural model. Equation (16) represents a blend of distributions in which the possible trades are weighted by the probability of a day with no news $(1 - \alpha)$, a day with good news $(\alpha (1 - \delta))$ or a day with bad news $(\alpha\delta)$. Conditional on the assumption that this process occurs independently across days, Easley *et al.* (1996) and Easley *et al.* (1997) obtain the parameter vector estimates via maximum likelihood estimation. Therefore, the parameters for each of the trading intervals and for each stock in the sample are estimated by using maximum likelihood estimation. Following Easley *et al.* (1996) and Easley *et al.* (1997), PIN is computed as:

$$PIN=\frac{\alpha\mu}{\alpha\mu+2\varepsilon}. \qquad (17)$$

Table 5 presents the cross sectional means and standard deviations of the PIN estimates by pound volume quintile for both the LSE and BATS Chi-X. There are three main noticeable aspects of the results as presented, and all three are consistent with the preceding hypotheses on informed trading evolution across the two trading venues. Firstly, with only the exception of one trading interval (13:00 – 16:00hrs), Chi-X's overall PIN estimates are higher than those of the LSE; even for the said interval, the difference is not statistically significant. The stock-level estimates also show that Chi-X's level of informed trading activity is higher for most stocks and during most trading intervals. Most of the instances where LSE's PINs are higher than the corresponding ones for Chi-X are shown to be for those stocks and intervals where the ILS measure of price leadership implies that the LSE leads the information incorporation/price discovery process. This is unsurprising since informed trading is based on information about the fundamental value of instruments and ILS is based on eliminating noise to obtain a clean measure of which price series incorporates information about the fundamental value first. The overall mean PIN estimate for the LSE trading periods are 0.118, 0.133, 0.126, 0.073, 0.110, 0.108 and 0.125 for the respective seven trading intervals. Whereas the corresponding estimates for Chi-X are 0.162, 0.204, 0.171, 0.140, 0.162, 0.091 and 0.30. The differences between the corresponding estimates are statistically significant at the 0.01 level. The higher level of informed trading on Chi-X, which is consistent with the wider spreads (see Figure 1) on the exchange, could be explained by the activity of HFTs.

INSERT TABLE 5 ABOUT HERE

The property of HFT/AT that enhances the informativeness of trades on entrant high-tech markets is perhaps the speed of order submission, cancellation and transaction. By being able to trade at a fast pace, even on public signals, HFTs rapidly eliminate arbitrage opportunities and thus enhance price discovery (see Brogaard *et al.* 2014; Chaboud *et al.* 2014). Since the measures of price discovery used are based on which platform impounds new information into the price of instruments ahead of the competition, Chi-X with the typically faster traders should have the larger IS, CS and ILS estimates. This is a fairly sound explanation for why a Chi-X's order book (CXE) could favourably compete with the LSE's main order book (SETS) for price discovery given that SETS has significantly higher trading volumes.

A second striking feature of the results in Table 5 is the noticeable rise in the PIN estimates during the final half hour of trading. For example, the average LSE PIN during the 16:00 – 16:30hrs interval for the highest volume stocks is about 287% (0.157) of the PIN (0.055) during the previous interval (13:00 – 16:00hrs). The trend holds for the top three quintiles in the case of the LSE and for all of the quintiles in the case of Chi-X. The differences between the two periods' PIN estimates are also statistically significant at the 0.01 level for all the pound volume quintiles and the overall stock estimate. This development is consistent with the hypothesis that informed trading activity is stepped up during the last half-hour of continuous trading. On the LSE this trend applies to the highest volume stocks whose traders are the main users of the VWAP trading arrangement at the close of the trading day. Therefore it is not surprising that the trend holds for the top three quintiles on the LSE. This view is reinforced by the fact that the PIN estimates for the lower volume stocks (Quintiles 1 and 2) that are seldom traded via the VWAP mechanism experience are lower in the final trading interval than the previous trading interval. Furthermore, the LSE's average PIN for the highest volume stocks (0.157) during the final period is greater than that of Chi-X (0.112) for the same period at about 140% of the latter's estimate. Such significant rise in informed trading can explain why the LSE is able to increase its share of price discovery in the final half hour of trading.

The final key feature of the results in Table 5 relate to the PIN estimates during the early trading on both platforms. Consistent with prior expectations, the first hour of trading records high PIN estimates across all stocks trading on Chi-X, and to a lesser extent on the LSE. The case of the LSE is interesting because the lower volume stocks' trades appear to be more informative than those of the higher volume ones. This is consistent with Ibikunle's (2015) findings regarding the pattern of incorporation of information across the day. Since a large percentage (over 30%) of the daily information for the highest volume stocks would have already been incorporated in uncrossing prices yielded by the opening call auction, the early trades in the high volume stocks are unlikely to be as informative as those in the lower volumes ones. Overall, the expectation that there is a direct link between informed trading and intraday price discovery is holding. The next section formally explores this link along with several others already discussed in the preceding sections.

5.3. Determinants of New entrant Price discovery

So far I have presented several lines of argument in order to explain the price discovery dynamics observed in this section. It is imperative that these arguments be substantiated with some empirical evidence. Therefore, in this sub-section, I conduct a multivariate analysis using daily variables, which are computed from ultra-high frequency data. The empirical approach employed in this section includes computing a series of panel estimations relating the ILS to informed trading, dark trading and algorithmic trading through stock-day regressions. Panel estimations are run for each of the five quintiles and combined 47 stocks for each of the two markets – Chi-X and the LSE. The panel estimations are computed in two ways: (i) one-stage panel least squares regressions with panel corrected standard errors (PCSE) and fixed effects (stock and date); and (ii) panel generalised method of moments (GMM) instrumental variables (IV) regression with PCSE in order to obtain heteroscedasticity and autocorrelation robust standard errors.[22] The GMM models with IV are employed specifically to tackle the likelihood of the endogeneity of dark and

[22] Additionally, Newey-West standard errors are also obtained with virtually no differences observed across all results.

informed trading. However, the standard panel OLS with fixed effects results are also reported for one reason. Evidence from previous papers (see as an example, Comerton-Forde & Putniņš 2015) on the order execution approaches of traders implies that endogeneity is more of a concern when causally relating dark (and informed) trading to liquidity rather than to price discovery. This view is vindicated by the multivariate panel regression results presented for both the panel least squares and GMM estimations. The corresponding results for both estimation approaches are strikingly similar; indeed in some cases, the values are unchanged. The main difference in the results is with regards to the likelihood of obtaining statistically significant results. For the GMM estimations, we are more likely to obtain statistically significant results, although the explanatory powers of the models are virtually identical. Thus, it appears, consistent with Comerton-Forde and Putniņš (2015), that endogeneity does not significantly affect the least squares estimation results.

The panel regression estimated is of the following form:

$$ILS_{it} = \alpha + \beta_{PIN} PIN_{it} + \beta_{DARK} DARK_{it} + \beta_{HFT} HFT_{it} + \sum_{k=1}^{3} \varphi_k V_{kit} + \varepsilon_{it} \tag{18}$$

where ILS_{it} is the price discovery proxy, information leadership share, for stock i on day t, PIN_{it} is the proxy for informed trading for stock i on day t and $DARK_{it}$ is the log of pound volume of dark trades for stock i on day t.[23] HFT_{it} serves as a proxy for algorithmic trading and is measured as the ratio of messages to trades for stock i on day t. V_{kit} is a set of k control variables which includes log of pound trading volume, share of trading volume and log of effective spread. In the GMM IV estimations, a two-stage least squares (2SLS) weighting is used to weight the various GMM moment conditions to be satisfied by the parameters of interest. The moment conditions are

[23] Two other proxies are also used for dark trading: the first is the proportion of dark Chi-X trades in the stock relative to lit Chi-X trades in the stock for each day; and the second is the proportion of Chi-X's dark trades relative to the rest of the market trading the stock on each day. The results obtained from these two other variables are qualitatively similar to the ones presented in this version of the paper.

restricted to those that could be written as an orthogonality condition between the residuals of Equation (18) and its right-hand side variables.

Now turning to the task of identifying good candidates for both $DARK_{it}$ and PIN_{it}. The instruments must satisfy the condition of being correlated with the relevant variable to be instrumented and also be largely uncorrelated with e_{it} in Equation (18) above. The IVs are selected by extending an approach first employed by Hasbrouck and Saar (2013) and subsequently employed by several others such as Buti *et al.* (2011) and Degryse *et al.* (2015). The approach involves using trading variables in other similar stocks (specifically the averages) on day *t* as an instrument for the relevant trading variable in stock *i*. In this paper, I attempt an improvement on this approach in order to maximise the potential for the instruments to be orthogonal to the error term. As conceded by Hasbrouck and Saar (2013), the average across stocks may be correlated with stock *i* but it is also as likely to be correlated with the error term in Equation (18) for example.

