

## **Trade Size Clustering and the Cost of Trading at the London Stock Exchange**

### **Abstract**

For the London Stock Exchange, this paper investigates differences in trading costs between market maker (off-book) and order book trades, in the context of clustering in trade sizes and prices. We report several substantial findings. Even after controlling for differences in trade size, the realised spread measure is lower for off-book trades. For the order book, trade size clustering is not associated with differences in transaction costs nor with differences in the information content of trades. For the off-book market, trades in clustered (popular) sizes carry significantly more information than non-clustered trades. Despite the significant differences in the price impact estimates between the order book and off-book, we show that traders placing large orders off-book are still better off than trading via the order book as they benefit from a large discount from the current midpoint price. Additionally, we highlight that price and size clustering tend to occur simultaneously rather than being substitutes in this market setting.

**Keywords:** Trade-size clustering, transaction costs, price impact, LSE

**JEL:** G20, G12

## 1. Introduction

The magnitude and significance of trade size clustering in financial markets has recently attracted attention (see Alexander and Peterson, 2007 and Moulton, 2005). Hodrick and Moulton (2009) demonstrate that liquidity has three dimensions; (1) price, (2) timing and (3) quantity. Therefore, investors need to focus not only on the price-time substitution effect, but also need to consider optimal traded quantities. That is, to the extent that investors fail to accommodate size in their optimal trade allocation decisions, their costs will increase. In particular, while previous studies focus on the extent of size clustering as well as the interaction of price clustering with size clustering, we extend this literature by examining the trade direction of size-clustered trades separately for liquidity demanders and liquidity providers. The latter allows us to study effective transaction costs and information flows separately for buy and sell orders and to establish to what extent the tendency of size clustering is linked to differences in market structure.

This paper focuses on trading costs at the London Stock Exchange (LSE). The unique structure of the LSE provides a further motivation for this study. In particular, although the main market at the LSE is a hybrid market, its market structure differs from the NYSE in relation to the interaction between market makers and order-book participants.<sup>1</sup> At the LSE, the downstairs market (the order book) is independent of the upstairs (off-book) market (involving market-makers), which implies that the market makers have no obligation to offer quotes on the order book. In addition, there is limited interaction between the upstairs market and the order book. Trades in the LSE upstairs market are privately negotiated and as a result there are no minimum tick restrictions. The latter also implies that the time priority rule is not

---

<sup>1</sup> At the NYSE, dealers are obliged to trade on the order book and are generally considered an integral part of the order book.

relevant, as upstairs market participants are able to front-run other investors by applying a very small incremental price improvement (see also Harris, 1991).

The upstairs market facilitates large trade sizes, as these trades would cause an adverse price effect in the order-book (see Bessembinder and Venkataraman, 2004).<sup>2</sup> When controlling for trade size differences, the information content of trades is greater for trades executed on the downstairs element of the London Stock Exchange Trading System (SETS) than for trades executed on the upstairs market (see Jain et al., 2010).<sup>3</sup> The latter finding is also confirmed by Gajewski and Gresse (2007) for the order books of Euronext Paris and the LSE, suggesting that informed trades are routed to the downstairs market. Further, Gajewski and Gresse (2007) show that small and medium-sized trades pay lower execution costs on the LSE order book than do same-size trades conducted in the upstairs market. We extend this literature in several respects. First, we provide evidence on trade execution costs net of the price impact component of trades, separately for the downstairs and the upstairs markets of the LSE. Related to this, we study execution costs separately for liquidity demanders and liquidity providers. Second, we examine execution cost differences on the LSE in light of the findings of Alexander and Peterson (2007) and Moulton (2005) that there are significant differences in execution and price impact costs between trades in popular sizes and those in non-clustered sizes. Hence, our primary aim is to study differences in transaction costs across different market structures with regard to their relationship to trade size clustering.

---

<sup>2</sup> Transparency is important in upstairs markets. Bernhardt et al. (2004) show that market participants that trade outside the order book tend to receive substantial discounts (price improvements) for larger trades when trading with a smaller number of dealers.

<sup>3</sup> SETS is the main electronic order book market in the LSE and the main trading platform for the most liquid securities.

Overall, we investigate three main hypotheses. First, we hypothesize that in a setting like SETS at the LSE, in which the minimum tick size is applied to the order book but not for the upstairs market, market participants in the latter market anchor their prices to the established minimum tick size applied on the order book. Second, we investigate the hypothesis that clustered trades (i.e. popular sizes) are associated with higher price impact than non-clustered trades. Third, we test the hypothesis that price clustering and size clustering are determined endogenously. A significant part of the contribution lies in the fact that we study the above hypotheses separately for liquidity providers and liquidity demanders

We use the effective half spread and its components as a measure of trading costs. We show that even after controlling for differences in trade size, the realised spread measure (which is net of the adverse selection costs) is lower for the off-book trades. Most importantly however, we investigate the level of trading costs and the informational content of trades that are associated with the optimal price/optimal size trading decisions. We do so in a setting that allows us to study market maker trades separately from non-exchange member trades. We document extensive size clustering in both markets, which persists when controlling for trade sizes. For the downstairs market, while the majority of assets trade in multiples of 500 shares, trade size clustering is not associated with differences in transaction costs nor with differences in the information content of trades. For the upstairs market, trades in the clustered (popular) sizes carry significantly more information than non-clustered trades. We also show that traders placing large orders via the upstairs market are still better off than trading via the order book. The latter finding holds mainly for buy orders and the benefits derived from trading in the upstairs market are greater for larger trade sizes.

In a second layer of analysis, we use a 2-stage least squares (2SLS) model to study whether price clustering and size clustering are substitutes or complements. We show that pricing in the upstairs market follows the notional minimum tick size rules (these are only binding for

the downstairs market), hence, prices cluster in multiples of the minimum tick size. The intraday distributions of price and size clustering exhibit a high correlation coefficient and this finding is validated with a multivariate model. We show that when trading frequency increases, market participants trade at clustered prices and clustered sizes, which supports the hypothesis that in busy trading periods, maximising liquidity comes at the expense of maximising utility from trading at exact prices and trade sizes. The latter is also confirmed at the end-of-quarter periods and is in contrast with the findings of Moulton (2005).

The remainder of this paper is organised as follows. Section 2 discusses the literature on size and price clustering. Section 3 outlines the data and Section 4 discusses the methodology. Section 5 presents the results and discusses the findings, and Section 6 concludes the paper.

## **2. Literature review and positioning of the paper**

In an ideal world, an investor is able to trade at the exact desired price, the exact desired size and at an exact time (see Harris, 2003, p. 398). Theoretically, Hodrick and Moulton (2009) demonstrate that deviations from the optimum trade price and optimum trade size impose shadow costs on portfolio managers. Moulton (2005) further argues that investors who are not able to optimize all three dimensions of liquidity need to consider the trade-offs regarding substituting the optimum price with the optimum size. To date, the literature is mainly concerned with deviations from the optimum price and offers little evidence on the costs faced by investors when there exist trade-offs in substituting optimum prices with optimum sizes.

Previous studies on price clustering have mainly focused on the motivation of market participants to concentrate their terms of trading in specific prices. Four theories have been widely cited as potential explanations for the clustering of final digits. Firstly, the negotiation theory (Harris, 1991) states that investors tend to reduce their terms of trading when there is

an increased need to execute trades. The negotiation hypothesis implies that price clustering will increase in periods of abnormally heavy trading. Secondly, the price resolution hypothesis (Ball et al., 1985) proposes that there is a difference between the “true” price of the asset and its observable price which is determined by the amount of information about that asset, its price level and variability. As a result, the greater the level of information, the larger the available price set should be. Thirdly, Goodhart and Curcio (1991) trace behavioural elements in price clustering, which might reflect a general tendency to use round numbers. Fourthly, Christie and Schultz (1994) and Christie, Harris and Schultz (1994) attribute price clustering at NASDAQ to the anticompetitive behaviour of dealers who implicitly colluded to increase market making revenues. Numerous empirical studies document price clustering in the equity, derivatives and foreign exchange markets (see Ahn et al., 2005, ap Gwilym and Verousis, 2010, Chung and Chiang, 2006, Narayan et al., 2011 and Sopranzetti and Datar, 2002).

A second stream of literature is generally consistent with the view that trade prices and trade sizes are determined simultaneously. Moulton (2005) shows that investors demonstrate an increased need to trade in exact quantities at year ends. Also, when investors choose to trade at exact quantities, the price impact of trades is greater, which reflects an increased need to trade at exact quantities rather than an increase in the available size set. The price-size substitution effect documented by Moulton (2005) for the foreign exchange market is not supported by the findings of Alexander and Peterson (2007) for stocks trading at the NYSE and NASDAQ markets. These authors show that price and size clustering are supplements hence occur simultaneously, which is consistent with the negotiation hypothesis of price clustering (Harris, 1991). The latter hypothesis is partially examined in ap Gwilym and Meng (2010), who report that size clustering in the FTSE100 futures market is inversely related to transaction frequency and is influenced by similar determinants to these identified in the price

clustering literature. Finally, Blau et al. (2012) show that short sellers' single aim is to exploit information inefficiencies, hence they are less concerned with the indirect negotiation and adverse price movement costs. The authors report that short selling positions demonstrate less size clustering and price clustering than non-short selling trades, which confirms the hypothesis that short sellers are more focused on private information than negotiation costs.

In light of this literature, we study three main themes. The first theme investigates to what extent differences in market structure affect price and size clustering. Previous literature has documented price clustering on the order book (see Ahn et al., 2005) and on a dealer market (see Christie and Schultz, 1994). However, the market structure of the LSE allows us to compare clustering of prices and sizes at downstairs and upstairs markets for the same assets. We hypothesize that in a setting like the main trading platform of the LSE, in which the minimum tick size is applied to the order book (SETS) but not for the upstairs market, market participants in the latter market anchor their prices to the established minimum tick size applied on the order book. Our second theme relates to the findings of Alexander and Peterson (2007) and Moulton (2005) regarding differences in the effective cost and price impact of clustered trades (i.e. in popular sizes) versus non-clustered trades. We investigate differences in trading costs within the different LSE market structures in relation to differences in trade size clustering. As in Alexander and Peterson (2007) and Moulton (2005), we hypothesize that clustered trades are associated with higher price impact than non-clustered trades.

