Market Microstructure Invariance in the FTSE 100

Michael Bowe[1,2], Efthymios Rizopoulos[1,3], S. Sarah Zhang[1,4]

Preliminary Version

Please do not cite without permission of the authors

Abstract

We examine market microstructure invariance relationships in FTSE 100 index constituent stocks. We formulate a generalised version of the invariance model to account for the impact of intermediation and order flow imbalances on returns variance. Our results provide qualified support for microstructure invariance, indicating that trade counts are proportional to trading activity (to the power of 0.5), when observations are averaged across days. Invariance relationships receive more support among stocks with high average volatility. However, notions of trading activity that imply a proportionality between trading volume and returns variance are more precise when estimating certain invariance relationships, especially so for equities with low average volatility. Excluding time intervals with extreme volatility reveals a correlation between the number of trades and returns variance.

Key words: Market Microstructure, Trading Invariance, Order Size, Trading Volume, Volatility, High Frequency Trading, Business Time

EFM Classification: 330, 360

[1] University of Manchester, Alliance Manchester Business School, Booth St E, Manchester, M13 9SS, UK
[2] E-mail: michael.bowe@manchester.ac.uk;
[3] Corresponding Author & Presenter. E-mail: efthymios.rizopoulos@manchester.ac.uk;
[4] E-mail: sarah.zhang@manchester.ac.uk

1. Introduction

Market microstructure invariance maintains that each capital market has a distinct business time[1] scale in which it operates and during which risk transfers occur. Market microstructure invariance investigates the relationship between what Kyle and Obizhaeva (2016) define as bets (asset-specific risk transfers with small or no correlation with market risk as customarily defined) and business time rather than exploring the relationship between trading volume and business time. In this context, Kyle and Obizhaeva (2016) specify trading activity as the product of trading volume (in local currency units) and return volatility in business time and research its effect on different market microstructure characteristics. Intuitively, market microstructure invariance hypothesises that market microstructure characteristics, such as order size, order arrival rate, price impact, bid-ask spreads and price resilience remain approximately constant when estimated in business time. In a recent paper, Andersen et al. (2015) extend this theory by studying whether similar invariance principles, which they name intraday trading invariance, apply to trades.

We extend the analysis of Kyle and Obizhaeva (2016) on market microstructure invariance together with that of Andersen et al. (2015) on intraday trading invariance and investigate its validity in a novel equity market context, specifically FTSE 100 stocks traded on the London Stock Exchange (LSE). Our contribution is threefold: First, we amend the market microstructure invariance model specification for bets to facilitate its empirical application in different market settings. Specifically, our proposed modification not only accommodates the effect of intermediation and order flow imbalances on returns variance, but it is also required in order to empirically implement the specific way in which our transaction dataset records executed trades. We utilise the modified model to test for an invariance relationship between the number of trades and trading activity following Andersen et al. (2016). We compare our findings using the Kyle and Obizhaeva (2016) invariance specification of trading activity to two well-known alternatives in the literature implied by Clark (1973) and Ané

[1] Business time is also referred to as operational time, economic time or information time The concept of using alternative time references to calendar time when estimating trading activity is a salient feature of the market microstructure literature. According to Hasbrouck (1999), a fundamental aspect of such "time deformation" invokes a differentiation between business time, in which a particular system develops and calendar time in which someone observes it. Empirical papers employ data, aggregated for short real time spans that incorporate shifting intervals of business time. The estimated returns over these periods are expected to follow combinations of business time distributions. Bochner, S. (1955) proposes that time deformation can be specified in terms of subordinated stochastic processes, whereas Clark, P. K. (1973) and Tauchen, G. E. & Pitts, M. (1983) imply that the relationship between business and calendar time is normally expressed as a function of latent or observed variables.

and Geman (2000), respectively. Second, we allow for a more general invariance relationship between the number of trades and trading activity by relaxing the assumption of a monopolistic market maker. This enables us to accommodate different institutional market arrangements and crucially methods of reporting transaction data into consideration. The stipulated theoretical 2/3 proportionality between the number of bets and trading activity proposed by Kyle and Obizhaeva (2013) and the empirical 2/3 proportionality between number of trades and trading activity as formulated in Andersen et al. (2016), are implied by our specification in the market environments they analyse. Finally, we conduct one of the first tests of intraday trading invariance for equity markets using a subset of 25 equities from the FTSE 100 stocks trading on the LSE.

In this regard, it is important to note the difference in the nature of our data and that in Andersen et al. (2015). Our analysis uses tick data, recording any large, liquidity demanding trades executed against multiple passive limit orders as separate trades. In contrast, Andersen et al.'s (2015) data only captures that part of the marketable order which is executed at the best bid or ask price, recording it as a single trade. Thus, our data biases are quite different: Andersen et al. (2015) may not capture the full executed order, as they record only trades transacted at the best bid and ask price, but the data does not suffer from order splitting. As our data can split a single marketable order, we may experience enhanced intermediation, an inflated number of trades and lower trade sizes. We accommodate these intermediation and data recording differences in the proposed extension of the theoretical invariance model.

Our principal empirical findings are as follows. Corroborating invariance, the number of trades is indeed proportional to trading activity. However, rather than the 2/3 proportionality documented in Andersen et al. (2016), we find proportionality to the power of 1/2 for the majority of stocks. Based on the extended model we propose, this suggests that an order is intermediated an average of 2 to 2.56 before it is fully executed, with the percentage of returns variance attributable to order flow imbalances ranging between 75% and 100%, respectively. As the arrival rate of trades measures market velocity, this intuitively means that the FTSE 100 stock market is approximately 50% slower in business time compared to the E-min S&P 100 future contracts market, the subject of analysis in Andersen et al. (2015).

The second principal finding is that this proportionality holds somewhat better when the notion of trading activity introduced by Clark (1973) is used, though the invariance model is more precise for stocks with high average volatility. This suggests a proportionality between returns variance and trading volume, consistent with Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983) and other papers on "mixture" of distributions hypothesis. Also, it shows that the correlation between trading volume and returns variance for most stocks is higher than that between trade size and returns variances and that between trade size and number of trades. When minutes with extreme volatility are excluded, invariance model becomes less accurate, while returns variance appears to be proportional to number of trades (i.e. when the notion of trading activity by Ané and Geman (2000) is employed). The model of Clark (1973) still yields significant results. This highlights that market participant's change trading behaviour based on the innovations in returns variance upon arrival of new information and that the correlations between the underlying variables are altered.

We make an important contribution to the literature about which distribution family best describes variations in returns and their relationship with other variables of trading activity, mainly trading volume (either measured in trade counts or number of securities traded). Some papers argue that price changes follow a Pareto-type distribution in calendar time and different distribution in business time, when the latter is measured in volume terms. Mandelbrot and Taylor (1967), Plerou et al. (2000) and Bouchaud et al. (2008) are representatives of this category. Other papers allow real time intervals of their sample to incorporate intervals of informational time and support that returns over those intervals follow a mixture of informational-time distributions. Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983), Gallant et al. (1992) and Andersen (1996) support this "mixture" of distributions hypothesis (MDH), defined as the joint hypothesis between returns and volume. Karpoff (1987) provides an excellent survey of empirical and theoretical papers concerned with the relationship between trading volume and price movements till that date. In contrast, Jones et al. (1994), Ané and Geman (2000), Dufour and Engle (2000) assert that that number of trades and not trading volume is a better proxy for business time, due to their connection with price volatility.[2]

The notion of invariance in finance and economics is not a recent innovation. The implicit use of invariance principles dates back to 1950s and the irrelevance of capital structure proposition in Modigliani and Miller (1958). In mathematics and theoretical physics invariance is a property of a system that does not change when transformed under a valid condition. The speed of light (Einstein,

[2] All these papers treat trading volume or number of trades as proxies for the business time clock, but they do not examine them simultaneously with price changes in the context of business time. The introduction of market microstructure invariance by Kyle and Obizhaeva (2013) attempts to switch the insights of the time deformation literature and solve this problem.

1905; Einstein, 1920) and the area of a defined shape that does not vary relative to Euclidean plane isometries which remains constant are established invariant characteristics (Hunt, 1996).

The paper proceeds as follows. Section 2 reviews related literature. Section 3 explains the generalised theoretical invariance model. Section 4 focuses on the methodology and the main and alternative empirical hypotheses. Section 5 highlights the characteristics and descriptive statistics of the dataset used in this paper. Section 6 presents and discusses the empirical results. Section 7 concludes.

2. Relevant Literature

2.1 Time deformation literature

Time deformation literature begins with the intuition that execution of individual trades leads to microscopic, normally and independently distributed, price fluctuations that are incremental to daily price changes. More specifically, every time new information becomes available, the variables of trading activity (e.g. number of trades, volume, transaction rate[3], quote revision frequency) and consequently prices shift. Engle (2000) explains that the arrival rate of information represents the speed at which business/economic time passes and that a blunt measure of this rate is obtained when transaction times and prices are analysed simultaneously. Intuitively, given that price changes follow high-kurtosis distributions[4], the observed fat tails can be interpreted from the perspective of a business clock that ticks at different velocities compared to real "wall-clock" time to allow for differentiation of trade execution speed across defined time periods (Kyle and Obizhaeva, 2010).

a) Measuring time deformation: Pareto distributions

Mandelbrot and Taylor (1967) are the first to examine the distribution of stock price changes (i.e. price volatility) by measuring time in volume of transactions. They argue that price changes have stable Pareto (power-law) distribution during defined calendar time intervals and Gaussian distribution, when these fixed intervals are measured in transactions time[5]. Similar to Mandelbrot and Taylor (1967), Gopikrishnan et al. (1998) analyse the probability distribution of stock price

[3] Time between trades is defined as the reciprocal of transaction rate

[4] Distributions with sharper peak around mean and fatter tails compared to normal distribution

[5] Mandelbrot and Taylor (1967) reach this conclusion by introducing for the first time the concept of subordination (i.e. time change stochastic process inside another stochastic process) when modelling financial returns

movements and deduce that they display asymptotic Pareto distribution with an exponent close to 3. Moving one step forward, Plerou et al. (2000) maintain that price fluctuations follow a complicated diffusion process in which the diffusion constant is related to the number of transactions for a specific time interval and the variance of price fluctuations for all transactions. The number of transactions and variance of price fluctuations between consecutive trades follow a power law distribution with a mean value of the exponent around 3. Provided that there is market impact is linear in trade size[6] and trades are i.i.d., the authors suggest that the Pareto distribution tails of price movements are attributed only to variance, whereas the number of transactions is responsible for the long-range correlations of volatility.

Alternatively, Gabaix et al. (2006) identify a dependence of empirical price changes on square-root price impact of i.i.d. trades based on a specified version of a model introduced by Torre and Ferrari (1998)[7]. They state that the estimation of price impact and its connection to order size is problematic due to the joint endogeneity of order flow and returns[8]. Bouchaud et al. (2008) adopt the Pareto distribution of price movements, but they propose an original model to explain how changes in supply and demand affect prices. They underline that order flow is a *"highly persistent long-memory process[9]"* and assert that the informativeness of prices stems more from supply and demand than external news. Intuitively, this suggests that any price revision must strongly rely upon the past history of order flow so that the market remains efficient.

a) Mixture of distribution Hypothesis

Other business time papers differentiate from the Pareto distribution approach. Those papers investigate the "mixture" of distributions hypothesis (MDH) as the joint hypothesis between returns and volume. The theory introduces an explicit way to model the impact of information on prices and volume by expressing the respective variables as a function of the arrival rate of information to the market. Clark (1973) suggests that the number of minor price innovations per day exhibits a log-normal distribution and that both volume and price innovations are triggered by the same

[6] Prices move upward or downward proportionally to the trade size
[7] Zhang (1999) and Gabaix et al.(2003) also propose a model with square-root price impact.
[8] Loeb (1983) allows for the exogeneity of order flow by using bids on various size blocks of stock. Both Torre (1997) and Grinold and Kahn (1999) mention square root price impact fits best this type of data.
[9] Autocorrelation in order flow decays very slowly

information arrival process[10]. Assuming identical distribution of distinctive price increments, he argues that return variance is proportional to trading volume when using the latter as an approximate transactions time clock. Both, Epps and Epps (1976) and Westerfield (1977) confirm this proportionality, though they underline that the theory of stable Paretian distributions proposed by Mandelbrot and Taylor (1967) cannot be ruled out.

Tauchen and Pitts (1983) add that the positive relation between returns variance and trading volume is subject to a fixed number of traders[11], while Harris (1987) maintain that prices and volume unfold at homogenous rates in business time[12]. Gallant et al. (1992), using a non-semiparametric specification, report a positive relationship between volatility and volume, both conditional on their past observations. Richardson and Smith (1994), introduce a direct test for the MDH and examine different distributional properties for the rate of information flow. They conclude that the bivariate distribution of price changes and volume is not as strong as previous studies suggest and that the distributional properties of the information flow rate approach those of log-normal distribution. Andersen (1996) incorporates the market microstructure setting of Glosten and Milgrom (1985) in the MDH arguing that the full dynamic representation of the stochastic volatility process for the rate at which information arrives in the market performs better compared to the standard one. Bollerslev and Jubinski (1999), investigate the implication of the MDH in terms of the long-memory characteristics of the rate of external news and find that the information arrival processes exhibit a sluggish hyperbolic rate of decay. Finally, Liesenfeld (2001) generalises the MDH standard model by allowing both the number of information arrivals and the sensitivity to new information to be dynamic over time. The revised model more accurately explains stock price fluctuations, whereas trading volume appears to be mainly affected by the information arrivals count.

b) Time deformation and market events

Hasbrouck (1999) approaches time deformation as a common feature in the rates at which market process events such as orders, quote or trade frequency occur. While time deformation defined in this way can only be estimated dependent upon a specified time horizon, the author finds a positive

[10] Clark (1973) implies that movements in returns are caused by a joint distribution between trading volume and prices conditional on current information

[11] This assumption is rational for mature markets. If the number of traders is evolving, then the average trading volume grows in a linear fashion with the number of traders.

[12] Harris (1986) shows that the relationship between returns variance and trading volume is also present in the cross-section of stocks.

correlation in the long-term, although there is no persistent proportionality in the count intensities for different types of events. Jones et al. (1994) discover that the number of transaction per se, and not their size (i.e. volume), creates daily volatility. They state that the volume does not contain any extra information other than the one included in the trades count. In line with Jones et al. (1994), Ané and Geman (2000) report that the aggregate number of trades is a better business time clock than volume for generating independently and identically distributed Gaussian intraday returns. Dufour and Engle (2000), based on the VAR specification of Hasbrouck (1991), analyse the impact of time duration between successive transactions on the process of price formation. Their findings indicate that whenever these waiting times decline, trades have greater impact on prices, the latter adjust faster to trade-related information and the positive serial correlation of signed trades increases. They claim that active markets, where the increased participation of informed traders leads to high trading activity, are illiquid.

Estimating a fully specified time deformation model for stock market dynamics is challenging. A number of the aforementioned papers suggest infinite variance distributions for price innovations; others use distributions with finite variance. Results regarding which distribution best describes price changes are inconclusive. Also, studies based on the MDH suggest different ways of approximating the mixing variable in the subordinate stochastic processes. While trading volume or the number of trades are the prevailing proxies, both are considered somewhat imperfect. Finally, all the above papers analyse only volume or price changes, not both, in the context of business time. Thus, their simultaneous inclusion in a time series model is not undertaken. The introduction of the notion of market microstructure invariance by Kyle and Obizhaeva (2013) aims to resolve these problems.

2.2 Market Microstructure Invariance

In market microstructure invariance theory, Kyle and Obizhaeva (2016) propose that trading activity constitutes a risk game, transferring risk (in the form of a bet) from one market participant to another. Here, a bet is a decision to initiate a long-term position of a certain size in a specific security. Each bet can be executed either in its entirety or by sequentially placing orders, with each order potentially being segmented into small trades executed at different prices. To accentuate the idea of risk transfer, invariance measures trading activity as the product of expected volume (in currency units) and return volatility, both expressed in business time, and defines the business-time clock as ticking at the arrival rate of bets in the market. The arrival rate of bets is generated by a compound Poisson process (to the extent that the number of bets per calendar day follow a Poisson distribution) and price innovations stemming from bets follow a symmetric distribution around zero, with their unsigned size being log-normally distributed under the assumption of linear price impact.

Market microstructure invariance complements existing theoretical models of market microstructure[13] based on the notion that order flow imbalances generate price fluctuations and which develop measures of market depth or liquidity. Currently, there is neither a consolidated framework to construct empirical measures for order flow imbalances, nor to provide accurate forecasts regarding the differentiation of price impact across stocks. As a result, the empirical proxies, for testing the relationship between price changes, order flow imbalances and their connection to stock characteristics are imperfect (e.g. Breen et al. (2002)). The invariance principle attempts to bridge the gap between theoretical market microstructure models and their empirical counterparts by imposing "cross-sectional restrictions" that facilitate both the empirical assessment of the former and the implementation of liquidity measures that are contingent on order flow imbalances.

Andersen et al. (2015) extend market microstructure invariance theory and investigate whether the invariance of bets applies also to trades. As trades are by definition, fundamental components of bets, one can ask whether trades exhibit similar invariance relationships proposed for bets during trading days. Since the constituent variables of trading activity are measured at high frequencies, Andersen et al. (2016) define the invariance principle as "intraday trading invariance". To that end, using tick by tick data on the E-mini S&P 500 future contracts, they test empirically, on a minute by minute basis whether the trade arrival rate is proportional to 2/3 power of trading activity (i.e. the product of expected trading volume and returns volatility). They conclude that invariance properties appear to be present in the specific market, but they highlight the possibility that these results are not universal to every asset class and market.

[13] More generally, market microstructure invariance consists of three principles proposed to hold across business time and securities: "invariance of bets", "invariance of transactions costs" and "invariance of market efficiency and resilience "The three empirical hypotheses are thoroughly explained in Kyle and Obizhaeva (2013). This paper only considers the invariance of bets.

3. Model

3.1 Notation

According to Kyle and Obizhaeva (2013), market microstructure invariance is based on the assumption that over short calendar time periods the bet arrival rate N_B (number of bets per unit of time) can be approximated by a compound Poisson process. Assuming that the inventories of intermediaries evolve in a bounded way, over longer periods the bets exhibit negative serial correlation and the bet arrival rate and distribution of bet size vary with trading activity. The bet arrival rate N_B measures market velocity. The signed size of bets (in shares, positive for buys and negative for sells) is represented by the probability distribution of a random variable $\tilde{Q}$ (number of shares) with a positive sign for purchases and a negative sign for sales, where $E\{\tilde{Q}\}$ is approximately zero.

Kyle and Obizhaeva (2013) assume that on average one unit of bet volume $\bar{V} := N_B \cdot E\left|\tilde{Q}\right|$ results in ζ units of total volume V (i.e. $\zeta - 1$ units of intermediation volume per one unit of bet volume). The expected total trading (market) volume during a specific calendar time is given by the following equation (each buy-sell bet pair is considered only once):

$$V := \frac{\zeta}{2} \cdot N_B \cdot E\left|\tilde{Q}\right| \qquad \textbf{(1)}$$

where N_B is the bet arrival rate and $E\,|\,\tilde{Q}\,|$ the average bet size

Given the aforementioned assumption and equation (1), the "expected bet volume" (stocks per unit of time) $\bar{V}$ can be specified by the following expression:

$$\bar{V} := N_B \cdot E\left|\tilde{Q}\right| = \frac{2}{\zeta} \cdot V \qquad \textbf{(2)}$$

where ζ is the "intermediation multiplier"

The value of ζ, the intermediation multiplier, depends on the number of intermediaries in the market. A greater value means that more agents intermediate a bet transfer. In equation (2), $\zeta = 1$

would mean that there are no intermediaries in the market. When there is only one market maker (i.e. market makes monopoly, $\zeta = 2$), the expected bet volume equals the expected market volume ($\bar{V} := V$). A value of $\zeta = 3$ equivalently implies two market makers, while for $\zeta > 3$ there are multiple intermediaries in the market. In Kyle and Obizhaeva (2013) and Andersen et al. (2016), the working assumption is that there is only one market maker and thus the expected bet volume equals the expected market volume ($\bar{V} := V$).

Following Kyle and Obizhaeva (2013), we argue that during trading days stock price fluctuations lead to a percentage variance of respective returns, denoted as σ^2. A function, χ^2, of the variation in asset prices is caused without trading, by information updating relating to prices and does not require trading. The other portion ψ^2, occurs in response to order flow imbalances during trading. Assuming that order flow imbalances solely result from bets, we can define "trading volatility" as:

$$\bar{\sigma} := \psi \cdot \sigma \tag{3}$$

where $\bar{\sigma}$ is the standard deviation of returns that stems from order flow imbalances related to bets.

Finally, if P (currency units per stock[14]) represents the stock price in currency units, then trading volatility in currency units is given by:

$$P \cdot \bar{\sigma} := \psi \cdot P \cdot \sigma \tag{4}$$

Following Andersen et al. (2016), we examine the relationship between the number of trades and trading activity, in order to provide empirical evidence relating to the first principle of market microstructure invariance, the invariance of bets. This empirical hypothesis implies that *"the pound distribution of risk in currency units transferred by bets is the same for all stocks when the risk transferred by a bet is measured in units of business time[15]"* (Kyle and Obizhaeva, 2013, p 6).