Hence I aim to minimise this likelihood by first employing the initial averages of the trading variables across stocks in the same quintile, as already defined in earlier sections, in a panel least squares framework by regressing each of the endogenous variables on their corresponding cross-sectional stock averages and the other control variables; I then collect the residuals and employ them as IVs in the GMM estimation. The IVs yielded have the desirable properties of being highly correlated with the endogenous variables while being virtually uncorrelated with the residuals in Equation (18). The reason for the lack of correlation with the Equation (18) residuals is that the common cross-sectional component in the stock averages has been stripped off in explaining the changes in the endogenous variables leaving behind the stock-dependent factors not explained by the cross-sectional average. The two IVs for $DARK_{it}$ and PIN_{it}, named $DARKRES_{it}$ and $PINRES_{it}$ respectively are highly correlated with the respective endogenous variables (averages of 0.596 and 0.984 respectively, which are higher than the 0.521 average in Hasbrouck and Saar, 2013) because they are components of the dark and informed trading variables. However, they have very low average correlation coefficients with the residual in Equation (18); the respective average correlation coefficients for $DARKRES_{it}$ and $PINRES_{it}$ with ε_{it} are -0.006 and -0.002 respectively – effectively nil correlation. These two aforementioned properties implicitly render both $DARKRES_{it}$ and $PINRES_{it}$ very valuable IVs in the IV modelling framework. Given that the starting hypothesis of this paper is that pricing efficiency is the main driver of order flow shares, the share of trading volume is also instrumented in the manner described above. The resultant IV, $SHARERES_{it}$ is highly correlated with the original endogenous variable with an average correlation coefficient of 0.88, while the correlation coefficient with ε_{it} is a lowly average of -0.001. This latter coefficient is again effectively zero.[24]

Tables 6 – 9 report the panel least squares and GMM estimation results for all stock quintiles as well as for the overall sample. The estimation results are presented for both Chi-X and the LSE separately in two panels each for each of the tables. The results are strikingly similar across all four tables, strongly indicating that the panel least squares results are not driven by the effects of endogenously determined variables. Based on the evidence presented, the most consistently significant determinant of price discovery share in London markets is the level of informed trading in a market. The higher the proportion of informed trades in a market, the higher its share of price discovery. However, the effect of informed trading on price discovery share appears to be higher for the LSE; indeed the LSE coefficients are usually, on average, more than twice the size of the Chi-X estimates. For example, the Chi-X estimates for the full sample are

[24] For completeness, IVs based solely on Hasbrouck and Sarr's (2013) approach are also employed in the 2SLS/GMM estimations in this paper, all the results are qualitatively similar to the ones presented.

2.07, 2.06, 2.01 and 2.06 for Tables 6 to 9 respectively, whereas the corresponding values for the LSE are 5.29, 5.25, 5.22 and 5.25. All values are statistically significant at 0.0000 level. Thus informed trading activity on the LSE elicits higher efficient price impact than on Chi-X.

There are two possible explanations for this result. The first is that the relatively consistent trade size on Chi-X makes it harder to spot informed trades in the Chi-X order flow than it is on the LSE with relatively more discrimination in trade sizes across stocks. A possible second reason is the higher level of noise generated by the effects of algorithmic trading on Chi-X relative to the LSE. The averages of the HFT metric employed are 27.51 and 23.84 for Chi-X and LSE respectively, suggesting a higher level of algorithmic trading on Chi-X. The higher level of algorithmic trading could blunt informed trading impact making it harder to identify informed transactions since arbitrageurs have to sift through a larger volume of orders and trades. The HFT coefficient estimates in Tables 6 to 9 overwhelmingly suggest that high frequency trading impairs Chi-X's share of price discovery for the lower volume stocks. The full sample estimates are also statically significant in Tables 6, 7 and 9, and they are all negative. Contrarily, when the corresponding coefficient estimates are statistically significant for the LSE, they are all positive. This suggests that the level of algorithmic trading on Chi-X generally impairs its share of price discovery, while algorithmic trading on the LSE appears to spur price discovery shares. The only exceptions to this view is in the case of the most active stocks, Quintile 5 stocks in the tables. The coefficient estimates are significant and positive in Tables 6 and 9, they are also positive for the other two tables. This implies that, consistent with Brogaard *et al.* (2014) and Chaboud *et al.* (2014), for the highest trading stocks on Chi-X and all stocks trading on the LSE, HFT/AT improves pricing efficiency. One possible reason for the asymmetric effect of HFT activity in the case of Chi-X trades is the ratio of HFT activity generated to the actual volume of trades eventually executed at the two venues. Lower ratio of trades to quotes could imply higher difficulty in identifying information content of trade.

INSERT TABLES 6, 7, 8 and 9 ABOUT HERE

Now turning to the impact of dark trading on price discovery, it is obvious based on the evidence collected that the volume of dark trading adversely affects a venue's share of price discovery (compare with the reported negative impact on liquidity as reported by Degryse *et al.* 2015). Although most of the quintile-based estimates are not statistically significant. With the exception of the GMM results, only three of full sample estimates are statistically significant. Indeed all statistically significant estimates are negative and imply that dark trading does not increase a venue's share of price discovery. This impact of dark trading on share of price discovery partly explains why Chi-X's spreads are consistently wider than those of the LSE throughout the trading day. There are several reasons why dark trading on an exchange could hamper price discovery potential of that exchange's quotes. The first obvious reason is the lack of transparency occasioned by the inability of the public to observe a dark order's details prior to submission. Since only the submitter of a dark order has knowledge of its details, whatever information contained therein, no matter how small, will remain hidden until execution. And when such orders fail to execute the market will be none wiser regarding its existence.

Secondly, Zhu (2014) argues that under natural conditions uninformed traders are more likely to trade in the dark, and informed traders on the lit exchange. Although Zhu (2014) also argues that this self-selection, induced by differences in execution risk for informed and liquidity traders, should help reduce noise in the exchange as a whole and thus improve price discovery, an increasing level of interaction with a dark pool could lead to reduction in price discovery for that platform. This is because, as shown by Comerton-Forde and Putniņš (2015), dark orders are generally less informative than lit trades. Therefore, as traders increase their share of dark trading in Chi-X's dark pool relative to the share of informed trades in Chi-X's lit market, Chi-X's share of price discovery should reduce. Evidence from financial media suggests that most of the traders now piling into Chi-X's dark order book are in fact large institutional traders (asset managers) executing largely uninformed liquidity-induced orders (see as an example, Financial Times, 2015). As modelled by Zhu (2014), because liquidity orders are driven by idiosyncratic needs of the individual traders, they are less correlated than informed orders, which are correlated with the value of stocks. Hence, for an uninformed trader the risks associated with dark trading is lower than for informed traders, who could end up on the heavier side of the market thereby suffering non-execution or costly delays.

Thirdly, Degryse *et al.* (2015) maintains that the negative effect of dark trading on market quality characteristics such as liquidity is consistent with a 'cream skimming' effect, with dark order books attracting mainly uninformed orders thus increasing adverse selection costs on visible markets. This view is consistent with the generally wider spreads observed for Chi-X's lit market (see Figure 1).

There are two other findings in Tables 6, 7, 8 and 9 that should be touched upon. As speculated upon earlier in Section 5.1 and consistent with Huang (2002), it appears that the LSE's share of trading volume either has none or has a negative effect on its share of price discovery, whereas there is solid evidence in all but Table 8 that Chi-X's share of price discovery is positively linked with its share of trading volume. As stated earlier, the most likely explanation for this phenomenon is the presence of institutional trading priority arrangements at the LSE. These arrangements appear to invalidate the theoretically-established link between trading activity and price discovery (see as an example, Biais *et al.* 1999). The relevant idiosyncratic structure relates to LSE operating a hybrid market structure, where parallel downstairs and upstairs markets exist in parallel. Although the information revealed by the upstairs market is non-negligible as found by Armitage and Ibikunle (2015), when it is related to the level of trading value it contributes to the LSE total, it is very low. Thus, per unit pound the LSE's upstairs market reveals far lower information than most trading mechanisms.

Finally, the results consistently uphold the expectation, based on theory, that spreads are inversely related to price discovery (cf. Taylor 2011). Since liquidity is inextricably linked to market efficiency (see Chordia *et al.* 2008), one should expect the price discovery share of a venue to increase with increasing liquidity. Thus, when the spreads narrow, the price discovery process should improve. Furthermore, the farther apart bid and ask quotes are, the more onerous is the task of determining an equilibrium price.

5.4. Competition for Order flow: market share

The starting hypothesis in this paper is that informational quality of quotes is critical to the acquisition of high tech entrant market share. It is only fitting that this proposition be formally tested. Therefore, in order to directly test this hypothesis Equation (19) is estimated for Chi-X. As is the case with Equation (18), panel least squares and GMM estimation methods are employed.

$$Marketshare_{it} = \alpha + \beta_{ILS} ILS_{it} + \beta_{Tradesize} Tradesize_{it} + \beta_{Trades} Trades_{it} + \beta_{Volatility} Volatility_{it} + \varepsilon_{it} \tag{19}$$

GMM is used on account for possible endogeneity of the main variable of interest, ILS, which proxies the quality of quotes emanating from the trading process on Chi-X. The three other explanatory variables are selected in order to ensure consistency with existing literature regarding the determinants of trading venue market share (see for example, Kwan *et al.* 2015). $Marketshare_{it}$ corresponds to the log of share of pound volume of stock i traded on day t on Chi-X, ILS_{it} is the information leadership share of Chi-X with respect to stock *i* on day *t*. $Tradesize_{it}$ is the log of average daily trade size in pounds traded on Chi-X for stock *i* on day *t*, *Trades* is the log of number of transactions executed on Chi-X with respect to stock *i* on day *t*, while $Volatility_{it}$ is the log of standard deviation of the intraday mid-point price return for stock *i* on day *t*. With regards to the endogeneity of the ILS variable, the instrument employed is derived as described in the case of the three earlier instruments computed in Section 5.4.[25]

INSERT TABLE 10 ABOUT HERE

Table 10 presents the results from the estimation of Equation (19). Consistent with the expectation that informational quality of quotes is an important determinant of market share, the ILS variable is highly significant in three of four estimations, the exception being the stock fixed effects OLS estimation. Trade size and number of transactions are also significant determinants of market share. Volatility is only significant in the stock fixed effects estimation and is shown to have a negative impact on Chi-X's market share; this is consistent with lit exchange estimates obtained by Kwan *et al.* (2015).

6. Conclusion

This paper investigates the intraday evolution of the distribution of price discovery and informed trading between the two largest equity trading venues in Europe by modelling ultra-high frequency data for 47 of the most active European stocks.

[25] The Hasbrouck and Saar (2013) IV procedure is also adopted for robustness, the results are qualitatively similar to the ones obtained based on my IV selection process.

The LSE is still clearly the platform of choice for trading for most investors in FTSE 100 stocks. However, price discovery estimates presented in this paper imply that the superior trading volumes have not resulted in price leadership. The main contributions in this paper are six-fold.

