# Trade Time Clustering

Jeffrey R. Black[1], Pankaj K. Jain[2], and Wei Sun[3]

**Abstract**

This paper introduces a new measure, time clustering, which uses relative volumes and trade durations to measure the concentration of trading in the time dimension. We then examine the determinants, associations, and effects of time clustering during both stable and volatile market conditions. Using intraday analysis, we find that both high information flow and low trading cost lead to time clustering in a stable market while only low trading cost is the dominated driver of time clustering in a volatile market. We also show that buy-side liquidity has stronger effects on time clustering in a stable market while sell-side liquidity has stronger effects on time clustering in a volatile market. We confirm that time clustering is associated with high price impact and high price volatility in both stable and volatile markets, and further we show that time clustering is associated with low variance ratio, suggesting that trading concertation contributes to price discovery and improves market efficiency. We also find that the effects of increased time clustering include higher ISO proportion, lower Odd-lot proportion, and larger number of trading exchanges, suggesting that informed traders trade more aggressively and split orders by exchange instead of by size as time clustering increases. Additionally, we demonstrate that in stable market time clustering reduces subsequent HFT.

JEL Classification: G12

Keywords: Trade time clustering; Time clustering pattern; Trading cost

[1] Assistant Professor, Fogelman College of Business & Economics, University of Memphis, jrblack@memphis.edu

[2] Professor of Finance, Fogelman College of Business & Economics, University of Memphis, pjain@memphis.edu

[3] Ph.D. candidate, Fogelman College of Business & Economics, University of Memphis, wsun1@memphis.edu

# Trade Time Clustering

**Abstract**

This paper introduces a new measure, time clustering, which uses relative volumes and trade durations to measure the concentration of trading in the time dimension. We then examine the determinants, associations, and effects of time clustering during both stable and volatile market conditions. Using intraday analysis, we find that both high information flow and low trading cost lead to time clustering in a stable market while only low trading cost is the dominated driver of time clustering in a volatile market. We also show that buy-side liquidity has stronger effects on time clustering in a stable market while sell-side liquidity has stronger effects on time clustering in a volatile market. We confirm that time clustering is associated with high price impact and high price volatility in both stable and volatile markets, and further we show that time clustering is associated with low variance ratio, suggesting that trading concertation contributes to price discovery and improves market efficiency. We also find that the effects of increased time clustering include higher ISO proportion, lower Odd-lot proportion, and larger number of trading exchanges, suggesting that informed traders trade more aggressively and split orders by exchange instead of by size as time clustering increases. Additionally, we demonstrate that in stable market time clustering reduces subsequent HFT.

JEL Classification: G12

Keywords: Trade time clustering; Time clustering pattern; Trading cost

## 1. Introduction

Three factors identify the processing of a trade: quantity, price, and time. Several studies document trade clustering in two of those dimensions, namely size and price clustering. Among others, Moulton (2005) provides an explanation for the time variation of trade-size clustering, Anand and Chakaravarty (2007) examine how size clustering affects price discovery in the options markets, and Alexander and Peterson (2007) argue that informed traders prefer trading with medium-sized transactions. Meanwhile, other studies examine price clustering. For example, both Ohta (2006) and Ascioglu, Comerton-Forde, and McInish (2007) investigate price clustering on the Tokyo Stock Exchange, showing that prices cluster in an intraday pattern. However, time clustering, or the concentration of volume in the time dimension, has yet to be measured or examined in the literature. In this paper we develop a new measure based on both trade volume and the time between transactions to examine how trades cluster in time, the determinants of that clustering, the associations, and the effects thereof.

Clustering is the process of grouping a collection of objects in the same "cluster" that are closer to each other than those in other clusters. Unlike clustering in a physical space that has three dimensions, time clustering of trades considers just one dimension. Thus, in this paper, time between two trades is employed as the key determinant of grouping trades into a cluster. Because we aim to measure the clustering of unique trades, rather than simply the volume of shares traded, it is critical to control for the volume of each security. Therefore, using the percentage volume of each trade divided by the time between trades, we construct an appropriate measure of trade clustering. Our time clustering measure can be used on multiple levels. At the individual trade level, it indicates how concentrated a trade is in the time dimension by taking into account both relative volume and time between trades. This measure can also be aggregated to describe the level of trade clustering over time intervals. For example, we aggregate clustering for each 5-minute interval by calculating the root of the sum squared trade-level time clustering measures for our intraday analysis.

We examine the determinants of time clustering using 5-minute interval aggregation. We find that in stable, non-volatile markets time clustering is positively associated with lagged order imbalance, lagged buy- and sell-side depth as well as lagged trading cost. This implies that information is a key driver of time clustering in a stable market. We further find that in a stable market the lagged buy-side liquidity has larger impact on time clustering than lagged sell-side liquidity. We also investigate the determinants in a volatile market, and find that time clustering is positively associated with lagged sell-side depth and negatively associated with lagged trading cost, suggesting that in a volatile market trades are concentrated in periods of greater sell-side liquidity, including more depth and lower trading cost. We also find that the effect of lagged quoted spread on time clustering is significantly larger than the effect of lagged effective spread, suggesting that in volatile market trading concertation is more likely to be caused by small or noise investors who have less negotiation power (see Bessembinder and Kaufman, 1997, for the discussion on the difference between quoted spread and effective spread).

According to stealth theory, informed traders attempt to hide their information while minimizing the trading cost at the same time by splitting large trades into medium trades (Kyle, 1985; Barclay and Warner, 1993). Both Alexander and Peterson (2007) and Anand and Chakravarty (2007) show evidence that size clustering varies over time and its effects become stronger when trading volume is higher. However, usually due to the short life of information, informed traders are willing to trade their information as soon as possible. Chiyachantana and Jain (2009) find that institutions choose to trade aggressively and not to divide their orders while being informed. The tradeoff between trading speed and transaction costs is a key element of trading, but the empirical analysis of that tradeoff remains inconclusive. This paper sheds light on that tradeoff by identifying different drivers of time clustering in different market conditions, helping us understand why trading becomes more or less concentrated in certain periods.

Next, we examine the associations of trade time clustering. We find that in both stable and volatile markets, time clustering is positively related with price impact and price volatility, but negatively related with variance ratio, the measure of price inefficiency. Previously, the literature has shown that the timing of a trade affects many factors of market microstructure. For instance, Harris (1991) establishes that price clustering is related to transaction frequency, showing that the two are negatively related. Huang and Masulis (2003) examine the relation between trading activity and stock price volatility on the London Stock Exchange and show that trade frequency can positively affect price volatility. We confirm that time clustering is associated with high price impact and high price volatility in both stable and volatile markets, and further we show that time clustering is associated with low variance ratio, suggesting that trading concertation contributes to price discovery and improves market efficiency.

When examining the effects of time clustering, we find that higher levels of time clustering are followed by higher ISO (Intermarket Sweep Orders) trades proportion, lower Odd-lot trades proportion, and increased number of exchanges under both stable and volatile market conditions, indicating a relation between trading concentration and order splitting. This result also implies that trade clustering might be a signal of informed institutions trading as the subsequent ISO orders increases (Chakravarty, Jain, Upson, and Wood, 2012). Additionally, we show that when the market is stable time clustering is negatively related with lead order cancellation rate, implying that trading concentration reduce subsequent HFT behaviors. This result is consistent with the argument of Easley, Lopez de Prado, and O’Hara (2012) that informed trading can be toxic to HF market makers.

Asymmetric information theory argues that trades convey information. Dufour and Engle (2000) support this argument by providing evidence that trading intensity is associated with the existence of informed traders. They also describe the role of time in the process of price formation and liquidity by showing the positive relation of trading intensity with spreads and price impact. This paper extends Dufour and Engle’s results in three directions. First, instead of using trade duration we construct a new measure, time clustering, which captures the concentration of trading activity; second, we investigate determinants, associations, and effects of time clustering using intraday analysis; and third, we examine the new measure under both stable and volatile market conditions to provide a deeper view of the process of price information.

The remainder of the paper is organized as follows. Section 2 presents relevant literature and the development of hypotheses. Section 3 describes the construction of our new measure, methodology and data used in the analysis. Section 4 documents the empirical results. Finally, Section 5 contains our summarization and conclusions.

## 2. Literature review and hypotheses development

### *2.1 Intraday Pattern of time clustering*

There are lots of research investigating the intraday pattern of liquidity. For example, McInish and Wood (1992) show that bid-ask spreads have a reverse "J" shape of intraday, suggesting that bid-ask spread is at highest level while the market opening, then decreases and reaches the second highest point while the market closing. Chung, Van Ness and Van Ness (1999) show a "U" shape pattern of the intraday bid-ask spread. However, Upson and Van Ness (2016) state that since 2011 the intraday pattern of bid-ask spread has changed to an "S" shape, suggesting that the highest bid-ask spread exists at the market opening and the lowest bid-ask spread exists at the market closing. They argue that the low bid-ask spread showing at the market closing is due to the extensive trades made by traders, especially HFT traders who want to keep a zero or close to zero inventory position at the end of the day. The dramatical changes in liquidity during marketing opening and closing are usually associated with aggregative trading activities, indicating a high level of trading concentration. Thus, our first hypothesis is as below.

***Hypothesis I:*** *The intraday pattern of time clustering is U-shape.*

*2.2. Determinants of time clustering*

Time clustering of trade might vary widely in different conditions due to the various strategies applied by different market participants. Dufour and Engle (2000) show that market is more active when information arrives, and also suggest that informed traders only trade on information while uninformed traders' trade is independent to information. Lee (1992) states that informed traders use market orders to get immediate execution to realize their information. All these informed traders' strategies might increase the level of trade clustering while information coming out. Therefore, our hypothesis II is as follows.

***Hypothesis II:*** *Time clustering is positively related with information flow.*

On the other hand, informed investors intend to hide the information and lower the trading cost by trading at the time while market liquidity is high. Admati and Pfleiderer (1988) theoretically show that informed traders tend to trade during times of heavy trading to hide their trades. To lower the trading cost, institution investors also have incentive to split their orders. Keim and Madhavan (1995) show that institutional traders split the orders and fill them over time to reduce their trading cost. Chiyachantana, Jain, Jiang, and Wood, (2004) support this argument by showing that institutional traders fill the orders over days. Heston and Sadka (2008) find that trading volume has seasonal patterns and abnormal returns from seasonal trading strategy is related with transaction cost. All these findings suggest that some certain trading strategies target at the time when market liquidity is high and trading cost is low. Therefore, while considering the relation between liquidity and trade time clustering, the following hypothesis is expected to hold.

***Hypothesis III-A:*** *Time clustering is positively related with order depth.*

***Hypothesis III-B:*** *Time clustering is negatively related with trading cost*

*2.3. Time clustering and trade processing*

According to asymmetric information theory, trades themselves convey information and thus as liquidity supplier market makers adjust the price to reflect the changes in the fundamental value of stocks while learning the information from informed trades. Chiyachantana, Jain, Jiang and Wood (2004) show that price impact is highly related with block (institutional) transactions, suggesting a positive relation between price impact and trading concentration. Dufour and Engle (2000) provide another evidence on supporting this relation by stating that price impact increases with the decline of time duration of transactions. Hence, our fourth hypothesis is that time clustering is positively associated with higher price impact, and thus contributes to price discovery, Additionally, time clustering improves market efficiency.

***Hypothesis IV-A:*** *Time clustering is positively associated with price impact.*

***Hypothesis IV-B:*** *Time clustering is positively associated with market efficiency.*

Volatility is proven to be related with trade frequency (Huang and Masulis, 2003). Some studies also find positive and significant relation between trading volume and price volatility. Using Nasdaq data, Jones, Kaul, and Lipson (1994) argue that number of trades beyond size of trades explain most of the volatility-volume relation, while Chan and Fong (2000) find that price volatility is more related to trades size rather than number of trades. However, the relation between price volatility and trading volume is confirmed. As we construct our measure of time clustering mainly based on trading volume and frequency, we expect there is a strong and positive relation between price volatility and time clustering. This expectation leads to our fifth hypothesis.

***Hypothesis V:*** *Time clustering is positively associated with price volatility.*

*2.4. Effects of time clustering*

Prior research on trade size document that informed investors tend to split their orders to hide the information from the market. Alexander and Peterson (2007) provide evidence on stealth trading by showing that the price impact of medium-sized rounded trades is greater than large rounded trades. O'Hara, Yao and Ye (2014) also argue that informed traders use odd-lot trades to avoid detection by showing that odd-lot trades dominate a large proportion of trades and contribute a lot to price discovery. On the other hand, due to the restriction of the time of information informed traders usually have incentive to realize the profit of the information quickly by using aggressive trading strategy. Chakravarty, Jain, Upson and Wood (2012) find that informed traders tend to use Intermarket Sweep Orders (ISO), which are designed to provide more efficient and faster execution. While considering time clustering as an indicator of informed trading, the concentration level of trading reveals the level of information content of trades and thus affects subsequent informed trading strategy selection.

