Order aggressiveness and flash crashes

KHALADDIN RZAYEV
University of Edinburgh, United Kingdom

GBENGA IBIKUNLE*
University of Edinburgh, United Kingdom
European Capital Markets Cooperative Research Centre, Pescara, Italy

Abstract We present a novel framework illustrating the links between order aggressiveness and flash crashes. Our framework involves a trading sequence beginning with significant increases in aggressive sell orders relative to aggressive buy orders until instruments' prices fall to their lowest levels. Thereafter, a rise in aggressive buy orders propels prices back to their pre-crash levels. Using a sample of S&P 500 stocks trading during the May 6 2010 flash crash, we show that our framework is correctly specified and provide a basis for linking flash crashes to aggressive strategies, which are found to be more profitable during flash crashes.

JEL Classification: G14; G15; G18

Keywords: order aggressiveness, algorithmic trading, asymmetric information, extreme price movement, high-frequency data, logistic regression.

---

* Corresponding author. Contact information: University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh EH8 9JS, United Kingdom; e-mail: Gbenga.Ibikunle@ed.ac.uk; phone: +441316515186.
We thank Jake Ansell, Thierry Foucault, Albert Menkveld, Vito Mollica, Talis Putnins, Tom Steffen and colleagues at the University of Edinburgh for helpful comments. We are also grateful to the seminar participants at the June 21 2017 Market Microstructure Workshop at Università della Svizzera italiana in Lugano, Switzerland and the June 1 2017 Finance Seminar at the University of Edinburgh for constructive comments.

1. Introduction

Flash crashes are characterised by high price volatility, a significant negative return in instruments' prices and are defined by a sharp price reversal (see Aldridge 2010, Easley et al. 2011). The most notable flash crash in recent history occurred on May 6, 2010 (see Kirilenko et al. 2017). On this day, market indices such as the S&P 500, the Dow Jones Industrial Average, the Russell 2000, and the Nasdaq 100, fell significantly before rebounding within an extremely short period of time.

In the aftermath of the May 6 flash crash there has been a widespread concern that trading strategies commonly deployed by the fastest traders in financial markets – the so-called high frequency traders (HFTs) – induce or worsen price crashes.[2] Kirilenko et al. (2017) argue that, although there may be no evidence of HFTs causing the May 6 flash crash, they nevertheless exacerbated it by demanding immediacy. The immediacy demanded at a heightened pace in a liquidity-constrained environment appeared to have led to an unbearably high level of order flow toxicity, thereby worsening the price crash.[3] The aggressiveness of HFTs in demanding liquidity could therefore be argued to be a major contributing factor to the extent of the price crash recorded on May 6, 2010. However, to date, there has been no study directly linking order aggressiveness[4] to flash crashes, with no constraints placed on market agents. This paper addresses this gap in the literature.

This paper differs from existing studies (see as examples, Easley et al., 2011; Jacob Leal et al., 2016; Kirilenko et al., 2017) in that the links we draw between order aggressiveness

---

[2] About five months after the flash crash, on September 30 2010, the Commodity Futures Trading Commission (CFTC) and the Securities and Exchange Commission (SEC) released a study identifying an automated program executing the sale of 75,000 E-mini S&P 500 futures contracts as the main trigger for the flash crash (see SEC 2010).

[3] Easley et al. (2011) highlight the key role played by order flow toxicity in the occurrence of the flash crash; they also propose a measure of order flow toxicity, which they call the Volume-Synchronized Probability of Informed Trading (VPIN).

[4] We define aggressive orders in line with the classification approach of Biais et al. (1995); specifically, aggressive orders are defined with respect to their sizes and tendency to cross the spread.

and flash crashes make no assumptions regarding liquidity constraints in the market.[5] Specifically, in order to investigate the empirical links between order aggressiveness and flash crashes, we first extend the approach of Menkveld (2013), developed to decompose the trading profit in a normal market environment into its spread and positioning components. Menkveld (2013) illustrates the decomposition of traders' profits by presenting two extreme cases – aggressive and passive market making trading strategies. The framework shows that traders adopting aggressive trading strategies incur losses during normal trading days and, therefore, the majority of traders – about 80% – tend to deploy passive market making trading strategies. The losses reported for aggressive traders on normal trading days is due to incoming market orders adversely selecting aggressive orders in the market (see also Glosten and Milgrom 1985). We extend Menkveld's (2013) two-stage approach in order to show how aggressive trading strategies affect the price discovery process in financial markets.

Our framework involves a trading sequence beginning with significant increases in aggressive sell orders relative to aggressive buy orders until instruments' prices fall to their lowest levels. Thereafter, a rise in aggressive buy orders propels prices back to their pre-crash levels. Using the predictions of the framework, we highlight the role of order aggressiveness in extreme price movements, such as flash crashes, and argue that order aggressiveness can lead to flash crashes.[6] This also implies that flash crashes can be predicted by analysing the evolution of order aggressiveness in financial markets. In this case, our framework shows that even in a liquid trading environment where there are no significant liquidity constraints, order

[5] Jacob Leal et al. (2016) also develop an agent-based model of a limit-order book to show the impact of HFT on financial markets; their HFTs are assumed to deploy only predatory high frequency trading strategies (aggressive trading strategies). They conclude that aggressive HFTs are culpable in flash crashes. Consistent with Jacob Leal et al. (2016), Mcinish et al. (2014) show that the aggressive behaviour of Intermarket Sweep Orders contributed to the May 6, 2010 flash crash.

[6] This argument is also motivated by the results of Griffiths et al. (2000) and Wuyts (2011), who show that aggressive orders have price impacts larger than those of other trades.

aggressiveness can create an environment of severe illiquidity such that prices become extremely volatile, as evident during the May 6, 2010 event.

Furthermore, the framework shows that profits in aggressive trading strategies are positive and large during extreme price movements such as flash crashes, and therefore the fraction and the number of aggressive orders should be higher during these periods when compared with normal trading periods. We decompose the profits of aggressive traders into their spread and positioning components and similar to Menkveld (2013), we show that traders are confronted with a position profit and, inevitably, a spread loss when they trade aggressively. However, unlike during normal trading periods, when markets are volatile, the position profit eclipses the spread loss, thus making aggressive trading ultimately profitable during periods of high price volatility. Since our framework, involving a three-stage aggressive trading strategy, which results in a price collapse and a subsequent sharp price reversal, mimics the form of a flash crash, we argue that aggressive trading strategies can cause flash crashes. This implies that we can obtain advance information about the likely onset of flash crashes by computing the relative weights of aggressive trading/orders. We test the foregoing arguments and framework predictions using ultra-high frequency trading data for the components of the S&P 500 stock index affected by the May 6 flash crash. The empirical results obtained are completely in line with the predictions of our framework.

Firstly, we find that a significant imbalance in order aggressiveness favouring sell orders ensues in the run-up to and during the flash crash. We document a significant increase in the number of aggressive sell orders relative to aggressive buy orders in the run-up to and during the flash crash until instruments' prices plummeted to their troughs. The increase in aggressive sell orders with no corresponding rise in aggressive buy orders precipitated the crash in instruments' prices. This finding is very important, since the total number of aggressive orders could be high; however, a significant price crash will only occur if aggressive sell orders significantly outstrip aggressive buy orders. This result is consistent with the official reporting following the flash crash.

Secondly, we link the evolution of order aggressiveness to the flash crash within an econometric framework, showing that increased order aggressiveness is related to the May 6 2010 flash crash; hence, order aggressiveness could provide a signal about the onset of the future flash crashes.

Thirdly, we show that aggressive trading is significantly more profitable during periods of high price volatility such as flash crashes, than during normal trading periods. We find that an informed trader could earn up a cumulative return in excess of 1,482 basis points (bps) based on our analysis of a sample of flash crash-affected stocks, this is significantly higher than possible during the non-flash crash periods. Consistent with this finding, the fraction of aggressive buy and sell orders during the May 6, 2010 flash crash is higher than the fraction of these kinds of orders during other periods under investigation. The actual number of aggressive sell and buy limit orders during the flash crash is also remarkably higher than during the surrounding periods (before and after the flash crash). Our results are robust to alternative estimation approaches and model specifications, including estimation frequencies. Overall, the empirical results show that our framework is correctly specified and the arguments we present valid in the case of the flash crash we examine.

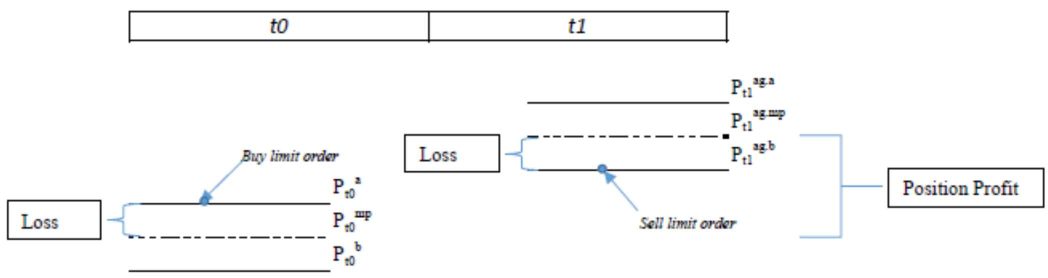
The remainder of this paper is structured as follows. Sections 2 and 3 present our framework/approach and data, respectively. Section 4 presents and discusses the empirical analysis and results, while Section 5 concludes.

2. The approach

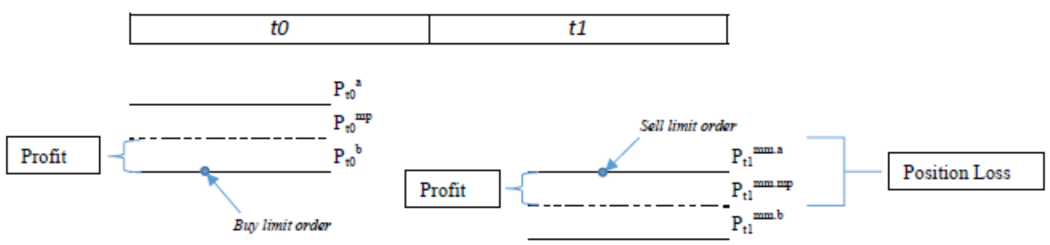
2.1 Motivation

Griffiths et al. (2000) and Wuyts (2011) find that aggressive orders generate larger price impacts. Given this finding, there is a case to be made for aggressive orders being culpable in inducing extreme price movements, such as flash crashes. However, this argument raises an interesting question about why aggressive orders do not always cause flash crashes, given that they are likely to be submitted repeatedly on any given day in financial markets. In order to examine this question and demonstrate the potential relationship between order aggressiveness and flash crashes, we extend the approach of Menkveld (2013). Following Sofianos (1995), Menkveld (2013) decomposes the profit of traders into two components: the spread component and the positioning component. Menkveld's (2013) framework focuses on two extreme cases involving aggressive trading on the one hand and passive market making on the other, by using a two-stage approach:

Aggressive trading strategy



Passive market making strategy



where P_{t0}^{a} is the ask price at time t_0, P_{t0}^{b} is the bid price at time t_0, P_{t0}^{mp} is the mid-price at time t_0, $P_{t1}^{ag.a}$, $P_{t1}^{ag.b}$ and $P_{t1}^{ag.mp}$ are the ask price, the bid-price and the mid-price at time t_1 under aggressive trading strategy, respectively, and $P_{t1}^{mm.a}$, $P_{t1}^{mm.b}$ and $P_{t1}^{mm.mp}$ are the ask price, the bid-price and the mid-price at time t_1 under passive (market-making) trading strategy, respectively.

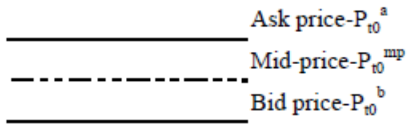
In the first extreme case, i.e. aggressive trading strategy, a trader consumes liquidity in order to pursue a fundamental value change, and then quickly follows this with a sell order. By submitting a buy limit order at the ask price and a sell limit order at the bid price, the trader will make a spread loss at t_0 and t_1, but will make a position profit at the end of the trading session. The trader will adopt this trading strategy if she expects a large position profit at the end of the trading session – this is necessary to compensate for the spread losses incurred from the first and second trading stages. However, adverse selection is a potential risk here, as the position profit could be negative if the trader's orders are adversely selected by an informed market order (see Glosten and Milgrom 1985). Consistent with this argument, Menkveld (2013) finds that position profit is negative in the Dutch stock market during normal trading periods – periods of no or very low price volatility. In the second extreme case, i.e. the passive market making strategy, a trader acting as a market maker makes a profit from the spread in the first and second trading session, and a loss from her position at the end of trading.

In this paper, we alter the strategies above and further extend the framework to decompose the profit of traders. Specifically, we employ a three-stage approach and alter the order of submitted orders to show the relationship between order aggressiveness and flash crashes; what this means is that while Menkveld's (2013) framework begins with a buy order, our approach begins with a sell limit order. Furthermore, we add the relative weights concept to this approach in to obtain the predictability of flash crashes.

2.2 Our three-stage approach

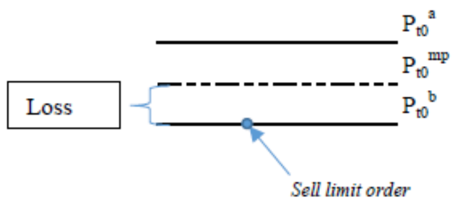
Trading at t_0

Traders submit sell limit orders at t_0 by following one of two trading strategies (passive and aggressive), while the subsisting bid and ask prices, with mid-price P_{t0}^{mp}, are set before traders come to the market:



$$P_{t0}^{mp} = \frac{P_{t0}^{a} + P_{t0}^{b}}{2} \quad (1)$$

We assume that a trader will submit a sell limit order at the prevailing best bid price if she wants to adopt an aggressive trading strategy, or a trader will submit a sell limit order at ask price is she wants to adopt a passive market making strategy. We focus on one of these extreme cases, an aggressive trading strategy, as we aim to illustrate the relationship between order aggressiveness and flash crashes. By submitting a sell limit order at the bid price, a trader will make a loss at t_0. The trading sequence is illustrated below:



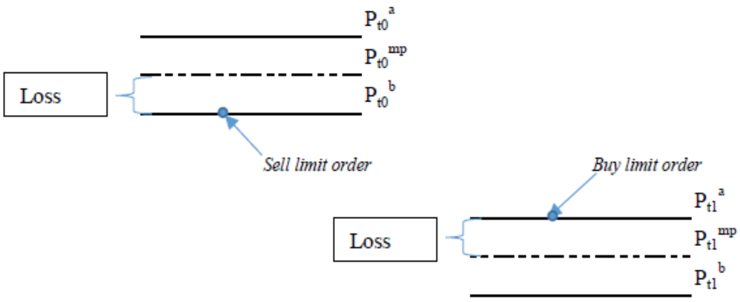
The loss of our hypothetical aggressive trader is therefore given as:

$$\pi_{t0}^{ag} = P_{t0}^{b} - P_{t0}^{mp} \quad (2)$$

Trading at t_1

Inevitably, different types of trading strategies in t_0 will have different impacts on ask and bid prices. This implies that bid and ask prices at t_1 will be different under either of the two extreme (passive and aggressive) strategies/cases. By submitting an aggressive sell limit order at t_0, the trader consumes liquidity, which in turn induces a price change. An aggressive trading strategy will therefore have a downward pulling effect on bid and ask prices, leading to bid and ask prices going down at t_1. However, if an aggressive order is adversely selected by an incoming informed market order, the price will go up at t_1 and the aggressive trader will incur a significant position loss. We therefore concentrate on the case where an aggressive order is not adversely selected. This is important for our framework to mimic the price evolution during a flash crash, i.e. price falls significantly from t_0 to t_1.