The third theme is related to the trade-off between size and price clustering. We hypothesize that price clustering and size clustering are determined endogenously which will confirm previous findings relating to the dimensions of liquidity (see Hodrick and Moulton, 2009). The negotiation theory implies that if investors want to reduce execution costs, both price and size clustering will increase simultaneously. Also, price and size clustering will increase in

periods of high trading intensity. In contrast, Moulton (2005) shows that size clustering decreases at quarter ends. In a second level of analysis, we are able to test the above hypotheses, and an important part of the contribution lies in doing this separately for liquidity providers and liquidity demanders.

### **3. Market structure and data**

The data are obtained from the LSE historical data service and include all trades and quotes reported on the exchange in 2005. The microstructure of the main trading platform of the LSE comprises of SETS (a central order book) and the upstairs market (a network of market makers who provide dealer services on a voluntary basis and have no interaction with the order book). Under the LSE rules, the choice of trading venue rests entirely with the buy-side of the market (see Friederich and Payne, 2007). Thus, customers wishing to trade on SETS can choose whether to post a limit or market order on the order book or execute their order against the market makers' inventory in the upstairs market. The order book is fully transparent as customers are able to observe its full depth, nevertheless trades are anonymous. Trades executed off-book are privately negotiated, hence quotes are provided on a bilateral basis only and there is no minimum price increment. The microstructure of the LSE is different from the trading arrangements of the NYSE in which market makers are generally considered a part of the order book. On the NYSE, trades negotiated off-book are exposed on the order book, while market makers' prices are constrained by the order book.

We identify downstairs (order-book) and upstairs (market maker) trades. The minimum price increment is a function of price level and varies across different market segments. All cancelled, zero-volume, zero-price, out-of-hours and out-of-date trades are removed. In order to detect outliers, we delete all trades that report either a price change of greater than 25% or a price which differs from the present midquote by more than 25% (Bessembinder, 1997). In



order to avoid stale pricing and missing data problems, we select firms that report at least one trade per hour for the sample period, separately for the upstairs and the downstairs markets. We focus on the minimum tick sizes of xx.25p and xx.50p. The resulting sample of 36 assets consists of firms trading on the main Stock Exchange Electronic Trading Service (SET1 and SET2).<sup>4</sup>

## 4. Methods

### 4.1. The magnitude of size and price clustering

In order to identify the extent of size clustering, we regress the percentage of trades conducted in size  $i$  against size dummy variables for the most popular sizes. We calculate the percentage of orders traded at lots of 100 shares as equal-weighted averages across all stocks in the sample and then regress the percentage of orders traded at each lot against certain size multiples (see Alexander and Peterson, 2007):

$$\text{LnSizePer}_v = \alpha + \beta_1 D_{500,v} + \beta_2 D_{1000,v} + \beta_3 D_{5000,v} + \beta_4 \text{LnSize}_v + \varepsilon_v . \quad (1)$$

$\text{LnSizePer}_v$  is the percentage of orders traded at size  $v$ , separately for the upstairs and the downstairs markets.  $D_{500,v}$ ,  $D_{1000,v}$  and  $D_{5000,v}$  are dummy variables that equal 1 if the trade size  $v$  is a multiple of 500, 1000 and 5000 shares respectively.  $\text{LnSize}_v$  is the natural logarithm of trade size  $v$ .

---

<sup>4</sup> LSE opens at 8:00 and closes at 16:30. We form 8 hourly intervals (hours starting 8:00 – 15:00). The remaining interval is 30 minutes. 250 trading days are identified. In total, 41 assets met the selection criterion, but 5 assets were dropped from the sample as they were the only assets traded with a minimum tick of one pound.

Since trades conducted at the upstairs market do not have to adhere to the minimum tick size restrictions, we test the significance of price clustering in the upstairs market as follows:

$$\text{LnPricePer}_z = \alpha + \beta_1 D_{0,z} + \beta_2 D_{0.25,z} + \beta_3 D_{0.50,z} + \beta_4 D_{0.75,z} + \beta_5 \text{LnPrice}_z + \varepsilon_z . \quad (2)$$

$\text{LnPricePer}_z$  is the percentage of orders traded at a decimal price  $z$  on the upstairs market.  $D_{0,z}$ ,  $D_{0.25,z}$ ,  $D_{0.50,z}$  and  $D_{0.75,z}$  are dummy variables that equal 1 if the decimal in the traded price  $z$  is 0, 0.25, 0.50 and 0.75 respectively.  $\text{LnPrice}_z$  is the natural logarithm of price  $z$ .

For the classification of sell and buy orders, we employ the tick trade identification algorithm (see Finucane, 2000).<sup>5</sup> Trades that are conducted above the previous traded price are classified as buys, while trades that are conducted below are classified as sells (both taking the investor perspective). For the upstairs market, LSE Rule 3012.2 instructs market makers to report trades when dealing with investors (i.e. non exchange members). Hence, for the upstairs market, assuming that investors are always the trade initiators (see Friederich and Payne, 2007), trades that are classified as buys (under the above method) are effectively buy orders from the perspective of the liquidity provider (the market makers) and sell orders from the perspective of the liquidity demander (non-market maker participants). In order to study trades from the perspective of the liquidity demander, we reverse the trade sign for the upstairs market.<sup>6</sup> For the downstairs market, the trade is classified from the perspective of the trade initiator i.e. the liquidity demander.

---

<sup>5</sup> In contrast to microstructure studies on the LSE using data from an earlier time period (see Jain et al., 2010), the trade sign is not provided by the exchange in datasets for more recent years.

<sup>6</sup> Often, in trades conducted between two market makers, it is the seller who reports the trade. However, market maker-to-market maker trades, which account for less than 1% of the total sample, are deleted from the sample, hence this potential problem does not apply to the current study.

Alexander and Peterson (2007) report that size clustering increases with trade size. In order to control for these differences, we test for size clustering across trade sizes, using binomial z-tests for the equality of proportions below and above multiples of 500 shares. For example, for trades of 500 shares, we select trades conducted at 400 and 600 shares and test for the equality of proportions. Alexander and Peterson (2007) iterate this method up to 20,000 shares. However, we restrict our calculations to up to 10,000 shares, as this captures more than 83% of the sample and guarantees increased trading frequency around trade multiples.

#### 4.2. Size clustering and trading costs

We calculate the effective half-spread (*EHS*) as a consistent transaction cost measure that also accounts for trading inside the quotes (Bessembinder and Kaufman, 1997):<sup>7</sup>

$$EHS_{i,tick} = 100D_{i,tick}(P_{i,tick} - M_{i,tick})/M_{i,tick} . \quad (3)$$

$EHS_{i,tick}$  is the effective half-spread measure per asset  $i$ , calculated for all trades at irregular time intervals.  $D_{i,tick}$  equals one for buy orders and negative one for sell orders.  $P_{i,tick}$  is the trade price and  $M_{i,tick}$  is the prevailing midpoint at the time of the trade. For the downstairs trades, the *EHS* is calculated based on the most recent quotes. However, for the upstairs market, as LSE rules allow market makers to report trades within three minutes of their execution, we calculate the *EHS* by selecting the most recent quotes that have been posted at least three minutes before the trade and not more than sixty minutes prior to the trade.

---

<sup>7</sup> The *EHS* is always positive for buys (sells) conducted above (below) the midprice. However, for the upstairs market, we expect the *EHS* to be negative as the trade sign is reversed. In the latter cases, the more negative the *EHS*, the smaller the implied cost of trade.

In order to control for trade size differences, we calculate test statistics for the equality of means using trade-size levels. Bessembinder (1997) notes that the *EHS* may give inconclusive results in a situation where market makers widen spreads in reaction to trades which potentially contain private information. In order to assess this possibility, we decompose the *EHS* into its informative component and the market cost component. We measure price impact (*PI*), the component of the spread that accounts for superior information, as follows:<sup>8</sup>

$$PI_{i,tick} = 100D_{i,tick}(M_{i,tick+5} - M_{i,tick})/M_{i,tick} . \quad (4)$$

Similar to Alexander and Peterson (2007), we measure *PI* as the difference in quote midpoints between the current trade and first trade reported at least 5 minutes after the announcement of the trade. As with the calculation of the *EHS*, for upstairs trades  $M_{i,tick}$  denotes the midpoint that has been posted at least three minutes before the trade and not more than sixty minutes prior to the trade. Finally, the realised spread (*RHS*) component is measured as follows:<sup>9</sup>

$$RHS_{i,tick} = EHS_{i,tick} - PI_{i,tick} . \quad (5)$$

---

<sup>8</sup> Irrespective of considering the liquidity demander or supplier, *PI* reflects the opportunity cost for conducting a trade, hence the more positive the *PI* the greater the informational content of the trade and the greater the benefit from conducting the trade.

<sup>9</sup> The more negative the *RHS*, the greater the benefit derived for the investor, irrespective of the trade classification, as it implies that either the *EHS* is smaller than the *PI* or that the *EHS* is more negative than the negative *PI*.

The *RHS* measure is thus a more appropriate measure of market making revenues as it is net of losses to informed traders (see Bessembinder, 1997). EHS, PI and RHS estimates are subsequently aggregated to equal-weighted averages for each day,  $t$ .