According to Kyle and Obizhaeva (2013), the trading volatility (in currency units) in one unit of business time ($P \cdot \bar{\sigma} \cdot N_B^{-1/2}$) multiplied by the distribution of signed bet size ($\tilde{Q}$) measures both, the

14 Kyle and Obizhaeva (2013) investigate US market and thus they measure price in dollars per stock

15 As before in our case it is the distribution of risk in currency units and not the dollar distribution of risk as in Kyle and Obizhaeva (2013)

direction and size of the risk transferred by a bet per unit of business time. Intuitively, this means that in one unit of business time (N_B^{-1}) a bet of size ($P \cdot \tilde{Q}$) in currency units generates a standard deviation of mark-to-market gains or losses in currency units equal to $P \cdot |\tilde{Q}| \cdot \bar{\sigma} \cdot N_B^{-1/2}$. The invariance of bets principle states that there exists a random variable, $\tilde{I}$, with a distribution equal that of $P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot N_B^{-1/2}$, which does not change across stocks:

$$\tilde{I} \sim P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot N_B^{-1/2}, \ \tilde{I} := P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot N_B^{-1/2} \qquad \textbf{(5)}$$

where $\tilde{I}$ is a market microstructure invariant

The invariance equation in (5) is based on two assumptions. First, the intermediation multiplier is $\zeta = 2$, which implies a monopolistic market where the expected bet volume equals the expected market volume ($\bar{V} := V$). Second, all fluctuations in returns stem from order flow imbalances (i.e. $\psi = 1$). We argue that these hypotheses empirically do not hold in every market and asset category. For this reason we generalise the invariance of bets in (5) as follows:

$$\tilde{I} := P \cdot \tilde{Q} \cdot \bar{\sigma} \cdot N_B^{-\frac{\xi}{1-\xi}} \qquad \textbf{(6)}$$

where ξ is a parameter such that $0 \le \xi \le 1$

The invariance specification in (6) is similar to the generalised invariance relationship introduced by Kyle and Obizhaeva (2010) as an alternative invariance hypothesis. For $\xi = 0$, equation in (6) implies that the bet size remains constant, while there is a proportionality between trading activity and number of bets. In that case the invariance specification captures some of the properties suggested by Gabaix et al. (2006) and Hasbrouck (2009). For $\xi = 1$, the number of bets per day remains constant, while the bet size is proportional to trading activity. In that case, equation in (6) resembles the model of Amihud (2002) and describes the common trading knowledge. In contrast to the specification of Kyle and Obizhaeva (2010), where ξ is a random parameter, here we allow ξ to be a function of the intermediation multiplier ζ and the percentage of order flow imbalances attributed to bets ψ. More specifically, ξ is defined as follows:

$$\xi = 1 - \frac{4\psi}{3\zeta} \quad \textbf{(7)}$$

where ζ is the intermediation multiplier and ψ the percentage of order flow imbalances attributed to bets.

From (7) it can easily be inferred that for $\psi = 1$ and $\zeta = 2$, the parameter ξ is equal to $1/3$ and thus the relationship in (6) becomes the invariance hypothesis in (5). Intuitively, by introducing the parameter ξ, we assume that empirically, the risk transferred by a bet is not proportional to the square root of bet arrival rate, but to the bet arrival rate in the power of $\xi/(1-\xi)$. This parameter depends on the intermediation multiplier ζ and the percentage of order flow imbalances attributed to bets ψ.

Proceeding, the "expected trading activity" W is represented by the product of the expected trading volume $P \cdot V$ in currency units and the volatility of returns σ in calendar time:

$$W := P \cdot V \cdot \sigma \quad \textbf{(8)}$$

Intuitively, expected trading activity W can be seen to be the standard deviation of mark-to-market gains or losses in currency units on all expected trading volume in one unit of calendar time (i.e. a measure of "total risk transfer" per unit of calendar time). Analogously, "expected bet activity" $\bar{W}$ is defined as:

$$\bar{W} := P \cdot \bar{V} \cdot \bar{\sigma} \quad \textbf{(9)}$$

where $P \cdot \bar{V}$ is expected bet volume in currency units and $\bar{\sigma}$ is trading volatility.

In addition, the expected trading and bet activity are connected by the following equality (proof is presented in Appendix I-1):

$$\overline{W} := \frac{2\psi}{\zeta} \cdot W \quad \textbf{(10)}$$

Expanding equation (9) in terms of total volume V, price P, volatility σ and expected trading activity W, yields the following expression (proof is presented in Appendix I-2):

$$\frac{2}{\zeta}\cdot \mathrm{W} = P \cdot N_B \cdot \mathrm{Q}_B \cdot \sigma \tag{11}$$

where $\mathrm{Q}_B = E\,|\,\tilde{\mathrm{Q}}\,|$

Also, from the generalised invariance relationship in (6), taking expectations produces the following equation:

$$E\,|\,\tilde{\mathrm{I}}\,| := \mathrm{P}\cdot E\,|\,\tilde{\mathrm{Q}}\,|\cdot\bar{\sigma}\cdot \mathrm{N}_\mathrm{B}^{-\frac{\xi}{1-\xi}} \tag{12}$$

Solving for $E\,|\,\tilde{\mathrm{Q}}\,|$ and substituting from (4) for price P and volatility σ gives,

$$\mathrm{Q}_B = \mathrm{I}\cdot \mathrm{P}^{-1}\cdot\psi^{-1}\cdot\sigma^{-1}\cdot \mathrm{N}_\mathrm{B}^{\frac{\xi}{1-\xi}} \tag{13}$$

where $\mathrm{Q}_B = E\,|\,\tilde{\mathrm{Q}}\,|$ and $\mathrm{I} := E\,|\,\tilde{\mathrm{I}}\,|$, the mean of the invariant $\tilde{I}$ distribution

Replacing Q_B in (11) with the expression in (13), we obtain the invariance of bets in model terms (proof is presented in Appendix I-3):

$$\mathrm{I} := \frac{2\psi}{\zeta}\cdot\frac{W}{N_B^{\frac{1}{1-\xi}}} \overset{(7)}{=} \frac{2\psi}{\zeta}\cdot\frac{W}{N_B^{\frac{3\zeta}{4\psi}}} \tag{14}$$

Since $\tilde{I}$ has an invariant distribution, its mean I is a constant which does not depend on the trading activity W, intermediation multiplier ζ, fraction ψ or the bet arrival rate N_B.

Thus, from equation (14) it can be inferred that N_B is proportional to $\frac{2\psi}{\zeta}\cdot W^{\frac{4\psi}{3\zeta}}$.

As a special case, if all asset prices change as a result from order flow imbalances stemming from bets, then $\psi = 1$. If, in addition, the intermediation multiplier is $\zeta = 2$, then from the relationship in (14), it follows:

$$\mathrm{I} := \frac{\mathrm{W}}{\mathrm{N}_B^{3/2}} \tag{15}$$

Equation (15) is precisely the invariance relationship suggested by Kyle and Obizhaeva (2013) and Andersen et al. (2016).

4. Methodology

4.1 Main Hypothesis

The relationship in (14) constitutes the basis of our empirical invariance hypothesis. We need to transform this proportional relationship between bets and trading activity into trading terms, in order to test it empirically. According to Andersen et al. (2016), intraday trading invariance "stems directly" from market microstructure invariance and the connection between bets and trades. In fact, intraday trading invariance can be obtained from the market microstructure invariance hypothesis for bets by imposing an additional requirement that the average number of trades per bet remains constant and assuming that the fixed cost of making a trade is also constant.

Given that market microstructure invariance implies an invariant distribution of $\tilde{I}$, using logarithms of means or means of logarithms for the relevant variables of trading activity should only imply a marginal difference. For example, we could take logarithms in Equation (14) and continue with the derivation of the model for empirical analysis. However, following Andersen et al. (2016), we first estimate the logarithms of variables and then their means.

From the expression in (6) it is clear that $\tilde{I} := P \cdot \tilde{\mathrm{Q}} \cdot \bar{\sigma} \cdot N_B^{-\frac{\xi}{1-\xi}}$. Substituting from (4) for price P and volatility σ, then $\tilde{I} := \psi \cdot P \cdot \sigma \cdot \tilde{\mathrm{Q}} \cdot N_B^{-\frac{\xi}{1-\xi}}$.

If logarithms and expectations are applied, the following liner representation is produced:

$$E\{\log \tilde{I}\} = p + q_B + \frac{s}{2} + \log \psi - \frac{\xi}{1-\xi} n_B \tag{16}$$

where p is the logarithm of prices, $\tilde{q}_B$ is the logarithm of the signed trade size $\tilde{\mathrm{Q}}$, s is the expected logarithmic value of variance σ^2, n_B is the expected logarithmic value of the number of bets N_B, q_B is the expected value of $\tilde{q}_B$.

Solving equation (16) for n_B and q_B yields the following expressions which characterise the implications of the invariance of bets. These expressions hold across time and stocks (proof is presented in Appendix I-4):

$$n_{\mathrm{B}} := \underbrace{c_1 + c_2}_{c} + (1-\xi)w = c + \frac{4\psi}{3\zeta}w \tag{17}$$

where n_B is the logarithm of the number of bets N_B, $c_1 = -(1-\xi)E\{\log \tilde{I}\}$ is a constant term that refers to the expected logarithmic value of the invariant variable $\tilde{I}$, $c_2 := (1-\xi)\left[\log\psi + \log\left(\frac{2}{\zeta}\right)\right]$ is a constant term that refers to the logarithmic value of the square root of ψ^2 and the intermediation multiplier ζ, w is the logarithm of expected trading activity

$$q_B := \underbrace{c_1 + c_2}_{c} + \xi w - \left(p + \frac{s}{2}\right) = c + \left(1 - \frac{4\psi}{3\zeta}\right)w - \left(p + \frac{1}{2}s\right) \tag{18}$$

where q_B is the expected value of the logarithm of the signed trade size $\tilde{\mathrm{Q}}$, p is the logarithm of prices, s is the expected logarithmic value of variance σ^2 and similar to (17) $c_1 := (1-\xi)E\{\log \tilde{I}\}$ is a constant term that refers to the expected logarithmic value of the invariant variable $\tilde{I}$, $c_2 := -(1-\xi)\log\psi + \xi\log\left(\frac{2}{\zeta}\right)$ is a constant term that refers to the logarithmic value of the square root of ψ^2 and the intermediation multiplier ζ, w is the logarithm of expected trading activity.

In equation (17) and (18), c is a constant which is defined in different ways and takes distinct values across equations (for example $c := -(1-\xi)E\{\log \tilde{I}\} + (1-\xi)\left[\log\psi + \log\left(\frac{2}{\zeta}\right)\right]$ in (17), while in (18) $c = (1-\xi)E\left\{\log \tilde{I}\right\} - (1-\xi)\log\psi + \xi\log\left(\frac{2}{\zeta}\right)$).

Both equations formalise the intuition that as trading activity w increases, holding s and p fixed, $\frac{4\psi}{3\zeta}$ of the increase can be attributed to the arrival rate of bets (speed of business time) and $1-\frac{4\psi}{3\zeta}$ to the magnitude of bets. This relationship between the size and number of bets is the cornerstone of market microstructure invariance. For instance, if $\zeta = 2$ then $\frac{2}{3}$ of the increase can be attributed to the arrival rate of bets and $\frac{1}{3}$ to the magnitude of bets. This result is consistent with the findings of both Kyle and Obizhaeva (2013) and Andersen et al. (2016).

The market microstructure invariance equations (17) and (18) refer to low frequency bets. To test the invariance relationship between the number of trades and trading activity, we transform the specific equations using intraday trading variables that are directly observed in a defined sample of $d = 1, \ldots\ldots, D$ trading days and $t = 1, \ldots\ldots, T$ minute intervals within each day:

$$I_{dt} = P_{dt} \cdot Q_{dt} \cdot \sigma_{dt} \cdot N_{dt}^{-\frac{\xi}{1-\xi}} \tag{19}$$

$$E\{\log \tilde{I}_{dt}\} = p_{dt} + q_{dt} + \frac{s_{dt}}{2} + \log \psi - \frac{\xi}{1-\xi} n_{dt} \tag{20}$$

where P_{dt} is the average price over minute t of day d, Q_{dt} (number of shares) is the average trade size over minute t of day d, σ_{dt} is the expected returns volatility over minute t of day , N_{dt} (number of trades per unit of calendar time) is the expected number of trades over minute t of day d, while p_{dt}, q_{dt}, s_{dt} and n_{dt} are the logarithmic forms of the variables

As in Andersen et al. (2016), we assume that synergies between trading strategies of market participants means that the distribution of $\tilde{I}$ retain the same characteristics as in market microstructure invariance (i.e. identical and independent distribution across time). Consequently, I_{dt} and $E\{\log \tilde{I}_{dt}\}$ in equations (19) and (20) respectively, are assumed to remain constant over day d or time-of-day t. Also, the realisations of the employed variables are taken to be directly observable or easily estimated. Market participants select the expected trading size Q_{dt} in a way that reflects their trading requirements and prevailing cirnumstances in the market (i.e. their expectations for the number of trades N_{dt} and the volatility of returns σ_{dt}). However, the expected values of these variables vary considerably for each security as compared to market expectations. Thus their realised values are different from a priori expectations. As a result, good estimators for the number of trades, the volatility of returns and the trading volume are required in order to test the intraday trading hypothesis.

Following closely the methodology of Andersen et al. (2016), active trading participants acquire real-time information on the state of the market and consequently they are able to form unbiased expectations for the number of transactions N_{dt} , trading volume V_{dt} and volatility σ_{dt} over the next minute. For this purpose, 1-minute intervals are used for the estimation of the variables[16]. In this context, the logarithms of a large number of one-minute observations are aggregated and averaged over various days. Averages are taken to mitigate the effect of sampling variation and measurement error. It is assumed that the large fluctuations during the day that remain after aggregation, provide an indication of fluctuations in the market's expectations. The goal is not to investigate whether intraday trading invariance relationship holds for every time interval, but rather during the whole time span.

If $\tilde{n}_{dt}$ is the log number of trades for minute t of day d observed in transactions data, then by averaging the observations for this interval across all the days in the sample:

$$n_t = \frac{1}{D}\cdot\sum_{d=1}^{D}\tilde{n}_{dt} \ , \quad t = 1,......,D \qquad \textbf{(21)}$$

where n_t is the average log trades count for intraday interval t,

In a similar fashion we specify the variables s_t, q_t, υ_t and w_t as follows:

$$s_t = \frac{1}{D}\cdot\sum_{d=1}^{D}\tilde{s}_{dt} \ , \quad q_t = \frac{1}{D}\cdot\sum_{d=1}^{D}\tilde{q}_{dt} \ , \quad \upsilon_t = \frac{1}{D}\cdot\sum_{d=1}^{D}\tilde{\upsilon}_{dt} \ , \quad w_t = \frac{1}{D}\cdot\sum_{d=1}^{D}\tilde{w}_{dt} \ , \quad t = 1,......,D \qquad \textbf{(22)}$$

where s_t is the average log realised volatility, q_t is the average log average trading size, υ_t is the average trading volume, w_t is the average log trading activity, for intraday interval t, respectively.

Alternatively, instead of averaging across all days, the number of trades can be averaged across all intraday intervals on trading day d :

[16] Lower frequencies yield measures with smaller errors, they also cause an upward bias in estimators.

$$n_d = \frac{1}{T} \cdot \sum_{t=1}^{T} \tilde{n}_{dt} \ , \ d = 1,......,T \tag{23}$$

The variables s_d, q_d, υ_d and w_d are specified in analogous fashion:

$$s_d = \frac{1}{T} \cdot \sum_{t=1}^{T} \tilde{s}_{dt} \ , \ q_d = \frac{1}{T} \cdot \sum_{t=1}^{T} \tilde{q}_{dt} \ , \ \upsilon_d = \frac{1}{T} \cdot \sum_{t=1}^{T} \tilde{\upsilon}_{dt} \ , w_d = \frac{1}{T} \cdot \sum_{t=1}^{T} \tilde{w}_{dt} \tag{24}$$

As in Andersen et al. (2001), the realised volatility σ_{dt} is calculated from 10-second returns,. Specifically, after filtering the data for outliers, we obtain prices in each 10-second time mark are obtained by taking the log average of the respective bid and ask quotes. Given that tick-by-tick data is not generally provided in continuously-spaced distinct time points, a number of bid and ask quotes are not available at the specific 10-second time mark. In some cases, the required midpoints are obtained by linearly interpolating between the previous and next available midpoint. We estimate the "continuously-compounded returns" using the difference between the 10th and 1st midpoints in each 10-second time interval. The realised volatility σ_{dt} estimator for each minute is then defined as the sum of squared 10-second returns (i.e. six squared returns per minute).

The returns are computed from bid and ask quotes and not from trade prices to avoid bid–ask bounce and stale prices effects that might yield a biased realised volatility estimator (Zhou, 1996; Andersen et al., 2000). The 10-second mark for the estimation of realised volatility is chosen to account for microstructure noise. Although the defined estimator suffers from measurement error in terms of actual local volatility, the sum of realised volatility across different one-minute intervals ensures its overall accuracy. This is the same "error diversification principle" as in Andersen et al. (2016).

Finally, we eliminate minutes without trades (i.e. $N_{dt} = 0$) or 10-second time points during which the realised volatility is zero from the sample, as otherwise the log specification is not feasible. Despite the fact that intraday trading invariance should hold in principle for any selected sample, omitting observations based on the above criterion will theoretically induce an increased bias in the estimator. Nevertheless, including periods with no or low trading activity in the sample might cause several complications for intraday trading invariance in the specific time intervals.

Based on the above analysis, we now model the main invariance hypothesis under investigation. This implies proportionality between the number of trades n_j and trading activity w_j, as follows:

Model 1:

$$n_j := c + \frac{4\psi}{3\zeta} w_j + u_j^n \quad \textbf{(25)}$$

where j represents either different intraday intervals (i.e. $j = t,\ t = 1, \ldots.., T$) or distinct trading days (i.e. $j = d,\ t = 1, \ldots.., D$) and u_j^n are the regression residuals.

4.2 Alternative Hypotheses

As already explained in literature review, there is a variety of time deformation and microstructure theories that examine the relationship between trading activity and volatility. Clark (1973) investigates the relationship between the trading volume and returns volatility and concludes that the expected volume can be used as a proxy for business clock, implying that it is directly proportional to return variation (i.e. $\sigma_{dt}^2 \sim V_{dt}$). This proportionality can be expressed in logarithmic terms as: $s_{dt} = c + \upsilon_{dt} = c + n_{dt} + q_{dt}$. Based on this we specify the following first alternative to the main hypothesis with similar representation as in (25) (proof is presented in Appendix I-5):

Model 2 (Clark, 1973):

$$n_j := c + \frac{4\psi}{3\zeta}\left[\frac{\zeta}{2\psi} w_j - \frac{3\zeta}{4\psi} q_j\right] + u_j^n \quad \textbf{(26)}$$

For $\zeta = 2$, equation (26) becomes the model analysed in Andersen et al. (2016):

$$n_j = c + \frac{2}{3}\left[w_j - \frac{3}{2} q_j\right] + u_j^n \quad \textbf{(27)}$$

In contrast to Clark (1973), other papers argue that the trade count is a better proxy for a business clock. For example, Ané and Geman (2000), based on earlier work by Jones et al. (1994), report a significant relationship between the number of trades and returns variations. Their empirical hypothesis indicate that the expected number of transactions is proportional to return variations (i.e. $\sigma_{dt}^2 \sim N_{dt}$). This proportionality is expressed in logarithmic terms as: $s_{dt} = c + n_{dt}$. As such, we propose the following second alternative to the main hypothesis with a similar representation as in (25) (proof is presented in Appendix I-6):

Model 3 (Ané and Geman, 2000):

$$n_j := c + \frac{4\psi}{3\zeta}\left[\frac{\zeta}{2\psi}w_j - \frac{\zeta}{2\psi}q_j\right] + u_j^n \qquad \textbf{(28)}$$

For $\zeta = 2$, equation (25) becomes the model proposed by Andersen et al. (2016):

$$n_j = c + \frac{2}{3}\left[w_j - q_j\right] + u_j^n \qquad \textbf{(29)}$$

Given that intraday invariance equation (19), it can be inferred that if the expected trade size (Q_{dt}) does not vary, there is a proportionality between expected return volatility and transactions count ($\sigma_{dt}^2 \sim N_{dt}$) and between expected return volatility and trading volume ($\sigma_{dt}^2 \sim V_{dt}$). In this case, the principles of Clark (1973), Ané and Geman (2000) and intraday trading invariance are equivalent. On the contrary, if there is a correlation between variations in trade size and variations in return volatility and number of trades, then the aforementioned proportionality is no longer present. This is more likely to occur if the traders actively control for their risk exposures in business time, which in turn leads to a systematic variation of trade size with volume and volatility. In this case, the three theories will have different implications.

4.3 Hypotheses testing

We examine two different variations of the invariance and alternative hypotheses. The first variation corresponds to the invariance relationship introduced by Kyle and Obizhaeva (2013) and tested empirically for trades by Andersen et al. (2016):

$H_0 : \beta_1 = 2/3$

$H_1 : \beta_1 \neq 2/3$

The specific proportionality is based on the assumption that there is only one market maker (i.e. $\zeta = 2$) and that all price changes stem from order flow imbalances (i.e. $\psi = 1$)

The second variation conjectures that the relationship between the number of trades and trading activity is $1/2$. Thus,

$H_0 : \beta_2 = 1/2$

$H_1 : \beta_2 \neq 1/2$

Intuitively this means that approximately, an order is intermediated in the market $\zeta = 2.56$ times and that only $\psi = 0.96$ of the price changes is attributed to order flow imbalances. The rest 4% is coming from overnight announcements[17].

5. The FTSE 100 Data

In this paper, we use time-stamped tick data from Thompson Reuters Tick History for 25 stocks listed on FTSE 100 and traded on LSE (Appendix II-Table 1). The dataset includes tick-by-tick information on the best bid and ask quotes, prices, and trading volume (in shares), between 1st January 2007 and 31st December 2009 (3 years)[18]. The stocks in our sample are picked based on two criteria. First, they are regular constituents of FTSE 100 during the employed period (i.e. they have not been dropped from the index at any point). Thus, any survivorship bias that may affect the results is eliminated. Second, they represent 51% of the total market capitalisation of the index. After applying the first criterion, we rank the remaining 71 stocks based on their market capitalisation from high to low. Then, the stocks are classified into five groups, with the first group having the stocks with the highest market capitalisation and the fifth the stocks with the lowest market capitalisation

[17] The value of the percentage of returns variations due to order flow imbalances is in line with Bouchaud et al. (2008), who assert that the majority of return variations come from the supply and demand imbalances.