During the second trading stage, the aggressive trader submits an aggressive buy limit order at the ask price:



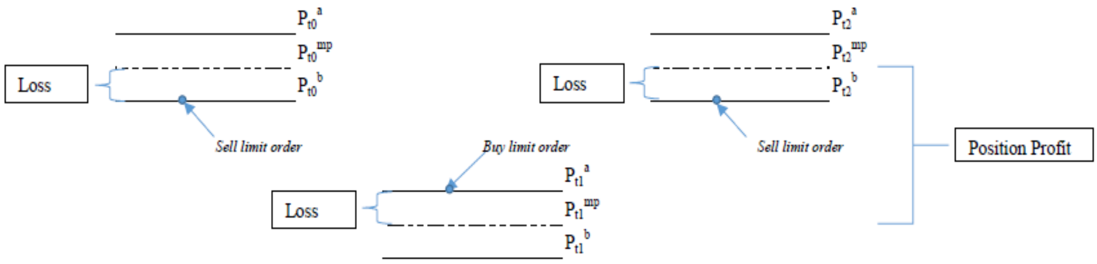
The submission of a buy order at the ask price will again lead to the trader incurring losses at t_1. The payout at this stage will be:

$$\pi_{t1}^{ag} = P_{t1}^{mp} - P_{t1}^{a} \tag{3}$$

Trading at t_2

As earlier stated, the deployed trading strategies will have varying impacts on ask and bid prices. An aggressive trading strategy at t_1 will generate an increasing pressure on bid and ask prices, thus bid and ask prices will appreciate subsequently at t_2 and reach initial position (t_0). If an aggressive order is adversely selected by an incoming informed market order, the price will go down at t_2 and our hypothetical trader will again incur significant position loss. Therefore, we again assume that an aggressive order is not adversely selected, to mimic the price evolution during a flash crash, i.e. price rebounds from t_1 to t_2 and attains the pre-flash crash level.

We further assume that the asset price at time t_2 will be equal to the asset price at time t_0. This is necessary for the sequence of events/price evolution to be consistent with a flash crash; i.e. a sudden/sharp fall in the price of an asset and a full rebound in price shortly afterwards:



By submitting a sell limit order at t_2's bid price, the aggressive trader makes a profit from her position and incurs losses from the bid-ask spread. Thus, her position profit and spread loss are as follows:

$$\text{Position Profit} \qquad \pi_{t2}^{ag.p} = P_{t2}^{mp} - P_{t1}^{mp} \tag{4}$$

$$\text{Spread Loss} \qquad \pi_{t2}^{ag.ba} = P_{t2}^{b} - P_{t2}^{mp} \tag{5}$$

$$\text{Total Profit} \qquad \pi_{t2}^{ag} = P_{t2}^{b} - P_{t1}^{mp} \tag{6}$$

To sum up these trading strategies thus far, we can examine the profitability of an aggressive trading strategy. By combining the above equations, we generate the following equations for an aggressive trading strategy, assuming that the bid and ask prices at time t_2 equal the bid and ask prices at time t_0:

$$\begin{aligned} \pi^{ag} &= (P_{t0}^{b} - P_{t0}^{mp}) + (P_{t1}^{mp} - P_{t1}^{a}) + (P_{t0}^{b} - P_{t1}^{mp}) = \\ &= (P_{t0}^{b} - P_{t0}^{mp}) + (P_{t0}^{b} - P_{t1}^{a}) = \\ &= 2P_{t0}^{b} - P_{t0}^{mp} - P_{t1}^{a} \end{aligned} \tag{7}$$

Typically, a trader should pay the clearing fee and the aggressive exchange fee (usually imposed by exchanges on traders consuming liquidity) when she adopts an aggressive trading strategy. For simplicity, we assume that these fees are zero. As seen from Equation 7, the position profit of an aggressive trading strategy is high if there is a sharp reduction in the asset price at $t_1 (P^a_{t1})$. The interesting point is that this type of sharp reduction is consistent with the extreme price movements documented in the case of flash crashes. Therefore, we argue that although Menkveld (2013) shows that position profit is negative during normal trading days, it might be large and positive during extreme price movements. This implies that this kind of extreme price movement could be profitable for some traders. This argument is consistent with Brogaard et al. (2014), who show that although HFTs do not cause extreme price movements such as flash crashes, these types of price movements is more profitable for HFTs. The argument raises an interesting question about why traders fail to always adopt an aggressive trading strategy and therefore obtain large and positive profitable positions or, more specifically, are there some other conditions that ensure that traders become aggressive? We argue that there should be other conditions, which are not necessarily directly linked with the traders themselves, which may lead to traders choosing an aggressive trading strategy. The important point to note is that the price decrease in t_1 should be very sharp in order to compensate for the losses from the spread. As already stated, that there will be a position loss if the order submitted by an aggressive trader is adversely selected by an incoming informed market order (see also Glosten and Milgrom 1985). Therefore, traders must be sure that they do not face adverse selection risk when attempting an aggressive trading strategy. Indeed, this argument explains Brogaard et al.'s (2014) view regarding the profitability of extreme price movements for HFTs. The ability of HFTs to make hay of volatile trading conditions as described above is not far-fetched. Hirschey (2017) argues that HFTs can anticipate buying and selling pressure, which could help them avoid being adversely selected when deploying aggressive trading strategies. Indeed, Hirschey (2017) finds that HFTs' aggressive sales and purchases consistently lead those of other investors. This implies that the framework we illustrate above is more likely to be successfully deployed when it is implemented at a high frequency.

2.3 Order aggressiveness and flash crashes

Thus far, we have demonstrated the price evolution under an aggressive trading strategy. It is important to note that the framework discussed above simply illustrates the pricing process under the assumption of zero adverse selection risk. It implies that the price discovery process is not affected by the behaviour of other traders in this approach. The sequence of aggressive trading strategy we describe is therefore mainly useful for understanding the contribution of order aggressiveness to flash crashes. Although, the sequence of orders is not based on the May 6 2010 flash crash, the aggressive trading strategy shares two notable characteristics with the May 6, 2010 flash crash. Firstly, the price movement under this strategy exactly mimics the price movements in the US financial markets during the flash crash, i.e. asset prices collapse and rebound very rapidly within a very short period of time. Secondly, the SEC (2010) finds that a large amount of seller-initiated E-mini contracts executed by algorithmic traders triggered the flash crash. Our approach also begins with a sell limit order. Inspired by these two commonalities, we argue that an aggressive trading strategy can cause flash crashes under certain conditions, mainly when aggressive traders do not face adverse selection risk. In order to further test our argument, we use real data and examine the aggressiveness of order flows during, and prior to, the May 6, 2010 flash crash. If, indeed, the predictions of our framework are consistent with the flash crash, then, firstly, there should be an excessive sell order aggressiveness in financial markets, which will create a downward pulling effect on prices. Thereafter, the excessive aggressiveness should shift to the buy side and as a result, prices will rise. This sequence of events implies that order aggressiveness should provide a signal about the onset of the future flash crashes. Secondly, the fraction and number of aggressive buy and sell orders during the May 6, 2010 flash crash should be higher than the fraction of aggressive buy and sell orders during the surrounding periods. This is simply because, as we have shown, aggressive orders are more profitable during these periods. It is very important to note that we do not argue that the three-stage aggressive trading strategy we illustrate in this paper is the reason of the May 6 flash crash. Rather, we argue that order aggressiveness prior and during the May 6, 2010 flash crash contributed to the crash (see SEC 2010).

2.4 Relative weight of an aggressive trading strategy

In each period, traders have different alternative trading strategies, and they can switch according to the profitability of each strategy. Inspired by Brock and Hommes (1998), we use the strategies' attractiveness function and assume that the choice among the alternative strategies available to traders depends on this function:

$$U_t^{st} = \pi_t^{st} + \eta U_{t-1}^{st} \tag{8}$$

where η represents the 'memory parameter', and st shows the type of strategy, i.e. extreme passive market making, extreme aggressive trading strategy, and others. For the extreme case $\eta = 0$, traders/HFTs have no memory. It is important to note that the attractiveness function of being inactive equals zero. Following one of the well-known discrete choice models, logit model, as developed by Manski and McFadden (1981), the relative weights of an extreme aggressive trading strategy can be determined as:

$$\varphi_t^{ag} = \left(\frac{e^{\xi \times U_t^{ag}}}{\sum e^{\xi \times U_t^{st}}}\right) \tag{9}$$

(see Appendix B for proof) where ξ is an intensity of switching parameter. The higher the value of ξ, the higher the fraction of traders who prefer the strategy with the highest fitness function. Brock and Hommes (1998) interpret ξ as a (bounded) rationality parameter. Westerhoff (2008) and Pellizzari and Westerhoff (2009) also use the same concept to compute the relative weights of trading strategies. Our proposal is to interpret this parameter as an inverse measure of adverse selection risk. It implies that if adverse selection risk is low, then this parameter will be high and traders will have more propensity to adopt an aggressive trading strategy.

As previously stated, extreme aggressive trading strategies can induce flash crashes under certain conditions, i.e. when there is no adverse selection risk. The implication here is that we can obtain advance information about flash crashes by computing the relative weight of an extreme aggressive trading strategy. Thus, we hypothesize that the probability of flash crashes in the next period will increase if the relative weight of an aggressive trading strategy in the next period (φ_t^{ag}) increases. It implies that this approach can be used to directly predict flash crashes.

3. Data

3.1 Sample selection

In order to empirically test our hypotheses, as developed above, we focus on the biggest and most reported flash crash in the recent financial markets history, the May 6, 2010 flash crash experienced in the U.S. markets. The flash crash was one of the most turbulent periods in U.S. financial markets history and has been considered to be the most harmful flash crash to date, during which the biggest intraday point decline in the history of the Dow Jones Industrial

Average was recorded. Not only major indices, but also options, exchange-traded funds, and individual stocks, suffered from the May 6, 2010 flash crash.[7]

The data employed consists of ultra-high frequency tick-by-tick data for a selection of 53 S&P 500 stocks sourced from the Thomson Reuters Tick History (TRTH) database. Appendix C contains a detailed list of all stocks that are analysed. We obtain data for all messages recorded for May 6 2010, but focus mainly on the period between 1:30 PM and 4 PM, since the flash crash started around 2:32 PM and lasted for about 36 minutes (see SEC 2010). In the data, each traded price and volume is recorded with a time stamp to the nearest 1/1000th of a second (millisecond). The following variables are included in the dataset: Reuters Identification Code (RIC), date, timestamp, price, volume, bid price, ask price, bid volume, and ask volume.

Although the S&P 500 index consists of 500 large companies listed on the NYSE and NASDAQ, we select only the 53 stocks deemed to have been severely affected by the flash crash. Furthermore, we select S&P 500 stocks for analysis purposes because SEC (2010) also examines the impact of the flash crash on individual stocks by using selected data from this index. SEC (2010) shows that a large trader executing a sell program for 75,000 E-mini S&P 500 index futures contracts triggered the flash crash of May 6 2010. As the performance of this index future is directly linked with the S&P 500 stocks, it is reasonable to select the components of S&P 500 for our analysis.

Once the raw data is obtained, we determine the prevailing best bid and best ask quotes for each transaction by using the order flow as downloaded. We then follow Chordia et al. (2001) and Ibikunle (2015) in applying a standard set of exclusion criteria to the data, thus deleting all inexplicable observations which might arise due to errors in data entry.

[7] According to SEC (2010), the May 6, 2010 flash crash lasted for approximately 36 minutes and could be viewed as consisting of two halves: (1) prices collapse and reach their lowest levels from 2:32 PM to 2:45 PM, (2) prices rebound and reach their pre-crash levels from 2:46 PM to 3:08 PM.

3.2. Sample Description

In order to better observe the dynamics of stocks during the flash crash, we classify the sample into three periods: before the flash crash (from 1:30 PM to 2:32 PM), the flash crash period (from 2:32:01 PM to 3:08 PM), and after the flash crash (from 3.08:01 PM to 4:00 PM).

INSERT TABLES 1 AND 2 ABOUT HERE

Panel A of Table 1 presents the summary statistics of trading activities of the selected stocks. We observe a marked increase in average per minute transactions and trading volume during the flash crash, followed by a fall after the flash crash. This volatility is consistent with the modelled effects of the flash crash as presented in our framework. Prior to the flash crash, the average per minute trading volume is about 1 million. This increases by 161% during the flash crash and afterward falls by approximately 14%. Furthermore, the average per minute number of transactions and dollar trading volume during the flash crash are about three times higher than before the flash crash. After attaining the highest levels, average transaction and dollar trading volumes per minute fall by about 13%. We compute statistical tests to show the differences in trading volume and dollar trading volume between the period of the flash crash and surrounding periods. In Table 1's Panel B, we present the p-values of different statistical approaches, testing for the null that there is no difference between the trading activity during the flash crash and non-flash crash periods. For robustness, we construct two-sample t-tests and pairwise Wilcoxon-Mann-Whitney U tests. Both methods show that the difference between these two periods is statistically significant. Given that, in the market microstructure literature, changes in trade sizes are thought to reflect the changing composition of the traders/participants in a market, one may assume that the fraction of traders that submit aggressive orders increases during flash crash.