#### 4.3. Size and price clustering: supplements or substitutes?

We estimate a 2SLS model to test the hypothesis that when investors are faced with the problem of optimising the price and size of their trades, they do so simultaneously. In order to investigate the outcome for the above hypothesis in the upstairs and the downstairs market, we estimate separate regressions for each market structure. The fact that market makers are not obliged to adhere to the minimum tick restrictions necessitates the use of different definitions of price clustering for these markets. We define price clustering on the order-book as the occurrence of trades in integers, while in the upstairs market, prices are defined as clustered when reported at multiples of the minimum tick size. We use the following 2SLS model:

$$SzClst_{i,t} = a_1 + \beta_1 PrClst_{i,t} + \beta_2 Frq_{i,t} + \beta_3 AR_{i,t} + \beta_4 LnTrdSz_{i,t} + \beta_5 LnMV_{i,t} + \beta_6 OEQ_i + \varepsilon_{i,t}. \quad (6)$$

$$PrClst_{i,t} = a_1 + \beta_1 SzClst_{i,t} + \beta_2 SQRTFrq_{i,t} + \beta_3 AR_{i,t} + \beta_4 LnTrdSz_{i,t} + \beta_5 LnMV_{i,t} + \beta_6 PrcLvl_{i,t} + \beta_7 Tck_i + \varepsilon_{i,t}$$

(7)

*SzClst* is the percentage of trades conducted at multiples of 500 shares and *PrClst* is the percentage of trades conducted at integers (downstairs market) or at multiples of the

minimum tick size (upstairs market). *Frq* is the daily number of trades for each asset, separately for each market structure. Moulton (2005) argues that the total number of trades per interval reflects increased liquidity needs, thus we expect a positive sign on *Frq*. *SQRTFrq* is the inverse square root of the daily number of trades. Harris (1991) shows that transaction frequency is positively related to price clustering, hence, the expected sign for *SQRTFrq* is negative.<sup>10</sup> *AR* is the average absolute daily return calculated over hourly intervals. In order to control for possible overnight return jumps, we exclude the open return. In several studies, price clustering is found to increase during high volatility periods (see Ahn et al., 2005 and ap Gwilym et al., 1998). Also, Moulton (2005) and ap Gwilym and Meng (2010) show that volatility is also a positive determinant for size clustering.

*LnTrdSz* denotes the natural logarithm of the average daily trade size for each asset (separately for order-book and market maker trades). Price clustering has been found to be positively correlated with trade size (Harris, 1991), implying that since the economic impact of the minimum tick size is greater for small trades, investors will use a coarser set of prices for large trades. *OEQ* is a binary dummy variable that takes the value of 1 when a stock trade occurs within the last two weeks of a calendar quarter end. Moulton (2005) shows that size clustering decreases at quarter-end periods which reflects the internal monitoring of trading firms. *LnMV* is the natural logarithm of market capitalisation for each firm.<sup>11</sup> Price clustering is expected to decrease with firm size (Harris, 1991). *PrcLvl* denotes the average daily price level for each asset and market segment. Ahn et al (2005) suggest that *PrcLvl* is positively related to price clustering and Harris (1991) suggests that a positive sign is expected because minimum tick size rules are more restrictive for lower-priced stocks. Finally, *Tck* is a control variable that takes the value of 1 for stocks with a minimum tick of xx.50p and 0 for stocks trading at xx.25p.

---

<sup>10</sup> In Moulton (2005), this variable is not significant for FX transactions.

<sup>11</sup> Data on firm values are obtained from DataStream.

## 5. Results

### 5.1. Descriptive statistics

Table 1 presents the descriptive statistics. As expected, assets with the smaller tick size trade at relatively lower prices, with small differences in average price between the upstairs and downstairs markets. The role of the upstairs market as a facilitator of block trades becomes apparent when looking at the size columns across both tick sizes. For the smaller tick size (.25p), mean daily trade size significantly differs from 11,770.95 for order-book trades to 43,518.56 for off-book trades. For the .50p tick size, average price levels are at approximately 760p compared to 280p for the smaller tick, whereas the equal-weighted average daily trade size is 3,739.36 for order-book and 21,232.23 for off-book trades. The distribution of trades at each price category shows that more than one-fifth of off-book trades are conducted at the implicit minimum tick. Finally, Harris (1991) documents that price clustering increases with an increasing price level and in Table 1 we show that the mean price varies with the final integer. In particular, both mean price and size measures increase at whole digits, thus reflecting the use of a smaller set of decimal prices for higher priced assets and larger size trades.

\*\*\*Insert Table 1 about here\*\*\*

## 5.2. Results for the magnitude of size and price clustering

### 5.2.1. The distribution of trade size and trade price

Figure 1 presents the distribution of trade size for order-book (Panel A) and market maker trades (Panel B).<sup>12</sup> Size frequency is a decreasing function of trade size and tends to be less concentrated for the downstairs market. Three distinctive sizes emerge: multiples of 500, 1000 and 5000 shares. Trades of 1000 shares are the most popular size. The size distributions for the downstairs and upstairs markets are investigated using Equation 1 and the results are presented in Table 2.

\*\*\*Insert Figure 1 about here\*\*\*

Table 2 (Panel A) verifies the findings of Figure 1 that trade sizes are clustered at multiples of 500, 1000 and 5000 shares. Also, the coefficient estimates for order-book and market maker trades are very similar and show strong clustering at multiples of 500 shares. Hence, for the remainder of the paper we define size clustering on the basis of multiples of 500 shares which enables investigation of the association between size clustering and price clustering.

Table 2 (Panel B) also presents the results for the price clustering distribution for market maker trades (Equation 2).<sup>13</sup> Recall that the upstairs market at the LSE has no minimum tick size rule. As hypothesized, there is a strong clustering of traded prices at the notional ticks of

---

<sup>12</sup> We present the distribution of trade size only up to 10,000 shares in order to conserve space.

<sup>13</sup> For the downstairs market, prices can only take up to four decimal values.



.25p, .50p, .75p and integers. The coefficients for the price multiples are all highly significant and price clustering increases for higher price levels.

\*\*\*Insert Table 2 about here\*\*\*

### 5.2.2. The impact of trade size on size clustering

Table 3 controls for the inverse relation between size and size frequency identified in Table 2. We focus on the distribution of size clustering around multiples of 500 shares. We test for the equality of differences between multiples of 500 shares and their close round equivalents (for example 900 and 1100 against 1000 shares).<sup>14</sup> Table 3 shows that even when controlling for differences in trade size, size clustering is highly statistically significant.

\*\*\*Insert Table 3 about here\*\*\*

In testing for size clustering differences between the downstairs and upstairs markets, we confirm that there is a significant difference in the distributions of sizes across these markets, which holds for both buy and sell orders. The latter is of particular importance, as buy and sell off-book trades reflect the direct negotiation of trades with market makers in the upstairs market. Despite controlling for trade size, we find that trading in the upstairs market is concentrated in a smaller set of sizes.

---

<sup>14</sup> That is, we employ the z-score in order to test the null that the relative frequency of clustered sizes is less than or equal to 1/3 of the overall trading frequency around multiples of clustered trades (e.g. around multiples of 500 shares, see also Alexander and Peterson, 2007). We conduct individual tests for each multiple of 500 shares and its counterparts and for each asset. The results are then aggregated to determine a z-score for each asset and a z-score for all assets.

### 5.3. Results for size clustering and trading costs

#### 5.3.1. Effective half spread

This section studies the hypothesis that clustered trades (those in popular sizes) are associated with greater price impact than non-clustered trades. Table 4 presents trading costs for a single trip trade, computed separately for clustered sizes (multiples of 500, denoted  $S_zClst$ ) and non-clustered sizes ( $S_z$ ) for the downstairs and upstairs markets. We also present the results separately for buy and sell orders to capture overall trade costs in both market structures.

\*\*\*Insert Table 4 about here\*\*\*

For the downstairs market, the EHS estimates for each asset show that most assets demonstrate a significant difference in the transaction cost estimates between clustered and non-clustered trades. Nevertheless, no clear pattern emerges and the difference becomes insignificant at the aggregate level. Overall, the EHS is similar for clustered and non-clustered trades for both buys and sells. For the downstairs market, on a per asset basis, buy (sell) orders tend to be filled below (above) the current midprice, resulting in smaller execution costs for assets trading off-book.<sup>15</sup>

For the upstairs market buy orders, 16 of 36 assets exhibit a statistically significant difference in EHS between clustered and non-clustered trades. In aggregate however, there is no statistically significant difference between clustered and non-clustered buy trades. For the sell orders however, a clear pattern emerges whereby 32 assets exhibit lower transaction costs for

---

<sup>15</sup> As both markets are investigated from the perspective of the liquidity demander, the results are comparable across market structures.

clustered trades than non-clustered trades. In total, a statistically significant difference of 3.7 basis points exists between sell orders in popular and non-popular sizes.

### 5.3.2. Price impact

We further investigate the above findings by decomposing the spread into its components, PI and RHS. Table 5 presents PI estimates that account for the price reaction to trading activity and reflect adverse selection costs. For the downstairs market, on the buy (sell) side, 19 (23) assets exhibit a positive and statistically significant difference for PI between clustered trades and non-clustered trades. This is consistent with the findings of Alexander and Peterson (2007) that trades in popular sizes tend to convey more information than unrounded trades. Overall however, the results for the downstairs market show that sell orders have a significant PI difference between clustered and non-clustered trades at the 10% level. For the upstairs market, the PI of non-clustered trades is often negative and overall clustered trades also exhibit greater information content than non-clustered trades. This difference is a statistically significant 2.7 basis points for buy orders.

\*\*\*Insert Table 5 about here\*\*\*

### 5.3.3. Realized half-spread

Table 6 presents the results for the RHS component of the EHS. For the downstairs market, even though the (overall) results are significant at the 10% level for sell orders, no clear pattern emerges. For the upstairs market, differences in the RHS between clustered and non-clustered trades are significant for 68 of 72 cases. This is an expected outcome because (as demonstrated above) trading off-book is advantageous in terms of trading further away from the current midprice and also in terms of off-book trades exhibiting larger price impact.

Regarding the effect of size clustering, clustered trades demonstrate lower RHS and buy (sell) orders show that on average a statistically significant difference of 3.5 (4.5) basis points exists.