[18] In contrast to Andersen et al. (2015), Thompson Reuters Tick History does not report the aggregate ticks for each price level. As a result, the number of trades demonstrates also the order flow stemming from intermediation in the supply side. According to Andersen et al. (2015), this will lead to an increase of the transaction counts and decrease of trade size from the levels that invariance as a concept implies. However, reporting trades with same price as one quantity does not consider the case when the demand side requires more than the offered quantity of shares. In this case, their limit order at the top of the order book is matched with a different marketable order which may or not be at a same price with the first order. Thus, the intermediation of the demand side will not manifest itself. If we aggregate the trades at same price that will deflate the number of trades that refer to the same marketable offer and inflate the trade size.

respectively. Five stocks are selected from each group so that our sample of 25 stocks accounts for more than 50% of the market capitalisation of the index.

We consider only trades for these stocks, between 8am and 4.30pm, five days a week, because only during this time span London Stock Exchange is open for continuous trading. Also, we exclude from our sample 30 days that correspond to holidays or other days with low trading activity (low volume and short trading hours), which leads to a total of 754 trading days. Each trading day is further divided into 510 one-minute intervals. During these intervals, we aggregate the observations for trading volume V , number of trades N and average trade size Q so that we estimate their one-minute values. The realised return volatility σ for each minute is computed based on the approach explained in subchapter 4.1.

The descriptive statistics of the dataset of 25 FTSE 100 stocks are depicted in Table 2 and Table 3 in Appendix II. Table 2 includes 1-minute averages and their standard deviations, with variables estimated using equations (21) and (22). Table 3 presents daily averages and their standard deviation, with variables estimated using equations (23) and (24). The stocks are reported in order, based on their market capitalisation, with RDSA having the highest and REX the lowest market capitalisation. Average annualised volatility for each stock is slightly less when volatility is estimated across all intraday intervals on trading day (Table 3), but standard deviation of the estimated values increases. However, it remains between 0.23 to 0.42 for all stocks, regardless of whether volatility is estimated by averaging observations across days for each minute (Table 2) or across intraday intervals on a day (Table 3). The case is similar for the other variables, which have minor differences in their mean values in Table 2 and Table 3 and higher standard deviation of means, when the respective variables are estimated across intraday intervals. Intuitively, the variables appear on average to fluctuate more during the course of trading day than across days for the same 1-minute interval.

The intraday innovations of trading volume, V , number of trades N , average trade size Q and annualised realised volatility σ, for all 25 stocks, are presented, in Figure 1. The substantive graphs illustrate the respective variables for each minute between 8:00 and 16:30 (calendar time). The observations for each variable are averaged for every 1-minute interval across all trading days in the dataset, using the equations in (21) and (22). There is a clear upward trend for all stocks between 13.30 and 15.00, more obvious in the number of trades N and annualised realised volatility, σ. That is because the specific time span coincides with the start of active trading in the US stock exchanges and/or because during this period most macroeconomic announcements and/or companies' announcements regarding key investor data are being realised. This fashion is less apparent for the trading volume, V , though the majority of the stocks exhibit an upturn in trading volume after 13:00. As for the average trade size Q, there are spikes between 13:30 and 15:00 for some specific stocks, but they are not persistent. This can be attributed either to same reasons as the other variables or to stock specific characteristics. The trends in the trading volume V and average trade size Q can be better observed in Figure 2 in Appendix II, in which the stocks with extreme values for each variable have been plotted in different graphs. In total, the intraday patterns for trading volume V , number of trades N and annualised realised volatility σ are mostly similar, while average trade size Q seems to be affected more by individual stock characteristics.

If we average daily the intraday observations of the same variables during the whole time span, then the time series in Figure 3 are produced. The number of trades and the volatility of returns increase during the financial crisis (September 2008-February 2009) for the majority of the stocks, while the trading volume and the average trade size are higher during the period before the financial crisis[19]. Also, during the sample period the number of trades, return volatility and, to some degree, the trading volume, appear to follow similar pattern, in relation to the occurrence of specific major events, as compared to that of the average trade size, which follows an opposing direction of change. For example, the overnight injection of liquidity by the ECB in August 2007 leads to an increase in the number of trades, the returns volatility and the trading volume. Similar spikes are present in the respective graphs of these variables during the Société Générale scandal in January 2008 and the bankruptcy of Lehman Brothers in September 2008. In contrast, the average trade size decreases during the same events for the majority of stocks. Finally, all variables show a drop during the end of each year, which is more obvious in the number of trades.

[19] Similar to the averages across days the trends in the trading volume V and average trade size Q can be better observed in Figure 4 in Appendix II, in which the stocks with extreme values for each variable have been plotted in different graphs

Figure 1-The figure shows averages across days for the number of trades N_t, volume V_t, trade size Q_t and annualised volatility σ_t for all stocks in the sample per one minute interval.

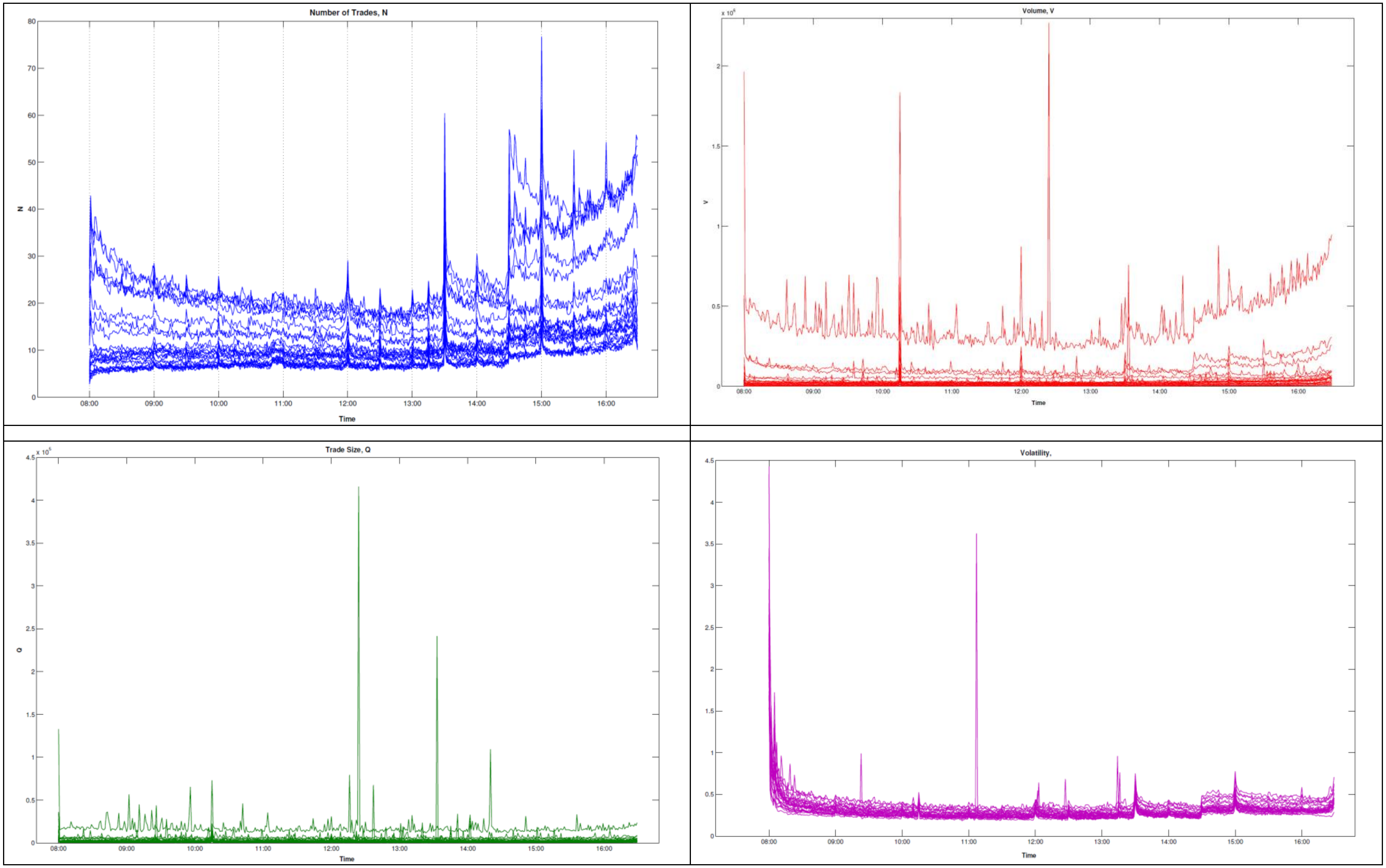
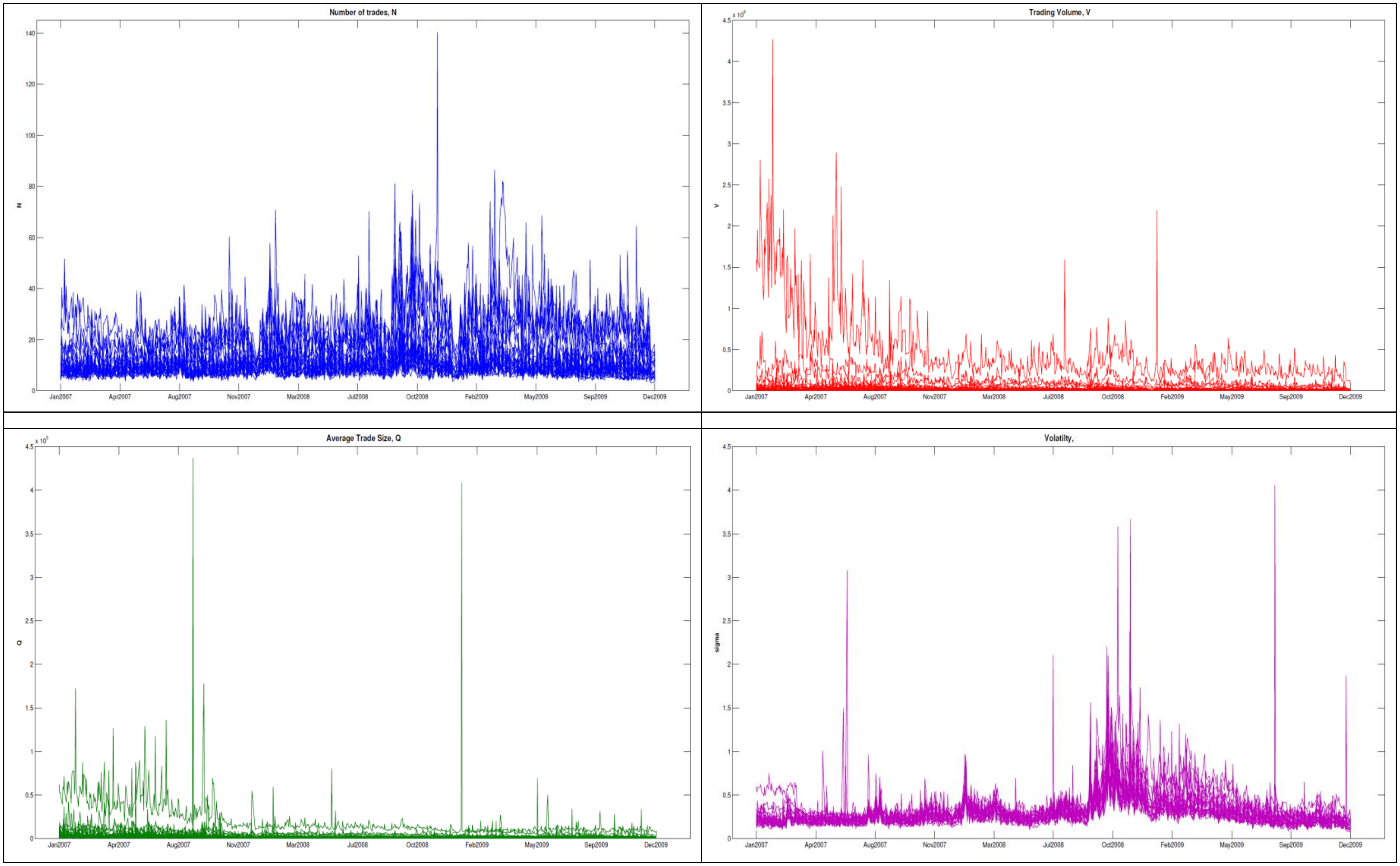


Figure 3- The figure shows averages intraday for the number of trades N_d, volume V_d , trade size Q_d and annulaised volatility σ_d for all stocks in the sample per day.



6. Empirical Results

This section focuses on an analysis and comparison of the empirical analysis of the invariance relationship in (25) and the alternative models of Clark (1973) and Ané and Geman (2000), in (26) and (28), respectively. We start the analysis by investigating which of the two assumptions (i.e. $\beta_1 = 2/3$ or $\beta_2 = 1/2$) produces a proportionality that empirically describes the invariance relationship (i.e. the relationship between the number of trades and trading activity) in our sample. Subsequently, we compare the OLS regression coefficients of the Clark (1973) and Ané and Geman (2000) models with those from the invariance model, to see which of the three models best capture the market microstructure properties in this specific market. We do this by employing variables estimated as average across days based on equations (21) and (22) and variables calculated as averages intraday based on equations (23) and (24).

6.1 Number of trades and trading activity for each intraday interval: Main hypothesis

First, we investigate if there exists a proportionality between the number of trades and trading activity based on the invariance model in (25). Results of OLS regressions for the number of trades regarding the 25 FTSE 100 stocks are presented in Table 4. The underlying variables are averages of respective observations for 1-minute and 5-minutes intervals across all days in the sample, as defined by equations (21) and (22). The constant term and coefficients are the same when testing for 2/3 and 1/2 proportionality, as the regression model for invariance does not change based on our theoretical model. Thus, in Table 4, we only report one regression result as estimates for models for a 2/3 and 1/2 proportionality are the same.

The estimated coefficients do not confirm the $2/3$ proportionality between the transaction counts and trading activity in the specific sample. The null hypothesis for $\beta_1 = 2/3$ is rejected for all stocks in both 1-minute and 5-minutes intervals at 1% significance level. On the contrary, the null hypothesis for $\beta_2 = 1/2$ is accepted for one third of the stocks based on 1-minute intervals and 20 out of 25 stocks based on 5-minute intervals at 5% significance level. In our sample, as already explained, minute intervals that do not include any trades and during which realised volatility is zero are omitted. The exclusion of periods with zero trades and minute intervals subject to zero realised volatility generates an upward bias in the estimated averages and estimated volatilities in the next intervals, respectively.

Table 5 in Appendix II reports the percentage of intervals excluded when averaging for 1-minute and 5-minutes for each stock. As the percentage of exclusions in 1-minute intervals is greater, in turn estimation becomes less accurate. This would explain why the adjusted R-squared for 1-minute intervals, especially for some stocks, is lower as compared to the equivalent for 5-minutes intervals, and why the coefficients (i.e. estimated proportionality) do not converge to a certain value. In contrast to Andersen et al. (2016), who report the same $2/3$ proportionality for 1-minute and 5-minutes intervals, we find that in our sample the relationship between the number of trades and trading activity is not consistent in the high-frequency (i.e. 1-minute) intervals. The reason is most likely that FTSE 100 stocks are less actively traded compared to the E-mini S&P future contracts analysed in Andersen et al. (2016). The number of observations with zero trades and intervals with zero realised volatility are fewer in their sample, thus their results suffer from smaller bias in high frequencies. Averaging the observations in the low frequency 5 minutes intervals, yields a on average less significant constant term. Given that the constant term represents in part the expected logarithmic value of the invariant variable $\tilde{I}$ (i.e. $E\{\log \tilde{I}\}$), this means that the specific variable is also not significantly different from zero or that is equal to $-\log\psi - \log\left(\frac{2}{\zeta}\right)$, where ζ is the intermediation multiplier and ψ the percentage of order flow imbalances attributed to trades. Consequently, based on the invariance model, only 17 stocks out of the 25 stocks examined, show a robust $1/2$ relationship between the number of trades and trading activity. Consistent with Kyle and Obizhaeva (2013) conjecture, market capitalisation does not appear to be a factor affecting the invariance relationship, as stocks with different market capitalisation exhibit similar $1/2$ proportionality.

Taking into account the findings in Table 4, we conjecture that the invariance relationship in our sample is closer to $1/2$ rather than 2/3. Intuitively, this means that half of the variations in trading activity are attributed to the arrival rate of trades and half to the average trade size. Therefore, the invariance model in (25) takes the following expression:

Model 1 (Invariance):

$$n_j = c + \frac{1}{2} w_j + u_j^n \quad \textbf{(30)}$$

Moving one step forward we now compare the OLS regression coefficients estimated for the models of Clark (1973) and Ané and Geman (2000) with those obtained for the invariance model. Accordingly, we re-specify the models of Clark (1973) in equation (26) and Ané and Geman (2000) in equation (28), respectively, so as to ensure similar $2/3$ and $1/2$ representations. The purpose here is to decide which of the underlying theories that are represented by these models better captures the microstructure properties in the specific sample. The resulting specifications for the models of Clark (1973) and Ané and Geman (2000) are:

Model 2 (Clark, 1973):

For $2/3$ proportionality:

$$n := c + \frac{2}{3}\left[w_j - \frac{3}{2} q_j \right] + u_j^n \quad \textbf{(31)}$$

For $1/2$ proportionality:

$$n := c + \frac{1}{2}\left[\frac{4}{3} w_j - 2 q_j \right] + u_j^n \quad \textbf{(32)}$$

Model 3 (Ané and Geman, 2000):

For $2/3$ proportionality:

$$n_j := c + \frac{2}{3}\left[w_j - q_j \right] + u_j^n \quad \textbf{(33)}$$

For $1/2$ proportionality:

$$n_j := c + \frac{1}{2}\left[\frac{4}{3} w_j - \frac{4}{3} q_j \right] + u_j^n \quad \textbf{(34)}$$

Table 4-OLS Regression results for Model 1 (Invariance). Variables estimated as averages across days. Stocks are grouped by market capitalization. ***c*** is the constant term of the invariance model in (25). $\beta_2 = 4\psi / 3\zeta$ is the coefficient (i.e. proportionality) of the invariance model in (25) and refers to the null hypothesis of 1/2 proportionality. Significance against 2/3 or 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis of 1/2 proportionality is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute				***5 minutes***		
Groups	***Stocks***	***c***	***β_2***	***$\bar{R}^2$***		***c***	***β_2***	***$\bar{R}^2$***
Group 1 Highest Mkt Cap	**RDSA**	-0.0919* (0.0385)	**0.4928** (0.0103)	0.8194		-0.1501 (0.1154)	0.5417* (0.0186)	0.8934
	BP	0.2414*** (0.0449)	**0.5090** (0.0086)	0.8721		0.2451 (0.1458)	**0.5338** (0.0186)	0.8904
	HSBA	0.3556*** (0.0439)	**0.4877** (0.0084)	0.8680		0.3944* (0.1521)	**0.5159** (0.0193)	0.8763
	GSK	0.0917 (0.0480)	0.5409*** (0.0103)	0.8443		0.2550 (0.1604)	**0.5404** (0.0224)	0.8525
	VOD	0.7540*** (0.0505)	0.4481*** (0.0106)	0.7772		0.8613*** (0.1710)	**0.4749** (0.0226)	0.8139
Group 2 Upper Middle Mkt Cap	**BLT**	-0.0285 (0.0446)	0.5468*** (0.0083)	0.8957		0.0150 (0.1546)	0.5617** (0.0191)	0.8956
	BG	0.2110*** (0.0530)	**0.4892** (0.0121)	0.7639		0.5369** (0.1725)	**0.4851** (0.0251)	0.7873
	XTA	0.4553*** (0.0450)	0.4510*** (0.0091)	0.8288		0.6519*** (0.1559)	**0.47585** (0.0199)	0.8491
	NG	0.1792*** (0.0463)	0.5261* (0.0127)	0.7721		0.4076* (0.1654)	**0.5242** (0.0281)	0.7751
	STAN	0.4137*** (0.0470)	0.4466*** (0.0110)	0.7648		0.6002*** (0.1480)	**0.4744** (0.0216)	0.8266
Group 3 Middle Mkt Cap	**EMG**	0.5408*** (0.0469)	0.4054*** (0.0133)	0.6464		0.5842*** (0.1591)	**0.4690** (0.0266)	0.7546
	OML	1.0820*** (0.0422)	0.2826*** (0.0152)	0.4054		0.9239*** (0.1489)	0.4327* (0.0280)	0.7023
	WPP	0.2805*** (0.0447)	**0.5087** (0.0130)	0.7512		0.4007* (0.1582)	**0.5230** (0.0274)	0.7826
	BLND	0.5727*** (0.0514)	0.4256*** (0.0157)	0.5908		0.6169** (0.1909)	**0.4841** (0.0331)	0.6786
	RR	0.4673*** (0.0404)	0.4513*** (0.0125)	0.7176		0.5413*** (0.1520)	**0.4960** (0.0268)	0.7714
Group 4 Lower Middle Mkt Cap	**CCL**	0.2531*** (0.0305)	**0.5157** (0.0104)	0.8278		0.1597 (0.1148)	0.5605** (0.0220)	0.8655
	SMIN	0.4539*** (0.0389)	0.4480*** (0.0146)	0.6503		0.3997** (0.1420)	**0.5170** (0.0298)	0.7477
	SHP	0.3269*** (0.0395)	**0.4800** (0.0124)	0.7454		0.3214* (0.1263)	**0.5253** (0.0238)	0.8281
	IPR	0.4932*** (0.0461)	0.4571** (0.0154)	0.6331		0.5484** (0.1661)	**0.5040** (0.0314)	0.7180
	IMT	0.4036*** (0.0486)	**0.4827** (0.0138)	0.7065		0.5722*** (0.1714)	**0.4990** (0.0289)	0.7459
Group 5 Lowest Middle Mkt Cap	**SVT**	0.2871*** (0.0380)	0.5322* (0.0147)	0.7194		0.2360 (0.1393)	0.5712* (0.0302)	0.7798
	CNE	0.4461*** (0.0400)	0.4343*** (0.0146)	0.6357		0.3371* (0.1516)	**0.5155** (0.0300)	0.7443
	JMAT	0.3525*** (0.0396)	**0.4959** (0.0155)	0.6688		0.2338 (0.1439)	**0.5585** (0.0306)	0.7670
	SGE	0.6104*** (0.0362)	0.4324*** (0.0156)	0.6022		0.3467* (0.1351)	**0.5496** (0.0303)	0.7649
	REX	0.5797*** (0.0338)	0.4233*** (0.0140)	0.6417		0.4587*** (0.1295)	**0.5163** (0.0284)	0.7656

Tables 6 and 7 depict the results of OLS regressions for the number of trades regarding the 25 FTSE 100 stocks based on the models of Clark (1973) and Ané and Geman (2000), respectively. The underlying variables here are also averages of respective observations for 1-minute and 5-minutes intervals across all days in the sample. As in the invariance model, the estimation of variables in low frequency 5-minutes intervals is more accurate than that of 1-minute, because more observations are included in the analysis. This is clearly demonstrated by the adjusted R-squared, which improves for both models when we average the observations in lower frequency. Both models fit the data better compared to the invariance model, their constant term is significant, while they can predict both $2/3$ and $1/2$ relationships (for the same stocks), especially in 5-minutes intervals. The representation of models that refers to $1/2$ proportionality yields coefficients with lower standard errors. This is an indication that our assumption of $1/2$ and not $2/3$ relationship between the number of trades and trading activity is appropriate for the specific sample. Also, specifying the models of Clark (1973) and Ané and Geman (2000) in invariance terms simplifies their empirical testing, while the high values of adjusted R-squared for the models show that invariance, as a principle, improves their performance. This is in line with Kyle and Obizhaeva (2013), who state that invariance is a principle that can complement and be applied to different models in way that facilitates their empirical testing.