***Hypothesis VI:*** *Time clustering affects subsequent informed trading strategy.*

Beside the influence on the trading strategy of institution investors and informed traders, who are considered as liquidity consumers, time clustering could also affect the strategy selection of liquidity providers, market makers. Based on volume instead of time, Easley, Lopez de Prado, and O'Hara (2012) introduce the Volume-synchronized Probability of Informed Trading (VPIN) to estimate flow toxicity and shed a light on the relation between informed transaction volume and High Frequency Trading (HFT). They state that the ability of High Frequency (HF) market makers to control their position risk depends on their strategy against informed (adverse selection) trading on their passive order. They also state that sometimes the excessive informed trading (toxicity) even forces the HF market maker to shut down their operations. Hence, our seventh hypothesis is as follows.

***Hypothesis VII:*** *Time clustering reduces HFT.*

## 3. Methodology and Data

In this section, we describe the construction of our time clustering measure, the statistical models of tests, the methodology of sample selection, and the characteristics of stocks and market conditions in our sample.

### *3.1.Measure of trade time clustering*

We use transaction data from Daily TAQ (Trade and Quote) to construct our trade time clustering measure. First, we identify time distance for each trade as the difference between time of this trade and time of previous trade.

$$D_{it} = T_t - T_{t-1} \tag{1}$$

In other word, if the time distance of a trade is small, it should be considered as being clustered with the previous trade. Then we compute the percentage volume of a trade as the volume of this trade divided by the total daily volume of this stock.

$$\mathrm{PV_{it}} = \frac{V_{\mathrm{it}}}{\sum_{d} \mathrm{V_{it}}} \tag{2}$$

For each trade, we calculate the trade-level time clustering measure as the percentage volume of a trade dividend by its time distance.

$$\mathrm{TC_{it}} = \frac{\mathrm{PV_{it}}}{\mathrm{D_{it}}} \tag{3}$$

Finally, we construct an aggregated time clustering measure for each 5-minute interval for each stock as the root of the sum of squared trade time clustering.

$$\mathrm{TC_{i5m}} = \sqrt{\sum_{5m} \mathrm{TC_{it}}^2} \tag{4}$$

We apply the sum of squared trade time clustering for aggregating so that the time clustering measure spikes when trade volume is grouped together over time.

*Insert Figure 1 about here*

Figure 1 shows an example of time clustering explaining that how the aggregated time clustering is calculated and how it can demonstrate the differences among several trading scenarios. The total trading daily volume in this example is 3,000 shares (except the scenario in Chart D which has total 5,000 shares), the first trades exit at 10:05 AM for the scenario in Chart A and 10:01 for the scenarios in other Charts. The following trades happen between 10:01 AM and 10:30 AM. To make it more visible and easier to understand, the time clustering measure in this sample is aggregated in 30-minute interval for period of 10:00 - 10:30, using the same methodology as our 5-minute aggregation. Chart A shows that during 10:05 am to 10:30 am there are 6 trades with the same distance of 5 minutes. Trades of 1, 3, and 5 have same volume of 999 share, while trades of 2, 4 and 6 have same volume of 1 share. Chart B shows that during 10:01 am to 10:03 am there are 3 trades with same volume of 1,000 shares and same distance of 1 minutes. Chart C shows that during 10:01 am to 10:05 am there are 5 trades with same volume of 600 shares and same distance of 1 minutes. Chart D shows a similar analogous scenario like Chart C, but the volume of each trade increases to 1,000 shares. For these four scenarios, the aggregated time clustering measures are 0.0019, 0.0096, 0.0075, and 0.0075 respectively. The most concentrated trading scenario, presented in Chart B, has the highest time clustering. Even though scenario in Chart D has larger trading volume, it has same time clustering as scenario in Chart C due to the same concentrated level (the same percentage volume and the same time distance of each trade). Chart E shows a scenario that has a "jump" trade during the trading period. Trades of 1, 2, 4, and 5 have same volume of 500 shares while trade 3 jumps to 1,000 shares. Compared with scenario in Chart C, which has time clustering of 0.0075, this scenario in Chart E has time clustering of 0.0079, indicating a higher level of concentration.

This example clearly shows that the measure of time clustering could capture authentic changes in both trades size and trades frequency by presenting that with the same trading volume, time clustering increases as the trading frequency increases, and that while frequency holds still, equalized increase in trading volume might not affect clustering level. Also, it suggests that this time clustering measure could capture the variance in trades size by showing that time clustering increases with a “jumpping” trade, which reflects more concentrated trading activity. Eventually, the comparisons among these five Charts demonstrate that our measure of time clustering is a more powerful proxy to reflect the diversity of trading activities than trade size or trade frequency alone because of its combined construction.

*3.2.Models*

As we aggregate time clustering in each 5-minute interval, we introduce a dummy variable for each of these 5-minute intervals to test the intraday pattern of time clustering. We also introduce dummy variables for weekdays to control the intra-week effects. Then we regress time clustering on the dummy variables of 5-minute intervals and weekdays. The regression model is shown as Model 5 below. To avoid perfect multicollinearity, we omit the first 5-minute interval of day and the first day of week.

$$TimeClustering_{it} = \alpha + \sum_{t=1}^{77} \beta_{1\mathrm{t}} 5minute_{it} + \sum_{t=1}^{4} \beta_{2\mathrm{n}} Weekday_{itn} + \varepsilon_{it} \quad (5)$$

To examine the intraday determinants, associations, and effects of time clustering, we construct the tests through timeline. Figure 2 shows the process flaw diagram of these tests. First, we regress time clustering at the time interval of t on the informed trading measures and liquidity measures at the previous time interval (t-1) to determine the drivers of trading concentration. Then, we use time clustering at the time interval of t as independent variable to examine the relation between time clustering and the market measures, such as price impact, price efficiency, and price volatility, at the same time interval (t). Finally, we test the effects of time clustering at the time interval of t on liquidity, order splitting, and HFT behaviors at the next time interval (t+1). Models of 6, 7 and 8 represent the regression equations used to test the determinants, associations, and effects of time clustering, respectively.

*Insert Figure 2 about here*

In Model 6, we regression time clustering at time interval of t on the determinants at the time interval of t-1. The determinants include information flow, liquidity and trading cost. Chan and Fong (2000) find that order imbalance plays an import role to explain daily price movements, suggesting that order imbalance could be considered as an indicator of informed trades. Brown, Walsh and Yuen (1997) and Chordia and Subrahmanyam (2004), who investigate the relation between imbalance and stock returns, also provide empirical evidences to this argument. Additionally, Chan and Lakonishok (1993) argue that buy orders contain more information shares than sell orders because institutional investors cannot sell the stocks that they do not own. Therefore, the time clustering of trades should reflect the information imbalance of buy-sell orders. We use *Order Imbalance*, which is the absolute value of the share volume difference between the National Best Bid order (NBB) and National Best offer order (NBO), as a proxy of informed trades. We compute the order imbalance for each NBBO and aggregate the time-weighted order imbalances for each 5-minute interval during market operating hours.

$$TimeClustering_{it} = \alpha + \beta_1 Determinant_{i(t-1)} + \beta_2 TimeClustering_{i(t-1)} + \sum \beta_{3n} ControlVariables_{itn} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it} \quad (6)$$

Using model 6, we also examine whether liquidity and trading cost lead to trading concentration, which is measure by time clustering. We follow the methods of Holden and Jacobsen (2014) to construct the liquidity and trading cost measures. To measure stock liquidity, we use both *NBB Depth* and *NBO Depth*, which are share volumes aggregated in 5-minute interval using time-weighting scheme. We use percent *Quoted Spread*, which is computed as the difference between the nature logarithms value of National Best Offer (NBO) price and the nature logarithms value of National Best Bid (NBB) price, to measure the observed trading cost. This measure is also aggregated in 5-minute interval using time-weighting scheme. We use percent *Effective Spread*, which is calculated as the difference between the nature logarithms value of transaction price and the nature logarithms value of the average of NBO and NBB prices, to measure the real transaction cost. For each 5-minute interval, the effective spreads are aggregated using share volume weighting scheme.

To control the autocorrelation effects, we include *Time Clustering* in the previous time interval in Model 6. Since the measure of time clustering is constructed from percentage daily volume capturing the abnormalities in trades size, we add *Size Volatility* as a control variable in the regressions. The variable of *Size Volatility* is computed as the standard deviation of trades size in each 5-minute interval. The control variables also include *Total Share Volume*, which is calculated as the total transaction share volume in each 5-minute interval. We use the characteristics of the exchange-traded fund (ETF) of SPY, including time clustering of SPY, share volume of SPY, and price volatility of SPY, to control the effects from the market. Price volatility of SPY is computed as the price range of SPY in a 5-minute interval divided share volume-weighted average price of SPY in the same interval. To apply the intraday analysis, we also control the fixed effects of stocks and trading days in the regression.

To examine the role of time clustering in the process of trades, we investigate the relation between time clustering and *Price Impact*, and the relation between time clustering market quality measures, including *Price Efficiency* and *Price Volatility*. Price impact is the permanent component of trading cost, representing the information content of trades. We use percent *Price Impact* to explore the relation between time clustering and the information content of trades. To compute the percent *Price Impact*, we calculate the average of NBB and NBO prices of a trade as $M_k$ and the average of NBB and NBO prices 5 minutes after the trade as $M_{(k+5)}$. Then we use the variable of buy-sell direction times the difference between the nature logarithms value of $M_{(k+5)}$ and the nature logarithms value of $M_k$ to get the percent *Price Impact* for each trade. For each 5-minute interval, the price impacts are aggregated using share volume weighting scheme. *Price Volatility* is calculated as the price range in a 5-minute interval divided by the share volume weighted average price in the same interval. To measure *Price Efficiency*, we follow Lo and MacKinlay (1988) to generate 5-mintute interval variance ratio. We calculate the variances of trade returns for each 5-minute interval and 1-minute interval, respectively, and then calculate a variance ratio using the 5-minute interval variance divided by the aggregated 1-minute interval variances. According to the efficient market theory, this variance ratio should be equal to 1. Thus, we use the absolute nature logarithms value of the variance ratio to indicate price inefficiency.

$$Association_{it} = \alpha + \beta_1 TimeClustering_{it} + \sum \beta_{2n} ControlVariables_{itn} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it} \quad (7)$$

In Model 7, *Price Impact*, *Price Volatility* and *Price Efficiency* at the time interval of t are regressed on *Time Clustering* at the same time interval, respectively. The control variables include *Size Volatility*, *Total Share Volume* and the proxies of the characteristics of SPY. The fixed effects of stocks and trading days are also controlled in Model 7.

We use the regression in Model 8 to investigate the effects of time clustering at the time interval of t on order splitting and HFT behaviors at the time interval of t+1. Since Intermarket Sweep Order (ISO) is the trading approach that provides more efficient and faster execution (Chakravarty, Jain, Upson and Wood, 2012), we introduce *ISO trade proportion*, which is computed as ISO trades share volume divided by the total transaction share volume in each 5-minute interval, as a proxy for anti-order-splitting. O'Hara, Yao and Ye (2014) state that informed traders use Odd-lot trades to split their orders and hide their information. Therefore, we calculate *Odd-lot trade proportion* by dividing Odd-lots trades share volume with the total transaction share volume in each 5-minute interval as another measure of order splitting. Another strategy of splitting orders is to break up orders by exchange. We introduce *Number of Exchanges* to represent this splitting strategy.

$$Affected_{i(t+1)} = \alpha + \beta_1 TimeClustering_{it} + \sum \beta_{2n} ControlVariables_{i(t+1)n} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{i(t+1)} \quad (8)$$

Prior studies demonstrate that limiter order cancellation rapidly increases over recent years because the High Frequency traders use it to either check the market condition or probe orders of opposite-side traders (Hasbrouck and Saar, 2013; O'Hara, 2015; Van Ness, Van Ness, and Watson, 2015). Additionally, while investigating the relation between HFT and market liquidity, some studies explore another trading strategy, which is limit order modification, used by high frequency traders (Conrad, Wahal, and Xiang, 2015; Nikolsko-Rzhevska and Nikolsko-Rzhevska, working paper). Following these research, we measure HFT behaviors using *Cancellation Rate*, which is computed as the total number of cancelled orders divided by the total number of executed orders in each 5-minute interval, and *Replacement Rate*, which is computed as the total number of replaced orders divided by the total number of executed orders in each 5-minute interval.

In Model 8, we regress the affected variables, including measures of liquidity, order splitting, and HFT behaviors, at the time interval of t+1 on *Time Clustering* at the time interval of t to examine the effects of time clustering. The control variables include *Size Volatility*, *Total Volume* and the proxies of the characteristics of SPY. The fixed effects of stocks and trading days are also controlled.