Table 2 presents the order submission summary statistics for our sample of stocks. Although average per minute trading volume increases sharply during the flash crash, the average volume of shares submitted in bid and ask orders over the same frequency decline during the flash crash. Firstly, this is consistent with what we would expect in t_1, following liquidity consumption in t_0. Secondly, when the ratio of shares in orders to trading volumes is calculated, we find that the ratio is 5.3 before the flash crash, indicating that approximately one in five submitted shares in the orders submitted is executed prior to the flash crash. The ratio quickly falls to 1.8 during the flash crash and increases 2.1 afterwards. Thus, the rate of order execution quickens during the flash crash as the search for liquidity intensifies. The estimate of 1.8 share in order to trade ratio shows that more than half of shares in orders submitted during the flash crash are executed. This result further supports our argument that traders become more aggressive during the flash crash or, at the very least, the proportion of aggressive traders in the market increases during the flash crash.

4. Empirical analyses, results and discussions

Our aim in this section is to formally test hypotheses arising from our three central framework arguments. The first argument suggests that excessive aggressiveness in trading is culpable in the inducement of flash crashes; this implies a significantly increased volume of aggressive sell and buy orders in the period leading up to and during the flash crash. More specifically, our framework predicts that, firstly, there should be an excessive sell aggressiveness in the first half of the flash crash and this aggressiveness will create a downward pressure on prices. Then, the buy side should subsequently become more aggressive, which will inevitably create an upward pressure on prices. Secondly, our framework predicts that order aggressiveness should provide a signal about the likely onset of the flash crashes; hence, the second hypothesis we examine is that aggressive orders are statistically significant predictors of flash crashes. Thirdly, the framework suggests that aggressive orders might be more profitable during extreme price movements such as flash crashes. The implication here is that the fraction and number of aggressive orders in the lead up to and during flash crashes should be higher than the fraction and number of aggressive orders during other trading periods surrounding flash crashes.

4.1. The evolution of order aggressiveness

In order to proceed with the test of the arguments/hypotheses above, we need to identify an appropriate indicator or proxy for aggressive orders. This is required to be able to compute interval-based fractions and volume of aggressive orders in the market. For consistency with the existing literature, we employ an established approach as developed by Biais et al. (1995) to categorise limit orders according to their aggressiveness for our empirical analysis. The acceptance of this classification scheme in the market microstructure literature is underscored by its relatively wide use (see as examples Degryse et al. 2005, Griffiths et al. 2000, Hagströmer et al. 2014). The Biais et al. (1995) order classification algorithm involves dividing buy and sell orders into six groups by their level of aggressiveness; Category 1 orders are the most aggressive orders, while Category 6 orders are the least aggressive. A Category 1 buy order has a bid price higher than the best ask price and a quantity larger than the quantity available at the best ask price at its time of submission. These kinds of buy orders would normally walk across the order book. A Category 2 buy order has a bid price equal to the best ask price but has a target quantity exceeding the prevailing depth at the best ask price. Category 3 buy orders also have bid prices equal to the best ask prices, however their target quantities do not exceed the prevailing depth at the best ask price. The bid price of Category 4 buy orders is higher than the best bid price but less than the best ask price. The quantity of this order is not necessary for categorisation purposes. Categories 5 and 6 buy orders are the least aggressive. Like the

Category 4 buy order, there are no quantity requirements for categorising Category 5 buy orders, however the bid prices of these orders are equal to the best bid prices. All buy orders not otherwise categorised above are classified as Category 6 orders; specifically, the prices of these orders are less than the best bid prices. Based on their classification, Category 4, 5 and 6 orders are not usually immediately executed, and are therefore considered passive.

The categorisation for the sell orders mirror those of the buy orders. The ask prices of the Category 1 sell orders are less than the best prevailing bid price and their sizes exceed the depths at the current best bid prices. The ask prices of the Category 2 and 3 sell orders equal to the best bid price. Furthermore, the target quantities of Category 2 orders are higher than the quantities available at the best bid prices, whereas the quantity of Category 3 sell orders are not. Consistent with the categorisation of buy orders, the prices of Category 4 sell orders lie within the best bid-ask spread, i.e. less than prevailing best ask prices. The prices of Category 5 sell orders equal the best ask price, while the remaining orders are classified as Category 6 sell orders. The prices of this latter group of sell orders are higher than the prevailing best ask prices.

Degryse et al. (2005) show that the most aggressive order types (Categories 1 and 2) execute immediately and cause a price movement. Although Category 3 orders are less aggressive than the first two classes of orders, they still usually result in prompt transactions, therefore these three types of orders (Categories 1, 2 and 3) can be considered as aggressive orders (see Degryse et al. 2005, Foucault 1999). Thus, we focus on the first three types of orders. Specifically, we compute the sum of fractions of the aggressive order categories for the May 6, 2010 flash crash, as well as for the normal periods surrounding the flash crash. We then compare the volumes within a statistical framework to determine whether the fraction of aggressive orders during the flash crash is higher than the fraction of the same types of orders during normal periods.

Figure 1 presents the evolution of order aggressiveness during the day of the flash crash. We use 1-minute time intervals to construct both panels of the panels in the figure. In Panel A, we employ the standard errors of the cross-sectional means to construct 99% confidence bands for the order aggressiveness estimates in Panel A, to show the upper and lower bounds of the fraction of aggressive orders during the flash crash.

INSERT FIGURE 1 ABOUT HERE

As evident in Panel A, the fraction of aggressive orders almost tripled during the flash crash from about 8% at 1:30 PM to 21.36% at 2:43 PM. The proportion of aggressive orders during the flash crash is, on average, higher than the surrounding time intervals. This finding suggests that aggressive trading activity is more prominent during the flash crash than in the surrounding periods. This result is consistent with the view that since aggressive orders might be more profitable during periods of extreme price movements, traders tend to show more aggressive behaviour during such periods. Furthermore, in Figure 1, we observe that the first of the two peaks of aggressive trading occurs just prior to the onset of the flash crash at about 2:23 PM, when the fraction of aggressive orders attains about 20.71% of the total order volume. This appears to underscore our intuition regarding the predictive power of aggressive orders with respect to flash crashes. We discuss the results of our formal test of this assertion in the next section.

Panel B makes the important distinction between buy and sell aggressive orders. Consistent with our framework's predictions, the sell side is more aggressive from 2:17 PM to 2:45 PM and then the buy side becomes more aggressive until 2:58 PM. This is not unexpected since SEC (2010) show that prices reached their lowest levels at 2:45 PM and the start to increase thereafter. This shows that the predictions of our framework are consistent with the empirical evidence and the arguments we make are valid in the case of the flash crash we examine. A clearer view of the balance between sell and buy aggressive orders is presented in Figure 2.

INSERT FIGURE 2 ABOUT HERE

Consistent with the results of Figure 1, Figure 2 shows that, as predicted by our framework, there is a significant increase of aggressive sell orders until the stocks' price attained their lowest levels during the flash crash (at 2:45 PM) and thereafter the number of aggressive sell orders are outstripped by the number of aggressive buy orders until the prices reverted back to their pre-crash levels. Furthermore, Figure 2 shows that we observe a peak in aggressive order imbalance (the difference between aggressive sell and buy orders) at 2:17 PM; this implies that as predicted by our framework, order aggressiveness could provide an indication of the onset of flash crashes.

However, it is important to note that, based on our predictions, a high fraction of aggressive orders during some specific days alone is not enough to cause an extreme price movement such as a flash crash; flash crashes are more likely induced by a large amount of aggressive orders. Therefore, we also need to examine the number of aggressive orders during the flash crash day in order to adequately investigate the prediction made in our framework.

INSERT FIGURE 3 ABOUT HERE

Figure 3 presents the evolution of the number of aggressive orders on May 6, 2010. As evident in the figure, there is a noteworthy rise in the number of aggressive orders as we approach the epicentre of the crash. The number of aggressive orders increases by about 6 times from the number at 2:00 PM (10,586/minute) to 62,760/minute at 2:43 PM, then falls precipitously to about 24,000/minute thereafter. Consistent with the data on the fraction of aggressive orders, we also observe a peak in the number of aggressive orders prior to the onset of the flash crash, at 2:22 PM (45,050/minute). This implies that the number of aggressive orders could be a signal for flash crashes. Furthermore, as evident in Figure 2, we observe an excessive level of sell order aggressiveness from 2:17 PM to 2:45 PM and an excessive buy order aggressiveness thereafter. A review of the balance between aggressive sell and buy orders is useful in clarifying the changing of order dominance between the two order types. Thus, we compute aggressive order imbalance by the numbers of orders.

INSERT FIGURE 4 ABOUT HERE

Similar to the picture painted in Figure 2, Figure 4 shows that the predictions of our framework are completely in line with the evolution of the number of buy and sell orders during a real flash crash. We observe a surge in sell order aggressiveness prior to and during the first half of the flash crash until the price levels of instruments reached their minimum levels. Thereafter, the number of aggressive buy orders start to increase relative to the number of aggressive sell orders until the prices regain their pre-crash levels. The implications of the findings presented in Figure 2 and Figure 4 are significant, since the total number of aggressive orders could be high for a number of reasons; however, a flash crash is unlikely to ensue if there are no significant differences in the fractions and numbers of aggressive buy and sell orders.

Figure 1 and Figure 3 show that, as predicted by our framework, aggressive trading activity is more prominent during the flash crash than in the surrounding periods, given that they are likely to be more profitable during periods of heightened price volatility, such as a flash crash. In order to formalise this, we compute number of aggressive orders for the flash crash and surrounding periods, and then examine the statistical significance of difference between the flash crash period (2:32 – 3:08PM) and the two surrounding intervals (1:30 – 2:32PM and 3:08 – 4PM).

INSERT TABLE 4 ABOUT HERE

As evident in Table 4, the mean of the number of aggressive orders during the flash crash is about 62% higher than the mean of the number of aggressive orders during the surrounding periods. This finding is in line with our arguments that aggressive orders could be more profitable during flash crashes and therefore, traders tend to submit more aggressive orders during these periods. We address the issue of profitability of aggressive orders during the flash crash later on in this section. Panel B in Table 4 shows that the difference between the number of aggressive orders during the flash crash and surrounding periods is statistically significant.

Thus far, the univariate empirical results presented have been generally consistent with the predictions of our framework concerning the relationship between order aggressiveness and flash crashes. Firstly, there is a significantly increased level of sell order aggressiveness prior to and during the first half of the flash crash and then, buy order aggressiveness gradually outstrips sell order aggressiveness. Secondly, the number and the fraction of aggressive orders appear to signal or induce liquidity constraints leading to the flash crash. Thirdly, the number and the fraction of aggressive orders attain their highest levels during the flash crash and is in line with our argument that these types of orders might be more profitable during extreme price movements. Although the initial results suggest that our hypothesis on the predictive power of aggressive orders for flash crashes has merit, it is imperative that these results are formally tested within a multivariate framework.

4.2. Multivariate Analysis

Next, we test whether the number of aggressive orders signal future flash crashes within a multivariate framework. Specifically, we estimate the following regression model with stock-specific variables:

$$\begin{aligned} FC_{it} &= \alpha + \beta_{NAO} NAO_{it} + \beta_{\ln V} \ln V_{it} + \beta_{VPIN} VPIN_{it} + \beta_{VLT} VLT_{it} \\ &+ \beta_{OIB} OIB_{it} + \beta_{BAS} BAS_{it} + \beta_{MF} MF_{it} + \varepsilon_{it} \end{aligned} \tag{10}$$

where FC_{it} is a binary dependent variable and time, t, equals one-second.[8] We employ two cases of the Model (10). Firstly, we use the standard logit model; in this step, our aim is testing whether the number of aggressive sell and buy orders provide a signal about future flash crashes. In the logit model, FC_{it} equals one for the pre-flash crash period (2:17 PM to 2:32 PM). This construction allows us to capture whether aggressive orders' build-up ahead of the flash crash helps predict it. Secondly, we employ the multinomial logit mode; by using this model, our main aim is to further test whether the number of aggressive sell and buy orders provide a signal about the future flash crashes. The multinomial modelling approach allows us to concurrently examine the relationship between both the pre-flash crash and flash crash periods on the one hand and contemporaneous order aggressiveness on the other. Thus, in the multinomial estimation of Model (10), FC_{it} equals one for the pre-flash period (2:17 PM – 2:32 PM), two for the flash crash period (2:32:01 PM – 3:08 PM) and zero otherwise. *NAO* is the number of aggressive orders obtained by using the order classification scheme described above. We estimate the above regression for sell (NASO) and buy (NABO) aggressive orders separately in order to capture the marginal impact of each type of order. Estimating the depth of the impact of each order type is important since according to the literature and our framework, aggressive sell orders should play a more important role in flash crashes (see SEC 2010). As already noted, the first three categories of orders are earmarked as aggressive orders. This is the most important variable in our study, and according to our arguments, we expect to see a positive relationship between the number of aggressive orders and the pre-flash crash (FC_{it}=1) period (see also Griffiths et al. 2000, Mcinish et al. 2014, Wuyts 2011).

Apart from the key variable, we employ some control variables in order to strengthen the consistency of our results. *lnV* is the natural logarithm of the number of shares traded for

[8] For robustness, we also employ five-second interval analysis and obtain qualitatively similar results. For parsimony, the results of the five-second estimation results are not presented; however, they are available on request.

one/five second interval. This proxy is used to control for the effect of trading volume. The *VPIN* metric is introduced as a real-time indicator of order flow toxicity. *VPIN* is a modified version of the Easley et al. (1996) and Easley et al. (1997) probability of an informed trade (PIN) metric and is proposed by Easley et al. (2011) as a measure of the probability of an informed trade in a high frequency environment. Easley et al. (2011) and Easley et al. (2012) highlight the role of order flow toxicity in the May 6, 2010 flash crash.[9] Easley et al. (2012), Easley et al. (2011) argue that *VPIN* can be used to predict flash crashes. By contrast, Andersen and Bondarenko (2014) show that *VPIN* is a poor predictor for flash crashes after controlling for volume. Therefore, including *VPIN* as a control variable in Model (10) offers another opportunity to examine the flash crash predictability potentials of *VPIN*. In addition to *VPIN*, *OIB* is also employed to control for the order flow toxicity. Note that multicollinearity is not an issue here, since the correlation coefficient between *VPIN* and *OIB* is very low, at 0.054 (see Table 3). SEC (2010), Kirilenko et al. (2017), and Easley et al. (2011), show that a large order imbalance was one of the contributing factors to the May 6, 2010 the flash crash, hence the inclusion of order imbalance as an explanatory variable is completely in line with the literature. *OIB* is calculated as the absolute value of the difference between the number of buy and sell trades, divided by the total number of trades (see Chordia et al. 2008). In order to obtain *OIB*, trades must first be classified into buys and sells. Generally, three types of trade classification schemes are used to classify trades; these are the tick rule, the Lee and Ready (1991) algorithm, and Easley et al. (2012), Easley et al. (2011) bulk volume classification (BVC) method. In this study, we employ the Lee and Ready (1991) algorithm for order classification.[10] Chakrabarty

[9] Computing VPIN requires determining the number of buckets to be employed for volume classification and a buy/sell trade classification method. We use 200 buckets for volume classification, because Wu et al. (2013), who examine 16,000 various parameter combinations for evaluating the effectiveness of VPIN, concludes that 200 buckets yield optimal results. Buy and sell volumes are computed using the BVC approach proposed by Easley et al. (2011).

