Low latency trading and comovement of order flow and prices

EkkehartBoehmer
Ekkehart.boehmer@edhec.edu

R.L.Shankar
rlshankar@gmail.com

EDHEC Business School
One George Street #07-02
Singapore, 049145
Singapore

November 18, 2013

Preliminary and incomplete

Abstract

We examine the impact of low latency / high frequency trading in equities on measures of the market's susceptibility to systemic shocks, including commonalities in order flow, liquidity, and volatility. Using order-level data around a natural experiment at the National Stock Exchange of India, which in 2010 has introduced features that promote low latency trading, we find that more low latency trading reduces commonality in returns, liquidity, and volatility. These declines are more pronounced for algorithmic order flow and for large-cap firms.

Low latency trading and comovement of order flow and prices

November 18, 2013

Preliminary and incomplete

Abstract

We examine the impact of low latency / high frequency trading in equities on measures of the market's susceptibility to systemic shocks, including commonalities in order flow, liquidity, and volatility. Using order-level data around a natural experiment at the National Stock Exchange of India, which in 2010 has introduced features that promote low latency trading, we find that more low latency trading reduces commonality in returns, liquidity, and volatility. These declines are more pronounced for algorithmic order flow and for large-cap firms.

1. Introduction

Hasbrouck and Saar (2013) define latency as "the time it takes to learn about an event, generate a response and have the exchange act on the response". Interactions in financial markets, they note, happen increasingly in the millisecond environment. High frequency traders (HFT) specializing in low-latency strategies dominate message traffic and trading in most markets. This explosive growth in HFT activities has raised serious concerns about the effect of low latency and algorithmic trading (AT) on financial markets. Much of the extant literature on AT focusses on individual securities. In this study, we investigate the impact of latency on comovement in order flows and liquidity. Examining this question is important as it offers insights on whether faster trading makes markets more susceptible to systemic risk. If faster trading leads to greater commonality in trading strategies and perhaps in liquidity, then the associated increase in correlation of order flow and liquidity could magnify systemic shocks to liquidity and thus increase systemic risk in financial markets. Conversely, as Chordia et al. (2013) note in their review article, lower latency may merely encourage faster trading without fundamentally changing either the strategies employed by traders or the underlying economics of financial markets. This view suggests that lower latency might not have any impact on commonality in liquidity or order flows.

We provide the first direct evidence on the impact of low-latency trading on comovement in order flows and liquidity. We use a natural experiment at the National Stock Exchange (NSE) of India.[1] Direct Market Access (DMA) was introduced in Indian markets in April 2008. This allowed institutional clients to directly access the exchange's trading system using their brokers'

[1]World Federation of Exchanges.(2012) reports that NSE is the largest exchange globally when ranked by number of trades in equity shares. During 2011-12, a total of 1.4 billion trades were executed at NSE compared to 1.37 billion trades at NYSE Euronext (US) and 1.26 billion trades at NASDAQ OMX. In terms of value of shares traded, NSE is ranked lower at 27th.

infrastructure, but without their manual intervention. To reduce external latency, NSE introduced co-location services in January 2010. Subsequently, a broker could rent servers situated within NSE's premises. Our analysis focusses on this step, which is designed to support and facilitate low latency strategies. Market participants have eagerly embraced these innovations–within fifteen months of launching co-location facilities, 60% of incoming orders at NSE were from co-located servers.[2] We use order-level NSE data that cover certain periods before and after the introduction of colocation facilities. Our data identify the originator for each incoming message as either AT (algorithmic trader) or non-AT.

We formally define net order flow as the difference between marketable buy orders and marketable sell orders. These are orders whose limit price is set at or better than the opposite-side quoted price (orders to buy at or above the best ask and orders to sell at or below the best bid). As these orders demand immediacy, net order flow measures the prevailing directional imbalance. We find that orders emanating from AT have lower net order flow commonality than those emanating from non-AT. Introduction of colocation facilities leads to a significant reduction in order flow comovement for both AT and non-AT.

Order flow is naturally related to prices and liquidity, so we investigate how co-location affects the commonality in these variables. Commonality in returns, volatility and certain measures of liquidity experience a significant decline around colocation. This effect tends to be most pronounced for AT order flow and for larger-cap firms. Taken together, our findings are not consistent with the notion that more low latency trading increases systemic risk by accentuating comovement in order flows, returns, or liquidity.

[2] The changing landscape of India's equity markets, Live Mint, April 26, 2011}

A substantial literature studies commonality in liquidity and order flow. Hasbrouck and Seppi (2001) document the presence of a common factor underlying order flows; this factor explains more than two-thirds of commonality in returns. Harford and Kaul (2005) also find correlated stock order flow to be an important driver of comovement in returns and, to a lesser extent, commonality in liquidity. Intertemporal changes in liquidity tend to covary across financial assets (Chordia et al., 2000; Hasbrouck and Seppi, 2001; Huberman and Halka, 2001). Acharya and Pedersen (2005) develop a liquidity-adjusted capital asset pricing model in which such commonality in liquidity is a priced risk factor. Intuitively, investors would demand a higher return premium for assets whose liquidity displays a higher covariance with that of the market.

There is little theoretical guidance and virtually no empirical evidence on the impact of automation on common cross-firm variation in order flow or liquidity. Biais et al. (2013) advance an equilibrium model of trading that permits a continuum of fast traders. They postulate that an increase in fraction of fast traders leads to an increase in market impact of trades and a decrease in expected gains from trades. At the extreme, for high values of this fraction, slow traders might be evicted from the market. Trading protocols that support low-latency trading could thus lead to crowding out of slow traders and equivalently, higher concentration of trading among fast traders. This in turn increases the likelihood that a shock to capital (or information) would impact multiple securities, leading to higher commonality in liquidity and order flows. This hypothesis derives support from extant studies on sources of comovement. Coughenour and Saad (2004) document that a stock's liquidity co-moves with that of other stocks handled by the same specialist firm. Evidently, shared capital and information among specialists within a firm play a key role. Koch et al. (2012) find that stocks owned by mutual funds that themselves experience liquidity shocks have higher commonality in liquidity. Cespa and Foucault (2013) postulate that liquidity providers glean information about an asset from other assets. This feedback loop implies that any liquidity shock to a single stock could potentially lead to a large drop in market-wide liquidity. The impact of this feedback mechanism could be further amplified if liquidity provision gets concentrated in the hands of a small group of HFT market makers. Menkveld's (2013) examination of a single HFT market maker at Chi-X lends further support to this hypothesis. He finds that the HFT market maker typically chooses not to carry a significant inventory position; trades arising out of such inventory control have an impact on market prices. Alternatively, low latency trading can impact comovement in order flows through HFT's collective and correlated responses to macroeconomic shocks or other public signals, such as those derived from machine-readable news (see Jones, 2013).

Our study complements literature that examines the impact of faster trading on market quality. Using a change in market structure as an exogenous instrument, Hendershott et al. (2011) establish that AT improves market liquidity. Hendershott and Riordan (2012) show that ATs in Deutsche Boerse continuously monitor the market for liquidity and strategically act as either consumers or suppliers of liquidity. Chaboud et al. (2009) find that algorithmic trades in currency markets tend to correlated; however, they do not find any evidence that AT increases market volatility. Brogaard (2010) also finds no evidence to support the hypothesis that HFT activity increases volatility. Boehmer et al. (2012) provide cross-country evidence on impact of AT using data from 39 exchanges. They find that on average AT activity increases market liquidity, informational efficiency of prices and volatility; the last result is found to be robust to a wide range of volatility measures. Hasbrouck and Saar (2013) document that increase in low-latency trading activity lowers short-term volatility and quoted spreads. We show that order flow, liquidity, and volatility have common factors that appear to be driven by AT and are more pronounced by actual AT order flow.