### *3.3. Data and Sample Description*

Chiyachantana, Jain, Jiang and Wood (2004) find that the price impact of institutional (block) trades varies under different market conditions. The price impact is higher for of institutional buy orders in bullish market while it is higher for institutional sell orders in bearish market. They argue that these institutional trades are on the same side of market and liquidity demanders, and thus have higher trading cost. Trade time clustering might have the same pattern as price impact of institutional trades because time clustering is highly related with institutional trading. Therefore, following Chiyachantana, Jain, Jiang and Wood (2004), we examine trade time clustering under both stable and volatile market conditions to investigate the underlying drivers and the potential effects. The first period that we select as stable market is December $1^{st}$ – $15^{th}$ of 2017. This period contains 10 trading days and 780 5-minute trading intervals. Then we select February $1^{st}$ – $15^{th}$ of 2018 as the period of volatile market, which also has 10 trading days and total 780 5-minute trading intervals.

To contract the sample of time clustering, we follow the method introduced by Chakravarty, Jain, Upson and Wood (2012). First, we select only regular stocks with CRSP share code of 10 or 11, and then sort them by their market capitulations on November $30^{th}$ 2017 into 3 groups, big, medium, and small, respectively. Finally, we form the sample of 600 stocks by selecting the top 200 stocks in each of the three groups.

Figure 3 presents the 5-miunte interval time clusterings in both years of 2017 and 2018. Figure 3 shows that time clustering is high and close to 400 in the first 5-minute interval while market is opening, then rapidly drops under 200 in the second 5-minute interval, and starts to climb in the 73th 5-minute interval (3:30 PM – 3:35 PM) till reaching the highest point above 800 in the last 5-minute interval while market is closing. This intraday “J” shape pattern of time clustering is apparent in both years of 2017 and 2018. Figure 3 also shows that time clustering in 2017 is slightly higher than in 2018 except in the last two 5-minute intervals. To eliminate the dramatic changes in time clustering and the effects of market opening and closing, we cut the sample by the two big gaps shown in Figure 3, which are the gap between the first and the second 5-minute intervals and the gap between the 76th and the 77th 5-minute intervals. Therefore, in the following analysis we focus on the intraday time period of 9:35 AM – 15:50 PM, which is from the second 5-minute interval to the 76th.

*Insert Figure 3 about here*

Table 1 reports the summary statistics of the sample during the market time of 9:35 AM – 15:50 PM, which excludes the market open-close period. Columns 1 and 2 in Table 1 present the summary statistics of all the variables for subsamples of December 2017 and February 2018, respectively. Column 3 presents the summary statistics for the whole sample. The last Column shows the differences in the variable means of subsamples between 2017 and 2018. The average 5-minute *SPY Price Volatility* is only 12.2 basis points (BPS) in December 2017, but it increases by 44.5 BPS (about 360%) in February 2018, indicating that the market is much more volatile in the second time period. The stock price volatility, which is the price range divided by volume-weighted average price in each 5-minute interval, is about 12.4 BPS or 44% higher in 2018 than in 2017, being consistent with the pattern of SPY price volatility and indicating a volatile market of 2018.

*Insert Table 1 about here*

Besides volatility measures, Table 1 also presents the summary statistics for other measures of market characteristics. Total share transaction volume, trading costs, price impact, ISO proportion, number of exchanges, HFT activities are all significant higher in the volatile period of February 2018 than in the stable period of December 2017. However, the increased trading activities do not lead to an increase in time clustering. Time clustering significantly drops about 7% in 2018 compared to in 2017. One possible reason for the decrease in time clustering is that noise traders, who employ small size orders, become more active during the volatile period. Additionally, NBO depths significantly drop more than 10% from 2017 to 2018, suggesting that liquidity providers try to control their trading risk by reducing order size during the volatile period.

Table 2 presents the Pearson correlations among the market measures. The highest correlation for *Time Clustering* is only 8.1% with *Effective Spread*. Although the time clustering measure is generated from percentage share volume, its correlations with *Total Share Volume* and *Size Volatility* are only -0.8% and 1.2%, respectively, indicating that the construction of *Time Clustering* successfully dissimilate it to trades volume measures. The correlations among *NBB Depth*, *NBO Depth*, and *Imbalance Depth* are all about or above 68%. The two trading cost measures, *Quoted Spread* and *Effective Spread* also have a high correlation with a coefficient of 60%. Additionally, the two measures are positively correlated to *Price Volatility* having coefficients of 66.5% and 54.9%, respectively, suggesting that volatile market is associated with high trading cost.

*Insert Table 2 about here*

## 4. Results

### *4.1. The Intraday Pattern of Time Clustering*

Several prior research show that liquidity presents a certain intraday pattern (McInish and Wood, 1992; Chung, Van Ness and Van Ness, 1999; Upson and Van Ness, 2016), suggesting that trading activities have calendar time pattern in a trading day. To examine the intraday time pattern of time clustering, we assign dummy variables for all the 78 5-minute intervals during market operating. To test the weekly pattern, we also introduce dummy variables for the 5 weekdays. As comparison, we omit the dummy variable of the first 5-minute interval and the dummy variable of the first day of week. Additionally, we add a dummy variable to indicate the subsample in February 2018 to investigate the difference between stable market and volatile market.

Using the sample during the full market operating hours, we report the regression result generated by Model 5 in Table 3, and show that all the coefficients of intraday 5-minute intervals are significantly negative except the last two, which are significantly positive, indicating that time clustering in the first 5-minute interval is significant higher than in the rest 5-minute intervals except the last two. This result is consistent with the shape shown in Figure 3, confirming the intraday "J" shape pattern of time clustering. Table 3 also shows that compared to the rest days of a week, both Wednesday and Thursday have positive and bigger coefficients, although Wednesday's coefficient is not significant. This result suggests a reverse "U" shape of weekly pattern of time clustering.

*Insert Table 3 about here*

*4.2. Determinants of Time Clustering*

Trading concentration is driven by either incoming information flow (Dufour and Engle, 2000) or high liquidity with low trading cost, which attracts institutional and informed traders. Lee (1992) argues that informed traders have incentive to use market order to implement the information quickly, while Admati and Pleiderer (1988) shows that informed traders tend to hide their information by placing orders at the time when the liquidity is abundant. Additionally, some research (such as Keim and Madhave, 1995; Chiyachantana, et al., 2004; Heston and Sadka, 2008) state that intuitional trading activities are highly related with trading cost. To determine the market characteristics that may drive trade time clustering, we employ the intraday analysis with fixed effects of both stock and date being controlled, and run the OLS regression of Model 6 for the measure of information flow, the measures of liquidity, and the measures of trading cost, respectively.

Table 4 reports the results of the regressions of time clustering on order imbalance, which is the measure of information flow. Columns 1-3 present the results for the 2017 subsample, the 2018 subsample, and the whole sample, respectively. The relation between lagged order imbalance in the 5-minute interval of t-1 and time clustering in the 5-minute interval of t is positive and statistically significant for stable market (2017) subsmaple, suggesting that information flow represented by order imbalance lead to higher time clustering. However, this relation becomes weaker and insignificant in volatile market. To investigate the difference in the coefficients between stable and volatile markets, we introduce a dummy variable named as *Volatile* to indicate the volatile market (2018), and the interaction between *Volatile* and lagged order imbalance. The coefficient of the interaction term is -0.184, which is statistically significant, indicating that the relation between time clustering and information flow is weaker in volatile market than in stable market.

*Insert Table 4 about here*

Table 5 reports the regression results of time clustering on both liquidity measures and trading cost measures. Columns 1-4 present the results for the subsample of 2017, Columns 5-8 present the results for the subsample of 2018, and Columns 9-12 present the results for the whole sample. We use the National Best Bid (NBB) depth and National Best Offer (NBO) depth to represent the liquidity of buy and sell sides, respectively. The coefficient of NBB and NBO depths are all positive, and only the coefficient of NBB for 2018 subsample is insignificant. This significantly positive relation between order depths and time clustering suggests that liquidity is one of the determinants of time clustering. The results of the whole sample show that the coefficients of the interactions between volatile market and order depths are significantly negative, indicating that the effects of liquidity on time clustering drop in volatile market. Beside the significant drop of effects from 2017 to 2018, the differences in the coefficients between NBB depth and NBO depth is also noticeable. Further, we employ the method introduced by Clogg, Petkova, and Haritou (1995) to examine the equality of the effects between NBB depth and NBO depth on time clustering, and confirm that the effects are significantly different for both 2017 and 2018 subsamples. However, the differences in the effects of liquidity on time clustering are in two opposite directions. For the 2017 subsample the effect of NBB depth is larger than NBO depth, while for the 2018 subsample the effect of NBO depth is larger than NBB depth. This result implies that trading concentration is more likely to be driven by buy-side liquidity in a stable market, while in a volatile market trading concentration has a stronger relation with sell-side liquidity.

*Insert Table 5 about here*

Table 5 also show the results of regressing time clustering on lagged trading cost measures, including percent quoted spread and percent effective spread. All the coefficients of trading cost measures are significantly negative except the coefficient of quoted spread for the 2017 subsample, which is negative but not significant, suggesting that lower trading cost will increase time clustering. Additionally, employing the cross-model test on the coefficients (Clogg, Petkova, and Haritou, 1995), we find that the effects between effective spread and quoted spread on time clustering are significantly different for 2018 subsample. This result suggests that in volatile market low quoted spread is more attractive than low effective spread. Petersen and Fialkowski (1994) document that quoted spread is not accurate as effective spread to measure the actual transaction cost. Bessembinder and Kaufman (1997) also state that effective spread captures the cost for trades either within or outside the quotes, suggesting that effective spread reflects the transaction cost of trade made by investors who have more negotiation power. Therefore, that in the volatile market quoted spread has larger effect on time clustering than effective spread implies that trading concertation in a volatile market is more likely to be caused by small or noise investors who have less negotiation power.

Based on previous results, we indicate the determinants of time clustering, including information flow, liquidity and trading cost. Further, we introduce a horse race test among these factors to examine if there is a dominating factor. To avoid the multicollinearity issue, we must drop *NBB depth* and *NBO depth*, because the information flow measure, *Order Imbalance*, is generated directly from NBB and NBB orders. Also, we separate *Quoted Spread* and *Effective Spread* in different models due to the same reason, as they are shown to have a high correlation in Table 3. Hence, we test two horse race models, including one with *Order Imbalance* and *Quoted Spread* and the other one with *Order Imbalance* and *Effective Spread*. Columns 1-2 in Table 6 present the results of the test for 2017 subsample, showing that the coefficients of *Order Imbalance* are positively significant in both the two models and that the coefficients of *Quoted Spread* is insignificant in the first model but the coefficients of *Effective Spread* is significantly negative in the second model. This result is consistent with that of Table 5, suggesting that in a stable market time clustering is driven by both information flow and trading cost measured by effective spread.

*Insert Table 6 about here*

Columns 3-4 in Table 6 reports the result of horse race test for 2018 subsample, showing that the coefficients of *Order Imbalance* are insignificant in both the two models and that both the coefficients of *Quoted Spread* and *Effective Spread* are negatively significant. Additionally, Columns 5-6 present negative and significant coefficients of the interaction between *Volatile* and *Order Imbalance*. This result suggests that in a volatile market trading cost rather than information flow is the dominated driver of trading concentration. That might be due to either a reduction in exploration of private information or an increase in noise trading related with uncertain information in the volatile market.

*4.3. Associations of Time Clustering*

In this section, we examine the associations of time clustering to explore how time clustering impacts trade processing. Chiyachantana, Jain, Jiang and Wood (2004) show that price impact is positively associated with institutional (block) trades, shedding light on the relation between time clustering and price impact. Some research, like Brennan and Subrahmannyam (1996) and O'Hara (2003), state that informed trading has higher price impact and thus contribute to price discovery. Therefore, we would expect a positive relation between time clustering and price efficiency through the path of price impact. Additionally, we investigate the relation between time clustering and price volatility.

*Insert Table 7 about here*

Columns 1-3 in Table 7 report the results of regressing price impact, variance ratio, and price volatility on time clustering, respectively, for 2017 subsample. The coefficient of *Price Impact* in Column 1 is positive and significant, suggesting that trading concentration is associated with high price impact and thus contributes to price discovery. The coefficient of *Variance Ratio* in Column 2 is significantly negative, implying that trading concentration is associated with low variance ratio and thus improves market efficiency. Trades with high time clustering, either low duration or large size, are also proven to be associated with high price volatility. (Chan and Fong, 2000; Huang and

Masulis, 2003; Jones, Kaul, and Lipson,1994; Muravyev and Picard, working paper). Our finding lends support the prior studies by showing that time clustering is positively and significantly associated with price volatility, which is measured as the price range divided by the volume-weighted average price.

Columns 4-7 of Table 7 present the results for 2018 subsample, which are consistent with the results of 2017 subsample, showing that time clustering is significantly and positively associated with both price impact and price volatility, but significantly and negatively related with variance ratio. The coefficient of time clustering on price impact is 1.55e-7, indicating that one standard deviation move of time clustering links with a change in percent price impact of 0.00012, or 1.2 BPS. Given that the average percent price impact is 0.00122, or 12.2 BPS, the link between time clustering and price impact is economically significant.

### *4.4. Effects of Time Clustering*

As discussed in the previous section, time clustering is associated with high price impact, suggesting that trading concertation is an implement of information processing. When considering time clustering as an indicator of informed trading, another interesting question is how the market reacts, or in other words, what the effects of time clustering are on the market. In this section, we analyze the effects of trade time clustering on market characteristics, including order splitting and high frequency trading (HFT).

