Dynamics of the limit order book and the volume-volatility relation

Manh Cuong Pham[1], Heather Margot Anderson*[2], Huu Nhan Duong[2], and Paul Lajbcygier[2]

[1] *Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, United Kingdom*

[2] *Monash Business School, Monash University, Wellington Road, Clayton, VIC 3800, Australia*

December 1, 2021

Abstract

We study how characteristics of the limit order book (LOB) affect return volatility and the volume-volatility relation at the transaction level. As important extensions to prior research, we allow for serial dependency of volatility on volume and for the volume-volatility relation to vary over time and depend on the dynamics of LOB characteristics. We find strong evidence that the LOB contains information about the volume-volatility relation and the return volatility of trades. The impact of LOB characteristics on return volatility is conveyed via two channels: a direct channel that is predominantly attributable to lagged LOB information, and an indirect channel that is transmitted via the volume-volatility relation and is mainly due to the current LOB slope.

Keywords: Volume-volatility relation; Limit order book; Order book slope; Order flow.

JEL classification: G10, G12.

Declarations of interest: None.

*Corresponding author. Tel: +61 3 9905 8462, Fax: +61 3 9905 5474.

Email addresses: m.c.pham@lancaster.ac.uk (M.C. Pham); heather.anderson@monash.edu (H.M. Anderson); huu.duong@monash.edu (H.N. Duong); paul.lajbcygier@monash.edu (P. Lajbcygier).

Acknowledgements: We are grateful for useful comments and suggestions from Neil Burgess, Catherine Forbes, Bonsoo Koo, Gael Martin, Ha Nguyen (discussant), Esen Onur, Tālis Putniņš, Elvira Sojli, Vu Le Tran (discussant), Farshid Vahid, conference participants at the 5th Vietnam International Conference in Finance 2018 (Ho Chi Minh city), the 4th International Conference on Accounting and Finance 2018 (Danang), the Monash Financial Markets Workshop 2018 (Monash University), and seminar participants at Monash University, Lancaster University and University of Bristol. We thank Rohan Fletcher for research assistant work. The research was funded by the CSIRO-Monash Superannuation Research Cluster, a collaboration across institutions and industry, among stakeholders of the retirement system in the interest of better outcomes for all. We would like to acknowledge the support of the Monash node of the NECTAR research, an initiative of the Australian government's Super Science Scheme and the Education Investment Fund cloud who provided IT infrastructure. Pham also acknowledges financial support from the ESRC-FWF bilateral grant titled "Bilateral Austria: Order Book Foundations of Price Risks and Liquidity: An Integrated Equity and Derivatives Markets Perspective", Grant Ref: ES/N014588/1 and the Austrian Science Fund (FWF): Research project: I-2762-G27. We are grateful to the Securities Industry Research Centre of Asia-Pacific (SIRCA) for providing the data used in our study. All remaining errors are our own.

1 Introduction

Electronic limit order books (LOBs) have become the dominant market trading platform in recent years, replacing traditional specialists or quote-driven trading platforms in many major financial markets around the globe (Bloomfield et al., 2005, Goettler et al., 2009, Malinova and Park, 2013). Coming with the popularity of limit order trading is increased interest in studying the properties of limit order book markets and their role in explaining high frequency price and return volatilty dynamics. While many theoretical and empirical studies examine the composition of order flows in LOB markets,[1] research into the use of LOB information to predict future returns and return volatility is relatively sparse, despite considerable evidence that the LOB drives the price process.[2]

We study how characteristics and dynamics of the limit order book (LOB) affect return volatility and the volume-volatility relation at the transaction level. Investigation of the relationship between trading volume and price volatility has been an area of active research in finance for a long time. According to Karpoff (1987), the study of the volume-volatility relation is important because it provides evidence on how information flows into the market, how it is processed and disseminated by the trading activities of market participants, and hence how it affects the price formation process.

Market microstructure theories provide several suggestions for a positive volume-volatility relation, and these include the arrival of new information that generates both price and volume movements (e.g. Clark, 1973, Harris, 1987, Andersen, 1996), disagreement among investors about asset values (e.g. Grundy and McNichols, 1989, Shalen, 1993, Banerjee and Kremer, 2010), and strategic trading activities by informed and uninformed traders in an asymmetric trading environment (e.g. Kyle, 1985, Holden and Subrahmanyam, 1992).

Meanwhile, many empirical studies (e.g. Ahn et al., 2001, Downing and Zhang, 2004, Chan

[1]These studies explore possible answers to questions such as which types of orders (limit vs. market) are often used by different types of investors (informed vs. uninformed), when these orders are used, and why. See, for example, Glosten (1994), Biais et al. (1995), Ranaldo (2004), Foucault et al. (2005), Wald and Horrigan (2005), Bloomfield et al. (2005), Anand et al. (2005), Kaniel and Liu (2006), Goettler et al. (2009), Roşu (2009), Chaboud et al. (2021).

[2]For example, Foucault et al. (2007), Nolte (2008), Pascual and Veredas (2010) and Jain and Jiang (2014) find that a wider bid-ask spread leads to higher future volatility. Thicker book depths help mitigate the return volatility of incoming orders (Ahn et al., 2001, Pascual and Veredas, 2010, Jain and Jiang, 2014). Further, a larger LOB slope is associated with a decrease in return volatility (Næs and Skjeltorp, 2006, Jain and Jiang, 2014, Tian et al., 2019), and a weaker correlation between the number of trades and return volatility (Næs and Skjeltorp, 2006, Jain and Jiang, 2014).

and Fong, 2006, Pascual and Veredas, 2010, Chevallier and Sévi, 2012, Carlin et al., 2014, Haugom et al., 2014, Valenzuela et al., 2015, Clements and Todorova, 2016, Tian et al., 2019) document the determinants of volatility, rather than the volume-volatility relation, because they all assume that the dependence of volatility on volume is constant and does not vary with other factors.

The objective of this paper is to examine the role that LOB information plays in explaining not only high frequency return volatility but also the volume-volatility relation. Unlike most existing studies which assume that the volatility-volume relation is fully contemporaneous (e.g. Jones et al., 1994, Downing and Zhang, 2004, Chan and Fong, 2006, Næs and Skjeltorp, 2006, Chevallier and Sévi, 2012, Clements and Todorova, 2016, Bollerslev et al., 2018), we allow for serial dependency of volatility on volume as implied by theoretical work of Copeland (1976), Shalen (1993) and Banerjee and Kremer (2010), and by empirical work of Manganelli (2005), Xu et al. (2006), Carlin et al. (2014), and Tian et al. (2019). We further allow the volume-volatility relation to vary over time and depend on the dynamics of LOB characteristics such as the bid-ask spread, the market depth at the inner quotes, and the LOB slope. The slope tracks how the quantity of stocks supplied in the LOB changes as a function of the limit price, and hence it provides a parsimonious summary of LOB information, not just at, but also beyond the best quotes.

We conduct our volume-volatility analysis at a transaction or tick-by-tick level, rather than at daily or lower frequencies used in most previous studies. The advantage of tick-by-tick analysis is that it fits the frameworks of most theoretical studies and enables a deeper understanding of how information from trades is incorporated into prices (e.g. Easley and O'Hara, 1987, Shalen, 1993, Holden and Subrahmanyam, 1992). Furthermore, a tick-by-tick analysis offers a natural remedy for the undetermined causality between volatility and its explanatory variables, which results from the fixed-time aggregation of trades and prices (Hasbrouck, 1995). It also helps avoid an information loss that comes from aggregating trades and prices over a fixed time interval that might bias estimation results (Engle, 2000, Manganelli, 2005, Russell and Engle, 2005). Acknowledging the random nature of trade arrival times at the transaction level, we follow Engle (2000) and Xu et al. (2006) in employing time-consistent measures of volume and volatility that are adjusted for time between trades. Our analysis also accommodates potential asymmetries between the bid and ask order books, and it controls for the effects of the order flow prior to a trade.

We examine the Australian limit order book market using a tick-by-tick dataset of stocks in the S&P/ASX200 index from July to December 2014. We find strong evidence that the LOB contains significant information about the volume-volatility relation and the return volatility of trades. The dependence of return volatility on trading volume is positive but dynamic and strongly related to LOB information, rather than contemporaneous and constant as typically assumed by most empirical studies. Both return volatility and the volume-volatility relation of a trade are positively associated with the bid-ask spread but negatively correlated with the market depth at the best quotes and the slope of the LOB prior to the transaction. These results highlight the dynamic nature of the volume-volatility relation, in contrast to previous findings in the literature that assume a constant volume-volatility relation (Ahn et al., 2001, Næs and Skjeltorp, 2006, Foucault et al., 2007, Nolte, 2008, Pascual and Veredas, 2010, Haugom et al., 2014, Jain and Jiang, 2014, Tian et al., 2019).

We find that the overall impact of LOB characteristics on the return volatility of an incoming trade is conveyed via two channels; a *direct* channel that is predominantly attributable to lagged LOB information, and an *indirect* channel that transfers through the volume-volatility relation and is mainly contributed by the prevailing LOB information right before the trade. This is because the direct (indirect) channel captures the partial effect of the LOB on return volatility without (with) the knowledge of a trade's volume. If the volume of a trade is not known, the recent past order book attributes, which contain information about recent past trading volumes and volatility (which in turn are correlated with the volume and volatility of the current trade), play a critical role in predicting future trading volume and return volatility. However, if the volume that a trade demands is known then the LOB characteristics immediately before the trade provide the most recent and relevant information about recent supply in the market, and this affects volatility via the second channel. In addition, the effects of LOB information on return volatility, either direct or indirect, depend on a stock's liquidity.

We observe significant asymmetries in the effects of the bid versus ask order books on return volatility, as in the studies by Ahn et al. (2001), Engle and Patton (2004), Harris and Panchapagesan (2005), Kalay and Wohl (2009) and Cenesizoglu et al. (2016), and our work shows that these asymmetries also affect the volume-volatility relation. We find that the order book of the opposite side to the direction of an incoming trade is more relevant for predicting **both** return volatility and the volume-volatility relation than the same side order book.

We also find that the bid-ask spread and market depth either switch signs or become less significant once we control for the LOB slope. This finding could arise because the LOB slope summarizes the LOB information at all quote levels, and thereby captures and dominates the information contained in the bid-ask spread and the market depth at the best quotes. Thus, the LOB slope is a key determinant of return volatility and the volume-volatility relation.

Our contributions to the literature are twofold. First, we extend research on the volume-volatility relation by showing that the dynamics of the LOB are important factors that drive a dynamic relationship between return volatility and trading volume at a high frequency tick-by-tick level. Existing studies primarily assume that the volume-volatility relation is either fully contemporaneous or does not vary with other characteristics that may explain volatility. In addition, they mainly examine the volume-volatility relation at a low frequency such as daily (see, amongst others, Næs and Skjeltorp, 2006, Chevallier and Sévi, 2012, Carlin et al., 2014, Wang and Wu, 2015, Clements and Todorova, 2016, Tian et al., 2019). Complementing prior findings, our study highlights the dynamic nature of the volume-volatility relation which is strongly dependent on the LOB characteristics at a transaction level.

Second, our work also contributes to the literature that examines the information content of the LOB, by providing empirical evidence that shows how characteristics of the LOB, in particular the slope, can *change* the volume-volatility relation. We also find strong evidence of asymmetries between the effects of the bid and ask order books, as noted elsewhere (e.g. Ahn et al., 2001, Engle and Patton, 2004, Harris and Panchapagesan, 2005).

The current study extends the closely related work of Næs and Skjeltorp (2006) that also investigates the informativeness of LOB attributes about the volume-volatility relation in several aspects. First, while Næs and Skjeltorp (2006) only consider the contemporaneous dependence of volatility on volume, we allow this relationship to be dynamic, as highlighted in the theoretical studies of Copeland (1976), Jennings et al. (1981), Shalen (1993) and Banerjee and Kremer (2010). Second, Næs and Skjeltorp (2006) do not control for the *direct* impact of LOB characteristics on volatility when estimating how the LOB information alters the volume-volatility relation. Consequently, they tend to overestimate the effects of the LOB characteristics on the volume-volatility relation. Our study, on the other hand, explicitly accounts for such direct impact, and it shows different mechanisms through which LOB information affects volatility, depending on whether the effect is direct or indirect. In addition, we allow for asymmetric effects of bid versus ask order books on the volume-volatility relation, which are not considered in Næs and Skjeltorp (2006); indeed, we find strong evidence of such asymmetries. Finally, instead of using lower frequency (daily) data as in Næs and Skjeltorp (2006), we employ high-frequency tick-by-tick data to conduct our analysis. This provides deeper insights into the price formation process, avoids information loss and mitigates the "undetermined causality" issue between volatility and its predictors due to fixed-time aggregation, as discussed further in subsection 2.2.

The rest of the paper is organized as follows. Section 2 details a general empirical framework that we employ to investigate the informativeness of the LOB characteristics about the volume-volatility relation. Section 3 describes the data, and discussions of the main empirical results follow in Section 4. In Section 5, we provide a graphical rationale for the informativeness of the LOB slope, and Section 6 offers some concluding remarks.

2 The volume-volatility relation

2.1 A general empirical framework

Let $\sigma_{i,t}$ and $v_{i,t}$ be the volatility and volume associated with stock i at time t (or during the time interval t). We study the volume-volatility relation using the following equation:

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{q} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t}, \tag{1}$$

where $Monday_{i,t}$ (a dummy variable for Monday) and $hour1_{i,t}$ (a dummy variable for the first trading hour (10:10:00-11:00:00) of a day) allow for the Monday and opening effects on volatility, $\sum_{j=1}^{q} \alpha_j \sigma_{i,t-j}$ captures the persistence in volatility, $x_{i,t}$ is a vector of explanatory variables that potentially influence the volume-volatility relation as well as the return volatility itself, $y_{i,t}$ is a vector of control variables that allow for the effects of past order flow information and $\eta_{i,t}$ is a zero-mean disturbance term,

Equation (1) allows for both *contemporaneous* and *lagged* effects of trading volume on volatility, in contrast to many previous empirical studies that only examine the *contemporaneous* correlation between volume and volatility (e.g. Chan and Fong, 2006, Næs and Skjeltorp, 2006, Shahzad et al., 2014, Wang and Wu, 2015, Bollerslev et al., 2018, Duong et al., 2018). Our accommodation of *dynamics* in volume-volatility dependence is consistent with various traditional finance theories. For example, the sequential arrival of information hypothesis (SAIH) of Copeland (1976) and Jennings et al. (1981) implies that there is a lead-lag relationship between volume and volatility, which results from sequential (rather than simultaneous) dissemination of information to market participants. Similarly, theoretical models that feature heterogeneity in investors' beliefs about asset prices, due to either asymmetric private information (Shalen, 1993) or differences of opinions about public information (e.g. Harris and Raviv, 1993, Banerjee and Kremer, 2010), show that traders' over-reaction to a change in trading activities gives rise to both contemporaneous and serial dependencies of volatility on volume. Meanwhile, the microstructure model of Hasbrouck (1991a,b) suggests that microstructure imperfections such as price discreteness and inventory control might induce lagged adjustments in stock prices to a trade's information, implying that past trading volumes could be informative about future prices and volatility. Our inclusion of dynamics in Equation (1) is also consistent with empirical work by Manganelli (2005), Xu et al. (2006), Nolte (2008), Carlin et al. (2014), Do et al. (2014), and Tian et al. (2019) that finds significant current and lagged volume effects on return volatility.

More importantly, Equation (1) allows the marginal effect of volume at time $t-k$ on return volatility at time t to vary with $x_{i,t-k}$ via the function $\beta_{0,k} + \delta_k' x_{i,t-k}$,[3] which relaxes the assumption of a constant volume-volatility relation that is implicit in most previous empirical work (e.g. Jones et al., 1994, Chan and Fong, 2000, Downing and Zhang, 2004, Pascual and Veredas, 2010, Carlin et al., 2014, Clements and Todorova, 2016, Tian et al., 2019). Our specification of this marginal effect as a linear function of $x_{i,t-k}$ in Equation (1) is similar to the econometric methodology used by Dufour and Engle (2000) to investigate the informativeness of trade arrival times for explaining the price impact of a trade. It is also similar to work in Avramov et al. (2006) that finds a link between selling activity and the well known asymmetry in daily volatility that is known as the "leverage effect" in individual stock returns.[4]

[3]We focus on the case for which the marginal effect of volume $v_{i,t-k}$ on return volatility $\sigma_{i,t}$ is a linear function of $x_{i,t-k}$ to simplify our analysis, and leave the study of (potentially nonparametric) extensions of this functional form for future research.

[4]Bollerslev et al. (2018) also use a similar specification and find that the volume-volatility elasticities of the S&P500 equity index and U.S. Treasury bonds become weaker around public news announcements and when there is a more disagreement in beliefs amongst investors.

The $\beta_{0,0} + \delta_0' x_{i,t}$ term in Equation (1) captures the contemporaneous impact of volume on volatility, $\beta_{0,k} + \delta_k' x_{i,t-k}$ ($k \geq 1$) captures lagged impacts, and $\sum_{k=0}^{p}[\beta_{0,k} + \delta_k' x_{i,t-k}]$ measures the cumulative impact. Meanwhile, $x_{i,t-k}$ influences return volatility $\sigma_{i,t}$ via two channels - via its *direct* impact on the latter (captured by γ_k), and via its *indirect* effect that alters the volume-volatility relation (captured by δ_k).

2.2 Proxies for volatility and volume

We examine the dynamic volume-volatility relation at a tick-by-tick or transaction level of detail. Although the use of transaction data is widespread in the market microstructure literature (e.g. Hasbrouck, 1991a,b, Dufour and Engle, 2000, Barclay et al., 2003) and in the duration-volatility modeling literature (e.g. Engle, 2000, Renault and Werker, 2011, Renault et al., 2014), most research on the volume-volatility relation has worked with daily or lower frequency data[5] and only a few studies, including Harris (1987), Manganelli (2005), Xu et al. (2006) and Nolte (2008), provide an examination of the volume-volatility relation at a transaction level. This is quite surprising, given that most theoretical studies on the topic develop their analysis at a tick-by-tick level (see, for example, Kyle, 1985, Easley and O'Hara, 1987, Holden and Subrahmanyam, 1992, Shalen, 1993).

Our tick-by-tick empirical analysis of the volume-volatility relation leads to a deeper understanding of how prices adjust to absorb the information from trades because, as highlighted in theoretical work by Diamond and Verrecchia (1987) and Easley and O'Hara (1992), the existence (absence) of each individual trade may signal the existence (absence) of news events and the presence (absence) of informed traders in the market, and this is informative about price formulation. In addition, our tick-by-tick analysis avoids bias (discussed in Engle (2000), Manganelli (2005), Russell and Engle (2005)), due to a loss of information that results from the aggregation of trades and prices over a fixed time interval. It also provides a natural solution to "undetermined causality" between volatility and its predictors that has been recognized in prior studies that employ

[5]See, amongst others, Jones et al. (1994), Andersen (1996), Chan and Fong (2000), Avramov et al. (2006), Næs and Skjeltorp (2006), Giot et al. (2010), Chevallier and Sévi (2012), Carlin et al. (2014), Do et al. (2014), Shahzad et al. (2014), Tian et al. (2019). Also see Karpoff (1987) for a detailed survey of related work. There are also several studies that examine the volume-volatility relation at an intradaily level such as Ahn et al. (2001) and Bollerslev et al. (2018) (15 minutes), Pascual and Veredas (2010) (1 and 5 minutes), Duong and Kalev (2014) (30 minutes), and Jain and Jiang (2014) (1 minute).

lower frequency data (see subsection 2.4 for a more detailed discussion).

Volatility is typically measured over a fixed time interval such as an hour, a day or a week, depending on the frequency at which data are sampled. Two fixed-interval volatility measures that are common in the volume-volatility literature include (i) the absolute size of residuals from an autoregression of returns (e.g. Jones et al., 1994, Chan and Fong, 2000, Avramov et al., 2006, Næs and Skjeltorp, 2006); and (ii) realized volatility measures[6] (e.g. Chan and Fong, 2006, Giot et al., 2010, Chevallier and Sévi, 2012, Shahzad et al., 2014, Tian et al., 2019). However, when working with tick-by-tick data, researchers face a challenging issue which is that transactions arrive in the market at irregularly spaced times. This complicates the measurement of the volatility of each trade because the time between consecutive trades is not the same. Engle (2000) and Xu et al. (2006) suggest adjusting volatility by trade durations to obtain *volatility per unit of time*, which provides a natural and meaningful measure of volatility in tick-by-tick empirical analyses. In this study, we estimate the volatility per unit of time for a transaction by dividing the absolute size of the residual from the following regression by the duration of the trade:

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{q} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}. \tag{2}$$

where $r_{i,t}$ denotes the return of the t-th trade in stock i, defined as the change in the natural logarithms of the bid-ask midpoint following the trade and is quoted as a percentage i.e. $r_{i,t} = 100(\ln(q_{i,t+1}) - \ln(q_{i,t}))$, where $q_{i,t}$ is the midpoint of the bid and ask quotes immediately before the t-th trade. $Day_{k,i,t}$ are day-of-week dummies, and $hour_{k,i,t}$ are time-of-day dummy variables.[7] Lagged returns ($r_{i,t-k}$) are used to control for the autocorrelation in the return series. The incorporation of $x_{i,t}$ and $y_{i,t}$ into Equation (2) allows for their possible power in explaining returns, which then ensures that the effects of $x_{i,t}$ and $y_{i,t}$ on volatility obtained from Equation (1) are genuine and not driven by the impact of $x_{i,t}$ and $y_{i,t}$ on returns.[8] We define $\sigma_{i,t}$ by $\sigma_{i,t} := |\widehat{\epsilon}_{i,t}|/T_{i,t}$,

[6] The most common of these is daily realized volatility, which is often calculated as the sum of squared five-minute returns over a day, but there are many different types of realized volatility measures. See the survey article by McAleer and Medeiros (2008) for discussion on such measures and their properties.

[7] Each trading day in the Australian Securities Exchange (ASX) is partitioned into six hourly intervals: 10:10-11:00, 11:00-12:00, 12:00-13:00, 13:00-14:00, 14:00-15:00 and 15:00-16:00. All trades in the first 10 minutes of each trading day are excluded from the analysis to avoid the effects of the ASX opening procedure. The first five hourly dummies are included in Equation (2), while the last trading hour serves as the base category.

[8] The inclusion of $x_{i,t}$ and $y_{i,t}$ in the return Equation (2) ensures that $\widehat{\epsilon}_{i,t}$, whose absolute value scaled by trade duration provides a proxy for the volatility of the trade, will be orthogonal to $x_{i,t}$ and $y_{i,t}$. Therefore, the effects of $x_{i,t}$ and $y_{i,t}$ on volatility in Equation (1), if any, will not be due to correlation between $x_{i,t}$, $y_{i,t}$ and the unexpected returns. Similarly, the two dummy variables in Equation (1) capture time-of-day and day-of-week effects that are not captured by Equation (2).

where $T_{i,t}$ is the duration of the t-th trade which measures the time (in seconds) between the $(t-1)$-th and t-th trades, and we use $\sigma_{i,t}$ as our proxy for volatility in Equation (1).[9]

Two popular proxies for volume in the volume-volatility literature are the number of trades and the average trade size during a fixed time interval. While both measures are found to be positively related to return volatility, prior studies often document that the number of trades, which essentially captures trading intensity, is far more informative about return volatility than is average trade size (e.g. Chan and Fong, 2006, Næs and Skjeltorp, 2006, Chevallier and Sévi, 2012). When working with transaction data, previous studies usually measure the volume of a trade by the number of shares executed by the trade (e.g. Hasbrouck, 1991a,b, Manganelli, 2005, Nolte, 2008), which coincides with the average trade size since the number of trades at a transaction time is always one.[10] Motivated by work in Engle (2000) and Xu et al. (2006) that adjusts variables computed in transaction time to account for calendar time, we employ a time-consistent measure of volume $v_{i,t}$, called *volume per unit of time*, which is defined as $V_{i,t}/T_{i,t}$, where $V_{i,t}$ is the number of shares traded (times 1000) divided by the total number of shares outstanding right before the t-th trade in stock i[11] and $T_{i,t}$ is the duration of the trade (as defined above).