First, I extend the price discovery literature by explicitly measuring the share of price discovery in relation to informed trading across the trading day. The findings show that price discovery is intraday time-varying (cf. Taylor 2011; Frijns *et al.* 2015). Results show that there are intraday variations in the price leadership between the LSE and Chi-X, with two of the employed price discovery metrics implying that Chi-X leads the price discovery process throughout the trading day. However, according to the ILS measure, the LSE leads the process of incorporating information about the fundamental value of stocks in the case of high and low volume stocks during two of the trading intervals examined, although the leads are not statistically significant in both cases. The variation in price leadership is closely linked with informed trading on both platforms across the day; econometric confirmation of this statement is presented in the paper. Informed trading on the LSE is found to lead to higher impact on efficient price than on Chi-X.

Secondly, this study presents evidence of asymmetric impact of algorithmic trading between established and new trading venues. Specifically, results suggest that algorithmic/high frequency trading impairs price discovery in the case of lower volume Chi-X traded instruments. For the highest trading stocks and those traded on the LSE the influence of algorithmic/high frequency trading is found to be mostly positive.

Thirdly, the impact of dark trading on lit price discovery is found to be negative in the case of a high tech entrant. This impact is linked to the higher level of liquidity risk for informed traders in dark pools; hence, informed traders gravitate towards lit venues leading to dark trades becoming inherently less informative than lit trades.

Fourthly, there appears to be a disconnection between the LSE's superior trading activity and its ability to lead the price discovery process across the day. Indeed, estimates suggest that LSE's share of trading volume is negatively linked to its share of price discovery. This phenomenon appears linked to the institutional parallel market trading arrangements on the LSE. The LSE operates a hybrid trading venue, which allows for delayed reporting of large institutional trades executed away from the downstairs SETS limit order book. The dealers responsible for executing the institutional trades away from the downstairs platform also have no obligation to post quotes, thereby making the upstairs institutional trades significantly less informative (cf. Armitage & Ibikunle 2015).

The fifth contribution of this paper is that estimates suggest that the LSE increases its share of price discovery across the day to close the day strongly despite still trailing Chi-X. This again could be linked to preferential institutional trading arrangements on the LSE. On the LSE, large institutional VWAP orders could be submitted for execution at the close. These trades only specify quantity and execute at the volume weighted average trading price for the trading day. Thus, their submission is an expression of trading intentions, which ultimately make no contribution to price discovery. It is therefore likely that traders attempt to influence the closing price, in order to obtain favourable VWAPs, by posting either informative or distorting quotes/orders during the minutes leading to the close. This view is underscored by the fact that the most active trading period on the LSE is the period leading to and around the close (see also Ibikunle, 2015).

Finally, evidence suggests that there is a greater concentration of informed traders on Chi-X than on the LSE. This further explains why Chi-X favourably contests with the LSE for price leadership despite the former's lower trading activity in the stocks examined. PIN analysis shows that Chi-X is more likely to host an informed trade than the LSE is. This implies that the bulk of

LSE participants are liquidity traders, while those on Chi-X are more proportionally informed. This is in line with financial media reports claiming that BATS Chi-X Europe aims to attract retail (liquidity/noise) traders with its attainment of RIE status in 2013 (see Financial-Times 2013).

The evidence in this paper should be of interest to platform operators, policy makers and trading market participants. Entrant high-tech markets are changing the landscape of instruments trading in Europe, as they have already done in the US. This paper shows that by attracting informed traders, high-tech entrants can favourably compete for price leadership with established superior volume platforms. Such level of competition help to secure further market share as suggested by this paper's analysis of market share determinants.

In conclusion, technological innovations and MiFID have both proven to be game changers with respect of order flow and price discovery distribution in Europe, and there is no telling what further changes directives like MiFID can bring, especially with a major facelift in the shape of MiFID II already in the works. The new regulations are expected to have far reaching implications for entrant high-tech markets because some of their provisions are being aimed at reducing the influence of HFTs and dark pools. Evidence in this paper suggests that curtailing the influence of HFTs and dark pools will not lead to significant reductions in the new entrants' contribution to efficient price discovery in European markets.

References

Admati, A., Pfleiderer, P., 1988. A theory of intraday patterns: volume and price variability. The Review of Financial Studies 1, 3-40

Aitken, M., Cumming, D., Zhan, F., 2014. Trade Size, High Frequency Trading, and Co-Location Around the World. The European Journal of Finance, Forthcoming

Armitage, S., Ibikunle, G., 2015. Price Discovery in London's Upstairs Equity Market. In: Univeristy of Edinburgh Business School Working Paper, Edinburgh

Bacidore, J.M., Sofianos, G., 2002. Liquidity provision and specialist trading in NYSE-listed non-U.S. stocks. Journal of Financial Economics 63, 133-158

Baillie, R.T., Geoffrey Booth, G., Tse, Y., Zabotina, T., 2002. Price discovery and common factor models. Journal of Financial Markets 5, 309-321

Barclay, M.J., Hendershott, T., 2003. Price Discovery and Trading After Hours. The Review of Financial Studies 16, 1041-1073

Barclay, M.J., Warner, J.B., 1993. Stealth trading and volatility: Which trades move prices? Journal of Financial Economics 34, 281-305

Biais, B., Hillion, P., Spatt, C., 1999. Price Discovery and Learning during the Preopening Period in the Paris Bourse. The Journal of Political Economy 107, 1218-1248

Boehmer, E., Fong, K., Wu, J., 2013. International Evidence on Algorithmic Trading. In: AFA 2013 San Diego Meetings, San Diego

Brogaard, J., 2010. High Frequency Trading and its Impact on Market Quality. Kellogg School of Management, Chicago, Illinois

Brogaard, J., Hendershott, T., Riordan, R., 2014. High-Frequency Trading and Price Discovery. Review of Financial Studies 27, 2267-2306

Buti, S., Rindi, B., Werner, I.M., 2011. Diving into dark pools. In: Working paper

Campbell, J.Y., Lo, A.W., Mackinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, New Jersey.

Carrion, A., 2013. Very fast money: High-frequency trading on the NASDAQ. Journal of Financial Markets 16, 680-711

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., Vega, C., 2014. Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. The Journal of Finance 69, 2045-2084

Chordia, T., Roll, R., Subrahmanyam, A., 2008. Liquidity and market efficiency. Journal of Financial Economics 87, 249-268

Comerton-Forde, C., Putniņš, T.J., 2015. Dark trading and price discovery. Journal of Financial Economics 118, 70-92

Cumming, D., Zhan, F., Aitken, M., 2012. High frequency trading and end-of-day price dislocation. York University Working Paper, Toronto, Ontario

Degryse, H., de Jong, F., Kervel, V.v., 2015. The Impact of Dark Trading and Visible Fragmentation on Market Quality. Review of Finance 19, 1587-1622

Easley, D., Hvidkjaer, S., O'Hara, M., 2002. Is Information Risk a Determinant of Asset Returns? The Journal of Finance 57, 2185-2221

Easley, D., Kiefer, N.M., O'Hara, M., 1996. Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow. The Journal of Finance 51, 811-833

Easley, D., Kiefer, N.M., O'Hara, M., 1997. One Day in the Life of a Very Common Stock. The Review of Financial Studies 10, 805-835

Easley, D., Lopez de Prado, M.M., O'Hara, M., 2011. The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes, and the Probability of Informed Trading. Journal of Portfolio Management 37, 118-128

Eun, C.S., Sabherwal, S., 2003. Cross-Border Listings and Price Discovery: Evidence from U.S.-Listed Canadian Stocks. The Journal of Finance 58, 549-575

Financial-Times, 2013. BATS Europe given UK exchange status. In: Financial Times. The Financial Times Limited, London

Financial-Times, 2015. Dark pool pitfalls getting deeper. In: FT Markets Stafford P (Ed.). The Financial Times Limited, London

Foucault, T., Menkveld, A.J., 2008. Competition for Order Flow and Smart Order Routing Systems. The Journal of Finance 63, 119-158

Friederich, S., Payne, R., 2007. Dealer Liquidity in an Auction Market: Evidence from the London Stock Exchange. The Economic Journal 117, 1168-1191

Frijns, B., Gilbert, A., Tourani-Rad, A., 2010. The dynamics of price discovery for cross-listed shares: Evidence from Australia and New Zealand. Journal of Banking & Finance 34, 498-508

Frijns, B., Indriawan, I., Tourani-Rad, A., 2015. Macroeconomic news announcements and price discovery: Evidence from Canadian–U.S. cross-listed firms. Journal of Empirical Finance 32, 35-48

Gonzalo, J., Granger, C.W.J., 1995. Estimation of common long-memory components in cointegrated systems. Journal of Business and Economic Statistics 13, 27-35

Hagströmer, B., Nordén, L., 2013. The diversity of high-frequency traders. Journal of Financial Markets 16, 741-770

Harris, F.H.d., McInish, T.H., Shoesmith, G.L., Wood, R.A., 1995. Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets. The Journal of Financial and Quantitative Analysis 30, 563-579

Harris, F.H.d., McInish, T.H., Wood, R.A., 2002. Common factor components versus information shares: a reply. Journal of Financial Markets 5, 341-348

Hasbrouck, J., 1995. One Security, Many Markets: Determining the Contributions to Price Discovery. The Journal of Finance 50, 1175-1199

Hasbrouck, J., Saar, G., 2013. Low-latency trading. Journal of Financial Markets 16, 646-679

Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does Algorithmic Trading Improve Liquidity? The Journal of Finance 66, 1-33

Hirschey, N., 2013. Do High-Frequency Traders Anticipate Buying and Selling Pressure? London Business School Working Paper

Huang, R.D., 2002. The Quality of ECN and Nasdaq Market Maker Quotes. The Journal of Finance 57, 1285-1319

Huang, R.D., Stoll, H.R., 1997. The Components of the Bid-Ask Spread: A General Approach. The Review of Financial Studies 10, 995-1034

Hupperets, E.C.J., Menkveld, A.J., 2002. Intraday analysis of market integration: Dutch blue chips traded in Amsterdam and New York. Journal of Financial Markets 5, 57-82