Our work is also part of a growing empirical literature that examines comovement in intertemporal changes of order flows, returns and liquidity. Comovement in asset returns has been widely studied; Barberis et al. (2005) provide an excellent overview. Hasbrouck and Seppi (2001), Harford and Kaul (2005), and Corwin and Lipson (2011) discover dominant common factors underlying order flows; these factors are found to have a significant impact on comovement in asset returns. Commonality in liquidity has been documented by Chordia et al. (2000), Hasbrouck and Seppi (2001) and Huberman and Halka (2001). However, very little is known about factors driving comovement in liquidity. Extant studies attribute a key role to specialists (Coughenour and Saad, 2004) and institutional investors (Koch et al., 2012). Cespa and Foucault (2013) postulate that liquidity shocks to a single stock could potentially lead to a large drop in market-wide liquidity and accentuate commonality in liquidity.

In the context of AT, only Chaboud et al. (2009) and Huh (2011) look at commonality in the context of low latency or algorithmic trading. Chaboud et al. find that the correlation of strategies is greater among computerized traders than among humans, but their analysis is limited to the FX market. Huh examines the NYSE hybrid market, which, arguably, attracts new algorithmic traders, and finds that liquidity commonality increases around this event. Our experiment differs in that we use order-level data that more clearly identify trading intention than the NYSE trade-level data. Perhaps more importantly, our data also clearly identify which order messages come from algorithmic traders. Moreover, AT-related events on the Indian market may be easier to interpret than those at the NYSE. This is because the Indian equity market is represented by only two exchanges. Importantly, aside from typical order processing automation, neither one had AT-related features prior to the NSE co-location event. In contrast, the U.S. equity market is highly decentralized and almost every U.S. traded equity security could be traded algorithmically in a number of different markets. As such, the NYSE's hybrid transformation may not provide novel AT-related features to traders, and instead represent its catching up with technology. We believe that these differences across experiments make it interesting to analyze the resulting differences in how AT affects commonalities in order flow and market quality.

The rest of the paper is organized as follows. In Section 2, we provide an overview of Indian equity markets, discuss our data sources and identify the event periods. In Section 3, we discuss our research design and present results for commonality in order flow. In Section 4, we discuss comovement in returns and volatility. We elaborate on commonality in liquidity and liquidity imbalance in Section V. We conclude in the last section.

2. Overview of Data and Trading Environment in India

Trading in Indian equity markets is concentrated in two national exchanges: National Stock Exchange (NSE) and Bombay Stock Exchange (BSE). There are a number of regional exchanges, which witness much lower participation. Dark pools are legally not permitted. Hence, unlike their Western counterparts, Indian markets are not very fragmented. While BSE is the oldest stock exchange in Asia, NSE has in recent years emerged as the more dominant exchange in India. Total value of shares traded at NSE is roughly five times that of BSE.[3]Globally, NSE is the largest exchange when ranked by number of trades in equity shares

[3]Source: World Federation of Exchanges. 2012 report

NSE uses anonymous electronic limit order book (LOB) systems for trading in both spot and derivatives markets. There are no designated market makers at NSE. Brokers or clients of brokers can enter orders through their trading terminals. Orders are stored in the LOB based on a price-time priority rule. These orders are then continuously matched. The usage of identical platforms for trading of underlying shares, futures and call options make trading in these securities more integrated. This also greatly facilitates proprietary trading as identification of arbitrage and statistical arbitrage opportunities across markets is easier.

2.1 Identification of event periods

Direct Market Access (DMA) was introduced in India in April 2008. Institutional clients could now directly enter orders in the exchange's trading system; while this had to be done using a brokers' infrastructure, it didn't require their manual intervention. Removal of this broker interface led to the formal launch of Algorithmic Trading in India. Algorithms used by trading desks for market making and trading activities require fast response times as such opportunities might be extremely short-lived. As part of its efforts to reduce external latency, NSE introduced co-location services in January 2010. A broker could subsequently rent servers situated within NSE's premises. This led to a drastic reduction in latency; NSE reports latency levels of less than ten milliseconds[4]. To further facilitate HFT trading, NSE also provides tick-by-tick market data feed.

We employ an event study approach in our analysis. Since DMA represents a fundamental shift in the way markets function, it is possible that participants would have adapted to this innovation with a significant lag. Instead, we focus on the second milestone, namely

[4]Source: NSE website, http://www.nseindia.com/technology/content/tech_intro.htm.

introduction of co-location facilities. With systems already in place, participants would have required a lower acclimatization period. Hence, this event might be more appropriate for our research design.

Our event study approach requires that we obtain data from representative periods before and after the event. For our pre-event sample, we use two-weeks of data two months prior to introduction of colocation. For our post-event sample, we obtain data for two weeks from three distinct time periods: two months, four months and eight after the event. Hence, we use a total of eight weeks of data. While results from the last period might be subject to confounding factors, it provides a sufficiently long window for equilibrium results to emerge.

2.2 Data and sample construction

For our current analysis, we use a rich proprietary database obtained from NSE. This dataset contains complete order book and trade data for periods mentioned earlier. Our data are unique for several reasons. First, each message that arrives at the exchange - order entry, cancellation, or modification - identifies its originator as AT or non-AT. This provides the cleanest identification of algorithmic activity. For a period of three weeks in January 2008, Hendershott and Riordan (2012) use similar data from Deutsche Borse. Second, access to the complete database of orders and trades permits us to reconstruct the entire limit order book in event time. This combined with the unambiguous identification of AT lets us compute order and liquidity imbalance measures by trader types at desired frequencies. Third, each trade in our database is accompanied by the matched buy order number and sell order number. This lets us compute benchmark prices not just at the time of trade, but also at the time of order entry. Hence, robust measures of market liquidity such as effective spreads or implementation shortfall can be reliably constructed.

We work with a sample of 150 stocks that was selected from a universe of over 1400 stocks that were traded at the beginning of our pre-event period. We first select fifty stocks that are members of NSE's key benchmark index, S&P CNX Nifty. It is a market-capitalization weighted index that is adjusted for free-float. It contains 50 stocks which represent 24 sectors of the economy. These stocks form the first group in our analysis, namely Index stocks. We then select another 100 stocks from those that are traded in the derivatives segment[5]. Our decision to focus on this group is motivated by two factors. First, stocks in the derivatives segment are more likely to witness HFT/AT activity. Second, while there are no ``circuit-breakers" for these stocks, there are clearly specified circuit-breaker rules for stocks that are not traded in the derivatives segment. This differential treatment of stocks could potentially impact cross-sectional inferences. As of 31st October 2009, there were 128 non-index stocks that had derivatives traded on them; from these, we select the hundred largest by market capitalization. These are further sorted into two groups based on their market capitalization as on 31st Oct 2009.

[5]NSE has laid out clear guidelines for adding stocks to the derivatives segment; the dominant criterion is liquidity. To be specific, NSE computes “quarter sigma” for each stock; a stock’s quarter-sigma order size refers to the order size (in value terms) that is required to cause a change in the stock price equal to one-quarter of its standard deviation. A stock is eligible for the F&O segment only if this amount is above a certain minimum threshold (currently, set at INR 500, 000).

3. Commonality in order flow

If colocation facilitates more low-latency trading, and low-latency traders use the same or correlated signals in developing their strategies, then the resulting order flow could be correlated as well. In this section, we address this possibility by measuring commonality in order flow and examining how it changes as low latency trading becomes more intense.

We follow Lee et al. (2004) and formally define order flow or order imbalance as the difference between marketable buy and sell orders[6]. These are defined as respectively orders to buy at or above the best ask and orders to sell at or below the best bid. These orders demand immediacy and hence the measure reflects the prevailing directional imbalance in the market. Since we reconstruct the entire limit order book for every book event, we are able to accurately measure the best bid and ask; hence, there is no ambiguity in construction of these measures. We normalize order flow with total liquidity demanded (i.e., sum of marketable orders) for the stock during that interval. We also compute these measures separately for AT and non-AT.