\*\*\*Insert Table 6 about here\*\*\*

#### 5.3.4. Summary of results for size clustering and trading costs

Tables 4-6 show that, for the downstairs market, several significant differences exist between clustered (popular sizes) and non-clustered buy and sell orders. In aggregate, the significance either generally disappears or is not strong enough to imply a systematic pattern. It is worth noting that orders in popular sizes carry more information than non-clustered orders (albeit at the 10% significance level).

For the upstairs market however, an interesting and clear pattern emerges. We show that traders submitting orders off-book generally receive better execution costs than in the order-book market. Clustered trades executed in the upstairs market exhibit, in aggregate, lower costs than non-clustered trades. However, the results for buyer-initiated trades are not statistically significant.

Decomposing the spread demonstrates that clustered buy orders exhibit a statistically significant difference in their PI estimates of 2.7%. For the sell orders, while the positive difference in the PI remains, the results are not significant. This finding is important as it shows that while buy orders in popular sizes are not associated with greater execution costs, they tend to carry greater information content. On the other hand, when investors are on the sell side, trades in popular sizes resemble liquidity orders (or orders with limited speculative nature) as they are generally not associated with a statistically significant price impact.

Finally, the RHS estimates show that investors are better off trading upstairs because the combination of low execution costs with a larger price impact results in negative RHS estimates. Net transaction costs are significantly higher for non-clustered trades than clustered trades, a finding that holds for both buy and sell orders. When combining the results of the PI with the RHS estimates, we show that when investors trade at multiples of 500 shares, these trades tend to reveal more information, confirming our hypothesis. However, investors are better off trading at the upstairs market because they receive better execution costs even after controlling for the price impact.

### 5.3.5. The effect of trade size on trading cost estimates

Table 7 considers the effect that trade size variability may have on the cost and price impact findings. We report the aggregate results in three different trade size categories: 1 to 1500 shares, 1501 to 3000 and 3001 or more shares.<sup>16</sup> With regard to the downstairs market, one important finding is that spread and price impact measures are relatively constant across all trade size categories.<sup>17</sup> Also, as expected, there is no statistically significant difference between clustered and non-clustered trades for buy or sell orders in any size category.

\*\*\*Insert Table 7 about here\*\*\*

---

<sup>16</sup> 50% of downstairs (upstairs) trades are below 1000 (2450) shares and two-thirds of downstairs (upstairs) trades are below 3000 (6700) shares. We arbitrarily use the categories above that capture the trade size effect and also, at least partially, offer a comparison of sizes between the downstairs and upstairs markets. We present the aggregate results only in order to conserve space.

<sup>17</sup> Overall however, PI increases with increasing trade size levels (see also Jiang et al., 2009 for NYSE and NASDAQ stocks)

For the upstairs market buy trades for trade sizes up to 1500 shares, the spread and price impact estimates remain at similar levels for buy and sell orders. Nevertheless, for larger trades, both components of the EHS increase substantially. For trades up to 3000 shares, there is no statistically significant difference for RHS between clustered and non-clustered trades. For orders over 3000 shares, clustered trades exhibit smaller RHS than non-clustered trades, a finding that is supported by a high PI component for the clustered trades. For sell orders, the RHS is significantly smaller for clustered trades than non-clustered trades for each trade size category, and as with buy orders, it becomes relatively large for the larger trade sizes. Nevertheless, the PI component of the spread remains statistically similar between sell orders conducted in popular and less-popular sizes. It follows that the decision to sell at popular sizes is not associated with differences in the information content of trades.

#### 5.4. Results for size and price clustering: supplements or substitutes?

##### 5.4.1. Two-stage least squares

For the third and final theme, we investigate the hypothesis that a substitution effect may exist between price and size clustering. An indication of the relationship between price and size clustering is offered in Figure 2. It is clear that both the upstairs and downstairs markets indicate a strong correlation between price and size clustering. The correlation coefficient (for intraday intervals) between size clustering and integer price clustering is 0.90 for the downstairs market and 0.82 for the upstairs market.

\*\*\*Insert Figure 2 about here\*\*\*

Table 8 presents the results of the Hausman test for endogeneity between price and size clustering using Equations 6 and 7. Results for both markets are highly significant; hence, a 2SLS model is implemented to account for the endogeneity between price and size clustering.<sup>18</sup>

\*\*\*Insert Table 8 about here\*\*\*

Table 9 presents the 2SLS regression estimates. The results for both markets show that trading at exact prices has a positive association with trading at exact sizes. This is consistent with the findings of Alexander and Peterson (2007). In the upstairs market, in which negotiations are more direct, it is clear that a simultaneous (or complementary) effect between price and size clustering exists, hence the positive signs for both price and size clustering variables. However, for the downstairs market, it is shown that when investors increase trading at exact sizes they tend to use a wider set of prices which compensates the loss of optimality.

\*\*\*Insert Table 9 about here\*\*\*

The results from the frequency variables are consistent with the previous literature, though only significant for the downstairs trades. The latter is inconsistent with the negotiation theory in price clustering (Harris, 1991) and the information-arrival hypothesis in size clustering (ap Gwilym and Meng, 2010). However, as ap Gwilym et al. (1998) and Alexander and Peterson (2007) indicate, trading in busy periods leads to the use of a smaller trade size

---

<sup>18</sup> For reasons of brevity we only present the relevant coefficients.

and trade price set. There is no such relationship for the upstairs market, which may provide an indication of its more fragmented nature.

#### 5.4.2. Other determinants of size and price clustering

Table 9 also documents the determinants of size and price clustering outside the price, size and time dimensions discussed by Hodrick and Moulton (2009). In particular, we show that several strong similarities exist between the upstairs and the downstairs market with regard to the decision to trade at exact sizes. Size clustering increases with the volatility of returns and with trade size, and decreases with market value. The results for the end-of-quarter dummy are also positive and marginally significant which is in contrast to the findings of Moulton (2005). This supports the hypothesis that investors at the end-of-quarter trading periods choose to focus on liquidity at the expense of exact trade prices and trade sizes. These results support the negotiation hypothesis for size and price clustering.

The results on price clustering determinants show that price clustering increases during volatile periods in the downstairs market but not in the upstairs market. For the downstairs market, we verify that price clustering increases with the trade size and decreases for higher valued firms, as the negotiation theory implies. For the upstairs market, we show that larger trades are more accurately priced. The results on the price level are consistent with the previous literature and are similar for both market structures (see also Ahn et al., 2005). Finally, we show that trading at integers is more prevalent for the larger tick size at the downstairs market, while the opposite is true for the upstairs market. The latter result indicates that participants deviate from trading at the implicit minimum tick when it becomes too costly to trade i.e. when it is relatively large.



## 6. Conclusions

Hodrick and Moulton (2009) show that if portfolio managers concentrate on the price-time substitution effect and omit the effects of trade size, they will impose additional costs on their portfolio allocation strategies. Despite the importance of the trade size (quantity) dimension, it is only recently that studies have begun to explore the full extent of the size-price relationship in financial markets. Alexander and Peterson (2007) show that price and size clustering occur simultaneously, which is consistent with the trade size negotiation cost hypothesis, similar to the price clustering negotiation hypothesis (Harris, 1991).

This paper explores size and price clustering in the upstairs (market maker) and downstairs (order book) markets of the LSE. This exchange offers an optimal setting for capturing the trading arrangements of market makers, given that the upstairs market operates independently from the downstairs market and market makers have no obligation to facilitate trading on the order book. Using a trade identification algorithm, we are able to document the size and price clustering patterns on the buy and sell sides.

We show that there is extensive size clustering in both the downstairs and upstairs markets of the LSE, which is expressed in multiples of 500 shares and shows strong resemblance to trade price clustering. We show that even after controlling for the trade size effect, the realised spread measure, which is net of adverse selection costs, is lower for off-book trades. We also document a significant difference in realized costs between trades in clustered (popular) and non-clustered sizes in both the upstairs and downstairs markets. For the downstairs market, while the majority of assets are trading in multiples of 500 shares, trade size clustering is not associated with differences in realized costs nor with differences in the information content of trades. For the upstairs market, trades in popular sizes carry significantly more information than trades in non-clustered sizes. However, we show that traders executing large orders via the upstairs market are still better off than trading via the order book. The latter finding holds

mainly for buy orders and the benefits derived from trading in the upstairs market increase for larger trade sizes.

Finally, we study the interaction of price clustering and trade size clustering and identify their determinants for each market structure. We confirm the hypothesis that trade size and trade price are determined endogenously. In particular, we report that in the upstairs market, price and size clustering complement each other rather than being substitutes, after controlling for transaction frequency. For the downstairs market, trade size clustering is also a positive function of the time and price dimensions of liquidity but price clustering follows a different pattern. At the quarter-end periods, market participants choose to focus on liquidity at the expense of exact trade sizes, which contrasts with Moulton (2005). For the downstairs market, the decision to trade at exact quantities and/or exact prices is largely in line with the negotiation hypothesis (Harris, 1991). For the upstairs market, the latter implication only holds for the decision to trade at exact trade sizes.

Our analysis of the upstairs and downstairs markets strongly confirms previous findings (Moulton, 2005) that liquidity entails three dimensions: namely, price, size and time. Most importantly, our evidence suggests that trade size clustering is strongly associated with the price impact measures. Market participants that fail to account for the information content of trades at multiples of 500 units are exposed to high adverse selection costs.