Ibikunle, G., 2015. Opening and closing price efficiency: Do financial markets need the call auction? Journal of International Financial Markets, Institutions and Money 34, 208-227

Jiang, C.X., Likitapiwat, T., McInish, T.H., 2012. Information Content of Earnings Announcements: Evidence from After-Hours Trading. Journal of Financial and Quantitative Analysis 47, 1303-1330

Jovanovic, B., Menkveld, A.J., 2012. Middlemen in Limit-Order Markets. New York University Working Paper

Kadapakkam, P.-R., Misra, L., Tse, Y., 2003. International Price Discovery for Emerging Market Stocks: Evidence from Indian GDRs. Review of Quantitative Finance and Accounting 21, 179-199

Kirilenko, A.A., Kyle, A.S., Samadi, M., Tuzun, T., 2011. The Flash Crash: The Impact of High Frequency Trading on an Electronic Market. SSRN eLibrary Working Paper

Korczak, P., Phylaktis, K., 2010. Related securities and price discovery: Evidence from NYSE-listed Non-U.S. stocks. Journal of Empirical Finance 17, 566-584

Kwan, A., Masulis, R., McInish, T.H., 2015. Trading rules, competition for order flow and market fragmentation. Journal of Financial Economics 115, 330-348

Kyle, A.S., 1985. Continuous Auctions and Insider Trading. Econometrica 53, 1315-1335

Lee, C.M., Ready, M.J., 1991. Inferring trade direction from intraday data. The Journal of Finance 46, 733-746

Lieberman, O., Ben-Zion, U., Hauser, S., 1999. A characterization of the price behavior of international dual stocks: an error correction approach. Journal of International Money and Finance 18, 289-304

Lok, E., Kalev, P.S., 2006. The intraday price behaviour of Australian and New Zealand cross-listed stocks. International Review of Financial Analysis 15, 377-397

Malinova, K., Park, A., 2014. The impact of competition and information on intraday trading. Journal of Banking & Finance 44, 55-71

Menkveld, A.J., 2013. High frequency trading and the new market makers. Journal of Financial Markets 16, 712-740

O'Hara, M., Ye, M., 2011. Is market fragmentation harming market quality? Journal of Financial Economics 100, 459-474

Pagano, M., 1989. Trading Volume and Asset Liquidity. The Quarterly Journal of Economics 104, 255-274

Pascual, R., Pascual-Fuster, B., Climent, F., 2006. Cross-listing, price discovery and the informativeness of the trading process. Journal of Financial Markets 9, 144-161

Putniņš, T.J., 2013. What do price discovery metrics really measure? Journal of Empirical Finance 23, 68-83

Roope, M., Zurbruegg, R., 2002. The intra-day price discovery process between the Singapore Exchange and Taiwan Futures Exchange. Journal of Futures Markets 22, 219-240

Schacht, K., Cronin, C., Allen, C.J., Preece, R., 2009. Market Microstructure: The Impact of Fragmentation under the Markets in Financial Instruments Directive. CFA Institute Centre for Financial Market Integrity, Charlottesville, VA

Stock, J.H., Watson, M.W., 1988. Testing for Common Trends. Journal of the American Statistical Association 83, 1097-1107

Su, Q., Chong, T.T.-L., 2007. Determining the contributions to price discovery for Chinese cross-listed stocks. Pacific-Basin Finance Journal 15, 140-153

Taylor, N., 2011. Time-varying price discovery in fragmented markets. Applied Financial Economics 21, 717-734

Wang, J., Yang, M., 2011. Housewives of Tokyo versus the gnomes of Zurich: Measuring price discovery in sequential markets. Journal of Financial Markets 14, 82-108

Yan, B., Zivot, E., 2010. A structural analysis of price discovery measures. Journal of Financial Markets 13, 1-19

Zhang, S.S., 2012. Need for Speed: An Empirical Analysis of Hard and Soft Information in a High Frequency World. Karlsruhe Institute of Technology Working Paper

Zhu, H., 2014. Do Dark Pools Harm Price Discovery? Review of Financial Studies 27, 747-789

Figure 1: Liquidity per minute for FTSE 100 Stocks

Liquidity proxies per minute are computed for 47 FTSE 100 stocks trading on the London Stock Exchange and BATS Chi-X between 1st July 2014 and 28th November 2014. The Quoted spread is the difference between the prevailing ask and bid prices at the time of the last transaction at every 1 minute mark for each stock; the spreads are then averaged cross-sectionally across stocks. The Effective spread is measured as twice the absolute value of the difference between the last transaction price at every 1 minute interval and the corresponding midpoint of the prevailing ask and bid prices at the time of that transaction; the spreads are then averaged cross-sectionally across stocks. The time covered is from 08:00:00hrs- 16:30:00hrs London local time. Quintiles are computed on the basis of daily pound volume across the sample period.

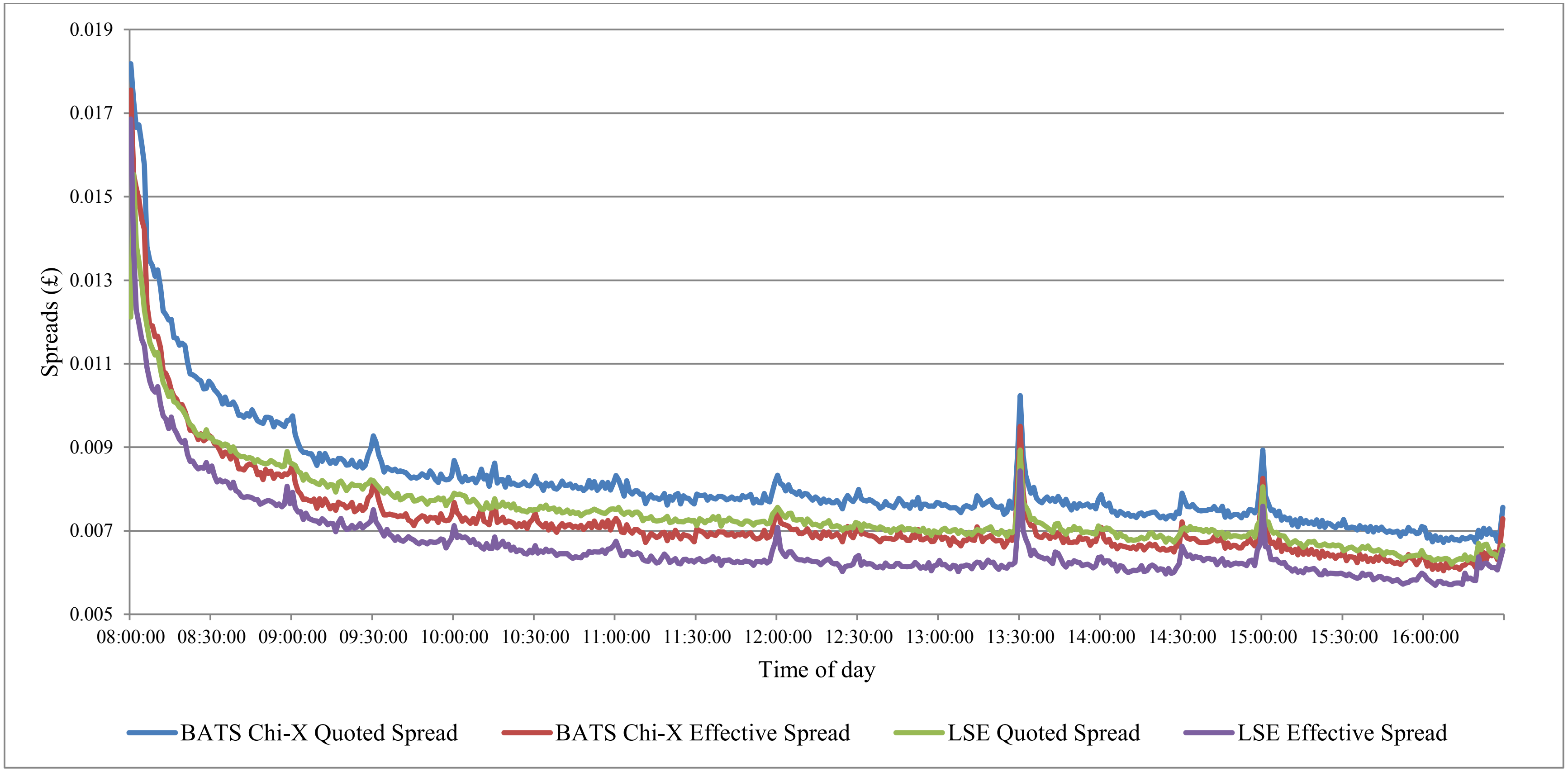

Figure 2: Tree Diagram for the Easley, Kiefer and O'Hara (1996, 1997)

A corresponds to the probability of an information event, δ represents the probability that a low signal ensues, μ is the arrival rate of informed orders, and ε is the arrival rate of uninformed orders. The nodes to the left of the thick vertical line occur only once a day.