2.3 Limit order book characteristics and the volume-volatility relation

Prior studies demonstrate the informativeness of LOB information about the price formation process. For example, Hasbrouck (1991a) documents that a wider bid-ask spread is associated with an increase in the price impact or return of an incoming trade. Similarly, Foucault et al. (2007) develop a theoretical model that predicts a positive relationship between the bid-ask spread and future volatility, which is strongly supported by empirical evidence provided in Næs and Skjeltorp (2006), Foucault et al. (2007), Nolte (2008), Pascual and Veredas (2010) and Jain and Jiang (2014). Meanwhile, an increase in market depth leads to a decline in the price impact and return

[9]Our proxy for volatility is analogous to the daily volatility measure used by Jones et al. (1994), Chan and Fong (2000), Avramov et al. (2006) and Næs and Skjeltorp (2006), and like these daily measures it is time-consistent, but at the transaction level. Note that a tick-by-tick version of realized volatility is not defined and cannot be computed.

[10]Although simultaneous transactions can occur, they typically result from the matching of one big market order against several opposite side limit orders and they are usually considered as one big trade in empirical analyses.

[11]Standardizing the number of shares in a trade by the number of outstanding shares right before the trade helps facilitate comparison between different stocks by putting them on roughly the same footing. We obtain qualitatively similar results without this standardization.

volatility of future trades (Ahn et al., 2001, Pascual and Veredas, 2010, Brogaard et al., 2015). In addition, Pham et al. (2020) document that the use of market depth information right before a trade significantly improves the prediction of the immediate price impact of the trade.

Several studies have shown that LOB information beyond the inner quotes is also informative about returns and volatility. Pascual and Veredas (2010) document a significant negative dependence of the ex-post informational volatility of the latent efficient price process on depth beyond the best quotes. Likewise, Næs and Skjeltorp (2006), Duong and Kalev (2014), Jain and Jiang (2014) and Tian et al. (2019) show that a steeper LOB curve is associated with lower return volatility, while Kalay and Wohl (2009) find that their buying pressure measure, which is calculated from the slopes of the bid and ask order books, is predictive of future returns. Another study by Cao et al. (2009) shows that the limit order book information after the first level is moderately informative, contributing to 22% of the price discovery of Australian stocks. Similarly, Kozhan and Salmon (2012) demonstrate that LOB information beyond the best prices can predict future (but in-sample) price adjustments, even though a simple trading strategy does not lead to out-of-sample economic profits. Meanwhile, Cenesizoglu et al. (2016) find significant asymmetric effects of the bid and ask slopes on price dynamics, which if ignored by traders when designing trading strategies could lower their daily profits by about 25 basis points.

In this study, we investigate the role played by the LOB in explaining the dynamic volume-volatility relation and ultimately the return volatility of trades at a tick-by-tick level. Following prior literature, the LOB characteristics that we examine consist of (i) the relative bid-ask spread, $Spread_{i,t}$, defined as the quoted spread divided by the mid-quote right before a trade; (ii) the market depth available at the inner quotes, $Depth_{i,t}$, defined as the total number of shares available at the best bid and ask prices (times 1000) and standardized by the total number of shares outstanding right before a trade; and (iii) the slope of the LOB immediately before a trade. The latter variable captures the steepness of the limit order book, and it essentially measures how the quantity of stocks supplied in the LOB changes as a function of the limit price. The slope measure summarizes the LOB information at all limit price levels, whereas the first two attributes (i.e. bid-ask spread and market depth) only capture the LOB information at the best quotes.

Following Næs and Skjeltorp (2006), we compute the LOB slope for stock i immediately before transaction time t or the t-th trade as:

$$Slope_{i,t} = \frac{BidSlope_{i,t} + AskSlope_{i,t}}{2}, \tag{3}$$

where $BidSlope_{i,t}$ and $AskSlope_{i,t}$ respectively denote the slopes of the bid and ask order books and are given by

$$BidSlope_{i,t} = \frac{1}{100N^B}\left\{\frac{v_1^B}{|p_1^B/p_0 - 1|} + \sum_{\tau=1}^{N_B-1}\frac{v_{\tau+1}^B/v_\tau^B - 1}{|p_{\tau+1}^B/p_\tau^B - 1|}\right\}, \quad \text{and} \tag{4}$$

$$AskSlope_{i,t} = \frac{1}{100N^A}\left\{\frac{v_1^A}{p_1^A/p_0 - 1} + \sum_{\tau=1}^{N_A-1}\frac{v_{\tau+1}^A/v_\tau^A - 1}{p_{\tau+1}^A/p_\tau^A - 1}\right\}, \tag{5}$$

where N^B and N^A are the total number of bid and ask prices (tick levels) containing orders of stock i right before time t, and τ designates tick levels for stock i that have positive share volumes at that time. The best bid (ask) price is denoted by p_1^B (p_1^A) and corresponds with $\tau = 1$, and p_0 denotes the best bid-ask midpoint immediately prior to time t. The quantities v_τ^B and v_τ^A denote the natural logarithms of the *accumulated* total share volume at each tick level τ on the bid and ask sides right before time t, so that if we define V_τ^B (V_τ^A) as the total share volume demanded (supplied) at p_τ^B (p_τ^A), then $v_\tau^B = \ln\left(\sum_{j=1}^{\tau} V_j^B\right)$ measures the natural logarithm of the total share volume demanded at p_τ^B or higher, and $v_\tau^A = \ln\left(\sum_{j=1}^{\tau} V_j^A\right)$ measures the natural logarithm of the total share volume supplied at p_τ^A or lower. Intuitively, the bid (ask) slope measures the percentage change in the bid (ask) volume relative to the percentage change in the corresponding bid (ask) price, which is averaged across all limit price levels in the bid (ask) order book, and the LOB slope is an equally weighted average of the bid and ask slopes. For each point in transaction time t, we use the 10 best bid and ask quotes, together with the share volumes queued at these quotes right before time t, to calculate the LOB slope.[12] We do not include information relating to undisclosed or hidden orders in this calculation of the LOB slope.

Market microstructure studies highlight the importance of trade direction or trade type (i.e. buy vs. sell) in explaining price dynamics (e.g. Hasbrouck, 1991a,b, Dufour and Engle, 2000, Barclay et al., 2003). In particular, an unexpected purchase (sale) results in a significant increase (decrease) in a stock's price. Meanwhile, Ahn et al. (2001), Engle and Patton (2004), Harris and Panchapagesan (2005), Kalay and Wohl (2009) and Cenesizoglu et al. (2016) document that there are significant asymmetries between the bid and ask sides of the LOB that are important when

[12] We use all available quotes with positive volumes to compute the slope if less than 10 levels of the best bid and ask quotes are available for stock i at time t. We also undertake some robustness analysis in subsection 4.4, that uses different sets of LOB information (of 5 and 20 best bid and ask levels) to compute the slope measures.

explaining price dynamics. We use trade direction information to allow for potential asymmetric effects of the bid and ask order books, by splitting the $Depth_{i,t}$ and $Slope_{i,t}$ measures into the corresponding bid and ask quantities, and interacting them with trade indicators.[13]

Previous studies find that information about past order flow has explanatory power for return volatility (e.g. Chan and Fong 2000, Chan and Fong 2006, Shahzad et al. 2014), so we incorporate several variables into our return volatility and volume-volatility models to control for this. Our controls also partly mitigate endogeneity issues associated with the joint determination of trading volume and LOB variables. The vector of control variables that we employ is $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$, where $T_{i,t}$ is the duration of the t-th trade, while $N_{i,t}$, $ATS_{i,t}$, $OIB_{i,t}$ and $QTT_{i,t}$ respectively measure the number of transactions, the average trade size (times 10^6 and divided by the total number of shares outstanding), the order imbalance (defined as the number of buys minus the number of sells), and the quote to trade ratio (defined as the total number of order submissions, revisions and cancellations divided by the number of trades) during the 5-minute interval right before the t-th trade. See Table A.1 in the Appendix for a complete list of all variables used in this study.

Overall, the main vector of interest $x_{i,t}$ in Equation (1) contains LOB characteristics, consisting of $Spread_{i,t}$, $Depth_{i,t}$ and $Slope_{i,t}$ if a combined LOB is considered, or $Spread_{i,t}$, $BV_{i,t}B_{i,t}$, $BV_{i,t}S_{i,t}$, $AV_{i,t}B_{i,t}$, $AV_{i,t}S_{i,t}$, $BidSlope_{i,t}B_{i,t}$, $BidSlope_{i,t}S_{i,t}$, $AskSlope_{i,t}B_{i,t}$, and $AskSlope_{i,t}S_{i,t}$ if we allow for a separation of the bid and ask order books. Note that the imposition of the restrictions that $p = 0$ and $\delta_0 = 0$ on Equation (1) gives a constant contemporaneous volume-volatility relation model similar to those examined in the existing literature. The imposition of $p = 0$ on Equation (1) gives an "*endogenous*" contemporaneous volume-volatility relation model similar to the model estimated by Næs and Skjeltorp (2006) for daily data.[14] We use the word "endogenous" here to indicate that the volume-volatility relation is no longer constant but dependent on the LOB characteristics. For the full model (1) that allows the dynamic dependence of return volatility on

[13]The resulting set of LOB attributes consist of $BV_{i,t}B_{i,t}$, $BV_{i,t}S_{i,t}$, $AV_{i,t}B_{i,t}$, $AV_{i,t}S_{i,t}$, $BidSlope_{i,t}B_{i,t}$, $BidSlope_{i,t}S_{i,t}$, $AskSlope_{i,t}B_{i,t}$, and $AskSlope_{i,t}S_{i,t}$, where $BV_{i,t}$ ($AV_{i,t}$) is the bid (ask) depth volume, $BidSlope_{i,t}$ ($AskSlope_{i,t}$) is the bid (ask) order book slope, and $B_{i,t}$ ($S_{i,t}$) is a buy (sell) indicator that equals 1 if the t-th trade is a buy (sell), and 0 otherwise.

[14]We note that Næs and Skjeltorp (2006) do not control for the direct impact of LOB information on volatility. Specifically, they first compute the sample correlation between the daily number of trades and daily volatility in every month, and then regress this monthly correlation series on the monthly averages of the LOB attributes for a panel of all stocks in their sample.

trading volume to be dependent on the LOB attributes, we assume that the model can be truncated at $p = 5$ lags, as is typically assumed in the microstructure literature (e.g. Hasbrouck, 1991a,b, Dufour and Engle, 2000, Xu et al., 2006). We also truncate the lags of returns and volatility in Equations (1) and (2) at $q = 12$, as typically done in previous studies (e.g. Avramov et al., 2006, Chan and Fong, 2006, Chevallier and Sévi, 2012).

2.4 A note on causality

Previous empirical studies on the volume-volatility relation have found two-way causal relationships between volatility and other variables such as trading volume and LOB information. This arises from the use of low frequency data that has been constructed by aggregating over higher frequency observations (Hasbrouck, 1995, Barclay et al., 2003, Benos and Sagade, 2016). For example, a large transaction can have a big impact on security prices and increase price volatility, which then sends signals to the market and affects the aggressiveness of quotes, trading intensity and the volume of subsequent trades (Easley and O'Hara, 1987, 1992, Dufour and Engle, 2000). If data for these high frequency events is aggregated, the (lower frequency) volatility and volume measures will be jointly determined and contemporaneously correlated, making it difficult to disentangle the causal relationship between the two. Similarly, the time aggregation of information relating to quotes and trades leads to undetermined causality between volatility and LOB information, as empirically observed in Næs and Skjeltorp (2006). Given that trading and quoting activities often arrive sequentially, Hasbrouck (1995) suggests that shortening the sampling time interval might mitigate this issue as it reduces contemporaneous correlation induced by time aggregation.

Motivated by Hasbrouck's (1995) suggestion, the current study utilizes tick-by-tick data, and it assumes that at this level of time resolution there are Granger-causal relationships running from trading volumes and LOB characteristics to return volatility, and that the LOB information Granger-causes the volume-volatility relation of trades. Intuitively, we assume that the state of the LOB is predictive of how incoming trades affect prices and return volatility. Given that our LOB attributes (trading volumes) are known right before (at) the execution of a trade, whereas the return and volatility of the trade can only be realized ex-post, these assumptions are intuitive and reflect the chronological operation of an electronic LOB market.

We note that our use of transaction data and the above assumptions does not completely rule out the potential endogeneity of trading volumes and LOB characteristics. This is because these variables are correlated with traders' unobserved liquidity needs and their information sets, which certainly influence return volatility. However, it would be difficult to find appropriate instruments for volume and LOB characteristics because (i) such instruments would need to be measured at a tick-by-tick level; and (ii) they must only affect return volatility indirectly through trading volume and LOB information (i.e. they must be uncorrelated with the error of the volatility equation), which is most unlikely given the dynamics of trading and price formation. Therefore, instead of finding possible instruments, we ameliorate the effects of this potential endogeneity by augmenting our models with lags of return volatility as well as lags of variables that account for the order flow information prior to a trade (see discussions in the previous subsections). This is common practice in the volume-volatility literature, and we have confidence that under our tick-by-tick setting that helps attenuate time-aggregation contemporaneous correlations, these lagged variables will capture sufficient information about traders' unobserved characteristics to minimize the endogeneity and joint determination problems and maintain the validity of our results.

3 Data

3.1 The Australian stock market

The Australian Securities Exchange (ASX) is amongst the 15 largest listed exchanges in the world by market capitalization. The ASX has operated as a purely electronic order-driven market since 1991. Orders submitted to the ASX follow a price-time priority, as in most other electronic LOB markets. In particular, limit orders are queued and ranked in the LOB first by price priority and then in chronological order. Meanwhile, market orders, which are orders with the highest price priority, are executed at the best available price immediately upon submissions. The LOB is updated instantaneously whenever an order submission, revision, cancellation, or execution occurs. The submitted price of an order must be in multiples of the minimum tick size, which is pre-specified by the exchange and is dependent on the price level of the security. The tick size is currently AUD$ 0.001, 0.005, and 0.01 for stock prices that are below AUD$ 0.1, from AUD$ 0.1 but below AUD$

2, and from AUD$ 2, respectively. A typical trading day consists of two sessions: a pre-market session from 7:00am to 10:00am Australian Eastern Standard Time (AEST), and a normal trading session from 10:00am to 4:00pm AEST. The first 10 minutes of the normal trading session are opening auctions. There is also a closing single price auction between 4:10pm and 4:12pm during which the daily closing price for each stock is determined (see http://www.asx.com.au).

3.2 The data

We investigate the informativeness of the LOB about the dynamic volume-volatility relation using stocks in the S&P/ASX200 index between 1 July and 31 December 2014. This sample period is chosen to avoid the confounding complications resulting from major upgrades of the technology infrastructure (including all main trading and post-trade systems) of the ASX after 2015 (ASX, 2015). The S&P/ASX200 index is the primary stock market index that serves as the main investment benchmark in Australia and it constitutes about 80% of Australia's sharemarket capitalization. We follow the ASX's classification to partition these stocks into three groups: "**Large cap**" which contains stocks in the S&P/ASX50 index, "**Mid cap**" which contains stocks in the S&P/ASX100 index but outside the S&P/ASX50 index, and "**Small cap**" which contains the remaining stocks in the S&P/ASX200 index. There were 198 stocks in our sample, consisting of 49 large cap, 50 mid cap, and 99 small cap stocks.[15]

We collect two datasets from the Securities Industry Research Centre of Asia-Pacific (SIRCA) database. The first dataset records details on every order submitted to the Australian central LOB, including the stock code, the order type (order submission, order revision, order cancellation and execution), the date and time (to millisecond precision), the order price, the order volume (number of shares), the order value (dollar value), and the order qualifiers.[16] Each new order is assigned a unique identification number (ID) so that it can be tracked from its initial submission through any

[15] We exclude two stocks, namely WES (Wesfarmers Limited - Large cap) and NWS (News Corporation - Small cap) from our analysis, since our database did not record data for these stocks during the sample period, even though they were listed and traded throughout the period. In addition, there are 6 stocks that were delisted during the sample period. We do not remove them from our sample since our analysis is conducted on a stock-by-stock basis and we still have a large sample size for these stocks (of more than 4000 transactions). Nevertheless, excluding these stocks negligibly affects our results.

[16] Each limit or non-market order has a qualifier indicating the order direction (buy or sell order). Trade and market orders also have qualifiers that declare their various qualitative properties. Examples include "Bi" ("Si") qualifiers that signal buyer-initiated (seller-initiated) trades, while "XT" denotes a cross trade and "CX" signal trades that are executed in an Australian dark pool called Centre Point.

revision, cancellation or execution. We extract all trades that occur within the continuous trading session in the lit market (from 10:10:00 to 16:00:00) and discard all transactions executed in the opening auction (i.e. either during 10:00:00-10:10:00 or with "AC" qualifiers that define auction trades) or in dark pools. We classify trades into buyer-initiated and seller-initiated trades based on the direction of the (market) orders that initiate the trade. Since one large buy (sell) market order can be matched against several limit orders queuing on the sell (buy) side of the LOB and appear as multiple instantaneous transactions that have zero durations, we follow standard practice in the literature (e.g Dufour and Engle, 2000, Nowak and Anderson, 2014, Renault et al., 2014) and aggregate same-direction trades executed at the same time into one "big" trade by calculating volume-weighted average prices and summing up the volumes of small trades. We use the order book dataset to compute our control variables that allow for the effects of order flow during the 5-minute interval immediately prior to each trade (i.e. the number of transactions ($N_{i,t}$), the average trade size ($ATS_{i,t}$), the order imbalance ($OIB_{i,t}$), and the quote to trade ratio ($QTT_{i,t}$)).

The second dataset contains detailed information on stock code, date, time, and the best bid/ask quotes and volumes up to 20 levels in the LOB. We remove all observations with either (i) a negative bid/ask quote or volume at any level, (ii) a bid quote higher than ask quote at any level, (iii) a positive bid or ask quote but with zero volume at any level, (iv) a zero bid or ask quote but with a positive volume at any level, or (v) a bid (ask) quote at level j lower (higher) than or equal to the bid (ask) quote at level $k > j$.[17] The transaction data are merged with the bid and ask quotes data to work out the bid-ask midpoint, the bid-ask spread, the depth volume at the best bid and ask quotes, and the LOB slope immediately before each transaction. Finally, we collect daily data on the numbers of shares outstanding for each stock from the DatAnalysis Premium database. We winsorize all variables at the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentiles on a stock-by-stock basis[18] to avoid the effects of outliers, The winsorization filters out another four stocks, leaving us

[17]The last filtering criterion ensures that bid (ask) quotes are in strictly decreasing (increasing) order as one moves further away from the best, i.e. level 1, bid (ask) quote. However, it is worth noting that some stocks, especially the illiquid stocks, might not have all 20 levels of the best bid/ask quotes and volumes available at some point in time, but only 5 or 10, for example, levels instead. In such a case, entries for the bid/ask prices and volumes of the remaining levels are displayed as 0. These observations are still valid and hence will not be removed if they pass the first four aforementioned filtering criteria ((i) - (iv)).

[18]Note that we estimate the return volatility of a trade as $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$, where $\widehat{\epsilon}_{i,t}$ is the residual obtained from an autoregressive model in Equation (2) of winsorized returns, and $T_{i,t}$ is the winsorized duration of the trade. We do not winsorize $\sigma_{i,t}$ since it would be effectively a double winsorization.

with a final sample of 194 stocks, consisting of 49 large cap, 48 mid cap, and 97 small cap stocks.[19]

Table 1 provides some cross-sectional summary statistics of trades and order book characteristics for the constituent stocks of the S&P/ASX200 index during July-December 2014 that we study. Consistent with previous empirical evidence, the transaction returns for all stock groups have a mean of almost zero percent. In conformance with well-documented stylized facts (e.g Manganelli, 2005, Xu et al., 2006, Jondeau et al., 2015, Pham et al., 2020), larger cap stocks trade more frequently and consequently have significantly smaller trade durations, reflecting their higher levels of liquidity. Moreover, they tend to trade in a smaller volume, either in the number of shares or per unit of time. Thus, the larger cap stocks have smaller volatility, as evidenced by the smaller absolute return per unit of time - a raw proxy for return volatility per unit of time.

<<INSERT TABLE 1 ABOUT HERE>>

Regarding the LOB characteristics, larger cap and more liquid stocks on average have a smaller relative spread, as reported elsewhere (e.g. Dufour and Engle, 2000, Næs and Skjeltorp, 2006). Interestingly, large cap stocks have significantly fewer shares supplied at the inner quotes than do mid and small cap stocks, possibly because the former are much more heavily traded so that more depth at the best quotes is absorbed. While the number of shares or depth available at the best bid and ask quotes are roughly equal for large cap stocks, significantly more shares are queued at the best bid than at the best ask for mid and small cap stocks. For all stock groups, the average amount of shares supplied at the best quotes is much larger than the average volume demanded by a trade, implying that the majority of transactions do not move the best bid or ask prices and hence have zero returns - an observation that is also documented by Dufour and Engle (2000), Renault et al. (2014), Pham et al. (2020), amongst others. The bid, ask and overall order book slopes are larger for more liquid stocks, suggesting that for these stocks more shares are queued closer to the inner bid/ask quotes, making their LOB steeper, which is consistent with the findings of Næs and Skjeltorp (2006) and Duong and Kalev (2014). Moreover, the LOB slope appears slightly higher on the bid or buy side than on the ask or sell side.

[19] After winsorization, the entire returns series ($r_{i,t}$) for DJS (David Jones Limited - Mid cap), ENV (Envestra Limited - Mid cap), AQA (Aquila Resources Limited - Small cap), and GFF (Goodman Fielder Limited - Small cap), are identically equal to zero, so that the return volatility estimates defined in subsection 2.2 are also identically zero and the volume-volatility regressions cannot be performed. Thus, we exclude these stocks from our analysis. Note that results for the unwinsorized series qualitatively similar to those in the main text and are available upon request.

The order flow characteristics are also in conformance with the liquidity of stocks. More specifically, more capitalized stocks trade more frequently in a smaller volume than less capitalized stocks, as can be seen from the number of trades and the average trade size during a 5-minute interval. Larger stocks also have a slightly higher quote to trade ratio, which, coupled with higher trading intensity, suggests that bigger stocks attract more attention and more intensive quoting activities from market participants than do smaller stocks, as expected. There are on average more purchases than sales for all stock groups, and the imbalance between buying and selling activities tends to increase with the liquidity level of stocks.

4 Results and discussion

In this section, we empirically examine the role played by the LOB characteristics in explaining the return volatility and the volume-volatility relation of trades. We begin with an investigation of the interaction between the LOB slope, (which summarizes LOB information at all quote levels), and the volume-volatility relation (subsection 4.1). Then, we zoom in on the information content of the LOB at the best bid and ask prices (subsection 4.2). Subsection 4.3 compares the predictive power of the LOB slope and LOB information at the best quotes, and subsection 4.4 provides a series of sensitivity analyses to assess the robustness of our results.