Our research design builds on Chordia et al.'s (2000) market model for liquidity which was subsequently extended to order flows by Harford and Kaul (2005). Each security's order flow is modeled as a linear function of market-wide order flow and a set of control variables. To examine the impact of automation on order-flow commonality, we analyze how firms' order flow beta - defined as the sensitivity of a firm's order flow to market order flow - changes with introduction of colocation facilities. To be specific, we estimate the below market model for each firm for each period:

[6] A related measure of order imbalance that is widely used is signed order flow (see for instance, Hasbrouck and Seppi, 2001). This measures the difference between buyer-initiated and seller-initiated trades with signs themselves being inferred using an algorithm. In separate unreported robustness checks, we verify that our findings are robust to this alternate definition of order imbalance.

$$\Delta OF_{it} = \alpha_i + \beta_i \Delta OF_{\mathrm{m,it}} + \gamma_i X_{it} + \varepsilon_{it} \tag{1}$$

where ΔOF_{it} refers to change in order flow for firm i and $\Delta OF_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide order flow. The latter is computed as the average of $OF_{\mathrm{m,it}}$ for individual securities. While computing the market-wide order flow for firm i, we exclude firm i's contribution. Control variables include lagged values of both market and firm returns. Lagged market returns are included to control for index arbitrage strategies and feedback effects from returns to order flow (Harford and Kaul, 2005). To avoid issues related to bid-ask bounce, all returns are computed from the bid-ask midpoint prevalent at the end of the interval.

We follow Hasbrouck and Seppi (2001) and sample data at 15-minute intervals. To account for deterministic intra-day effects, we standardize all variables with mean and standard deviation corresponding to the firm-interval of day combination. To clarify, suppose Δy_{itk} refers to an observation for firm *i* on day *t* and time-interval *k*. This value is standardized by the mean and standard deviation of *y* estimated for firm *i* and interval *k* across all days. During the period considered in our analysis, NSE did not have any call auction at opening. This creates the well-documented spike in opening spreads. More critically for our study, this also creates an impression of high comovement as spreads for all stocks start narrowing during the subsequent period. Hence, we remove the first observation for each day.

Table II presents results of our analysis for order flows. The market model is estimated separately for each stock and for each period. For sake of brevity, we report only the cross-sectional median of β_i for the pre-event period and the median of change in β_i during each of our post-event periods. Our tests of significance are based on the non-parametric Wilcoxon signed-rank test. We defer discussion on specification issues to later sections. As in Kamara et al.

(2008), we use both liquidity beta and adjusted R^2 from market model as measures of comovement.

In Panel A of Table II, we document results for orders emanating from all market participants, i.e. AT and non-AT combined. The median order flow beta for the full sample (across stocks) is estimated to be 0.380 during the pre-event period. The null hypothesis that the median beta is zero is rejected at conventional levels of significance; further, 96.7% of estimated beta co-efficients are positive and significant at 95% confidence level (based on untabulated computations). The explanatory power of the model, as measured by the cross-sectional median of adjusted R^2, is 26.1%. These results suggest the presence of significant - both economical and statistical - commonality in order flow; buys (sells) in a stock tend to be highly correlated with buys (sells) in other stocks. Examining the various sub-samples, we find that that while commonality is lower for non-index stocks, the difference isn't substantial.

We next turn our attention to comovement in order flow of AT and non-AT; these results are presented in Panels B and C respectively. Order flow of non-AT exhibits higher commonality than that of AT. While median order flow beta of non-AT for the entire sample is about twice that of AT, median adjusted R^2 is about four-fold. Lower commonality of order flow for AT traders may be consistent with the behavioral characterization of commonality by Barberis et al. (2005). Whatever the reason for the commonality, our results suggest that competition among algorithmic traders, and their cross-sectional arbitrage activities, reduce commonality.

To examine the impact of latency on commonality, we estimate the market model of order flow separately for each of our three post-event periods. For each stock, we then compute the difference between estimates of order flow beta (and adjusted R^2) obtained from the post-event period and the pre-event period. These results are also reported in Table II. The change in beta is significant and negative during all periods. Examining changes in R^2 offers a similar interpretation: the explanatory power of the market model has decreased substantially post co-location. In Panels B and C, we find that this result holds for both AT and non-AT. These results strongly suggest that order flow comovement has decreased after introduction of colocation services.

A potential problem with statistical inference outlined above is that the signed ranked test assumes estimation errors to be independent across firm-specific equations. To address this issue, we estimate the market model using panel regression techniques and allow the error term to have a time-component that is common across all firms. As we are primarily interested in the average order flow beta for a group, we constrain all firms within a group to have the same beta. Additionally, we capture the impact of co-location facilities through an event-specific dummy variable. To be specific, we estimate the following model:

$$\Delta OF_{it} = \alpha_i + \beta \Delta OF_{m,it} + \lambda \delta_t \Delta OF_{m,it} + \gamma X_{it} + \varepsilon_{it} \tag{2}$$

$$\varepsilon_{it} = \vartheta_t + \zeta_{it}$$

where δ_t is a dummy variable that takes value of one for post-event periods and zero otherwise. We also let this dummy variable interact with our control variables. The error term in the above specification has a factor that is common to all stocks (ϑ_t). As this specification renders OLS estimation inappropriate, we estimate the model using panel data techniques. t-statistics based on Rogers standard errors that account for heteroskedasticity and autocorrelation are used for statistical inference (Petersen, 2009).

Table III reveals that results from panel regression are economically and statistically similar to those obtained from the cross-section of firm-level regressions. For instance, β and λ for the full sample are estimated to be respectively 0.386 and -0.186 for the full sample; both estimates are statistically significant. The corresponding estimates from individual regressions are 0.380 and -0.181 (cf. Table II). We conclude that our inferences are not unduly influenced by any cross correlation in errors. Chordia et al. (2000) arrive at a similar conclusion by examining the residuals from individual firm regressions. While we estimate equivalent panel regressions for other models, we do not report them for sake of brevity.

4. Commonality in returns and volatility

There is an extensive body of literature that studies comovement in asset returns; Barberis et al. (2005) provide an excellent overview. Hasbrouck and Seppi (2001) document that the common factor underlying order flows explains more than two-thirds of commonality in returns. Harford and Kaul (2005) also find correlated stock order flow to be an important driver of comovement in returns, and to a lesser extent, liquidity. If factors driving returns exhibit stochastic volatility, then volatility of returns would also exhibit some comovement. Kelly et al. (2012) provide evidence to support this hypothesis. Using a market model similar to the one employed in earlier sections, they find that market volatility explains about 38% of variation in firm-level volatility; the volatility beta is over 0.9. A natural extension of our work is then to examine (a) if low latency trading impacts commonality in returns, volatility and liquidity and (b) if any of these effects can be attributed to reduction in order flow comovement.

4.1 Tests of commonality

We estimate the following univariate market model for each firm for each period:

$$y_{it} = \alpha_i + \beta_i y_{m,it} + \varepsilon_{it} \qquad (3)$$

where y_{it} refers to either the return on stock i or a proxy for its volatility and $y_{m,it}$ refers to the contemporaneous market variable. As before, while estimating the market variable for firm i, we exclude its contribution. All prices used are bid-ask midpoints. Data is sampled at 15-minute intervals; to correct for intra-day variations, returns for each period are standardized with the mean and standard-deviation for the firm-interval of day combination.

Panel A of Table IV presents our results for returns. It is evident that returns display high comovement. Cross-sectional median of Adjusted R^2 is 34.3% for index stocks; median beta is 0.590 and is statistically significant. There isn't a discernible size effect in return betas. To examine the impact of latency on commonality, we compare the estimate of return beta (and adjusted R^2) obtained from the post-event period with that from the pre-event period for each stock in our sample. Medians of these changes, along with inferences from a signed ranked test, are also reported in Table IV. During the first post-event period, median return beta for index stocks declines by about 25% and adjusted R^2 by 40% (compared to the pre-event levels); these differences are significant at conventional levels of significance. We obtain similar results for most period- firm-group combinations.