[10] For robustness, we also compute OIB using the other two methods and employ them in Model (10), the inferences drawn from those estimations are unchanged irrespective of which OIB computation approach we use.

et al. (2015), in their comparative analysis of the aforementioned trade classification methods, conclude that the Lee and Ready (1991) algorithm method is a more accurate trade classification method than competing methods.

VLT is the one/five-second standard deviation of mid-price returns; this variable is introduced to control for trading volatility.[11] Prior contributions report extreme price volatility during the May 6, 2010 flash crash day (see as examples Easley et al. 2011, Easley et al. 2012, Kirilenko et al. 2017, SEC 2010). Furthermore, an increase in the volatility of an instrument's price will increase its market risk, leading to a larger price impact as well as extreme price movements. *BAS* is the one/five second spread between the best ask and best bid prices, and is a proxy for liquidity. *BAS* tends to be narrow when liquidity is high; hence, under liquidity constraints, i.e. when *BAS* is wide, we therefore expect a larger price impact (see Borkovec et al. 2010). *MF* corresponds to market fragmentation. Madhavan (2011) and Golub et al. (2012) show that market fragmentation is one of the factors that contribute to flash crashes, and Menkveld and Yueshen (2017) underscore and further explain the results of Madhavan (2011). In this study, the inverse of the Herfindahl-Hirschman Index is used for capturing how fragmented each stock is across various venues for each corresponding interval.[12]

INSERT TABLE 3 ABOUT HERE

Table 3 presents the correlation matrix of the explanatory variables; the low correlation coefficient estimates suggest that multicollinearity is not an issue with the regression model.

The results for both the logit and multinomial logit models' estimations are presented in Table 5 and Table 6 respectively.

INSERT TABLE 5 ABOUT HERE

---

[11] We employ mid-price returns in order to reduce bid-ask bounce (see Avramov et al. 2006).

[12] The index is defined as: $HHI_t = \sum_{k=1}^{K} \left(s_t^k\right)$, where s_t^k is volume share of venue *k* on day *t*. The value of the index ranges from 0 to 1; higher value implies less fragmentation.

The results presented in Table 5 show that, as predicted by our framework, aggressive orders are positively linked with the pre-flash crash period. The result holds for a combination of buy and sell aggressive orders as well as for each type of aggressive orders separately. The positive and statistically significant coefficients suggest the predictive power of the number of aggressive orders for flash crashes. An essential point to note is that the relationship between aggressive orders and pre-flash crash period is statistically significant even after controlling for volume, liquidity, order flow toxicity and volatility. This finding is important given recent findings by Andersen and Bondarenko (2014), showing that a *popular* metric for order flow toxicity, the *VPIN* metric, developed by Easley et al. (2012), Easley et al. (2011), is a poor predictor for flash crashes once trading activity is controlled for. The practical implication of this finding is that traders seeking to avoid the adverse effects of a flash crash must act quickly to do so. However, their actions could be inevitably endogenous, leading to a self-fulfilling prophecy, as their actions could exacerbate what might already be proving to be a challenging and increasingly illiquid trading environment. As already noted, according to the existing literature and the predictions of our approach, we expect that sell orders to play a more important role in the flash crash (SEC 2010) and therefore, estimation separate regressions for aggressive sell and buy orders may provide more insightful results. This expectation is by the magnitude of the coefficient estimates and explanatory power for both the buy and sell aggressive orders estimations. Firstly, the coefficient estimate for aggressive sell orders is 2.3 times higher than the coefficient for the number of aggressive buy orders. Secondly, according to the *McFadden's* R^2, the model with the sell order has a higher explanatory power.

The estimated coefficients for all the other explanatory variables, except *MF* (market fragmentation), are also significantly correlated with the pre-flash crash period; however, the aggressive orders variables (*NAO*, *NASO* and *NABO*) are the only positive and statistically significant predictors. As already noted, our model allows us to test the flash crash predictability potential of *VPIN* after controlling for trading activity, liquidity and volatility. Our findings show that *VPIN* is negatively correlated with the pre-flash period; increases in the value of the *VPIN* metric does not provide a signal about extreme volatility. This is in some ways an unsurprising result, since Andersen and Bondarenko (2014) also show that *VPIN* is negatively correlated with future short-term volatility after controlling for trading activity. The explanatory power of the standard logit model reported for the *NAO*, *NASO* and *NABO* regressions using *McFadden's* R^2, are 2.5%, 2.9% and 2.51% respectively. This is also unsurprising because of the following two reasons. Firstly, we employ one-second frequency for the estimations.[13] Secondly, although *McFadden's* R^2 is a similar measure of the goodness of fit to the classic R^2, the value of *McFadden's* R^2 tend to be remarkably lower than the value of R^2 (see David and Peter 1979).

INSERT TABLE 6 ABOUT HERE

Table 6 presents the results for the multinominal logit model estimation. We employ this model to test the consistency of the standard logit model and in order to examine the relationship between contemporaneous order aggressiveness on the one hand and the pre-flash crash and the flash crash period on the other. This approach expectedly leads to a higher model explanatory power for the multinominal logit model estimation (*McFadden's* R^2 of 6.9% and 6.6% for the number of aggressive sell and buy orders, respectively) when compared with the standard logit model estimation reported in Table 5. Firstly, the findings in Table 6 are generally consistent with the results we present in Table 5; all the aggressive orders variables are positively and significantly related with the pre-flash crash period, which suggests that the number of aggressive orders can provide a signal about an impending flash crash. Furthermore, consistent with the findings from Table 5, the number of aggressive sell orders play a more

[13] *McFadden's* R^2 rises to about 4.5% when we estimate the regression at five-second frequencies; the results are not presented for parsimony, but are available on request.

important role in the flash crash. The only difference in the results is that while market fragmentation (*MF*) is not statistically significant in the standard logit model, it is significantly and positively correlated with the pre-flash period in the multinominal logit model. This implies that prior market fragmentation is related to flash crashes (see also Madhavan 2011, Menkveld and Yueshen 2017). The second set of results in Table 6, based on the flash crash period itself, are also interesting. The results show that the *NAO*, *NASO* and *NABO* are positively and significantly correlated with the flash crash period even after controlling for volume, liquidity and volatility. The positive and statistically significant estimates of the aggressive orders variables appear to confirm that increases in aggressive orders make flash crashes more likely to ensue. Specifically, the results suggest that the probability of flash crashes at time t rises as the number of aggressive orders increases at the same time. The evidence is in line with our approach that order aggressiveness plays an important role in flash crashes.

The regression results above, documenting the relationship between order aggressiveness and flash crashes, are consistent with the previous literature since they show that aggressive orders have a larger price impact than non-aggressive orders and that aggressive trading behaviour contributes to flash crashes (see as examples Griffiths et al. 2000, Mcinish et al. 2014, Wuyts 2011).

The estimated coefficient estimates for all the other explanatory variables in Table 6 are also consistent with the existing literature on flash crashes. For example, the market toxicity metric, *VPIN*, has a statistically significant and positive relationship with the flash crash period. Taken together with the metric's documented relationship with the pre-flash crash period, the implication here is that while *VPIN*, may be a poor predictor of flash crashes when trading activity is controlled for (see also Andersen and Bondarenko 2014), it nevertheless is positively correlated with flash crashes themselves. This suggests that market toxicity has a direct relationship with the flash crash; this evidence is in line with findings of Easley et al. (2012), Easley et al. (2011) that market toxicity plays an important role in the flash crash. Volatility exhibits a statistically significant and positive relationship with the flash crash. The positive coefficient is consistent with the stream of the market microstructure literature that states that an increase in the volatility of stock prices causes a larger price impact, since extreme price movements and flash crashes are characterized by extreme price volatility (see as examples Easley et al. 2011, Kirilenko et al. 2017, SEC 2010). One plausible explanation of this positive relationship is that an increase in the volatility of stock prices increases the market risk, which in turn leads to larger spreads and extreme price movements.

The literature identifies order imbalance as one of the instigators of the May 6, 2010 flash crash (see as examples Easley et al. 2011, Kirilenko et al. 2017, SEC 2010). Furthermore, Sun and Ibikunle (2016) find that order imbalance has information content and there is a significant and positive relationship between order imbalance and price impact in a high frequency trading environment. Thus, the positive relationship between *OIB* and the flash crash reported in Table 6 is unsurprising and is in line with the literature. The bid-ask spread, *BAS*, is also positively and statistically significantly related with the May 6, 2010 flash crash. This result is again unsurprising because existing literature finds that orders have a larger price impact when the bid-ask spread is wide (see Aitken and Frino 1996) and, as already enumerated, liquidity constraints contribute to extreme price movements in the market. Furthermore, Borkovec et al. (2010), SEC (2010), and Menkveld and Yueshen (2017) find that the spread during the May 6, 2010 flash crash was uncharacteristically wide. Market fragmentation, *MF*, exhibits a statistically significant and positive relationship with the flash crash as well; this result can be justified that market fragmentation is important in explaining the anatomy of the flash crash. This result underscores the results of Madhavan (2011), Golub et al. (2012), and Menkveld and Yueshen (2017) that show that the flash crash is linked directly to market structure. When liquidity is fragmented across several venues, immediate access to counterparties becomes slightly more challenging given that orders may now need to be routed through several other channels in order for them to be filled.

4.3. Directional returns during the flash crash

We now turn our attention to the third mainline argument derived from our framework, which is that aggressive orders are more profitable during flash crashes. Earlier, we observe an increase in the volume of aggressive orders during the flash crash, we interpret this to be in response to their profitability during such periods. However, we also note that such increases may relate to the unwinding of untenable positions that arise as a result of extreme swings in instruments' valuations during a flash crash. In order to examine the veracity of our argument regarding the profitability of aggressive orders, we follow the approach proposed by Ederington and Lee (1995) to compute hypothetical returns attributable to an informed trader active during the flash crash and its surrounding periods (see also Caminschi and Heaney 2014, Frino et al. 2017).

We estimate simple returns for each stock and sign the returns using a directional parameter ($DIR_{t,s}$), based on the assumption that the informed trader holds private information regarding the trajectory of the stocks' prices she trades. We define the directional return for each one-minute interval as

$$DR_{t,s} = R_{t,s} * DIR_{t,s} \tag{11}$$

where, $R_{t,s}$ represents simple return for stock *s* for time *t*. In order to define the directional parameter ($DIR_{t,s}$), firstly we compute the returns of each stock for the flash crash period (from 14:32 PM to 15:08 PM) ($R_{fc,s}$). The direction factor, $DIR_{t,s} = 1$ if $R_{fc,s} > 0$, $DIR_{t,s} = -1$ if $R_{fc,s} < 0$, and $DIR_{t,s} = 0$ if $R_{fc,s} = 0$. $DIR_{t,s} = 1\ (-1)$ indicates that the trader takes a long (short) position at time *t* for stock, *s*. We compute the average directional return, ADR_t, as the average of adjusted returns for all stocks for each one-minute interval. The cumulative average directional return, $CADR_t$ from 1:30 PM to 4:00 PM is estimated using the average directional returns.

INSERT FIGURE 5 AND TABLE 7 ABOUT HERE

Figure 5 reports the hypothetical returns attainable through aggressive (directional) trading in 53 selected S&P 500 stocks around the May 6, 2010 flash crash. Panel A shows the simple returns adjusted for direction of price movement over the flash crash period averaged across all 53 stocks, while Panel B shows the cumulative average direction-adjusted returns for the same stocks. As presented in Panel A, there are positive and significant directional returns during the flash crash. Remarkably, as predicted by our framework, the positive directional return is gained during the second half of the flash crash and only ends at the end of the flash crash at about 3:08 PM. The cumulative directional returns in Panel B shows the clear and continuous trend in adjusted returns during the flash crash period. This and the stabilisation of the cumulative returns following the conclusion of the flash crash support our arguments about the profitability of aggressive orders during periods of extreme price movements like flash crashes. The overall cumulative returns accruable to an informed trader during the flash crash is in excess of 1,482 basis points.

Table 7 reports the average direction-adjusted returns in 10-minute batches. Consistent with the insights from Figure 5, there is a positive and statistically significant adjusted returns, which commences in the second half of the flash crash and continues until the end of the flash crash. All estimated directional returns outside of the flash crash period are not statistically significant.

Overall, the directional returns analysis yields consistent results with the predictions of our framework, implying that aggressive orders are significantly more profitable during extreme price movements like flash crashes.

5. Conclusion

In this paper, we develop a new framework for understanding the role of aggressive orders in flash crashes by extending the approach of Menkveld (2013). We then use ultra-high frequency data from 53 S&P 500 stocks affected by the May 6, 2010 flash crash to test the arguments motivated by the framework. The selection of the May 6, 2010 flash crash for our investigation is motivated by its recognition as the most significant flash crash in recent financial markets history. Our main framework predictions/arguments are as follows. Firstly, there should be a significant increase in sell order aggressiveness prior to and during the first half of flash crashes, i.e. until instruments' price levels hit their lowest values and then the balance of order aggressiveness should shift to the buy side in the second half of the flash crash, i.e., until the prices re-attain their pre-crash levels). Secondly, our framework predicts that order aggressiveness is culpable in flash crashes and, therefore, flash crashes could be predicted by evaluating changes in the number or proportion of aggressive orders in the market. Thirdly, aggressive orders should be more profitable during extreme price movements and thus traders tend to submit orders that are more aggressive during those periods.