## References

- Ahn, H., Cai, J. & Cheung, Y. L. (2005). Price Clustering on the Limit Order Book: Evidence from the Stock Exchange of Hong Kong. *Journal of Financial Markets* 8, 421-451.
- 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.
- ap Gwilym, O. & Meng, L. (2010). Size Clustering in the FTSE100 Index Futures Market. *Journal of Futures Markets* 30, 432-443.
- ap Gwilym, O., Clare, A. & Thomas, S. (1998). Extreme Price Clustering in the London Equity Index Futures and Options Markets. *Journal of Banking & Finance* 22, 1193-1206.
- ap Gwilym, O. and Verousis, T. (2010). Price Clustering and Underpricing in the IPO Aftermarket. *International Review of Financial Analysis* 19, 89-97.
- Ball, C., Torous, W. & Tschoegl, A. (1985). The Degree of Price Resolution - The Case of the Gold Market. *Journal of Futures Markets* 5, 29-43.
- Bernhardt D., Dvoracek, V., Hughson, E. & Werner, I M. (2004). Why do larger orders receive discounts on the London Stock Exchanges? *Review of Financial Studies* 18, 1343-1368.
- Bessembinder, H. (1997). The Degree of Price Resolution and Equity Trading Costs. *Journal of Financial Economics* 45, 9-34.
- 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.
- Bessembinder, H. & Venkataraman, K. (2004). Does an Electronic Stock Exchange Need an Upstairs Market?. *Journal of Financial Economics* 73, 3-36.
- Blau, B. M., Van Ness, B. F. & Van Ness, R. A. (2012). Trade-Size and Price Clustering: the Case of Short Sales and the Suspension of Price Tests. *Journal of Financial Research* 35, 159-182.

Christie, W.G. & Schultz, P.H. (1994). Who do Nasdaq Market Makers Avoid Odd-Eighth Quotes?. *Journal of Finance* 49, 1813-1840.

Christie, W.G., Harris, J.H. & Schultz, P.H. (1994). Why did Nasdaq Market Makers Stop Avoiding Odd-Eighth Quotes?. *Journal of Finance* 49, 1841-1860.

Chung, H. & Chiang, S. (2006). Price Clustering in E-Mini and Floor-Traded Index Futures. *Journal of Futures Markets* 26, 269-295.

Finucane, T. J. (2000). A Direct Test of Methods for Inferring Trade Direction from Intra-Day Data. *Journal of Financial and Quantitative Analysis* 35, 553-576.

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

Gajewski, J.F. & Gresse, C. (2007). Centralized Order Books Versus Hybrid Order Books: A Paired Comparison of Trading Costs on NSC (Euronext Paris) and SETS (London Stock Exchange). *Journal of Banking and Finance* 31, 2906-2924.

Goodhart, C. & Curcio, R. (1991). The Clustering of Bid/Ask Prices and the Spread in the Foreign Exchange Market. London School of Economics, Financial Market Group, Discussion Paper 110.

Harris, L. (2003). *Trading and Exchanges: Market Microstructure for Practitioners*. New York: Oxford University Press.

Harris, L. (1991). Stock Price Clustering and Discreteness. *Review of Financial Studies* 4, 389-415.

Hodrick, L. S. & Moulton, P. (2009). Liquidity: Considerations of a Portfolio Manager. *Financial Management* 38, 59-74.

Jain, P., Jiang, C., McInish, T. & Taechapiroontong, N. (2010). Informed Trading in Parallel Auction and Dealer Markets: The Case of the London Stock Exchange in Greg N. Gregoriou ed., *The Handbook of Trading: Strategies for Navigating and Profiting from Currency, Bond, and Stock Markets*: McGraw-Hill, NYC, NY, 2010, Ch. 2, 23-38.

Jiang, C., Kim, J-C. & Wood, R. (2009). Adverse Selection Costs for NASDAQ and NYSE after Decimalisation. *International Review of Financial Analysis* 18, 205-211.

Moulton, P. (2005). You Can't Always Get What You Want: Trade Size Clustering and Quantity Choice in Liquidity. *Journal of Financial Economics* 78, 89-119.

Narayan, P. K., Narayan, S., Popp, S. & D’Rosario, M. (2011). Share Price Clustering in Mexico. *International Review of Financial Analysis* 20, 113-119.

Sopranzetti, B. & Datar, V. (2002). Price Clustering in Foreign Exchange Spot Markets. *Journal of Financial Markets* 5, 411-417.

Table 1: Descriptive statistics								
	Assets with minimum tick xx.25				Assets with minimum tick xx.50			
Market	Downstairs		Upstairs		Downstairs		Upstairs	
	Price	Size	Price	Size	Price	Size	Price	Size
Panel A: Integers								
Perc.	30.9		6.4		56.7		12.2	
Mean	282.9	13475.5	282.5	165207.8	765.9	4070.3	765.9	61045.6
SD	23.3	25947.7	41.9	1030609.7	47.4	6693.6	111.2	547139.0
Panel B: xx.50								
Perc.	27.4		5.3		43.3		9.2	
Mean	282.9	12110.7	282.2	121038.1	744.0	3499.6	751.2	40057.9
SD	23.2	22845.2	23.1	921929.7	36.9	5686.8	45.7	353255.7
Panel C: xx.25								
Perc.	20.4		4.2		.		2.0	
Mean	277.1	11099.4	278.9	81768.0	.	.	758.7	25698.6
SD	21.2	20958.0	22.4	404655.6	.	.	47.4	184315.6
Panel D: xx.75								
Perc.	21.28		4.2		.		1.77	
Mean	277.1	10998.1	278.4	88910.5	.	.	757.0	68247.3
SD	21.3	20846.6	22.4	506031.7	.	.	48.8	1522364.3
Panel E: Totals								
Days	250		250		250		250	
Perc.	82.3		17.7		85.0		15.0	
Price/Size	280.6	11770.9	280.6	43518.6	760.3	3739.4	760.4	21232.2
Buy (Sell)	51.2 (48.8)		48.1 (51.9)		51.2 (48.8)		48.8 (51.2)	
Perc. denotes the proportion of trades at each price category. For the downstairs market, Perc across Panels A-D add up to 100%. For the upstairs market, the sum of Perc across Panels A-D is always less than 100 and the difference denotes the percentage that trades in other end-digits. Mean and SD denote the average mean price / trade size and standard deviation estimates per asset for each sub-sample respectively, estimated as an equal-weighted average at each sub-sample. Price and size denote the equal-								

weighted average price and traded size for each sub-sample respectively.

Table 2: Size and price clustering regression estimates							
Panel A: Size clustering							
	Coefficients						
Market	Constant	D <sub>500</sub>	D <sub>1000</sub>	D <sub>5000</sub>	.	LnSize	Adj R <sup>2</sup>
A. Downstairs	2.709*** (1529.5)	3.322*** (1499.4)	1.017*** (382.1)	1.929*** (922.7)	.	-0.968*** (-4239.5)	0.70
B. Upstairs	2.793*** (805.5)	3.732*** (951.5)	0.962*** (205.4)	1.748*** (418.2)	.	-0.964*** (-1927.5)	0.68
Panel B: Price clustering (off-book trades only)							
	Constant	D <sub>0</sub>	D <sub>50</sub>	D <sub>25</sub>	D <sub>75</sub>	LnPrice	Adj R <sup>2</sup>
B. Upstairs	-1.714*** (-170.7)	2.573*** (898.2)	2.347*** (745.8)	1.681*** (382.5)	1.668*** (375.6)	0.213*** (131.2)	0.45
<p><i>LnSizePer<sub>i</sub></i> is the percentage of orders traded at size <i>i</i>, separately for the upstairs and downstairs markets. <i>D<sub>500</sub></i>, <i>D<sub>1000</sub></i> and <i>D<sub>5000</sub></i> are dummy variables that equal 1 if the trade size <i>i</i> is a multiple of 500, 1000 and 5000 shares respectively. <i>LnSize<sub>i</sub></i> is the natural logarithm of trade size <i>i</i>. <i>LnPricePer<sub>i</sub></i> is the percentage of orders traded at a decimal price <i>i</i> on the upstairs market. <i>D<sub>0</sub></i>, <i>D<sub>0.25</sub></i>, <i>D<sub>0.50</sub></i> and <i>D<sub>0.75</sub></i> are dummy variables that equal 1 if the decimal in the traded price <i>i</i> is 0, 0.25, 0.50 and 0.75 respectively. <i>LnPrice<sub>i</sub></i> is the natural logarithm of price <i>i</i>. *, **, *** denote significance level at 10%, 5% and 1% respectively.</p>							

Table 3: Size clustering variation (z- and t-tests)				
Order type	Buy		Sell	
Market Asset	Downstairs	Upstairs	Downstairs	Upstairs
1	15.96***	22.33**	16.34***	24.02***
2	11.67***	14.79**	12.90***	16.27**
3	15.29***	32.29**	15.85***	34.05***
4	13.41***	10.14**	12.89***	11.96**
5	10.36***	17.00***	10.16***	18.48**
6	12.46***	20.58**	12.94**	22.58***
7	11.08***	23.44*	11.21***	23.48**
8	15.41**	29.79***	15.84***	30.77***
9	13.69***	28.62***	14.66***	30.77***
10	15.06***	18.61***	15.63**	23.13**
11	15.19***	21.46***	15.57***	21.73***
12	12.07**	35.24***	12.99**	38.53***
13	12.76***	21.25**	12.10**	22.42***
14	13.62***	14.20**	13.95***	18.47**
15	12.10***	19.62***	12.42***	21.20***
16	15.62**	41.93***	15.48**	44.43***
17	13.50***	21.44***	13.59***	24.06**
18	12.44***	21.89***	13.25***	24.47***
19	23.80***	27.87***	24.81***	32.98***
20	10.79***	11.20**	10.31***	13.46***
21	19.02***	22.01***	19.48**	23.71***
22	21.65**	46.36***	19.72*	49.74***
23	14.07***	22.70***	14.19***	26.30***
24	12.83***	13.25**	12.81***	15.50**
25	13.68**	17.41***	14.12**	19.40***



26	11.87**	9.67**	12.85***	12.05***
27	12.08**	9.14**	11.92***	11.14***
28	17.23***	15.50***	17.69**	20.73***
29	14.40***	13.52***	14.26***	15.27**
30	15.83***	19.16***	15.69***	22.07***
31	13.90***	26.65***	13.86***	28.41***
32	13.07***	21.30**	12.57**	23.67***
33	14.03***	34.29***	14.31***	38.38***
34	15.89**	34.58***	16.59**	38.28***
35	13.84***	19.90***	13.77**	24.18***
36	13.81***	13.92**	14.13**	18.44***
Total (t-test)	14.26	22.03	14.47	24.57
	(-5.00)***		(-6.36)***	

Each cell reports the percent of total trading frequency trading in multiples of 500 shares. Individual asset significance tests are conducted using the z-test. In the last two rows, the t-test is used to test for differences in means between the order-book and off-book trades. \*, \*\*, \*\*\* denote significance level at 10%, 5% and 1% respectively.