4.1 Order book slope and the volume-volatility relation

This subsection investigates the information content of the LOB across different best quote levels by examining how the LOB slopes, computed according to Equations (3)-(5) using the 10 best quotes, explain return volatility and the volume-volatility relation at a transaction level for stocks in the S&P/ASX200 index during July-December 2014. We estimate Equation (1) for each individual stock in our sample, and then report the median coefficient and the proportions of coefficients that are negatively and positively significant at a 5% level for each of the three stock groups. The results for a combined LOB are reported in Table 2, while those for separate bid and ask sides of the LOB are shown in Table 3.[20]

[20]Note that since the median operator is not additive, "Lag 0" and "$\sum_{1:p}$" median coefficients generally do not add up to that of "$\sum_{0:p}$". The mean coefficients, which preserve additivity but are more prone to outliers, are qualitatively similar to the reported median coefficients, and are available upon request.

Panel A of Table 2 contains an analysis of the volume-volatility relation formulated in Equation (1) which is fully exogenous and contemporaneous, as typically assumed in most previous studies. Consistent with prior findings, we observe a strong positive contemporaneous relation between trading volume and return volatility for all stock groups which is statistically significant at a 5% level for almost all stocks, even after controlling for the effects of the LOB slope and other order flow characteristics. As expected, the return volatility of a trade is negatively related to the slope of the LOB immediately before the trade, with statistical significance obtained for the majority of stocks, which is consistent with the findings of Næs and Skjeltorp (2006), Duong and Kalev (2014), Jain and Jiang (2014) and Tian et al. (2019). A larger order book slope implies steeper LOB curves where more shares are supplied closer to the inner bid/ask quotes. Consequently, the larger the LOB slope, the better the LOB is able to absorb a given amount of shares demanded from an incoming trade, and the smaller the price impact and the return volatility of the trade. The negative relation between return volatility and the LOB slope becomes weaker, in magnitude, for larger cap stocks, reflecting the fact that more liquid stocks have a steeper slope (Næs and Skjeltorp, 2006). Hence, for a given change, e.g. a one unit increase in the LOB slope, the price of more liquid stocks moves less to accommodate a given trading volume, resulting in lower volatility.

<<INSERT TABLE 2 ABOUT HERE>>

We also find strong evidence in support of the predictability of the order flow information about future return volatility at the tick-by-tick level.[21] Consistent with the findings of Xu et al. (2006), Manganelli (2005), Russell and Engle (2005) and Nolte (2008), a shorter trade duration increases the return volatility of the trade. The result lends support to Easley and O'Hara's (1992) theory which implies that shorter time between trades or higher trading intensity is a signal of more private news and a higher fraction of informed traders present in the market. Since the increased presence of informed investors constrains liquidity traders from entering the market, possibly via toxic order flows that adversely select the latter (Easley et al., 2012), trades with shorter durations have larger impacts on prices, leading to higher volatility. The average size of trades that are executed during a 5-minute interval before a trade is positively related to the return volatility of the trade, with statistical significance observed for most stocks. This suggests that

[21] We do not report the estimated coefficients for the order flow characteristics in subsequent analyses to save space but note that the results are qualitatively similar to those being discussed here.

past trading volumes are predictive of future volatility, making the volume-volatility relation both path-dependent and dynamic.

The return volatility of a trade is inversely dependent on the number of trades during a 5-minute interval prior to a trade - a proxy for the trading frequency prior to the trade, which appears to be inconsistent with the findings of most previous studies. This surprising observation can be explained as a result of both measures of trading intensity, namely trade duration ($T_{i,t}$) and the number of trades ($N_{i,t}$), being included in the regression.[22] Results from an unreported experiment in which trade duration is removed from the volatility equation show a positive and significant relation between return volatility and the number of trades. While reaffirming the findings of previous work, this outcome suggests that the most recent information about trading intensity, captured by $T_{i,t}$, appears to be more relevant than and dominate the older and more distant information, proxied by $N_{i,t}$ in explaining future volatility.

In conformance with Chan and Fong (2000), Chan and Fong (2006) and Shahzad et al. (2014), there is a positive link between the order imbalance of trades ($OIB_{i,t}$) and volatility which is statistically significant for a fair proportion of stocks in our sample (more than 36% for all groups), suggesting that trade order imbalance has some predictive power about future return volatility. Meanwhile, there is a strong positive dependence of the volatility of a trade on the quote to trade ratio ($QTT_{i,t}$) which measures the quoting activities during a 5-minute interval before the trade. Quote to trade ratios have increased considerably in today's fast trading environment, as a consequence of the dominance of algorithmic and high frequency traders (HFTs) who utilize their speed advantage to split and submit many orders that are subsequently canceled very quickly (e.g. SEC, 2010, Hasbrouck and Saar, 2013, Conrad et al., 2015, O'Hara, 2015). Our result suggests that HFT activities tend to increase future return volatility, as also noted by Boehmer et al. (2020).

Panel B of Table 2 reports the results relating to the relaxation of the assumption of a constant contemporaneous volume-volatility relation. The results show that the positive dependence of return volatility on the trading volume of a trade is negatively associated with the slope of the LOB right before the trade, with statistical significance observed for more than 77% of all stocks. Thus, the positive volume-volatility relation is neither constant nor exogenous as typically assumed

[22] This result is not due to multicollinearity, since the average correlation between $N_{i,t}$ and $T_{i,t}$ ($\ln(T_{i,t})$) across all stocks in our sample is just -0.21 (-0.29). This correlation, although significant, is well away from 1, and in addition, our estimated values of the $N_{i,t}$ and $T_{i,t}$ coefficients are individually statistically significant.

in most prior studies but it becomes weaker as the LOB becomes steeper and more concentrated around the best quotes. Our result shows that the negative dependence of the volume-volatility relation on the LOB slope prevails strongly at a tick-by-tick level, which complements and supports similar findingb by Næs and Skjeltorp (2006) and Jain and Jiang (2014) using lower-frequency data.

The dependence of the contemporaneous volume-volatility relation on the LOB slope is genuine since the direct impacts of the slope and other order flow attributes on volatility are already controlled for (provided in Table 2(B) below the interaction term). In fact, a comparison of the coefficients on $Slope_{i,t}$ between Panels A and B suggests that allowing the volume-volatility relation to vary with the LOB slope reduces the direct effects of the slope on return volatility, even though the direct effects are still strong and significant for the majority of stocks. It follows from Panel B that the LOB slope right before a trade negatively affects the return volatility of the trade through two channels: a direct channel (captured by the coefficients on $Slope_{i,t}$) and an indirect channel (captured by the coefficients on $v_{i,t}Slope_{i,t}$) that changes the volume-volatility relation. Intuitively, a steeper LOB with more shares allocated around the inner quotes can better absorb the liquidity demand of a future trade, which directly reduces the volatility of the trade. In addition, a steeper LOB also weakens the volume-volatility relation of the trade, lowering the trade's volatility further. Meanwhile, the direct effects of trading volume on return volatility (captured by the coefficients on $v_{i,t}$) remain strongly positive after one allows the volume-volatility relation to depend on the LOB slope, even though the proportions of significant volume coefficients are slightly smaller for medium and small cap stocks.

We now investigate the results for a volume-volatility relation that is dynamically dependent on the LOB slope (see Panel C of Table 2). In support of the theories of Copeland (1976), Jennings et al. (1981), Shalen (1993), and Banerjee and Kremer (2010) and the empirical work of Manganelli (2005), Xu et al. (2006), Nolte (2008), Carlin et al. (2014), and Do et al. (2014), Panel C shows that return volatility is positively correlated with both current and lagged trading volumes, implying that this relation is indeed dynamic. In addition, the positive dependence of return volatility on volume also varies with the dynamics of the LOB slope, with larger order book slopes (i.e. steeper LOBs) weakening the volume-volatility relation. However, most of the effects on the volume-volatility relation are attributable to the LOB slope immediately before a trade (see the "Lag 0" coefficients), while the contribution of the past order book information, albeit of expected sign, is of much smaller magnitude and of much less statistical significance. In contrast, most of the direct impact of the LOB slope on future return volatility comes from the lagged information (see the "$\sum_{1:p}$" coefficients), while the coefficients measuring the direct effects of the slope right before a trade (i.e. the "Lag 0" coefficient estimates) on volatility are usually of opposite and unexpected signs. All else being equal, the cumulative influence, either direct or indirect, of the LOB slope prior to a trade on the return volatility of the trade is smaller in magnitude for more liquid stocks (compare large cap stock with mid and small cap stocks).

The contrasting results between the direct and indirect effects of the LOB slope are interesting and can be explained as follows: In order to predict the return volatility of an incoming trade without knowing the volume of that trade, one needs to make use of all past trading information (including past order book characteristics, past trading volumes and past volatilities) to draw a likely and sensible picture of the relation between the past information, the expected volume of the trade and its future volatility. That is, all past information is exploited to form a prediction of the future transaction's volume, based on which its return volatility is predicted. However, if one knew the volume that the trade would demand, then the information that is currently being supplied right before the execution of the trade contained in the LOB would be more relevant than the past order book information for the determination of how the trade would move prices. As a result, it seems reasonable that the direct impact of the LOB slope on return volatility, which does not take into account the information about the volume of current trade, is mostly contributed by the lagged effects, whereas its indirect effects on volatility, which channel through the volume-volatility relation and incorporate the current volume information, are primarily driven by the current state of the LOB right before the trade.

We now turn to an analysis of the dynamic volume-volatility relation that allows for possible asymmetries between the bid and ask sides of the LOB. The results are reported in Table 3. First, as expected, when an incoming trade is a buy (sell), it is the opposite (i.e. ask (bid)) side of the LOB that is more important for determining the impact that the trade will have on prices. Specifically, for a given trading volume the steeper the slope of the ask (bid) order book immediately before a purchase (sale), the smaller the return volatility (see the coefficients on $AskSlope_{i,t}B_{i,t}$ ($BidSlope_{i,t}S_{i,t}$) in Panels A, B and C) and the weaker the positive dependence of volatility on the volume (see the coefficients on $v_{i,t}AskSlope_{i,t}B_{i,t}$ ($v_{i,t}BidSlope_{i,t}S_{i,t}$) in Panels B and C) of the trade. These results not only complement the corresponding findings presented in Table 2, but are also stronger than the latter in terms of statistical significance, which is generally in conformance with previous studies (e.g. Næs and Skjeltorp, 2006, Brogaard et al., 2015) that show steeper and deeper markets with more shares supplied near the inner quotes support liquidity and mitigate the price impact and return volatility of trades. These results demonstrate the value in separating analyses depending on which side of the LOB we are concerned.

<<INSERT TABLE 3 ABOUT HERE>>

In contrast to the opposite-side LOB slope, we find that a steeper slope of the same side LOB as the direction of a trade tends to strengthen the volume-volatility relation (see the coefficients on $v_{i,t}BidSlope_{i,t}B_{i,t}$ and $v_{i,t}AskSlope_{i,t}S_{i,t}$ in Panels B and C) of that trade, even though the proportions of stocks that have significant coefficients are generally lower. Meanwhile, the direct impact of the same side LOB slope on the volatility of the trade appears ambiguous (see the proportions of significant coefficients on $BidSlope_{i,t}B_{i,t}$ and $AskSlope_{i,t}S_{i,t}$ in Panels A and B). These results highlight the asymmetric effects between the same-side and opposite-side LOB slopes on return volatility and the volume-volatility relation.

Panel C provides strong evidence of the predictability of the dynamics of the bid and ask slopes for the dynamic volume-volatility relation, with most of the predictive power contributed by the current slope information right before the execution of a trade, which is consistent with the results from a combined LOB presented in Table 2. While the negative indirect impact of the bid (ask) slope on the future return volatility of a sell (buy), which is channeled through the dynamic volume-volatility relation, remains economically and statistically significant for the vast majority of stocks (see the coefficients on $v_{i,t}AskSlope_{i,t}B_{i,t}$ ($v_{i,t}BidSlope_{i,t}S_{i,t}$) in Panel C), its direct effects on volatility, albeit of expected negative signs, are of much less statistical significance (compare the coefficients on $AskSlope_{i,t}B_{i,t}$ ($BidSlope_{i,t}S_{i,t}$) in Panel C of Table 3 with the corresponding results for $Slope_{i,t}$ in Panel C of Table 2).

Table 3 indicates that the order type and direction of trade contain useful information about the return volatility of the trade, which is in agreement with previous studies that find strong evidence supportive of the important role played by trading or quoting directions in explaining the price formation process (e.g. Hasbrouck, 1991a,b, Dufour and Engle, 2000, Barclay et al., 2003).

We observe some asymmetries in the effects of the bid versus ask LOB slopes on return volatility as well as on the volume-volatility relation. Incorporating the trade direction information and these asymmetries between the bid and ask sides of the LOB significantly improves the in-sample fit of the volatility regressions, with the adjusted R^2 measures increasing by 7-8 percentage points (or about 40-45%), relative to those for a combined LOB in Table 2.

Overall, the results presented in Tables 2 and 3 show that the dynamics of the LOB information summarized by the LOB slope are informative about the volume-volatility relation and ultimately the return volatility of trades. The positive dependence of volatility on volume is dynamic, path dependent and negatively related to the LOB slope. A larger same-side (opposite-side) slope prior to a trade increases (decreases) the volume-volatility relation of the trade. The effects of market depth on return volatility and the volume-volatility relation are asymmetric between the bid and ask sides of the LOB and tend to be smaller in magnitude for large cap stocks.

4.2 Spread, depth and the volume-volatility relation

The previous subsection examines the informativeness of the LOB information across multiple price levels about return volatility and the volume-volatility relation. In this subsection, we focus on the LOB information at the inner quotes and investigate how the bid-ask spread and market depths at best bid and ask quotes explain return volatility and the volume-volatility relation.

The results for an investigation of the power of the LOB characteristics at the inner quotes for explaining the tick-by-tick volume-volatility relation for a combined LOB are presented in Table 4. Panel A reports the coefficient estimates of a volatility regression under the assumption that the volume-volatility relation is exogenous and fully contemporaneous. Consistent with the findings discussed in the previous subsection and in prior literature, there is a strong positive contemporaneous dependence of return volatility on trading volume that is statistically significant for almost all stocks. The best level (i.e. lowest ask and highest bid) of the LOB has predictive power about future return volatility in that the wider the bid-ask spread prior to a trade, the larger the volatility (per unit of time) of the trade. This positive relation between spread and volatility is significant at a 5% level for the majority of stocks in three groups and is consistent with the theoretical model of Foucault et al. (2007) and the empirical findings of Hasbrouck (1991a), Næs and Skjeltorp (2006), Foucault et al. (2007), Nolte (2008), Pascual and Veredas (2010), Haugom et al. (2014), and Jain and Jiang (2014). Meanwhile, the return volatility of a trade is negatively dependent on the prevailing quoted depth right before the trade. This result is intuitive because larger market depths available at the best bid and ask prices are better able to accommodate a trade of a given size, resulting in fewer quote revisions and consequently lower price impact and volatility of the trade (e.g. Ahn et al., 2001, Jain and Jiang, 2014, Brogaard et al., 2015, Pham et al., 2020).

<<INSERT TABLE 4 ABOUT HERE>>

We now examine the results for the contemporaneous volume-volatility relation that is allowed to vary with the LOB information at the inner quotes (see Panel B of Table 4). Consistent with the theory of Foucault et al. (2007), there is strong evidence that the contemporaneous volume-volatility relation is significantly related to the LOB characteristics at the best level. In particular, the coefficients on $v_{i,t}Spread_{i,t}$ show that the positive dependence of return volatility on the trading volume of a trade becomes stronger, the larger the bid-ask spread right before the trade. In contrast, the $v_{i,t}Depth_{i,t}$ coefficients show that larger supplies of shares at the best bid/ask price weaken the volume-volatility relation. Both results are statistically significant at the 5% level for the majority of stocks (more than 68% and 79% respectively), and they are stronger, in terms of the magnitude of the coefficients, for stocks with higher capitalization.

Similar to the results in Table 2, allowing the volume-volatility relation to be endogenously related to the bid-ask spread and market depth weakens the direct effects of these order book characteristics on return volatility (see the coefficients on $Spread_{i,t}$ and $Depth_{i,t}$ in Panel B in comparison to those in Panel A), even though the direct effects are still strong and significant for a big proportion of stocks, especially with regard to the bid-ask spread.[23] This finding implies that a more liquid order book market (which has deeper depths and/or narrower bid-ask spreads) reduces trading volatility via two channels: by its direct impact on volatility and by its indirect

[23]The change of sign of the coefficient on market depth $Depth_{i,t}$ for the large cap stocks from negative (-0.994) in Panel A to positive (+1.205) in Panel B appears counterintuitive, but it could be explained by the dominance of the positive direct impact of same-side depth over the negative direct impact of opposite-side depth on return volatility, as will be shown in Table 5(B). This result suggests that the indirect effects of market depth on return volatility, which are transmitted through the volume-volatility relation, outweigh its direct impact, so overall the return volatility of a trade in a large cap stock is negatively related the market depth prior to the trade, as seen in Table 4(A).

effect that is transmitted through and weakens the volume-volatility relation. Unlike the results shown in Panel B of Table 2, the direct impact of trading volume on volatility, captured by β_0 in Equation (1), changes sign from positive to negative for large and mid cap stocks and becomes much less significant for small stocks following the relaxation of the constant volume-volatility relation (comparing the coefficients on $v_{i,t}$ in Panels A and B). This result suggests that the well-documented positive association between trading volume and return volatility seems to be driven by the LOB characteristics.

We now allow the volume-volatility relation to be dynamically dependent on the LOB information at the best level. The results in Panel C of Table 4 indicate the dynamic nature of the volume-volatility relation, which is also related to the dynamics of the bid-ask spread and market depth. Larger bid-ask spreads and smaller depths available at the best quotes are associated with a stronger positive dependence of return volatility not only on current trading volumes, as discussed in Panel B, but also on lagged volumes. Similar to the results in Table 2, while most of the direct impact of the bid-ask spread and market depth on return volatility comes from the lagged LOB information (see the coefficients on $Spread_{i,t}$ and $Depth_{i,t}$), their negative indirect impact on future return volatility is primarily contributed by the current LOB characteristics that prevail right before a trade, as demonstrated by the coefficients on $v_{i,t}Spread_{i,t}$ and $v_{i,t}Depth_{i,t}$. This result is in conformance with the findings of Pham et al. (2020), who show that a comparison of the volume of a trade with the prevailing market depth information right before the trade is of particular relevance for identifying whether the trade results in any immediate impact on prices. These authors show that the incorporation of the depth information into an immediate price impact model significantly enhances the forecast accuracy of the model. Unlike the LOB slope, both direct and indirect effects of the bid-ask spread and market depth on the volatility of an incoming trade increase, in magnitude, with a stock's liquidity (as proxied by stock market capitalization).

Table 5 reports the results of an investigation in which we allow for possible asymmetries between the LOB in explaining return volatility and the volume-volatility relation. Similar to the results in Table 3, it is the opposite side of the LOB that is more predictive of the future return volatility of a trade. In particular, for an incoming purchase (sale) of a given volume, the larger the amount of shares available at the best ask (bid) quote immediately before that purchase (sale), the smaller the return volatility (see the coefficients on $AV_{i,t}B_{i,t}$ ($BV_{i,t}S_{i,t}$) in Panels A, B and C) and the weaker the positive dependence of volatility on the volume (see the coefficients on $v_{i,t}AV_{i,t}B_{i,t}$ ($v_{i,t}BV_{i,t}S_{i,t}$) in Panels B and C) of the trade. In contrast, both return volatility and the volume-volatility relation of a trade is positively related to the market depth available on the same side of the LOB as the direction of a trade, as demonstrated by the coefficients on $BV_{i,t}B_{i,t}$ and $AV_{i,t}S_{i,t}$ in Panels A and B, as well as those for $v_{i,t}BV_{i,t}B_{i,t}$ and $v_{i,t}AV_{i,t}S_{i,t}$ in Panels B and C. These results are consistent with the findings in the prior literature on order aggressiveness (e.g. Biais et al., 1995, Ranaldo, 2004, Aitken et al., 2007, Duong and Kalev, 2013) in that investors tend to submit more (less) aggressive orders when the same-side (opposite-side) market depth increases since the non-execution risk of an incoming limit order is higher (lower). As more aggressive orders typically have a larger impact on prices (Biais et al., 1995, Duong and Kalev, 2013, Brogaard et al., 2019), it follows that larger same-side (opposite-side) market depth increases (decreases) the future trading volatility.

<<INSERT TABLE 5 ABOUT HERE>>

In conformance with the results from the previous tables, the direct effects of the bid and ask depths on return volatility become weaker once one allows for the endogeneity of the volume-volatility relation. Panel C of Table 5 shows that the dynamics of the bid and ask depths play a significant role in explaining future return volatility. While the direct impact of the bid and ask depths on return volatility mostly comes from their lagged information, their current information right before the execution of a trade is the main driver of the volume-volatility relation which constitutes their indirect impact on volatility. Furthermore, the higher the stock capitalization, the bigger the cumulative impact of the best bid and ask depths, either direct or indirect, on return volatility. Similar to the earlier comparison of Table 2 and Table 3 results, Table 5 shows that consideration of bid/ask and direction of trade asymmetries enhances the in-sample explanatory power of the volatility regressions relative to those in Table 4, leading to an average increase of about 3-4 percentage points (or about 20-25%) in the adjusted R^2.

4.3 Spread, depth, slope and the volume-volatility relation

Subsections 4.1 and 4.2 demonstrate the influence of LOB dynamics on return volatility and the volume-volatility relation, with subsection 4.1 examining the explanatory power of the LOB slope (which uses quotes at many levels), and subsection 4.2 focussing on the bid-ask spread and market depth (based on best level quotes). In this subsection, we investigate which portions of order book information play a more important role in explaining the positive dependence of return volatility on trading volume and when. The results of this analysis are reported in Table 6 for a combined LOB, and in Table 7 where we allow for the separation of the bid and ask order books.[24]

<<INSERT TABLES 6 & 7 ABOUT HERE>>

In conformance with the results discussed in previous subsections, the return volatility of a trade is positively related to both contemporaneous and lagged trading volumes while being negatively dependent on the LOB slope prior to the trade, with statistical significance observed for the majority of stocks. In addition, the dynamic volume-volatility relation is not constant but varies inversely with the dynamics of the LOB slope, which is in agreement with Næs and Skjeltorp (2006) and Duong and Kalev (2014). The more concentrated the (opposite-side) order book is around the inner quotes or the larger the (opposite-side) book slope prior to an incoming trade, the more able the market is to absorb the trade. Consequently, there are fewer price revisions, resulting in lower return volatility and a weaker volume-volatility relation. Since more highly capitalized stocks typically have a steeper LOB (see Table 1), it follows that the effects of the LOB slope on return volatility, either direct or indirect, should decrease with stocks' liquidity, which is indeed what we observe (when we compare mid and small cap stocks with large cap stocks).