In the formal test of the relationship between the number of aggressive orders and the pre-flash crash period, the empirical results are consistent with the predictions of our framework. Firstly, we find a significant increase in sell order aggressiveness prior to and during the first half of the May 6 2010 flash crash, thereafter the balance of order aggressiveness swings to the buy side, with traders submitting more aggressive buy orders relative to aggressive sell orders. The sell side is more aggressive until prices plummet to their lowest levels and then, the buy side becomes more aggressive in the run-up to prices regaining their pre-crash levels. Secondly, we find that the number of aggressive orders is positively and significantly related to the pre-flash crash period; thus, the number or level of aggressive orders may signal flash crashes. Thirdly, the fraction and the number of aggressive orders during the flash crash are higher than the fraction and the number of orders during the surrounding periods due to the significantly larger (than other periods) profits accruable to informed investors during the flash crash. We estimate that for the stocks in our sample, an informed investor during the flash crash could achieve a return on his portfolio in excess of 1,482 bps, a return far larger than accruable during surrounding periods. This finding supports our argument that aggressive orders are more profitable markets are volatile and hence, traders tend to submit orders that are more aggressive during such periods.

Our findings should not be misconstrued as an endorsement of policies aimed at limiting aggressive orders or aggressive trading behaviours in financial markets. While we acknowledge that aggressive traders can induce extreme price movements, aggressive trading in itself could be a symptom of deeper underlying structural issues, which are not the focus of this study.

References

Aitken, M. and Frino, A. (1996) 'Execution costs associated with institutional trades on the Australian Stock Exchange', *Pacific-Basin Finance Journal,* 4(1), 45-58.

Aldridge, I. (2010) *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems* 2nd ed., Wiley.

Andersen, T. and Bondarenko, O. (2014) 'VPIN and the flash crash', *Journal of Financial Markets,* 17, 1-46.

Avramov, D., Chordia, T. and Goyal, A. (2006) 'The Impact of Trades on Daily Volatility', *The Review of Financial Studies,* 19(4), 1241-1277.

Biais, B., Hillion, P. and Spatt, C. (1995) 'An empirical analysis of the limit-order book and order flow in the Paris Bourse', *The Journal of Finance,* 50, 1655-1689.

Borkovec, M., Domowitz, I., Serbin, V. and Yegerman, H. (2010) 'Liquidity and Price Discovery in Exchange-Traded Funds: One of Several Possible Lessons from the Flash Crash', *The Journal of Index Investing,* 1(2), 24-42.

Brock, W. and Hommes, C. (1998) 'Heterogeneous beliefs and routes to chaos in a simple asset pricing model', *Journal of Economic Dynamics and Control,* 22(8), 1235-1274.

Brogaard, J., Carrion, A., Riordan, R., Shkilko, A. and Sokolov, K. (2014) *High Frequency Trading and Extreme Price Movements*, unpublished.

Caminschi, A. and Heaney, R. (2014) 'Fixing a leaky fixing: short-term market reactions to the London pm gold price fixing', *Journal of Futures Markets,* 34, 1003-1039.

Chakrabarty, B., Moulton, P. and Shkilko, A. (2015) 'Evaluating Trade Classification Algorithms: Bulk Volume Classification versus the Tick Rule and the Lee-Ready Algorithm', *Journal of Financial Markets,* 25, 52-79.

Chordia, T., Roll, R. and Subrahmanyam, A. (2001) 'Market Liquidity and Trading Activity', *Journal of Finance,* 56(2), 501-530.

Chordia, T., Roll, R. and Subrahmanyam, A. (2008) 'Liquidity and market efficieny', *Journal of Financial Economics,* 87, 249-268.

David, S. H. and Peter, R. H. (1979) *Behavioural travel modelling,* London: Croom Helm.

Degryse, H., Jong, F., Ravenswaaij, M. and Wuyts, G. (2005) 'Aggressive Orders and the Resiliency of a Limit Order Market', *Review of Finance,* 9(2), 201-242.

Easley, D., De Prado, M. and 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(2), 118-129.

Easley, D., De Prado, M. and O'Hara, M. (2012) 'Flow Toxicity and Liquidity in a High-frequency World', *The Review of Financial Studies,* 25(5), 1457-1493.

Easley, D., Kiefer, N., O'Hara, M. and Paperman, J. (1996) 'Liquidity, Information, and Infrequently Traded Stocks', *The Journal of Finance,* 51(4), 1405-1436.

Easley, D., Kiefer, N. M. and O'Hara, M. (1997) 'One Day in the Life of a Very Common Stock', *The Review of Financial Studies,* 10(3), 805-835.

Ederington, L. H. and Lee, J. H. (1995) 'The Short-Run Dynamics of the Price Adjustment to New Information', *the Journal of Financial and Quantitative Analysis,* 30(1), 117-134.

Foucault, T. (1999) 'Order flow composition and trading costs in a dynamic limit order market', *Journal of Financial Markets,* 2, 99-134.

Frino, A., Ibikunle, G., Mollica, V. and Steffen, T. (2017) 'The impact of commodity benchmark on derivatives markets: The case of the dated Brent assessment and Brent futures', *journal of Banking & Finance*, 1-17.

Glosten, L. and Milgrom, P. (1985) 'Bid, ask, and transaction prices in a specialist market with heterogeneously informed agents', *Journal of Financial Economics,* 14, 71-100.

Golub, A., Keane, J. and Poon, S. (2012) *High Frequency Trading and Mini Flash Crashes*, unpublished.

Griffiths, M., Smith, B., Turnbull, A. and White, R. (2000) 'The costs and determinants of order aggressiveness', *Journal of Financial Economics,* 56, 65-88.

Hagströmer, B., Nordén, L. and Zhang, D. (2014) 'How Aggressive Are High-Frequency Traders?', *The Financial Review,* 49, 395-419.

Hirschey, N. (2017) *Do High-Frequency Traders Anticipate Buying and Selling Pressure?*, London Business School: unpublished.

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

Jacob Leal, S., Napoletano, M., Roventini, A. and Fagiolo, G. (2016) 'Rock around the clock: An agent-based model of low- and high-frequency trading', *Journal of Evolutionary Economics,* 26(1), 49-76.

Kirilenko, A., Kyle, A., Samadi, M. and Tuzun, T. (2017) 'The Flash Crash: High Frequency Trading in an Electronic Market', *The Journal of Finance, Forthcoming*.

Lee, C. and Ready, M. (1991) 'Inferring Trade Direction from Intraday Data', *The Journal of Finance,* 46(2), 733-746.

Luce, D. (1959) *Individual Choice Behavior,* New York: Wiley.

Madhavan, A. (2011) *Exchange-Traded Funds, Market Structure and the Flash Crash*, unpublished.

Manski, C. and McFadden, D. (1981) *Structural analysis of discrete data with econometric applications,* Cambridge: MIT Press.

McFadden, D. (1974) *Frontiers in Econometrics,* New York: Academic Press.

Mcinish, T., Upson, J. and Wood, R. (2014) 'The Flash Crash: Trading Aggressiveness, Liquidity Supply, and the Impact of Intermarket Sweep Orders', *Financial Review,* 49(3), 481-509.

Menkveld, A. and Yueshen, B. (2017) *The Flash Crash: A Cautionary Tale about Highly Fragmented Markets*, unpublished.

Menkveld, A. J. (2013) 'High frequency trading and the new market makers', *Journal of Financial Markets,* 16(4), 712-741.

Pellizzari, P. and Westerhoff, F. (2009) 'Some effects of transaction taxes under different microstructures', *Journal of Economic Behavior and Organization,* 72(3), 850-863.

SEC (2010) *Findings Regarding the Market Events of May 6, 2010*, unpublished.

Sofianos, G. (1995) *Specialist gross trading revenues at the New York Stock Exchange*, New York Stock Exchange: unpublished.

Sun, Y. and Ibikunle, Y. (2016) 'Informed trading and the price impact of block trades : A high frequency trading analysis', *International Review of Financial Analysis*.

Train, K. (2002) *Discrete Choice Methods with Simulation,* Cambridge: Cambridge University Press.

Westerhoff, F. (2008) 'The Use of Agent-Based Financial Market Models to Test the Effectiveness of Regulatory Policies', *Journal of Economics and Statistic,* 228(2-3), 195-227.

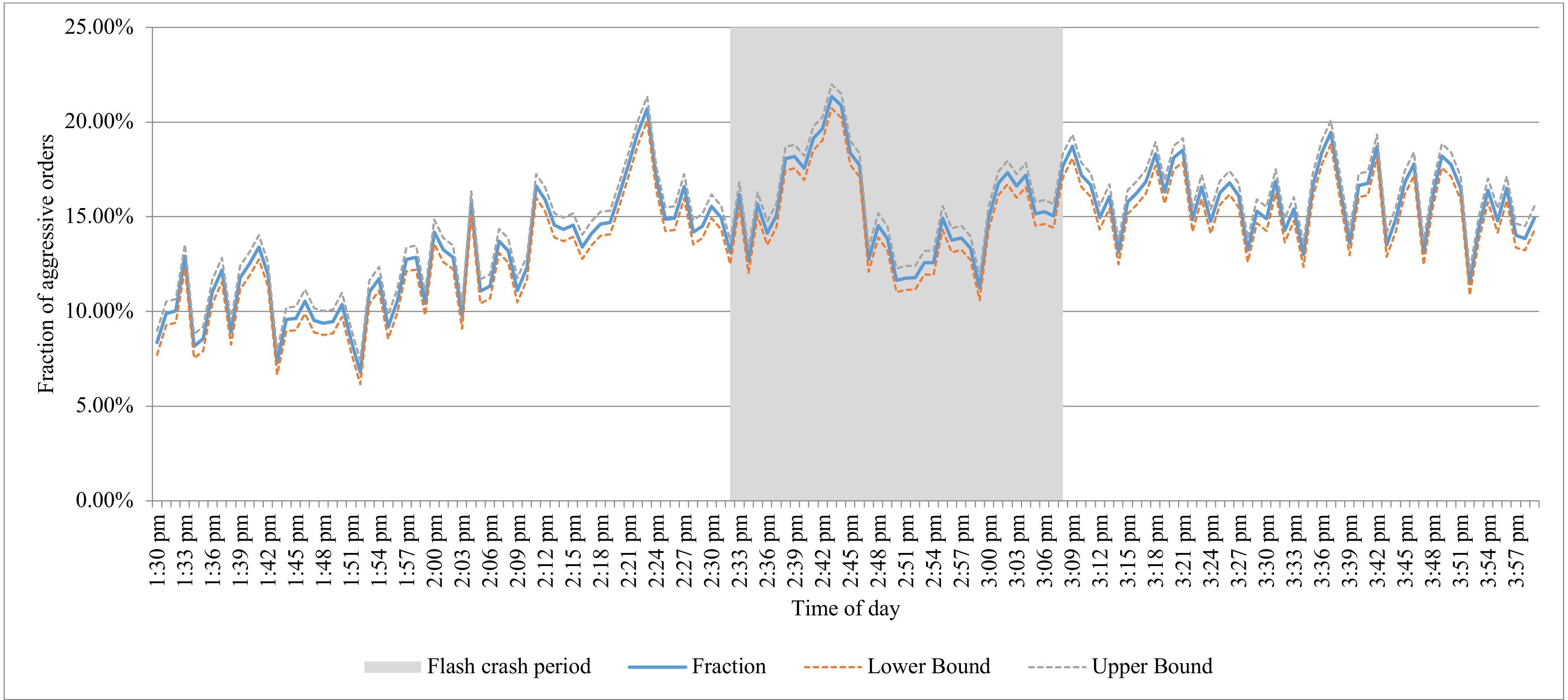
Wu, K., Bethel, W., Leinweber, D., Rübel, J. and Gu, M. (2013) 'A big data approach to analyzing market volatility', *Algorithmic Finance,* 2(3-4), 241-267.

Wuyts, G. (2011) 'The impact of aggressive orders in an order-driven market: a simulation approach', *The European Journal of Finance,* 18, 1015-1038.

Figure 1. Intraday evolution of the fraction of aggressive orders

Panels A and B depict the minute-by-minute evolution of the fraction of aggressive orders for 53 S&P 500 stocks affected by the May 6 2010 flash crash; Panel B presents the fraction of aggressive orders when disaggregated into buys and sells, as well as the fraction of all aggressive orders, while Panel presents only the fraction of all aggressive orders. 99% confidence bands are constructed for Panel A using the means of the minute-by-minute fractions of aggressive orders across the stocks in the sample. The sample period covers 1:30 PM to 4 PM May 6, 2010. The shaded area indicates the flash crash period.

Panel A. Fraction of total aggressive orders



Panel B. Fraction of total, buy and sell aggressive orders

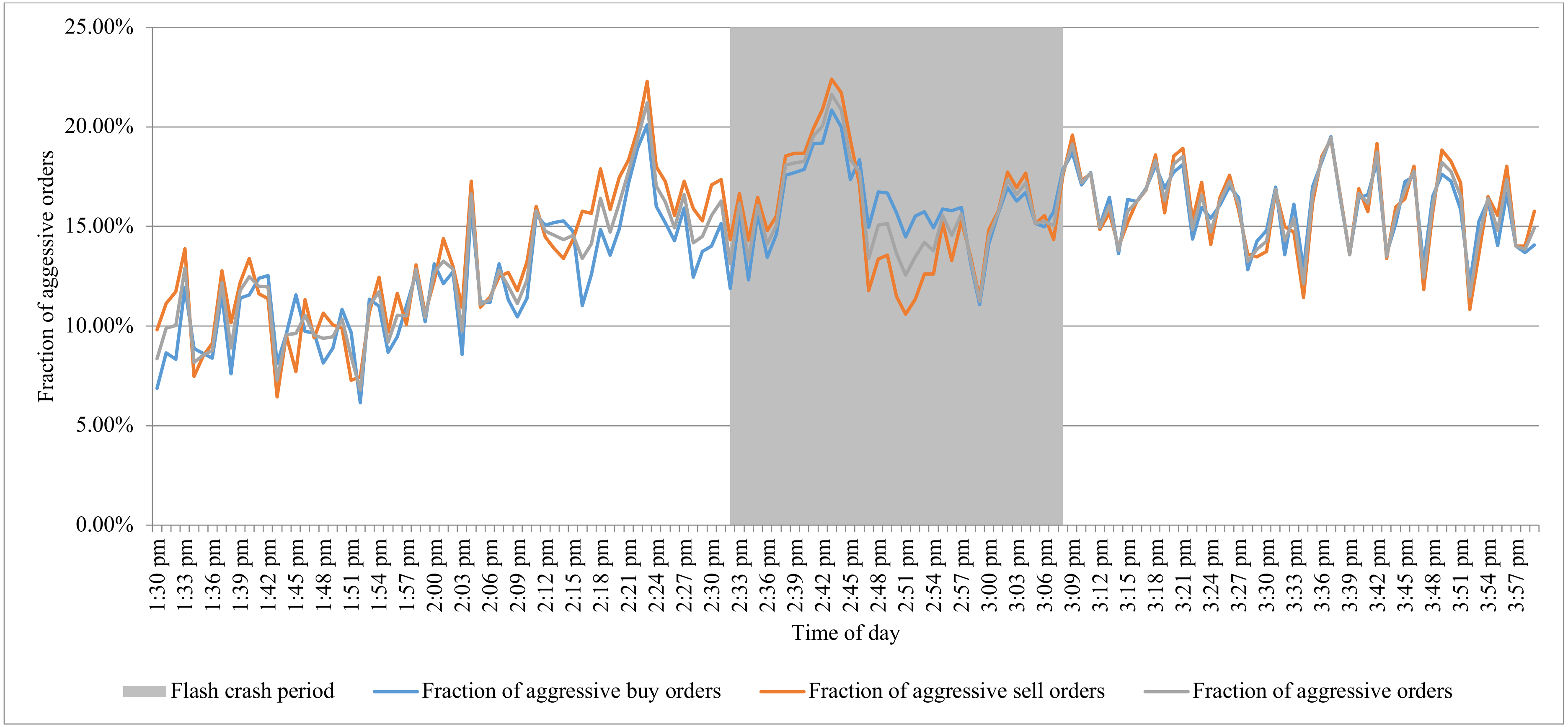


Figure 2. Intraday evolution of aggressive order imbalance I

The figure presents the minute-by-minute evolution of aggressive order imbalance (difference between the fractions of aggressive sell and buy orders) for 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. The shaded are indicates the flash crash period. The shaded area indicates the flash crash period.