Panels B of Table IV present results for absolute returns, our first proxy for volatility. We work with total returns and not with residuals from the market model or a factor model. This decision is motivated by Kelley et al.'s (2012) finding that idiosyncratic volatility accounts for a majority (96%) of variation in stock's volatility and that both total volatility and residual volatility (after accounting for factors) effectively possess the common factor structure. The median of volatility sensitivity is 0.349 for index stocks; the median explanatory power is 11.7%. We conclude that stocks in our sample display weak, albeit significant, commonality in volatility. We next investigate the impact of colocation. As with returns, we find a significant decline in volatility comovement. In Panel C, we examine if our results for volatility are robust to the proxy used for measuring volatility. Specifically, we use the range measure, defined as the difference between high and low prices scaled by the average price for the interval. While this measure yields higher estimates for comovement, results on the impact of colocation facilities are qualitatively similar.

4.2 Drivers of comovement

Our findings from previous sections suggest that both return and volatility comovement decrease with low latency. We next examine if this reduction is related to decline in order flow commonality. Kamara et al. (2008) undertake a similar exercise; they study if comovement in asset returns is related to commonality in liquidity measures by regressing return beta with liquidity beta. We adapt their framework and estimate the following cross-sectional regression between change in return/volatility beta ($\Delta\beta_i^{Mkt}$) and order flow beta ($\Delta\beta_i^{OF}$):

$$\Delta\beta_i^{Mkt} = \alpha + \lambda\Delta\beta_i^{OF} + \eta X_i + \varepsilon_i \tag{4}$$

where $\Delta\beta_i^{Mkt}$ refers to change in beta estimated from (3) around the co-location period, $\Delta\beta_i^{OF}$ refers to change in order flow beta estimate from (1) and X_i refers to control variables, namely change in average stock price and average dollar trading volume.

Panel A of Table V presents our analysis for returns. For sake of brevity, we report results only for the full sample. We reject the hypothesis that λ is zero at conventional levels of significance. For the first period after co-location, we estimate λ to be 0.419. The model has also reasonable explanatory power: adjusted R^2 is 23.3%. Hence, we conclude there is a positive association between changes in return beta and changes in order flow beta in the cross-section of firms. As noted by Kamara et al. (2008), this specification does suffer from an errors-in-variable problem which causes λ to be biased towards zero. In Panels B and C, we present results for our volatility proxies. While the association appears weaker (R^2s are lower), the co-efficient on order flow beta is statistically significant.

We had earlier documented that comovement in both AT and non-AT order flow decline post co-location. We next examine if either of them has a stronger association with reduction in comovement of returns or volatility. We estimate a variant of (4); instead of change in order flow beta, we use change in AT order flow beta or change in non-AT order flow beta as our explanatory variable. These results are also presented in Table V. It is evident that the strong association between order flow and return/volatility comovement is driven primarily by the changing dynamics of non-AT order flow.

5. Liquidity and liquidity imbalance

In this section, we define the various measures of liquidity used in this study. The relative quoted spread (RQS) measures the cost of executing a full round-trip trade executed at quoted prices: buying at the ask price and selling at the bid price. For a stock *i*, Relative Quoted Spread (RQS) over any time interval *t*, is computed as

$$RQS_{it} = \frac{\sum_{j=1}^{N} \tau_{it_j} \frac{Ask_{it_j} - Bid_{it_j}}{M_{it_j}}}{\sum_{j=1}^{N} \tau_{it_j}} \qquad (5)$$

where Ask_{it_j} and Bid_{it_j} are the best ask and bid prices at time t_j respectively, τ_{it_j} is the time-interval for which the quote is active (or the time between two consecutive order events), M_{it_j} is the mean of bid and ask prices. Specifically, RQS for any interval is constructed as the time-weighted average of relative bid-ask spreads, where the weight is the duration for which the quote is active.

As RQS considers only the best quotes on either side, it ignores that orders can execute at prices that differ from the quotes. The Relative Effective Spread (RES) corrects for this concern by measuring the difference between the actual average execution price and the prevailing mid-quote at the time of order entry. Hence, it is more appropriate for measuring the true cost of executing a market or a marketable limit order. It is computed as

$$RES_{it_j} = q_{it_j} \frac{P_{it_j} - M_{it_j}}{M_{it_j}} \tag{6}$$

where q_{it_j} is a signed indicator variable that takes a value $+1$ for buyer-initiated trade and -1 for seller initiated trade, P_{it_j} is the trade price and M_{it_j} is the quote midpoint at the time the initiating order enters the book. RES for an interval is then computed as the value-weighted average across all trades. We expect the difference between RQS and RES to be small, because there are no hidden orders inside the quote at the NSE and the market is automated and fast. The main difference between the two measures comes from the different weighting schemes. RQS is time weighted, reflecting the expected execution costs of a randomly arriving trader. RES represents the actual out-of-pocket ex-post execution costs of the typical trader.We present some basic descriptive statistics about these spread measures in Table I. While RQS measures the round-trip cost, RES as defined above measures the cost only for one leg of the trade.Quoted spreads in Indian markets are relatively tight; they are under 5 basis points for index stocks.

RQS and depth at five ticks - consolidated depth at the best quote and four ticks behind the best quote - are used to verify the robustness of our results. Depth for any interval is again computed as the time-weighted average of depth at different points in time.

5.1 Empirical Results

We estimate Chordia et al.'s (2000) market model of liquidity. Each security's liquidity is modeled as a linear function of market liquidity and a set of control variables. We estimate the following model for each firm for each event period:

$$\Delta \mathrm{y_{it}} = \alpha_\mathrm{i} + \beta_\mathrm{i}\Delta \mathrm{y_{m,it}} + \gamma_\mathrm{i}\mathrm{X}_{it} + \varepsilon_\mathrm{it} \tag{7}$$

where $\Delta \mathrm{y_{it}}$ refers to change in liquidity measure for stock *i* and $\Delta \mathrm{y_{m,it}}$ refers to the contemporaneous change in market-wide liquidity. The latter is computed as the average of y_{it} for individual securities. While estimating the model for firm *i*, we exclude its contribution while computing the market-wide liquidity. This definition of market liquidity is widely used (Acharya and Pedersen, 2005; Comerton-Forde et al., 2010) and is often employed in the context of liquidity commonality (Chordia et al., 2000; Coughenour and Saad, 2004; Kamara et al, 2008; Koch et al., 2012). Control variables include contemporaneous market return and change in stock *i*'s absolute returns. These variables are included to alleviate the possibility of obtaining spurious results that stem from any association between liquidity measures and returns and volatility (Chordia et al. 2000). To avoid issues related to bid-ask bounce, all returns are computed from the bid-ask midpoint prevalent at the end of the interval.

Table VI presents results of our analysis for Relative Effective Spreads (RES). The median liquidity beta for the full sample is estimated to be 0.065 during the pre-event period. The null hypothesis that the median beta is zero is rejected at conventional levels of significance. The explanatory power of the model, as measured by the cross-sectional median of adjusted R^2, is 3.8%. These results suggest the presence of weak, albeit statistically significant liquidity commonality in Indian markets. To better anchor these findings, we compare our results with those reported by Chordia et al. (2000). They use daily data on over thousand NYSE stocks and estimate average liquidity beta to be 0.28. The median adjusted R^2 of the market model is 0.3%. Hasbrouck and Seppi (2001) examine commonality in Dow 30 stocks by conducting principal component analysis on intra-day data. They document that the first component explains 4.7% of variation in levels of RES. We conclude that the level of commonality found in Indian markets during the pre-event period is comparable to other developed markets.

The well-documented size effect in comovement is also evident in our results. Index stocks have a median liquidity beta of 0.104; median adjusted R^2 of market models for these stocks is 4.9%. While these stocks have the lowest spreads, their spreads react most to contemporaneous change in market-wide spreads. Non-index stocks have lower betas, with beta being the lowest for the smallest firms in our sample.