The model of Clark (1973), accepts the null hypothesis of $\beta_2 = 1/2$ for the majority of stocks (19 out of 25) in 5-minutes interval, while the model of Ané and Geman (2000) accepts it only for 1/3 of the stocks at the same frequency. Also, the model of Ané and Geman (2000) performs better for lower market capitalisation stocks. Among the three models, the model of Clark (1973) appears to have higher estimation precision in predicting the required $1/2$ proportionality[20], notably for the very high capitalisation stocks of RDSA and BP. An interesting fact is that all three models predict $1/2$ proportionality for certain stocks[21]. Intuitively, this means that for the specific stocks the number of trades and volume are proportional to returns variation, while the number of trades is also proportional to trading activity. This can be attributed to stock specific characteristics the

[20] It empirically predicts 1/2 proportionality for two more stocks compared to the invariance model, while its adjusted R- squared is higher.

[21] The only stock for which the required proportionality is not empirically predicted by any of the three models is OML.

Table 6- OLS Regression results for Model 2 (Clark, 1973). Variables estimated as averages across days. Stocks are grouped by market capitalization. $\boldsymbol{c}$ is the constant term in both (31) and (32). β_1 is the coefficient of Clark's model specification in (31), referring to a 2/3 proportionality. β_2 is the coefficient of Clark's model specification in (32) referring to a 1/2 proportionality. Significance against 2/3 or 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis (β_1=2/3 or β_2=1/2) is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute					***5 minutes***			
Groups	***Stocks***	***c***	***β₁***	***β₂***	$\bar{R}^2$		***c***	***β₁***	***β₂***	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	6.1758*** (0.0529)	0.7025*** (0.0084)	0.5269*** (0.0063)	0.9328		6.0397*** (0.0678)	**0.6757** (0.0159)	**0.5068** (0.0119)	0.9470
	BP	7.7796*** (0.0456)	0.6997*** (0.0065)	0.5248*** (0.0049)	0.9582		7.5912*** (0.0801)	**0.6697** (0.0168)	**0.5023** (0.0126)	0.9405
	HSBA	7.4500*** (0.0391)	0.6862*** (0.0058)	0.5147*** (0.0044)	0.9644		7.2880*** (0.0598)	**0.6593** (0.0138)	**0.4945** (0.0103)	0.9577
	GSK	6.9618*** (0.0502)	0.7263*** (0.0083)	0.5447*** (0.0063)	0.9371		6.5978*** (0.0785)	**0.6525** (0.0205)	**0.4893** (0.0154)	0.9093
	VOD	8.4832*** (0.0992)	0.6175*** (0.0109)	0.4631*** (0.0082)	0.8636		8.5106*** (0.1336)	0.6151* (0.0201)	0.4613* (0.0151)	0.9022
Group 2 Upper Middle Mkt Cap	**BLT**	6.4200*** (0.0257)	0.6931*** (0.0050)	0.5198*** (0.0038)	0.9740		6.2402*** (0.0330)	**0.6489** (0.0124)	**0.4867** (0.0093)	0.9646
	BG	6.4807*** (0.0693)	**0.6678** (0.0112)	**0.5008** (0.0084)	0.8750		6.2751*** (0.0935)	0.6043** (0.0233)	0.4532** (0.0175)	0.8696
	XTA	5.6749*** (0.0388)	0.6115*** (0.0079)	0.4586*** (0.0059)	0.9223		5.7106*** (0.0419)	0.5905*** (0.0179)	0.4429*** (0.0134)	0.9153
	NG	7.0318*** (0.0856)	0.7005** (0.0121)	0.5254** (0.0091)	0.8681		6.7254*** (0.1424)	**0.6375** (0.0279)	**0.4781** (0.0209)	0.8376
	STAN	5.8203*** (0.0472)	0.6207*** (0.0083)	0.4656*** (0.0063)	0.9160		5.8275*** (0.0590)	0.5929*** (0.0174)	0.4447*** (0.0130)	0.9202
Group 3 Middle Mkt Cap	**EMG**	6.4432*** (0.1140)	0.5836*** (0.0148)	0.4377*** (0.0111)	0.7526		6.6737*** (0.1487)	0.5972* (0.0268)	0.4479* (0.0201)	0.8305
	OML	5.4404*** (0.1781)	0.3669*** (0.0182)	0.2752*** (0.0137)	0.4424		7.2265*** (0.2073)	0.5355*** (0.0276)	0.4016*** (0.0207)	0.7882
	WPP	7.1162*** (0.1050)	0.7009* (0.0144)	0.5257* (0.0108)	0.8230		6.8630*** (0.1531)	**0.6638** (0.0292)	**0.4978** (0.0219)	0.8360
	BLND	6.2059*** (0.1283)	0.6263* (0.0189)	0.4697* (0.0142)	0.6836		6.2907*** (0.1542)	**0.6328** (0.0335)	**0.4746** (0.0251)	0.7786
	RR	6.7851*** (0.1150)	0.6204** (0.0146)	0.4653** (0.0110)	0.7802		6.8781*** (0.1597)	**0.6202** (0.0278)	**0.4651** (0.0209)	0.8306
Group 4 Lower Middle Mkt Cap	**CCL**	5.6603*** (0.0691)	0.6921* (0.0121)	0.5191* (0.0091)	0.8647		5.5263*** (0.0823)	**0.6841** (0.0223)	**0.5131** (0.0167)	0.9030
	SMIN	5.9425*** (0.1245)	**0.6336** (0.0183)	**0.4752** (0.0137)	0.7019		6.2390*** (0.1532)	**0.6797** (0.0305)	**0.5098** (0.0228)	0.8311
	SHP	6.4656*** (0.1060)	0.6988* (0.0160)	0.5241* (0.0120)	0.7899		6.2962*** (0.1312)	**0.6789** (0.0276)	**0.5091** (0.0207)	0.8572
	IPR	6.9978*** (0.1492)	**0.6316** (0.0183)	**0.4737** **(**0.0137)	0.7009		7.0746*** (0.2010)	**0.6322** (0.0327)	**0.4741** (0.0245)	0.7869
	IMT	6.0077*** (0.0899)	0.7056* (0.0162)	0.5292* (0.0121)	0.7889		5.7213*** (0.1064)	**0.6458** (0.0308)	**0.4844** (0.0231)	0.8126
Group 5 Lowest Middle Mkt Cap	**SVT**	5.8940*** (0.1030)	**0.6751** (0.0163)	**0.5063** (0.0123)	0.7702		5.9539*** (0.1317)	**0.6854** (0.0289)	**0.5140** (0.0217)	0.8475
	CNE	4.8999*** (0.1131)	0.5947*** (0.0205)	0.4460*** (0.0154)	0.6223		5.2562*** (0.1242)	**0.6673** (0.0352)	**0.5005** (0.0264)	0.7806
	JMAT	5.2180*** (0.1089)	0.5966*** (0.0180)	0.4475*** (0.0135)	0.6841		5.5385*** (0.1287)	**0.6467** (0.0305)	**0.4850** (0.0229)	0.8159
	SGE	6.5369*** (0.1807)	0.5537*** (0.0203)	0.4152*** (0.0152)	0.5946		7.5599*** (0.2234)	**0.6753** (0.0314)	**0.5065** (0.0236)	0.8202
	REX	6.1611*** (0.1639)	0.5669*** (0.0203)	0.4252*** (0.0152)	0.6049		6.8040*** (0.1999)	**0.6455** (0.0321)	**0.4841** (0.0240)	0.8002

Table 7- OLS Regression results for Model 3 (Ané and Geman, 2000). Variables estimated as averages across days. Stocks are grouped by market capitalization. ***c*** is the constant term in both (33) and (34). β_1 is the coefficient of Ané and Geman's model specification in (33), referring to a 2/3 proportionality. β_2 is the coefficient of Ané and Geman's model specification in (34), referring to a 1/2 proportionality. Significance against 2/3 or 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute					***5 minutes***			
Groups	***Stocks***	***c***	β_1	β_2	$\bar{R}^2$		***c***	β_1	β_2	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	3.6204*** (0.0263)	0.6330*** (0.0088)	0.4748*** (0.0066)	0.9113		3.6656*** (0.0164)	0.6287* (0.0167)	0.4716* (0.0126)	0.9332
	BP	4.7418*** (0.0215)	0.6331*** (0.0071)	0.4748*** (0.0053)	0.9396		4.7618*** (0.0138)	0.6214* (0.0173)	0.4660* (0.0130)	0.9272
	HSBA	4.5595*** (0.0181)	0.6186*** (0.0065)	0.4640*** (0.0049)	0.9467		4.6021*** (0.0096)	0.6109*** (0.0157)	0.4582*** (0.0118)	0.9376
	GSK	4.2366*** (0.0219)	**0.6658** (0.0088)	**0.4994** (0.0066)	0.9193		4.2107*** (0.0119)	0.6144* (0.0210)	0.4608* (0.0158)	0.8940
	VOD	5.3737*** (0.0484)	0.5580*** (0.0107)	0.4185*** (0.0080)	0.8429		5.5129 (0.0420)	0.5648*** (0.0212)	0.4236*** (0.0159)	0.8750
Group 2 Upper Middle Mkt Cap	**BLT**	3.9329*** (0.0103)	0.6453*** (0.0061)	0.4840*** (0.0046)	0.9570		3.9559*** (0.0167)	0.6222** (0.0146)	0.4666** (0.0109)	0.9475
	BG	3.9844*** (0.0302)	0.6130*** (0.0112)	0.4598*** (0.0084)	0.8546		4.0743*** (0.0154)	0.5627*** (0.0240)	0.4220*** (0.0180)	0.8443
	XTA	3.5807*** (0.0141)	0.5556*** (0.0083)	0.4167*** (0.0062)	0.8990		3.7700*** (0.0224)	0.5500*** (0.0187)	0.4125*** (0.0140)	0.8956
	NG	4.3703*** (0.0413)	**0.6529** (0.0117)	**0.4897** (0.0088)	0.8588		4.3425*** (0.0422)	0.5990* (0.0279)	0.4492* (0.0209)	0.8199
	STAN	3.6424*** (0.0221)	0.5659*** (0.0093)	0.4244*** (0.0070)	0.8798		3.8170*** (0.0105)	0.5527*** (0.0190)	0.4145*** (0.0143)	0.8929
Group 3 Middle Mkt Cap	**EMG**	4.0371*** (0.0561)	0.5253*** (0.0141)	0.3940*** (0.0106)	0.7307		4.3115*** (0.0479)	0.5511*** (0.0268)	0.4133*** (0.0201)	0.8066
	OML	3.7753*** (0.0961)	0.3427*** (0.0171)	0.2571*** (0.0129)	0.4394		4.8297*** (0.0912)	0.5000*** (0.0278)	0.3750*** (0.0209)	0.7613
	WPP	4.3854*** (0.0506)	0.6380* (0.0135)	0.4785* (0.0102)	0.8135		4.3596*** (0.0472)	**0.6126** (0.0285)	**0.4594** (0.0214)	0.8199
	BLND	3.8800*** (0.0608)	0.5592*** (0.0176)	0.4194*** (0.0138)	0.6657		4.0541*** (0.0418)	0.5792* (0.0336)	0.4344* (0.0252)	0.7458
	RR	4.2765*** (0.0567)	0.5671*** (0.0135)	0.4253*** (0.0101)	0.7761		4.4472*** (0.0551)	0.5759** (0.0275)	0.4319** (0.0206)	0.8128
Group 4 Lower Middle Mkt Cap	**CCL**	3.5075*** (0.0324)	0.6279*** (0.0113)	0.4709*** (0.0085)	0.8590		3.4935*** (0.0230)	**0.6390** (0.0222)	**0.4792** (0.0167)	0.8912
	SMIN	3.7560*** (0.0603)	0.5805*** (0.0165)	0.4354*** (0.0123)	0.7098		3.9365*** (0.0557)	**0.6229** (0.0302)	**0.4671** (0.0226)	0.8081
	SHP	3.9337*** (0.0481)	0.6227*** (0.0142)	0.4670** (0.0106)	0.7916		3.9553*** (0.0396)	**0.6218** (0.0260)	**0.4663** (0.0195)	0.8502
	IPR	4.4354*** (0.0755)	0.5818*** (0.0169)	0.4363*** (0.0127)	0.6991		4.5693*** (0.0770)	0.5871* (0.0322)	0.4404* (0.0242)	0.7661
	IMT	3.6941*** (0.0381)	0.6323* (0.0149)	0.4742* (0.0112)	0.7799		3.7001*** (0.0178)	0.5937* (0.0301)	0.4453* (0.0225)	0.7940
Group 5 Lowest Middle Mkt Cap	**SVT**	3.7855*** (0.0514)	**0.6398** (0.0153)	**0.4799** (0.0115)	0.7750		3.8171*** (0.0464)	**0.6474** (0.0293)	**0.4856** (0.0220)	0.8284
	CNE	3.1316*** (0.0497)	0.5446*** (0.0179)	0.4084*** (0.0134)	0.6457		3.3250*** (0.0280)	**0.6102** (0.0332)	**0.4576** (0.0249)	0.7696
	JMAT	3.4412*** (0.0538)	0.5747*** (0.0168)	0.4310*** (0.0126)	0.6976		3.5975*** (0.0414)	**0.6177** (0.0305)	**0.4633** (0.0229)	0.8024
	SGE	4.3086*** (0.0948)	0.5227*** (0.0182)	0.3921*** (0.0137)	0.6170		4.8287*** (0.1021)	**0.6318** (0.0309)	**0.4739** (0.0232)	0.8047
	REX	4.0248*** (0.0800)	0.5316*** (0.0174)	0.3987*** (0.0131)	0.6468		4.3771*** (0.0820)	0.6011* (0.0304)	0.4508 (0.0228)	0.7940

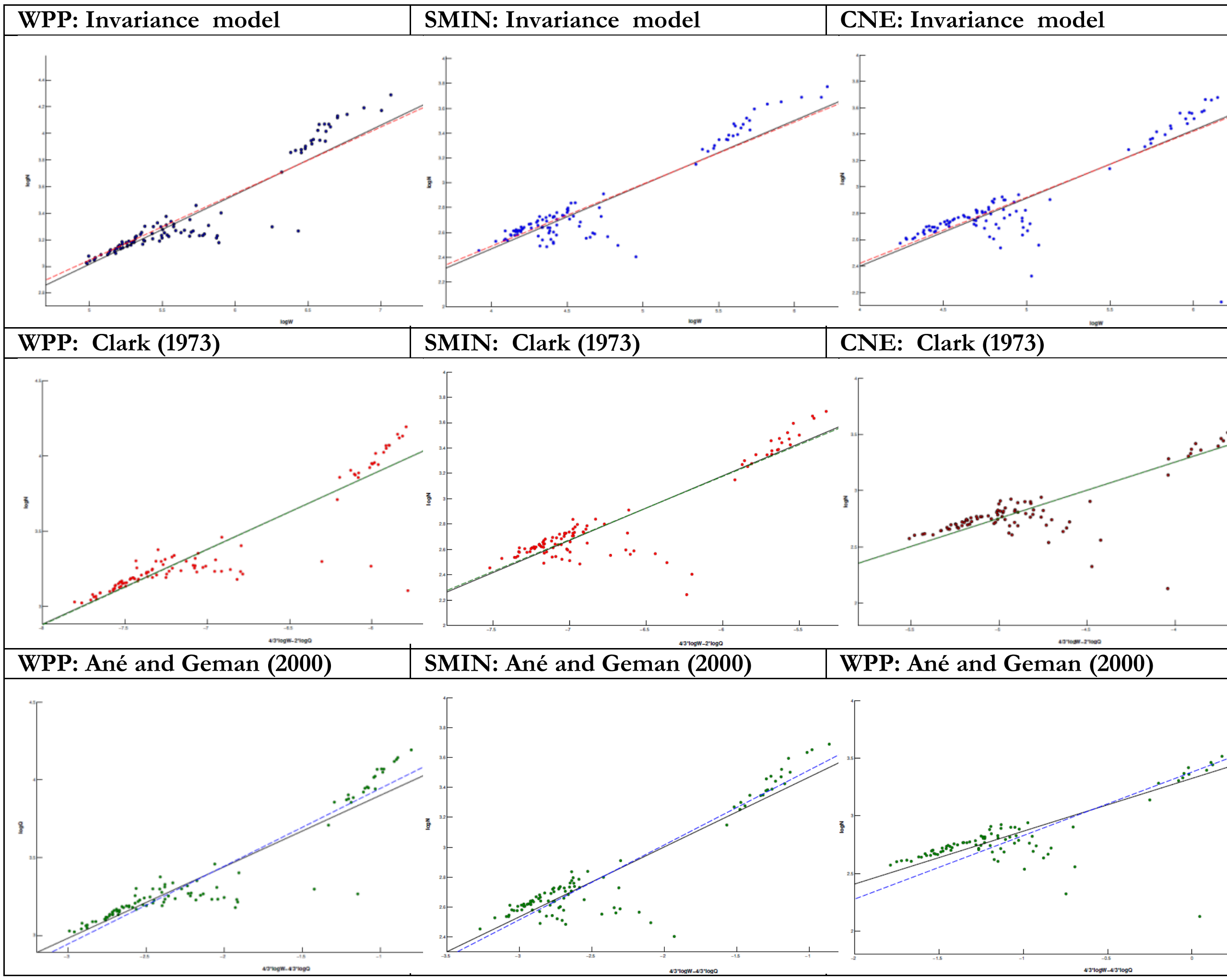
discussion of which exceeds the intended scope of this paper. The stocks (WPP, SMIN, SHP, CNE and SGE) for which all three models empirically predict the $1/2$ proportionality have medium to low market capitalisation (Groups 3 to 5).

To further examine which model predicts more accurately the $1/2$ proportionality between the number of trades and trading activity in the sample, we plot the number of trades (n_t) against trading activity (w_t) for invariance model, (n_t) against $\frac{4}{3}w_t - 2q_t$ for the model of Clark (1973) and (n_t) against $\frac{4}{3}w_t - \frac{4}{3}q_t$ for the model of Ané and Geman (2000). Here we report only three stocks (WPP, SMIN and CNE), for which all models generate $1/2$ proportionality. In that way the comparison between the models is straightforward. The respective variables are averages of 5-minutes interval across all days[22]. Figure 5 shows scatter plots of the number of trades against the respective notions of trading activity. All three models predict the slope of regression line (i.e. coefficient β) to be $1/2$, represented by the dashed line in each respective plot. The actual regression line for each model is depicted as a solid black line. All plots show that the model of Clark (1973) predicts more accurately the $1/2$ proportionality (the predicted dashed line almost coincides with the actual solid black regression line). The invariance model is the second best in predicting the required proportionality, whereas the model of Ané and Geman (2000) is the least accurate. The scatterplots confirm the findings of OLS regressions. Even for the stocks that all models accept the $1/2$ proportionality, the model of Clark is more precise, with smaller standard errors and higher adjusted R-squared.

Taking into account the descriptive statistics in Table 2 (Appendix II), the model of Ané and Geman (2000) does not yield significant coefficients in terms of $1/2$ proportionality when the number of trades per minute is on average very high (e.g. for stocks BLT, BP, HSBA, VOD). OML is the only stock for which none of the models is able to predict the required relationship between trade counts and trading activity. A possible explanation is that OML is the stock with the lowest on average price, with the second highest average volatility and trade size, as well as very high average trading

[22] As stated above, the low frequency 5-minutes intervals produce more accurate estimation than high frequency 1-minute intervals. Thus, the regressions for 5-minutes intervals will give a more distinct and accurate difference between the invariance model and the other two models.