On one hand, informed traders tend to realize the information quickly by using Intermarket Sweep Order (ISO) due to its original design. On the other hand, Informed traders want to hide their information to reduce the trading cost by splitting a large order into small size orders (like Odd-lot orders) and sending them to multiple exchanges. However, while information is revealed by high time clustering, the informed traders and the competitors might change their trading strategies. Table 8 presents the results of regressing led *ISO Proportion*, *Odd-lot Proportion*, and *Number of Exchanges* on *Time Clustering*, showing that the coefficients of led ISO trades proportion and number of exchanges are positive and statistically significant while the coefficient of Odd-lot trades proportion is significantly negative for both stable and volatile markets. The results suggest that while information being disclosed by trading concertation, informed traders tend to trade more aggressively and quickly by increasing ISO trades and reducing Odd-lot trades. The positive relation between led number of exchanges and time clustering implies that informed traders still tend to send their orders to multiple exchanges, which can reduce trading cost and speed transactions.

*Insert Table 8 about here*

Informed trading, especially the aggressive trading causing trades concentration, has significant impact on HFT (Easley, Lopez de Prado, and O'Hara, 2012). To examine the relation between led HFT activities and time clustering, we regress led *Replacement Ratio* and *Cancellation Ratio* on time clustering and report the regression result in Table 9. Columns 1 and 2 of Table 9 show that both order replacement ratio and cancellation ratio are negatively related with time clustering in the stable market of 2017, even though the coefficient of replacement ratio lacks significance, suggesting that time clustering reduces subsequent HFT activities.

However, the effects of time clustering on replacement ratio and cancellation ratio both drop significantly in 2018 (Columns 5 and 6) and become insignificant (Columns 3 and 4). One possible reason for the relation between HFT and time clustering becoming weaker and insignificant is that as we discussed previously in this paper, time clustering in a volatile market is partly caused by noise trading which does not contain information. Alternatively, the reason also could be that HF traders have already made changes in their strategies as a reaction to the increased risk in a volatile market and thus weak the relation with time clustering.

*Insert Table 9 about here*

### *4.5. Robustness Tests*

To test the robustness of our trade time clustering measure and previous results, we compute time clustering of Exchange Traded Funds (ETFs) and examine its determinates, associations and effects. There exist lots of studies shedding light on EFTs trading. For example, while investigating intraday herding for ETFs, Gleason, Mathur, Peterson (2004) state that as a sector with aggregate information ETFs trading does not contain private information and thus lacks herding. Bernile, Hu, and Tang (2016) find that ETFs abnormal returns are related with pre-released of government information. Another study done by Ben-David, Franzoni, and Mousswi (2018) demonstrates that ETFs provide intraday liquidity and satisfy short-term traders' liquidity demand. According to these studies, time clustering of EFT might be related with liquidity, trading cost and pre-released macro-news instead of private information.

We select the top 200 ETFs excluding SPY by the market capitalization ranking, construct the time clustering measure for these ETFs using the same method as for regular stocks, and remove the observations before 9:35 AM and after 3:50 PM. The numbers of observations are 149,458 and 149,857, during the periods of December 1$^{st}$-15$^{th}$, 2017, and February 1$^{st}$-15$^{th}$, 2018, respectively. The mean and standard deviation of ETF *Time Clustering* for the subsample of 2017 are 61.163 and 483.678, respectively, and for the subsample of 2018 the mean and standard deviation are 51.340 and 554.366, respectively. The difference in the means of ETF time clustering between 2017 and 2018 is statistically significant.

Panel A in Table 10 presents the results of regressing lagged liquidity measures and trading cost measures on ETF *Time Clustering*. Columns 1 and 2 show that only the coefficient of lagged *NBB Depth* is statistically significant for subsample of 2017, indicating that in a stable market ETF time clustering is driven by buy-side liquidity supply but not sell-side. Columns 5 and 6 show that the coefficients of both *NBB Depth* and *NBO Depth* are positive and significant, suggesting that in a volatile market both buy-side and sell-side liquidity increase leads to higher time clustering. These findings are consistent with Ben-David, Franzoni, and Mousswi (2018), who argue that ETFs attract investors looking for short-term liquidity.

We next examine the relation between time clustering and lagged *Order Imbalance*, which represents one-sided trading pressure (Marshall, Nguyen, and Visaltanachoti, 2013) rather than private information flow for ETFs. The results of order imbalance show that in the stable market (Columns 3 and 4) the relation between time clustering and lagged order imbalance is not significant while the relation is positive and statistically significant in the volatile market (Columns 7 and 8). Additionally, in an untabulated test, we investigate the relation of time clustering with lagged buy-order imbalance (where NBB depth is larger than NBO depth) and sell-order imbalance (where NBB depth is smaller than NBO depth) , respectively, and find that the coefficient of lagged buy-order imbalance is positive and significant in the stable market of 2017 but insignificant in the volatile market of 2018 while the coefficient of lagged buy-order imbalance is insignificant in the stable market of 2017 but significantly positive in the volatile market of 2018. These findings suggest that ETF time clustering is related with buy-side trading pressure in a stable market but sell-side trading pressure in a volatile market.

Panel A of Table 10 also show that the coefficients of lagged *Quoted Spread* and *Effective Spread* are both negative and significant in the stable market of 2017, but insignificant in the volatile market of 2018. This result indicates that ETF time clustering is also driven by low trading cost in a stable market. However, as discussed in previous section in a volatile market ETF time clustering is more likely to driven by noise investors who have less negotiation power, and thus the relation with trading cost is insignificant.

Next, we examine the associations of ETF time clustering and report the results in Panel B of Table 10. Columns 1-3 show the result for the subsample of 2017, indicating that ETF time clustering are significantly associated with high price impact, low variance ratio and high price volatility. This result is consistent with the result of regular stocks. However, for the volatile market only the relation between time clustering and variance ratio is statistically significant

(Column 5), suggesting that time clustering is associated with high market efficiency. Column 4 of Panel B demonstrates that the coefficient of price impact is negative and insignificant, suggesting that when market is volatile ETF time clustering is not related with price discovery. This result is consistent with the finding of Clifford, Fulkerson and Jordan (2014), who state that return chasing of ETFs is due to naïve extrapolation bias, suggesting that ETFs investors are more likely to be uninformed traders. The result in Column 6 suggests that the relation between ETF time clustering and price volatility is positive but lack of significance.

Panel C of Table 10 presents the results of regressing led order splitting measures and HFT measures on ETF time clustering. Columns 1-5 report the results for the subsample of 2017, showing that as ETF time clustering increases ISO proportion increase, Odd-lot trade proportion decreases, number of exchanges increases, and both replacement and cancellation ratios decrease, which are consistent with the results of regular stocks. Columns 6-10 report the results for the subsample of 2018, showing that the coefficient of ISO proportion is not significant, the coefficient of Odd-lot proportion is significantly negative, the coefficient of number of exchanges is positive and significant, and both the coefficients of replacement and cancellation ratios are insignificant, most of which are consistent with the results of regular stocks except the result of ISO proportion.

## 5. Conclusion

In this paper, we construct a new measure of trade clustering, called time clustering and computed as relative volume over trade time distance for each trade, to capture effects from both trades size and trades frequency. To investigate the determinants, associations and effects of trade clustering, we compute an aggregated measure as the root of the sum of squared trade time clustering for each 5-minute interval. We use Daily TAQ (Trade and Quote) data and separate the sample into two periods, December 2017 and February 2018, by market conditions. This is the first paper to address that the drivers and effects of trade clustering are differentiated by differing market conditions.

We show that there exists a “J” shape for the intraday pattern of time clustering. We apply OLS regression with fixed effects of both stock and date to examine the relation between time clustering and other trade characteristics. We show that time clustering is positively related with lagged NBB depth, NBO depth, and order imbalance, but negatively related with lagged quoted spread and effective spread. These findings suggest that trading concentration is related with high information flow, high liquidity, and low trading cost. Further, we differentiate the effects of these market characteristics on time clustering by market conditions. We show that the relation between time clustering and information flow is significantly weaker in a volatile market. We also demonstrate that buy-side liquidity (NBB depth) has a larger impact on time clustering than sell-side liquidity (NBO depth) in a stable market, while the relation of time clustering with sell-side liquidity is stronger than the relation with buy-side liquidity in a volatile market. Additionally, we show that in a volatile market the effect of quoted spread on time clustering is bigger than that of effective spread, suggesting that in the volatile market time clustering is more likely to be driven by noise investors with less negotiation power.

To investigate the associations of trade time clustering, we regress price impact, price efficiency, and price volatility on time clustering, respectively. The results show that time clustering is positively associated with price impact and price volatility, but negatively associated with variance ratio, implying that trading concentration increases price discovery, improves market efficiency and also raises market volatility.

Finally, we analyze the effects of trade time clustering on order splitting and HFT behaviors, and find that time clustering increases the use of ISO orders and splitting orders by exchanges but decreases the use of Odd-lot orders, and that only in the stable market time clustering can reduce HFT behaviors measured by cancellation rate.

Conclusively, trade time clustering is driven by information flow, liquidity, and trading cost, associated with high price impact, price efficiency, and price volatility, and has significant impacts on order splitting and HFT behaviors.

## References


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

Alexander, G.J., Peterson, M.A., 2007. An analysis of trade-size clustering and its relation to stealth trading. Journal of Financial Economics 84, 435-471.

Anand, A., Chakravarty, S., 2007. Stealth trading in options markets. Journal of Financial and Quantitative Analysis 42, 167-188.

Ascioglu, A., Comerton-Forde C., McInish, T.H., 2007. Price clustering on Tokyo Stock Exchange. The Financial Review 42, 289-301.

Boehmer, B, Boehmer, E., 2003. Trading your neighbor's ETFs: Competition or fragmentation? Journal of Banking and Finance 27, 1667-1703.

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

Ben-David, I., Franzoni, F., Moussawi, R., 2018. Do ETFs increase volatility? Journal of Finance 73, 2471-2535.

Bernile, G., Hu, J., Tang, Y., 2016. Can information be locked up? Informed trading ahead of macro-news announcements. Journal of Financial Economics 121, 496-520.

Bessembinder, H., Kaufman, H.M., 1997. A Comparison of Trade Execution Costs for NYSE and NASDAQ-Listed Stocks. Journal of Financial and Quantitative Analysis 32, 287-310.

Brennan, M., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441-464.

Brown, P., Walsh, D., Yuen, A., 1997. The interaction between order imbalance and stock price. Pacific-Basin Finance Journal 5, 539-557.

Chakravarty, S., Jain, P.K., Upson, J., Wood, R., 2012. Clean sweep: Informed trading through Intermarket Sweep Orders. Journal of Financial and Quantitative Analysis 47, 415-435.

Chan, K., Fong, W., 2000. Trade size, order imbalance, and the volatility-volume relation. Journal of Financial Economics 57, 247-273.

Chan, L.K.C., Lakonishok, J., 1993. Institutional trades and intraday stock price behavior. Journal of Financial Economics 33, 173-199.

Chiyachantana, C.N., Jain, P.K., Jiang, C., Wood, R.A., 2004. International evidence on institutional trading behavior and price impact. Journal of Finance 59, 869-898.

Chiyachantana, C.N., Jain, P.K., 2009. Institutional Trading Frictions. Working paper.

Chordia, T., Subrahmanyam, A., 2004. Order imbalance and individual stock returns: Theory and evidence. Journal of Financial Economics 72, 485-518.

Clifford, C.P., Fulkerson, J.A., Jordan, B., 2014. What drives ETF flows? Financial Review 49, 619-642

Clogg, C. C., Petkova, E., Haritou, A., 1995. Statistical methods for comparing regression coefficients between models. American Journal of Sociology, 100, 1261-1293.

Copeland, T.E., Galai, D., 1985. Information effects on the bid - ask spread. Journal of Finance 38, 1457-1469.

Dufour, A., Engle, R.F., 2000. Time and the price impact of a trade. Journal of Finance 6, 2467-2498.

Easley, D., Lopez de Prado, M.M., O'Hara, M., 2012. Flow toxicity and liquidity in a high-frequency world. Review of Financial Studies 25, 1457-1493.

Franz, D.R., Rao, R.P., and Tripathy, N., 1995. Informed trading risk and bid-ask spread changes around open market stock repurchase in the NASDAQ market. Journal of Financial Research 18, 311-327.

Gleason, K.C., Mathur, I., Peterson, M.A., 2004. Analysis of intraday herding behavior among the sector ETFs. Journal of Empirical Finance 11, 681-694.

Harris, L., 1991. Stock price clustering and discreteness. Journal of Financial and Quantitative Analysis 25, 291-306.

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

Heston, S.L., Sadka, R., 2008. Seasonality in the cross-section of stock returns. Journal of Financial Economics 87, 418-445.

Huang, R.D., Masulis, R.W., 2003. Trading activity and stock price volatility: evidence from London Stock Exchange. Journal of Empirical Finance 10, 249-269.