Market	Downstairs				Upstairs			
Order type	Buy		Sell		Buy		Sell	
Asset	Sz	SzClst	Sz	SzClst	Sz	SzClst	Sz	SzClst
1	0.027	0.028**	0.026	0.026	-0.040	-0.032	-0.022	-0.061***
2	0.036	0.034***	0.034	0.030***	-0.027	-0.050***	-0.007	-0.041***
3	0.039	0.038*	0.035	0.035	-0.079	-0.104	-0.009	-0.038***
4	0.034	0.033***	0.031	0.028***	-0.039	-0.042	0.009	-0.043***
5	0.027	0.028***	0.026	0.026	-0.019	-0.043***	-0.011	-0.055***
6	0.037	0.035***	0.033	0.032***	-0.013	-0.068***	-0.035	-0.088***

7	0.030	0.032***	0.027	0.030***	-0.026	-0.085***	-0.036	-0.052**
8	0.030	0.031***	0.029	0.029	-0.031	-0.047	-0.046	-0.079***
9	0.036	0.034***	0.035	0.033	-0.059	-0.056	-0.027	-0.078***
10	0.050	0.049***	0.044	0.042***	-0.084	-0.076	0.004	-0.056***
11	0.023	0.021***	0.021	0.019***	-0.031	-0.053***	-0.009	-0.028***
12	0.094	0.088***	0.091	0.088***	-0.068	-0.062	-0.037	-0.061***
13	0.033	0.033	0.030	0.029***	-0.046	-0.047	-0.004	-0.014**
14	0.056	0.055	0.049	0.050	-0.061	-0.061	-0.039	-0.071
15	0.034	0.034	0.032	0.028***	-0.037	-0.052**	-0.027	-0.031
16	0.118	0.110***	0.111	0.109**	-0.062	-0.051	-0.076	-0.106***
17	0.028	0.031***	0.028	0.028	-0.031	-0.067***	-0.035	-0.039
18	0.032	0.030***	0.030	0.028***	-0.034	-0.046**	-0.020	-0.058***
19	0.078	0.076***	0.074	0.072***	-0.062	-0.070***	-0.022	-0.059***
20	0.030	0.030	0.028	0.029	-0.035	-0.040	-0.007	-0.035***
21	0.033	0.032***	0.030	0.030***	-0.022	-0.048***	-0.023	-0.053***
22	0.212	0.210	0.219	0.214***	-0.148	-0.109***	-0.212	-0.158
23	0.028	0.028	0.027	0.027	-0.027	-0.030	-0.015	-0.051***
24	0.031	0.030**	0.028	0.025***	-0.009	0.028***	-0.032	-0.070***
25	0.035	0.033***	0.033	0.031***	-0.021	-0.031***	-0.010	-0.043***
26	0.059	0.056	0.081	0.125***	0.009	-0.046	-0.086	-0.178*
27	0.027	0.027	0.021	0.023***	-0.034	-0.037	-0.004	-0.034***
28	0.050	0.049***	0.047	0.042***	-0.039	-0.038	-0.007	-0.045***
29	0.035	0.034	0.030	0.031***	-0.017	-0.020	-0.032	-0.113***
30	0.036	0.035***	0.034	0.034**	-0.029	-0.037***	-0.016	-0.064***
31	0.036	0.036	0.032	0.032	-0.083	-0.068	0.014	-0.023***
32	0.035	0.035	0.034	0.034	-0.016	-0.099***	-0.056	-0.086*
33	0.041	0.043***	0.040	0.039	-0.036	-0.035	-0.038	-0.088***
34	0.083	0.068***	0.083	0.077***	-0.074	-0.037***	-0.040	-0.126***
35	0.036	0.041	0.036	0.037	-0.043	-0.034	-0.027	-0.095***
36	0.047	0.045***	0.044	0.038***	-0.040	-0.036	-0.014	-0.068***
Total	0.047	0.046	0.045	0.045	-0.042	-0.051	-0.029	-0.066***

Effective half-spread is calculated as:  $EHS_{i,t} = 100D_{i,t}(P_{i,t} - M_{i,t})/M_{i,t}$ .  $D_{i,t}$  equals one for buy trades and negative one for sell orders.  $P_{i,t}$  is the trade price and  $M_{i,t}$  is the midpoint that prevailed at the time of the trade.  $SzClst$  ( $Sz$ ) refers to average effective spreads for trades of sizes (not) of multiples of 500 shares. The t-test is used to test for differences in means between clustered and non-clustered trades separately for order-book and off-book trades. \*, \*\*, \*\*\* denote significance level at 10%, 5% and 1% respectively.

Table 5: Price impact estimates								
Market	Downstairs				Upstairs			
Order type	Buy		Sell		Buy		Sell	
Asset	Sz	SzClst	Sz	SzClst	Sz	SzClst	Sz	SzClst
1	0.026	0.021***	0.015	0.018***	-0.010	0.034***	0.011	0.029***
2	0.028	0.036***	0.021	0.021	0.009	0.018	0.008	0.027***
3	0.037	0.046***	0.020	0.029***	-0.014	-0.011	0.031	0.070***
4	0.030	0.030	0.019	0.022**	0.000	0.033***	0.016	0.021
5	0.021	0.020	0.016	0.017	0.013	0.030***	0.003	0.026***
6	0.026	0.021***	0.023	0.033***	0.029	0.067***	-0.010	-0.001
7	0.026	0.034***	0.021	0.029***	0.010	0.027**	0.003	0.032***
8	0.022	0.023	0.022	0.031***	0.010	0.084***	-0.009	-0.021
9	0.032	0.041***	0.022	0.023	-0.013	0.040	0.034	0.027***
10	0.040	0.052***	0.029	0.049***	-0.013	0.004**	0.007	0.038***
11	0.015	0.016**	0.009	0.010	-0.007	-0.005	0.005	0.020***
12	0.056	0.061***	0.044	0.061***	-0.002	0.010**	0.008	0.006

13	0.026	0.028*	0.018	0.024***	-0.008	0.008***	0.005	0.024***
14	0.042	0.039	0.043	0.064***	0.019	0.061	-0.026	0.031*
15	0.024	0.022	0.020	0.018	-0.007	0.000	0.009	0.024***
16	0.061	0.082***	0.053	0.072***	0.012	0.026*	-0.013	-0.004
17	0.029	0.042***	0.022	0.025	0.009	0.022*	0.002	0.023***
18	0.025	0.033***	0.021	0.023	-0.001	0.028***	0.004	0.023***
19	0.044	0.061***	0.030	0.045***	-0.007	-0.006	0.008	0.015***
20	0.027	0.026	0.026	0.024	0.003	0.056***	0.010	0.013
21	0.022	0.022***	0.021	0.024	0.008	0.022***	-0.007	-0.002*
22	0.115	0.139***	0.094	0.143***	-0.011	0.044**	-0.032	-0.028
23	0.025	0.023	0.021	0.021	0.001	0.013***	0.002	-0.001
24	0.030	0.019***	0.027	0.028	0.030	0.112***	-0.016	-0.004
25	0.024	0.025	0.024	0.027***	0.007	0.013***	0.000	0.002
26	0.043	0.047	0.040	0.053***	0.066	0.065	-0.093	-0.105
27	0.020	0.014***	0.016	0.024***	0.006	0.013***	0.004	0.014***
28	0.036	0.036***	0.024	0.029	0.000	0.008***	0.007	0.008
29	0.023	0.006***	0.033	0.049***	0.038	0.092***	-0.025	-0.038**
30	0.026	0.023***	0.017	0.021***	0.003	0.012***	0.000	0.003
31	0.032	0.030	0.022	0.035***	-0.028	0.016***	0.041	0.074***
32	0.022	0.015***	0.031	0.037***	0.051	0.092***	-0.026	-0.017
33	0.031	0.029	0.025	0.040***	0.012	0.050***	0.019	0.016
34	0.056	0.076***	0.045	0.058***	-0.017	0.045***	0.014	-0.010*
35	0.026	0.035***	0.018	0.020	0.008	0.032***	0.000	-0.016**
36	0.036	0.045***	0.032	0.039***	0.003	0.034***	0.003	0.018***
Total	0.033	0.037	0.027	0.036*	0.006	0.033***	0.000	0.009

Price impact is calculated as:  $P_{i,t} = 100D_{i,t}(M_{i,t+5} - M_{i,t})/M_{i,t}$ .  $D_{i,t}$  equals one for buy trades and negative one for sell orders.  $M_{i,t}$  is the midpoint at time  $t$ .  $SzClst$  ( $Sz$ ) refers to the average price impact estimates for trades of sizes (not) of multiples of 500 shares. The t-test is used to test for differences in means between clustered and non-clustered trades separately for order-book and off-book trades. \*, \*\*, \*\*\* denote significance level at 10%, 5% and 1% respectively.