Figure 3-The graphs represent **scatterplots of** the number of trades (n_t) against trading activity (w_t) for invariance model, $\frac{4}{3}w_t - 2q_t$ for the model of Clark (1973) and $\frac{4}{3}w_t - \frac{4}{3}q_t$ for the model of Ané and Geman (2000)



volume. The interaction of these variables breaks any proportionality of trades count and volume with returns variation and trades count with trading activity. Finally, the invariance model is the only model that accepts the null hypothesis for $1/2$ proportionality for stocks with high average volatility, whereas the other alternative models are less precise when high volatility is present. The only exception is stock BLT, which has the highest number of trades on average. The invariance model predicts a higher than $1/2$ proportionality between the number of trades and trading activity for the specific stocks. This means that on average the specific number of trades should be stemming from a lower total risk transfer, or given the number of trades and trading activity, that the total risk should be transferred faster for the specific stock. The fact that this stock exhibits only proportionality between the trading volume and volatility means that changes in trade size are less correlated with innovations in returns volatility and number of trades than invariance implies. This finding is in line with Andersen et al. (2016), who argue that invariance indicate no proportionality between number of trades and trading volume to returns variation, when there exists a correlation between the innovation in trade size and returns volatility and number of trades, respectively.

In general, empirically the invariance model yields coefficients that are higher than the coefficients of the alternative models in our sample. This can be attributed to the expression of both volume and volatility in business time and not simply volume (either expressed as trade counts or number of securities traded) as in the other models. Intuitively, given the specification of the models under investigation, the inclusion of average trade size in the regressions leads to a higher number for the relationship between the number of trades and trading activity. Also, the required proportionality for some stocks is achieved with more than one models, which highlights the impact of stock specific risk and other stock specific characteristics on the proportionality and correlations of the underlying variables. The extent to which these individual characteristics and correlations of the underlying variables affect the relationship between the number of trades and trading activity is very interesting for future research.

6.2 Number of trades and trading activity for each day

Based on the invariance model in (25), we now re-examine the $2/3$ and $1/2$ relationships between the trade counts and trading activity. Instead of averaging the observations across days, the underlying variables are now intraday averages of respective observations for 1-minute and 5-minutes intervals, as defined by equations (23) and (24). Table 7 displays results of OLS regressions for the number of trades regarding the 25 FTSE 100 stocks. The constant term and coefficients here are also the same when testing for 2/3 and 1/2 proportionality, as the regression model for invariance does not change based on our theoretical model.. Thus, we report only the results that refer to the null hypothesis of $\beta_2 = 1/2$ for brevity.

Similar to the previous findings, the null hypothesis for $\beta_1 = 2/3$ is rejected for all stocks in both 1-minute and 5-minutes intervals at 1% significance level. All coefficients in 1-minute and most in 5-

minute intervals do not confirm the $1/2$ proportionality either. The null hypothesis for $\beta_2 = 1/2$ is accepted only for three stocks (XTA, STAN and IMT) in 5-minute interval averages. Consistent with the previous scenario, the adjusted R-squared is higher for 5-minutes, while the constant term is significantly different from zero for all stocks. Upon inspection of Table 7, there is no apparent proportionality between trades counts and trading activity that will hold for all or the majority of the stocks in the sample, even when we calculate variables in low frequency 5-minutes that suffer from less estimation errors. However, certain group of stocks can predict different proportionalities. For instance, the predicted proportionality for some stocks could be $2/5\,(=0.4)$, while for others $4/9$ (≈ 0.444).

The fact that the coefficients do not converge to a value that is constant across all stocks in the sample, when the underlying variables are intraday averages, is intriguing. A possible explanation is that innovations in the underlying variables for some stocks are more apparent when averaging intraday than across days and this may affect the estimated coefficients[23]. Our invariance model proposes a $\frac{4\psi}{3\zeta}$ proportionality between the number of trades and trading activity, where ζ is the intermediation multiplier and ψ denotes the percentage of volatility in prices that comes from order flow imbalances. For certain stocks, the number of intermediaries is high in some days and lower in some others for the same time interval. When averaging observations across days the impact of extreme number of intermediaries fades away. The same holds for the percentage of price volatility that comes from order flow imbalances. This smoothing in the values of ζ and ψ is less evident when the averages are estimated intraday, as any impact of extreme values will drive the daily average to decrease (increase). Consequently, the ratio ψ/ζ is lower and the estimated proportionality becomes smaller. That would explain why in some stocks the estimated relationship is lower than the expected 1/2. Finally, some days in certain stocks include more intervals with zero trades and/or zero realised volatility than others, even for 5-minutes. Thus, estimated daily averages will be upward biased for the specific days, a fact that in turn leads to less accurate estimation of proportionality. Averaging observations across day yields more accurate estimations as the possibility of the same interval to have zero trades or realised volatility across days is smaller.

[23] We have already discussed in the descriptive statistics that the standard deviation of the underlying variable s is higher when averaging observation intraday.

Table 7- OLS Regression results for Model 1 (Invariance). Variables estimated as averages intraday. Stocks are grouped by market capitalization. $\boldsymbol{c}$ is the constant term of the invariance model in (25). $\beta_2 = 4\psi / 3\zeta$ is the coefficient (i.e. proportionality) of the invariance model in (25) . β_2 refers to the null hypothesis of 1/2 proportionality. Significance against 2/3 or 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis of 1/2 proportionality is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute				***5 minutes***		
Groups	***Stocks***	***c***	***β₂***	$\bar{R}^2$		***c***	***β₂***	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	0.8024*** (0.0405)	0.2510*** (0.0107)	0.4243		0.6983*** (0.0717)	0.4035*** (0.0115)	0.6186
	BP	2.1341*** (0.0466)	0.1367*** (0.0085)	0.2568		3.0518*** (0.1178)	0.1720*** (0.0150)	0.1486
	HSBA	1.4713*** (0.0737)	0.2626*** (0.0139)	0.3210		1.1396*** (0.1262)	0.4195*** (0.0160)	0.4782
	GSK	1.5134*** (0.0451)	0.2287*** (0.0095)	0.4368		1.9049*** (0.0805)	0.3084*** (0.0112)	0.5030
	VOD	2.1230*** (0.0405)	0.1521*** (0.0081)	0.3212		2.7311*** (0.0989)	0.2235*** (0.0129)	0.2840
Group 2 Upper Middle Mkt Cap	**BLT**	0.8058*** (0.0653)	0.3884*** (0.0121)	0.5792		0.9229*** (0.1062)	0.4491*** (0.0131)	0.6097
	BG	0.6593*** (0.0461)	0.3841*** (0.0105)	0.6412		0.8326*** (0.0653)	0.4415*** (0.0095)	0.7421
	XTA	0.2266*** (0.0628)	**0.4966** (0.0127)	0.6717		0.4030*** (0.1066)	**0.5078** (0.0136)	0.6483
	NG	0.6496*** (0.0362)	0.3973*** (0.0100)	0.6769		0.9893*** (0.0476)	0.4250*** (0.0081)	0.7862
	STAN	0.4883*** (0.0466)	0.4251*** (0.0109)	0.6701		0.4485*** (0.0660)	**0.4959** (0.0096)	0.7792
Group 3 Middle Mkt Cap	**EMG**	0.9996*** (0.0311)	0.2706*** (0.0086)	0.5666		1.1155*** (0.0551)	0.3789*** (0.0091)	0.6951
	OML	1.3401*** (0.0278)	0.1834*** (0.0095)	0.3307		1.1517*** (0.0604)	0.3866*** (0.0112)	0.6107
	WPP	1.0495*** (0.0493)	0.2774*** (0.0142)	0.3371		0.7903*** (0.0731)	0.4535*** (0.0127)	0.6300
	BLND	1.1633*** (0.0323)	0.2399*** (0.0096)	0.4513		1.1979*** (0.0586)	0.3819*** (0.0101)	0.6550
	RR	1.1035*** (0.0291)	0.2492*** (0.0088)	0.5149		1.2206*** (0.0555)	0.3742*** (0.0098)	0.6617
Group 4 Lower Middle Mkt Cap	**CCL**	0.7971*** (0.0306)	0.3315*** (0.0103)	0.5777		0.8413*** (0.0453)	0.4283*** (0.0087)	0.7638
	SMIN	0.9344*** (0.0233)	0.2658*** (0.0085)	0.5673		0.9833*** (0.0378)	0.3920*** (0.0079)	0.7676
	SHP	1.0163*** (0.0317)	0.2567*** (0.0098)	0.4748		0.9821*** (0.0478)	0.3977*** (0.0090)	0.7232
	IPR	0.9517*** (0.0288)	0.3009*** (0.0095)	0.5702		0.9526*** (0.0476)	0.4266*** (0.0090)	0.7507
	IMT	0.4586*** (0.0389)	0.4652*** (0.0111)	0.7005		0.5400*** (0.0578)	**0.5041** (0.0097)	0.7808
Group 5 Lowest Middle Mkt Cap	**SVT**	0.8086*** (0.0224)	0.3275*** (0.0086)	0.6580		0.7835*** (0.0351)	0.4505*** (0.0076)	0.8251
	CNE	0.9313*** (0.0283)	0.2560*** (0.0102)	0.4554		0.9592*** (0.0442)	0.3907*** (0.0087)	0.7272
	JMAT	0.7252*** (0.0256)	0.3490*** (0.0099)	0.6227		0.6324*** (0.0379)	0.4724*** (0.0080)	0.8216
	SGE	1.2418*** (0.0219)	0.1517*** (0.0088)	0.2823		1.2513*** (0.0470)	0.3414*** (0.0104)	0.5890
	REX	1.0235*** (0.0272)	0.2318*** (0.0108)	0.3773		0.9686*** (0.0494)	0.4002*** (0.0108)	0.6471

We also test the $2/3$ and $1/2$ relationships using the models of Clark (1973) and Ané and Geman (2000) when the underlying variables are daily averages of respective observations for 1-minute and 5-minutes intervals for each day in the sample. Tables 8 and 9 in Appendix II present the results ofOLS regressions for the number of trades regarding the 25 FTSE 100 stocks. The null hypothesis for $2/3$ and $1/2$ proportionalities is rejected for all stocks in both 1-minute and 5-minutes intervals, even if the data fits both models better than invariance model (higher adjusted R-squared). As we explained before, averaging observations intraday reveals certain problems in terms of measurement error and sampling variation which biases the estimated coefficients. It appears that the models of Clark (1973) and Ané and Geman (2000) are affected more than the invariance model. In a sense, the alleged proportionality between the number of trades and trading activity is more accurately predicted by the invariance model when we use intraday averages, at least for some stocks, compared to the other two models.

6.3 Exclusion of market opening

Figure 1 clearly reveals the presence of excess volatility in the first minutes of trading activity for all FTSE 100 stocks of our sample. This is consistent with Areal and Taylor (2002) who find that the FTSE-100 market is more volatile when the market opens. Also, we show that invariance model is more accurate for stocks with higher on average[24] volatility than the alternative models. For these two reasons to examine the impact of volatility on proportionality, we re-run the OLS regressions for all three models in both 1-minute and five minute intervals. The estimated coefficients are presented in Table 10 for 1-minute and Table 11 for 5-minutes intervals. The underlying variables are averages of respective observations across all days in the sample. Upon inspection of both tables, it is obvious that the exclusion of extreme volatility improves the adjusted R-squared for all three models, more significantly for the invariance model. The model of Ané and Geman (2000) is more precise both in 1-minute and 5-minutes intervals, while the invariance model does not confirm the null hypothesis for $\beta_2 = 1/2$, especially for the low frequency 5-minutes intervals. Also, the constant term for invariance model in 5-minutes interval is not significant from zero for the majority of stocks. The model of Clark (1973) continues to be accurate, though there is a shift in terms of stocks for which it can empirically predict the required proportionality.

[24] The estimated averages for volatility are affected by the extreme values at the first five minutes of trading.

It is likely that as more traders enter the market, volatility becomes less extreme, as they adjust the number of trades and/or trading volume upon arrival of new information. A possible explanation for the deviations from invariance principle in the sample that is free of extreme volatility is that the trade size is less correlated with the returns volatility and number of trades and returns volatility becomes proportional to trading volume, either measured in number of trades or number of securities traded. This finding is different to what Andersen et al. (2016) reports for the E-mini S&P 500 future contracts market. In our sample the invariance principle holds only when extreme volatility is included which consequently indicates that market participants will adjust the trade size more actively when the prices are more volatile.

The invariance model is the only to predict empirically $1/2$ proportionality for OML stock, which is a stock for which both volatility and trade size are high. Excluding the first five minutes allows the invariance principle to manifest itself in the specific stock, for which possibly market participants actively change the trading size based on changes in returns volatility. For stocks with low number of trades and/or trading volume on average, there exists a proportionality of returns variance with both trading and number of trades, while for very high capitalisation stocks market participants seem to adjust only there number of trades based on returns volatility whenever new information arrives in the market. That explains why the model of Ané and Geman (2000) is more accurate than other models in high market capitalisation stocks.

Generally, the model of Clark (1973) is more accurate in capturing the market microstructure properties in our sample, either when we include minutes with extreme volatility or not. This result is consistent with the findings of Epps and Epps (1976), Westerfield (1977), Tauchen and Pitts (1983) and other similar papers that support the proportionality between the returns variance and the trading volume. Including minutes with extreme volatility, allows the invariance principle to manifest itself for the majority of stocks, especially for stocks with high on average volatility. In stocks for which both the invariance and Clark's (1973) model are accurate, it is probable that market participants revise trade size upon arrival of information more through the trading volume than the number of trades. When extreme volatility is omitted, the model of Ané and Geman (2000) is more precise and invariance model becomes less accurate in predicting empirically $1/2$ proportionality. This signifies a shift in trading behaviour when there is a great change in volatility. Also, this can be attributed to a change in the correlations between the underlying variables (i.e. between trade size and number of trades, trade size and return variance, return variance and trading volume and return

Table 10- OLS Regression results for 1-minute intervals excluding first 5 minutes of trading. Variables estimated as averages across days. Stocks are grouped by market capitalization. $\boldsymbol{c}$ is the constant term and β_2 is the coefficient of the models in (30), (32) and (34), respectively. Significance against 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis of 1/2 is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute								
		Invariance			***Clark***			***Ane-Geman***		
Groups	***Stocks-t***	***c***	$\boldsymbol{\beta_2}$	$\bar{R}^2$	***c***	$\boldsymbol{\beta_2}$	$\bar{R}^2$	***c***	$\boldsymbol{\beta_2}$	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	-0.3302*** (0.0299)	0.5586*** (0.0080)	0.9059	6.2997*** (0.0494)	0.5413*** (0.0059)	0.9445	3.7325*** (0.0215)	**0.5020** (0.0054)	0.8934
	BP	0.0291 (0.0330)	0.5515*** (0.00637)	0.9370	7.8707*** (0.0414)	0.5342*** (0.0044)	0.9669	4.8100*** (0.0173)	0.4909* (0.0043)	0.8904
	HSBA	0.1171*** (0.0312)	0.5352*** (0.0060)	0.9402	7.5159*** (0.0357)	0.5218*** (0.0040)	0.9712	4.6249*** (0.0135)	0.48081*** (0.0036)	0.8763
	GSK	-0.1747*** (0.03099)	0.6000*** (0.0067)	0.9414	7.0786*** (0.0416)	0.5587*** (0.0052)	0.9586	4.3127*** (0.0161)	0.5208*** (0.0048)	0.8525
	VOD	0.5124*** (0.0411)	**0.5013** (0.0087)	0.8681	8.6806*** (0.0922)	0.4789** (0.0076)	0.8880	5.5196*** (0.0428)	0.4417*** (0.0071)	0.8139
Group 2 Upper Middle Mkt Cap	**BLT**	-0.2729*** (0.0298)	0.5938*** (0.0056)	0.9578	6.4882*** (0.0211)	0.5294*** (0.0031)	0.9832	3.9755*** (0.0072)	**0.5015** (0.0032)	0.8956
	BG	-0.0796* (0.0361)	0.5575*** (0.0082)	0.9008	6.6586*** (0.0558)	0.5217** (0.0068)	0.9222	4.0845*** (0.0227)	0.4860* (0.0063)	0.7873
	XTA	0.1989*** (0.0280)	**0.5047** (0.0057)	0.9400	5.7912*** (0.0299)	0.4756*** (0.0045)	0.9562	3.6388*** (0.0097)	0.4405*** (0.0043)	0.8491
	NG	-0.0378 (0.0294)	0.5881*** (0.0081)	0.9132	7.1654*** (0.0680)	0.5389*** (0.0072)	0.9172	4.4562*** (0.0303)	0.5066*** (0.0065)	0.7751
	STAN	0.1077*** (0.0315)	0.5205** (0.0074)	0.9079	5.9362*** (0.0384)	0.4804*** (0.0051)	0.9468	3.7301*** (0.0160)	0.4505*** (0.0050)	0.8266
Group 3 Middle Mkt Cap	**EMG**	0.3057*** (0.0345)	0.4750* (0.0098)	0.8231	6.7118*** (0.0943)	0.4633*** (0.0092)	0.8344	4.2092*** (0.0442)	0.4252*** (0.0083)	0.7546
	OML	0.9024*** (0.0368)	0.3509*** (0.0133)	0.5795	5.8308*** (0.1630)	0.3047*** (0.0125)	0.5405	4.0468*** (0.0863)	0.2925*** (0.0115)	0.7023
	WPP	0.0973** (0.0311)	0.5649*** (0.0091)	0.8853	7.3083*** (0.0854)	0.5448*** (0.0088)	0.8840	4.5051*** (0.0389)	**0.5011** (0.0078)	0.7826
	BLND	0.3643*** (0.0390)	**0.4927** (0.0119)	0.7719	6.4458*** (0.1052)	**0.4953** (0.0116)	0.7835	4.0264*** (0.0482)	0.4494*** (0.0104)	0.6786
	RR	0.3026*** (0.0285)	**0.5054** (0.0089)	0.8655	6.9720*** (0.0943)	**0.4825** (0.0090)	0.8514	4.4005*** (0.0441)	0.4463*** (0.0079)	0.7714
Group 4 Lower Middle Mkt Cap	**CCL**	0.1721*** (0.0217)	0.5470*** (0.0074)	0.9151	5.7668*** (0.0549)	0.5321*** (0.0072)	0.9152	3.5733*** (0.0247)	0.4862* (0.0064)	0.8655
	SMIN	0.2928*** (0.0281)	**0.5120** (0.0106)	0.8237	6.0978*** (0.1060)	**0.4916** (0.0117)	0.7784	3.8634*** (0.0486)	0.4560*** (0.0099)	0.7477
	SHP	0.1495*** (0.0271)	0.5392*** (0.0086)	0.8870	6.7091 (0.0886)	0.5508 (0.0100)	0.8574	4.0739*** (0.0374)	**0.4965** (0.0082)	0.8281
	IPR	0.2764*** (0.0321)	0.5333** (0.0108)	0.8299	7.2996*** (0.1210)	**0.5008** (0.0111)	0.8009	4.6286*** (0.0578)	0.4675*** (0.0097)	0.7180
	IMT	0.1385*** (0.0322)	0.5610*** (0.0092)	0.8817	6.1635*** (0.0757)	0.5494*** (0.0102)	0.8519	3.7987*** (0.0294)	**0.5030** (0.0086)	0.7459
Group 5 Lowest Middle Mkt Cap	**SVT**	0.1656*** (0.0268)	0.5827*** (0.0104)	0.8615	6.0084*** (0.0820)	**0.5191** (0.0098)	0.8489	3.8607*** (0.0391)	**0.4951** (0.0087)	0.7798
	CNE	0.3030*** (0.0285)	**0.4899** (0.0104)	0.8147	5.1389*** (0.0921)	**0.4775** (0.0125)	0.7424	3.2581*** (0.0384)	0.4407*** (0.0104)	0.7443
	JMAT	0.2191*** (0.0284)	0.5518*** (0.0111)	0.8302	5.4207*** (0.0877)	0.4716** (0.0108)	0.7895	3.5567*** (0.0416)	0.4562*** (0.0097)	0.7670
	SGE	0.5156*** (0.0287)	**0.4772** (0.0124)	0.7466	6.7799*** (0.1534)	**0.4350** (0.0129)	0.6931	4.4554*** (0.0785)	0.4121*** (0.0113)	0.7649
	REX	0.4285*** (0.0241)	**0.4901** (0.0100)	0.8264	6.4463*** (0.1427)	0.4511*** (0.0133)	0.6968	4.2150*** (0.0653)	0.4286*** (0.0106)	0.7656