After controlling for the LOB slope, both direct and indirect effects of the bid-ask spread and market depth on return volatility and the volume-volatility relation either switch signs or become much less significant, as compared with the corresponding results previously reported in Tables 4 and 5. This observation can be explained by the fact that the LOB slope, by definition, encompasses the information about the LOB both at and beyond the best quote level. Since the LOB outside the inner quotes is informative about future returns and volatility (Ahn et al., 2001, Kalay et al., 2004, Næs and Skjeltorp, 2006, Kalay and Wohl, 2009, Pascual and Veredas, 2010, Duong and Kalev, 2014, Jain and Jiang, 2014, Tian et al., 2019), the information contained in the LOB slope appears to dominate the bid-ask spread and the market depth in explaining the

[24]For brevity, we only report in Table 7 the estimates of the LOB attributes that are of the opposite side to the direction of a trade (e.g. $AskSlope_{i,t}B_{i,t}$) from the volatility equation. The results for the LOB characteristics that are of the same side as the direction of a trade (e.g. $BidSlope_{i,t}B_{i,t}$) are of less interest and are often less statistically significant. A complete table of results is available upon request.

return volatility and the volume-volatility relation of trades. In fact, this result is in harmony with work by Næs and Skjeltorp (2006), who show that the contemporaneous correlation between daily volatility and the number of trades within a day becomes negatively (positively) related to the bid-ask spread (total depth in the LOB) after the LOB slope is taken into account. It is also consistent with Pascual and Veredas (2010), who find that the ex-post informational volatility of the latent efficient price process is positively (negatively) dependent on the depth available at (beyond) the best quotes, especially when one realizes that the information of the depth beyond the best quotes is incorporated in the LOB slope.

Overall, the results in subsections 4.1-4.3 highlight the dynamic nature of the volume-volatility relation which is positive and varies with the dynamics of the LOB. The dependence of return volatility on the trading volume of a trade is positively associated with the bid-ask spread but negatively correlated with the market depth at the best quotes and the slope of the LOB prior to the transaction. Since the LOB slope, by definition, captures the information contained in the bid-ask spread and the market depth at the best quotes, it acts as the dominant explanatory factor of the volume-volatility relation and the return volatility of a trade. The impact of the LOB characteristics on the future return volatility of a trade depends on the liquidity of stocks and is transmitted through two channels: a direct channel that is mainly contributed by the lagged order book information, and an indirect channel that transfers the effects via the volume-volatility relation and is primarily driven by current order book information that prevails immediately before the trade. Table 7 shows that there are also asymmetries between the influence of the bid and ask order books on return volatility and the volume-volatility relation, with the opposite-side order book possessing the dominant predictive power about the return volatility of an incoming trade.

4.4 Robustness

We use the LOB slopes that are calculated using 10 best bid and ask levels from the LOB. An interesting and natural question is whether the slopes become more or less informative about the return volatility and the volume-volatility relation of trades if they are computed from different sets of the LOB information. To answer this question, we employ different bid and ask levels (5 and 20) from the LOB to calculate the slope measures, and then reexamine our analysis. The results of this exercise are reported in Table A.2 in the Appendix for a combined LOB, and in Table A.3 for an order book that is separated into bid and ask sides.[25]

Overall, the results from Tables A.2 and A.3 are qualitatively similar to those reported in Tables 6 and 7, respectively, with both the dynamic volume-volatility relation and return volatility strongly negatively associated with the dynamics of the slopes of the (opposite-side) order book. The indirect effects of the (opposite-side) order book slope on return volatility, which are transmitted through the volume-volatility relation, tend to decrease with stocks' liquidity (when comparing mid and small cap stocks with large cap stocks) and mainly stem from the slope information that is available right before a trade. The direct effects of the LOB slope on return volatility are also inversely related to stocks' liquidity; however, they are of less statistical significance than the indirect effects and are mainly explained by lagged slope information (see Table A.2). These direct effects even play a much smaller statistical role than the corresponding indirect effects when one allows for potential asymmetries between the bid and ask order books (see Table A.3).

The LOB slope dominates the bid-ask spread and the market depth in explaining the return volatility and the volume-volatility relation of a trade. Nevertheless, the market depth at the best quotes possesses significant predictive power for volatility, especially for mid and small cap stocks, when 20 best quote levels are used to construct the LOB slope (see the coefficients for $v_{i,t}Depth_{i,t}$ in Panel B of Table A.2, and those for $v_{i,t}BV_{i,t}S_{i,t}$ and $v_{i,t}AV_{i,t}B_{i,t}$ in Panel B of Table A.3).

There is, however, an interesting observation that is worth highlighting. While the impact of the LOB slope on the volume-volatility relation, or equivalently the indirect influence of the slope on return volatility, becomes stronger (in magnitude) for large cap stocks when more order book information is employed to construct the slope measure, this impact is biggest for mid and small cap stocks when the LOB slope is computed using 10 best bid and ask levels. In addition, the proportions of significant coefficients for the mid and small cap stocks are also remarkably lower for the slope measure computed using the 20 best quotes from the book. These results suggest that for almost all stocks, the sixth to tenth best levels of the LOB possess significant predictive

[25] To save space, we only tabulate the results for the case where the volume-volatility relation is allowed to be dynamic and endogenously related to the bid-ask spread, the market depth at the best quotes, and the slope of the LOB right before a trade. In addition, we only report the estimated coefficients for the LOB attributes and their interactions with trading volume $v_{i,t}$ in the volatility equation. In Table A.3, only the coefficients for the attributes of the opposite side to the trade direction (e.g. $v_{i,t}AskSlope_{i,t}B_{i,t}$) are reported. The estimates for other variables are of less interest and are qualitatively similar to the corresponding ones reported in the main text. A complete table of results is available upon request.

power about future return volatility in addition to that contained in the first five best quotes. Quotes and depths that are queued beyond the tenth best level (and up to the twentieth best level) are informative about volatility only for highly liquid stocks but not for less liquid stocks. The reason for this is that for illiquid stocks, quotes outside the 10 best levels are likely stale orders. Consequently, the inclusion of these levels in the slope calculation reduces the relevance of the LOB slope measure for less liquid stocks, which possibly explains the observed improvements in the ability of the depth (at best quotes) measure to predict volatility and the volume-volatility relation for these stocks.

The second set of robustness checks the sensitivity of the informativeness of the LOB information to different winsorization cut-off levels. In order to avoid the effect of the outliers, in previous analyses all variables are winsorized, on a stock-by-stock basis, at the 1^{st} and 99^{th} quantiles (i.e. 2% winsorization). We now redo our analyses (with the LOB slope constructed from the 10 best bid and ask levels) adopting two different winsorization cut-off levels, namely the 0.5^{th}-99.5^{th} quantiles (i.e. 1% winsorization) and the 2^{nd}-98^{th} quantiles (i.e. 4% winsorization).[26] The results of this investigation, respectively reported in Tables A.4 and A.5 in the Appendix, again qualitatively resemble those reported in Tables 6 and 7, suggesting that our main finding that the slope of the (opposite-side) LOB is an important determinant of the dynamic volume-volatility relation is robust to different winsorization levels. It is, however, noted that this main finding generally becomes more (less) statistically significant when the 4% (1%) winsorization window is employed.

5 Why is the order book slope informative?

The previous analyses highlight the significant information content of the LOB slope about the return volatility and the volume-volatility relation of trades. In order to interpret the informativeness of the order book slope given a lack of theoretical guidance, Næs and Skjeltorp (2006)

[26]In addition to the two stocks discussed in footnote 15, another stock GFF (Goodman Fielder Limited - Small cap) is removed after the 0.5^{th}-99.5^{th} winsorization for the reason explained in footnote 19, leaving us with a sample of 197 stocks (49 Large cap, 50 Mid cap, and 98 Small cap). Meanwhile, additional eleven stocks, namely TLS (Telstra Corporation Limited - Large cap), ALZ (Australand Property Group - Mid cap), DUE (DUET Group - Mid cap), DJS (David Jones Limited - Mid cap), ENV (Envestra Limited - Mid cap), AQA (Aquila Resources Limited - Small cap), CMW (Cromwell Property Group - Small cap), GFF (Goodman Fielder Limited - Small cap), HZN (Horizon Oil Limited - Small cap), SIP (Sigma Pharmaceuticals Limited - Small cap), and TEN (Ten Network Holdings Limited - Small cap), are removed after the 2^{nd}-98^{th} winsorization, resulting in a sample of 187 stocks (48 Large cap, 46 Mid cap, and 93 Small cap).

conduct an empirical analysis to identify factors that can explain the slope. Based on an empirical observation that there is a significant negative relation between the average monthly LOB slope and the variation in the analysts' monthly earnings forecasts, these authors suggest that the LOB slope acts as a proxy for disagreements amongst investors. The more traders disagree about the true value of a stock, the wider the range of prices and volumes of the limit or market orders that they will submit, resulting in a less concentrated LOB with a more gentle slope. This conjecture of Næs and Skjeltorp (2006) seems to fit in with a strand of theoretical studies that demonstrate that disagreements amongst investors about asset values are a key factor contributing to the positive correlation between trading volumes and absolute price changes. These disagreements may result from either private information asymmetry (e.g. Grundy and McNichols, 1989, Shalen, 1993) or differences of opinions about public information (e.g. Harris and Raviv, 1993, Kandel and Pearson, 1995, Banerjee and Kremer, 2010). An empirical study of Carlin et al. (2014) also finds that both trading volume and return volatility become larger following an increase in investors' disagreement. Similarly, Wang and Wu (2015) document that the contemporaneous impact of the number of trades on price volatility varies across different corporate bond groups that are classified according to the dispersion of analysts' earnings forecasts, and it is typically larger for bonds that have higher analyst disagreement. Since investor heterogeneity is a driver of the positive dependence of volatility on volume, the informativeness of the LOB slope about the volume-volatility relation can be reasonably explained if the slope is indeed a proxy for the heterogeneity of investors as suggested by Næs and Skjeltorp (2006).

In this study, we do not aim to empirically test the above Næs and Skjeltorp's (2006) conjecture, which is connected to the theoretical prediction of Harris and Raviv (1993), Shalen (1993), Banerjee and Kremer (2010), amongst others. Instead, we provide an intuitive graphical illustration that not only directly explains why return volatility and the volume-volatility relation are negatively associated with the slope of the (opposite-side) LOB, but also complements Næs and Skjeltorp's (2006) conjecture.

Consider a market order submitted to the LOB of a hypothetical stock A that immediately results in a trade. Suppose that right before the execution of the market order, the LOB of stock A has the best bid quote of $P_0 - s/2$ and the best ask quote of $P_0 + s/2$, implying that the prevailing mid quote is P_0 and the quoted bid-ask spread is s. To obtain a clearer and simplified picture of how the slope of the LOB that prevails immediately prior to the trade affects the price at which the trade is transacted and the volume-volatility relation of the trade, we assume that (i) the market order that leads to the trade is buyer-initiated so that the ask side of the LOB is relevant for the execution of the order; (ii) the depths queued on the ask order book right before the trade are nicely allocated such that the ask book can be smoothly illustrated by an increasing straight line starting from the best ask;[27] (iii) the last transacted price of stock A is P and it is no greater than the prevailing best ask quote $P_0 + s/2$; (iv) the size of the market order, V_{buy}, is larger than the depth available at the best ask; and (v) the market order is very aggressive such that it walks up the LOB and is fully executed.

<<INSERT FIGURE 1 ABOUT HERE>>

Figure 1 illustrates how the price of stock A adjusts to accommodate the market buy order or the purchase. We consider two scenarios. The first is one for which the ask order book of stock A right before the execution of the purchase has less shares queued close to the best ask quote and hence is relatively flat. The ask order book in this scenario is illustrated with a dashed black line labeled as "Ask order book 1", with Ask Depth_1 shares available at the best ask. The second scenario is one for which the ask order book prior to the trade is more concentrated around the best ask and has a bigger slope, which has the best ask depth of Ask Depth_2 ($>$ Ask Depth_1) and is presented by a solid black line with an "Ask order book 2" label. Note that the "Ask order book 1" ("Ask order book 2") can also be viewed as the state of the LOB for stock A when there is a high (low) degree of disagreement amongst traders whose orders are placed over a wide (narrow) range of prices. Thus, the two scenarios under consideration here are compatible with Næs and Skjeltorp's (2006) suggestion.

From Figure 1, the execution of the buy of size V_{buy} moves the price of stock A to P_1 (P_2) under the first (second) scenario from the previous transaction price P. Clearly, the absolute change in the price of stock A, which is a proxy for volatility that is widely used in the literature, is smaller in the second scenario where the slope of the ask order book that prevails right before the purchase is larger (i.e. $|P_2 - P| < |P_1 - P|$), and this explains the negative correlation between volatility and the LOB slope.

[27] Strictly speaking, the limit order book has a non-decreasing piece-wise linear shape.

To see how the LOB slope affects the volume-volatility relation, consider a hypothetical increase of ΔV_{buy} in the volume of the purchase from V_{buy} to V'_{buy}. This pushes the price of stock A further to P'_1 (P'_2) under the first (second) scenario, implying an increase in the price of ΔP_1 (ΔP_2), relative to the previous price when the size of the purchase is V_{buy}. The impact of the increase in the buying volume on the stock price, which is essentially a measure of the volume-volatility relation, is $\Delta P_1/\Delta V_{\mathrm{buy}}$ ($\Delta P_2/\Delta V_{\mathrm{buy}}$) in the first (second) scenario. Since $\Delta P_1/\Delta V_{\mathrm{buy}} > \Delta P_2/\Delta V_{\mathrm{buy}}$, it follows that the volume-volatility relation becomes weaker the larger the LOB slope is prior to the trade.

The main intuition underlying the informativeness of the LOB slope discussed above still holds without the aforementioned simplifying assumptions. Assumption (i) is imposed without a loss of generality so that we only need to focus on the relevant side of the order book. If the market order is a sale, a qualitatively similar graph based on the bid order book can be employed. Assumption (ii) is also trivial and is added to assist the drawing of the graph. It is easy to verify that the above argument from Figure 1 remains valid if we use the commonly observed piece-wise linear limit order book instead. Assumption (iii) is also not an unreasonable assumption given that the majority of trades are executed at the inner quotes, since trading volume is often much smaller than the quoted market depth at the best level (see Table 1). If the previous transaction before the purchase of size V_{buy} illustrated in Figure 1 was a sale, it was certainly executed against the bid order book at a price less than the best ask price. If the last transaction was a purchase, it was very likely transacted either at $P_0 + s/2$ (when either (a) the best ask prior to that transaction was also $P_0 + s/2$ and the trading volume was less than the depth at $P_0 + s/2$; or (b) the best ask prior to the transaction was less $P_0 + s/2$ and the trading volume was less than the cumulative ask depths up to $P_0 + s/2$), or at one tick lower (when the best ask prior to the transaction was one tick lower than $P_0 + s/2$ and the trading volume was exactly equal to the ask depth). However, there is a possibility that the last transacted price might be larger than $P_0 + s/2$, which happens if after the last transaction, there were submissions of limit sell orders that pushed the best ask back to $P_0 + s/2$. Even in this case, the main idea from Figure 1 still holds in general.

Unlike the first three assumptions, assumption (iv) is quite strong and often unrealistic. It is imposed to facilitate the delivery of the main intuition, but it can be relaxed. So far, we have treated the trading volume V_{buy} of the purchase as known and given, but it should be a random variable whose value depends on an investor's liquidity needs and/or information and belief set. If $V_{\text{buy}} \leq \text{Ask Depth}_1 < \text{Ask Depth}_2$, the market buy order will be executed at the best ask $P_0 + s/2$ under both scenarios, suggesting that the volatility of the trade will be the same for both situations. However, as $\text{Ask Depth}_2 > \text{Ask Depth}_1$, the probability of V_{buy} being larger than Ask Depth_1 is higher than the probability that V_{buy} is bigger than Ask Depth_2, all else being equal, which implies that the main idea from Figure 1 remains true in a probabilistic sense and so too on average. Similarly, assumption (v) is rather strong since large aggressive market orders are less often seen in empirical data (e.g. Griffiths et al., 2000, Ranaldo, 2004, Duong and Kalev, 2013). However, if the probability of such large aggressive orders is the same in both scenarios (i.e. for both ask order books 1 and 2), then the main intuition from Figure 1 still holds on average without assumption (v).

In summary, we present in Figure 1 as a graphical rationale for the negative dependence of the return volatility and the volume-volatility relation of a trade on the prevailing LOB slope right before the trade. This negative dependence is empirically found in the current study and previous research (see, amongst others, Næs and Skjeltorp, 2006, Duong and Kalev, 2014, Jain and Jiang, 2014, Tian et al., 2019), and it reaffirms the informativeness of the LOB information about the price formation process.

6 Conclusions

This research extends prior literature on the volume-volatility relation by highlighting the significant information content of the LOB about the return volatility and the volume-volatility relation of individual trades. While most existing studies in the literature assume a constant and fully contemporaneous volume-volatility relation, we find strong evidence that the positive dependence of return volatility on the trading volume of a trade is dynamic. In addition, the volume-volatility relation is positively correlated with the bid-ask spread but negatively related to the market depth at the best quotes and the LOB slope prior to the transaction. The dynamics of the LOB characteristics also play a significant role in explaining future return volatility. While their direct impact on volatility is primarily contributed by their lagged information, it is their current information right before a trade that drives the volume-volatility relation, which captures their indirect impact on volatility.

We find significant asymmetries between the effects of the bid and ask order books on return volatility and on the volume-volatility relation, with the LOB of the opposite side to the direction of an incoming trade being particularly informative about the return volatility of the trade. The LOB slope plays the dominant role in explaining return volatility and the volume-volatility relation even when we incorporate the information from the bid-ask spread and the market depth at the best quotes. We justify our finding that return volatility and the volume-volatility relation are negatively associated with the LOB slope with a simple intuitive graphical illustration, which is compatible with prior explanations of the informativeness of the order book slope in the literature.

References

Ahn, H., Bae, K., and Chan, K. (2001). Limit orders, depth, and volatility : Evidence from the stock exchange of Hong Kong. *The Journal of Finance*, 56(2):767–788.

Aitken, M., Almeida, N., deB. Harris, F. H., and McInish, T. H. (2007). Liquidity supply in electronic markets. *Journal of Financial Markets*, 10(2):144 – 168.

Anand, A., Chakravarty, S., and Martell, T. (2005). Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. *Journal of Financial Markets*, 8(3):288–308.

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

ASX (2015). Review of competition in clearing australian cash equities. *ASX Submission*, available at https://treasury.gov.au/sites/default/files/2019-03/C2015-007_ASX.pdf.

Avramov, D., Chordia, T., and Goyal, A. (2006). The impact of trades on daily volatility. *Review of Financial Studies*, 19(4):1241–1277.

Banerjee, S. and Kremer, I. (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance*, 65(4):1269–1302.

Barclay, M. J., Hendershott, T., and McCormick, D. T. (2003). Competition among trading venues: Information and trading on electronic communications networks. *The Journal of Finance*, 58(6):2637–2666.

Benos, E. and Sagade, S. (2016). Price discovery and the cross-section of high-frequency trading. *Journal of Financial Markets*, 30:54–77.

Biais, B., Hillion, P., and Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse. *The Journal of Finance*, 50(5):1655–1689.

Bloomfield, R., O'Hara, M., and Saar, G. (2005). The "make or take" decision in an electronic market: Evidence on the evolution of liquidity. *Journal of Financial Economics*, 75(1):165–199.

Boehmer, E., Fong, K. Y., and Wu, J. (2020). Algorithmic trading and market quality: International evidence. *Journal of Financial and Quantitative Analysis*. Forthcoming.

Bollerslev, T., Li, J., and Xue, Y. (2018). Volume, volatility, and public news announcements. *Review of Economic Studies*, 85(4):2005–2041.

Brogaard, J., Hagströmer, B., Norden, L., and Riordan, R. (2015). Trading fast and slow: Colocation and liquidity. *Review of Financial Studies*, 28(12):3407–3443.

Brogaard, J., Hendershott, T., and Riordan, R. (2019). Price discovery without trading: Evidence from limit orders. *The Journal of Finance*, 74(4):1621–1658.

Cao, C., Hansch, O., and Wang, X. (2009). The information content of an open limit-order book. *Journal of Futures Markets*, 29(1):16–41.

Carlin, B. I., Longstaff, F. A., and Matoba, K. (2014). Disagreement and asset prices. *Journal of Financial Economics*, 114(2):226–238.

Cenesizoglu, T., Dionne, G., and Zhou, X. (2016). Asymmetric effects of the limit order book on price dynamics. Available at SSRN: https://ssrn.com/abstract=2878945.

Chaboud, A., Hjalmarsson, E., and Zikes, F. (2021). The evolution of price discovery in an electronic market. *Journal of Banking & Finance*, 130:106171.

Chan, C. C. and Fong, W. M. (2006). Realized volatility and transactions. *Journal of Banking & Finance*, 30(7):2063–2085.

Chan, K. and Fong, W.-M. (2000). Trade size, order imbalance, and the volatility–volume relation. *Journal of Financial Economics*, 57(2):247–273.

Chevallier, J. and Sévi, B. (2012). On the volatility–volume relationship in energy futures markets using intraday data. *Energy Economics*, 34(6):1896–1909.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155.

Clements, A. E. and Todorova, N. (2016). Information flow, trading activity and commodity futures volatility. *Journal of Futures Markets*, 36(1):88–104.

Conrad, J., Wahal, S., and Xiang, J. (2015). High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics*, 116(2):271 – 291.

Copeland, T. E. (1976). A model of asset trading under the assumption of sequential information arrival. *The Journal of Finance*, 31(4):1149–1168.

Diamond, D. W. and Verrecchia, R. E. (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics*, 18(2):277–311.

Do, H. X., Brooks, R., Treepongkaruna, S., and Wu, E. (2014). How does trading volume affect financial return distributions? *International Review of Financial Analysis*, 35:190–206.

Downing, C. and Zhang, F. (2004). Trading activity and price volatility in the municipal bond market. *The Journal of Finance*, 59(2):899–931.

Dufour, A. and Engle, R. F. (2000). Time and the price impact of a trade. *The Journal of Finance*, 55(6):2467–2498.

Duong, H. N. and Kalev, P. S. (2013). Anonymity and order submissions. *Pacific-Basin Finance Journal*, 25:101–118.

Duong, H. N. and Kalev, P. S. (2014). Anonymity and the information content of the limit order book. *Journal of International Financial Markets, Institutions and Money*, 30:205–219.

Duong, H. N., Kalev, P. S., and Tian, X. (2018). Short selling, trading activity and volatility in corporate bond market. Available at SSRN: https://ssrn.com/abstract=3234262.

Easley, D., López de Prado, M. M., and O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies*, 25(5):1457–1493.

Easley, D. and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19(1):69–90.

Easley, D. and O'Hara, M. (1992). Time and the process of security price adjustment. *The Journal of Finance*, 47(2):577–605.

Engle, R. F. (2000). The econometrics of ultra-high frequency data. *Econometrica*, 68(1):1–22.

Engle, R. F. and Patton, A. J. (2004). Impacts of trades in an error-correction model of quote prices. *Journal of Financial Markets*, 7(1):1–25.

Foucault, T., Kadan, O., and Kandel, E. (2005). Limit order book as a market for liquidity. *Review of Financial Studies*, 18(4):1171–1217.

Foucault, T., Moinas, S., and Theissen, E. (2007). Does anonymity matter in electronic limit order markets? *Review of Financial Studies*, 20(5):1707–1747.

Giot, P., Laurent, S., and Petitjean, M. (2010). Trading activity, realized volatility and jumps. *Journal of Empirical Finance*, 17(1):168–175.

Glosten, L. R. (1994). Is the electronic open limit order book inevitable? *The Journal of Finance*, 49(4):1127–1161.

Goettler, R. L., Parlour, C. A., and Rajan, U. (2009). Informed traders and limit order markets. *Journal of Financial Economics*, 93(1):67–87.