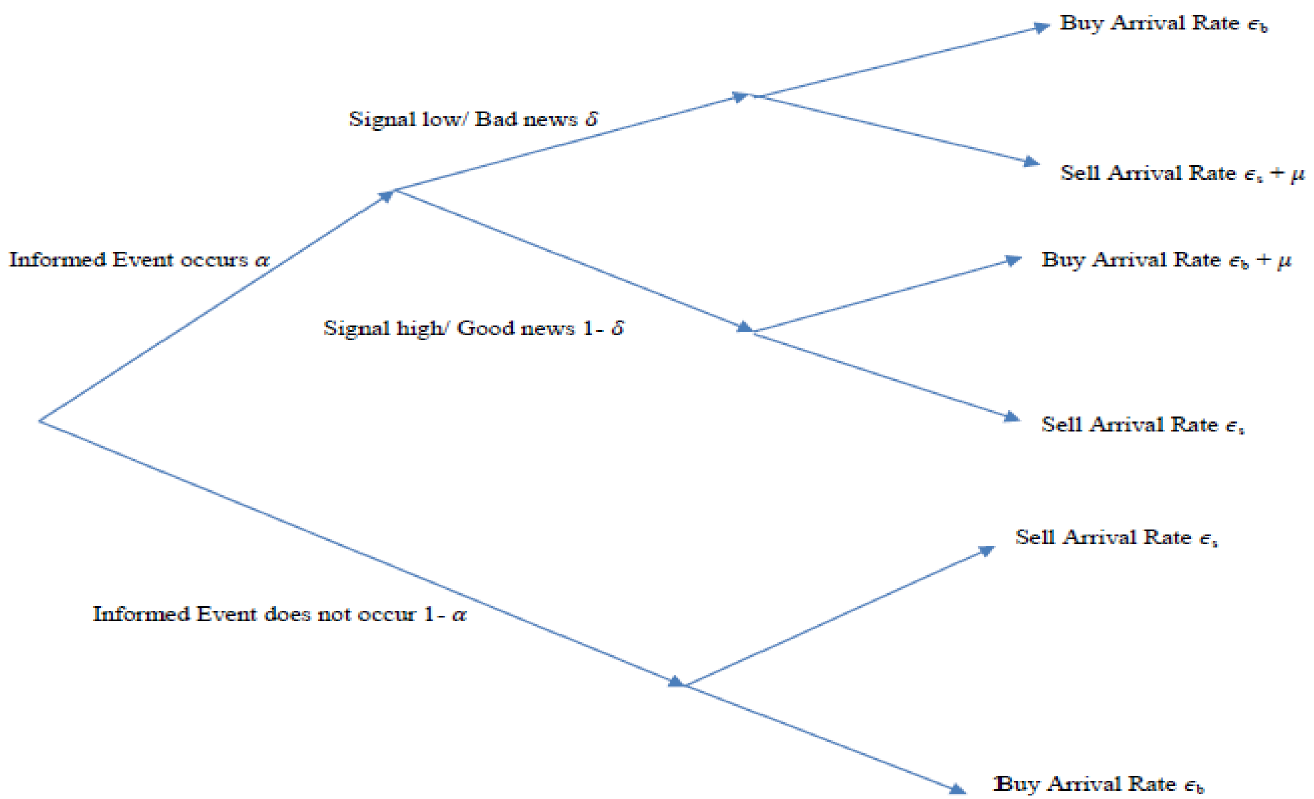

Table 1: Daily Trading Summary Statistics for FTSE 100 Stocks

Panels A and B present daily summary statistics for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books respectively. The sample period covers 1st July 2014 and 28th November 2014. The quintiles are computed on the basis of daily pound volume across the sample period.

Panel A: SETS

		Highest	4	3	2	Lowest	Overall
Number of Transactions	08:00 - 08:30	247.79	199.79	116.75	87.95	53.16	138.09
	08:30 - 09:00	176.48	143.69	86.44	65.99	40.72	100.57
	09:00 - 10:00	313.34	256.35	151.26	118.67	71.92	178.61
	10:00 - 12:00	536.04	426.65	257.34	196.56	122.32	301.47
	12:00 - 13:00	233.54	185.97	109.74	83.97	51.92	130.26
	13:00 - 16:00	1,084.61	831.7	480.5	360.55	228.41	584.27
	16:00 - 16:30	287.36	218.7	145.17	113.48	75.22	164.85
	All	2,879.16	2,262.86	1,347.21	1,027.17	643.66	1,598.11
Pound Volume (£'000,000)	08:00 - 08:30	8.77	5.11	3.42	1.94	1.04	3.95
	08:30 - 09:00	5.10	3.12	2.17	1.45	0.80	2.47
	09:00 - 10:00	8.73	5.6	4.01	2.61	1.42	4.37
	10:00 - 12:00	14.83	9.63	7.13	4.32	2.54	7.51
	12:00 - 13:00	6.71	4.36	2.99	1.85	1.18	3.34
	13:00 - 16:00	33.98	21.54	15.44	9	5.53	16.68
	16:00 - 16:30	10.02	7.64	4.66	2.93	1.89	5.3
	All	88.15	57.01	39.81	24.1	14.39	43.61
Average Trade Size (£'000)	08:00 - 08:30	35.41	25.59	29.28	22.01	19.54	28.58
	08:30 - 09:00	28.89	21.71	25.09	22.01	19.56	24.54
	09:00 - 10:00	27.87	21.84	26.49	21.99	19.74	24.46
	10:00 - 12:00	27.67	22.58	27.70	21.99	20.76	24.91
	12:00 - 13:00	28.75	23.44	27.22	22.03	22.64	25.61
	13:00 - 16:00	31.33	25.90	32.14	24.95	24.22	28.55
	16:00 - 16:30	34.87	34.94	32.09	25.83	25.13	32.15
	All	30.62	25.19	29.55	23.46	22.36	27.29

Panel B: CXE

		Highest	4	3	2	Lowest	Overall
Number of Transactions	08:00 - 08:30	142.52	102.02	51.42	37.55	26.79	70.36
	08:30 - 09:00	122.1	91.08	45.06	34.38	25.16	62.12
	09:00 - 10:00	229	167.37	84.17	66.25	47.18	116.15
	10:00 - 12:00	389.94	283.09	139.7	106.71	78.25	194.98
	12:00 - 13:00	204.87	145.43	71.81	55.68	41.95	101.6
	13:00 - 16:00	932.72	626.87	321.35	221.71	172.58	444.07
	16:00 - 16:30	53.85	43.85	19.70	12.82	11.24	27.60
	All	2,075.01	1,459.72	733.22	535.09	403.15	1,016.89
Pound Volume (£'000,000)	08:00 - 08:30	2.13	1.26	0.71	0.45	0.27	0.94
	08:30 - 09:00	1.8	1.13	0.63	0.43	0.27	0.83
	09:00 - 10:00	3.51	2.18	1.27	0.88	0.53	1.63
	10:00 - 12:00	6.07	3.84	2.19	1.45	0.93	2.82
	12:00 - 13:00	3.22	1.97	1.11	0.76	0.51	1.48
	13:00 - 16:00	16.76	9.83	5.87	3.49	2.33	7.45
	16:00 - 16:30	1.06	0.75	0.41	0.24	0.16	0.51
	All	34.55	20.97	12.2	7.7	4.99	15.67
Average Trade Size (£'000)	08:00 - 08:30	14.97	12.36	13.82	11.98	10.09	13.35
	08:30 - 09:00	14.78	12.44	14.07	12.42	10.63	13.39
	09:00 - 10:00	15.31	13.01	15.07	13.33	11.18	14.05
	10:00 - 12:00	15.56	13.58	15.66	13.54	11.84	14.47
	12:00 - 13:00	15.7	13.55	15.49	13.68	12.18	14.54
	13:00 - 16:00	17.97	15.68	18.27	15.72	13.5	16.79
	16:00 - 16:30	19.7	17.04	20.77	18.95	14.56	18.52
	All	16.65	14.36	16.63	14.38	12.39	15.41

Table 2: Information Shares of FTSE 100 Stocks for SETS and CXE

The table presents quintile mean information shares (IS) and standard deviations in parentheses for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The IS for both trading venues are computed per stock and for each interval on the basis of Eqs. (12) and (13). The information shares for each stock's trading periods are then cross-sectionally averaged across stocks for each period and for each quintile to obtain the IS for each quintile and for each trading period across the trading day. Wilcoxon/Mann-Whitney and Kruskal-Wallis tests are employed to test the null of no differences between the two venues' corresponding trading intervals. A venue's interval IS estimate which differ from the other venue's corresponding interval at the 0.001 level are denoted with *. † denotes the venue with the larger share of price discovery during a trading interval. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

		Highest	4	3	2	Lowest	Overall
London Stock Exchange	08:00 - 08:30	3.33* (1.73)	2.37* (1.34)	3.13* (1.48)	2.65* (1.28)	1.84* (0.82)	2.62* (1.44)
	08:30 - 09:00	38.39* (3.43)	37.43* (2.76)	36.73* (2.78)	34.71* (2.94)	36.77* (2.08)	36.76* (3.07)
	09:00 - 10:00	37.51* (2.15)	37.16* (2.07)	35.78* (3.63)	32.84* (2.47)	36.14* (1.13)	35.82* (2.93)
	10:00 - 12:00	32.78* (3.82)	29.69* (5.90)	37.64* (1.41)	26.41* (2.52)	29.56* (3.08)	29.86* (4.78)
	12:00 - 13:00	40.86* (4.09)	41.87* (2.15)	40.43* (2.77)	38.20* (4.52)	39.85* (1.39)	40.14* (3.47)
	13:00 - 16:00	25.08* (8.57)	33.37* (7.41)	22.81* (2.72)	23.83* (2.50)	24.41* (3.08)	25.77* (6.60)
	16:00 - 16:30	46.40* (0.45)	45.91* (1.19)	44.91* (1.91)	44.80* (1.11)	44.84* (1.11)	45.29* (1.44)
BATS Chi-X	08:00 - 08:30	96.67†* (1.73)	97.63†* (1.34)	96.87†* (1.48)	97.35†* (1.28)	98.16†* (0.82)	97.38†* (1.44)
	08:30 - 09:00	61.61†* (3.43)	62.57†* (2.76)	63.27†* (2.78)	65.29†* (2.94)	63.23†* (2.08)	63.24†* (3.07)
	09:00 - 10:00	62.49†* (2.15)	62.84†* (2.07)	64.22†* (3.63)	67.16†* (2.47)	63.86†* (1.13)	64.18†* (2.93)
	10:00 - 12:00	67.22†* (3.82)	70.31†* (5.90)	62.36†* (1.41)	73.59†* (2.52)	70.44†* (3.08)	70.14†* (4.78)
	12:00 - 13:00	59.14†* (4.09)	58.13†* (2.15)	59.57†* (2.77)	61.80†* (4.52)	60.15†* (1.39)	59.86†* (3.47)
	13:00 - 16:00	74.92†* (8.57)	66.63†* (7.41)	77.19†* (2.72)	76.17†* (2.50)	75.59†* (3.08)	74.23†* (6.60)
	16:00 - 16:30	53.60†* (0.45)	54.09†* (1.19)	55.09†* (1.91)	55.20†* (1.11)	55.16†* (1.11)	54.71†* (1.44)