We next turn our attention to the impact of colocation on liquidity commonality. Change in liquidity beta is significant only during the first period. However, the sign on the median of changes is negative. Examining changes in R^2 offers a similar interpretation: the median change in the explanatory power of the market model has decreased post co-location. While inferences based on liquidity beta and R^2 differ in their statistical significance, they broadly concur on the sign. These results strongly reject the hypothesis that commonality in liquidity has increased with introduction of colocation facilities. Examining results for sub-samples provides further insights. Most of reduction in commonality appears to be driven by index stocks. By either metric - R^2 or beta - we do not find any reduction in comovement for small stocks.

To verify if our central finding is sensitive to the measure of liquidity used in the analysis, we next estimate market models using RQS and Depth at five ticks as our proxies for liquidity. Table VII documents these results for various sub-samples. During the pre-event period, quoted spreads display higher commonality than RES. For index stocks, the median of cross-sectional beta is 0.149 and that of Adjusted R^2 is 5%. Depth exhibits much lesser commonality. These results are again in confirmation with those reported by Chordia et al. (2000) and Hasbrouck and Seppi (2001). A common interpretation of these findings is that market makers respond to systematic changes in liquidity by revising both quoted spreads and depth; however, the former is revised to a greater extent than the latter. This also explains the intermediate result that we obtain for RES. As a measure of implementation shortfall, it is driven by the underlying dynamics of both quoted spreads and depth.

Turning to the impact of colocation, we observe that commonality in RQS has decreased for all group-period combinations. Further, this decline is mostly significant. Decline in certain cases is as high as 80% relative to the pre-event levels. In sharp contrast, we find very little impact of latency on comovement of depths. Changes continue to be negative, but mostly insignificant. Considered in tandem, our results strongly suggest that colocation facilities do not increase commonality in liquidity.

5.2 Drivers of liquidity comovement

Literature on the drivers of liquidity commonality is scarce. Harford and Kaul (2005) document that commonality in order flow doesn't explain comovement in transaction costs. Domowitz et al. (2005) postulate that comovement in liquidity measures and returns are driven by economic forces that are fundamentally different. They classify orders into two types: liquidity demanding market orders and liquidity supplying limit orders. They argue that order type - and not its direction - solely determines the impact of incoming order on liquidity. Empirically, they establish that correlation in arrivals of order types is positively related to correlation in liquidity measures. We next study the impact of latency on comovement in liquidity imbalance and examine if it can explain changes in liquidity commonality.

Formally, we define liquidity imbalance over any time interval as the difference between liquidity demanded and supplied during that interval. Liquidity demanded in turn is defined as the total number of shares demanded via market and marketable orders; liquidity supplied is measured as the total number of shares supplied via non-marketable orders. To facilitate comparisons across stocks, we normalize imbalance measure with total liquidity supplied and demanded for the stock during that period. We also compute this imbalance measure separately for AT and non-AT.

We extend our earlier framework for order flow to liquidity imbalance. To be specific, we estimate the below model for each firm for each period:

$$\Delta LiqImb_{it} = \alpha_i + \beta_i \Delta LiqImb_{\mathrm{m,it}} + \gamma_i X + \varepsilon_{it} \quad (8)$$

where $\Delta LiqImb_{it}$ refers to the change in liquidity imbalance for firm i, $\Delta LiqImb_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide liquidity imbalance and X refers to the same set of control variables used earlier: contemporaneous market return and change in stock *i*'s absolute returns. As before, while estimating the model for firm *i*, we exclude its contribution to market-wide measures. We sample data at 15-minute intervals. All firm-specific and market-wide variables are standardized with mean and standard deviation corresponding to the firm-period of day combination.

Table VIII presents our results. In Panel A, we document results for orders emanating from all market participants, i.e. AT and non-AT combined. The cross-sectional median of imbalance beta is positive and statistically significant for all groups. Comovement in liquidity imbalance tends to be highest for index stocks. Medians of imbalance beta and Adjusted R^2 are 0.102 and 10.8% respectively. As with liquidity measures, we find a significant reduction in commonality in liquidity imbalance. Of course, this doesn't imply causation; we elaborate more on this later.

In Panels B and C of Table VIII, we present results separately for AT and non-AT. For index stocks, the cross-sectional medians of beta and R^2 are higher for AT; these differences, however, appear to be marginal. Statistical inference on the impact of latency depends on the metric used to measure commonality. While median change in R^2 is negative for both groups, it is (mostly) significant only for AT. Median change in imbalance beta is however mostly negative and insignificant for both AT and non-AT.

A potential concern with our earlier analysis is that the definition of liquidity supply is too conservative: we consider all non-marketable orders in computing liquidity supply. To alleviate this concern, we define a more aggressive measure of liquidity supply. To be specific, we consider total volume of liquidity supplied inside the prevailing best quotes. While the average commonality appears to be lower under this alternate definition of liquidity imbalance, results are qualitatively similar. Hence, these results are not tabulated here.

Our findings suggest that, at least under normal circumstances, access to higher speeds and lower latency doesn't necessarily exacerbate commonality in liquidity imbalances. This should alleviate concerns among regulators that algorithmic market-makers render trading systems more fragile by inflicting correlated shocks to liquidity imbalances across securities. In analyzing the impact of latency on market quality of single stocks, Hasbrouck and Saar (2013) note that episodes of extreme stress such as the flash crash on May 2010 might prompt key market makers to withdraw their liquidity supply. Such concerted actions could create sporadic bursts of heightened comovement in liquidity imbalance. However, as evidenced during the market crash on October 1987, widespread withdrawal of liquidity supply in face of extreme volatility isn't necessarily unique to electronic market makers.

We next investigate if a stock's liquidity beta is related to its liquidity imbalance beta. This helps us shed some light on linkages between comovement in liquidity and liquidity imbalance. We estimate the following cross-sectional regression between change in liquidity beta ($\Delta\beta_i^{Liq}$) and liquidity imbalance beta ($\Delta\beta_i^{LiqImb}$):

$$\Delta\beta_i^{Liq} = \alpha + \lambda\Delta\beta_i^{LiqImb} + \eta X_i + \varepsilon_i \tag{9}$$

where $\Delta\beta_i^{Liq}$ refers to change in liquidity beta estimated from (7) around the co-location period (for RES), $\Delta\beta_i^{LiqImb}$ refers to change in liquidity imbalance beta estimate from (8) and X_i refers to control variables, namely change in average stock price and average dollar trading volume. Table IX reports these results. We fail to reject the hypothesis that λ is zero at conventional levels of significance. Adjusted R^2 are very low too.

Finally, using the same model as in (9), we examine if liquidity comovement is associated with orderflow comovement. These results are presented in Table X. Again, we do not find any association between liquidity and order flow comovement. Our results suggest that sensitivity to changes in market liquidity is uncorrelated with sensitivity to changes in both market-wide liquidity imbalances and market-wide order flow.

6. Conclusions

While co-movement and the market impact of low latency / high frequency trading on market quality have independently attracted attention in recent years, there is little evidence on how the intensity of low latency trading is related to the commonality in order flow,liquidity, and volatility. In this paper, we provide the first direct evidence on this issue using data from a natural experiment at National Stock Exchange of India. Contrary to chabould et al. (2009) and to Huh (2011), we find that order flow from AT is less correlated than the order flow of traders not classified as AT. Reduction in latency, as represented by introduction of colocation facilities, leads to a significant reduction in order flow commonality for both trader categories. Comovement in other firm-specific attributes such as returns, volatility and liquidity also show a significant decline around this event. We show that the commonality in returns and volatility derives, at least in part, from commonality in order flow. Our findings are not consistent with the notion that more low latency trading increases systemic risk by accentuating comovement in order flows and liquidity. Instead, they are consistent with Chordia et al. (2013) who highlight that low latency trading is merely fast trading, without any fundamental changes in either the strategies employed by these traders or the underlying economics of financial markets.

References

Acharya, V.V., Pedersen, L. (2005), Asset pricing with liquidity risk, Journal of Financial Economics, 77 (2): 375-410.

Barberis, N., Shleifer, A., Wurgler, J. (2005), Comovement, Journal of Financial Economics 75, 283-317.