Table 6: Realised spread estimates								
Market	Downstairs				Upstairs			
Order type	Buy		Sell		Buy		Sell	
Asset	Sz	SzClst	Sz	SzClst	Sz	SzClst	Sz	SzClst
1	0.001	0.006***	0.011	0.007***	-0.032	-0.064***	-0.032	-0.090***
2	0.008	-0.002***	0.014	0.009***	-0.034	-0.067***	-0.015	-0.067***
3	0.002	-0.008***	0.015	0.007***	-0.053	-0.095***	-0.039	-0.100***
4	0.004	0.004	0.012	0.006***	-0.038	-0.072***	-0.007	-0.063***
5	0.006	0.008*	0.010	0.008	-0.032	-0.073***	-0.013	-0.077***
6	0.011	0.014**	0.010	-0.001***	-0.041	-0.129***	-0.025	-0.092***
7	0.004	-0.001***	0.007	0.001***	-0.036	-0.117***	-0.036	-0.077***
8	0.008	0.009	0.007	-0.002***	-0.040	-0.130***	-0.037	-0.058***
9	0.004	-0.007***	0.012	0.011	-0.046	-0.092***	-0.062	-0.106***
10	0.010	-0.003***	0.015	-0.006***	-0.072	-0.084**	-0.003	-0.085***
11	0.008	0.006***	0.011	0.010**	-0.023	-0.048***	-0.013	-0.048***
12	0.039	0.027***	0.047	0.028***	-0.066	-0.073	-0.046	-0.067***
13	0.007	0.005*	0.012	0.004***	-0.039	-0.055***	-0.009	-0.037***
14	0.015	0.018	0.007	-0.014***	-0.079	-0.118***	-0.014	-0.102***
15	0.011	0.011	0.012	0.010*	-0.031	-0.052***	-0.036	-0.056***
16	0.058	0.030***	0.058	0.038***	-0.072	-0.079	-0.060	-0.099***
17	0.000	-0.011***	0.006	0.004	-0.040	-0.088***	-0.038	-0.058**
18	0.007	-0.002***	0.008	0.005***	-0.034	-0.077***	-0.021	-0.076***
19	0.034	0.015***	0.044	0.027***	-0.055	-0.063***	-0.029	-0.073***

20	0.004	0.004	0.003	0.004	-0.039	-0.094***	-0.015	-0.049***
21	0.011	0.010**	0.009	0.006***	-0.030	-0.068***	-0.016	-0.050***
22	0.098	0.071***	0.126	0.072***	-0.144	-0.150	-0.200	-0.126
23	0.004	0.005	0.006	0.006	-0.028	-0.043***	-0.018	-0.051***
24	0.001	0.011***	0.001	-0.003***	-0.037	-0.080***	-0.017	-0.070***
25	0.011	0.009***	0.009	0.004***	-0.029	-0.044***	-0.010	-0.044***
26	0.016	0.010	0.044	0.077*	-0.059	-0.105***	0.004	-0.082***
27	0.006	0.013***	0.005	-0.002***	-0.040	-0.047***	-0.007	-0.046***
28	0.015	0.013	0.023	0.014***	-0.038	-0.043*	-0.014	-0.053***
29	0.011	0.028***	-0.003	-0.017***	-0.055	-0.115***	-0.005	-0.060***
30	0.010	0.012***	0.017	0.013***	-0.032	-0.048***	-0.016	-0.065***
31	0.004	0.006	0.011	-0.003***	-0.056	-0.085***	-0.024	-0.087***
32	0.013	0.021***	0.004	-0.003***	-0.067	-0.184***	-0.028	-0.067***
33	0.010	0.013	0.015	-0.001***	-0.048	-0.086***	-0.054	-0.102***
34	0.027	-0.008***	0.039	0.019***	-0.059	-0.081***	-0.052	-0.114***
35	0.010	0.006	0.018	0.017	-0.050	-0.064***	-0.027	-0.078***
36	0.011	0.000***	0.013	-0.001***	-0.043	-0.071***	-0.016	-0.082***
Total	0.014	0.009	0.018	0.010*	-0.048	-0.083***	-0.029	-0.074***
The realised spread is calculated as the difference between the effective half-spread and the price impact. $SzClst$ ( $Sz$ ) refers to the average realised spread estimates for trades of sizes (not) of multiples of 500 shares. The t-test is used to test for differences in means between clustered and non-clustered trades separately for order-book and off-book trades. *, **, *** denote significance level at 10%, 5% and 1% respectively.								

Table 7: The effect of trade size on spread measures								
Market	Downstairs				Upstairs			
Order type	Buy		Sell		Buy		Sell	
Asset	Sz	SzClst	Sz	SzClst	Sz	SzClst	Sz	SzClst
Panel A: 1 – 1500 Shares								
EHS	0.047	0.045	0.045	0.041	-0.037	-0.032	-0.014	-0.033**
PI	0.030	0.027	0.021	0.017	0.003	0.003	-0.002	-0.005
RHS	0.017	0.019	0.025	0.024	-0.041	-0.036	-0.013	-0.028***
Panel B: 1501 – 3000 Shares								
EHS	0.047	0.044	0.046	0.042	-0.036	-0.016**	-0.027	-0.043***
PI	0.032	0.028	0.029	0.025	0.003	0.020*	-0.002	-0.007
RHS	0.016	0.016	0.021	0.017	-0.040	-0.036	-0.024	-0.036***
Panel C: 3001 + Shares								
EHS	0.047	0.047	0.046	0.047	-0.077	-0.071	-0.075	-0.09
PI	0.037	0.040*	0.033	0.041	0.018	0.051***	0.015	0.021
RHS	0.011	0.007	0.013	0.006	-0.091	-0.121***	-0.090	-0.109**
<p><i>SzClst</i> (<i>Sz</i>) refers to average effective spreads for trades of sizes (not) of multiples of 500 shares. Effective half-spread is calculated as: <math>EHS_{i,t} = 100D_{i,t}(P_{i,t} - M_{i,t})/M_{i,t} \cdot D_{i,t}</math> equals one for buy trades and negative one for sell orders. <math>P_{i,t}</math> is the trade price and <math>M_{i,t}</math> is the midpoint that prevailed at the time of the trade. Price impact is calculated as: <math>PI_{i,t} = 100D_{i,t}(M_{i,t+5} - M_{i,t})/M_{i,t}</math>. The realised spread (RHS) is calculated as the difference between the effective half-spread and the price impact. The t-test is used to test for differences in means between clustered and non-clustered trades separately for order-book and off-book trades. *, **, *** denote significance level at 10%, 5% and 1% respectively.</p>								

Table 8: Endogeneity test

Market	Downstairs		Upstairs	
Dependent	Size Clust	Price Clust	Size Clust	Price Clust
PrcClst	0.13***		0.09***	
	(29.07)		(4.64)	
PrcRs	-0.10***		0.07***	
	(-16.06)		(3.39)	
SzClst		-1.66***		-2.25***
		(-6.46)		(-4.01)
SzRs		1.82***		2.51***
		(7.09)		(4.49)
Adjusted R <sup>2</sup>	0.41	0.70	0.42	0.24
N	9,000	9,000	9,000	9,000

*SzClst* is the percentage of trades conducted at multiples of 500 shares and *PrClst* is the percentage of trades conducted at integers (downstairs market) or at multiples of the minimum tick size (upstairs market). *PrcRs* and *SzRs* denote the residuals from the *PrClst* and *SzClst* variables from the first set of regressions respectively.

\*, \*\*, \*\*\* denote significance level at 10%, 5% and 1% respectively.

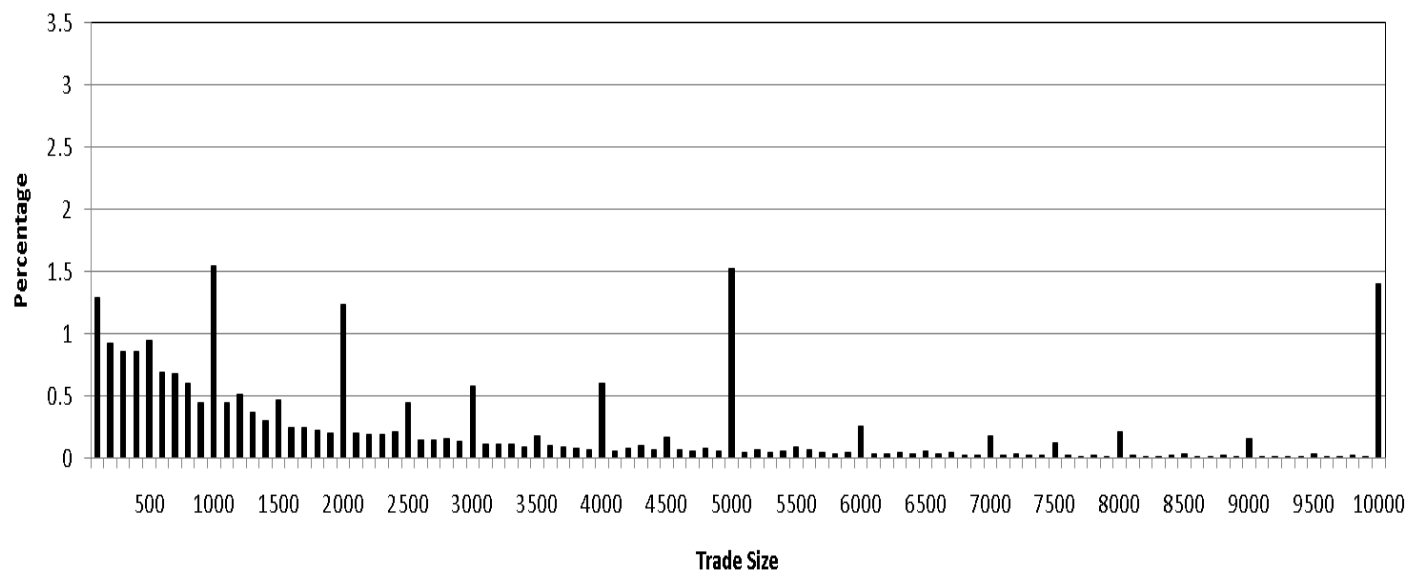


Table 9: 2SLS regression estimates				
Market	Downstairs		Upstairs	
Dependent	Size Clust	Price Clust	Size Clust	Price Clust
Constant	-27.99***	-73.37***	-20.03***	9.61