Jones, C., Kaul, G., Lipson, M., 1994. Transactions, volume and volatility. Review of Financial Studies 7, 631–651.

Keim, D.B., Madhavan, A., 1995. Anatomy of trading process: empirical evidence on the behavior of institutional traders. Journal of Financial Economics 37, 371-398.

Lee, C.M.C., 1992. Earnings news and small traders: an intraday analysis. Journal of Accounting and Economics 15, 265-302.

Lo, A.W., MacKinlay, A. C., Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41-66.

Marshall, B.R., Nguyen, N.H., Visaltanachoti, N., 2013. ETF arbitrage: Intraday evidence. Journal of Banking and Finance 37, 3486-3498.

McInish, T.H., Wood, R.A., 1990. An analysis of transactions data for the Toronto Stock Exchange: return patterns and end-of-the-day effect. Journal of Banking and Finance 14, 441-458.

McInish, T.H., Wood, R.A., 1991. Hourly returns, volume, trade size, and number of trades. Journal of Financial Research 14, 303-315.

McInish, T.H., Wood, R.A., 1992. An analysis of intraday patterns in bid/ask spreads for NYSE stocks. Journal of Finance 47, 753-764.

Muravyev,D., Picard, J., 2016. Does trade clustering reduce trading costs? Unpublished working Paper. Boston College.

O'Hara, M., 2003. Presidential Address: Liquidity and Price Discovery. Journal of Finance 58, 1335-1354.

O'Hara, M., 2015. High frequency market microstructure. Journal of Financial Economics 116, 257-270.

O'Hara, M., Yao, C., Ye, M., 2014. What's not there: Odd lots and market data. Journal of Finance 69, 2199-2236.

Ohta, W., 2006. An analysis of intraday patterns in price clustering on the Tokyo Stock Exchange. Journal of Banking and Finance 30, 1023-1039

Petersen, M. A., Fialkowski, D., 1994. Posted versus effective spreads: Good prices or bad quotes. Journal of Financial Economics 35, 269-292.

Table 1 Summary statistics

This table presents the sample statistics. The means of each variable are reported for years of 2017 and 2018 in columns 1 and 2, respectively, and then for the whole sample in column 3. Column 4 presents the differences in variable means between 2017 and 2018. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| Variable | | 2017 | 2018 | All | Difference (2017-2018) |
|---|---|---|---|---|---|
| Time Clustering | Mean | 144.644 | 134.612 | 139.620 | 10.031*** |
| | Std. | 961.345 | 752.124 | 862.949 | |
| | N | 439,297 | 440,676 | 879,973 | |
| Total Share Volume | Mean | 26,866 | 37,879 | 32,381 | -11,014*** |
| | Std. | 86,907 | 120,252 | 105,083 | |
| | N | 439,297 | 440,676 | 879,973 | |
| Size Volatility | Mean | 209.035 | 206.755 | 207.891 | 2.280 |
| | Std. | 1,121 | 1,325 | 1,228 | |
| | N | 429,387 | 432,142 | 861,529 | |
| Imbalance Depth | Mean | 4.533 | 5.143 | 4.838 | -0.611*** |
| | Std. | 37.251 | 71.743 | 57.185 | |
| | N | 433,834 | 435,044 | 868,878 | |
| NBB Depth | Mean | 9.483 | 9.706 | 9.595 | -0.223 |
| | Std. | 59.216 | 91.924 | 77.341 | |
| | N | 433,834 | 435,044 | 868,878 | |
| NBO Depth | Mean | 9.601 | 8.068 | 8.833 | 1.533*** |
| | Std. | 58.432 | 45.959 | 52.563 | |
| | N | 433,858 | 435,116 | 868,974 | |
| Quoted Spread | Mean | 0.00265 | 0.00313 | 0.00289 | -0.00048*** |
| | Std. | 0.00401 | 0.00420 | 0.00411 | |
| | N | 433,834 | 435,044 | 868,878 | |
| Effective Spread | Mean | 0.00154 | 0.00178 | 0.00166 | -0.00024*** |
| | Std. | 0.00299 | 0.00315 | 0.00308 | |
| | N | 434,817 | 435,751 | 870,568 | |
| Price Impact | Mean | 0.00108 | 0.00122 | 0.00115 | -0.00014*** |
| | Std. | 0.00389 | 0.00393 | 0.00391 | |
| | N | 434,759 | 431,186 | 865,945 | |
| Variance Ratio | Mean | 1.621 | 1.615 | 1.618 | 0.007*** |
| | Std. | 0.400 | 0.347 | 0.374 | |
| | N | 416,904 | 422,443 | 839,347 | |
| Price Volatility | Mean | 0.00281 | 0.00406 | 0.00344 | -0.00124*** |
| | Std. | 0.00305 | 0.00383 | 0.00352 | |
| | N | 439,297 | 440,676 | 879,973 | |
| ISO Proportion | Mean | 0.335 | 0.351 | 0.343 | -0.016*** |
| | Std. | 0.201 | 0.194 | 0.198 | |

| | | | | | |
|---|---|---|---|---|---|
| | N | 439,297 | 440,676 | 879,973 | |
| Odd-lot Proportion | Mean | 0.179 | 0.167 | 0.173 | 0.013*** |
| | Std. | 0.198 | 0.183 | 0.191 | |
| | N | 439,297 | 440,676 | 879,973 | |
| Number of Exchange | Mean | 7.846 | 8.293 | 8.070 | -0.447*** |
| | Std. | 3.134 | 3.126 | 3.138 | |
| | N | 439,297 | 440,676 | 879,973 | |
| Replacement Ratio | Mean | 3.801 | 4.470 | 4.142 | -0.669*** |
| | Std. | 10.598 | 12.513 | 11.619 | |
| | N | 373,325 | 388,911 | 762,236 | |
| Cancellation Ratio | Mean | 17.749 | 20.567 | 19.186 | -2.818*** |
| | Std. | 23.568 | 27.554 | 25.718 | |
| | N | 373,325 | 388,911 | 762,236 | |
| SPY Time Clustering | Mean | 10.297 | 7.316 | 8.806 | 2.981*** |
| | Std. | 5.508 | 4.857 | 5.401 | |
| | N | 750 | 750 | 1,500 | |
| SPY Volume | Mean | 872,756 | 1,955,627 | 1,414,192 | -1,082,871*** |
| | Std. | 837,280 | 1,596,695 | 1,384,738 | |
| | N | 750 | 750 | 1,500 | |
| SPY Price Volatility | Mean | 0.00122 | 0.00566 | 0.00344 | -0.00445*** |
| | Std. | 0.00163 | 0.00806 | 0.00622 | |
| | N | 750 | 750 | 1,500 | |

Table 2 Pearson's Correlation of the Market Measures

This table presents the Pearson product-moment correlations among the market measures, including (1) *Time Clustering*, (2) *Total Share Volume*, (3) *Size Volatility*, (4) *Imbalance Depth*, (5) *NBB Depth*, (6) *NBO Depth*, (7) *Quoted Spread*, (8) *Effective Spread*, (9) *Price Impact*, (10) *Variance Ratio*, (11) *Price Volatility*, (12) *ISO Proportion*, (13) *Odd-lot Proportion*, (14) *Number of Exchange*, (15) *Replacement Ratio*, (16) *Cancellation Ratio*, (17) *SPY Time Clustering*, (18) *SPY Volume*, and (19) *SPY Price Volatility*.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (2) | -0.008 | 1.000 | | | | | | | | | | | | | | | | |
| (3) | 0.012 | 0.323 | 1.000 | | | | | | | | | | | | | | | |
| (4) | 0.016 | 0.192 | 0.077 | 1.000 | | | | | | | | | | | | | | |
| (5) | 0.020 | 0.293 | 0.106 | 0.802 | 1.000 | | | | | | | | | | | | | |
| (6) | 0.021 | 0.306 | 0.113 | 0.679 | 0.732 | 1.000 | | | | | | | | | | | | |
| (7) | 0.081 | -0.159 | -0.041 | -0.013 | -0.035 | -0.040 | 1.000 | | | | | | | | | | | |
| (8) | 0.078 | -0.107 | -0.003 | 0.001 | -0.015 | -0.017 | 0.600 | 1.000 | | | | | | | | | | |
| (9) | 0.051 | -0.066 | -0.025 | -0.004 | -0.012 | -0.012 | 0.309 | 0.530 | 1.000 | | | | | | | | | |
| (10) | -0.009 | 0.004 | 0.005 | 0.004 | 0.005 | 0.005 | -0.052 | -0.032 | -0.025 | 1.000 | | | | | | | | |
| (11) | 0.066 | 0.101 | 0.029 | -0.001 | -0.010 | -0.009 | 0.359 | 0.314 | 0.190 | -0.009 | 1.000 | | | | | | | |
| (12) | -0.017 | -0.039 | -0.100 | -0.027 | -0.038 | -0.037 | -0.024 | -0.005 | 0.004 | -0.025 | -0.049 | 1.000 | | | | | | |
| (13) | -0.005 | -0.204 | -0.134 | -0.076 | -0.107 | -0.115 | 0.143 | 0.089 | 0.031 | -0.023 | -0.097 | 0.190 | 1.000 | | | | | |
| (14) | -0.058 | 0.307 | 0.097 | 0.055 | 0.096 | 0.104 | -0.470 | -0.343 | -0.170 | 0.070 | 0.020 | 0.012 | -0.322 | 1.000 | | | | |
| (15) | -0.007 | -0.056 | -0.017 | 0.001 | -0.002 | 0.000 | 0.126 | 0.086 | 0.043 | -0.026 | 0.016 | -0.020 | 0.024 | -0.208 | 1.000 | | | |
| (16) | -0.043 | -0.050 | -0.009 | 0.042 | 0.055 | 0.058 | 0.065 | 0.044 | 0.019 | -0.015 | -0.013 | -0.023 | -0.045 | -0.147 | 0.407 | 1.000 | | |
| (17) | 0.027 | 0.023 | 0.002 | 0.006 | 0.010 | 0.013 | -0.024 | -0.015 | 0.000 | 0.003 | -0.001 | 0.007 | 0.004 | 0.025 | -0.013 | -0.029 | 1.000 | |
| (18) | 0.009 | 0.113 | -0.002 | -0.002 | -0.005 | -0.009 | 0.108 | 0.078 | 0.042 | -0.006 | 0.332 | 0.061 | -0.041 | 0.085 | 0.043 | 0.033 | 0.111 | 1.000 |
| (19) | 0.004 | 0.074 | -0.004 | 0.000 | -0.002 | -0.006 | 0.056 | 0.042 | 0.023 | -0.005 | 0.192 | 0.058 | -0.026 | 0.065 | 0.023 | 0.023 | -0.012 | 0.636 |

Table 3 Intraday determinants of time clustering

This table shows the regression results of time clustering on dummy variables of weekdays and intraday 5-minute intervals. *Volatile Market* is a dummy variable which takes one if it is in February 2018, and zero if it is in December 2017. The fixed effects of stocks are controlled. t-statistics are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Time Clustering |
|---|---|
| Constant | 257.9*** |
| | (10.94) |
| Volatile Market | -13.47*** |
| | (-5.006) |
| 2nd.dayofweek | -2.221 |
| | (-0.797) |
| 3rd.dayofweek | 4.337 |
| | (1.418) |
| 4th.dayofweek | 7.350* |
| | (1.847) |
| 5th.dayofweek | 0.138 |
| | (0.0451) |
| 2nd.5-minute | -233.0*** |
| | (-9.165) |
| 3rd.5-minute | -224.8*** |
| | (-9.377) |
| 77th.5-minute | 120.8*** |
| | (4.668) |
| 78th.5-minute | 542.5*** |
| | (13.31) |
| 4th.5-minute to 76th.5-minute | Significantly Negative |
| Stock Fixed Effects | Yes |
| Observations | 915,948 |
| R-squared | 0.044 |

Table 4 Determinants of Time Clustering – Informed Trading

This table shows the regression results of time clustering on informed trading measure. Column 1 presents the result for the subsample in December 2017, when market is stable. Column 2 presents the results for the subsample in February 2018, when market is volatile. Column 3 presents the results for the whole sample. The regression equation used here is:

$$TimeClustering_{it} = \alpha + \beta_1 Determinant_{i(t-1)} + \beta_2 TimeClustering_{i(t-1)} + \sum \beta_{3n} ControlVariables_{itn} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it}$$