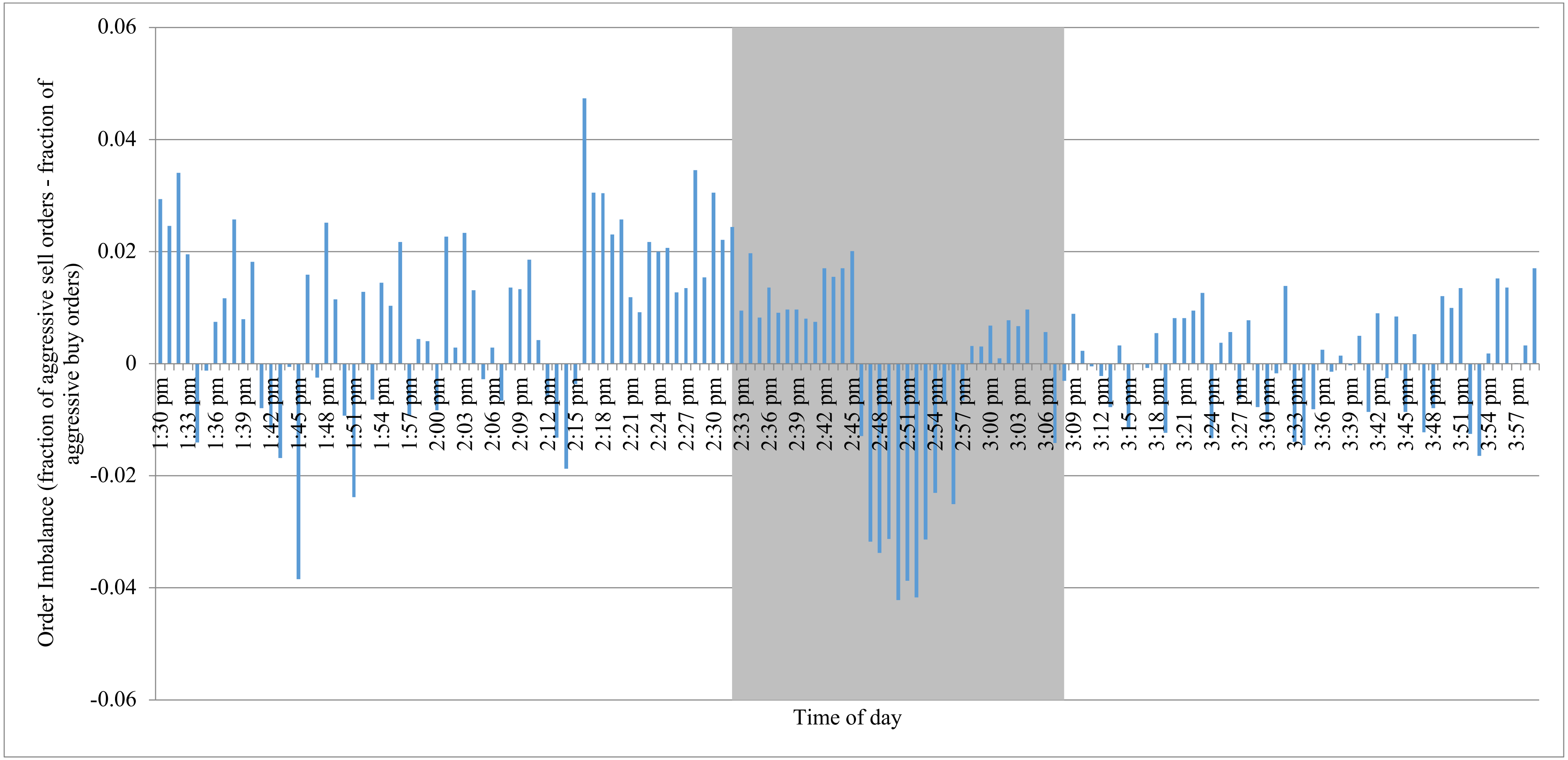


Figure 3. Intraday evolution of aggressive orders

The figure presents the minute-by-minute evolution of the numbers of total, sell and buy aggressive orders for 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. The shaded area indicates the flash crash period. The shaded area indicates the flash crash period.

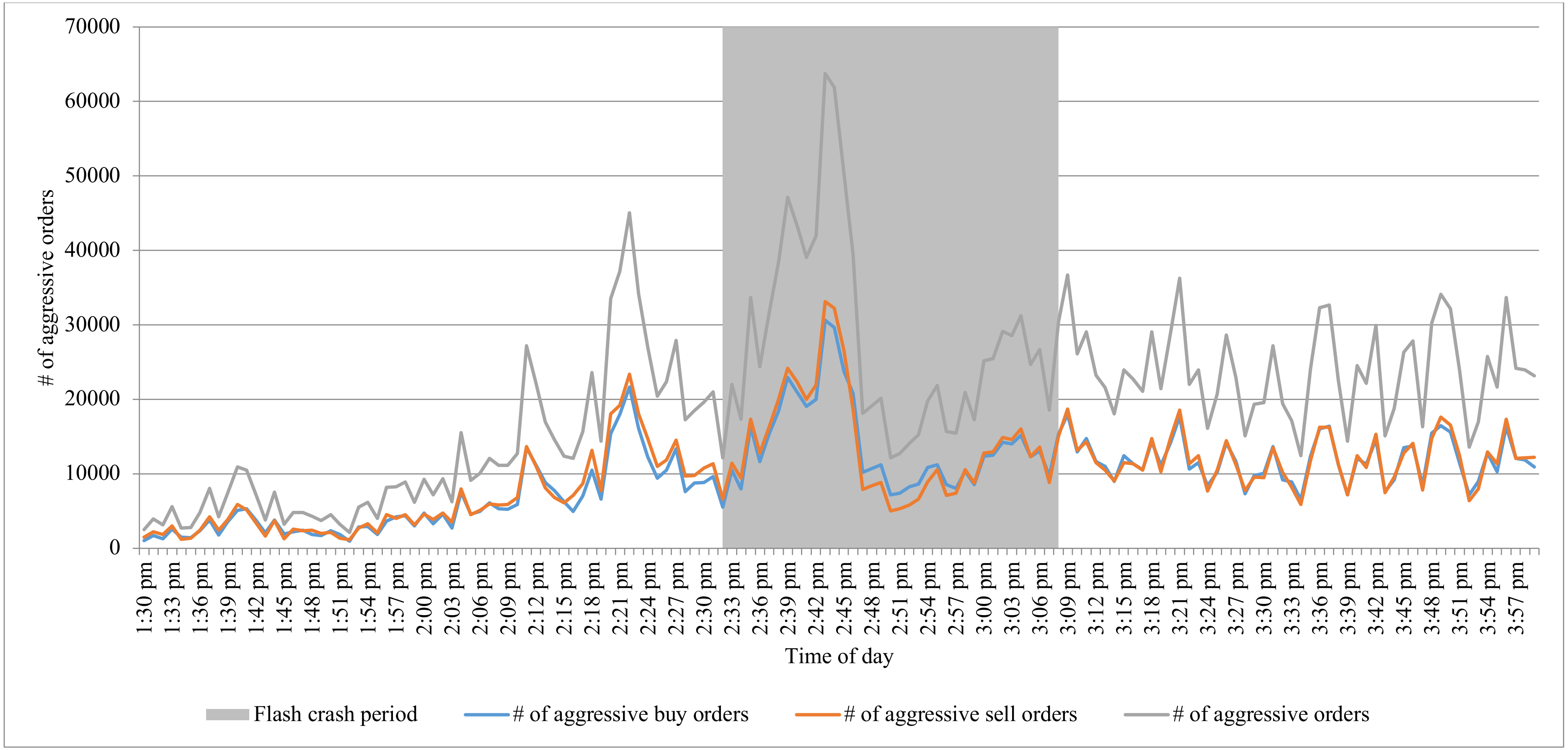


Figure 4. Intraday evolution of aggressive order imbalance II

The figure presents the minute-by-minute evolution of aggressive order imbalance (difference between the number of aggressive sell and buy orders) for 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. The shaded area indicates the flash crash period. The shaded area indicates the flash crash period.

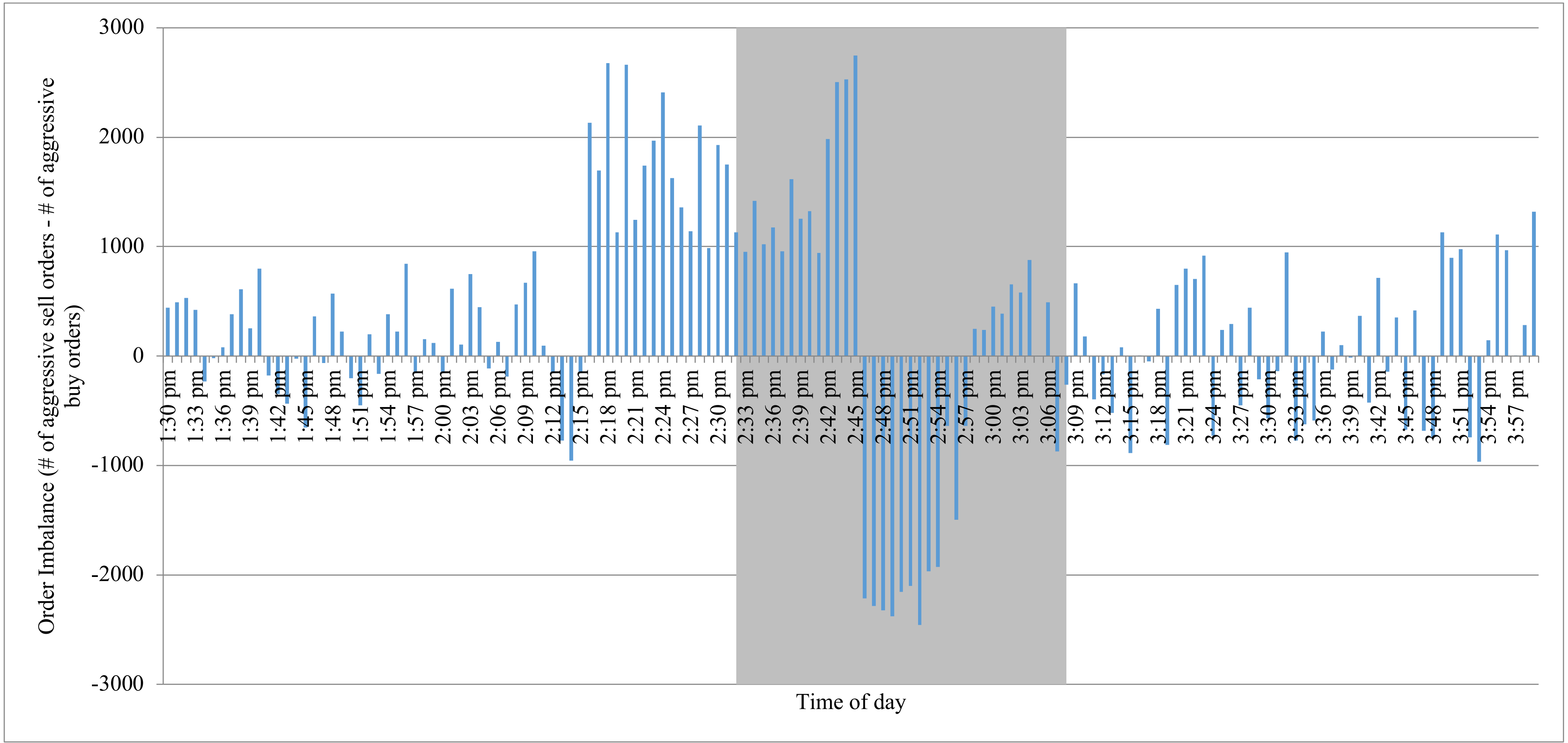
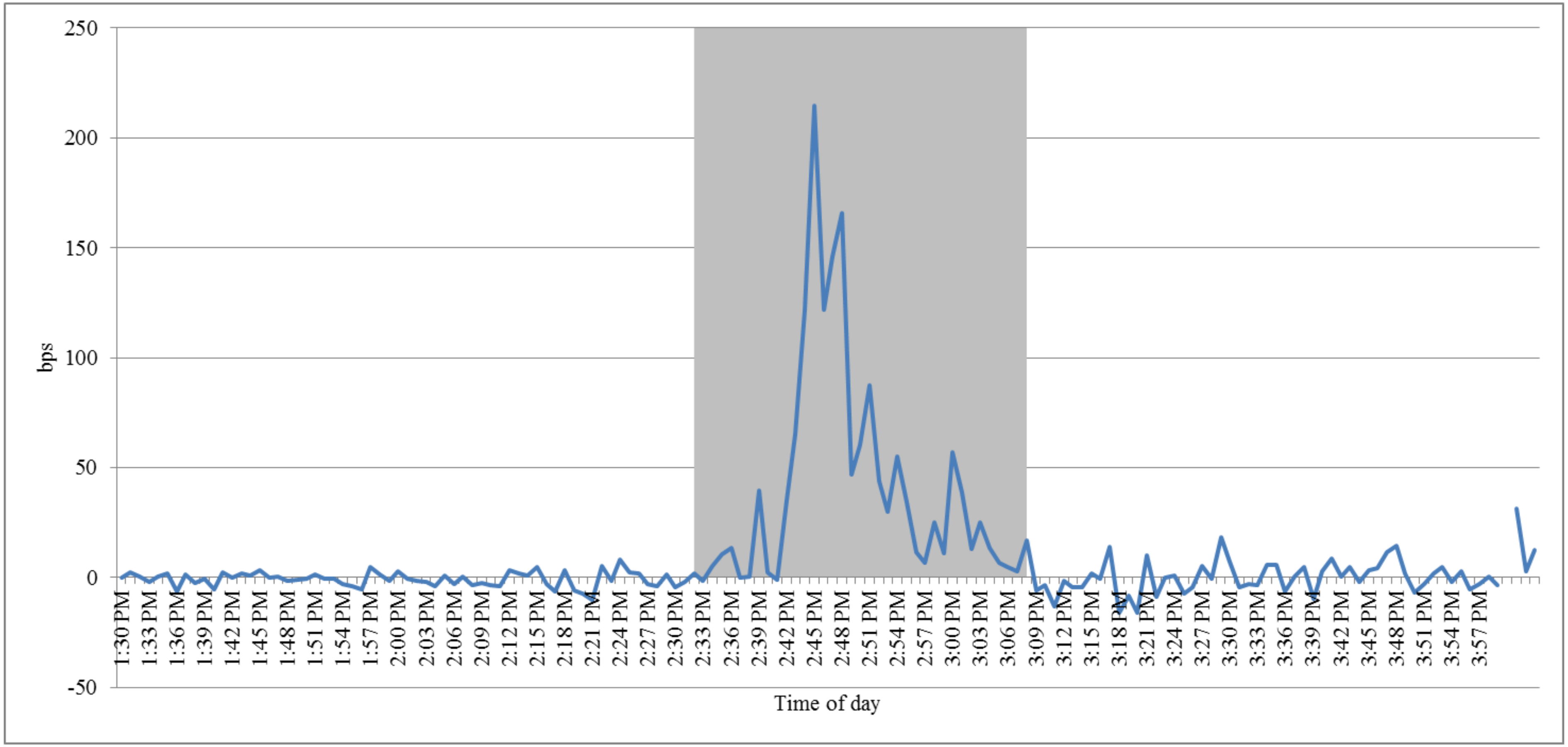