Griffiths, M. D., Smith, B. F., Turnbull, D. A. S., and White, R. W. (2000). The costs and determinants of order aggressiveness. *Journal of Financial Economics*, 56(1):65–88.

Grundy, B. D. and McNichols, M. (1989). Trade and the revelation of information through prices and direct disclosure. *Review of Financial Studies*, 2(4):495–526.

Harris, L. (1987). Transaction data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 22(2):127–141.

Harris, L. E. and Panchapagesan, V. (2005). The information content of the limit order book: Evidence from NYSE specialist trading decisions. *Journal of Financial Markets*, 8(1):25–67.

Harris, M. and Raviv, A. (1993). Differences of opinion make a horse race. *Review of Financial Studies*, 6(3):473–506.

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

Hasbrouck, J. (1991b). The summary informativeness of stock trades: An econometric analysis. *Review of Financial Studies*, 4(3):571–595.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*, 50(4):1175–1199.

Hasbrouck, J. and Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, 16(4):646 – 679.

Haugom, E., Langeland, H., Molnár, P., and Westgaard, S. (2014). Forecasting volatility of the U.S. oil market. *Journal of Banking & Finance*, 47:1–14.

Holden, C. W. and Subrahmanyam, A. (1992). Long-lived private information and imperfect competition. *The Journal of Finance*, 47(1):247–270.

Jain, P. and Jiang, C. (2014). Predicting future price volatility: Empirical evidence from an emerging limit order market. *Pacific-Basin Finance Journal*, 27:72–93.

Jennings, R. H., Starks, L. T., and Fellingham, J. C. (1981). An equilibrium model of asset trading with sequential information arrival. *The Journal of Finance*, 36(1):143–161.

Jondeau, E., Lahaye, J., and Rockinger, M. (2015). Estimating the price impact of trades in a high-frequency microstructure model with jumps. *Journal of Banking & Finance*, 61:S205–S224.

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

Kalay, A., Sade, O., and Wohl, A. (2004). Measuring stock illiquidity: An investigation of the demand and supply schedules at the TASE. *Journal of Financial Economics*, 74(3):461–486.

Kalay, A. and Wohl, A. (2009). Detecting liquidity traders. *Journal of Financial and Quantitative Analysis*, 44(1):29–54.

Kandel, E. and Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4):831–872.

Kaniel, R. and Liu, H. (2006). So what orders do informed traders use? *Journal of Business*, 79(4):1867–1913.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22(1):109–126.

Kozhan, R. and Salmon, M. (2012). The information content of a limit order book: The case of an FX market. *Journal of Financial Markets*, 15(1):1–28.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.

Malinova, K. and Park, A. (2013). Liquidity, volume and price efficiency: The impact of order vs. quote driven trading. *Journal of Financial Markets*, 16(1):104–126.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets*, 8(4):377–399.

McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, 27(1-3):10–45.

Næs, R. and Skjeltorp, J. A. (2006). Order book characteristics and the volume-volatility relation:

Empirical evidence from a limit order market. *Journal of Financial Markets*, 9(4):408–432.
Nolte, I. (2008). Modeling a multivariate transaction process. *Journal of Financial Econometrics*, 6(1):143–170.
Nowak, S. and Anderson, H. M. (2014). How does public information affect the frequency of trading in airline stocks? *Journal of Banking & Finance*, 44:26 – 38.
O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, 116(2):257 – 270.
Pascual, R. and Veredas, D. (2010). Does the open limit order book matter in explaining informational volatility? *Journal of Financial Econometrics*, 8(1):57–87.
Pham, M. C., Anderson, H. M., Duong, H. N., and Lajbcygier, P. (2020). The effects of trade size and market depth on immediate price impact in a limit order book market. *Journal of Economic Dynamics and Control*, 120:103992.
Ranaldo, A. (2004). Order aggressiveness in limit order book markets. *Journal of Financial Markets*, 7(1):53–74.
Renault, E., van der Heijden, T., and Werker, B. J. (2014). The dynamic mixed hitting-time model for multiple transaction prices and times. *Journal of Econometrics*, 180(2):233–250.
Renault, E. and Werker, B. J. (2011). Causality effects in return volatility measures with random times. *Journal of Econometrics*, 160(1):272–279.
Roşu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies*, 22(11):4601–4641.
Russell, J. R. and Engle, R. F. (2005). A discrete-state continuous-time model of financial transactions prices and times. *Journal of Business & Economic Statistics*, 23(2):166–180.
SEC (2010). Concept release on equity market structure. U.S. Securities & Exchange Commission, Release No.34-61458, File No. S7-02-10.
Shahzad, H., Duong, H. N., Kalev, P. S., and Singh, H. (2014). Trading volume, realized volatility and jumps in the Australian stock market. *Journal of International Financial Markets, Institutions and Money*, 31:414–430.
Shalen, C. T. (1993). Volume, volatility, and the dispersion of beliefs. *Review of Financial Studies*, 6(2):405–434.
Tian, X., Duong, H. N., and Kalev, P. S. (2019). Information content of the limit order book for crude oil futures price volatility. *Energy Economics*, 81:584–597.
Valenzuela, M., Zer, I., Fryzlewicz, P., and Rheinländer, T. (2015). Relative liquidity and future volatility. *Journal of Financial Markets*, 24:25–48.
Wald, J. K. and Horrigan, H. T. (2005). Optimal limit order choice. *Journal of Business*, 78(2):597–620.
Wang, J. and Wu, C. (2015). Liquidity, credit quality, and the relation between volatility and trading activity: Evidence from the corporate bond market. *Journal of Banking & Finance*, 50(0):183 – 203.
Xu, X. E., Chen, P., and Wu, C. (2006). Time and dynamic volume-volatility relation. *Journal of Banking & Finance*, 30(5):1535–1558.

Table 1: Summary statistics of trading activities and the order book characteristics

	Large cap	**Mid cap**	**Small cap**	**All stocks**
Number of stocks	49	48	97	194
Market capitalization ($AUD bn)	22.707	3.266	1.016	7.051
Shares outstanding (millions)	1748.409	856.875	482.700	894.969
Return (%) ($\times$100)	-0.003	0.007	-0.034	-0.016
Volume (thousand shares)	1.457	1.564	2.328	1.919
Duration (secs)	11.582	21.788	33.864	25.248
Absretpd	0.834	1.502	2.005	1.585
Volpd	81.681	84.777	93.033	88.123
Spread (%)	0.123	0.237	0.442	0.311
Bidvol (thousand shares)	56.436	155.741	100.329	102.953
Askvol (thousand shares)	56.861	66.036	84.114	72.758
Depth (thousand shares)	113.696	222.123	184.846	176.098
BidSlope	25.268	11.456	7.030	12.732
AskSlope	25.155	11.355	6.987	12.657
Slope	25.215	11.407	7.009	12.696
N	42.804	23.124	16.235	24.650
ATS (thousand shares)	1.639	1.767	2.710	2.206
OIB	0.635	0.451	0.427	0.486
QTT	13.763	13.004	12.349	12.868

This table presents summary statistics of trading activities and the order book characteristics for the constituent stocks of the S&P/ASX200 index in *July-December 2014*. These stocks are classified into three groups: "**Large cap**" which contains stocks in the S&P/ASX50 index, "**Mid cap**" which contains stocks in the S&P/ASX100 index but outside the S&P/ASX50 index, and "**Small cap**" which contains the remaining stocks in the S&P/ASX200 index. "Market capitalization" (in $AUD billion) is the market capitalization of firms as of 1 July 2014. "Shares outstanding" is the number of shares outstanding (in millions) right before a trade. "Return" (in %, and multiplied by 100) measures the change in log of the mid-quote right before a trade and the next trade. "Volume" is the number of shares (in thousands) traded in each trade. "Duration" (in seconds) is the time interval between two consecutive trades. "Absretpd" is the absolute return per unit of time, calculated as the absolute value of the return of a trade divided by its duration. "Volpd" is the share volume traded per unit of time, calculated as the volume (in thousands) of a trade divided by its duration. "Spread" (in %) is the relative spread (i.e. quoted spread as a % of the mid-quote right before a trade). "Bidvol", "Askvol" and "Depth" are respectively the total share volumes (in thousands) available at the best bid price, the best ask price, and both best bid and ask prices right before a trade. "BidSlope" ("AskSlope") is the slope of the bid (ask) side of the order book using 10 best bid/ask price levels right before a trade. "Slope" is the slope of the limit order book right before a trade, calculated as ("BidSlope" + "AskSlope")/2. "N" ("ATS") is the number of trades (the average trade size, in thousands) during a 5 minute interval right before a trade. "OIB" is the order imbalance, defined as the number of buys minus the number of sells during a 5 minute interval right before a trade. "QTT" is the quote to trade ratio during a 5 minute interval right before a trade. All variables for each stock are winsorized at the 1^{st} and 99^{th} quantiles to avoid the effects of outliers. All the statistics reported in the table (excepting those in the first line) are first computed for each individual stock and then equally averaged across all stocks.

Table 2: Slope and the volume-volatility relation: Combined limit order book

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%-_{5\%}$	$\%+_{5\%}$	**Med**	$\%-_{5\%}$	$\%+_{5\%}$	**Med**	$\%-_{5\%}$	$\%+_{5\%}$
Panel A: Constant contemporaneous volume-volatility relation										
$v_{i,t}$		0.937	0.0%	95.9%	0.874	0.0%	100.0%	1.012	0.0%	100.0%
$Slope_{i,t}$		-0.029	69.4%	24.5%	-0.299	85.4%	10.4%	-0.398	75.3%	17.5%
$\ln(T_{i,t})$		-0.633	100.0%	0.0%	-1.226	100.0%	0.0%	-1.469	100.0%	0.0%
$N_{i,t}$		-0.009	100.0%	0.0%	-0.036	100.0%	0.0%	-0.062	100.0%	0.0%
$ATS_{i,t}$		0.193	0.0%	98.0%	0.216	2.1%	95.8%	0.130	0.0%	99.0%
$OIBtr_{i,t}$		0.000	16.3%	36.7%	0.002	4.2%	41.7%	0.005	8.2%	41.2%
$QTT_{i,t}$		0.016	0.0%	100.0%	0.029	2.1%	97.9%	0.041	0.0%	100.0%
adj. R^2		0.159	-	-	0.174	-	-	0.179	-	-
Panel B: Endogenous contemporaneous volume-volatility relation										
$v_{i,t}$		3.980	0.0%	98.0%	8.420	2.1%	97.9%	4.941	0.0%	92.8%
$v_{i,t}Slope_{i,t}$		-0.132	77.6%	12.2%	-0.826	85.4%	8.3%	-0.577	82.5%	7.2%
$Slope_{i,t}$		-0.018	61.2%	30.6%	-0.207	79.2%	12.5%	-0.233	67.0%	21.6%
adj. R^2		0.169	-	-	0.182	-	-	0.185	-	-
Panel C: Endogenous dynamic volume-volatility relation										
$v_{i,t}$	Lag 0	3.865	2.0%	95.9%	8.265	2.1%	95.8%	4.931	0.0%	91.8%
	$\sum_{1:p}$	1.395	0.0%	77.6%	1.812	0.0%	47.9%	0.964	3.1%	40.2%
	$\sum_{0:p}$	7.013	0.0%	95.9%	11.046	2.1%	93.8%	6.116	0.0%	94.8%
$v_{i,t}Slope_{i,t}$	Lag 0	-0.128	77.6%	12.2%	-0.786	85.4%	8.3%	-0.574	81.4%	8.2%
	$\sum_{1:p}$	-0.034	65.3%	0.0%	-0.140	41.7%	0.0%	-0.111	39.2%	4.1%
	$\sum_{0:p}$	-0.205	81.6%	10.2%	-1.011	87.5%	4.2%	-0.737	82.5%	3.1%
$Slope_{i,t}$	Lag 0	0.017	30.6%	57.1%	-0.019	41.7%	39.6%	0.188	27.8%	56.7%
	$\sum_{1:p}$	-0.111	85.7%	2.0%	-0.247	70.8%	4.2%	-0.520	72.2%	12.4%
	$\sum_{0:p}$	-0.084	93.9%	0.0%	-0.380	87.5%	0.0%	-0.437	90.7%	1.0%
adj. R^2		0.174	-	-	0.193	-	-	0.194	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A to C is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t} = Slope_{i,t}$ is a potential predictor of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. The restriction $\delta_0 = 0$ is imposed in Panel A, and the coefficient lag length p is set to $p = 0$ in Panels A and B, while p is set to $p = 5$ in Panel C.

This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ from the volatility equation and those for $y_{i,t}$ in Panel A only, but a complete version of this table is available upon request. The regression is separately run for each stock, using Newey-West heteroskedasticity and autocorrelation consistent estimation. $\Sigma_{i:p}$ (in Panel C only) denotes the sum of the coefficients from lag i up to lag p. For brevity, we only report the median coefficients in "**Med**" column for each group. $\%-_{5\%}$ ($\%+_{5\%}$) indicates the proportion of estimates in each group that are significantly negative (positive) at a 5% level. "adj.R^2" denotes the adjusted R^2. Note that as the median operator is not additive, "Lag 0" and "$\sum_{1:p}$" median coefficients generally do not add up to that of "$\sum_{0:p}$".

Table 3: Slope and the volume-volatility relation: Bid vs. Ask sides

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: Constant contemporaneous volume-volatility relation										
$v_{i,t}$		1.390	0.0%	100.0%	1.295	0.0%	100.0%	1.417	0.0%	100.0%
$BidSlope_{i,t}B_{i,t}$		0.002	36.7%	55.1%	-0.017	39.6%	33.3%	0.019	25.8%	40.2%
$BidSlope_{i,t}S_{i,t}$		-0.039	93.9%	2.0%	-0.279	97.9%	2.1%	-0.344	90.7%	1.0%
$AskSlope_{i,t}B_{i,t}$		-0.037	89.8%	6.1%	-0.212	89.6%	4.2%	-0.337	82.5%	5.2%
$AskSlope_{i,t}S_{i,t}$		0.003	30.6%	55.1%	0.009	25.0%	37.5%	0.073	24.7%	52.6%
adj. R^2		0.235	-	-	0.252	-	-	0.244	-	-
Panel B: Endogenous contemporaneous volume-volatility relation										
$v_{i,t}$		6.916	0.0%	98.0%	11.525	2.1%	97.9%	5.443	0.0%	95.9%
$v_{i,t}BidSlope_{i,t}B_{i,t}$		0.101	2.0%	75.5%	0.327	4.2%	66.7%	0.446	0.0%	71.1%
$v_{i,t}BidSlope_{i,t}S_{i,t}$		-0.467	91.8%	0.0%	-1.569	93.8%	6.2%	-1.360	96.9%	0.0%
$v_{i,t}AskSlope_{i,t}B_{i,t}$		-0.467	95.9%	0.0%	-1.584	93.8%	2.1%	-1.418	92.8%	0.0%
$v_{i,t}AskSlope_{i,t}S_{i,t}$		0.157	0.0%	85.7%	0.382	6.2%	72.9%	0.583	1.0%	71.1%
$BidSlope_{i,t}B_{i,t}$		0.001	36.7%	51.0%	-0.045	45.8%	25.0%	-0.006	30.9%	29.9%
$BidSlope_{i,t}S_{i,t}$		-0.020	89.8%	4.1%	-0.154	83.3%	2.1%	-0.212	79.4%	6.2%
$AskSlope_{i,t}B_{i,t}$		-0.027	83.7%	10.2%	-0.113	83.3%	4.2%	-0.191	74.2%	9.3%
$AskSlope_{i,t}S_{i,t}$		0.002	36.7%	49.0%	-0.001	27.1%	27.1%	0.024	26.8%	43.3%
adj. R^2		0.270	-	-	0.277	-	-	0.268	-	-
Panel C: Endogenous dynamic volume-volatility relation										
$v_{i,t}$	Lag 0	6.580	0.0%	93.9%	11.348	2.1%	97.9%	5.317	0.0%	95.9%
	$\sum_{1:p}$	1.479	0.0%	77.6%	1.176	0.0%	52.1%	0.799	3.1%	39.2%
	$\sum_{0:p}$	8.276	0.0%	95.9%	12.024	2.1%	95.8%	6.493	0.0%	94.8%
$v_{i,t}BidSlope_{i,t}B_{i,t}$	Lag 0	0.102	2.0%	75.5%	0.318	4.2%	66.7%	0.471	0.0%	72.2%
	$\sum_{1:p}$	-0.045	46.9%	2.0%	-0.081	39.6%	2.1%	-0.180	43.3%	1.0%
	$\sum_{0:p}$	0.039	4.1%	46.9%	0.128	6.2%	43.8%	0.284	3.1%	44.3%
$v_{i,t}BidSlope_{i,t}S_{i,t}$	Lag 0	-0.468	89.8%	0.0%	-1.576	93.8%	4.2%	-1.354	96.9%	0.0%
	$\sum_{1:p}$	0.006	18.4%	18.4%	0.049	4.2%	18.8%	0.087	2.1%	22.7%
	$\sum_{0:p}$	-0.473	87.8%	0.0%	-1.481	93.8%	6.2%	-1.192	89.7%	0.0%
$v_{i,t}AskSlope_{i,t}B_{i,t}$	Lag 0	-0.471	95.9%	0.0%	-1.559	91.7%	2.1%	-1.440	92.8%	0.0%
	$\sum_{1:p}$	0.001	20.4%	14.3%	0.004	8.3%	20.8%	0.052	4.1%	17.5%
	$\sum_{0:p}$	-0.401	89.8%	4.1%	-1.511	93.8%	4.2%	-1.221	87.6%	0.0%
$v_{i,t}AskSlope_{i,t}S_{i,t}$	Lag 0	0.163	0.0%	87.8%	0.388	6.2%	75.0%	0.601	1.0%	73.2%
	$\sum_{1:p}$	-0.050	57.1%	0.0%	-0.167	47.9%	0.0%	-0.213	44.3%	0.0%
	$\sum_{0:p}$	0.043	0.0%	46.9%	0.142	10.4%	37.5%	0.327	5.2%	41.2%
$BidSlope_{i,t}B_{i,t}$	Lag 0	-0.010	44.9%	44.9%	-0.093	58.3%	29.2%	-0.125	46.4%	30.9%
	$\sum_{1:p}$	-0.020	38.8%	8.2%	0.020	22.9%	27.1%	-0.054	23.7%	25.8%
	$\sum_{0:p}$	-0.039	36.7%	4.1%	-0.036	29.2%	10.4%	-0.021	16.5%	11.3%
$BidSlope_{i,t}S_{i,t}$	Lag 0	-0.007	51.0%	40.8%	-0.137	50.0%	31.2%	-0.057	40.2%	35.1%
	$\sum_{1:p}$	-0.009	36.7%	6.1%	-0.020	25.0%	14.6%	-0.084	34.0%	8.2%
	$\sum_{0:p}$	-0.015	28.6%	2.0%	-0.100	37.5%	0.0%	-0.169	43.3%	2.1%
$AskSlope_{i,t}B_{i,t}$	Lag 0	-0.009	44.9%	40.8%	-0.037	41.7%	33.3%	0.048	29.9%	44.3%
	$\sum_{1:p}$	-0.041	46.9%	2.0%	-0.045	35.4%	4.2%	-0.196	47.4%	3.1%
	$\sum_{0:p}$	-0.041	40.8%	2.0%	-0.119	39.6%	2.1%	-0.242	49.5%	2.1%
$AskSlope_{i,t}S_{i,t}$	Lag 0	0.005	36.7%	49.0%	-0.025	41.7%	27.1%	-0.066	35.1%	40.2%
	$\sum_{1:p}$	-0.034	51.0%	8.2%	-0.044	33.3%	14.6%	-0.063	34.0%	18.6%
	$\sum_{0:p}$	-0.042	51.0%	2.0%	-0.066	22.9%	10.4%	-0.089	22.7%	9.3%
adj. R^2		0.272	-	-	0.288	-	-	0.274	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A to C is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t} = (BidSlope_{i,t}B_{i,t}, BidSlope_{i,t}S_{i,t}, AskSlope_{i,t}B_{i,t}, AskSlope_{i,t}S_{i,t})'$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. The restriction $\delta_0 = 0$ is imposed in Panel A, and the coefficient lag length p is set to $p = 0$ in Panels A and B, while p is set to $p = 5$ in Panel C. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ from the volatility equation only, but a complete version of this table is available upon request.

Table 4: Spread, depth and the volume-volatility relation: Combined limit order book

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: Constant contemporaneous volume-volatility relation										
$v_{i,t}$		0.937	0.0%	95.9%	0.847	0.0%	100.0%	1.022	0.0%	100.0%
$Spread_{i,t}$		3.607	20.4%	71.4%	5.573	10.4%	72.9%	3.029	21.6%	63.9%
$Depth_{i,t}$		-0.994	49.0%	36.7%	-2.994	58.3%	29.2%	-1.600	57.7%	24.7%
adj. R^2		0.155	-	-	0.171	-	-	0.179	-	-
Panel B: Endogenous contemporaneous volume-volatility relation										
$v_{i,t}$		-0.463	44.9%	38.8%	-0.908	45.8%	27.1%	0.444	19.6%	39.2%
$v_{i,t}Spread_{i,t}$		18.181	10.2%	75.5%	12.988	6.2%	79.2%	5.169	10.3%	68.0%
$v_{i,t}Depth_{i,t}$		-26.223	83.7%	0.0%	-13.381	85.4%	8.3%	-7.094	79.4%	1.0%
$Spread_{i,t}$		2.292	22.4%	67.3%	2.969	12.5%	66.7%	1.637	24.7%	54.6%
$Depth_{i,t}$		1.205	36.7%	51.0%	-1.617	56.2%	35.4%	-0.412	47.4%	35.1%
adj. R^2		0.162	-	-	0.180	-	-	0.184	-	-
Panel C: Endogenous dynamic volume-volatility relation										
$v_{i,t}$	Lag 0	-0.420	44.9%	40.8%	-0.872	45.8%	27.1%	0.460	19.6%	40.2%
	$\sum_{1:p}$	0.204	22.4%	24.5%	0.317	10.4%	12.5%	0.040	8.2%	10.3%
	$\sum_{0:p}$	-0.014	28.6%	28.6%	-0.272	22.9%	16.7%	0.392	8.2%	36.1%
$v_{i,t}Spread_{i,t}$	Lag 0	16.727	10.2%	75.5%	12.819	6.2%	75.0%	5.057	10.3%	67.0%
	$\sum_{1:p}$	8.092	0.0%	49.0%	2.885	2.1%	22.9%	0.783	3.1%	19.6%
	$\sum_{0:p}$	26.320	4.1%	85.7%	15.889	2.1%	79.2%	6.055	6.2%	68.0%
$v_{i,t}Depth_{i,t}$	Lag 0	-25.380	83.7%	0.0%	-13.103	85.4%	8.3%	-7.176	79.4%	0.0%
	$\sum_{1:p}$	-9.332	49.0%	12.2%	-3.680	43.8%	12.5%	-1.552	29.9%	1.0%
	$\sum_{0:p}$	-29.747	75.5%	4.1%	-18.056	85.4%	8.3%	-9.570	84.5%	1.0%
$Spread_{i,t}$	Lag 0	-7.793	63.3%	26.5%	-1.510	43.8%	29.2%	-4.518	54.6%	22.7%
	$\sum_{1:p}$	19.423	0.0%	71.4%	9.051	4.2%	56.2%	9.304	12.4%	66.0%
	$\sum_{0:p}$	14.797	0.0%	73.5%	7.596	0.0%	66.7%	4.718	2.1%	68.0%
$Depth_{i,t}$	Lag 0	61.806	14.3%	79.6%	25.541	27.1%	58.3%	25.687	7.2%	71.1%
	$\sum_{1:p}$	-61.657	79.6%	6.1%	-28.805	62.5%	12.5%	-32.570	74.2%	4.1%
	$\sum_{0:p}$	-13.354	59.2%	10.2%	-7.715	66.7%	6.2%	-4.852	77.3%	0.0%
adj. R^2		0.172	-	-	0.188	-	-	0.194	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A to C is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t} = (Spread_{i,t}, Depth_{i,t})'$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. See Table A.1 for definitions of the variables. The restriction $\delta_0 = 0$ is imposed in Panel A, and the coefficient lag length p is set to $p = 0$ in Panels A and B, while p is set to $p = 5$ in Panel C. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ from the volatility equation only, but a complete version of this table is available upon request.