All the variables are computed at trade level and then aggregated for each 5-minute interval. The informed trading measure is lagged *Order Imbalance*, which is the absolute value of the share volume difference between NBO and NBB orders. The control variables include *Lagged Time Clustering*, *Total Share Volume*, *Size Volatility, Time Clustering of SPY, Volume of SPY, and Price Volatility of SPY*. *Size Volatility* is the standard deviation of trade sizes. *Price Volatility* is the price range divided by volume-weighted average price. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Stable Market | Volatile Market | All |
|---|---|---|---|
| | (1) | (2) | (3) |
| Constant | -31.54*** | 7.709* | -21.45*** |
| | (-5.592) | (1.788) | (-4.457) |
| Lagged Order Imbalance | 0.189** | 0.0361 | 0.217*** |
| | (2.223) | (1.571) | (4.227) |
| Volatile | | | 13.24** |
| | | | (2.297) |
| Volatile * Lagged Order Imbalance | | | -0.184*** |
| | | | (-3.540) |
| Lagged Time Clustering | 0.0119** | 0.0133*** | 0.0140*** |
| | (2.489) | (2.604) | (4.018) |
| Total Share Volume | 0.000447** | 0.000185*** | 0.000254** |
| | (2.328) | (2.677) | (2.557) |
| Size Volatility | 0.00466 | 0.00769** | 0.00747** |
| | (0.700) | (2.162) | (2.137) |
| TC of SPY | 3.221*** | 1.881*** | 2.751*** |
| | (8.528) | (6.230) | (10.37) |
| Volume of SPY | 7.91e-06*** | 1.13e-05*** | 9.58e-06*** |
| | (3.904) | (8.039) | (7.901) |
| Price Volatility of SPY | -395.0 | 429.0** | 340.5** |
| | (-0.512) | (2.544) | (2.069) |
| Stock Fixed Effects | Yes | Yes | Yes |
| Date Fixed Effects | Yes | Yes | Yes |
| Observations | 413,319 | 416,576 | 829,895 |
| R-squared | 0.029 | 0.043 | 0.033 |

Table 5 Determinants of Time Clustering – Liquidity and Trading Cost

This table reports the regression results of time clustering on liquidity and trading cost measures. Columns 1-4 present the results for the subsample in December 2017, when market is stable. Columns 5-8 present the results for the subsample in February 2018, when market is volatile. Columns 9-12 present the results for the whole sample. The regression equation used here is:

$$TimeClustering_{it} = \alpha + \beta_1 Determinant_{i(t-1)} + \beta_2 TimeClustering_{i(t-1)} + \sum \beta_{3n} ControlVariables_{itn} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it}$$

All the variables are computed at trade level and then aggregated for each 5-minute interval. The dependent variable is *Time Clustering* in the 5-minute interval of t. The liquidity measure in Columns 1, 5, and 9 is lagged *NBB Depth*. The liquidity measure in Columns 2, 6, and 10 is lagged *NBO Depth*. The trading cost measure in Columns 3, 7, and 11 is lagged *Quoted Spread*. The d trading cost measure in Columns 4, 8, and 12 is lagged *Effective Spread*. All the liquidity measures and trading cost measures represent the 5-minute interval of t-1. The control variables include *Lagged Time Clustering*, *Total Share Volume*, *Size Volatility*, *Time Clustering of SPY, Volume of SPY, and Price Volatility of SPY*. *Size Volatility* is the standard deviation of trade sizes. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Stable Market – Dec. 2017 | | | | Volatile Market – Feb. 2018 | | | | All | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NBB | NBO | Qt. Spread | Ef. Spread | NBB | NBO | Qt. Spread | Ef. Spread | NBB | NBO | Qt. Spread | Ef. Spread |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Constant | -31.72*** | -31.72*** | -27.86*** | -29.11*** | 7.656* | 7.449* | 15.08*** | 10.10** | -21.92*** | -21.67*** | -14.52*** | -17.02*** |
| | (-5.626) | (-5.625) | (-5.150) | (-5.458) | (1.773) | (1.717) | (3.380) | (2.287) | (-4.558) | (-4.495) | (-2.999) | (-3.872) |
| Lag. Market | 0.347*** | 0.213** | -2,250 | -3,324*** | 0.0304 | 0.129** | -3,640*** | -2,597*** | 0.226*** | 0.204*** | -3,327** | -4,007*** |
| Measures | (4.270) | (2.425) | (-1.176) | (-2.956) | (1.539) | (2.230) | (-7.232) | (-3.774) | (6.643) | (4.194) | (-2.035) | (-3.777) |
| Volatile | | | | | | | | | 14.02** | 13.27** | 9.736 | 9.214 |
| | | | | | | | | | (2.435) | (2.298) | (1.474) | (1.620) |
| Vol.*Lag. | | | | | | | | | -0.189*** | -0.0883** | 898.6 | 2,198* |
| Measures | | | | | | | | | (-5.491) | (-2.288) | (0.567) | (1.681) |
| Lagged TC | 0.0118** | 0.0118** | 0.0118** | 0.0115** | 0.0133*** | 0.0133*** | 0.0132*** | 0.0134*** | 0.0139*** | 0.0139*** | 0.0139*** | 0.0137*** |
| | (2.484) | (2.484) | (2.477) | (2.505) | (2.604) | (2.602) | (2.590) | (2.609) | (4.015) | (4.014) | (4.006) | (4.028) |
| Volume | 0.000446* | 0.000451* | 0.000449* | 0.000449* | 0.000186** | 0.000186** | 0.000184** | 0.000185** | 0.000260** | 0.000257** | 0.000251* | 0.000252* |
| | (2.355) | (2.348) | (2.322) | (2.325) | (2.682) | (2.699) | (2.669) | (2.681) | (2.599) | (2.590) | (2.509) | (2.520) |
| Size Volt. | 0.00468 | 0.00460 | 0.00477 | 0.00479 | 0.00769** | 0.00768** | 0.00769** | 0.00767** | 0.00739** | 0.00740** | 0.00757** | 0.00754** |
| | (0.702) | (0.692) | (0.715) | (0.723) | (2.162) | (2.176) | (2.157) | (2.155) | (2.121) | (2.126) | (2.162) | (2.154) |
| SPY TC | 3.217*** | 3.210*** | 3.227*** | 3.222*** | 1.880*** | 1.877*** | 1.804*** | 1.848*** | 2.747*** | 2.744*** | 2.740*** | 2.749*** |
| | (8.551) | (8.508) | (8.518) | (8.553) | (6.229) | (6.226) | (5.966) | (6.121) | (10.36) | (10.35) | (10.27) | (10.34) |
| SPY Volm. | 7.98e- | 7.91e- | 8.13e- | 8.24e- | 1.13e- | 1.13e- | 1.18e- | 1.14e- | 9.55e- | 9.57e- | 9.96e- | 9.79e- |
| | (3.946) | (3.907) | (3.956) | (4.063) | (8.038) | (8.048) | (8.398) | (8.171) | (7.868) | (7.887) | (8.061) | (8.007) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPY Volt. | -420.6 | -392.3 | -419.9 | -428.9 | 428.4** | 428.4** | 349.2** | 408.3** | 338.1** | 340.2** | 283.4* | 319.8* |
| | (-0.546) | (-0.509) | (-0.549) | (-0.558) | (2.540) | (2.541) | (2.067) | (2.408) | (2.054) | (2.068) | (1.723) | (1.931) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observation | 413,319 | 413,331 | 413,319 | 413,939 | 416,576 | 416,619 | 416,576 | 416,973 | 829,895 | 829,950 | 829,895 | 830,912 |
| R-squared | 0.029 | 0.029 | 0.029 | 0.029 | 0.043 | 0.043 | 0.043 | 0.043 | 0.033 | 0.033 | 0.033 | 0.033 |

## Table 6 Determinants of Time Clustering – Horse Race

This table presents the regression results of time clustering on its determinants. Columns 1 and 2 present the results for the subsample in December 2017, when market is stable. Columns 3 and 4 present the results for the subsample in February 2018, when market is volatile. Columns 5 and 6 present the results for the whole sample. All the variables are computed at trade level and then aggregated for each 5-minute interval. The determinant variables in Columns 1, 3, and 5 include lagged *Order Imbalance* and *Quoted Spread*. The determinant variables in Columns 2, 4, and 6 include lagged *Order Imbalance* and *Effective Spread*. The control variables include *Lagged Time Clustering*, *Total Share Volume*, *Size Volatility*, *Time Clustering of SPY*, *Volume of SPY*, *and Price Volatility of SPY*. *Size Volatility* is the standard deviation of trade sizes. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Stable Market | | Volatile Market | | All | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | -27.70*** | -28.58*** | 15.09*** | 10.26** | -14.72*** | -16.94*** |
| | (-5.123) | (-5.368) | (3.381) | (2.318) | (-3.037) | (-3.866) |
| Lag. Order | 0.188** | 0.188** | 0.0361 | 0.0363 | 0.217*** | 0.217*** |
| Imbalance | (2.228) | (2.229) | (1.572) | (1.587) | (4.234) | (4.214) |
| Lag. Qt. Spread | -2,244 | | -3,640*** | | -3,309** | |
| | (-1.173) | | (-7.232) | | (-2.023) | |
| Lag. Ef. Spread | | -3,274*** | | -2,589*** | | -3,922*** |
| | | (-2.917) | | (-3.703) | | (-3.679) |
| Volatile | | | | | 10.36 | 9.565* |
| | | | | | (1.568) | (1.684) |
| Vol.*Lag. | | | | | -0.185*** | -0.185*** |
| Order Imb. | | | | | (-3.567) | (-3.545) |
| Vol.*Lag. | | | | | 856.8 | |
| Qt. Spread | | | | | (0.541) | |
| Vol.*Lag. | | | | | | 2,099 |
| Ef. Spread | | | | | | (1.586) |
| Lagged TC | 0.0118** | 0.0119** | 0.0132*** | 0.0133*** | 0.0139*** | 0.0139*** |
| | (2.478) | (2.488) | (2.590) | (2.601) | (4.005) | (4.016) |
| Volume | 0.000446** | 0.000447** | 0.000183*** | 0.000185*** | 0.000254** | 0.000255** |
| | (2.322) | (2.325) | (2.659) | (2.672) | (2.549) | (2.561) |
| Size Volatility | 0.00478 | 0.00473 | 0.00774** | 0.00771** | 0.00753** | 0.00748** |
| | (0.716) | (0.711) | (2.168) | (2.166) | (2.157) | (2.144) |
| SPY TC | 3.228*** | 3.222*** | 1.803*** | 1.850*** | 2.738*** | 2.747*** |
| | (8.537) | (8.525) | (5.965) | (6.116) | (10.27) | (10.31) |
| SPY Volume | 8.16e-06*** | 8.13e-06*** | 1.18e-05*** | 1.14e-05*** | 9.95e-06*** | 9.72e-06*** |
| | (3.969) | (4.004) | (8.404) | (8.157) | (8.071) | (7.975) |
| SPY Volatility | -426.7 | -437.8 | 348.7** | 401.1** | 281.7* | 314.6* |
| | (-0.557) | (-0.569) | (2.064) | (2.373) | (1.713) | (1.904) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 413,319 | 413,302 | 416,576 | 416,509 | 829,895 | 829,811 |
| R-squared | 0.029 | 0.029 | 0.043 | 0.043 | 0.033 | 0.033 |

## Table 7 Associations of Time Clustering – Price Impact, Price Efficiency and Price Volatility

This table reports the results of regressing price impact, price efficiency and price volatility on time clustering, respectively. Columns 1-3 present the results for the subsample in December 2017, when market is stable. Columns 4-6 present the results for the subsample in February 2018, when market is volatile. Columns 7-9 present the results for the whole sample. The regression equation used here is:

$$Association_{it} = \alpha + \beta_1 Timeclustering_{it} + \sum \beta_{2n} ControlVariables_{itn} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it}$$