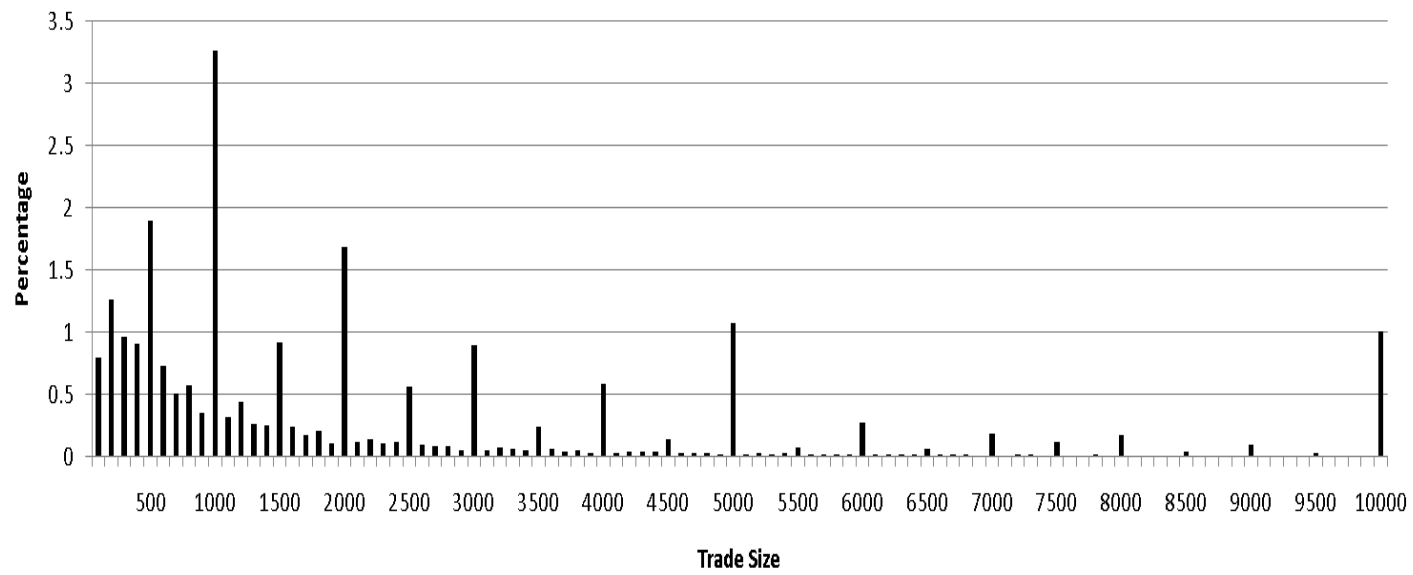
	(-32.19)	(-3.66)	(-15.35)	(1.53)
PrcClst	0.13***		0.11***	
	(28.60)		(6.05)	
SzClst		-3.01***		0.10***
		(-3.75)		(4.95)
Frq	0.001586***		0.000086	
	(16.47)		(0.18)	
SQRTFrq		-369.65***		-21.10
		(-2.83)		(-1.40)
AR	4.29***	25.38***	4.79***	-2.73**
	(9.82)	(5.20)	(12.36)	(-2.16)
LnTrdSz	4.75***	19.88***	5.79***	-2.24*
	(57.42)	(4.46)	(65.45)	(-1.91)
LnMV	-0.87***	-5.56***	-2.00***	0.45
	(-9.87)	(-5.77)	(-16.52)	(1.57)
OEQ	0.28*		0.52*	
	(1.84)		(1.85)	
PrcLvl		0.08***		0.03***
		(12.65)		(17.02)
Tck		8.20***		-4.19***
		(7.70)		(-4.25)
Adjusted R <sup>2</sup>	0.39	0.45	0.40	0.17
N	9,000	9,000	9,000	9,000
<p><i>SzClst</i> is the percentage of trades conducted at multiples of 500 shares and <i>PrcClst</i> is the percentage of trades conducted at integers (downstairs market) or at multiples of the minimum tick size (upstairs market). <i>Frq</i> is the daily number of trades for each contract, separately for the electronic and the upstairs market. Similarly, <i>SQRTFrq</i> is the inverse square root of the daily number of trades. <i>AR</i> is the average absolute daily return calculated over hourly intervals. <i>LnTrdSz</i> denotes the natural logarithm of the average daily trade size for each contract separately for electronic and market makers trades. <i>OEQ</i> is a binary dummy variable that takes the value of 1 when a stock trades at a calendar quarter end. <i>LnMV</i> is the natural logarithm of .market capitalisation for each firm. Firm values are obtained from DataStream.</p>				

Panel A: Downstairs market

*PrcLvl* denotes the average daily price level for each contract and market segment. *Tck* is a dummy variable that takes the value of 1 for stocks trading at a high tick size.  
\*, \*\*, \*\*\* denote significance level at 10%, 5% and 1% respectively.

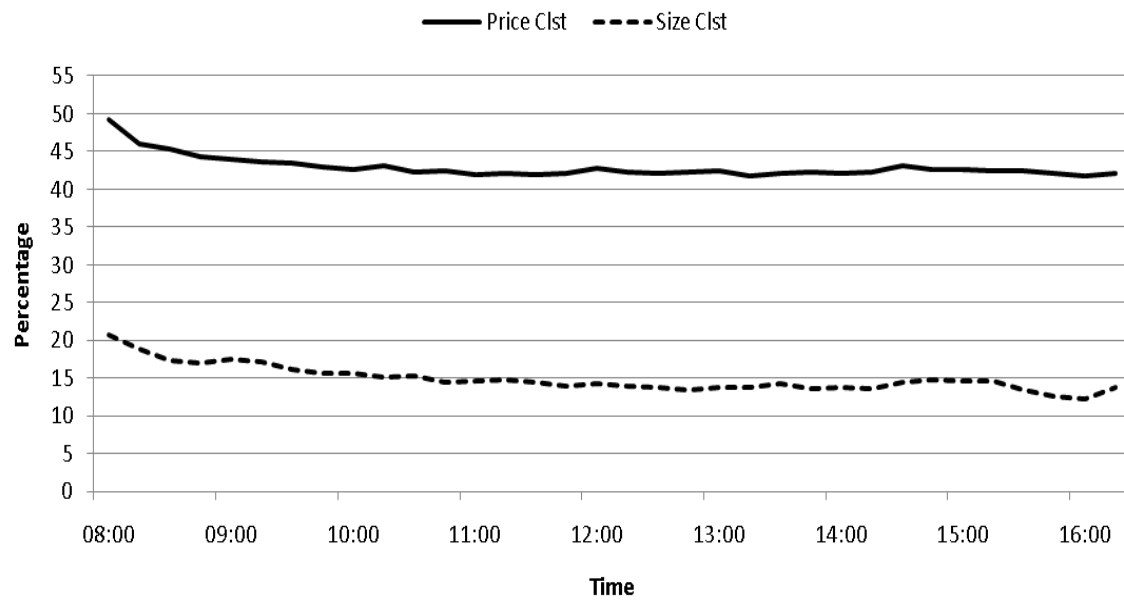


Panel B: Upstairs market

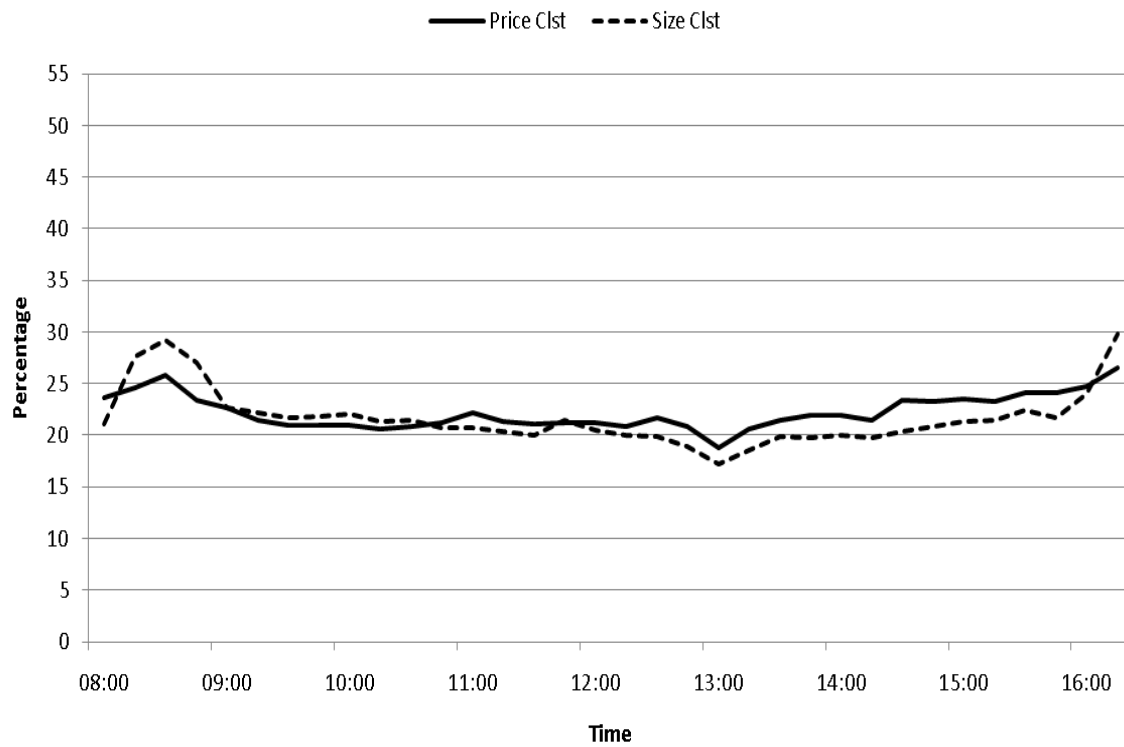


**Fig 1: Trade size distribution**

Panel A: Downstairs market (Price clustering: integers, size clustering: multiples of 500 shares)



Panel B: Upstairs market (Price clustering: minimum tick, size clustering: multiples of 500 shares)



**Fig 2: Intraday distribution of size and price clustering.**

For the downstairs market, price clustering refers to the percentage of trades that are conducted at integers. For the upstairs market, price clustering refers to the percentage of trades conducted at the notional minimum tick sizes (i.e. xx.25 and xx.50).