Table 3: Component Shares of FTSE 100 Stocks for SETS and CXE

The table presents quintile mean component shares (CS) and standard deviations in parentheses for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. CS estimates are computed for each stock and time interval by estimating the following vector error correction model (VECM):

$$\Delta P_t^{LSE} = \alpha_0^{LSE} + \alpha_1^{LSE}\hat{u}_{t-1} + \sum_{i=1}^{p}\gamma_i \Delta P_{t-i}^{LSE} + \sum_{i=1}^{p}\vartheta_i \Delta P_{t-i}^{BCE} + \varepsilon_t^{LSE},$$
$$\Delta P_t^{BCE} = \alpha_0^{BCE} + \alpha_1^{BCE}\hat{u}_{t-1} + \sum_{i=1}^{p}\gamma_i \Delta P_{t-i}^{BCE} + \sum_{i=1}^{p}\vartheta_i \Delta P_{t-i}^{LSE} + \varepsilon_t^{BCE}.$$

where P_t^{LSE} and P_t^{BCE} are observable price processes from the SETS and CXE respectively. The LSE and BATS Chi-X CS estimates for each stock and within each interval are then computed:

$$CS_j^{LSE} = \frac{\alpha_1^{BCE}}{\alpha_1^{BCE} - \alpha_1^{LSE}} \quad \text{and} \quad CS_j^{BCE} = 1 - CS_j^{LSE} \text{ respectively.}$$

Adjusted Median Chi-Square tests are employed to test the null of no differences between the two venues' corresponding trading intervals. A venue's interval CS estimate which differ from the other venue's corresponding interval at 0.05 level of statistical significance are denoted with *. † denotes the venue with the larger share of price discovery during a trading interval. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

		Highest	4	3	2	Lowest	Overall
London Stock Exchange	08:00 - 08:30	4.52* (0.77)	22.73* (12.95)	30.17* (18.42)	17.89* (8.92)	32.07 (27.02)	23.60* (19.23)
	08:30 - 09:00	– –	– –	– –	– –	– –	– –
	09:00 - 10:00	46.12* (2.37)	36.80* (1.92)	40.12* (0.03)	35.41* (1.87)	43.48* (4.48)	39.51* (4.99)
	10:00 - 12:00	– –	42.55* (0.00)	– –	16.15* (0.07)	15.54* (0.77)	21.18* (10.70)
	12:00 - 13:00	– –	43.61* (6.13)	50.90† (0.40)	35.60* (6.25)	– –	42.26* (8.27)
	13:00 - 16:00	12.04* (11.23)	23.08* (4.04)	22.03* (1.20)	26.15* (7.09)	27.16* (10.20)	24.52* (9.20)
	16:00 - 16:30	– –	– –	– –	– –	– –	– –
BATS Chi-X	08:00 - 08:30	95.48†* (0.77)	77.27†* (12.95)	69.83†* (18.42)	82.11†* (8.92)	67.93† (27.02)	76.40†* (19.23)
	08:30 - 09:00	– –	– –	– –	– –	– –	– –
	09:00 - 10:00	53.88†* (2.37)	63.20†* (1.92)	59.88†* (0.03)	64.59†* (1.87)	56.52†* (4.48)	60.49†* (4.99)
	10:00 - 12:00	– –	57.45†* (0.00)	– –	83.85†* (0.07)	84.46†* (0.77)	78.82†* (10.70)
	12:00 - 13:00	– –	56.39†* (6.13)	49.10†* (0.40)	64.40†* (6.25)	– –	57.74† (8.27)

13:00 - 16:00	87.96†* (11.23)	76.92†* (4.04)	77.97†* (1.20)	73.85†* (7.09)	72.84†* (10.20)	75.48†* (9.20)
16:00 - 16:30	– –	– –	– –	– –	– –	– –

Table 4: Information Leadership Shares of FTSE 100 Stocks for SETS and CXE

The table presents quintile mean information leadership shares (ILS) for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The ILS is computed as follows:

$$ILS_j^{LSE} = \frac{\left|\frac{IS_j^{LSE}}{IS_j^{BCE}}\frac{CS_j^{BCE}}{CS_j^{LSE}}\right|}{\left|\frac{IS_j^{LSE}}{IS_j^{BCE}}\frac{CS_j^{BCE}}{CS_j^{LSE}}\right| + \left|\frac{IS_j^{BCE}}{IS_j^{LSE}}\frac{CS_j^{LSE}}{CS_j^{BCE}}\right|} \qquad ILS_j^{BCE} = \frac{\left|\frac{IS_j^{BCE}}{IS_j^{LSE}}\frac{CS_j^{LSE}}{CS_j^{BCE}}\right|}{\left|\frac{IS_j^{LSE}}{IS_j^{BCE}}\frac{CS_j^{BCE}}{CS_j^{LSE}}\right| + \left|\frac{IS_j^{BCE}}{IS_j^{LSE}}\frac{CS_j^{LSE}}{CS_j^{BCE}}\right|}$$

where IL_j^{LSE} and IL_j^{BCE} correspond to the information leadership share with respect to stock *j* for SETS and CXE respectively. IS_j^{LSE} and IS_j^{BCE} represent the IS with respect to stock j for SETS and CXE respectively, while CS_j^{LSE} and CS_j^{BCE} correspond to the CS with respect to stock j for SETS and CXE respectively. Adjusted Median Chi-Square and Kruskal-Wallis tests are employed to test the null of no differences between the two venues' corresponding trading intervals. A venue's interval ILS estimate which differ from the other venue's corresponding interval at the 0.01 level is denoted with *. † denotes the venue that is fastest in impounding information about the fundamental value of stocks during a trading interval. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

		Highest	4	3	2	Lowest	Overall
London Stock Exchange	08:00 - 08:30	33.36 (17.80)	2.65* (3.66)	7.83* (13.70)	4.06* (7.28)	6.09* (11.26)	7.52* (13.48)
	08:30 - 09:00	–	–	–	–	–	–
	09:00 - 10:00	29.82* (0.50)	50.29† (3.40)	33.85* (0.72)	43.69* (2.52)	35.28* (7.97)	40.00* (8.42)
	10:00 - 12:00	–	8.13* (3.36)	–	75.00†* (1.43)	75.71†* (2.32)	58.63† (29.78)
	12:00 - 13:00	–	54.75†* (1.10)	26.62* (4.85)	47.29 (3.36)	–	41.64 (11.54)
	13:00 - 16:00	72.55†* (11.58)	72.42†* (16.49)	41.38* (4.74)	45.92 (16.63)	48.14 (23.86)	52.07† (21.50)
	16:00 - 16:30	–	–	–	–	–	–
BATS Chi-X	08:00 - 08:30	66.64†* (17.80)	97.35†* (3.66)	92.17†* (13.70)	95.94†* (7.28)	93.91†* (11.26)	92.48†* (13.48)
	08:30 - 09:00	–	–	–	–	–	–
	09:00 - 10:00	70.18†* (0.50)	49.71 (3.40)	66.15†* (0.72)	56.31†* (2.52)	64.72†* (7.97)	60.00†* (8.42)
	10:00 - 12:00	–	91.87†* (3.36)	–	25.00* (1.43)	24.29* (2.32)	41.37 (29.78)
	12:00 - 13:00	–	45.25* (1.10)	73.38†* (4.85)	52.71† (3.36)	–	58.36† (11.54)
	13:00 - 16:00	27.45* (11.58)	27.58* (16.49)	58.62†* (4.74)	54.08† (16.63)	51.86† (23.86)	47.93 (21.50)
	16:00 - 16:30	–	–	–	–	–	–

Table 5: Probability of Informed Trading Analysis

The table presents quintile mean probability of informed trading (PIN) and standard deviations in parentheses for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The PIN estimates are computed per stock and for each trading interval by using the Easley et al. (1996, 1997) PIN model. PIN parameters are computed for each stock and time interval by maximising the following likelihood function:

$$L((B,S)\,|\,\theta) = (1-\alpha)e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} + \alpha\delta e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-(\mu+\varepsilon)T}\frac{((\mu+\varepsilon)T)^S}{S!}$$
$$+\,\alpha(1-\delta)e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!}e^{-(\mu+\varepsilon)T}\frac{((\mu+\varepsilon)T)^B}{B!},$$

where B and S respectively correspond to the total number of buy and sell orders for the day within each trading interval. $\theta = (\alpha, \delta, \mu, \varepsilon)$ is the parameter vector for the model. α corresponds to the probability of an information event, δ is the conditional probability of a low signal of an information event, μ is the arrival rate of informed orders, and ε is the arrival rate of uninformed orders. The probability that a trade is informed for each stock and within each interval is then computed as:

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}.$$

Wilcoxon/Mann-Whitney tests are used to test the null of no differences among the time intervals and between the two trading venues. Trading interval mean PIN estimate which differ from the immediate past interval's mean PIN at 0.01 (0.05) level for the BATS Chi-X (LSE) trading venue are denoted with *. A venue's interval PIN value which differs from the other venue's corresponding interval's PIN at the 0.05 level is denoted with †. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