Table 5: Spread, depth and the volume-volatility relation: Bid vs. Ask sides

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: Constant contemporaneous volume-volatility relation										
$v_{i,t}$		1.254	0.0%	100.0%	1.238	0.0%	100.0%	1.416	0.0%	100.0%
$Spread_{i,t}$		4.564	16.3%	77.6%	6.546	10.4%	79.2%	3.508	13.4%	73.2%
$BV_{i,t}B_{i,t}$		12.233	26.5%	71.4%	4.921	22.9%	60.4%	3.530	20.6%	55.7%
$BV_{i,t}S_{i,t}$		-9.697	79.6%	6.1%	-8.736	91.7%	6.2%	-7.919	95.9%	0.0%
$AV_{i,t}B_{i,t}$		-6.753	75.5%	8.2%	-6.557	85.4%	10.4%	-5.477	89.7%	0.0%
$AV_{i,t}S_{i,t}$		13.568	6.1%	83.7%	4.846	14.6%	68.8%	3.768	13.4%	59.8%
adj. R^2		0.189	-	-	0.215	-	-	0.212	-	-
Panel B: Endogenous contemporaneous volume-volatility relation										
$v_{i,t}$		-0.357	46.9%	42.9%	-0.760	45.8%	25.0%	0.578	16.5%	40.2%
$v_{i,t}Spread_{i,t}$		18.988	10.2%	75.5%	15.285	6.2%	79.2%	5.354	6.2%	71.1%
$v_{i,t}BV_{i,t}B_{i,t}$		25.892	12.2%	63.3%	5.447	16.7%	50.0%	2.682	9.3%	38.1%
$v_{i,t}BV_{i,t}S_{i,t}$		-83.303	93.9%	0.0%	-36.415	87.5%	0.0%	-23.425	92.8%	0.0%
$v_{i,t}AV_{i,t}B_{i,t}$		-77.159	98.0%	0.0%	-37.500	87.5%	0.0%	-21.382	90.7%	0.0%
$v_{i,t}AV_{i,t}S_{i,t}$		30.906	4.1%	61.2%	4.567	10.4%	41.7%	1.473	12.4%	41.2%
$Spread_{i,t}$		3.339	16.3%	71.4%	4.062	10.4%	77.1%	2.527	15.5%	63.9%
$BV_{i,t}B_{i,t}$		10.111	26.5%	69.4%	2.572	25.0%	56.2%	2.567	21.6%	55.7%
$BV_{i,t}S_{i,t}$		-5.413	63.3%	14.3%	-5.117	70.8%	6.2%	-4.408	86.6%	1.0%
$AV_{i,t}B_{i,t}$		-2.899	57.1%	18.4%	-3.696	66.7%	12.5%	-2.851	67.0%	2.1%
$AV_{i,t}S_{i,t}$		10.542	8.2%	81.6%	3.769	16.7%	64.6%	2.735	11.3%	56.7%
adj. R^2		0.197	-	-	0.226	-	-	0.225	-	-
Panel C: Endogenous dynamic volume-volatility relation										
$v_{i,t}$	Lag 0	-0.257	44.9%	42.9%	-0.748	47.9%	25.0%	0.547	16.5%	40.2%
	$\sum_{1:p}$	-0.066	24.5%	14.3%	0.236	12.5%	10.4%	0.016	6.2%	10.3%
	$\sum_{0:p}$	-0.148	34.7%	26.5%	-0.528	33.3%	20.8%	0.447	8.2%	36.1%
$v_{i,t}Spread_{i,t}$	Lag 0	18.739	10.2%	75.5%	14.545	6.2%	77.1%	5.259	8.2%	71.1%
	$\sum_{1:p}$	9.219	0.0%	59.2%	3.085	2.1%	22.9%	0.986	2.1%	19.6%
	$\sum_{0:p}$	27.570	4.1%	87.8%	17.782	2.1%	77.1%	6.972	5.2%	71.1%
$v_{i,t}BV_{i,t}B_{i,t}$	Lag 0	33.340	12.2%	63.3%	4.931	10.4%	50.0%	2.422	9.3%	39.2%
	$\sum_{1:p}$	-7.249	32.7%	0.0%	-3.391	16.7%	0.0%	-2.095	23.7%	0.0%
	$\sum_{0:p}$	25.947	16.3%	55.1%	3.146	16.7%	31.2%	0.904	10.3%	18.6%
$v_{i,t}BV_{i,t}S_{i,t}$	Lag 0	-79.618	93.9%	0.0%	-35.032	85.4%	0.0%	-23.226	92.8%	0.0%
	$\sum_{1:p}$	-10.397	34.7%	20.4%	-2.641	29.2%	10.4%	-1.571	20.6%	2.1%
	$\sum_{0:p}$	-77.428	77.6%	2.0%	-42.287	83.3%	6.2%	-25.294	91.8%	0.0%
$v_{i,t}AV_{i,t}B_{i,t}$	Lag 0	-77.249	98.0%	0.0%	-35.970	91.7%	0.0%	-21.503	92.8%	0.0%
	$\sum_{1:p}$	-13.508	34.7%	18.4%	-7.367	39.6%	16.7%	-2.007	22.7%	3.1%
	$\sum_{0:p}$	-80.332	81.6%	2.0%	-47.534	89.6%	6.2%	-25.972	87.6%	0.0%
$v_{i,t}AV_{i,t}S_{i,t}$	Lag 0	30.841	4.1%	63.3%	4.478	8.3%	41.7%	1.299	10.3%	40.2%
	$\sum_{1:p}$	-8.706	22.4%	0.0%	-2.598	22.9%	2.1%	-1.830	12.4%	1.0%
	$\sum_{0:p}$	31.542	10.2%	46.9%	0.498	18.8%	31.2%	0.830	15.5%	22.7%
$Spread_{i,t}$	Lag 0	-3.653	55.1%	34.7%	2.411	37.5%	52.1%	-1.181	45.4%	34.0%
	$\sum_{1:p}$	14.497	8.2%	63.3%	2.927	16.7%	37.5%	5.575	14.4%	53.6%
	$\sum_{0:p}$	15.268	0.0%	71.4%	7.144	0.0%	64.6%	4.422	2.1%	66.0%
$BV_{i,t}B_{i,t}$	Lag 0	52.212	18.4%	71.4%	17.953	31.2%	54.2%	20.088	13.4%	55.7%
	$\sum_{1:p}$	-33.574	71.4%	14.3%	-16.805	54.2%	33.3%	-20.849	55.7%	8.2%
	$\sum_{0:p}$	-2.895	16.3%	14.3%	-0.780	0.0%	12.5%	-0.006	11.3%	12.4%
$BV_{i,t}S_{i,t}$	Lag 0	43.679	18.4%	71.4%	10.437	35.4%	50.0%	13.238	15.5%	51.5%
	$\sum_{1:p}$	-44.858	61.2%	4.1%	-21.941	47.9%	18.8%	-23.380	60.8%	5.2%
	$\sum_{0:p}$	-18.327	44.9%	8.2%	-10.195	45.8%	4.2%	-8.925	62.9%	0.0%
$AV_{i,t}B_{i,t}$	Lag 0	50.413	14.3%	77.6%	12.364	29.2%	52.1%	19.656	10.3%	62.9%
	$\sum_{1:p}$	-52.767	79.6%	0.0%	-22.814	50.0%	0.0%	-26.060	66.0%	3.1%
	$\sum_{0:p}$	-22.111	53.1%	6.1%	-12.336	52.1%	6.2%	-8.525	59.8%	0.0%
$AV_{i,t}S_{i,t}$	Lag 0	62.248	14.3%	77.6%	19.345	27.1%	58.3%	20.633	8.2%	62.9%
	$\sum_{1:p}$	-57.520	79.6%	8.2%	-20.757	58.3%	18.8%	-23.948	66.0%	8.2%
	$\sum_{0:p}$	-0.912	20.4%	2.0%	-1.112	14.6%	12.5%	-0.766	12.4%	8.2%
adj. R^2		0.210	-	-	0.236	-	-	0.234	-	-

Continued on next page

This table reports summary estimation results for all stocks in the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A to C is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t} = (Spread_{i,t}, BV_{i,t}B_{i,t}, BV_{i,t}S_{i,t}, AV_{i,t}B_{i,t}, AV_{i,t}S_{i,t})'$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. The restriction $\delta_0 = 0$ is imposed in Panel A, and the coefficient lag length p is set to $p = 0$ in Panels A and B, while p is set to $p = 5$ in Panel C. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ from the volatility equation only, but a complete version of this table is available upon request.

Table 6: Spread, depth, slope and the volume-volatility relation: Combined limit order book

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: Constant contemporaneous volume-volatility relation										
$v_{i,t}$		0.970	0.0%	98.0%	0.897	0.0%	100.0%	1.047	0.0%	100.0%
$Spread_{i,t}$		-6.335	65.3%	28.6%	-6.745	72.9%	10.4%	-4.682	60.8%	5.2%
$Depth_{i,t}$		5.203	20.4%	67.3%	5.320	14.6%	77.1%	2.025	22.7%	55.7%
$Slope_{i,t}$		-0.096	71.4%	26.5%	-0.572	89.6%	8.3%	-0.582	81.4%	6.2%
adj. R^2		0.159	-	-	0.176	-	-	0.181	-	-
Panel B: Endogenous contemporaneous volume-volatility relation										
$v_{i,t}$		10.725	10.2%	79.6%	22.215	6.2%	85.4%	10.652	0.0%	79.4%
$v_{i,t}Spread_{i,t}$		-32.536	61.2%	14.3%	-24.684	68.8%	6.2%	-7.339	59.8%	3.1%
$v_{i,t}Depth_{i,t}$		7.154	30.6%	40.8%	2.772	18.8%	35.4%	-0.682	35.1%	23.7%
$v_{i,t}Slope_{i,t}$		-0.342	75.5%	16.3%	-1.564	91.7%	6.2%	-1.088	73.2%	5.2%
$Spread_{i,t}$		-6.010	63.3%	28.6%	-5.722	68.8%	10.4%	-3.994	61.9%	7.2%
$Depth_{i,t}$		5.459	10.2%	71.4%	5.466	10.4%	79.2%	2.702	13.4%	58.8%
$Slope_{i,t}$		-0.077	71.4%	26.5%	-0.496	85.4%	8.3%	-0.489	80.4%	6.2%
adj. R^2		0.174	-	-	0.187	-	-	0.191	-	-
Panel C: Endogenous dynamic volume-volatility relation										
$v_{i,t}$	Lag 0	10.027	10.2%	77.6%	21.664	6.2%	85.4%	10.627	0.0%	80.4%
	$\sum_{1:p}$	0.256	28.6%	22.4%	3.037	14.6%	22.9%	1.777	3.1%	17.5%
	$\sum_{0:p}$	12.213	18.4%	67.3%	23.348	6.2%	79.2%	12.633	0.0%	76.3%
$v_{i,t}Spread_{i,t}$	Lag 0	-28.056	61.2%	14.3%	-23.886	64.6%	6.2%	-7.576	59.8%	3.1%
	$\sum_{1:p}$	6.042	18.4%	38.8%	-1.762	22.9%	16.7%	-1.310	9.3%	4.1%
	$\sum_{0:p}$	-21.467	49.0%	26.5%	-28.891	54.2%	8.3%	-9.907	48.5%	1.0%
$v_{i,t}Depth_{i,t}$	Lag 0	4.235	30.6%	40.8%	2.376	20.8%	35.4%	-0.892	34.0%	22.7%
	$\sum_{1:p}$	-8.079	28.6%	4.1%	-1.914	16.7%	6.2%	-0.615	14.4%	1.0%
	$\sum_{0:p}$	-3.300	36.7%	28.6%	-0.689	18.8%	22.9%	-1.016	32.0%	14.4%
$v_{i,t}Slope_{i,t}$	Lag 0	-0.327	73.5%	18.4%	-1.476	91.7%	6.2%	-1.078	73.2%	5.2%
	$\sum_{1:p}$	0.001	26.5%	28.6%	-0.088	25.0%	14.6%	-0.164	19.6%	1.0%
	$\sum_{0:p}$	-0.302	65.3%	22.4%	-1.740	81.2%	8.3%	-1.207	73.2%	2.1%
$Spread_{i,t}$	Lag 0	-5.538	59.2%	10.2%	-2.911	47.9%	12.5%	-0.755	38.1%	19.6%
	$\sum_{1:p}$	5.984	22.4%	40.8%	-9.140	35.4%	16.7%	-7.347	35.1%	8.2%
	$\sum_{0:p}$	3.380	26.5%	36.7%	-13.550	43.8%	12.5%	-6.910	47.4%	6.2%
$Depth_{i,t}$	Lag 0	53.795	12.2%	83.7%	24.443	16.7%	60.4%	21.384	4.1%	67.0%
	$\sum_{1:p}$	-61.702	77.6%	10.2%	-23.948	52.1%	16.7%	-22.482	66.0%	4.1%
	$\sum_{0:p}$	-5.272	30.6%	12.2%	-0.063	10.4%	22.9%	0.209	18.6%	11.3%
$Slope_{i,t}$	Lag 0	0.011	34.7%	51.0%	0.011	35.4%	41.7%	0.308	15.5%	53.6%
	$\sum_{1:p}$	-0.073	44.9%	18.4%	-0.372	58.3%	2.1%	-0.968	75.3%	2.1%
	$\sum_{0:p}$	-0.066	42.9%	24.5%	-0.636	62.5%	10.4%	-0.916	73.2%	0.0%
adj. R^2		0.185	-	-	0.203	-	-	0.204	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A to C is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t} = (Spread_{i,t}, Depth_{i,t}, Slope_{i,t})'$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. The restriction $\delta_0 = 0$ is imposed in Panel A, and the coefficient lag length p is set to $p = 0$ in Panels A and B, while p is set to $p = 5$ in Panel C. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ from the volatility equation only, but a complete version of this table is available upon request.

Table 7: Spread, depth, slope and the volume-volatility relation: Bid vs. Ask sides

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: Constant contemporaneous volume-volatility relation										
$v_{i,t}$		1.381	0.0%	100.0%	1.296	0.0%	100.0%	1.465	0.0%	100.0%
$Spread_{i,t}$		-1.790	49.0%	30.6%	-3.015	52.1%	18.8%	-2.188	44.3%	11.3%
$BV_{i,t}S_{i,t}$		9.272	16.3%	69.4%	5.580	10.4%	72.9%	0.970	20.6%	36.1%
$AV_{i,t}B_{i,t}$		8.400	18.4%	71.4%	6.384	8.3%	66.7%	1.531	14.4%	36.1%
$BidSlope_{i,t}S_{i,t}$		-0.065	83.7%	14.3%	-0.383	83.3%	4.2%	-0.433	83.5%	5.2%
$AskSlope_{i,t}B_{i,t}$		-0.065	81.6%	14.3%	-0.271	85.4%	8.3%	-0.366	76.3%	1.0%
adj. R^2		0.236	-	-	0.252	-	-	0.246	-	-
Panel B: Endogenous contemporaneous volume-volatility relation										
$v_{i,t}$		29.845	8.2%	87.8%	36.983	6.2%	91.7%	16.660	0.0%	87.6%
$v_{i,t}Spread_{i,t}$		-100.431	83.7%	8.2%	-52.209	83.3%	6.2%	-17.117	73.2%	5.2%
$v_{i,t}BV_{i,t}S_{i,t}$		65.062	16.3%	61.2%	36.890	10.4%	64.6%	4.138	13.4%	41.2%
$v_{i,t}AV_{i,t}B_{i,t}$		80.364	20.4%	63.3%	34.515	8.3%	60.4%	5.152	13.4%	35.1%
$v_{i,t}BidSlope_{i,t}S_{i,t}$		-0.820	85.7%	6.1%	-3.160	93.8%	6.2%	-2.671	91.8%	1.0%
$v_{i,t}AskSlope_{i,t}B_{i,t}$		-0.909	87.8%	4.1%	-2.933	91.7%	6.2%	-2.575	92.8%	0.0%
$Spread_{i,t}$		-1.719	51.0%	32.7%	-4.408	60.4%	18.8%	-1.985	50.5%	11.3%
$BV_{i,t}S_{i,t}$		11.473	10.2%	83.7%	6.105	8.3%	79.2%	2.166	9.3%	48.5%
$AV_{i,t}B_{i,t}$		11.553	12.2%	79.6%	8.011	4.2%	81.2%	3.102	9.3%	55.7%
$BidSlope_{i,t}S_{i,t}$		-0.045	83.7%	8.2%	-0.274	85.4%	2.1%	-0.266	78.4%	4.1%
$AskSlope_{i,t}B_{i,t}$		-0.049	81.6%	14.3%	-0.197	83.3%	4.2%	-0.227	70.1%	3.1%
adj. R^2		0.278	-	-	0.297	-	-	0.279	-	-
Panel C: Endogenous dynamic volume-volatility relation										
$v_{i,t}$	Lag 0	30.218	8.2%	87.8%	36.503	6.2%	91.7%	16.639	0.0%	88.7%
	$\sum_{1:p}$	1.260	28.6%	26.5%	1.103	14.6%	25.0%	0.796	4.1%	10.3%
	$\sum_{0:p}$	29.435	12.2%	81.6%	35.541	6.2%	89.6%	17.936	0.0%	86.6%
$v_{i,t}Spread_{i,t}$	Lag 0	-98.626	83.7%	8.2%	-52.148	85.4%	6.2%	-17.010	76.3%	4.1%
	$\sum_{1:p}$	4.377	18.4%	30.6%	0.676	16.7%	16.7%	0.338	7.2%	7.2%
	$\sum_{0:p}$	-96.019	75.5%	16.3%	-49.564	75.0%	6.2%	-16.107	64.9%	2.1%
$v_{i,t}BV_{i,t}S_{i,t}$	Lag 0	64.537	16.3%	61.2%	38.249	10.4%	64.6%	4.512	13.4%	42.3%
	$\sum_{1:p}$	-17.418	30.6%	2.0%	-9.876	18.8%	4.2%	-4.051	29.9%	0.0%
	$\sum_{0:p}$	56.117	18.4%	51.0%	26.766	10.4%	50.0%	1.275	16.5%	23.7%
$v_{i,t}AV_{i,t}B_{i,t}$	Lag 0	78.634	20.4%	63.3%	32.267	8.3%	62.5%	4.986	13.4%	35.1%
	$\sum_{1:p}$	-14.717	20.4%	0.0%	-6.871	20.8%	6.2%	-2.972	14.4%	0.0%
	$\sum_{0:p}$	52.137	20.4%	46.9%	21.558	2.1%	45.8%	1.849	15.5%	22.7%
$v_{i,t}BidSlope_{i,t}S_{i,t}$	Lag 0	-0.811	85.7%	6.1%	-3.186	93.8%	6.2%	-2.669	91.8%	1.0%
	$\sum_{1:p}$	0.033	8.2%	42.9%	0.140	4.2%	22.9%	0.180	1.0%	22.7%
	$\sum_{0:p}$	-0.800	81.6%	16.3%	-2.985	87.5%	6.2%	-2.149	81.4%	1.0%
$v_{i,t}AskSlope_{i,t}B_{i,t}$	Lag 0	-0.909	87.8%	4.1%	-2.961	91.7%	6.2%	-2.586	92.8%	0.0%
	$\sum_{1:p}$	0.016	8.2%	36.7%	0.056	2.1%	29.2%	0.128	1.0%	15.5%
	$\sum_{0:p}$	-0.839	81.6%	12.2%	-2.652	89.6%	6.2%	-2.273	85.6%	1.0%
$Spread_{i,t}$	Lag 0	-6.091	69.4%	6.1%	-6.988	60.4%	12.5%	-4.086	47.4%	13.4%
	$\sum_{1:p}$	3.844	14.3%	36.7%	0.914	22.9%	18.8%	0.166	12.4%	14.4%
	$\sum_{0:p}$	-3.644	24.5%	32.7%	-10.663	41.7%	14.6%	-4.809	33.0%	10.3%
$BV_{i,t}S_{i,t}$	Lag 0	57.390	10.2%	79.6%	15.468	25.0%	56.2%	14.929	5.2%	54.6%
	$\sum_{1:p}$	-62.621	61.2%	4.1%	-23.179	39.6%	6.2%	-15.343	47.4%	4.1%
	$\sum_{0:p}$	-11.812	24.5%	0.0%	-1.200	10.4%	14.6%	-0.047	15.5%	9.3%
$AV_{i,t}B_{i,t}$	Lag 0	59.366	6.1%	83.7%	24.653	8.3%	60.4%	20.213	4.1%	60.8%
	$\sum_{1:p}$	-53.029	63.3%	0.0%	-17.516	47.9%	6.2%	-14.848	52.6%	2.1%
	$\sum_{0:p}$	-4.502	24.5%	4.1%	1.018	4.2%	10.4%	0.258	12.4%	14.4%
$BidSlope_{i,t}S_{i,t}$	Lag 0	-0.022	55.1%	26.5%	-0.119	52.1%	25.0%	-0.043	33.0%	24.7%
	$\sum_{1:p}$	0.021	10.2%	38.8%	-0.008	12.5%	10.4%	-0.126	22.7%	4.1%
	$\sum_{0:p}$	-0.008	20.4%	30.6%	-0.203	20.8%	6.2%	-0.189	27.8%	0.0%
$AskSlope_{i,t}B_{i,t}$	Lag 0	-0.026	55.1%	34.7%	-0.063	52.1%	27.1%	0.014	26.8%	25.8%
	$\sum_{1:p}$	-0.003	16.3%	26.5%	-0.042	18.8%	12.5%	-0.162	30.9%	1.0%
	$\sum_{0:p}$	-0.032	32.7%	22.4%	-0.223	31.2%	8.3%	-0.238	35.1%	3.1%
adj. R^2		0.288	-	-	0.314	-	-	0.291	-	-