Table 11- OLS Regression results for 5-minutes intervals excluding first 5 minutes of trading. Variables estimated as averages across days. Stocks are grouped by market capitalization. $\boldsymbol{c}$ is the constant term and β_2 is the coefficient of the models in (30), (32) and (34), respectively. Significance against 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis of 1/2 is not rejected. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		5 minutes								
		Invariance			***Clark***			***Ané and Geman***		
Groups	**Stocks-t**	***c***	β_2	$\bar{R}^2$	***c***	β_2	$\bar{R}^2$	***c***	β_2	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	-0.4838*** (0.0865)	0.5975*** (0.0140)	0.9478	6.1726*** (0.0559)	0.5293** (0.0098)	0.9668	3.7039*** (0.0125)	**0.5018** (0.0096)	0.9649
	BP	-0.1543 (0.1116)	0.5862*** (0.0143)	0.9438	7.7458*** (0.0659)	0.5258* (0.0103)	0.9629	4.7914*** (0.0106)	**0.4961** (0.0100)	0.9607
	HSBA	-0.0762 (0.1111)	0.5769*** (0.0141)	0.9434	7.4140*** (0.0441)	0.5155* (0.0076)	0.9786	4.6194*** (0.0063)	**0.4891** (0.0079)	0.9746
	GSK	-0.2102 (0.1136)	0.6069*** (0.0159)	0.9359	6.7349*** (0.0646)	**0.5151** (0.0126)	0.9435	4.2258*** (0.0088)	**0.4954** (0.0120)	0.9445
	VOD	0.3833** (0.1328)	0.5397* (0.0176)	0.9037	8.7295 (0.1120)	**0.4854** (0.0126)	0.9364	5.6062*** (0.0337)	0.4572** (0.0127)	0.9284
Group 2 Upper Middle Mkt Cap	**BLT**	-0.4245*** (0.1071)	0.6173*** (0.0133)	0.9560	6.3044*** (0.0236)	**0.5036** (0.0066)	0.9830	3.9306*** (0.0108)	**0.4920** (0.0072)	0.9788
	BG	0.0707 (0.1263)	0.5547*** (0.0184)	0.9007	6.4190*** (0.0764)	**0.4789** (0.0142)	0.9188	4.1007 (0.0118)	0.4574** (0.0139)	0.9152
	XTA	0.2110* (0.1027)	0.5337* (0.0132)	0.9425	5.7954 (0.0296)	0.4684** (0.0094)	0.9611	3.7317 (0.0146)	0.4450*** (0.0093)	0.9581
	NG	-0.0179 (0.1153)	0.5987*** (0.0196)	0.9030	6.9189*** (0.1170)	**0.5055** (0.0172)	0.8966	4.4239*** (0.0325)	**0.4869** (0.0160)	0.9021
	STAN	0.1656 (0.1014)	0.5397*** (0.0149)	0.9295	5.9222*** (0.0459)	0.4645*** (0.0101)	0.9548	3.8234 (0.0073)	0.4447*** (0.0102)	0.9498
Group 3 Middle Mkt Cap	**EMG**	0.2191 (0.1127)	**0.5326** (0.0189)	0.8882	6.8546*** (0.1199)	**0.4713** (0.0162)	0.8944	4.3953*** (0.0366)	0.4459*** (0.0153)	0.8946
	OML	0.5955*** (0.1081)	**0.4972** (0.0204)	0.8558	7.4402 (0.1708)	0.4222*** (0.0170)	0.8599	4.9702*** (0.0715)	0.4053*** (0.01631)	0.8604
	WPP	0.0728 (0.1108)	0.5824*** (0.0193)	0.9014	7.0335*** (0.1247)	**0.5211** (0.0178)	0.8951	4.4355 (0.0363)	**0.4909** (0.0164)	0.8998
	BLND	0.2581 (0.1383)	0.5491* (0.0240)	0.8390	6.4169*** (0.1243)	**0.4937** (0.0203)	0.8558	4.1115*** (0.0322)	**0.4645** (0.0193)	0.8522
	RR	0.2292* (0.1039)	0.5537*** (0.0184)	0.9005	7.0399*** (0.1290)	**0.4852** (0.0169)	0.8923	4.5312*** (0.0418)	0.4606* (0.0156)	0.8975
Group 4 Lower Middle Mkt Cap	**CCL**	0.0135 (0.0775)	0.5916*** (0.0149)	0.9405	5.5935*** (0.0637)	**0.5249** (0.0129)	0.9428	3.5203*** (0.0169)	**0.4953** (0.0122)	0.9431
	SMIN	0.1179 (0.0905)	0.5800*** (0.0191)	0.9021	6.2985*** (0.1312)	**0.5175** (0.0196)	0.8749	3.9964*** (0.0429)	**0.4880** (0.0174)	0.8874
	SHP	0.0436 (0.0843)	0.5807*** (0.0160)	0.9298	6.4478*** (0.1099)	**0.5318** (0.0173)	0.9044	4.0224*** (0.0306)	**0.4966** (0.0150)	0.9164
	IPR	0.1806 (0.1121)	0.5767*** (0.0212)	0.8804	7.2809*** (0.1612)	**0.4982** (0.0196)	0.8652	4.6902*** (0.0579)	**0.4755** (0.0181)	0.8732
	IMT	0.1526 (0.1151)	0.5725*** (0.0195)	0.8959	5.8228*** (0.0899)	**0.5048** (0.0195)	0.8700	3.7242*** (0.0138)	**0.4784** (0.0175)	0.8823
Group 5 Lowest Middle Mkt Cap	**SVT**	0.0090 (0.0912)	0.6238*** (0.0198)	0.9083	6.0345*** (0.1051)	**0.5259** (0.0173)	0.9024	3.8697*** (0.0343)	**0.5065** (0.0162)	0.9073
	CNE	0.1314 (0.0996)	0.5594** (0.0198)	0.8889	5.3442*** (0.0955)	**0.5170** (0.0203)	0.8667	3.3577*** (0.0204)	**0.4808** (0.0180)	0.8770
	JMAT	0.0153 (0.0958)	0.6085*** (0.0204)	0.8733	5.6321*** (0.1015)	**0.4999** (0.0180)	0.8846	3.6465*** (0.0308)	**0.4858** (0.0169)	0.8920
	SGE	0.1491 (0.0871)	0.5977*** (0.0196)	0.9030	7.6752*** (0.1770)	**0.5176** (0.0187)	0.8847	4.9267*** (0.0749)	**0.4937** (0.0170)	0.8941
	REX	0.1993* (0.0816)	0.5769*** (0.0180)	0.9116	6.9497*** (0.1672)	**0.5006** (0.0201)	0.8612	4.4830*** (0.0625)	**0.4777** (0.0173)	0.8835

variance and number of trades). The dynamics of these variables are altered depending on the presence of extreme volatility. Finally, it is likely that stock specific characteristics and idiosyncratic risk play an important role in defining the relationship between the variables of trading activity.

7. Conclusion

In this paper we examine invariance principles of trading in stocks of FTSE 100 index. The findings are contrast somewhat with the analysis of Andersen et al. (2016) for the E-mini S&P 500 future contract market. Although, there exists a proportionality between the number of trades and trading activity, it differs from 2/3, while on certain criteria, model performance improves when trading activity is defined according to Clark (1973). Our analysis stems directly from market microstructure invariance and is based on a theoretical model that constitutes a generalised version of the model proposed by Kyle and Obizhaeva (2013). This extension of the initial market microstructure model aims at accommodating empirical phenomena in the stock market, as well as the way the trading is reported in the majority of the available databases. Also, this might provide a new and maybe a better way to measure liquidity based on how order flow imbalances impact prices in the stock market.

It still remains an open question whether invariance applies or not in different time periods in the specific stock market or for the same period in other markets. Also, it is interesting to explore whether the idiosyncratic risk of each security plays an important role and accordingly alter the invariance principles. In the same respect, future research should focus on including other individual characteristics of securities in the estimation of invariance relationships, as well as correlations of the underlying variables. In this paper, we highlight the importance of volatility in the required proportionality; thus, further investigation of how different measures of realised volatility affect the relationship between trade counts and trading activity could be useful. Finally, given the level of fragmentation in the stock market, future papers should focus on how the introduction of different trading platforms affects the invariance principles if any.

8. References

Amihud, Y. (2002) Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets,* 5(1)**,** 31-56.

Andersen, T., Bollerslev, T., Diebold, F. & Labys, P. (2000) Great realizations. *Risk,* 13**,** 105-108.

Andersen, T. G. (1996) Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance,* 51(1)**,** 169-204.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2001) The distribution of realized exchange rate volatility. *Journal of the American statistical association,* 96(453)**,** 42-55.

Andersen , T. G., Bondarenko, O., Kyle , A. S. & Obizhaeva, A. A. (2016) Intraday Trading Invariance in the E-Mini S&P 500 Futures Market. Working Paper.

Ané, T. & Geman, H. (2000) Order flow, transaction clock, and normality of asset returns. *The Journal of Finance,* 55(5)**,** 2259-2284.

Areal, N. M. & Taylor, S. J. (2002) The realized volatility of FTSE ‐ 100 futures prices. *Journal of Futures Markets,* 22(7)**,** 627-648.

Bochner, S. (1955) *Harmonic analysis and the theory of probability*: Berkeley: University of California Press.

Bollerslev, T. & Jubinski, D. (1999) Equity trading volume and volatility: Latent information arrivals and common long-run dependencies. *Journal of Business & Economic Statistics,* 17(1)**,** 9-21.

Bouchaud, J.-P., Farmer, J. D. & Lillo, F. (2008) How markets slowly digest changes in supply and demand. *Fabrizio, How Markets Slowly Digest Changes in Supply and Demand (September 11, 2008)*.

Breen, W. J., Hodrick, L. S. & Korajczyk, R. A. (2002) Predicting equity liquidity. *Management Science,* 48(4)**,** 470-483.

Clark, P. K. (1973) A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society***,** 135-155.

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

Einstein, A. (1905) On the electrodynamics of moving bodies. *Annalen der Physik,* 17(891)**,** 50.

Einstein, A. (1920) Relativity: The Special and General Theory, trans. *Robert W. Lawson (New York: Crown, 1961),* 26.

Engle, R. F. (2000) The Econometrics of Ultra-High-Frequency Data. *Econometrica***,** 1-22.

Epps, T. W. & Epps, M. L. (1976) The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. *Econometrica: Journal of the Econometric Society***,** 305-321.

Gabaix, X., Gopikrishnan, P., Plerou, V. & Stanley, H. E. (2003) A theory of power-law distributions in financial market fluctuations. *Nature,* 423(6937)**,** 267-270.

Gabaix, X., Gopikrishnan, P., Plerou, V. & Stanley, H. E. (2006) Institutional Investors and Stock Market Volatility. *Quarterly Journal of Economics,* 121(2).

Gallant, A. R., Rossi, P. E. & Tauchen, G. (1992) Stock prices and volume. *Review of Financial studies,* 5(2)**,** 199-242.

Glosten, L. R. & Milgrom, P. R. (1985) Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics,* 14(1)**,** 71-100.

Gopikrishnan, P., Meyer, M., Amaral, L. N. & Stanley, H. E. (1998) Inverse cubic law for the distribution of stock price variations. *The European Physical Journal B-Condensed Matter and Complex Systems,* 3(2)**,** 139-140.

Grinold, R. & Kahn, R. (1999) Active portfolio management. McGraw-Hill, New York.

Harris, L. (1986) Cross-security tests of the mixture of distributions hypothesis. *Journal of financial and Quantitative Analysis,* 21(01)**,** 39-46.

Harris, L. (1987) Transaction data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis,* 22(02)**,** 127-141.

Hasbrouck, J. (1991) Measuring the information content of stock trades. *The Journal of Finance,* 46(1)**,** 179-207.

Hasbrouck, J. (1999) Trading Fast and Slow: Security Market Events in Real Time.

Hasbrouck, J. (2009) Trading costs and returns for US equities: Estimating effective costs from daily data. *The Journal of Finance,* 64(3)**,** 1445-1477.

Hunt, B. (1996) *The geometry of some special arithmetic quotients*: Springer.

Jones, C. M., Kaul, G. & Lipson, M. L. (1994) Transactions, volume, and volatility. *Review of Financial Studies,* 7(4)**,** 631-651.

Karpoff, J. M. (1987) The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis,* 22(01)**,** 109-126.

Kyle, A. S. & Obizhaeva, A. A. (2010) Market microstructure invariants. *Available at SSRN 1687965*.

Kyle, A. S. & Obizhaeva, A. A. (2013) Market microstructure invariance: Theory and empirical tests. working paper, University of Maryland.

Kyle, A. S. & Obizhaeva, A. A. (2016) Market Microstructure Invariance: Empirical Hypotheses. *Econometrica,* 84(4)**,** 1345–1404.

Liesenfeld, R. (2001) A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics,* 104(1)**,** 141-178.

Loeb, T. F. (1983) Trading Cost: The Critical Link between Investment Information and Results. *Financial Analysts Journal***,** 39-44.

Mandelbrot, B. & Taylor, H. M. (1967) On the distribution of stock price differences. *Operations research,* 15(6)**,** 1057-1062.

Modigliani, F. & Miller, M. H. (1958) The cost of capital, corporation finance and the theory of investment. *The American economic review***,** 261-297.

Plerou, V., Gopikrishnan, P., Amaral, L. A. N., Gabaix, X. & Stanley, H. E. (2000) Economic fluctuations and anomalous diffusion. *Physical Review E,* 62(3)**,** R3023.

Richardson, M. & Smith, T. (1994) A direct test of the mixture of distributions hypothesis: Measuring the daily flow of information. *Journal of Financial and Quantitative Analysis,* 29(01)**,** 101-116.

Tauchen, G. E. & Pitts, M. (1983) The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society***,** 485-505.

Torre, N. (1997) BARRA market Impact model handbook. *BARRA Inc., Berkeley*.

Torre, N. & Ferrari, M. J. (1998) The market impact model. *Horizons, The Barra Newsletter,* 165.

Westerfield, R. (1977) The distribution of common stock price changes: An application of transactions time and subordinated stochastic models. *Journal of Financial and Quantitative Analysis,* 12(05)**,** 743-765.

Zhang, Y.-C. (1999) Toward a theory of marginally efficient markets. *Physica A: Statistical Mechanics and its Applications,* 269(1)**,** 30-44.

Zhou, B. (1996) High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics,* 14(1)**,** 45-52.

Appendix I

Proofs

1. Relationship between expected bet and trading activity

From the specification of bet volume we know that:

$$\bar{\mathrm{V}} := \frac{2}{\zeta} \cdot \mathrm{V} \Rightarrow \mathrm{P} \cdot \bar{\sigma} \cdot \bar{\mathrm{V}} = \frac{2}{\zeta} \cdot \mathrm{V} \cdot \mathrm{P} \cdot \bar{\sigma} \qquad \textbf{(i)}$$

Also, we know that the trading volatility in currency units is given by:

$$P \cdot \bar{\sigma} := \psi \cdot P \cdot \sigma \qquad \textbf{(ii)}$$

Combining equations (i) and (ii), we get the following expression:

$$\mathrm{P} \cdot \bar{\sigma} \cdot \bar{\mathrm{V}} = \frac{2}{\zeta} \cdot \mathrm{V} \cdot \mathrm{P} \cdot \psi \cdot \sigma = \psi \cdot \mathrm{P} \cdot \frac{2}{\zeta} \cdot \mathrm{V} \cdot \sigma \qquad \textbf{(iii)}$$

Finally, taking into account the definitions of expected bet and trading activity from (iii) it can be easily inferred that:

$$\overline{\mathrm{W}} := \frac{2\psi}{\zeta} \cdot \mathrm{W}$$

2. Expressing expected bet activity in terms of total volume $\boldsymbol{V}$, price $\boldsymbol{P}$, volatility $\boldsymbol{\sigma}$ and expected trading activity W

From the specification of expected bet activity we know that:

$$\overline{\mathrm{W}} = P \cdot \bar{V} \cdot \bar{\sigma} \qquad \textbf{(i)}$$

If now we substitute for total volume $\boldsymbol{V}$, price $\boldsymbol{P}$ and volatility $\boldsymbol{\sigma}$, then

$$\overline{\mathrm{W}} = P \cdot N_B \cdot E\left|\tilde{\mathrm{Q}}\right| \cdot \psi \cdot \sigma \qquad \textbf{(ii)}$$

Finally, using the relationship between expected bet and trading activity,

$$\frac{2}{\zeta} \cdot \mathrm{W} = P \cdot N_B \cdot \mathrm{Q}_B \cdot \sigma$$

3. Invariance relationship in terms of trading activity W and number of bets N_B

We have already proved that,

$$\frac{2}{\zeta} \cdot \mathrm{W} = P \cdot N_B \cdot \mathrm{Q}_B \cdot \sigma \qquad \textbf{(i)}$$

We also know that,

$$\mathrm{Q}_B = \mathrm{I} \cdot \mathrm{P}^{-1} \cdot \psi^{-1} \cdot \sigma^{-1} \cdot \mathrm{N}_{\mathrm{B}}^{\frac{\xi}{1-\xi}} \qquad \textbf{(ii)}$$

Combining the two relationships,

$$\frac{2}{\zeta} \cdot \mathrm{W} = \cancel{P} \cdot N_B \cdot \mathrm{I} \cdot \cancel{\mathrm{P}^{-1}} \cdot \psi^{-1} \cdot \cancel{\sigma^{-1}} \cdot \mathrm{N}_{\mathrm{B}}^{\frac{\xi}{1-\xi}} \cdot \cancel{\sigma} \qquad \textbf{(iii)}$$

Finally, we solve for I,

$$\mathrm{I} := \frac{2\psi}{\zeta} \cdot \frac{W}{N_B^{\frac{1}{1-\xi}}}$$

4. Invariance relationships in logs (proportionality between trading activity and number of bets and between trading activity and bet size)

Invariance of bets states that,

$$\tilde{I} := P \cdot \tilde{\mathrm{Q}} \cdot \bar{\sigma} \cdot N_B^{-\frac{\xi}{1-\xi}} \qquad \textbf{(i)}$$

If logarithms and expectations are applied then:

$$E\{\log \tilde{I}\} = p + q_B + \log \psi + \frac{s}{2} - \frac{\xi}{1-\xi} n_B \qquad \textbf{(ii)}$$

where $p := \log P$, $\tilde{q}_B := \log \tilde{\mathrm{Q}}$, $s := \log \sigma^2$, $n_B := \log N_B$, $q_B := E\{\log \tilde{q}_B\}$

Solving for q_B:

$$q_B = E\{\log \tilde{I}\} - p - \log\psi - \frac{s}{2} + \frac{\xi}{1-\xi} n_B \qquad \textbf{(iii)}$$

Also, from the relationship between bet and trading activity we know that,

$$\overline{\mathrm{W}} := \frac{2\psi}{\zeta} \cdot \mathrm{W} \Leftrightarrow P \cdot \bar{V} \cdot \bar{\sigma} = \frac{2\psi}{\zeta} \cdot \mathrm{W} \Leftrightarrow P \cdot N_B \cdot \mathrm{Q}_B \cdot \sigma = \frac{2}{\zeta} \cdot \mathrm{W} \qquad \textbf{(iv)}$$

Taking the logarithms in the above relationship, yields

$$p + \bar{\upsilon} + \frac{s}{2} := p + n_{\mathrm{B}} + q_{\mathrm{B}} + \frac{s}{2} := \log\left(\frac{2}{\zeta}\right) + w \qquad \textbf{(v)}$$

where $p := \log P$, $\bar{\upsilon} = \log \bar{\mathrm{V}}$, $s := \log \sigma^2$, $w := \log \mathrm{W}$, $n_B := \log N_B$, $q_B := E\{\log \tilde{q}\}$

If we substitute q_{B} in (v) from (iii) we obtain the invariance expression for n_{B}:

$$\cancel{p} + n_{\mathrm{B}} + E\{\log \tilde{I}\} \cancel{- p} - \log\psi \cancel{- \frac{s}{2}} + \frac{\xi}{1-\xi} n_B + \cancel{\frac{s}{2}} := \log\left(\frac{2}{\zeta}\right) + w$$

$$\frac{1}{1-\xi} n_{\mathrm{B}} + E\{\log \tilde{I}\} := \log\psi + + w \quad \Rightarrow n_{\mathrm{B}} := \underbrace{-(1-\xi)E\{\log \tilde{I}\}}_{c_1} + \underbrace{(1-\xi)\left[\log\psi + \log\left(\frac{2}{\zeta}\right)\right]}_{c_2} + (1-\xi)w$$

$$n_{\mathrm{B}} := \underbrace{c_1 + c_2}_{c} + (1-\xi)w \overset{\xi = 1 - \frac{4\psi}{3\zeta}}{=} c + \frac{4\psi}{3\zeta} w$$

where $c_1 := -(1-\xi)E\{\log \tilde{I}\}$ and $c_2 := (1-\xi)\left[\log\psi + \log\left(\frac{2}{\zeta}\right)\right]$

Similarly, in order to obtain the invariance expression for q_B, we start from the expression in (ii). Given that $\bar{\upsilon} = n_B + q_B$,

$$E\{\log \tilde{I}\} = p + q_B + \log\psi + \frac{s}{2} - \frac{\xi}{1-\xi} n_B := p + q_B + \log\psi + \frac{s}{2} - \frac{\xi}{1-\xi}\left(\bar{\upsilon} - q_B\right) \qquad \textbf{(vi)}$$

Also, from (v) it is obvious that,

$$\bar{\upsilon} := \log\left(\frac{2}{\zeta}\right) + w - p - \frac{s}{2} \qquad \textbf{(vii)}$$

Substituting (vii) into (vi),

$$E\{\log \tilde{I}\} := p + q_B + \log\psi + \frac{s}{2} - \frac{\xi}{1-\xi}\left(\log\left(\frac{2}{\zeta}\right) + w - p - \frac{s}{2} - q_B\right) \Rightarrow$$

$$E\{\log \tilde{I}\} := p + q_B + \log\psi + \frac{s}{2} - \frac{\xi}{1-\xi}\log\left(\frac{2}{\zeta}\right) - \frac{\xi}{1-\xi} w + \frac{\xi}{1-\xi} p + \frac{\xi}{1-\xi}\frac{s}{2} + \frac{\xi}{1-\xi} q_B \Rightarrow$$

$$E\{\log \tilde{I}\} := \frac{1}{1-\xi} p + \frac{1}{1-\xi} q_B + \log\psi - \frac{\xi}{1-\xi}\log\left(\frac{2}{\zeta}\right) + \frac{1}{1-\xi}\frac{s}{2} - \frac{\xi}{1-\xi} w \Rightarrow$$

$$\frac{1}{1-\xi} q_B = E\{\log \tilde{I}\} - \frac{1}{1-\xi} p - \log\psi + \frac{\xi}{1-\xi}\log\left(\frac{2}{\zeta}\right) - \frac{1}{1-\xi}\frac{s}{2} + \frac{\xi}{1-\xi} w$$

$$q_B := \underbrace{(1-\xi)E\{\log \tilde{I}\}}_{c_1} \underbrace{-(1-\xi)\log\psi + \xi\log\left(\frac{2}{\zeta}\right)}_{c_2} - \left(p + \frac{s}{2}\right) + \xi w \Rightarrow$$

$$q_B := \underbrace{c_1 + c_2}_{c} + \xi w - \left(p + \frac{s}{2}\right) \overset{\xi = 1 - \frac{4\psi}{3\zeta}}{=} c + \left(1 - \frac{4\psi}{3\zeta}\right) w - \left(p + \frac{1}{2}s\right)$$

where $c_1 := (1-\xi)E\{\log \tilde{I}\}$ and $c_2 := -(1-\xi)\log\psi + \xi\log\left(\frac{2}{\zeta}\right)$