		Highest	4	3	2	Lowest	Overall
London Stock Exchange	08:00 - 08:30	0.0899† (0.015)	0.0972† (0.009)	0.0983† (0.016)	0.1768† 0.0520	0.1204† (0.008)	0.1178† (0.042)
	08:30 - 09:00	0.0929† (0.018)	0.1953†* (0.028)	0.1083† (0.012)	0.1182†* (0.010)	0.1541†* (0.059)	0.1334† (0.048)
	09:00 - 10:00	0.1375* (0.028)	0.0742†* (0.013)	0.0920† (0.017)	0.1726†* (0.038)	0.0975* (0.015)	0.1142† (0.044)
	10:00 - 12:00	0.0649* (0.006)	0.0562†* (0.006)	0.0554†* (0.008)	0.0861†* (0.015)	0.0989† (0.016)	0.0726†* (0.020)
	12:00 - 13:00	0.0819* (0.027)	0.1505* (0.032)	0.1070†* (0.019)	0.1041†* (0.020)	0.1136† (0.020)	0.1104†* (0.032)
	13:00 - 16:00	0.0547†* (0.008)	0.0682* (0.010)	0.1314†* (0.013)	0.1292* (0.015)	0.1636* (0.011)	0.1077 (0.043)
	16:00 - 16:30	0.1572†* (0.027)	0.1288†* (0.033)	0.1498† (0.030)	0.0959†* (0.031)	0.0946†* (0.013)	0.1246† (0.038)
BATS Chi-X	08:00 - 08:30	0.1772† (0.049)	0.1361† (0.094)	0.1799† (0.080)	0.1577† (0.067)	0.1587† (0.063)	0.1617†* (0.073)
	08:30 - 09:00	0.1844† (0.088)	0.1448† (0.051)	0.3281†* (0.132)	0.1758† (0.080)	0.1901†* (0.106)	0.2037†* (0.114)
	09:00 - 10:00	0.1525 (0.036)	0.1951†* (0.052)	0.2609†* (0.093)	0.1412†* (0.058)	0.1113* (0.055)	0.1706†* (0.081)
	10:00 - 12:00	0.0608*	0.0861†*	0.1158†*	0.2185†*	0.2099†*	0.1398†*

	(0.033)	(0.029)	(0.044)	(0.090)	(0.069)	(0.087)
12:00 - 13:00	0.0847	0.1472*	0.2026†*	0.2266†	0.1506†*	0.1621†*
	(0.032)	(0.062)	(0.095)	(0.134)	(0.060)	(0.097)
13:00 - 16:00	0.0320†*	0.0536*	0.0766†*	0.1246*	0.1548	0.0907*
	(0.013)	(0.014)	(0.038)	(0.059)	(0.062)	(0.063)
16:00 - 16:30	0.1119†*	0.1813†*	0.2756†*	0.4119†*	0.4676†*	0.2960†*
	(0.039)	(0.036)	(0.096)	(0.068)	(0.038)	(0.148)

Table 6: Effects of informed, algorithmic and dark trading on information leadership I

The table reports panel least squares regression coefficient estimates using a stock-day panel, in which the dependent variable is the information leadership share (ILS), for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The estimated regressions is:

$$ILS_{it} = \alpha + \beta_{PIN} PIN_{it} + \beta_{DARK} DARK_{it} + \beta_{HFT} HFT_{it} + \sum_{k=1}^{3} \varphi_k V_{kit} + \varepsilon_{it}$$

where ILS_{it} is the price discovery proxy, information leadership share for stock *i* on day *t* and is as defined in Table 4. PIN_{it} is the proxy for informed trading for stock *i* on day *t* and is as defined in Table 5. $DARK_{it}$ is the log of pound volume of dark trades for stock *i* on day *t*. HFT_{it} serves as proxy for algorithmic trading and is the messages to trades ratio for stock *i* on day *t*. V_{kit} is a set of *k* control variables which include log of pound trading volume, share of trading volume and log of effective spread. The t-statistics are presented in parentheses and derived from panel corrected standard errors (PCSE). *, ** and *** correspond to statistical significance at 0.1, 0.05 and 0.01 levels respectively. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

	Panel A: BATS Chi-X Europe						Panel B: London Stock Exchange					
	Highest	4	3	2	Lowest	Full sample	Highest	4	3	2	Lowest	Full sample
Intercept	5.57 (0.50)	30.41 (2.93***)	28.74 (2.76***)	-43.87 (-2.06**)	20.33 (4.34***)	29.90 (6.57***)	42.93 (6.75***)	32.24 (4.04***)	11.00 (1.61)	-29.76 (-1.54)	-22.99 (-1.80*)	28.44 (3.20***)
PIN	1.47 (11.1***)	3.16 (14.0***)	1.68 (16.3***)	1.58 (11.6***)	5.12 (26.7***)	2.07 (28.3***)	3.72 (17.0***)	4.23 (15.7***)	9.65 (32.5***)	4.89 (18.7***)	13.41 (38.6***)	5.29 (34.7***)
Dark	0.53 (1.09)	-0.65 (-1.20)	-0.89 (-1.63)	-0.45 (-0.66)	-0.50 (-0.90)	-0.51 (-2.38**)	–	–	–	–	–	–
HFT	0.07 (2.05**)	-0.13 (-4.7***)	-0.01 (0.42)	-0.00 (-0.08)	-0.03 (-1.02)	-0.29 (-2.10**)	-0.03 (-0.85)	0.07 (1.41)	0.04 (1.03)	0.08 (1.61)	0.06 (1.82*)	0.05 (2.34**)
Effective Spread	-0.83 (-0.68)	-0.77 (-0.64)	-2.69 (-1.97**)	-3.25 (-2.7***)	-4.41 (-3.5***)	-2.68 (-5.3***)	-5.17 (-5.6***)	-2.86 (-3.0***)	-2.18 (-2.50**)	-3.37 (-3.5***)	0.67 (0.77)	-3.42 (-7.7***)
£Volume	0.28 (0.39)	0.38 (0.94)	1.05 (1.89*)	3.83 (2.94***)	0.70 (0.62*)	0.31 (1.77*)	-1.17 (-4.4***)	-0.03 (-0.09)	0.68 (1.96**)	2.90 (3.30***)	1.54 (2.77***)	-0.04 (-0.26)
Vol. Share	0.30 (2.85***)	0.33 (3.09***)	0.23 (2.09**)	0.25 (2.20**)	0.22 (2.19**)	0.30 (6.22***)	0.17 (2.24**)	-0.05 (-0.67)	-0.04 (-0.73)	-0.10 (-1.69*)	-0.05 (-1.20)	-0.00 (-0.12)
Adj. R^2	0.21	0.31	0.26	0.23	0.50	0.25	0.33	0.33	0.61	0.41	0.67	0.39

Table 7: Effects of informed, algorithmic and dark trading on information leadership II

The table reports panel least squares regression (with date fixed effects) coefficient estimates using a stock-day panel, in which the dependent variable is the information leadership share (ILS), for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The estimated regressions is:

$$ILS_{it} = \alpha + \beta_{PIN} PIN_{it} + \beta_{DARK} DARK_{it} + \beta_{HFT} HFT_{it} + \sum_{k=1}^{3} \varphi_k V_{kit} + \varepsilon_{it}$$

where ILS_{it} is the price discovery proxy, information leadership share for stock *i* on day *t* and is as defined in Table 4. PIN_{it} is the proxy for informed trading for stock *i* on day *t* and is as defined in Table 5. $DARK_{it}$ is the log of pound volume of dark trades for stock *i* on day *t*. HFT_{it} serves as proxy for algorithmic trading and is the messages to trades ratio for stock *i* on day *t*. V_{kit} is a set of *k* control variables which include log of pound trading volume, share of trading volume and log of effective spread. The t-statistics are presented in parentheses and derived from panel corrected standard errors (PCSE). *, ** and *** correspond to statistical significance at 0.1, 0.05 and 0.01 levels respectively. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

	Panel A: BATS Chi-X Europe						Panel B: London Stock Exchange					
	Highest	4	3	2	Lowest	Full sample	Highest	4	3	2	Lowest	Full sample
Intercept	5.38 (0.36)	41.27 (3.34***)	41.03 (3.02***)	-3.88 (-0.13)	42.21 (2.00**)	37.88 (7.94***)	39.71 (5.54***)	44.20 (4.78***)	23.83 (2.96***)	-18.82 (-0.72)	-27.45 (-1.8*)	32.27 (9.87***)
PIN	1.44 (10.1***)	3.22 (13.7***)	1.63 (14.2***)	1.65 (11.3***)	5.10 (24.2***)	2.06 (27.9***)	3.70 (15.7***)	4.20 (14.1***)	9.62 (29.5***)	4.92 (17.2***)	13.40 (35.1***)	5.25 (34.0***)
Dark	0.34 (0.59)	-0.98 (-1.62)	-1.28 (-1.90*)	-0.43 (-0.56)	-0.15 (-0.24)	-0.75 (-3.4***)	–	–	–	–	–	–
HFT	0.07 (1.49)	-0.14 (-4.4***)	-0.01 (-0.37)	-0.02 (-0.32)	-0.03 (-0.80)	-0.05 (-3.1***)	-0.04 (-0.82)	0.08 (1.37)	0.06 (0.12)	0.01 (0.13)	0.08 (2.45**)	0.04 (1.86*)
Effective Spread	-0.31 (-0.19)	-1.76 (-1.21)	-3.02 (-1.97**)	-3.10 (-2.29**)	-4.34 (-3.2***)	-3.30 (-6.3***)	-5.44 (-5.3***)	-3.43 (-3.2***)	-3.29 (-3.4***)	-3.98 (-3.8***)	1.03 (1.10)	-3.70 (-8.2***)
£Volume	0.29 (0.73)	0.18 (0.46)	0.61 (1.01)	1.70 (0.98)	-0.74 (-0.56)	0.14 (0.81)	-1.16 (-4.0***)	-0.25 (-0.65)	0.10 (0.26)	2.43 (2.09**)	1.72 (2.64***)	-0.18 (-1.27)
Vol. Share	0.45 (3.60***)	0.30 (2.17**)	0.32 (2.36**)	0.42 (3.11***)	0.24 (2.18**)	0.32 (6.45***)	0.23 (2.52**)	-0.18 (-2.07**)	-0.06 (-1.01)	-0.07 (-0.94)	-0.06 (-1.37)	-0.02 (-0.68)
Adj. R^2	0.22	0.32	0.26	0.25	0.51	0.26	0.34	0.34	0.61	0.42	0.68	0.39

Table 8: Effects of informed, algorithmic and dark trading on information leadership III