All the variables are computed at trade level and then aggregated for each 5-minute interval. The dependent variable in Columns 1, 4, and 7 is percent *Price Impact* in the 5-minute interval of t. The dependent variable in Columns 2, 5, and 8 is *Price Efficiency*, which is the absolute nature logarithms value of the variance ratio of the 5-minute interval of t. The dependent variable in Columns 3, 6, and 9 is *Price Volatility*, which is price range divided by volume-weighted average price in the 5-minute interval of t. We regress these market quality measures on *Time Clustering* in the same 5-minute interval of t. The control variables include *Total Share Volume, Size Volatility, Time Clustering of SPY, Volume of SPY, and Price Volatility of SPY*. *Size Volatility* is the standard deviation of trade sizes. *Price Volatility* is the standard deviation of trade returns. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Stable Market | | | Volatile Market | | | All | | |
|---|---|---|---|---|---|---|---|---|---|
| | Price Impact | Price Efficiency | Price Volatility | Price Impact | Price Efficiency | Price Volatility | Price Impact | Price Efficiency | Price Volatility |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Constant | 0.000771*** | 1.622*** | 0.00298*** | 0.000911*** | 1.645*** | 0.00246*** | 0.000928*** | 1.629*** | 0.00268*** |
| | (32.10) | (753.3) | (92.69) | (34.11) | (718.3) | (52.30) | (41.46) | (821.9) | (75.45) |
| Time Clustering | 8.53e-08*** | -3.24e-06*** | 1.19e-07*** | 1.55e-07*** | -2.84e-06* | 2.00e-07*** | 9.15e-08*** | -3.15e-06*** | 1.34e-07*** |
| | (2.824) | (-2.848) | (3.062) | (3.985) | (-1.770) | (4.462) | (2.904) | (-2.908) | (3.144) |
| Volatile | | | | | | | -0.000233*** | 1.05e-07 | -9.68e-05* |
| | | | | | | | (-7.444) | (0.0579) | (-1.769) |
| Volatile*TC | | | | | | | 4.99e-08 | 0.000404 | 3.30e-08 |
| | | | | | | | (1.087) | (0.145) | (0.598) |
| Volume | 1.30e-10 | -1.47e-08** | 6.44e-09*** | -2.10e- | -6.70e-09** | 6.07e-09*** | -7.58e-11 | -7.79e-09** | 6.43e-09*** |
| | (1.332) | (-2.407) | (4.187) | (-2.895) | (-2.333) | (5.001) | (-1.423) | (-2.189) | (4.908) |
| Size Volatility | -2.01e-08** | 4.15e-07 | -2.84e-08 | -4.99e-09 | 2.31e-07** | -3.49e-08 | -1.29e-08* | 2.58e-07 | -3.82e-08* |
| | (-2.173) | (1.117) | (-1.341) | (-0.733) | (2.445) | (-1.579) | (-1.929) | (1.525) | (-1.811) |
| SPY TC | 6.89e-06*** | -0.000410*** | 3.12e-05*** | -5.25e- | -0.000168 | -2.60e-05*** | 2.57e-06*** | -0.000338*** | 5.07e-06*** |
| | (6.032) | (-3.994) | (25.75) | (-4.043) | (-1.543) | (-22.74) | (3.133) | (-4.617) | (5.603) |
| SPY Volume | 6.25e-11*** | -5.76e-10 | 4.66e-10*** | 7.25e-11*** | 1.23e-10 | 7.46e-10*** | 6.79e-11*** | -0 | 6.56e-10*** |
| | (5.697) | (-0.644) | (45.70) | (10.01) | (0.303) | (62.47) | (11.16) | (-0.00747) | (64.46) |
| SPY Volatility | 0.00484 | 0.0432 | -0.0171*** | -0.00395*** | 0.0344 | -0.0315*** | -0.00448*** | 0.0642 | -0.0303*** |
| | (1.182) | (0.114) | (-6.198) | (-4.421) | (0.542) | (-27.53) | (-5.332) | (1.036) | (-27.35) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 425,151 | 416,904 | 429,387 | 423,034 | 422,443 | 432,142 | 848,185 | 839,347 | 861,529 |
| R-squared | 0.060 | 0.006 | 0.290 | 0.052 | 0.006 | 0.252 | 0.053 | 0.005 | 0.273 |

## Table 8 Effects of Time Clustering – Order Splitting

This table presents the results of regressing order splitting measures on time clustering. Columns 1 and 2 present the results for the subsample in December 2017, when market is stable. Columns 3 and 4 present the results for the subsample in February 2018, when market is volatile. Columns 5 and 6 present the results for the whole sample. The regression model is:

$$Affected_{i(t+1)} = \alpha + \beta_1 Timeclustering_{it} + \sum \beta_{2n} ControlVariables_{i(t+1)n} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it}$$

All the variables are computed at trade level and then aggregated for each 5-minute interval. The determinant variable is *Time Clustering* in the 5-minute interval of t. The order splitting measures are led *ISO Proportion* of the 5-minute interval of t+1 in Columns 1, 4, and 7, led *Odd-lot Proportion* of the 5-minute interval of t+1 in Columns 2, 5, and 8, led N*umber of Exchanges* of the 5-minute interval of t+1 in Columns 3, 6, and 9,respectively. The control variables include *Total Share Volume, Size Volatility, Time Clustering of SPY, Volume of SPY, and Price Volatility of SPY*. *Size Volatility* is the standard deviation of trade sizes. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Stable Market | | | Volatile Market | | | All | | |
|---|---|---|---|---|---|---|---|---|---|
| | Led ISO | Led Odd-lot | Led N. of Exc. | Led ISO | Led Odd-lot | Led N. of Exc. | Led ISO | Led Odd-lot | Led N. of Exc. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Constant | 0.361*** | 0.153*** | 9.328*** | 0.369*** | 0.153*** | 8.576*** | 0.368*** | 0.153*** | 8.880*** |
| | (172.0) | (96.90) | (412.2) | (159.2) | (91.88) | (305.5) | (165.6) | (92.68) | (355.9) |
| TC | 1.30e-06*** | -1.72e-06*** | 2.52e-05*** | 9.32e-07* | -1.97e-06*** | 3.23e-05*** | 2.22e-06*** | -1.41e-06*** | 3.10e-05*** |
| | (3.205) | (-2.771) | (3.596) | (1.874) | (-3.422) | (3.298) | (3.807) | (-2.771) | (3.611) |
| Volatile | | | | | | | 0.0182*** | -0.00976*** | 0.186*** |
| | | | | | | | (5.266) | (-4.109) | (4.748) |
| Vol.* TC | | | | | | | -2.50e-06*** | -6.90e-07 | -1.11e-05 |
| | | | | | | | (-3.202) | (-0.818) | (-0.919) |
| Led Volume | 3.46e-08 | -6.27e-08** | 1.90e-06*** | -1.49e-08 | -3.65e-08** | 4.33e-07* | 2.22e-08 | -4.02e-08** | 1.00e-06*** |
| | (1.229) | (-2.483) | (3.045) | (-0.779) | (-2.365) | (1.689) | (1.291) | (-2.469) | (2.609) |
| Led Size Volt. | -1.28e-05*** | -6.51e-06*** | 8.83e-06 | -8.06e-06** | -4.18e-06* | 2.14e-05 | -1.04e-05*** | -5.41e-06*** | 1.80e-05 |
| | (-3.479) | (-3.726) | (0.932) | (-2.134) | (-1.783) | (1.530) | (-3.360) | (-2.897) | (1.549) |
| Led SPY TC | 0.000714*** | -0.000656*** | 0.0303*** | 0.000944*** | -0.000156*** | 0.0225*** | 0.000780*** | -0.000474*** | 0.0267*** |
| | (10.94) | (-12.43) | (36.83) | (13.18) | (-2.921) | (31.15) | (16.25) | (-11.66) | (42.54) |
| Led SPY Volm. | 3.16e-09*** | -2.11e-09*** | 1.10e-07*** | 4.19e-09*** | -2.97e-09*** | 1.66e-07*** | 3.54e-09*** | -2.76e-09*** | 1.44e-07*** |
| | (6.182) | (-5.495) | (19.89) | (11.41) | (-12.49) | (35.04) | (11.63) | (-12.93) | (32.53) |
| Led SPY Volt. | -0.339* | -0.737*** | 3.828** | 0.565*** | -0.00137 | 8.026*** | 0.564*** | 0.000549 | 8.094*** |
| | (-1.752) | (-4.671) | (2.205) | (11.64) | (-0.0381) | (17.28) | (11.82) | (0.0151) | (17.60) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 417,937 | 417,937 | 417,937 | 421,395 | 421,395 | 421,395 | 839,332 | 839,332 | 839,332 |
| R-squared | 0.083 | 0.304 | 0.682 | 0.110 | 0.307 | 0.711 | 0.081 | 0.291 | 0.686 |

## Table 9 Effects of Time Clustering – High Frequency Trading

This table presents the results of regressing high frequency trading (HFT) measures on time clustering. Columns 1 and 2 present the results for the subsample in December 2017, when market is stable. Columns 3 and 4 present the results for the subsample in February 2018, when market is volatile. Columns 5 and 6 present the results for the whole sample. The regression model is:

$$Affected_{i(t+1)} = \alpha + \beta_1 Timeclustering_{it} + \sum \beta_{2n} ControlVariables_{i(t+1)n} + FE\ of\ Stock + FE\ of\ Day + \varepsilon_{it}$$

All the variables are computed at trade level and then aggregated for each 5-minute interval. The determinant variable is *Time Clustering* in the 5-minute interval of t. The liquidity measures are led *Replacement Ratio* of the 5-minute interval of t+1 in Columns 1, 3, and 5, and led *Cancellation Ratio* of the 5-minute interval of t+1 in Columns 2, 4, and 6, respectively. The control variables include led *Total Share Volume, Size Volatility, Time Clustering of SPY, Volume of SPY, and Price Volatility of SPY*. *Size Volatility* is the standard deviation of trade sizes. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

| | Stable Market | | Volatile Market | | All | |
|---|---|---|---|---|---|---|
| | Led Repl. | Led Canc. | Led Repl. | Led Canc. | Led Repl. | Led Canc. |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 5.374*** | 17.43*** | 4.784*** | 19.51*** | 6.311*** | 20.61*** |
| | (30.89) | (58.33) | (34.77) | (59.15) | (29.18) | (58.79) |
| TC | -1.33e-05 | -4.34e-05* | 1.58e-05 | -4.27e-05 | -0.000132** | -4.40e-05 |
| | (-1.510) | (-1.846) | (0.203) | (-1.062) | (-2.483) | (-1.272) |
| Volatile | | | | | -4.43e-06*** | -2.57e-05*** |
| | | | | | (-4.145) | (-4.693) |
| Vol.* TC | | | | | 5.28e-05* | 0.000322*** |
| | | | | | (1.895) | (2.779) |
| Led Volume | -4.49e-06*** | -2.70e-05*** | -3.75e-06*** | -2.31e-05*** | -0.00111 | -0.0130** |
| | (-4.416) | (-3.812) | (-4.014) | (-4.447) | (-0.465) | (-2.067) |
| Led Size Volt. | 7.79e-05*** | 0.000391*** | 3.90e-05 | 0.000258* | 1.68e-07*** | 1.66e-07*** |
| | (3.151) | (2.744) | (1.393) | (1.691) | (6.628) | (3.186) |
| Led SPY TC | 0.00701** | 0.00945 | -0.0197*** | -0.0711*** | -31.45*** | -60.21*** |
| | (1.999) | (1.031) | (-5.619) | (-7.776) | (-8.862) | (-7.767) |
| Led SPY Volm. | 2.36e-07*** | 5.01e-07*** | 1.31e-07*** | 3.52e-08 | -1.568*** | -1.523*** |
| | (6.193) | (6.126) | (4.445) | (0.599) | (-5.418) | (-2.732) |
| Led SPY Volt. | 9.183 | -34.31 | -29.24*** | -47.67*** | 0.000297*** | 7.73e-06 |
| | (0.840) | (-1.209) | (-8.026) | (-6.272) | (2.765) | (0.111) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 364,779 | 364,779 | 380,818 | 380,818 | 745,597 | 745,597 |
| R-squared | 0.115 | 0.181 | 0.266 | 0.193 | 0.166 | 0.167 |

## Table 10 Robustness Tests – ETF Time Clustering

This table presents the results of tests for time clustering of exchange traded fund (ETF). In Panel A, we report the results of regressing *ETF Time Clustering* on its determinants, including lagged *NBB Depth*, *NBO Depth*, *Order Imbalance, Quoted Spread* and *Effective Spread*. Columns 1-4 present the results for the subsample in December 2017, when market is stable. Columns 5-6 present the results for the subsample in February 2018, when market is volatile. Panel B presents the results of regressing time clustering associations, including *Price Impact*, *Price Efficiency*, and *Price Volatility,* on time clustering. Columns 1-3 in Panel B present the results for the subsample in December 2017. Columns 4-6 in Panel B present the results for the subsample in February 2018. Panel C presents the results of regressing order splitting and HFT measures, including led *ISO Proportion*, *Odd-lot Proportion, Number of Exchanges, Replacement Ratio,* and *Cancellation Ratio,* on time clustering. Columns 1-5 in Panel C present the results for the subsample in December 2017. Columns 6-10 in Panel C present the results for the subsample in February 2018. All the variables are computed at trade level and then aggregated for each 5-minute interval. The characteristics of the ETF of SPY are introduced as control variables of market condition. t-Statistics based on standard errors adjusted for stock clustering are reported in parentheses. ***, **, and * represent statistical significance at the 1%, 5%, and 10%, respectively.