Biais, B., Foucault, T., Moinas, S. (2013), Equilibrium fast trading, Working paper, HEC Paris.

Boehmer, E., Fong, Y.L., Wu, J. (2012), International evidence on algorithmic trading, Working paper, EDHEC.

Brogaard, J. (2010), High frequency trading and its impact on market quality, Working paper, University iof Washington.

Brogaard, J., Hendershott, T., Riordan, R. (2013), High frequency trading and price discovery, Working paper, UC Berkeley.

Cespa, G., Foucault, T. (2013), Illiquidity contagion and liquidity crashes, Working paper, HEC Paris.

Chaboud, A., Chiquoine, B., Hjalmarsson, E., and Vega, C. (2009), Rise of the machines: Algorithmic trading in the foreign exchange market, Working paper, FRB.

Chordia, T., Goyal, A., Lehmann, B.N., Saar, G. (2013), High frequency trading, Journal of Financial Markets, 16(4): 637-45.

Chordia, T., Roll, R., Subrahmanyam, A. (2000), Commonality in liquidity, Journal of Financial Economics, 65 (1): 111-30.

Comerton-Forde, C., Hendershott, T., Jones, C.M., Moulton, P.C., Seasholes, M.S. (2010) Time variation in liquidity: The role of market maker inventories and revenues, Journal of Finance, 65: 295-331.

Corwin, S.A., Lipson, M.L. (2011), Order characteristics and the sources of commonality in prices and liquidity, Journal of Financial Markets, 14(1):47-81.

Coughenour, J.F., Saad, M.M. (2004), Common market makers and commonality in liquidity, Journal of Financial Economics, 73:37-69.

Harford, J., Kaul, A. (2005), Correlated order flow: Pervasiveness, sources, and pricing effects, Journal of Financial and Quantitative Analysis, 40, 29-55.

Hasbrouck, J., Saar, G. (2013), Low-latency trading, Journal of Financial Markets, 16(4): 646-79.

Hasbrouck, J., Seppi, D.J. (2001), Common factor in prices, order flows, and liquidity, Journal of Financial Economics, 59 (3): 383-411.

Hendershott, T., Jones, C.M., Menkveld, A. (2011), Does algorithmic trading improve liquidity? Journal of Finance, 66: 1-33.

Hendershott, T., Riordan, R. (2012), Algorithmic trading and the market for liquidity, Working paper, UC Berkeley.

Huberman, G., Halka, D. (2001), Systematic liquidity, Journal of Financial Economics, 24(2):161-78.

Huh, Yesol (2011), Algorithmic trading and liquidity commonality, Working paper, Stanford.

Jones, C.M. (2013), What do we know about high-frequency trading? Working paper, Columbia.

Kamara, A., Lou, X., Sadka, R.(2008), The divergence of liquidity commonality in the cross-section of stocks, Journal of Financial Economics 89, 444-466.

Kelly, B.T., Lustig, H.N., Nieuwerburgh, S.N. (2012), The volatility factor structure, . Working Paper, Chicago Booth.

Koch, A., Ruenzi, S., Starks, L. (2012), Commonality in liquidity: A demand-side explanation, Working paper, UT Austin.

Kondor, P. (2009), Risk in dynamic arbitrage: Price effects of convergence trading, Journal of Finance, 64(2):638-58.

Lee, Y., Liu, Y., Roll, R., and Subrahmanyam, A. (2004), Order imbalances and market efficiency: Evidence from the Taiwan Stock Exchange, Journal of Financial and Quantitative Analysis, 39: 327-41.

Menkveld, A. (2013), High frequency trading and the new market makers, Journal of Financial Markets, 16(4): 712-40.

Oehmke, M. (2009), Gradual arbitrage, Working paper, Columbia University.

Petersen, M.A. (2009), Estimating standard errors in finance panel data sets: Comparing approaches, Review of Financial Studies, 22(1):435-80.

Table I: Descriptive Statistics

This table presents some basic descriptive statistics for the sample that we consider in our study. Return is measured from quote midpoints sampled at 15-minute intervals. RQS and RES refer respectively to relative quoted (roundtrip) spreads and relative effective (half) spreads; these are defined in equations (5) and (6). %Liq Dem and %Liq Sup refer to percentage liquidity demanded and supplied by AT. The category Big refers to index stocks in our sample. Non-index stocks are classified into two categories - Med and Small- based on their market capitalization at the beginning of our first period.

	Period I: Pre co-location			Period II: Two months after co-location			Period III: Four months after co-location			Period IV: Eight months after co-location		
	Big	Med	Small	Big	Med	Small	Big	Med	Small	Big	Med	Small
Return (bp)	1.4	0.8	1.4	0.2	-0.1	0	-1.1	-0.8	-1.4	0.9	0.6	0.6
RQS (bp)	4.9	9.7	10.7	3.7	7.2	8.7	4.1	7.9	10.1	3.8	6.5	8.1
RES (bp)	3.4	6.1	6.8	2.5	4.4	5.2	2.9	4.8	6	2.7	3.9	4.8
% Liq Dem	12	9.2	9.5	14.6	9.9	8.8	13.8	12.6	9	14.6	12.4	10.5
%LiqSupp	64.1	56.5	57.3	64.7	60.1	63.3	72.1	66.3	64.3	73.3	65.1	61.4

Table II: Commonality in order flow

We test how colocation facilities impact commonality in order flow. Order flow is defined as the difference between marketable buy and sell orders, normalized by the total marketable orders. We estimate the following time-series model for each firm for each period: $\Delta OrdImb_{it} = \alpha_i + \beta_i \Delta OrdImb_{\mathrm{m,it}} + \gamma_i X_{it} + \varepsilon_{it}$, where $\Delta OrdImb_{it}$ refers to the change in order flow for firm i, $\Delta OrdImb_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide liquidity imbalance of orders, excluding firm i's imbalance. Control variables include lagged values of both market and firm return. We also estimate the model separately for AT and non-AT; these are presented in Panels B and C respectively. The group Big refers to index stocks in our sample. Non-index stocks are classified into two categories – Med and Small- based on their market capitalization at the beginning of first period. Inferences are based on non-parametric signed-rank test. AR2 refer to cross-sectional median of adjusted R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i/ Adj R^2 along with results of inference tests bases on a signed-rank test. $^{***}/^{**}$denote respectively significance at 99%, and 95% confidence levels.

	Period I: Pre co-location		Event I: Two months after co-location		Event II: Four months after co-location		Event III: Eight months after co-location	
	β_i	$AR2$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$
Panel A: All orders								
Full	0.380***	26.1	-0.181***	-9.1***	-0.110***	-7.2***	-0.201***	-10.2***
Big	0.444***	29.4	-0.214***	-12.8***	-0.111***	-8.0***	-0.235***	-13.0***
Med	0.340***	23.6	-0.158***	-8.6***	-0.101***	-6.1***	-0.162***	-9.5***
Small	0.378***	25.8	-0.182***	-8.5***	-0.116***	-7.6***	-0.205***	-10.8***
Panel B: AT orders								
Full	0.191***	5.7	-0.126***	-3.4***	-0.075***	-2.8***	-0.150***	-3.8***
Big	0.275***	9.9	-0.189***	-6.4***	-0.132***	-4.8***	-0.203***	-7.3***
Med	0.154***	4.0	-0.081***	-2.9	-0.013	-0.5	-0.083***	-2.5***
Small	0.15***	4.0	-0.089***	-2.4***	-0.075***	-2.6***	-0.123***	-2.2***
Panel C: Non-AT orders								
Full	0.351***	22.6	-0.142***	-6.9***	-0.087***	-6.1***	-0.186***	-7.5***
Big	0.392***	24.9	-0.163***	-8.3***	-0.075***	-6.3***	-0.206***	-8.1***
Med	0.315***	20.8	-0.117***	-5.2***	-0.072***	-5.4***	-0.149***	-6.6***
Small	0.356***	23.2	-0.149***	-7.1***	-0.112***	-7.2***	-0.217***	-11.4***