Figure 5. Directional returns

Panels A and B are minute-by-minute plots of average direction-adjusted returns and cumulative average direction-adjusted returns measures (in basis points) respectively for 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. The shaded area indicates the flash crash period.

Panel A. Average direction-adjusted returns



Panel B. Cumulative Average Adjusted Return

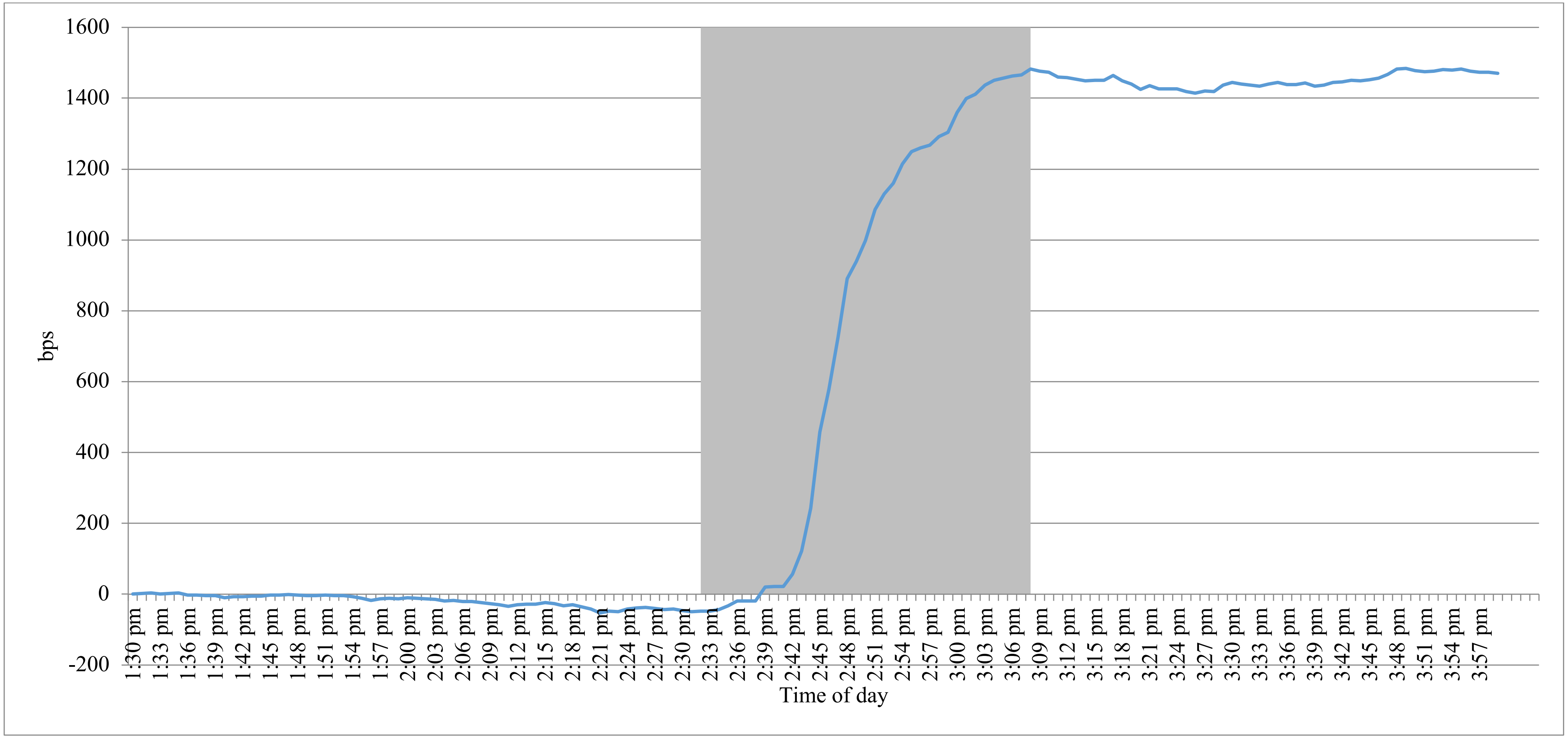


Table 1. Transactions' summary statistics and statistical tests

Panels A and B respectively present trading summary statistics and statistical tests of differences between the period of the flash crash and surrounding periods for 53 S&P 500 stocks affected by the May 6 2010 flash crash. The statistical tests conducted are two-sample t-tests and pairwise Wilcoxon-Mann-Whitney U tests. The sample period covers 1:30 PM to 4 PM May 6, 2010. The time series on May 6 2010 is divided into three: before the flash crash (from 1:30 PM to 2:32 PM), the flash crash period (from 2:32 PM to 3:08 PM), and after the flash crash (from 3.08 PM to 4 PM).

Panel A. Summary statistics

		Total transactions (000s)	Average per minute transactions (000s)
Number of Transactions	1:30 PM – 2:32 PM	186.6	3.0
	2:32 PM – 3:08 PM	329.9	8.9
	3:08 PM – 4 PM	405.8	7.8
	All	922.3	19.7

		Total trading volume (000s)	Average per minute trading volume (000s)
Trading Volume	1:30 PM – 2:32 PM	62878.8	1014.2
	2:32 PM – 3:08 PM	98185.5	2653.7
	3:08 PM – 4 PM	119209.9	2292.5
	All	280274.2	5960.4

		Total dollar trading volume ($'000,000)	Average per minute dollar trading volume ($'000,000)
Dollar Trading Volume	1:30 PM – 2:32 PM	2541.6	41.0
	2:32 PM – 3:08 PM	4332.2	117.1
	3:08 PM – 4 PM	5239.7	100.8
	All	12113.5	258.9

Panel B. Statistical tests

Trading volume	
Method	p-value
Two-Sample T tests	
Pooled	<0.0001
Satterthwaite	<0.0001
Wilcoxon-Mann-Whitney U tests	<0.0001
Dollar trading volume	
Method	**p-value**
Two-Sample T tests	
Pooled	<0.0001
Satterthwaite	<0.0001
Wilcoxon-Mann-Whitney U tests	<0.0001

Table 2. Order quoting summary statistics

Table presents order quoting summary statistics for 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. The time series on May 6 2010 is divided into three: before the flash crash (from 1:30 PM to 2:32 PM), the flash crash period (from 2:32 PM to 3:08 PM), and after the flash crash (from 3.08 PM to 4 PM).

		Total number of shares at the bid side (000,000s)	Average shares/minute at the bid side (000,000s)
Number of shares in orders submitted at the bid side	1:30 PM – 2:32 PM	168.4	2.7
	2:32 PM – 3:08 PM	93.3	2.5
	3:08 PM – 4 PM	130.0	2.5
	All	391.7	7.7

		Total number of shares/minute at the ask side (000,000s)	Average shares/minute at the ask side (000,000s)
Number of Shares in orders submitted at the ask side	1:30 PM – 2:32 PM	168.3	2.7
	2:32 PM – 3:08 PM	85.7	2.3
	3:08 PM – 4 PM	123.9	2.4
	All	377.8	7.4

Table 3. Correlation matrix of explanatory variables

The table presents the correlation matrix for the explanatory variables employed in the flash crash models. *NAO* is the number of aggressive orders, *NASO* is the number of aggressive sell orders, *NABO* is the number of aggressive buy orders, *VPIN* is the Volume-Synchronized Probability of Informed Trading, *VLT* is the standard deviation of the mid-price returns, *OIB* is the order imbalance, *BAS* is a bid-ask spread, *MF* represents market fragmentation, and *lnV* is the natural logarithm of the number of shares. The sample includes 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010.

	NAO	NASO	NABO	VPIN	VLT	OIB	BAS	MF	lnV
NAO	1								
NASO	0.91	1							
NABO	0.90	0.92	1						
VPIN	0.014	0.013	0.014	1					
VLT	0.093	0.094	0.093	0.146	1				
OIB	0.385	0.383	0.384	0.054	0.139	1			
BAS	-0.013	-0.012	-0.014	0.10	0.27	0.03	1		
MF	0.273	0.274	0.273	0.12	0.20	0.13	0.05	1	
lnV	0.394	0.395	0.394	0.08	0.30	0.27	0.05	0.58	1

Table 4. Summary statistics and statistical tests for aggressive orders/quotes

Panels A and B present aggressive orders summary statistics and statistical tests of differences between the period of the flash crash and surrounding periods respectively for selected 53 S&P 500 stocks. The statistical tests conducted test the null of equality between the number of aggressive orders during the flash crash and the surrounding periods. The sample period covers 1:30 PM to 4 PM May 6, 2010. The time series on May 6 2010 is divided into three: before the flash crash (from 1:30 PM to 2:32 PM), the flash crash period (from 2:32 PM to 3:08 PM), and after the flash crash (from 3.08 PM to 4 PM).

Panel A. Summary statistics

		Total aggressive sell orders (000s)	Average per minute aggressive sell orders (000s)
Number of Aggressive Sell Orders	1:30 PM – 2:32 PM	397.3	6.4
	2:32 PM – 3:08 PM	515.4	13.9
	3:08 PM – 4 PM	608.4	11.9
	All	1519.5	10.1

		Total aggressive buy orders (000s)	Average per minute aggressive buy orders (000s)
Number of Aggressive Buy Orders	1:30 PM – 2:32 PM	362.8	5.8
	2:32 PM – 3:08 PM	513.7	13.8
	3:08 PM – 4 PM	604.4	11.8
	All	1480.9	9.8

		Total aggressive orders (000s)	Average per minute aggressive orders (000s)
Number of Aggressive Orders	1:30 PM – 2:32 PM	760.1	12.2
	2:32 PM – 3:08 PM	1029.2	27.8
	3:08 PM – 4 PM	1212.8	23.7
	All	3002.2	20

Panel B. statistical tests

	Mean number of aggressive orders
Flash Crash Period	28204.1
Surrounding Period	17413.7
Difference	10790.4

	Number of aggressive orders
Method	p-value
Two-Sample T tests	
Pooled	<0.0001
Satterthwaite	<0.0001
Wilcoxon-Mann-Whitney U tests	<0.0001

Table 5. Standard logit model for one second frequency

The predictive power of the number of aggressive orders on flash crashes is estimated using the following model:

$$FC_{it} = \alpha + \beta_{NAO} NAO_{it} + \beta_{\ln V} \ln V_{it} + \beta_{VPIN} VPIN_{it} + \beta_{VLT} VLT_{it}$$
$$+ \beta_{OIB} OIB_{it} + \beta_{BAS} BAS_{it} + \beta_{MF} MF_{it} + \varepsilon_{it}$$

The table reports logit regressions' coefficient estimates using one second frequencies. Results for standard logit model estimations are presented for the number of aggressive orders, aggressive sell orders and aggressive buy orders in the second, third and fourth columns respectively. FC_{it} equals zero from 1:30 PM to 2:17 PM, and from 2:32 PM to 4:00 PM, while it takes the value of one from 2:17 PM to 2:32 PM. *NAO*, *NASO* and *NABO* are the number of aggressive orders, number of aggressive sell orders and number of aggressive buy orders, respectively, *lnV* is the natural logarithm of the number of shares, *VPIN* is the Volume-Synchronized Probability of Informed Trading, *VLT* is the standard deviation of the mid-price returns, *OIB* is the order imbalance, *BAS* is the prevailing bid-ask spread and *MF* represents market fragmentation. Standard errors are presented in parentheses. The sample includes 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. *** and ** correspond to statistical significance at the 0.01 and 0.05 levels, respectively.