Panel A Determinants of ETF Time Clustering

| | Stable Market | | | | Volatile Market | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Constant | 180.8*** | 180.7*** | 181.6*** | 181.2*** | 156.7*** | 156.7*** | 157.9*** | 157.4*** |
| | (39.65) | (39.40) | (39.46) | (39.82) | (33.00) | (32.83) | (35.22) | (33.35) |
| Lag. NBB | 0.0203*** | | | | 0.0336*** | | | |
| | (2.646) | | | | (3.826) | | | |
| Lag. NBO | | 0.00854 | | | | 0.0280*** | | |
| | | (0.894) | | | | (5.288) | | |
| Lag. Order | | | -0.00342 | -0.00345 | | | 0.0199*** | 0.0198*** |
| Imbalance | | | (-0.304) | (-0.307) | | | (3.262) | (3.239) |
| Lag. | | | -5,136** | | | | -2,993 | |
| Qt. Spread | | | (-2.375) | | | | (-0.986) | |
| Lag. | | | | -3,788** | | | | -1,724 |
| Ef. Spread | | | | (-2.320) | | | | (-1.533) |
| Lag. | 0.0235** | 0.0235** | 0.0235** | 0.0235** | 0.00497 | 0.00495 | 0.00496 | 0.00497 |
| ETF TC | (2.023) | (2.024) | (2.025) | (2.029) | (1.472) | (1.470) | (1.468) | (1.473) |
| Volume | 0.000145*** | 0.000144*** | 0.000142*** | 0.000142*** | 6.05e- | 6.03e- | 5.85e-05** | 5.88e-05** |
| | (2.697) | (2.705) | (2.643) | (2.646) | (2.618) | (2.611) | (2.523) | (2.535) |
| Size | 7.89e-05 | 8.30e-05 | 0.000103 | 0.000101 | 0.000164 | 0.000168 | 0.000193 | 0.000189 |
| | (0.203) | (0.218) | (0.265) | (0.261) | (0.591) | (0.605) | (0.686) | (0.673) |
| SPY TC | 0.967*** | 0.971*** | 0.974*** | 0.974*** | 0.769*** | 0.773*** | 0.770*** | 0.776*** |
| | (3.815) | (3.830) | (3.824) | (3.809) | (3.197) | (3.224) | (3.268) | (3.250) |
| SPY Volm. | 7.67e-06*** | 7.67e-06*** | 7.77e-06*** | 7.69e-06*** | 8.52e- | 8.51e- | 8.63e- | 8.56e- |
| | (4.968) | (4.954) | (5.009) | (4.954) | (5.880) | (5.882) | (5.718) | (5.847) |
| SPY Volt. | -1,754*** | -1,758*** | -1,761*** | -1,757*** | -342.3* | -342.5* | -347.7* | -342.9* |
| | (-3.042) | (-3.050) | (-3.049) | (-3.040) | (-1.895) | (-1.897) | (-1.894) | (-1.901) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 143,813 | 143,813 | 143,813 | 143,811 | 146,203 | 146,203 | 146,203 | 146,201 |
| R-squared | 0.031 | 0.031 | 0.031 | 0.031 | 0.016 | 0.016 | 0.016 | 0.016 |

Panel B Associations of ETF Time Clustering

| | Stable Market | | | Volatile Market | | |
|---|---|---|---|---|---|---|
| | Price Impact | Price Efficiency | Price Volatility | Price Impact | Price Efficiency | Price Volatility |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 0.000115*** | 1.604*** | 0.000549*** | 0.000197*** | 1.602*** | 0.000833*** |
| | (15.95) | (394.8) | (52.71) | (15.24) | (420.5) | (15.00) |
| ETF TC | 1.46e-08*** | -5.03e-06*** | 1.93e-08*** | -2.80e-09 | -1.81e-06** | 1.08e-08 |
| | (3.934) | (-3.814) | (2.813) | (-0.591) | (-1.994) | (1.531) |
| Volume | 6.63e-11** | -3.64e-09 | 1.21e-09*** | 5.61e-11* | -7.41e-09** | 1.47e-09*** |
| | (2.553) | (-1.072) | (3.968) | (1.937) | (-2.592) | (4.590) |
| Size Volatility | -7.83e-10 | -1.41e-07 | -6.03e-09*** | 5.38e-10 | 8.96e-08 | -4.92e-09 |
| | (-1.475) | (-0.721) | (-3.181) | (0.506) | (0.732) | (-1.122) |
| SPY TC | 1.57e-06*** | -0.000604*** | 7.33e-06*** | 7.51e-07 | 0.000237 | -1.42e-05*** |
| | (3.881) | (-3.113) | (8.911) | (0.762) | (1.600) | (-8.625) |
| SPY Volume | 0** | -3.49e-09 | 2.53e-10*** | 0*** | -2.12e-09*** | 4.56e-10*** |
| | (2.188) | (-1.595) | (21.66) | (3.535) | (-3.353) | (20.91) |
| SPY Volatility | 0.00252 | 0.544 | 0.00921*** | 0.00401*** | 0.238* | 0.00186 |
| | (1.494) | (0.357) | (5.040) | (5.601) | (1.863) | (0.984) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 147,358 | 144,931 | 148,119 | 148,315 | 148,451 | 149,518 |
| R-squared | 0.013 | 0.012 | 0.159 | 0.011 | 0.024 | 0.268 |

Panel C Effects of ETF Time Clustering

| | Stable Market | | | | | Volatile Market | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Led ISO | Led Odd-lot | Led N. of Exc. | Led Repl. | Led Canc. | Led ISO | Led Odd-lot | Led N. of Exc. | Led Repl. | Led Canc. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Constant | 0.327*** | 0.0653*** | 7.732*** | 24.91*** | 131.2*** | 0.209*** | 0.0755*** | 8.254*** | 19.94*** | 53.13*** |
| | (87.95) | (27.23) | (166.1) | (20.63) | (30.49) | (55.68) | (27.78) | (128.6) | (10.56) | (6.775) |
| ETF TC | 2.16e-06* | -3.91e-06*** | 9.38e-05*** | -0.000386* | -0.00197** | 7.67e-07 | -1.84e-06*** | 4.61e-05*** | 0.000182 | -0.000636 |
| | (1.758) | (-4.463) | (3.873) | (-1.819) | (-2.470) | (1.057) | (-3.962) | (3.469) | (0.528) | (-0.619) |
| Led Volume | 6.88e-08*** | -1.28e-08 | 8.63e-07** | -1.83e-05*** | -7.56e-05*** | 1.16e-08 | 6.53e-09** | -8.45e-08 | -1.75e-05*** | -9.38e-05*** |
| | (3.721) | (-1.190) | (2.211) | (-3.802) | (-3.823) | (1.192) | (2.442) | (-0.662) | (-3.467) | (-3.441) |
| Led Size Volt. | 1.04e-06** | -1.48e-06*** | 6.21e-06 | 0.000103*** | 0.000406*** | 1.92e-06*** | -1.29e-06*** | 1.43e-05*** | 0.000174*** | 0.000810** |
| | (2.244) | (-2.886) | (1.493) | (2.973) | (2.923) | (3.995) | (-5.724) | (4.175) | (2.790) | (2.429) |
| Led SPY TC | 0.00187*** | -0.000901*** | 0.0362*** | 0.110*** | -0.0392 | 0.00126*** | -0.000497*** | 0.0279*** | -0.277*** | -1.385*** |
| | (12.58) | (-8.720) | (21.57) | (2.847) | (-0.256) | (8.776) | (-6.438) | (19.01) | (-4.199) | (-4.990) |
| Led SPY Volm. | 4.85e-09*** | -6.39e-09*** | 1.87e-07*** | 4.73e-06*** | 1.19e-05*** | 5.26e-09*** | -5.06e-09*** | 2.29e-07*** | 3.04e-06*** | 5.16e-06** |
| | (5.598) | (-8.266) | (18.53) | (7.375) | (6.540) | (8.924) | (-13.40) | (24.12) | (5.501) | (2.599) |
| Led SPY Volt. | -0.136 | 0.433 | 4.450 | 143.2 | -9.591 | 0.163* | 0.0261 | 3.982*** | -203.6*** | -624.6*** |
| | (-0.387) | (1.538) | (1.500) | (1.163) | (-0.0300) | (1.961) | (0.703) | (4.592) | (-5.503) | (-3.360) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 145,714 | 145,714 | 145,714 | 86,576 | 86,576 | 147,400 | 147,400 | 147,400 | 109,894 | 109,894 |
| R-squared | 0.237 | 0.283 | 0.727 | 0.298 | 0.277 | 0.350 | 0.335 | 0.718 | 0.352 | 0.358 |

Figure 1. Scenarios of time clustering.
All the total trading volume of the examples is 3,000 shares except the example in Chart D of which the total volume is 5,000 shares. Only the example in Chart A has duration of 5 minutes for each trade, and all the other examples have the same time distance of 1 minute for each trade.

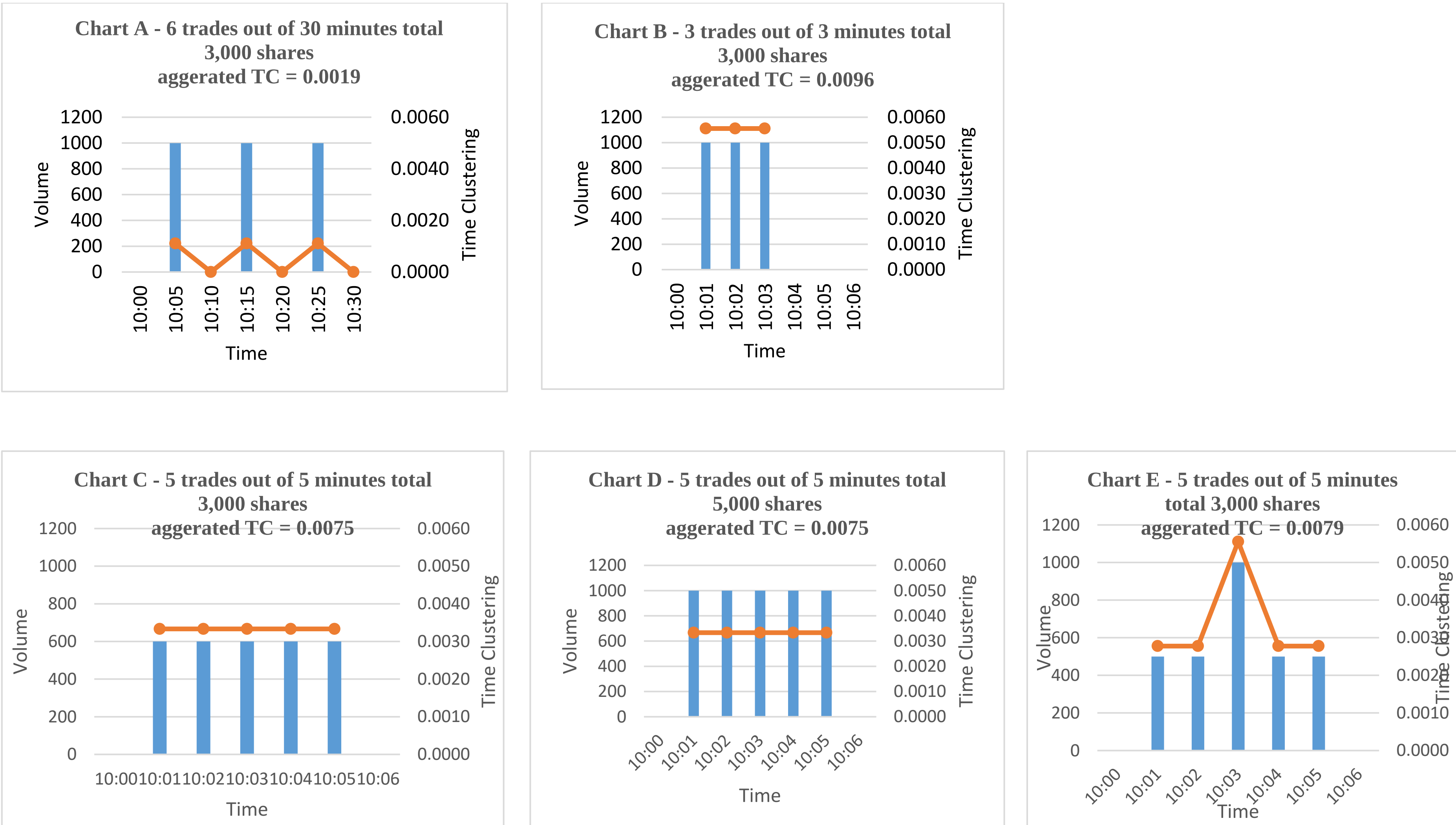

Figure 2. The timeline of determinants, associations, and effects of time clustering

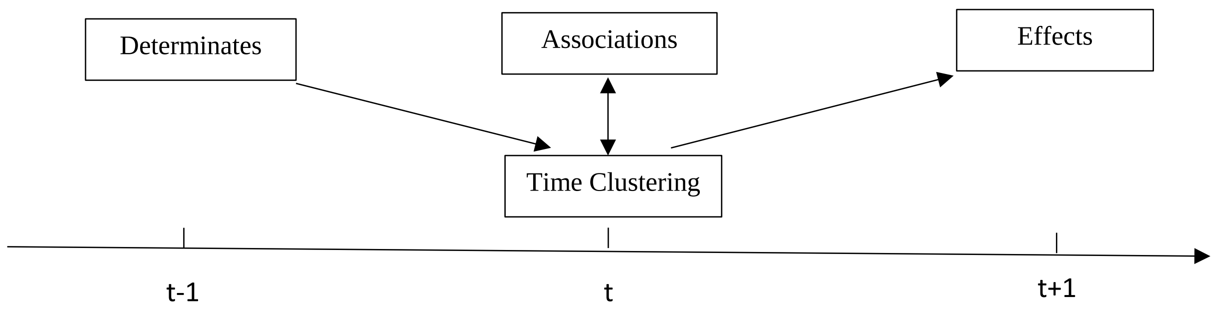

Figure 3. Intraday Pattern of Time Clustering

This figure presents the intraday pattern of time clustering. The axis label represents the serial numbers of 5-minute intervals from 9:30 to 16:00. The circle dots represent time clustering for year of 2017. The triangle dots represent time clustering for year of 2018.