5. Invariance representation of (Clark (1973)) theory

From the relationship between bet and trading activity we know that:

$$\overline{\mathrm{W}} := \frac{2\psi}{\zeta}\cdot \mathrm{W} \Leftrightarrow P\cdot \bar{V}\cdot \bar{\sigma} = \frac{2\psi}{\zeta}\cdot \mathrm{W} \Leftrightarrow P\cdot N_B \cdot \mathrm{Q}_B \cdot \sigma = \frac{2}{\zeta}\cdot \mathrm{W}$$

This can be expressed in log terms as

$$p + \bar{\upsilon} + \frac{s}{2} := \log\left(\frac{2}{\zeta}\right) + w \qquad \textbf{(i)}$$

Clark (1973) implies that $s = c + \bar{\upsilon} = c + n + q$, so (i) can be written as:

$$p + s - c + \frac{s}{2} := \log\left(\frac{2}{\zeta}\right) + w \Rightarrow w := -\underbrace{\left[\log\left(\frac{2}{\zeta}\right) + c\right]}_{c} + p + \frac{3}{2}s \Rightarrow w := -c + p + \frac{3}{2}s \qquad \textbf{(ii)}$$

If in (ii), we substitute s for $c + n + q$, then

$$w := -c + p + \frac{3}{2}(c + n + q) \Rightarrow w := \underbrace{p - \frac{1}{2}c}_{c} + \frac{3}{2}n + \frac{3}{2}q \Rightarrow$$

$$w := c + \frac{3}{2}n + \frac{3}{2}q \Rightarrow \frac{3}{2}n = -c + w - \frac{3}{2}q \Rightarrow$$

$$n_j = c + \frac{2}{3}\left[w_j - \frac{3}{2}q_j\right] + u_j^n \qquad \textbf{(iii)}$$

In order for (iii) to be comparable to the invariance model it needs to be transformed as follows:

$$n_j = c + \frac{2}{3}\frac{2\psi}{\zeta}\left[\frac{\zeta}{2\psi}w_j - \frac{3\zeta}{4\psi}q_j\right] + u_j^n = c + \frac{4\psi}{3\zeta}\left[\frac{\zeta}{2\psi}w_j - \frac{3\zeta}{4\psi}q_j\right] + u_j^n$$

6. Invariance representation of (Ané and Geman (2000)) theory

From the relationship between bet and trading activity we know that:

$$\overline{\mathrm{W}} := \frac{2\psi}{\zeta}\cdot \mathrm{W} \Leftrightarrow P\cdot \bar{V}\cdot \bar{\sigma} = \frac{2\psi}{\zeta}\cdot \mathrm{W} \Leftrightarrow P\cdot N_B \cdot \mathrm{Q}_B \cdot \sigma = \frac{2}{\zeta}\cdot \mathrm{W}$$

This can be expressed in log terms as

$$p+\bar{\upsilon}+\frac{s}{2}:=p+n+q+\frac{s}{2}:=\log\left(\frac{2}{\zeta}\right)+w \qquad \textbf{(i)}$$

Ané and Geman (2000) implies that $s=c+n$, so (i) can be written as:

$$p+n+q+\frac{1}{2}(c+n):=p+n+q+\frac{1}{2}c+\frac{1}{2}n:=\log\left(\frac{2}{\zeta}\right)+w\Rightarrow$$

$$w:=-\log\left(\frac{2}{\zeta}\right)+p+\frac{3}{2}n+q+\frac{1}{2}c\Rightarrow w:=\underbrace{p+\frac{1}{2}c-\log\left(\frac{2}{\zeta}\right)}_{c}+q+\frac{3}{2}n\Rightarrow$$

$$w:=c+q+\frac{3}{2}n\Rightarrow$$

$$n_j:=c+\frac{2}{3}\left[w_j-q_j\right]+u_j^n \qquad \textbf{(ii)}$$

In order for (ii) to be comparable to the invariance model it needs to be transformed as follows:

$$n_j:=c+\frac{2}{3}\frac{2\psi}{\zeta}\left[\frac{\zeta}{2\psi}w_j-\frac{\zeta}{2\psi}q_j\right]+u_j^n=c+\frac{4\psi}{3\zeta}\left[\frac{\zeta}{2\psi}w_j-\frac{\zeta}{2\psi}q_j\right]+u_j^n$$

Appendix II

Tables and Graphs

Table 1- Stocks and their abbreviations

Stock	Abbreviation	Stock	Abbreviation
ROYAL DUTCH SHELL PLC	**RDSA**	BRITISH LAND CO PLC	**BLND**
BP PLC	**BP**	ROLLS-ROYCE HOLDINGS PLC	**RR**
HSBC HOLDINGS PLC	**HSBA**	CARNIVAL PLC	**CCL**
GLAXOSMITHKLINE PLC	**GSK**	SMITHS GROUP PLC	**SMIN**
VODAFONE	**VOD**	SHIRE PLC	**SHP**
BHP BILLITON PLC	**BLT**	INTERNATIONAL POWER PLC	**IPR**
BG GROUP PLC	**BG**	IMPERIAL TOBACCO GROUP PLC	**IMT**
XSTRATA	**XTA**	SEVERN TRENT PLC	**SVT**
NATIONAL GRID PLC	**NG**	CAIRN ENERGY PLC	**CNE**
STANDARD CHARTERED PLC	**STAN**	JOHNSON MATTHEY PLC	**JMAT**
MAN GROUP PLC	**EMG**	SAGE GROUP PLC	**SGE**
OLD MUTUAL PLC	**OML**	REXAM PLC	**REX**
WPP PLC	**WPP**		

Source: Thompson Reuters Tick History

Table 2- Descriptive Statistics for each stock (1-minute averages across whole sample). The values of variables are estimated using equations (21) and (22). The numbers off brackets are means of variables calculated by aggregating their respective individual values for 3 years (for the whole sample) and then dividing them by 510 (the number of minutes during which the stocks are traded every day). The numbers in brackets are the standard deviations of the respective variables. The annualised volatility is calculated as follows: $\text{Annualised Volatility}=\sqrt{252\cdot 8.5\cdot 60}$

	RDSA	**BP**	**HSBA**	**GSK**	**VOD**	**BLT**	**BG**	**XTA**	**NG**	**STAN**	**EMG**	**OML**	**WPP**
Volatility (σ) Annualised volatility calculated from 10s returns	0.2282 (0.1081)	0.2733 (0.1042)	0.2674 (0.1181)	0.2500 (0.0954)	0.2717 (0.1946)	0.3801 (0.1942)	0.3303 (0.1112)	0.4192 (0.1741)	0.2576 (0.0832)	0.3585 (0.1369)	0.3611 (0.1440)	0.3984 (0.1841)	0.3053 (0.2037)
Volume (V) Average no. of stocks per minute (1,000 shares)	11.752 (6.0843)	115.28 (53.277)	95.852 (42.783)	32.903 (15.759)	392.49 (209.05)	39.260 (33.824)	23.266 (8.3435)	22.685 (7.9899)	19.833 (8.8957)	16.264 (6.4193)	29.341 (10.756)	57.0243 (18.811)	18.468 (6.8106)
Trades (N) Average no. of trades per minute	8.3277 (2.7266)	20.582 (8.2781)	23.109 (8.3199)	16.085 (5.9233)	21.533 (8.1503)	25.961 (10.459)	13.166 (3.9613)	21.484 (6.0226)	9.4114 (2.8117)	13.873 (3.6279)	9.7146 (2.4047)	8.7062 (1.8875)	9.7874 (2.9144)
Trade Size (Q) Average no. of stocks per trade per minute (1000 shares)	1.6907 (1.0428)	6.6921 (3.0662)	5.0798 (2.4897)	2.3540 (1.1503)	24.870 (28.459)	2.2633 (9.8618)	2.2252 (0.8406)	1.1959 (0.4655)	2.5606 (1.6331)	1.4646 (0.9458)	3.7615 (1.7925)	8.2803 (6.5072)	2.3003 (0.7826)
Price (P) Average price per trade per minute (£)	17.920 (0.0469)	5.4004 (0.0205)	7.6311 (0.2852)	12.225 (0.0289)	1.4500 (0.0060)	14.552 (0.0892)	9.8611 (0.0903)	22.003 (0.1312)	6.9470 (0.0472)	14.149 (0.0744)	4.3726 (0.0374)	1.1667 (0.0197)	5.5870 (0.0506)

	BLND	**RR**	**CCL**	**SMIN**	**SHP**	**IPR**	**IMT**	**SVT**	**CNE**	**JMAT**	**SGE**	**REX**
Volatility (σ) Annualised volatility calculated from 10s returns	0.3101 (0.1222)	0.2843 (0.1163)	0.2691 (0.1181)	0.2568 (0.0968)	0.2646 (0.1013)	0.2685 (0.1032)	0.2438 (0.0827)	0.2455 (0.0795)	0.2922 (0.1483)	0.2717 (0.1176)	0.2717 (0.0946)	0.2720 (0.1067)
Volume (V) Average no. of stocks per minute (1,000 shares)	12.187 (3.6733)	23.717 (7.7029)	3.9799 (1.8535)	6.9768 (2.4251)	9.5054 (3.3289)	22.297 (6.8696)	7.6664 (3.2553)	3.9121 (1.1501)	3.1667 (1.0310)	3.3789 (0.9266)	19.921 (7.1286)	11.702 (3.8533)
Trades (N) Average no. of trades per minute	9.7616 (2.4025)	9.0196 (2.4182)	8.2784 (3.6188)	6.6033 (1.5928)	8.0740 (2.3895)	8.4717 (2.1044)	11.063 (2.9579)	6.6650 (1.7494)	7.0265 (1.5143)	6.7220 (1.7167)	6.3263 (1.5215)	6.3174 (1.4498)
Trade Size (Q) Average no. of stocks per trade per minute (1000 shares)	1.4438 (0.4971)	3.1491 (1.2815)	0.5487 (0.2036)	1.3541 (1.0968)	1.4175 (0.7819)	3.2037 (1.3903)	0.9447 (0.9652)	0.6832 (0.3458)	0.5616 (0.3452)	0.5733 (0.1818)	3.8689 (2.7344)	2.2179 (1.0495)
Price (P) Average price per trade per minute (£)	8.4667 (0.0923)	4.2516 (0.0449)	19.723 (0.2121)	9.5537 (0.1047)	10.072 (0.1191)	3.5457 (0.0425)	19.934 (0.1399)	12.893 (0.1463)	22.626 (0.1524)	15.044 (0.1153)	2.1269 (0.0097)	3.9979 (0.0455)

Table 3-Descriptive Statistics for each stock (daily averages across whole sample). The values of variables are estimated using equations (23) and (24). The numbers off brackets are means of the variables are calculated by aggregating their respective individual values for 3 years (for the whole sample) and then dividing them by 754 (the number of trading days in the sample). The numbers in brackets are the standard deviations of the respective variables. The annualised volatility is calculated as follows: $\text{Annualised Volatility}=\sqrt{252\cdot 8.5\cdot 60}$

	RDSA	**BP**	**HSBA**	**GSK**	**VOD**	**BLT**	**BG**	**XTA**	**NG**	**STAN**	**EMG**	**OML**	**WPP**
Volatility (σ) Annualised volatility calculated from 10s returns	0.2221 (0.1123)	0.2795 (0.1078)	0.2564 (0.1280)	0.2471 (0.0825)	0.2858 (0.1902)	0.3755 (0.1951)	0.3204 (0.1320)	0.4168 (0.2380)	0.2472 (0.0950)	0.3472 (0.1813)	0.3513 (0.1903)	0.3931 (0.2743)	0.2914 (0.2535)
Volume (V) Average no. of stocks per minute (1,000 shares)	12.039 (6.4391)	115.58 (6.1140)	96.264 (51.863)	32.879 (21.140)	392.98 (284.57)	39.426 (42.371)	23.363 (11.197)	22.686 (14.442)	19.714 (12.619)	16.243 (10.247)	29.278 (19.234)	57.058 (26.054)	18.618 (7.6505)
Trades (N) Average no. of trades per minute	8.3725 (2.7169)	20.611 (8.3690)	23.170 (11.805)	16.147 (6.1196)	21.580 (8.9659)	26.018 (11.413)	13.185 (4.6213)	21.451 (9.2414)	9.4429 (3.2430)	13.785 (5.5930)	9.6883 (3.4068)	8.6645 (2.7770)	9.7682 (3.3392)
Trade Size (Q) Average no. of stocks per trade per minute (1000 shares)	1.7382 (1.5388)	6.6846 (5.1291)	5.0926 (4.0425)	2.3385 (1.8097)	24.858 (41.221)	2.2886 (12.261)	2.2383 (1.6735)	1.2046 (0.7819)	2.5296 (2.3256)	1.4845 (1.6176)	3.7978 (3.3972)	8.3158 (8.9451)	2.3374 (1.5070)
Price (P) Average price per trade per minute (£)	17.943 (1.7767)	5.4055 (0.4840	7.6420 (1.5825)	12.2217 (1.1746)	1.4493 (0.2020)	14.549 (2.8697)	9.8569 (1.7500)	22.017 (12.704)	6.9026 (0.8331)	14.159 (3.1602)	4.3650 (1.4925)	1.1595 (0.4464)	5.6493 (1.3043)

	BLND	**RR**	**CCL**	**SMIN**	**SHP**	**IPR**	**IMT**	**SVT**	**CNE**	**JMAT**	**SGE**	**REX**
Volatility (σ) Annualised volatility calculated from 10s returns	0.3034 (0.1251)	0.2785 (0.1382)	0.2656 (0.1227)	0.2513 (0.1434)	0.2619 (0.0858)	0.2653 (0.1217)	0.2372 (0.0994)	0.2371 (0.0952)	0.2815 (0.1849)	0.2652 (0.1353)	0.2682 (0.1087)	0.2647 (0.1228)
Volume (V) Average no. of stocks per minute (1,000 shares)	12.224 (4.9445)	23.767 (11.368)	4.0836 (2.1167)	6.0903 (5.1020)	9.3592 (6.0597)	22.389 (10.393)	7.6872 (6.0251)	3.9323 (1.9583)	3.2043 (2.4005)	3.3740 (1.8063)	19.920 (11.562)	11.718 (6.8156)
Trades (N) Average no. of trades per minute	9.7631 (3.0247)	9.0653 (2.7551)	8.5643 (2.7737)	6.6532 (1.8085)	8.1079 (2.5395)	8.4700 (2.6518)	10.916 (4.6128)	6.6712 (2.1117)	7.0751 (1.8987)	6.7524 (1.9706)	6.3659 (1.7958)	6.3342 (1.9247)
Trade Size (Q) Average no. of stocks per trade per minute (1000 shares)	1.4503 (0.7662)	3.1446 (2.3736)	0.5475 (0.3578)	1.3283 (1.8528)	1.3975 (1.4847)	3.2291 (2.3338)	0.9782 (1.8122)	0.6923 (0.6193)	0.5699 (0.7016)	0.5682 (0.3648)	3.8306 (3.9040)	2.2174 (1.8229)
Price (P) Average price per trade per minute (£)	8.4179 (3.9719)	4.2618 (0.8376)	19.807 (3.8067)	9.5122 (1.3484)	10.104 (1.4441)	3.5111 (0.8588)	19.934 (3.1871)	12.775 (1.8832)	22.599 (5.5533)	14.967 (3.1685)	2.1375 (0.3061)	4.0070 (1.0112)

Table 5-% percentage exclusions of 1-minute and 5-minute intervals based on the number of intervals that have zero number of trades and/or zero realised volatility

Stocks	% Exclusions (1 minute)	% Exclusions (5 minutes)
RDSA	28.33%	2.44%
BP	28.70%	7.15%
HSBA	21.30%	4.11%
GSK	30.34%	5.87%
VOD	24.21%	6.41%
BLT	11.30%	0.96%
BG	29.47%	4.80%
XTA	8.18%	0.46%
NG	42.81%	10.77%
STAN	23.86%	3.32%
EMG	32.28%	4.12%
OML	38.56%	7.11%
WPP	37.04%	7.25%
BLND	30.00%	3.64%
RR	32.95%	4.92%
CCL	34.30%	3.37%
SMIN	47.41%	9.58%
SHP	45.61%	10.35%
IPR	37.12%	5.03%
IMT	30.99%	3.82%
SVT	46.96%	9.21%
CNE	37.83%	5.02%
JMAT	41.83%	6.10%
SGE	48.60%	11.35%
REX	47.07%	10.21%

Figure 2- The figure shows averages across days for the volume V_t and trade size Q_t. The substantive graphs are divided in two groups. Group 1 refer to stocks with the highest (extreme) average values for volume and trade size and Group 2 to the remaining stocks.