Table III: Commonality in order flow: Panel regression

This table presents results of our analysis on how colocation impacts commonality in order flow using panel techniques. Order flow is defined as the difference between marketable buy and sell orders, normalized by the total marketable orders. We estimate the following model with firm fixed effects for each group in our sample:

$$\Delta OF_{it} = \alpha_i + \beta \Delta OF_{m,it} + \lambda \delta_t \Delta OF_{m,it} + \gamma X_{it} + \varepsilon_{it}$$
$$\varepsilon_{it} = \vartheta_t + \zeta_{it}$$

ΔOF_{it} refers to the change in order flow for stock i, $\Delta OF_{m,it}$ refers to the contemporaneous change in market-wide order flow excluding firm i's order flow and δ_t is a dummy variable that takes value of one for post-co-location periods. Control variables include lagged values of both market and firm return. We interact these control variables also with event dummies. Data is sampled at 15-minute intervals. The category Big refers to index stocks in our sample. Non-index stocks are classified into two categories - Med and Small- based on their market capitalization at the beginning of our first period. t-statistics based on Rogers standard errors are used for statistical inference. $^{***}/^{**}$ denote respectively significance at 99% and 95% confidence levels.

	Event I: Two months after co-location			Event II: Four months after co-location			Event III: Eight months after co-location		
	β	λ	R^2	β	λ	R^2	β	λ	R^2
Full	0.386***	-0.186***	21.9	0.386***	-0.100***	22.6	0.386***	-0.205***	21.1
Big	0.444***	-0.220***	25.0	0.444***	-0.111***	26.4	0.444***	-0.244***	23.9
sMed	0.339***	-0.160***	19.8	0.339***	-0.077***	20.5	0.339***	-0.166***	18.9
Small	0.373***	-0.178***	21.3	0.373***	-0.113***	21.1	0.373***	-0.206***	20.7

Table IV: Commonality in returns and volatility

This table presents results of our analysis on how colocation facilities impact commonality in returns and volatility. Absolute returns and range measures (high – low scaled by average price for that interval) are used as proxies for volatility. Returns are computed from the bid-ask midpoints. We estimate the following model for each firm for each period: $y_{it} = \alpha_i + \beta_i y_{m,it} + \varepsilon_{it}$, where y_{it} refers to return or volatility for firm *i*, $y_{m,it}$ refers to the contemporaneous market-wide measure excluding firm i's value. The group Big refers to index stocks in our sample. *N*on-index stocks are classified into two categories – Med and Small- based on their market capitalization at the beginning of first period. Inferences are based on non-parametric signed-rank test. AR2 refer to cross-sectional median of adjusted R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i/ Adj R^2 along with results of inference tests bases on a signed-rank test. ***/**denote respectively significance at 99%, and 95% confidence levels.

	Period I: Pre co-location		Event I: Two months after co-location		Event II: Four months after co-location		Event III: Eight months after co-location	
	β_i	$AR2$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$
Panel A: Returns								
Big	0.590***	34.3	-0.151***	-13.7***	-0.078***	--8.0***	-0.187***	-17.9***
Med	0.485***	24.9	-0.142***	-10.7***	-0.021	-1.6	-0.152***	-11.4***
Small	0.563***	29.7	-0.154***	-13.9***	-0.063***	-5.7***	-0.203***	-19.1***
Panel B: Absolute Returns								
Big	0.349***	11.7	-0.164***	-7.4***	-0.051***	-3.9***	-0.183***	-8.1***
Med	0.236***	5.1	-0.111***	-3.7***	-0.010	-0.0	-0.119***	-4.0***
Small	0.279***	7.3	-0.153***	-6.1***	-0.038**	-1.1*	-0.143***	-6.1***
Panel C: Range [High-Low scaled by average price]								
Big	0.472***	21.9	-0.198***	-15.5***	-0.067**	-6.4***	-0.217***	-16.1***
Med	0.333***	10.6	-0.120***	-6.0***	-0.005	-0.3	-0.197***	-7.6***
Small	0.391***	14.9	-0.177***	-11.7***	-0.041*	-2.6*	-0.192***	-11.4***

Table V: Drivers of commonality in returns and volatility

This table presents cross-sectional regressions of changes in order flow comovement on changes in commonality in returns and volatility. Specifically, we estimate the following cross-sectional regression: $\Delta\beta_i^{Mkt} = \alpha + \lambda\Delta\beta_i^{OF} + \eta X_i + \varepsilon_i$ where $\Delta\beta_i^{Mkt}$ refers to change in return or volatility beta around the co-location period, $\Delta\beta_i^{OF}$ refers to change in order flow beta and X_i refers to control variables, namely change in average stock price and average dollar trading volume. The analysis is done for the full sample. We report results separately for models estimated using order flow imbalance for all orders (All), orders emanating from AT (AT) and non-AT. Inferences are based on non-parametric signed-rank test. Adj R^2 refers to adjusted R^2 expressed in %; t-statistics are presented in parenthesis. $^{***}/^{**}$denote respectively significance at 99%, and 95% confidence levels.

	Period I: Two months before co-location			Period I: Two months after co-location			Period III: Four months after co-location		
	Panel A: Returns								
	All	AT	Non-AT	All	AT	Non-AT	All	AT	Non-AT
λ	0.419	0.114	0.432	0.379	0.167	0.345	0.458	0.234	0.461
(t-stat)	(6.69)	(1.72)	6.60	(5.64)	(2.74)	(5.13)	(7.33)	(3.34)	(7.81)
[Adj R^2]	[23.3]	[1.7]	[22.8]	[17.2]	[4.0]	[14.5]	[28.1]	[8.6]	[30.6]
	Panel B: Volatility (Abs Returns)								
	All	AT	Non-AT	All	AT	Non-AT	All	AT	Non-AT
λ	0.171	0.051	0.144	0.351	0.052	0.336	0.314	0.194	0.333
(t-stat)	(2.29)	(0.72)	(1.84)	(4.65)	(0.77)	(4.50)	(3.95)	(2.38)	(4.4)
[Adj R^2]	[2.1]	[-1.1]	[0.9]	[11.3]	[-1.5]	[10.5]	[13.3]	[7.6]	[15.3]
	Panel C: Volatility (High-Low scaled by average price)								
	All	AT	Non-AT	All	AT	Non-AT	All	AT	Non-AT
λ	0.229	0.044	0.242	0.366	0.025	0.377	0.285	0.056	0.331
(t-stat)	(2.29)	(0.46)	(2.33)	(3.80)	(0.29)	(3.99)	(2.72)	(0.52)	(3.32)
[Adj R^2]	[3.8]	[0.5]	[3.9]	[7.7]	[-1.4]	[8.6]	[9.8]	[5.5]	[11.9]

Table VI: Commonality in relative effective spread

This table presents results of our analysis on how colocation facilities impact commonality in relative effective spreads (RES). We estimate the following time series model for each firm for each period: $\Delta RES_{it} = \alpha_i + \beta_i \Delta RES_{\mathrm{m,it}} + \gamma_i X_{it} + \varepsilon_{it}$, where $\Delta \mathrm{RES}_{it}$ refers to the change in RES for firm i, $\Delta RES_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide RES. Data is sampled at 15-minute intervals; RES for an interval is computed as the value-weighted average of RES during that interval. Control variables include contemporaneous market return and change in stock i's absolute returns. The group Big refers to index stocks in our sample. *N*on-index stocks are classified into two categories – Med and Small- based on their market capitalization at the beginning of first period. Inferences are based on non-parametric signed-rank test. AR2 refer to cross-sectional median of adjusted R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i/ Adj R^2 along with results of inference tests bases on a signed-rank test. ***/**denote respectively significance at 99%, and 95% confidence levels.