Continued on next page

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A to C is

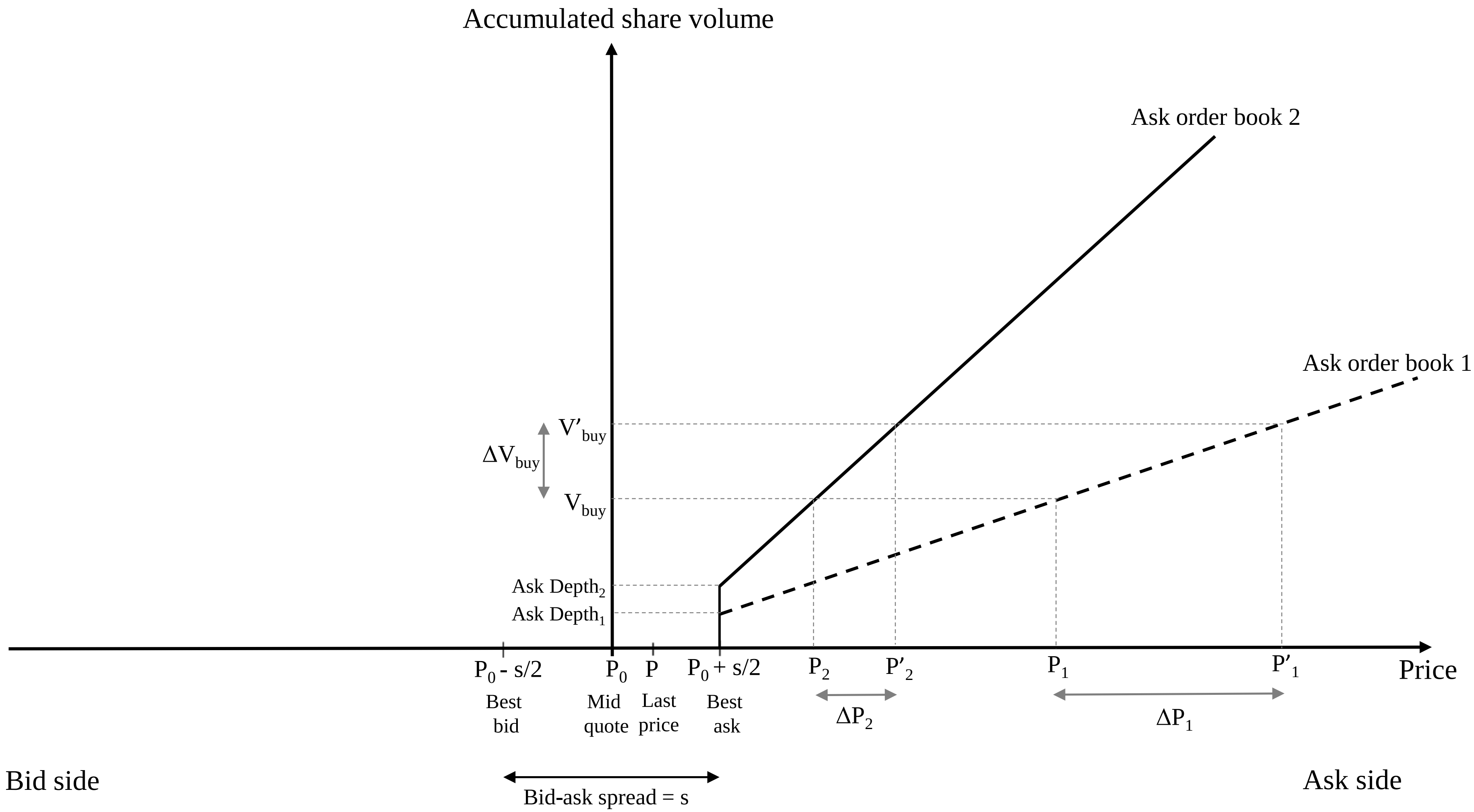
$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{p} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{p} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t} = (Spread_{i,t}, BV_{i,t}B_{i,t}, BV_{i,t}S_{i,t}, AV_{i,t}B_{i,t}, AV_{i,t}S_{i,t}, BidSlope_{i,t}B_{i,t}, BidSlope_{i,t}S_{i,t}, AskSlope_{i,t}B_{i,t}, AskSlope_{i,t}S_{i,t})'$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. The restriction $\delta_0 = 0$ is imposed in Panel A, and the coefficient lag length p is set to $p = 0$ in Panels A and B, while p is set to $p = 5$ in Panel C. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ of the order book that are of the opposite side to the direction of a trade (e.g. $AskSlope_{i,t}B_{i,t}$) from the volatility equation only, but a complete version of this table is available upon request.

Figure 1: Order book slope and the volume-volatility relation



Note: This figure depicts how the order book slope affects the volume-volatility relation, using the ask side of the order book as an illustration.

Appendix

Table A.1: Definitions of variables

Notation	Description
$r_{i,t}$	Return of the t-th trade in stock i: $r_{i,t} = 100(\ln(q_{i,t+1}) - \ln(q_{i,t}))$, where $q_{i,t}$ is the midpoint of the bid and ask quotes right before the trade
$T_{i,t}$	Time duration (in seconds) between the $(t-1)$-th and t-th trades
$\sigma_{i,t}$	Volatility *per unit of time* of the t-th trade: $\sigma_{i,t} = \lvert\widehat{\epsilon}_{i,t}\rvert / T_{i,t}$, where $\widehat{\epsilon}_{i,t}$ is the residual from the model of returns specified in Equation (2)
$v_{i,t}$	Volume *per unit of time* of the t-th trade: $v_{i,t} = V_{i,t}/T_{i,t}$, where $V_{i,t}$ is the number of shares traded (times 1000) divided by the total number of shares outstanding right before the trade, and $T_{i,t}$ is defined above
$x_{i,t}$	A vector of potential predictors of the volume-volatility relation of the t-th trade
$y_{i,t}$	A vector of control variables that allows for the effects of the order flow prior to the t-th trade
$Spread_{i,t}$	Relative spread, defined as quoted spread as a % of the mid-quote right before the t-th trade
$BV_{i,t}$	Total number of shares available at the best *bid* price (times 1000) divided by the total number of shares outstanding right before the t-th trade
$AV_{i,t}$	Total number of shares available at the best *ask* price (times 1000) divided by the total number of shares outstanding right before the t-th trade
$Depth_{i,t}$	Total number of shares available at the best *bid and ask* prices (times 1000) divided by total number of shares outstanding right before the t-th trade: $Depth_{i,t} = BV_{i,t} + AV_{i,t}$
$BidSlope_{i,t}$	Slope of the *bid* order book right before the t-th trade, defined in Equation (4) and calculated using the 10 best bid/ask price levels right before the trade
$AskSlope_{i,t}$	Slope of the *ask* order book right before the t-th trade, defined in Equation (5) and calculated using the 10 best bid/ask price levels right before the trade
$Slope_{i,t}$	Slope of the limit order book right before the t-th trade: $Slope_{i,t} = (BidSlope_{i,t} + AskSlope_{i,t})/2$
$B_{i,t}$	Buy indicator: equals 1 if the t-th trade is a purchase, 0 otherwise
$S_{i,t}$	Sell indicator: equals 1 if the t-th trade is a sale, 0 otherwise
$N_{i,t}$	Number of transactions during the 5-minute interval right before the t-th trade
$ATS_{i,t}$	Average trade size (times 10^6 and divided by the total number of shares outstanding) during the 5-minute interval right before the t-th trade
$OIB_{i,t}$	Order imbalance (= number of buys - number of sells) during the 5-minute interval right before the t-th trade
$QTT_{i,t}$	Quote to trade ratio (= total number of order submissions, revisions and cancellations divided by number of trades) during the 5-minute interval right before the t-th trade
$Day_{k,i,t}$	Day-of-week dummy variables, $k = 1, \cdots, 5$ for Monday till Friday
$Monday_{i,t}$	Dummy variable for Monday (same as $Day_{1,i,t}$)
$hour_{k,i,t}$	Time-of-day dummy variables, $k = 1, \cdots, 6$ for six hourly intervals: 10:10-11:00, 11:00-12:00, 12:00-13:00, 13:00-14:00, 14:00-15:00 and 15:00-16:00
$hour1_{i,t}$	Dummy variable for the first trading hour (10:10-11:00) of a day (same as $hour_{1,i,t}$)

Table A.2: LOB and the *endogenous dynamic* volume-volatility relation: Combined LOB

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: 5 best bid/ask price levels										
$v_{i,t}$	Lag 0	10.150	10.2%	77.6%	21.764	6.2%	87.5%	10.734	0.0%	80.4%
	$\sum_{1:p}$	0.103	28.6%	22.4%	3.304	14.6%	22.9%	1.851	2.1%	19.6%
	$\sum_{0:p}$	12.334	18.4%	67.3%	23.418	6.2%	79.2%	12.693	0.0%	76.3%
$v_{i,t}Spread_{i,t}$	Lag 0	-28.723	61.2%	14.3%	-24.056	66.7%	6.2%	-8.093	61.9%	3.1%
	$\sum_{1:p}$	6.057	18.4%	38.8%	-2.102	20.8%	16.7%	-1.323	11.3%	3.1%
	$\sum_{0:p}$	-23.129	51.0%	24.5%	-29.233	54.2%	8.3%	-10.510	50.5%	1.0%
$v_{i,t}Depth_{i,t}$	Lag 0	4.584	30.6%	42.9%	2.522	20.8%	35.4%	-0.617	32.0%	23.7%
	$\sum_{1:p}$	-8.132	28.6%	4.1%	-1.808	16.7%	8.3%	-0.641	14.4%	1.0%
	$\sum_{0:p}$	-2.972	36.7%	28.6%	-0.594	18.8%	27.1%	-0.607	32.0%	15.5%
$v_{i,t}Slope5_{i,t}$	Lag 0	-0.164	73.5%	18.4%	-0.747	91.7%	6.2%	-0.553	74.2%	5.2%
	$\sum_{1:p}$	0.000	26.5%	28.6%	-0.057	27.1%	14.6%	-0.102	21.6%	1.0%
	$\sum_{0:p}$	-0.155	65.3%	22.4%	-0.877	81.2%	6.2%	-0.598	73.2%	2.1%
$Spread_{i,t}$	Lag 0	-5.503	59.2%	10.2%	-2.795	47.9%	14.6%	-0.791	39.2%	19.6%
	$\sum_{1:p}$	6.498	22.4%	40.8%	-10.476	35.4%	16.7%	-7.859	38.1%	8.2%
	$\sum_{0:p}$	3.371	26.5%	36.7%	-15.003	43.8%	12.5%	-7.260	51.5%	6.2%
$Depth_{i,t}$	Lag 0	53.812	12.2%	83.7%	24.370	16.7%	60.4%	21.503	4.1%	67.0%
	$\sum_{1:p}$	-61.880	77.6%	10.2%	-23.904	52.1%	16.7%	-22.636	64.9%	4.1%
	$\sum_{0:p}$	-5.628	30.6%	12.2%	0.106	10.4%	22.9%	0.321	18.6%	15.5%
$Slope5_{i,t}$	Lag 0	0.005	34.7%	49.0%	0.004	35.4%	41.7%	0.148	16.5%	51.5%
	$\sum_{1:p}$	-0.035	44.9%	20.4%	-0.185	58.3%	2.1%	-0.490	76.3%	2.1%
	$\sum_{0:p}$	-0.033	42.9%	24.5%	-0.274	60.4%	10.4%	-0.477	75.3%	0.0%
adj. R^2		0.185	-	-	0.203	-	-	0.204	-	-
Panel B: 20 best bid/ask price levels										
$v_{i,t}$	Lag 0	9.454	10.2%	75.5%	8.260	8.3%	68.8%	5.721	0.0%	67.0%
	$\sum_{1:p}$	-0.816	30.6%	12.2%	0.534	14.6%	18.8%	0.634	8.2%	14.4%
	$\sum_{0:p}$	8.180	18.4%	65.3%	10.000	6.2%	58.3%	6.268	0.0%	55.7%
$v_{i,t}Spread_{i,t}$	Lag 0	-11.023	51.0%	16.3%	-2.512	31.2%	18.8%	-1.513	30.9%	13.4%
	$\sum_{1:p}$	10.137	10.2%	42.9%	0.722	12.5%	16.7%	0.151	6.2%	9.3%
	$\sum_{0:p}$	-4.340	32.7%	28.6%	-2.591	16.7%	20.8%	-1.027	21.6%	10.3%
$v_{i,t}Depth_{i,t}$	Lag 0	-6.680	40.8%	30.6%	-5.688	52.1%	12.5%	-3.294	53.6%	4.1%
	$\sum_{1:p}$	-10.074	38.8%	4.1%	-2.565	33.3%	4.2%	-1.100	21.6%	2.1%
	$\sum_{0:p}$	-17.178	44.9%	18.4%	-6.297	47.9%	8.3%	-4.809	54.6%	2.1%
$v_{i,t}Slope20_{i,t}$	Lag 0	-0.596	73.5%	18.4%	-1.112	81.2%	6.2%	-0.960	63.9%	4.1%
	$\sum_{1:p}$	0.028	14.3%	28.6%	-0.064	14.6%	12.5%	-0.075	16.5%	5.2%
	$\sum_{0:p}$	-0.527	65.3%	22.4%	-1.157	60.4%	10.4%	-1.081	55.7%	2.1%
$Spread_{i,t}$	Lag 0	-5.102	59.2%	8.2%	-1.950	43.8%	16.7%	0.890	26.8%	25.8%
	$\sum_{1:p}$	13.180	12.2%	44.9%	1.271	20.8%	20.8%	-1.556	22.7%	15.5%
	$\sum_{0:p}$	7.863	14.3%	38.8%	-1.634	27.1%	18.8%	-0.389	29.9%	16.5%
$Depth_{i,t}$	Lag 0	53.733	12.2%	83.7%	25.332	20.8%	60.4%	22.235	4.1%	66.0%
	$\sum_{1:p}$	-62.284	77.6%	8.2%	-29.216	56.2%	14.6%	-27.466	69.1%	4.1%
	$\sum_{0:p}$	-8.944	38.8%	4.1%	-2.518	33.3%	12.5%	-2.323	35.1%	3.1%
$Slope20_{i,t}$	Lag 0	0.021	34.7%	51.0%	0.024	31.2%	43.8%	0.747	7.2%	60.8%
	$\sum_{1:p}$	-0.087	32.7%	20.4%	-0.344	50.0%	8.3%	-1.294	66.0%	5.2%
	$\sum_{0:p}$	-0.064	32.7%	24.5%	-0.544	45.8%	10.4%	-0.592	50.5%	1.0%
adj. R^2		0.181	-	-	0.199	-	-	0.201	-	-

Continued on next page

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A and B is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{5} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{5} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\hat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\hat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t}$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. Panel A reports the results when $x_{i,t} = (Spread_{i,t}, Depth_{i,t}, Slope5_{i,t})'$, and Panel B reports the results when $x_{i,t} = (Spread_{i,t}, Depth_{i,t}, Slope20_{i,t})'$, where $Slope5_{i,t}$ ($Slope20_{i,t}$) is the slope of the LOB, calculated using 5 (20) best bid and ask price levels, right before the t-th trade. See Table A.1 and the notes of Table 2 for the definitions of other variables and notation. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ from the volatility equation only, but a complete version of this table is available upon request.

Table A.3: LOB and the *endogenous dynamic* volume-volatility relation: Bid vs. Ask sides

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: 5 best bid/ask price levels										
$v_{i,t}$	Lag 0	30.468	8.2%	87.8%	36.589	6.2%	93.8%	17.338	0.0%	88.7%
	$\sum_{1:p}$	1.273	26.5%	26.5%	1.443	14.6%	25.0%	0.834	3.1%	11.3%
	$\sum_{0:p}$	29.596	12.2%	81.6%	35.732	6.2%	93.8%	18.050	0.0%	87.6%
$v_{i,t}Spread_{i,t}$	Lag 0	-99.187	83.7%	8.2%	-52.642	85.4%	6.2%	-17.464	76.3%	4.1%
	$\sum_{1:p}$	4.741	18.4%	30.6%	-0.555	16.7%	16.7%	-0.309	8.2%	6.2%
	$\sum_{0:p}$	-93.647	73.5%	16.3%	-53.336	79.2%	6.2%	-16.600	67.0%	2.1%
$v_{i,t}BV_{i,t}S_{i,t}$	Lag 0	65.116	16.3%	61.2%	44.539	8.3%	68.8%	5.124	13.4%	44.3%
	$\sum_{1:p}$	-17.980	30.6%	0.0%	-9.775	16.7%	4.2%	-4.124	29.9%	0.0%
	$\sum_{0:p}$	56.671	18.4%	51.0%	27.645	8.3%	58.3%	1.983	15.5%	27.8%
$v_{i,t}AV_{i,t}B_{i,t}$	Lag 0	80.376	20.4%	63.3%	38.416	6.2%	62.5%	5.807	12.4%	39.2%
	$\sum_{1:p}$	-14.712	20.4%	0.0%	-7.363	20.8%	6.2%	-2.851	14.4%	0.0%
	$\sum_{0:p}$	52.582	20.4%	46.9%	26.468	0.0%	47.9%	2.981	12.4%	26.8%
$v_{i,t}BidSlope5_{i,t}S_{i,t}$	Lag 0	-0.407	85.7%	6.1%	-1.737	93.8%	6.2%	-1.392	91.8%	1.0%
	$\sum_{1:p}$	0.016	8.2%	42.9%	0.055	6.2%	20.8%	0.092	1.0%	21.6%
	$\sum_{0:p}$	-0.402	79.6%	16.3%	-1.699	87.5%	6.2%	-1.109	82.5%	1.0%
$v_{i,t}AskSlope5_{i,t}B_{i,t}$	Lag 0	-0.456	87.8%	4.1%	-1.532	93.8%	4.2%	-1.306	92.8%	0.0%
	$\sum_{1:p}$	0.009	8.2%	34.7%	0.034	2.1%	29.2%	0.052	2.1%	16.5%
	$\sum_{0:p}$	-0.424	81.6%	12.2%	-1.375	91.7%	6.2%	-1.147	85.6%	1.0%
$Spread_{i,t}$	Lag 0	-6.064	69.4%	6.1%	-7.006	58.3%	12.5%	-4.241	48.5%	14.4%
	$\sum_{1:p}$	3.794	14.3%	36.7%	-1.139	25.0%	18.8%	-0.479	13.4%	14.4%
	$\sum_{0:p}$	-3.643	24.5%	32.7%	-12.190	41.7%	14.6%	-5.701	36.1%	7.2%
$BV_{i,t}S_{i,t}$	Lag 0	57.307	10.2%	79.6%	15.224	25.0%	56.2%	14.002	5.2%	54.6%
	$\sum_{1:p}$	-62.426	61.2%	4.1%	-21.458	39.6%	6.2%	-17.467	46.4%	4.1%
	$\sum_{0:p}$	-11.762	24.5%	0.0%	-0.359	8.3%	14.6%	-0.351	16.5%	10.3%
$AV_{i,t}B_{i,t}$	Lag 0	59.425	6.1%	83.7%	24.629	10.4%	60.4%	20.368	4.1%	60.8%
	$\sum_{1:p}$	-52.693	63.3%	0.0%	-17.531	45.8%	6.2%	-13.444	48.5%	2.1%
	$\sum_{0:p}$	-4.515	24.5%	4.1%	0.758	4.2%	12.5%	0.447	12.4%	14.4%
$BidSlope5_{i,t}S_{i,t}$	Lag 0	-0.011	55.1%	26.5%	-0.060	52.1%	25.0%	-0.023	35.1%	24.7%
	$\sum_{1:p}$	0.009	12.2%	38.8%	-0.010	14.6%	10.4%	-0.068	23.7%	3.1%
	$\sum_{0:p}$	-0.004	20.4%	30.6%	-0.120	22.9%	6.2%	-0.098	28.9%	0.0%
$AskSlope5_{i,t}B_{i,t}$	Lag 0	-0.013	55.1%	34.7%	-0.032	50.0%	27.1%	0.007	26.8%	25.8%
	$\sum_{1:p}$	-0.002	16.3%	26.5%	-0.015	20.8%	10.4%	-0.087	29.9%	1.0%
	$\sum_{0:p}$	-0.017	32.7%	22.4%	-0.114	31.2%	8.3%	-0.116	35.1%	2.1%
adj. R^2		0.288	-	-	0.314	-	-	0.294	-	-
Panel B: 20 best bid/ask price levels										
$v_{i,t}$	Lag 0	25.639	8.2%	87.8%	14.567	8.3%	85.4%	10.238	0.0%	78.4%
	$\sum_{1:p}$	-0.308	28.6%	16.3%	0.773	12.5%	16.7%	0.484	4.1%	11.3%
	$\sum_{0:p}$	25.776	12.2%	75.5%	16.836	6.2%	81.2%	10.073	0.0%	75.3%
$v_{i,t}Spread_{i,t}$	Lag 0	-70.844	77.6%	10.2%	-11.350	54.2%	10.4%	-6.832	49.5%	11.3%
	$\sum_{1:p}$	6.504	10.2%	32.7%	1.790	12.5%	14.6%	0.409	7.2%	10.3%
	$\sum_{0:p}$	-54.066	69.4%	18.4%	-10.366	43.8%	10.4%	-4.508	39.2%	7.2%
$v_{i,t}BV_{i,t}S_{i,t}$	Lag 0	39.038	22.4%	51.0%	-1.402	35.4%	29.2%	-7.100	52.6%	6.2%
	$\sum_{1:p}$	-18.042	32.7%	2.0%	-6.790	20.8%	4.2%	-3.085	21.6%	0.0%
	$\sum_{0:p}$	12.826	24.5%	40.8%	-5.186	35.4%	16.7%	-12.292	49.5%	3.1%
$v_{i,t}AV_{i,t}B_{i,t}$	Lag 0	43.981	26.5%	53.1%	0.258	37.5%	29.2%	-4.797	41.2%	10.3%
	$\sum_{1:p}$	-13.644	24.5%	2.0%	-7.017	25.0%	6.2%	-1.576	10.3%	2.1%
	$\sum_{0:p}$	35.467	26.5%	34.7%	0.133	31.2%	22.9%	-5.220	43.3%	7.2%
$v_{i,t}BidSlope20_{i,t}S_{i,t}$	Lag 0	-1.588	85.7%	6.1%	-2.259	87.5%	6.2%	-1.871	84.5%	1.0%
	$\sum_{1:p}$	0.066	8.2%	44.9%	0.123	4.2%	18.8%	0.158	3.1%	13.4%
	$\sum_{0:p}$	-1.500	79.6%	16.3%	-2.421	79.2%	6.2%	-1.713	73.2%	1.0%
$v_{i,t}AskSlope20_{i,t}B_{i,t}$	Lag 0	-1.580	87.8%	4.1%	-2.529	89.6%	6.2%	-2.210	82.5%	2.1%
	$\sum_{1:p}$	0.044	8.2%	34.7%	0.083	6.2%	16.7%	0.015	4.1%	8.2%
	$\sum_{0:p}$	-1.459	81.6%	12.2%	-2.457	83.3%	6.2%	-2.041	73.2%	2.1%

Continued on next page

Table A.3 – *continued from previous page*

		Large cap (49 stocks)			**Mid cap** (48 stocks)			**Small cap** (97 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
$Spread_{i,t}$	Lag 0	-5.730	65.3%	4.1%	-3.363	47.9%	12.5%	-1.834	40.2%	16.5%
	$\sum_{1:p}$	13.154	8.2%	42.9%	2.147	16.7%	20.8%	1.308	10.3%	19.6%
	$\sum_{0:p}$	2.361	20.4%	36.7%	0.135	27.1%	16.7%	-0.556	19.6%	19.6%
$BV_{i,t}S_{i,t}$	Lag 0	57.306	8.2%	79.6%	12.443	27.1%	56.2%	13.834	5.2%	53.6%
	$\sum_{1:p}$	-62.679	63.3%	4.1%	-24.822	45.8%	10.4%	-18.149	53.6%	2.1%
	$\sum_{0:p}$	-13.791	26.5%	0.0%	-1.664	16.7%	8.3%	-2.397	20.6%	6.2%
$AV_{i,t}B_{i,t}$	Lag 0	59.429	4.1%	83.7%	23.756	10.4%	58.3%	19.773	3.1%	61.9%
	$\sum_{1:p}$	-53.466	69.4%	2.0%	-16.437	43.8%	2.1%	-18.226	53.6%	3.1%
	$\sum_{0:p}$	-6.722	30.6%	6.1%	-0.679	10.4%	16.7%	1.156	13.4%	15.5%
$BidSlope20_{i,t}S_{i,t}$	Lag 0	-0.039	53.1%	26.5%	-0.258	52.1%	22.9%	-0.098	32.0%	23.7%
	$\sum_{1:p}$	0.044	10.2%	36.7%	0.010	14.6%	14.6%	-0.115	23.7%	9.3%
	$\sum_{0:p}$	-0.003	20.4%	30.6%	-0.380	31.2%	4.2%	-0.271	35.1%	0.0%
$AskSlope20_{i,t}B_{i,t}$	Lag 0	-0.052	53.1%	36.7%	-0.127	52.1%	27.1%	0.011	22.7%	28.9%
	$\sum_{1:p}$	0.030	12.2%	30.6%	-0.031	18.8%	4.2%	-0.462	44.3%	3.1%
	$\sum_{0:p}$	-0.023	28.6%	22.4%	-0.366	39.6%	4.2%	-0.505	51.5%	4.1%
adj. R^2		0.280	-	-	0.286	-	-	0.276	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A and B is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{5} [\beta_{0,k} + \delta'_k x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{5} \gamma'_k x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t}$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. Panel A reports the results when $x_{i,t} = (Spread_{i,t}, BV_{i,t}B_{i,t}, BV_{i,t}S_{i,t}, AV_{i,t}B_{i,t}, AV_{i,t}S_{i,t}, BidSlope5_{i,t}B_{i,t}, BidSlope5_{i,t}S_{i,t}, AskSlope5_{i,t}B_{i,t}, AskSlope5_{i,t}S_{i,t})'$; Panel B reports the results when $x_{i,t} = (Spread_{i,t}, BV_{i,t}B_{i,t}, BV_{i,t}S_{i,t}, AV_{i,t}B_{i,t}, AV_{i,t}S_{i,t}, BidSlope20_{i,t}B_{i,t}, BidSlope20_{i,t}S_{i,t}, AskSlope20_{i,t}B_{i,t}, AskSlope20_{i,t}S_{i,t})'$, where $BidSlope5_{i,t}$ ($AskSlope5_{i,t}$, $BidSlope20_{i,t}$, $AskSlope20_{i,t}$) is the slope of the bid (ask, bid, ask) side of the LOB, calculated using 5 (5, 20, 20) best bid (ask, bid, ask) price levels, right before the t-th trade. See Table A.1 and the notes of Table 2 for the definitions of other variables and notation. This table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ of the order book that are of the opposite side to the direction of a trade (e.g. $v_{i,t}AskSlope_{i,t}B_{i,t}$ and $AskSlope_{i,t}B_{i,t}$) from the volatility equation only, but a complete version of this table is available upon request.