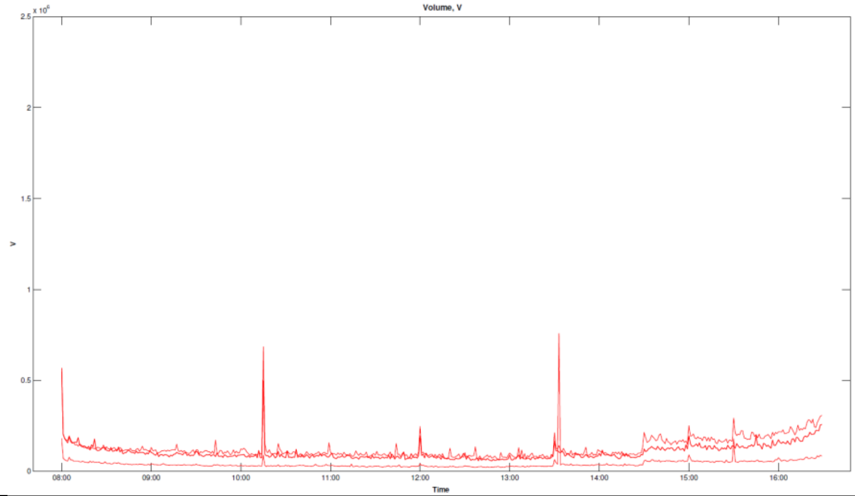
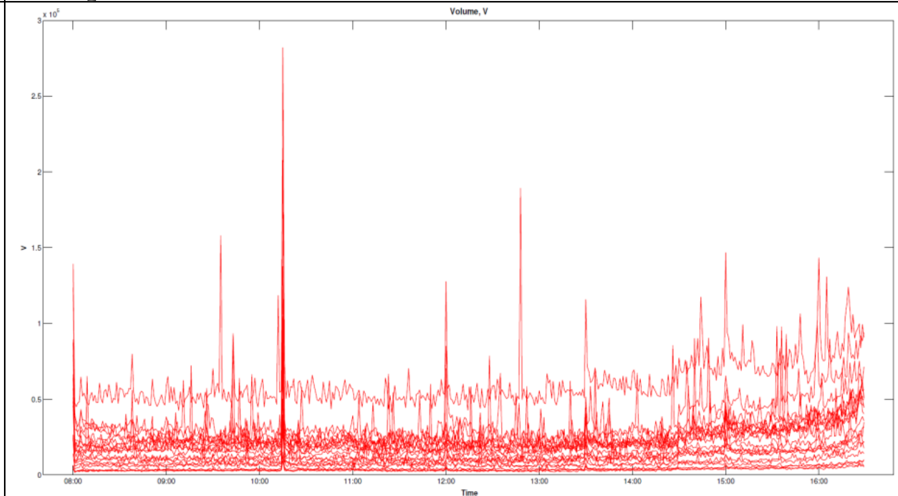
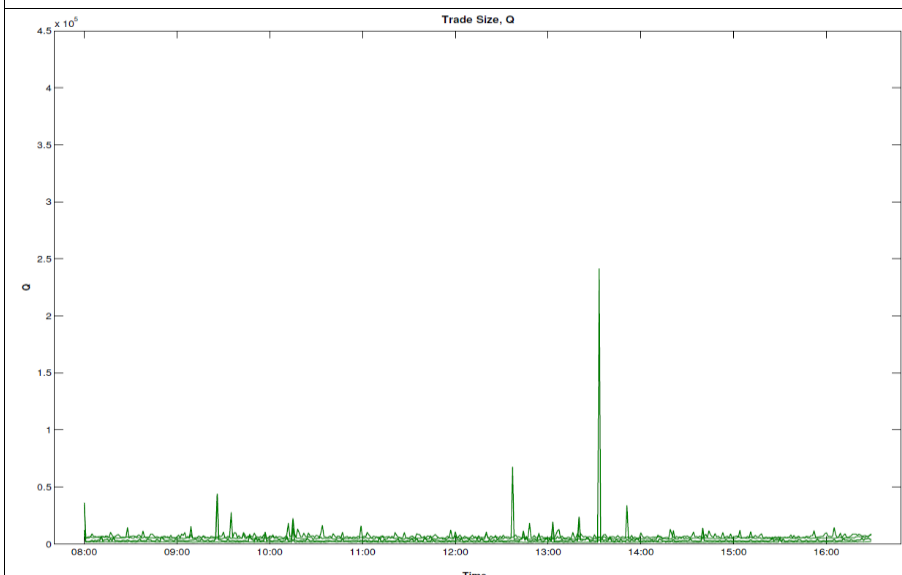
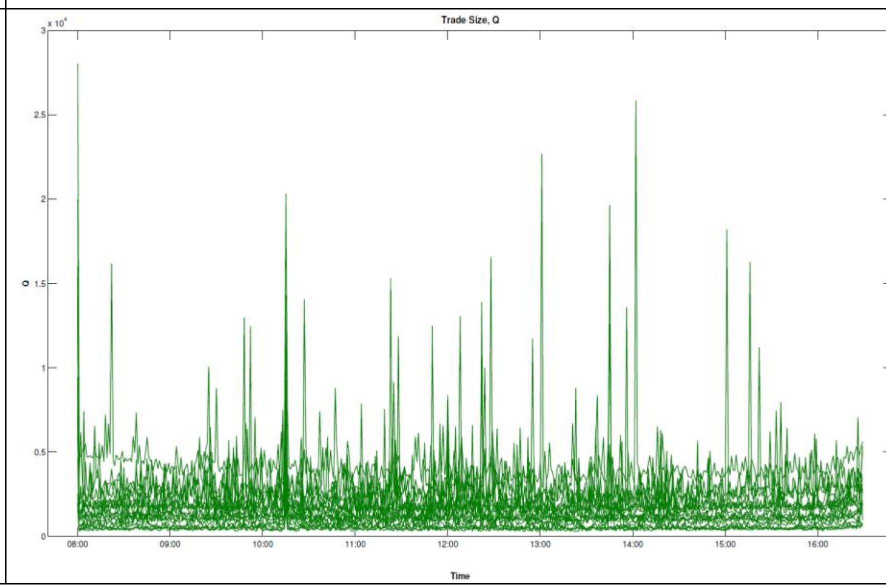
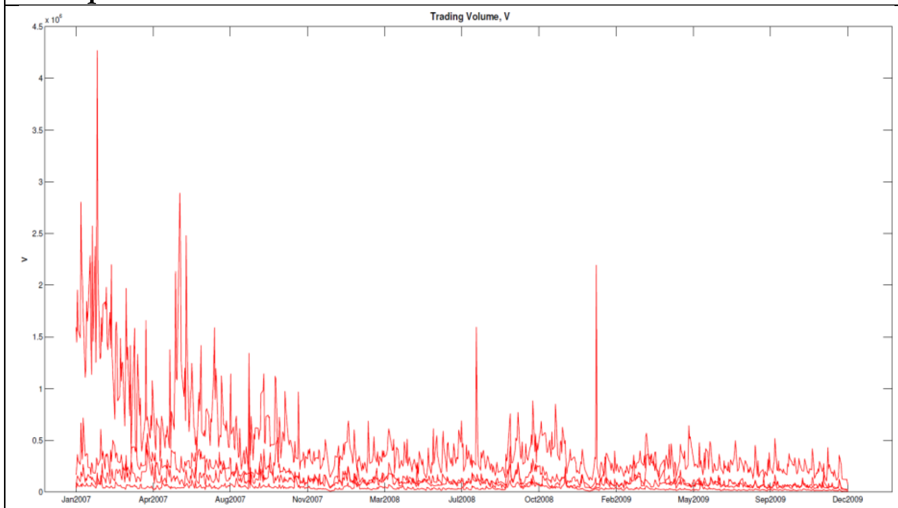
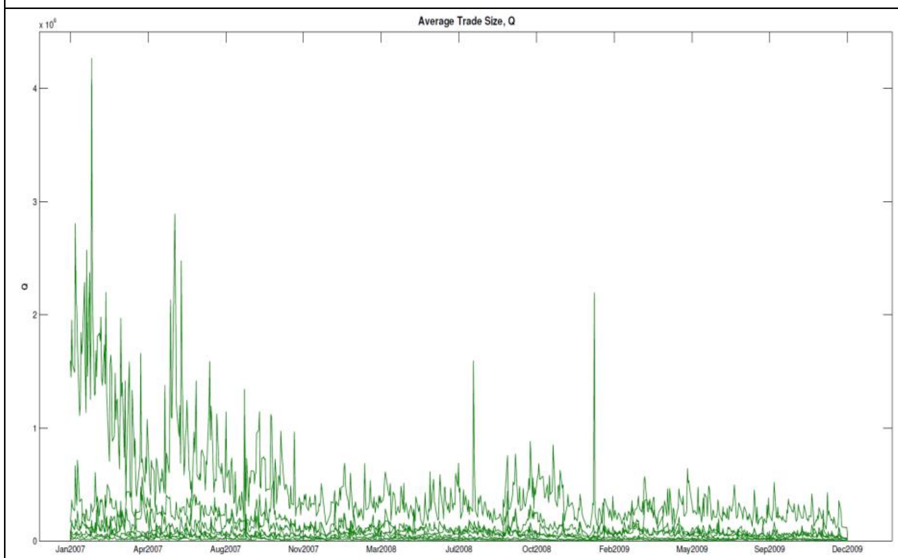
Group 1	**Group 2**
	
	

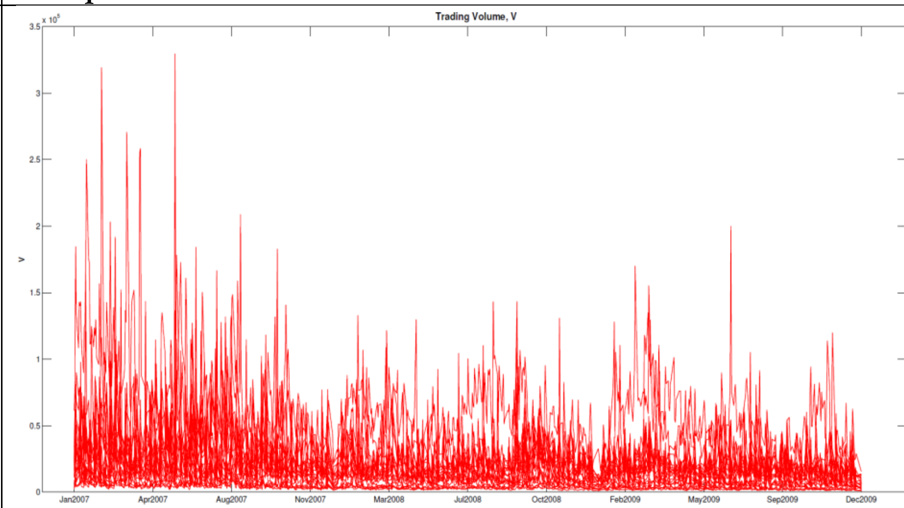
Figure 4- The figure shows averages intraday for the volume V_d and trade size Q_d for all stocks in the sample per day. The substantive graphs are divided in two groups. Group 1 refer to stocks with the highest (extreme) average values for volume and trade size and Group 2 to the remaining stocks.

Group 1





Group 2



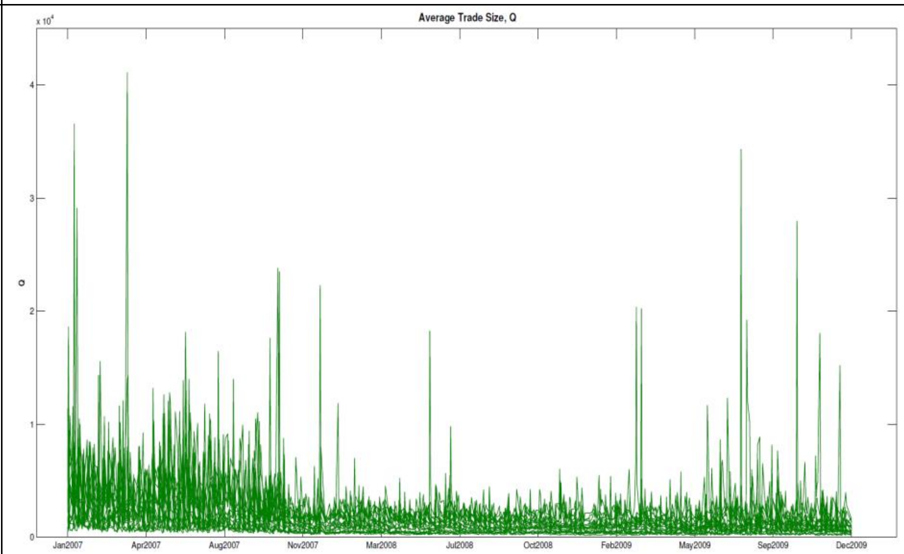


Table 8-OLS Regression results for Model 2 (Clark, 1973). Variables estimated as averages intraday. Stocks are grouped by market capitalization. $\boldsymbol{c}$ is the constant term of Clark's model specification in both (31) and (32). β_1 is the coefficient of Clark's model specification in (31), referring to a 2/3 proportionality. β_2 is the coefficient of Clark's model specification in (32), referring to a 1/2 proportionality. Significance against 2/3 or 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute					*5 minutes*			
Groups	***Stocks-t***	***c***	β_1	β_2	$\bar{R}^2$		***c***	β_1	β_2	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	3.9609*** (0.0608)	0.3489*** (0.0095)	0.2617*** (0.0071)	0.6411		5.0142*** (0.0337)	0.4321*** (0.0078)	0.3241*** (0.0059)	0.8030
	BP	6.1536*** (0.0851)	0.4665*** (0.0121)	0.3499*** (0.0091)	0.6644		6.5541*** (0.0393)	0.4501*** (0.0081)	0.3376*** (0.0061)	0.8032
	HSBA	6.5183*** (0.0732)	0.5457*** (0.0108)	0.4093*** (0.0081)	0.7709		6.5743*** (0.0340)	0.4925*** (0.0077)	0.3694*** (0.0058)	0.8444
	GSK	5.6859*** (0.0585)	0.5122*** (0.0096)	0.3842*** (0.0072)	0.7895		5.8937*** (0.0277)	0.4661*** (0.0071)	0.3496*** (0.0054)	0.8499
	VOD	5.3810*** (0.0973)	0.2762*** (0.0106)	0.2071*** (0.0080)	0.4717		7.1373*** (0.0539)	0.4076*** (0.0081)	0.3057*** (0.0061)	0.7723
Group 2 Upper Middle Mkt Cap	**BLT**	4.9495*** (0.0404)	0.4021*** (0.0078)	0.3016*** (0.0058)	0.7798		5.6137*** (0.0178)	0.4075*** (0.0064)	0.3056*** (0.0048)	0.8429
	BG	4.1371*** (0.0468)	0.2879*** (0.0074)	0.2159*** (0.0056)	0.6669		5.1753*** (0.02734)	0.3281*** (0.0066)	0.2461*** (0.0050)	0.7666
	XTA	4.1453*** (0.0741)	0.3002*** (0.0149)	0.2252*** (0.0112)	0.3507		5.1171*** (0.0342)	0.3311*** (0.0139)	0.2483*** (0.0105)	0.4280
	NG	4.8254*** (0.0781)	0.3860*** (0.0109)	0.2895*** (0.0082)	0.6237		5.5495*** (0.0440)	0.4055*** (0.0085)	0.3041*** (0.0064)	0.7513
	STAN	5.0605*** (0.0592)	0.4858*** (0.0103)	0.3644*** (0.0077)	0.7465		5.4955*** (0.0268)	0.4933*** (0.0077)	0.3700*** (0.0058)	0.8452
Group 3 Middle Mkt Cap	**EMG**	4.7499*** (0.0834)	0.3639*** (0.0108)	0.2729*** (0.0081)	0.6003		5.8852*** (0.0534)	0.4545*** (0.0096)	0.3409*** (0.0072)	0.7501
	OML	3.8276*** (0.1234)	0.2023*** (0.0126)	0.1517*** (0.0095)	0.2536		6.5385*** (0.0894)	0.4443*** (0.0119)	0.3332*** (0.0089)	0.6510
	WPP	5.5469*** (0.0999)	0.4860*** (0.0137)	0.3645*** (0.0103)	0.6264		6.2395*** (0.0473)	0.5442*** (0.0090)	0.4082*** (0.0067)	0.8307
	BLND	3.7137*** (0.0771)	0.2602*** (0.0113)	0.1952*** (0.0085)	0.4125		5.2793*** (0.0496)	0.4125*** (0.0107)	0.3093*** (0.0080)	0.6646
	RR	4.8482*** (0.0960)	0.3735*** (0.0122)	0.2801*** (0.0091)	0.5555		5.9468*** (0.0594)	0.4570*** (0.0103)	0.3427*** (0.0077)	0.7230
Group 4 Lower Middle Mkt Cap	**CCL**	4.1930*** (0.0706)	0.4305*** (0.0125)	0.3228*** (0.0094)	0.6129		4.9386*** (0.0371)	0.5214*** (0.0101)	0.3911*** (0.0076)	0.7806
	SMIN	3.9431*** (0.0925)	0.3371*** (0.0135)	0.2528*** (0.0101)	0.4516		5.0946*** (0.0563)	0.4498*** (0.0111)	0.3373*** (0.0083)	0.6857
	SHP	4.6189*** (0.1146)	0.4219*** (0.0173)	0.3164*** (0.0130)	0.4412		5.7176*** (0.0587)	0.5571*** (0.0122)	0.4178*** (0.0092)	0.7335
	IPR	4.7263*** (0.0911)	0.3525*** (0.0111)	0.2644*** (0.0083)	0.5722		5.9873*** (0.0575)	0.4547*** (0.0093)	0.3411*** (0.0070)	0.7610
	IMT	4.9233*** (0.0594)	0.5090*** (0.0105)	0.3817*** (0.0079)	0.7569		5.2806*** (0.0283)	0.5167*** (0.0080)	0.3875*** (0.0060)	0.8471
Group 5 Lowest Middle Mkt Cap	**SVT**	3.7135*** (0.0698)	0.3264*** (0.0109)	0.2448*** (0.0082)	0.5415		4.8332*** (0.0416)	0.4372*** (0.0090)	0.3279*** (0.0068)	0.7582
	CNE	2.5370*** (0.0493)	0.1630*** (0.0088)	0.1222*** (0.0066)	0.3147		3.8827*** (0.0333)	0.2739*** (0.0091)	0.2054*** (0.0068)	0.5478
	JMAT	4.0947*** (0.0802)	0.4095*** (0.0132)	0.3071*** (0.0099)	0.5620		4.9809*** (0.0428)	0.5128*** (0.0101)	0.3846*** (0.0075)	0.7756
	SGE	4.3190*** (0.1203)	0.3046*** (0.0135)	0.2285*** (0.0101)	0.4044		6.2826*** (0.0727)	0.4948*** (0.0102)	0.3711*** (0.0076)	0.7584
	REX	3.4196*** (0.1115)	0.2272*** (0.0138)	0.1704*** (0.0104)	0.2642		5.3380*** (0.0786)	0.4102*** (0.0125)	0.3077*** (0.0094)	0.5877

Table 9-OLS Regression results for Model 3 (Ané and Geman, 2000). Variables estimated as averages intraday. Stocks are grouped by market capitalization. $\boldsymbol{c}$ is the constant term of Ané and Geman's model specification in both (33) and (34). β_1 is the coefficient of Ané and Geman's model specification in (33), referring to a 2/3 proportionality. β_2 is the coefficient of Ané and Geman's model specification in (34), referring to a 1/2 proportionality. Significance against 2/3 or 1/2 proportionality is tested with a Wald test. Numbers in bold signify that the null hypothesis is accepted. $\bar{R}^2$ is the adjusted R-squared of the OLS regressions. * refers to 5%, ** to 1% and *** to 0.1% significance level.

		1 minute					***5 minutes***			
Groups	***Stocks-t***	***c***	β_1	β_2	$\bar{R}^2$		***c***	β_1	β_2	$\bar{R}^2$
Group 1 Highest Mkt Cap	**RDSA**	2.9830*** (0.0271)	0.4146*** (0.0089)	0.3110*** (0.0067)	0.7420		3.5650*** (0.0078)	0.4947*** (0.0071)	0.3711*** (0.0053)	0.8660
	BP	3.9360*** (0.0335)	0.3683*** (0.0114)	0.2762*** (0.0086)	0.5809		4.6917*** (0.0090)	0.5054*** (0.0107)	0.3790*** (0.0080)	0.7476
	HSBA	4.3023*** (0.0340)	0.5315*** (0.0122)	0.3986*** (0.0091)	0.7176		4.5887*** (0.0069)	0.5588*** (0.0085)	0.4191*** (0.0064)	0.8514
	GSK	3.7621*** (0.0224)	0.4761*** (0.0089)	0.3571*** (0.0067)	0.7928		4.1930*** (0.0046)	0.5015*** (0.0067)	0.3762** (0.0050)	0.8827
	VOD	3.9718*** (0.0445)	0.2490*** (0.0098)	0.1867*** (0.0074)	0.4590		5.2321*** (0.0198)	0.4235*** (0.0098)	0.3176*** (0.0073)	0.7147
Group 2 Upper Middle Mkt Cap	**BLT**	3.6972*** (0.0113)	0.4962*** (0.0064)	0.3722*** (0.0048)	0.8886		4.0734*** (0.0069)	0.5002*** (0.0055)	0.3752*** (0.0041)	0.9165
	BG	3.3614*** (0.0197)	0.3774*** (0.0070)	0.2831*** (0.0053)	0.7932		4.0178*** (0.0056)	0.4076*** (0.0059)	0.3057*** (0.0044)	0.8635
	XTA	3.3214*** (0.0265)	0.3980*** (0.0148)	0.2985*** (0.0111)	0.4896		3.9136*** (0.0187)	0.4161*** (0.0141)	0.3121*** (0.0106)	0.5359
	NG	3.7013*** (0.0343)	0.4574*** (0.0096)	0.3431*** (0.0072)	0.7527		4.1415*** (0.0116)	0.4561*** (0.0071)	0.3421*** (0.0053)	0.8448
	STAN	3.6084*** (0.0203)	0.5518*** (0.0083)	0.4139*** (0.0062)	0.8554		3.8179*** (0.0050)	0.5473*** (0.0062)	0.4105*** (0.0047)	0.9119
Group 3 Middle Mkt Cap	**EMG**	3.3654*** (0.0389)	0.3576*** (0.0097)	0.2682*** (0.0073)	0.6429		4.1551*** (0.0164)	0.4604*** (0.0087)	0.3453*** (0.0065)	0.7883
	OML	3.0310*** (0.0653)	0.2113*** (0.0116)	0.1585*** (0.0087)	0.3037		4.6774*** (0.0373)	0.4555*** (0.0113)	0.3417*** (0.0084)	0.6850
	WPP	3.7986*** (0.0525)	0.4832*** (0.0140)	0.3624*** (0.0105)	0.6113		4.2860*** (0.0154)	0.5671*** (0.0090)	0.4253*** (0.0068)	0.8398
	BLND	2.8545*** (0.0376)	0.2641*** (0.0108)	0.1980*** (0.0081)	0.4446		3.8641*** (0.0142)	0.4158*** (0.0103)	0.3118*** (0.0077)	0.6844
	RR	3.4806*** (0.0439)	0.3769*** (0.0104)	0.2827*** (0.0078)	0.6335		4.2727*** (0.0182)	0.4863*** (0.0089)	0.3647*** (0.0067)	0.7996
Group 4 Lower Middle Mkt Cap	**CCL**	3.0882*** (0.0301)	0.4744*** (0.0106)	0.3558*** (0.0080)	0.7262		3.4220*** (0.0079)	0.5346*** (0.0081)	0.4009*** (0.0061)	0.8514
	SMIN	3.0971*** (0.0403)	0.3973*** (0.0109)	0.2979*** (0.0082)	0.6378		3.7045*** (0.0161)	0.4900*** (0.0084)	0.3675*** (0.0063)	0.8180
	SHP	3.3801*** (0.0458)	0.4623*** (0.0135)	0.3468*** (0.0101)	0.6079		3.8712*** (0.0142)	0.5649*** (0.0090)	0.4237*** (0.0068)	0.8385
	IPR	3.5003*** (0.0460)	0.3711*** (0.0102)	0.2784*** (0.0077)	0.6362		4.3095*** (0.0206)	0.4760*** (0.0084)	0.3570*** (0.0063)	0.8103
	IMT	3.5251*** (0.0237)	0.5631*** (0.0089)	0.4223*** (0.0067)	0.8411		3.6883*** (0.0062)	0.5543*** (0.0070)	0.4157*** (0.0053)	0.8922
Group 5 Lowest Middle Mkt Cap	**SVT**	2.9086*** (0.0333)	0.3742*** (0.0097)	0.2807*** (0.0073)	0.6648		3.5592*** (0.0130)	0.4739*** (0.0076)	0.3554*** (0.0057)	0.8373
	CNE	2.3158*** (0.0277)	0.2445*** (0.0096)	0.1834*** (0.0072)	0.4602		3.1587*** (0.0094)	0.3545*** (0.0086)	0.2659*** (0.0065)	0.6929
	JMAT	3.1999*** (0.0338)	0.4966*** (0.0104)	0.3725*** (0.0078)	0.7501		3.5213*** (0.0106)	0.5552*** (0.0074)	0.4164*** (0.0056)	0.8818
	SGE	2.9861*** (0.0616)	0.2685*** (0.0119)	0.2014*** (0.0089)	0.4044		4.3694*** (0.0312)	0.4919*** (0.0094)	0.3689*** (0.0070)	0.7857
	REX	2.7793*** (0.0603)	0.2611*** (0.0131)	0.1958*** (0.0098)	0.3442		3.9425*** (0.0318)	0.4389*** (0.0116)	0.3292*** (0.0087)	0.6559