	Period I: Pre co-location		Event I: Two months after co-location		Event II: Four months after co-location		Event III: Eight months after co-location	
	β_i	$AR2$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$
Full	0.065***	3.8	-0.023**	-0.8	-0.012	-1.3***	0.007	-1.1***
Big	0.104***	4.9	-0.059***	-1.4***	-0.020	-1.3**	0.005	-1.6***
Med	0.036***	3.5	-0.014	0.0	-0.002	-2.2**	-0.009	-0.7
Small	0.029**	3.3	-0.005	0.2	-0.015	-0.7	0.016	-0.3

Table VII: Commonality in relative quoted spread and depth

This table presents results of our analysis on how colocation facilities impact commonality in relative quoted spreads (RQS) and depth. Depth refers to consolidated depth (bid and ask) at the best quote and four ticks behind the best quote. We estimate the following time series model for each firm for each period: $\Delta y_{it} = \alpha_i + \beta_i \Delta y_{\mathrm{m,it}} + \gamma_i X_{it} + \varepsilon_{it}$, where Δy_{it} refers to the change in RQS/Depth for firm i, $\Delta y_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide RQS/Depth. Data is sampled at 15-minute intervals; Depth and RQS for any interval are again computed as the time-weighted average of values at different event times. Control variables include contemporaneous market return and change in stock i's absolute returns. The group Big refers to index stocks in our sample. *N*on-index stocks are classified into two categories – Med and Small- based on their market capitalization at the beginning of first period. Inferences are based on non-parametric signed-rank test. AR2 refer to cross-sectional median of adjusted R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i/ Adj R^2 along with results of inference tests bases on a signed-rank test. ***/**denote respectively significance at 99%, and 95% confidence levels.

	Period I: Pre co-location		Event I: Two months after co-location		Event II: Four months after co-location		Event III: Eight months after co-location	
	β_i	$AR2$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$
Panel A: RQS								
Big	0.149***	5.0	-0.092***	-3.8***	-0.019	-3.5***	-0.079***	-3.6***
Med	0.112***	2.0	-0.058***	-1.2***	-0.041***	-0.9**	-0.014*	-1.4***
Small	0.079***	1.5	-0.062***	-0.6*	-0.066**	0.0	-0.039**	-1.0***
Panel B: Depth at five ticks								
Big	0.023*	1.3	-0.029	-0.8***	-0.006	-0.5	-0.003	-0.4*
Med	0.028**	1.2	-0.024	-0.1	-0.017*	-0.2	-0.034	-0.4
Small	0.042***	0.4	-0.056***	0.5	-0.020**	-0.0	-0.008	0.3

Table VIII: Commonality in liquidity imbalance

This table presents results of our analysis on how colocation facilities impact commonality in liquidity imbalance. Liquidity imbalance is defined as the difference between liquidity demanded and supplied. This measure is normalized with total liquidity supplied and demanded for the stock during the period. We estimate the following time series model for each firm for each period: $\Delta LiqImb_{it} = \alpha_i + \beta_i LiqImb_{\mathrm{m,it}} + \gamma_i X_{it} + \varepsilon_{it}$ where $\Delta LiqImb_{it}$refers to the change in liquidity imbalance in order for firm i, $\Delta LiqImb_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide liquidity imbalance of orders. Control variables include contemporaneous market return and change in stock i's absolute returns. Data is sampled at 15-minute intervals. The group Big refers to index stocks in our sample. *N*on-index stocks are classified into two categories – Med and Small- based on their market capitalization at the beginning of first period. Inferences are based on non-parametric signed-rank test. AR2 refer to cross-sectional median of adjusted R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i/ Adj R^2 along with results of inference tests bases on a signed-rank test. ***/**denote respectively significance at 99%, and 95% confidence levels.

	Period I: Pre co-location		Event I: Two months after co-location		Event II: Four months after co-location		Event III: Eight months after co-location	
	β_i	$AR2$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$	β_i-β_i^{pre}	$AR2$-$AR2_i^{pre}$
Panel A: All orders								
Big	0.102***	10.8	-0.047***	-1.6**	-0.056***	-2.1***	-0.099***	-3.5***
Med	0.061***	6.3	-0.049***	-1.4***	-0.020	-1.9***	-0.049***	-3.6**
Small	0.055***	8.4	0.005	-5.0***	-0.041**	-4.2***	-0.078***	-1.2**
Panel B: AT orders								
Big	0.114***	5.1	-0.071***	-3.2***	-0.050	-1.4**	-0.038***	-2.9***
Med	0.054**	3.2	-0.007	-2.5***	-0.014	-1.6***	-0.000	-2.9***
Small	0.042***	2.1	-0.054**	-2.1***	-0.032	-0.7***	-0.001	-1.4***
Panel C: Non-AT orders								
Big	0.094***	1.5	-0.039**	-0.5**	-0.004	-0.2	-0.021	-0.0
Med	0.022*	1.1	0.028	-0.6*	0.030	-0.2	0.063*	-0.2
Small	0.066***	0.8	-0.019*	-0.4	-0.031*	-0.5**	-0.024	-0.4*

Table IX: Liquidity beta and liquidity imbalance beta

This table presents cross-sectional regressions of changes in liquidity comovement (obtained for RES) on changes in commonality in liquidity imbalance. Specifically, we estimate the following cross-sectional regression: $\Delta\beta_i^{RES} = \alpha + \lambda\Delta\beta_i^{LiqImb} + \eta X_i + \varepsilon_i$ where $\Delta\beta_i^{RES}$ refers to change in liquidity beta around the co-location period, $\Delta\beta_i^{LiqImb}$ refers to change in liquidity imbalance beta and X_i refers to control variables, namely change in average stock price and average dollar trading volume. The analysis is done for full sample. We report results separately for models estimated using order flow imbalance for all orders (All), orders emanating from AT (AT) and non-AT. Inferences are based on non-parametric signed-rank test. Adj R^2 refers to adjusted R^2 expressed in %; t-statistics are presented in parenthesis. ***/**denote respectively significance at 99%, and 95% confidence levels.

	Period I: Two months before co-location			Period I: Two months after co-location			Period III: Four months after co-location		
	All	AT	Non-AT	All	AT	Non-AT	All	AT	Non-AT
λ	-0.114	0.047	0.098	-0.029	0.012	0.001	0.099	-0.033	-0.002
(t-stat)	(-1.08)	(0.50)	(1.07)	(-0.34)	(0.15)	(0.01)	(1.08)	(-0.42)	(-0.02)
[Adj R^2]	[0.5]	[-0.1]	[0.5]	[-1.9]	[-2.0]	[-2.0]	[-0.9]	[-1.6]	[-1.7]

Table X: Liquidity beta and order flow beta

This table presents results of cross-sectional regressions of liquidity comovement (obtained for RES) on changes in commonality in order flow. Specifically, we estimate the following cross-sectional regression: $\Delta\beta_i^{RES} = \alpha + \lambda\Delta\beta_i^{OF} + \eta X_i + \varepsilon_i$ where $\Delta\beta_i^{RES}$ refers to change in liquidity beta around the co-location period, $\Delta\beta_i^{OF}$ refers to change in order flow beta and X_i refers to control variables, namely change in average stock price and average dollar trading volume. The analysis is done for full sample. We report results separately for models estimated using order flow imbalance for all orders (All), orders emanating from AT (AT) and non-AT. Inferences are based on non-parametric signed-rank test. Adj R^2 refers to adjusted R^2 expressed in %; t-statistics are presented in parenthesis. $^{***}/^{**}$denote respectively significance at 99%, and 95% confidence levels.

	Period I: Two months before co-location			Period I: Two months after co-location			Period III: Four months after co-location		
	All	AT	Non-AT	All	AT	Non-AT	All	AT	Non-AT
λ	0.093	0.127	0.030	-0.018	-0.094	-0.049	0.080	0.073	0.048
(t-stat)	(1.01)	(1.50)	(0.31)	(-0.23)	(1.43)	(-0.64)	(0.94)	(0.86)	(0.59)
[Adj R^2]	[0.42]	[1.2]	[-0.2]	[-1.9]	[-0.6]	[-1.7]	[-1.1]	[-1.2]	[-1.5]