Variables	NAO	NASO	NABO
	4.98×10^{-3}***	1.66×10^{-2}***	7.12×10^{-3}***
	(2.18×10^{-4})	(7.27×10^{-4})	(3.12×10^{-4})
lnV	-1.23×10^{-2}***	-1.23×10^{-2}***	-1.23×10^{-2}***
	(2.37×10^{-3})	(2.37×10^{-3})	(2.37×10^{-3})
VPIN	-1.2531***	-1.2531***	-1.2530***
	(2.43×10^{-2})	(2.43×10^{-2})	(2.43×10^{-2})
VLT	-1460.6***	-1460.6***	-1460.5***
	(45)	(46.002)	(45.998)
OIB	-3×10^{-5}***	-3.11×10^{-5}***	-3×10^{-5}***
	(5.1×10^{-6})	(5.1×10^{-6})	(5.1×10^{-6})
BAS	-1.34***	-1.3432***	-1.3431***
	(6.48×10^{-2})	(6.48×10^{-2})	(6.48×10^{-2})
MF	1.70×10^{-3}	1.68×10^{-3}	1.7×10^{-3}
	(5.26×10^{-3})	(5.26×10^{-3})	(5.26×10^{-3})
Mc Fadden's R^2	0.025	0.0292	0.0251

Table 6. Multinomial logit model for one second frequency

The predictive power of the number of aggressive orders is estimated using the following model:

$$FC_{it} = \alpha + \beta_{NAO} NAO_{it} + \beta_{\ln V} \ln V_{it} + \beta_{VPIN} VPIN_{it} + \beta_{VLT} VLT_{it}$$
$$+ \beta_{OIB} OIB_{it} + \beta_{BAS} BAS_{it} + \beta_{MF} MF_{it} + \varepsilon_{it}$$

The table reports multinomial logit regressions' coefficient estimates using one second frequencies; Results for multinomial logit model estimations for the number of aggressive orders, the number of aggressive sell orders and the number of aggressive buy orders are presented in the second, third and fourth columns respectively. FC_{it} equals zero from 1:30 PM to 2:17 PM, and from 3:08 PM to 4:00 PM, while it takes the value of one from 2:17 PM to 2:32 PM (pre-flash crash period) and takes the value of two from 2:32 PM to 3:08 PM (the flash crash period). *NAO*, *NASO* and *NABO* are the number of aggressive orders, the number of aggressive sell orders and the number of aggressive buy orders, respectively, *lnV* is the natural logarithm of the number of shares, *VPIN* is the Volume-Synchronized Probability of Informed Trading, *VLT* is the standard deviation of the mid-price returns, *OIB* is the order imbalance, *BAS* is the prevailing bid-ask spread and *MF* represents market fragmentation. Standard errors are presented in parentheses. The sample includes 53 S&P 500 stocks affected by the May 6 2010 flash crash. The sample period covers 1:30 PM to 4 PM May 6, 2010. *** and ** correspond to statistical significance at the 0.01 and 0.05 levels, respectively.

	NAO		NASO		NABO	
Variables	FC = 1	FC = 2	FC = 1	FC = 2	FC = 1	FC = 2
	$6.30 \times 10^{-3***}$	$3.81 \times 10^{-3***}$	$2.09 \times 10^{-2***}$	$1.27 \times 10^{-2***}$	$9.0 \times 10^{-3***}$	$5.45 \times 10^{-3***}$
	(2.33×10^{-4})	(1.79×10^{-4})	(7.74×10^{-3})	(5.95×10^{-4})	(3.32×10^{-3})	(2.55×10^{-4})
lnV	$-2.75 \times 10^{-2***}$	$4.71 \times 10^{-2***}$	$-2.75 \times 10^{-2***}$	$4.71 \times 10^{-2***}$	$-2.74 \times 10^{-2***}$	$4.71 \times 10^{-2***}$
	(2.43×10^{-3})	(1.77×10^{-3})	(2.43×10^{-3})	(1.77×10^{-3})	(2.43×10^{-3})	(1.77×10^{-3})
VPIN	$-4.25 \times 10^{-1***}$	2.88^{***}	$-4.25 \times 10^{-1***}$	2.88^{***}	$-4.25 \times 10^{-1***}$	2.88^{***}
	(2.53×10^{-2})	(1.72×10^{-2})	(2.53×10^{-2})	(1.72×10^{-2})	(2.53×10^{-2})	(1.72×10^{-2})
VLT	-1045.2^{***}	1106.9^{***}	-1045.5^{***}	1106.8^{***}	-1045.0^{***}	1106.9^{***}
	(47.81)	(19.25)	(47.81)	(19.24)	(47.81)	(19.24)
OIB	$-2.00 \times 10^{-5***}$	$2.4 \times 10^{-5***}$	$-2.10 \times 10^{-5***}$	$2.4 \times 10^{-5***}$	$-2.10 \times 10^{-5***}$	$2.4 \times 10^{-5***}$
	(5.27×10^{-6})	(2.74×10^{-6})	(5.27×10^{-6})	(2.74×10^{-6})	(5.27×10^{-6})	(2.74×10^{-6})
BAS	$-5.2 \times 10^{-1***}$	2.05^{***}	$-5.2 \times 10^{-1***}$	2.05^{***}	$-5.2 \times 10^{-1***}$	2.05^{***}
	(6.94×10^{-2})	(3.08×10^{-2})	(6.94×10^{-2})	(3.08×10^{-2})	(6.94×10^{-2})	(3.08×10^{-2})
MF	$4.18 \times 10^{-2***}$	$1.71 \times 10^{-1***}$	$4.18 \times 10^{-2***}$	$1.71 \times 10^{-1***}$	$4.18 \times 10^{-2***}$	$1.71 \times 10^{-1***}$
	(5.36×10^{-3})	(4.24×10^{-3})	(5.36×10^{-3})	(4.24×10^{-3})	(5.36×10^{-3})	(4.24×10^{-3})
Mc Fadden's R^2	0.066		0.0702		0.0665	

APPENDIX A. Framework variables and definitions

VARIABLE	DEFINITION
P_t^{a}	Ask Price at time *t*.
P_t^{b}	Bid Price at time *t*.
P_t^{mp}	Mid-Price at time *t*.
$P_t^{ag.a}$	Ask Price at time *t* under aggressive trading strategy.
$P_t^{ag.b}$	Bid Price at time *t* under aggressive trading strategy.
$P_t^{ag.mp}$	Mid-Price at time *t* under aggressive trading strategy.
$P_t^{mm.a}$	Ask Price at time *t* under market-making strategy.
$P_t^{mm.b}$	Bid Price at time *t* under market-making strategy.
$P_t^{mm.mp}$	Mid-Price at time *t* under market-making strategy.
π_t^{ag}	Profit at time *t* under aggressive trading strategy.
$\pi_t^{ag.ba}$	Profit from bid-ask spread at time *t* under aggressive trading strategy.
$\pi_t^{ag.p}$	Profit from position at time *t* under aggressive trading strategy.
U_t^{st}	Attractiveness or fitness function of each type of strategy at time *t*.
η	Memory parameter.
ξ	Intensity of switching parameter.
φ_t^{ag}	Relative weight of aggressive trading strategy at time *t*.

APPENDIX B. Proof of the logit (discrete choice) model

Logit is the most popular and easiest discrete choice model and has been extensively used to compute the relative weights of each trading strategy (see as examples Jacob Leal et al. 2016, Pellizzari and Westerhoff 2009, Westerhoff 2008). Originally introduced by Luce (1959), it was further developed by McFadden (1974). McFadden (1974) and Train (2002) derive the logit choice probabilities model as a specific form of a discrete choice model. We introduce their proof step-by-step:

The choice under logit model is based on unobserved utility, and this utility is decomposed into two components, i.e. the known (V_{nj}) and unknown (ε_{nj}) parameters:

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \forall j \tag{B1}$$

McFadden (1974) shows that the unobserved utility under logit model is a distributed extreme value. Therefore, it is assumed that the unknown parameter (ε_{nj}) is an independently and identically distributed extreme value, and the density for each ε_{nj} is given as:

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}} \tag{B2}$$

and the cumulative distribution is:

$$F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}} \tag{B3}$$

By using the formula of unobserved utility and following McFadden (1974), we can say that the probability of selecting alternative *I* is determined as follows:

$$\begin{aligned} P_{ni} &= \Pr(U_{ni} \rangle U_{nj}, \forall j \neq i) = \Pr(V_{ni} + \varepsilon_{ni} \rangle V_{nj} + \varepsilon_{nj}, \forall j \neq i) = \\ &= \Pr(\varepsilon_{nj} \langle \varepsilon_{ni} + V_{ni} - V_{nj}, \forall j \neq i) \end{aligned} \tag{B4}$$

If ε_{ni} is considered given, this formula is the cumulative distribution for each ε_{nj} evaluated at $\varepsilon_{ni} + V_{ni} - V_{nj}$ and, according to Equation B3, it is $\exp(-\exp(-(\varepsilon_{ni} + V_{ni} - V_{nj})))$. Since ε_{nj} and ε_{ni} are independent, the cumulative distribution is the product of the individual cumulative distributions:

$$P_{ni} \mid \varepsilon_{ni} = \prod_{j \neq 1} e^{-e^{-(\varepsilon_{ni} + V_{ni} - V_{nj})}} \tag{B5}$$

Then we can derive the logit choice probabilities as follows:

$$P_{ni} = \int (\prod_{j\neq 1} e^{-e^{-(\varepsilon_{ni}+V_{ni}-V_{nj})}}) e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni} \tag{B6}$$

If s is ε_{ni}, then:

$$\begin{aligned} P_{ni} &= \int_{s=-\infty}^{+\infty} (\prod_{j\neq 1} e^{-e^{-(\varepsilon_{ni}+V_{ni}-V_{nj})}}) e^{-e_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni} = \int_{s=-\infty}^{+\infty} (\prod_{j\neq 1} e^{-e^{-(s+V_{ni}-V_{nj})}}) e^{-s} e^{-e^{-s}} ds = \\ &= \int_{s=-\infty}^{+\infty} (\prod_{j} e^{-e^{-(s+V_{ni}-V_{nj})}}) e^{-s} ds = \int_{s=-\infty}^{+\infty} \exp(-\sum_{j} e^{-(s+V_{ni}-V_{nj})}) e^{-s} ds = \\ &= \int_{s=-\infty}^{+\infty} \exp(-e^{-s} \sum_{j} e^{-(V_{ni}-V_{nj})}) e^{-s} ds \end{aligned} \tag{B7}$$

Define $t = \exp(-s)$ such that $-\exp(-s)ds = dt$. Note that as s approaches infinity then t approaches zero. Furthermore, as s approaches negative infinity, t becomes large. Using that,

$$\begin{aligned} P_{ni} &= \int_{+\infty}^{0} \exp(-t\sum_{j} e^{-(V_{ni}-V_{nj})})(-dt) = \int_{0}^{+\infty} \exp(-t\sum_{j} e^{-(V_{ni}-V_{nj})}) dt = \\ &= \frac{\exp(-t\sum_{j} e^{-(V_{ni}-V_{nj})})}{-\sum_{j} e^{-(V_{ni}-V_{nj})})} \Big|_{0}^{\infty} = \frac{1}{\sum_{j} e^{-(V_{ni}-V_{nj})})} = \frac{e^{V_{ni}}}{\sum_{j} e^{V_{nj}}} \end{aligned} \tag{B8}$$

As the first parameters of unobserved utility (V_{ni} and V_{nj}) are known, they can be specified as a linear in parameters. We follow Brock and Hommes (1998), Westerhoff (2008), and Pellizzari and Westerhoff (2009), and define fitness function (U_k^{st}) and the intensity of switching parameter (ξ) as our parameters for the specification of the representative utilities. Then we can obtain the relative weights of each trading strategy as:

$$V_{nj} = \xi \times U_t^{st} \tag{B9}$$

$$\varphi_t^{ag} = (\frac{e^{\xi \times U_t^{ag}}}{\sum e^{\xi \times U_t^{st}}}) \tag{B10}$$

APPENDIX C. List of the sample stocks

ISIN CODE	RIC	Security name
US0378331005	AAPL.OQ	Apple Inc.
US03073E1055	ABC.N	AmerisourceBergen Corp.
IE00B4BNMY34	ACN.N	Accenture plc
US0530151036	ADP.OQ	Automatic Data Processing Inc.
US0236081024	AEE.N	Ameren Corp.
US0015471081	AKS.N	AK Steel Holding Corp.
US0200021014	ALL.N	Allstate Corp.
US0231351067	AMZN.OQ	Amazon.com Inc.
US0325111070	APC.N	Anadarko Petroleum Corp.
US1101221083	BMY.N	Bristol-Myers Squibb Co.
US0846707026	BRKb.N	Berkshire Hathaway Inc.
US2058871029	CAG.N	ConAgra Brands Inc.
US1491231015	CAT.N	Caterpillar Inc.
US1651671075	CHK.N	Chesapeake Energy Corp.
US1567001060	CTL.N	CenturyLink Inc.
US1667641005	CVX.N	Chevron Corp.
US2635341090	DD.N	E I du Pont de Nemours and Co.
US2479162081	DNR.N	Denbury Resources
US2605431038	DOW.N	Dow Chemical Co.
US2786421030	EBAY.OQ	eBay Inc.
US2686481027	EMC.N	EMC Corp.
US30219G1085	ESRX.OQ	Express Scripts Holding Co.
US2971781057	ESS.N	Essex Property Trust Inc.
US3453708600	F.N	Ford Motor Co.
US3696041033	GE.N	General Electric Co.
US38259P7069	GOOG.OQ	Alphabet Inc. (Google Inc. Class C)
US4370761029	HD.N	Home Depot Inc.
US4282361033	HPQ.N	Hewlett-Packard Inc.
US4592001014	IBM.N	International Business Machines Corp.
US4581401001	INTC.OQ	Intel Corp.
US9255501051	JDSU.OQ	JDS Uniphase Corp.
US4781601046	JNJ.N	Johnson & Johnson
US1912161007	KO.N	The Coca Cola Co.
US5260571048	LEN.N	Lennar Corp.
US58155Q1031	MCK.N	McKesson Corp.
IE00BTN1Y115	MDT.N	Medtronic Plc.
US88579Y1010	MMM.N	3M Co.
US02209S1033	MO.N	Altria Group Inc.
US5949181045	MSFT.OQ	Microsoft Corp.
US68389X1054	ORCL.OQ	Oracle Corp.
US7134481081	PEP.N	PepsiCo Inc.
US7170811035	PFE.N	Pfizer Inc.
US7427181091	PG.N	Procter & Gamble Co.
US7181721090	PM.N	Philip Morris International Inc.
US7132911022	POM.N	Pepco Holdings Inc.
US8454671095	SWN.N	Southwestern Energy Co.
US8835561023	TMO.N	Thermo Fisher Scientific Inc.
US8825081040	TXN.N	Texas Instruments Inc.
US91324P1021	UNH.N	United Health Group Inc.
US9497461015	WFC.N	Wells Fargo & Co.
US9311421039	WMT.N	Wal-Mart Stores Inc.
US30231G1022	XOM.N	Exxon Mobil Corp.
US9884981013	YUM.N	Yum! Brands Inc.