The table reports panel least squares regression (with stock fixed effects) coefficient estimates using a stock-day panel, in which the dependent variable is the information leadership share (ILS), for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The estimated regressions is:

$$ILS_{it} = \alpha + \beta_{PIN} PIN_{it} + \beta_{DARK} DARK_{it} + \beta_{HFT} HFT_{it} + \sum_{k=1}^{3} \varphi_k V_{kit} + \varepsilon_{it}$$

where ILS_{it} is the price discovery proxy, information leadership share for stock i on day t and is as defined in Table 4. PIN_{it} is the proxy for informed trading for stock i on day t and is as defined in Table 5. $DARK_{it}$ is the log of pound volume of dark trades for stock i on day t. HFT_{it} serves as proxy for algorithmic trading and is the messages to trades ratio for stock i on day t. V_{kit} is a set of k control variables which include log of pound trading volume, share of trading volume and log of effective spread. The t-statistics are presented in parentheses and derived from panel corrected standard errors (PCSE). *, ** and *** correspond to statistical significance at 0.1, 0.05 and 0.01 levels respectively. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

	Panel A: BATS Chi-X Europe						Panel B: London Stock Exchange					
	Highest	4	3	2	Lowest	Full sample	Highest	4	3	2	Lowest	Full sample
Intercept	-7.64 (-0.52)	9.77 (0.67)	8.87 (0.69)	-51.23 (-2.24**)	-1.83 (-0.10)	9.42 (1.55)	56.36 (3.46***)	23.05 (1.40)	-7.02 (-0.62)	-30.64 (-1.43)	-11.12 (-0.77)	12.45 (2.02**)
PIN	1.46 (11.0***)	3.11 (13.9***)	1.68 (16.3***)	1.51 (11.3***)	5.05 (26.1***)	2.01 (27.8***)	3.66 (16.6***)	4.19 (15.5***)	9.68 (32.4***)	4.79 (18.3***)	13.52 (39.0***)	5.22 (34.2***)
Dark	0.42 (0.51)	-0.93 (-1.22)	-2.11 (-2.9***)	-0.35 (-0.50)	-0.40 (-0.70)	-0.33 (-1.09)	–	–	–	–	–	–
HFT	0.07 (1.54)	-0.04 (-0.96)	-0.03 (-0.65)	-0.01 (-0.23)	0.02 (0.57)	0.02 (0.86)	0.04 (0.71)	0.02 (0.24)	0.07 (1.34)	0.07 (1.16)	0.01 (0.24)	0.06 (2.06**)
Effective Spread	-4.20 (-1.24)	2.47 (1.02)	0.09 (0.04)	-0.49 (-0.20)	-2.75 (-1.33)	0.26 (0.25)	-5.59 (-3.1***)	-2.05 (-1.04)	0.39 (0.26)	-1.38 (-0.69)	-3.97 (-2.7***)	-3.23 (-4.0***)
£Volume	1.91 (1.31)	1.94 (1.40)	3.80 (2.94***)	4.11 (2.96***)	1.63 (1.29)	1.47 (2.91***)	-2.16 (-2.23**)	0.47 (0.45)	1.57 (2.40**)	2.88 (3.03***)	1.33 (2.13**)	0.96 (2.67***)
Vol. Share	0.05 (0.38)	0.15 (1.28)	0.24 (1.77*)	0.02 (0.16)	0.06 (0.55)	0.05 (0.78)	0.19 (2.33**)	-0.02 (-0.22)	-0.05 (-0.86)	-0.13 (-1.99**)	-0.01 (-0.14)	0.00 (0.01)
Adj. R^2	0.23	0.32	0.26	0.23	0.51	0.27	0.34	0.33	0.61	0.42	0.68	0.40

Table 9: Effects of informed, algorithmic and dark trading on information leadership IV

The table reports GMM regression coefficient estimates using a stock-day panel, in which the dependent variable is the information leadership share (ILS), for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The estimated regressions is:

$$ILS_{it} = \alpha + \beta_{PIN} PIN_{it} + \beta_{DARK} DARK_{it} + \beta_{HFT} HFT_{it} + \sum_{k=1}^{3} \varphi_k V_{kit} + \varepsilon_{it}$$

where ILS_{it} is the price discovery proxy, information leadership share for stock i on day t and is as defined in Table 4. PIN_{it} is the proxy for informed trading for stock i on day t and is as defined in Table 5. $DARK_{it}$ is the log of pound volume of dark trades for stock i on day t. HFT_{it} serves as proxy for algorithmic trading and is the messages to trades ratio for stock i on day t. V_{kit} is a set of k control variables which include log of pound trading volume, share of trading volume and log of effective spread. Appropriate instrumental variables (IVs) are obtained for $DARK_{it}$ and PIN_{it} by first collecting the within-quintile/full sample cross-sectional averages of the trading variables. $DARK_{it}$ and PIN_{it} are then each individually regressed on their corresponding cross-sectional stock averages and the other control variables in panel least squares frameworks. The residuals from these regression are each employed as IVs in the GMM estimation. The t-statistics are presented in parentheses and derived from panel corrected standard errors (PCSE). *, ** and *** correspond to statistical significance at 0.1, 0.05 and 0.01 levels respectively. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

	Panel A: BATS Chi-X Europe						Panel B: London Stock Exchange					
Pound volume Quintile	Highest	4	3	2	Lowest	Full sample	Highest	4	3	2	Lowest	Full sample
Intercept	6.18 (0.53)	36.42 (3.51***)	33.10 (3.02***)	-50.53 (-2.46**)	36.27 (4.39***)	32.59 (7.51***)	42.96 (6.75***)	32.28 (3.92***)	11.03 (1.63)	30.31 (4.67***)	0.37 (0.10)	28.52 (9.25***)
PIN	1.45 (10.8***)	3.23 (18.2***)	1.63 (16.0***)	1.63 (13.8***)	5.10 (28.5***)	2.06 (37.0***)	3.70 (16.7***)	4.22 (19.5***)	9.59 (33.9***)	4.95 (23.1***)	13.33 (43.1***)	5.25 (51.7***)
Dark	0.50 (0.95)	-1.04 (-1.87*)	-1.23 (-1.99**)	-1.41 (-1.83*)	0.08 (0.18)	-0.68 (-3.3***)	–	–	–	–	–	–
HFT	0.07 (2.04**)	-0.13 (-4.4***)	-0.02 (-0.65)	-0.02 (-0.43)	-0.02 (-0.55)	-0.03 (-2.25**)	-0.03 (-0.84)	0.07 (1.42)	0.04 (1.09)	0.03 (0.71)	0.02 (0.65)	0.05 (2.40**)
Effective Spread	-0.86 (-0.70)	-0.94 (-0.81)	-2.59 (-1.94**)	-2.94 (-2.30**)	-5.37 (-3.9***)	-2.79 (-5.8***)	-5.17 (-5.6***)	-2.87 (-2.58**)	-2.20 (-2.52**)	-3.69 (-3.4***)	0.94 (1.07)	-3.43 (-8.5***)
£Volume	0.30 (0.86)	0.44 (1.32)	1.18 (2.07**)	5.00 (3.71***)	-0.80 (-1.82*)	0.36 (2.14**)	-1.17 (-4.4***)	-0.03 (-0.08)	0.69 (1.97**)	0.14 (0.42)	0.71 (2.56**)	-0.04 (-0.26)
Vol. Share	0.30 (2.87***)	0.33 (3.23***)	0.24 (2.13**)	0.25 (2.32**)	0.21 (2.30**)	0.30 (6.87***)	0.17 (2.24**)	-0.05 (-0.67)	-0.04 (-0.74)	-0.05 (-0.87)	-0.05 (-1.30)	-0.00 (-0.11)
Adj. R^2	0.21	0.31	0.26	0.22	0.50	0.25	0.33	0.32	0.61	0.41	0.67	0.39

Table 10: Effects of share of price discovery on market share

The table reports regression coefficient estimates using a stock-day panel, in which the dependent variable is the pound volume Chi-X market share for 47 FTSE 100 stocks trading simultaneously on the London Stock Exchange's Stock Exchange Electronic Trading System (SETS) and the BATS Chi-X's CXE order books. The estimated regressions is:

$$Marketshare_{it} = \alpha + \beta_{ILS} ILS_{it} + \beta_{Tradesize} Tradesize_{it} + \beta_{Trades} Trades_{it} + \beta_{Volatility} Volatility_{it} + \varepsilon_{it}$$

where $Marketshare_{it}$ corresponds to the log of share of pound volume of stock *i* traded on day *t* on Chi-X, ILS_{it} is the information leadership share of Chi-X with respect to stock *i* on day *t* and is as defined in Table 4. $Tradesize_{it}$ is the log of average daily trade size in pounds traded on Chi-X for stock *i* on day *t*, $Trades_{it}$ is the log of number of transactions executed on Chi-X with respect to stock *i* on day *t*, while $Volatility_{it}$ is the log of standard deviation of the intraday mid-point price return for stock *i* on day *t*. The t-statistics are presented in parentheses and derived from panel corrected standard errors (PCSE). *, ** and *** correspond to statistical significance at 0.1, 0.05 and 0.01 levels respectively. Quintiles are computed on the basis of daily pound volume across the sample period 1st July 2014 and 28th November 2014.

Intercept	1.064 (2.42**)	1.122 (2.50**)	-1.863 (-3.66***)	1.055 (2.41**)
Information leadership	0.002 (6.94***)	0.002 (7.63***)	0.000 (0.29)	0.002 (7.80***)
Trade size	0.084 (5.43***)	0.082 (5.18***)	0.202 (11.26***)	0.084 (5.44***)
Transactions	0.123 (9.14***)	0.123 (9.02***)	0.207 (11.75***)	0.123 (9.11***)
Volatility	0.008 (0.71)	0.012 (1.02)	-0.05 (-3.97***)	0.008 (0.69)
Adj. R^2	0.24	0.28	0.56	0.24
Estimation Method	OLS	OLS	OLS	GMM
Fixed Effects	None	Date	Stock	None
Instrumental variables	NA	NA	NA	Yes