Table A.4: LOB and the *endogenous dynamic* volume-volatility relation: $0.5^{\text{th}} - 99.5^{\text{th}}$ winsorization

		Large cap (49 stocks)			**Mid cap** (50 stocks)			**Small cap** (98 stocks)		
		Med	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$	**Med**	$\%_{-5\%}$	$\%_{+5\%}$
Panel A: Combined limit order book										
$v_{i,t}$	Lag 0	6.712	10.2%	71.4%	11.514	8.0%	74.0%	7.276	1.0%	72.4%
	$\sum_{1:p}$	0.700	26.5%	20.4%	2.838	12.0%	20.0%	1.863	4.1%	20.4%
	$\sum_{0:p}$	7.360	16.3%	55.1%	15.619	8.0%	72.0%	8.760	0.0%	65.3%
$v_{i,t}Spread_{i,t}$	Lag 0	-10.180	46.9%	14.3%	-11.612	54.0%	8.0%	-4.296	38.8%	6.1%
	$\sum_{1:p}$	5.977	14.3%	38.8%	-1.483	18.0%	12.0%	-1.167	12.2%	5.1%
	$\sum_{0:p}$	-0.829	34.7%	22.4%	-11.986	42.0%	10.0%	-4.825	36.7%	2.0%
$v_{i,t}Depth_{i,t}$	Lag 0	-2.638	34.7%	28.6%	0.023	24.0%	26.0%	-1.518	36.7%	15.3%
	$\sum_{1:p}$	-5.677	30.6%	4.1%	-0.903	16.0%	8.0%	-0.462	15.3%	0.0%
	$\sum_{0:p}$	-4.775	36.7%	10.2%	-2.678	20.0%	18.0%	-1.592	35.7%	8.2%
$v_{i,t}Slope_{i,t}$	Lag 0	-0.245	67.3%	18.4%	-0.948	84.0%	6.0%	-0.715	69.4%	4.1%
	$\sum_{1:p}$	-0.005	24.5%	28.6%	-0.159	24.0%	8.0%	-0.169	19.4%	1.0%
	$\sum_{0:p}$	-0.267	59.2%	18.4%	-1.185	78.0%	8.0%	-0.883	62.2%	2.0%
$Spread_{i,t}$	Lag 0	-6.439	63.3%	8.2%	-2.256	42.0%	16.0%	-0.362	32.7%	19.4%
	$\sum_{1:p}$	8.843	24.5%	38.8%	-9.753	32.0%	16.0%	-6.381	37.8%	9.2%
	$\sum_{0:p}$	5.251	30.6%	36.7%	-13.720	46.0%	14.0%	-7.332	49.0%	9.2%
$Depth_{i,t}$	Lag 0	57.991	14.3%	83.7%	22.060	20.0%	58.0%	19.503	4.1%	63.3%
	$\sum_{1:p}$	-60.984	81.6%	12.2%	-16.504	54.0%	18.0%	-21.894	65.3%	6.1%
	$\sum_{0:p}$	-5.381	30.6%	10.2%	0.027	8.0%	24.0%	0.079	19.4%	10.2%
$Slope_{i,t}$	Lag 0	0.011	34.7%	53.1%	-0.027	34.0%	42.0%	0.397	15.3%	56.1%
	$\sum_{1:p}$	-0.089	49.0%	4.1%	-0.469	60.0%	2.0%	-1.143	78.6%	3.1%
	$\sum_{0:p}$	-0.062	40.8%	24.5%	-0.733	72.0%	6.0%	-0.938	79.6%	0.0%
adj. R^2		0.183	-	-	0.204	-	-	0.201	-	-
Panel B: Bid vs. Ask sides										
$v_{i,t}$	Lag 0	24.370	8.2%	85.7%	26.887	6.0%	86.0%	14.009	0.0%	87.8%
	$\sum_{1:p}$	0.160	22.4%	20.4%	0.963	10.0%	26.0%	0.797	4.1%	8.2%
	$\sum_{0:p}$	23.726	12.2%	79.6%	29.477	6.0%	86.0%	15.649	0.0%	85.7%
$v_{i,t}Spread_{i,t}$	Lag 0	-68.566	81.6%	8.2%	-38.571	82.0%	6.0%	-13.449	70.4%	4.1%
	$\sum_{1:p}$	5.175	10.2%	30.6%	0.936	18.0%	12.0%	0.129	6.1%	5.1%
	$\sum_{0:p}$	-70.439	63.3%	16.3%	-40.211	76.0%	6.0%	-12.769	60.2%	1.0%
$v_{i,t}BV_{i,t}S_{i,t}$	Lag 0	49.395	18.4%	57.1%	26.027	12.0%	62.0%	1.476	19.4%	29.6%
	$\sum_{1:p}$	-14.230	28.6%	0.0%	-4.620	16.0%	4.0%	-3.272	25.5%	0.0%
	$\sum_{0:p}$	33.308	16.3%	40.8%	16.441	12.0%	46.0%	-0.570	20.4%	22.4%
$v_{i,t}AV_{i,t}B_{i,t}$	Lag 0	53.901	18.4%	63.3%	19.920	10.0%	58.0%	3.758	14.3%	30.6%
	$\sum_{1:p}$	-9.068	20.4%	2.0%	-2.448	20.0%	6.0%	-2.613	15.3%	1.0%
	$\sum_{0:p}$	44.352	22.4%	44.9%	11.282	8.0%	44.0%	0.642	19.4%	17.3%
$v_{i,t}BidSlope_{i,t}S_{i,t}$	Lag 0	-0.680	85.7%	6.1%	-2.173	90.0%	8.0%	-1.967	83.7%	1.0%
	$\sum_{1:p}$	0.027	6.1%	40.8%	0.042	6.0%	24.0%	0.163	0.0%	22.4%
	$\sum_{0:p}$	-0.639	81.6%	14.3%	-1.976	82.0%	6.0%	-1.624	72.4%	1.0%
$v_{i,t}AskSlope_{i,t}B_{i,t}$	Lag 0	-0.840	87.8%	4.1%	-2.460	90.0%	4.0%	-1.768	85.7%	0.0%
	$\sum_{1:p}$	0.018	10.2%	28.6%	0.025	2.0%	22.0%	0.155	1.0%	14.3%
	$\sum_{0:p}$	-0.771	81.6%	12.2%	-2.120	86.0%	6.0%	-1.561	75.5%	2.0%
$Spread_{i,t}$	Lag 0	-6.476	67.3%	2.0%	-9.595	66.0%	8.0%	-5.380	52.0%	10.2%
	$\sum_{1:p}$	1.888	8.2%	40.8%	2.691	14.0%	28.0%	1.769	13.3%	19.4%
	$\sum_{0:p}$	-3.832	26.5%	36.7%	-11.362	42.0%	14.0%	-5.559	38.8%	12.2%
$BV_{i,t}S_{i,t}$	Lag 0	56.916	12.2%	79.6%	11.378	24.0%	52.0%	13.560	7.1%	53.1%
	$\sum_{1:p}$	-63.428	65.3%	6.1%	-15.658	42.0%	12.0%	-14.754	46.9%	5.1%
	$\sum_{0:p}$	-11.321	24.5%	4.1%	3.366	12.0%	14.0%	0.029	15.3%	6.1%
$AV_{i,t}B_{i,t}$	Lag 0	59.774	6.1%	83.7%	22.384	12.0%	60.0%	19.624	4.1%	61.2%
	$\sum_{1:p}$	-50.959	73.5%	0.0%	-13.017	44.0%	8.0%	-17.898	57.1%	3.1%
	$\sum_{0:p}$	-4.421	20.4%	6.1%	4.850	8.0%	12.0%	0.243	14.3%	15.3%

Continued on next page

Table A.4 – *continued from previous page*

$BidSlope_{i,t}S_{i,t}$	Lag 0	-0.016	49.0%	32.7%	-0.056	42.0%	30.0%	0.028	25.5%	36.7%
	$\sum_{1:p}$	0.009	8.2%	32.7%	-0.081	20.0%	6.0%	-0.257	24.5%	2.0%
	$\sum_{0:p}$	-0.015	24.5%	30.6%	-0.292	30.0%	6.0%	-0.264	31.6%	0.0%
$AskSlope_{i,t}B_{i,t}$	Lag 0	-0.018	51.0%	36.7%	-0.057	44.0%	24.0%	0.055	16.3%	37.8%
	$\sum_{1:p}$	-0.019	18.4%	18.4%	-0.072	34.0%	4.0%	-0.288	35.7%	2.0%
	$\sum_{0:p}$	-0.032	36.7%	22.4%	-0.250	36.0%	8.0%	-0.288	31.6%	3.1%
adj. R^2		0.280	-	-	0.308	-	-	0.283	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A and B is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{5} [\beta_{0,k} + \delta_k' x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{5} \gamma_k' x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t}$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. Panel A reports the results when $x_{i,t} = (Spread_{i,t}, Depth_{i,t}, Slope_{i,t})'$, and Panel B reports the results when $x_{i,t} = (Spread_{i,t}, BV_{i,t}B_{i,t}, BV_{i,t}S_{i,t}, AV_{i,t}B_{i,t}, AV_{i,t}S_{i,t}, BidSlope_{i,t}B_{i,t}, BidSlope_{i,t}S_{i,t}, AskSlope_{i,t}B_{i,t}, AskSlope_{i,t}S_{i,t})'$. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. Panel B of this table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ of the order book that are of the opposite side to the direction of a trade (e.g. $v_{i,t}AskSlope_{i,t}B_{i,t}$ and $AskSlope_{i,t}B_{i,t}$) from the volatility equation only, but a complete version of this table is available upon request.

Table A.5: LOB and the *endogenous dynamic* volume-volatility relation: $2^{\mathrm{nd}} - 98^{\mathrm{th}}$ winsorization

		Large cap (48 stocks)			**Mid cap** (46 stocks)			**Small cap** (93 stocks)		
		Med	$\%-_{5\%}$	$\%+_{5\%}$	**Med**	$\%-_{5\%}$	$\%+_{5\%}$	**Med**	$\%-_{5\%}$	$\%+_{5\%}$
Panel A: Combined limit order book										
$v_{i,t}$	Lag 0	23.064	12.5%	81.2%	29.335	6.5%	93.5%	16.189	0.0%	90.3%
	$\sum_{1:p}$	-2.133	45.8%	25.0%	1.034	17.4%	21.7%	1.220	5.4%	15.1%
	$\sum_{0:p}$	20.215	20.8%	66.7%	32.854	6.5%	89.1%	17.868	0.0%	80.6%
$v_{i,t}Spread_{i,t}$	Lag 0	-44.176	68.8%	16.7%	-35.335	87.0%	8.7%	-14.876	71.0%	2.2%
	$\sum_{1:p}$	18.447	10.4%	50.0%	0.696	13.0%	17.4%	-0.230	7.5%	7.5%
	$\sum_{0:p}$	-37.933	52.1%	25.0%	-33.733	67.4%	8.7%	-14.541	57.0%	1.1%
$v_{i,t}Depth_{i,t}$	Lag 0	6.591	29.2%	50.0%	6.578	23.9%	45.7%	-0.006	33.3%	25.8%
	$\sum_{1:p}$	-12.804	47.9%	0.0%	-3.194	28.3%	10.9%	-0.854	11.8%	1.1%
	$\sum_{0:p}$	-2.281	39.6%	29.2%	1.570	23.9%	26.1%	-0.854	31.2%	18.3%
$v_{i,t}Slope_{i,t}$	Lag 0	-0.625	79.2%	12.5%	-2.223	93.5%	6.5%	-1.402	86.0%	2.2%
	$\sum_{1:p}$	0.022	25.0%	47.9%	-0.035	17.4%	17.4%	-0.076	15.1%	6.5%
	$\sum_{0:p}$	-0.441	68.8%	20.8%	-2.492	87.0%	6.5%	-1.825	79.6%	1.1%
$Spread_{i,t}$	Lag 0	-5.461	66.7%	4.2%	-4.639	52.2%	8.7%	-1.428	41.9%	20.4%
	$\sum_{1:p}$	20.453	16.7%	50.0%	-7.943	34.8%	17.4%	-6.250	35.5%	8.6%
	$\sum_{0:p}$	14.913	25.0%	41.7%	-12.604	43.5%	17.4%	-6.849	41.9%	5.4%
$Depth_{i,t}$	Lag 0	53.248	10.4%	83.3%	25.476	21.7%	63.0%	22.611	3.2%	73.1%
	$\sum_{1:p}$	-68.405	79.2%	12.5%	-27.482	58.7%	13.0%	-24.085	68.8%	3.2%
	$\sum_{0:p}$	-12.011	45.8%	6.2%	-1.425	10.9%	21.7%	-0.061	16.1%	10.8%
$Slope_{i,t}$	Lag 0	0.008	35.4%	50.0%	-0.021	34.8%	41.3%	0.235	15.1%	53.8%
	$\sum_{1:p}$	0.009	29.2%	33.3%	-0.251	41.3%	6.5%	-0.886	67.7%	3.2%
	$\sum_{0:p}$	0.013	37.5%	37.5%	-0.523	52.2%	15.2%	-0.726	66.7%	1.1%
adj. R^2		0.191	-	-	0.206	-	-	0.210	-	-
Panel B: Bid vs. Ask sides										
$v_{i,t}$	Lag 0	44.845	8.3%	85.4%	47.101	6.5%	93.5%	22.920	0.0%	93.5%
	$\sum_{1:p}$	1.012	31.2%	16.7%	1.135	17.4%	17.4%	0.446	6.5%	9.7%
	$\sum_{0:p}$	46.745	16.7%	81.2%	44.248	6.5%	93.5%	21.951	0.0%	91.4%
$v_{i,t}Spread_{i,t}$	Lag 0	-107.285	83.3%	12.5%	-64.169	91.3%	6.5%	-23.020	81.7%	2.2%
	$\sum_{1:p}$	5.688	12.5%	33.3%	-0.116	15.2%	17.4%	0.262	6.5%	7.5%
	$\sum_{0:p}$	-106.186	81.2%	16.7%	-58.882	84.8%	6.5%	-21.771	71.0%	1.1%
$v_{i,t}BV_{i,t}S_{i,t}$	Lag 0	88.714	18.8%	66.7%	70.362	10.9%	76.1%	12.739	9.7%	49.5%
	$\sum_{1:p}$	-25.104	37.5%	0.0%	-17.585	30.4%	2.2%	-6.438	29.0%	0.0%
	$\sum_{0:p}$	64.403	16.7%	56.2%	50.435	10.9%	60.9%	6.339	14.0%	30.1%
$v_{i,t}AV_{i,t}B_{i,t}$	Lag 0	113.426	20.8%	72.9%	63.300	6.5%	78.3%	14.764	9.7%	47.3%
	$\sum_{1:p}$	-16.346	29.2%	0.0%	-10.552	26.1%	6.5%	-5.335	19.4%	0.0%
	$\sum_{0:p}$	88.396	22.9%	62.5%	48.395	2.2%	63.0%	8.092	8.6%	24.7%
$v_{i,t}BidSlope_{i,t}S_{i,t}$	Lag 0	-1.324	85.4%	8.3%	-4.290	93.5%	6.5%	-3.227	93.5%	1.1%
	$\sum_{1:p}$	0.075	4.2%	45.8%	0.292	2.2%	34.8%	0.273	0.0%	26.9%
	$\sum_{0:p}$	-1.175	79.2%	14.6%	-3.899	89.1%	6.5%	-2.806	87.1%	1.1%
$v_{i,t}AskSlope_{i,t}B_{i,t}$	Lag 0	-1.397	87.5%	6.2%	-4.014	93.5%	6.5%	-3.435	96.8%	0.0%
	$\sum_{1:p}$	0.037	6.2%	43.8%	0.141	0.0%	37.0%	0.196	1.1%	21.5%
	$\sum_{0:p}$	-1.396	83.3%	14.6%	-3.724	93.5%	6.5%	-3.040	90.3%	1.1%
$Spread_{i,t}$	Lag 0	-3.899	52.1%	16.7%	-5.595	47.8%	13.0%	-2.237	44.1%	12.9%
	$\sum_{1:p}$	5.643	18.8%	39.6%	-0.145	17.4%	21.7%	-1.663	18.3%	17.2%
	$\sum_{0:p}$	-0.891	20.8%	35.4%	-10.563	28.3%	17.4%	-3.724	29.0%	6.5%
$BV_{i,t}S_{i,t}$	Lag 0	59.239	12.5%	77.1%	18.286	21.7%	60.9%	15.339	6.5%	55.9%
	$\sum_{1:p}$	-68.301	62.5%	2.1%	-28.137	43.5%	4.3%	-18.968	51.6%	3.2%
	$\sum_{0:p}$	-19.035	33.3%	0.0%	-6.996	8.7%	6.5%	-1.476	15.1%	8.6%
$AV_{i,t}B_{i,t}$	Lag 0	60.950	2.1%	89.6%	23.818	8.7%	65.2%	21.373	3.2%	65.6%
	$\sum_{1:p}$	-46.679	54.2%	0.0%	-21.144	47.8%	2.2%	-19.823	59.1%	2.2%
	$\sum_{0:p}$	-7.451	20.8%	4.2%	-0.043	6.5%	6.5%	-0.289	10.8%	6.5%

Continued on next page

Table A.5 – *continued from previous page*

$BidSlope_{i,t}S_{i,t}$	Lag 0	-0.043	72.9%	12.5%	-0.150	69.6%	17.4%	-0.157	49.5%	7.5%
	$\sum_{1:p}$	0.066	6.2%	45.8%	0.165	4.3%	13.0%	0.000	8.6%	5.4%
	$\sum_{0:p}$	0.022	14.6%	35.4%	-0.025	13.0%	8.7%	-0.159	22.6%	0.0%
$AskSlope_{i,t}B_{i,t}$	Lag 0	-0.039	60.4%	27.1%	-0.208	69.6%	15.2%	-0.077	44.1%	14.0%
	$\sum_{1:p}$	0.026	10.4%	37.5%	0.036	4.3%	28.3%	-0.057	16.1%	7.5%
	$\sum_{0:p}$	-0.019	20.8%	31.2%	-0.047	23.9%	13.0%	-0.239	29.0%	2.2%
adj. R^2		0.296	-	-	0.324	-	-	0.309	-	-

This table reports summary estimation results for all stocks of the S&P/ASX200 index over *Jul-Dec 2014*. The estimated model in Panels A and B is

$$\sigma_{i,t} = \alpha_0 + \mu_1 Monday_{i,t} + \mu_2 hour1_{i,t} + \sum_{j=1}^{12} \alpha_j \sigma_{i,t-j} + \sum_{k=0}^{5} [\beta_{0,k} + \delta'_k x_{i,t-k}] v_{i,t-k} + \sum_{k=0}^{5} \gamma'_k x_{i,t-k} + \pi' y_{i,t} + \eta_{i,t},$$

where $\sigma_{i,t}$ is a proxy for return volatility *per unit of time* of the t-th transaction in stock i, which is estimated as the absolute value of the residual $|\widehat{\epsilon}_{i,t}|$ of the following autoregressive model of returns $r_{i,t}$ divided by the duration $T_{i,t}$ of the trade (i.e. $\sigma_{i,t} = |\widehat{\epsilon}_{i,t}|/T_{i,t}$):

$$r_{i,t} = \sum_{k=1}^{5} \psi_{i,k} Day_{k,i,t} + \sum_{k=1}^{5} \phi_{i,k} hour_{k,i,t} + \sum_{k=1}^{12} \rho_{i,k} r_{i,t-k} + \varphi' x_{i,t} + \lambda' y_{i,t} + \epsilon_{i,t}.$$

$x_{i,t}$ is a vector of potential predictors of the volume-volatility relation. $y_{i,t} = (\ln(T_{i,t}), N_{i,t}, ATS_{i,t}, OIB_{i,t}, QTT_{i,t})'$ is a vector of control variables that allow for the effects of the order flow prior to a trade. Panel A reports the results when $x_{i,t} = (Spread_{i,t}, Depth_{i,t}, Slope_{i,t})'$, and Panel B reports the results when $x_{i,t} = (Spread_{i,t}, BV_{i,t}B_{i,t}, BV_{i,t}S_{i,t}, AV_{i,t}B_{i,t}, AV_{i,t}S_{i,t}, BidSlope_{i,t}B_{i,t}, BidSlope_{i,t}S_{i,t}, AskSlope_{i,t}B_{i,t}, AskSlope_{i,t}S_{i,t})'$. See Table A.1 and the notes of Table 2 for the definitions of the variables and other notation. Panel B of this table reports the coefficient estimates for $v_{i,t}$ and $x_{i,t}$ of the order book that are of the opposite side to the direction of a trade (e.g. $v_{i,t}AskSlope_{i,t}B_{i,t}$ and $AskSlope_{i,t}B_{i,t}$) from the volatility equation only, but a complete version of this table is available upon request.