Low latency trading and the comovement of order flow, prices, and market conditions*

Ekkehart Boehmer
Ekkehart.boehmer@edhec.edu

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

R.L.Shankar
rlshankar@gmail.com

Institute for Financial Management and Research, India

April 4, 2014

Preliminary

Abstract

We examine the impact of algorithmic trading (AT) in equities on the comovement of order flow, returns, liquidity, and volatility to assess how AT affects the market's susceptibility to systemic shocks. Using order-level data around a natural experiment at the National Stock Exchange of India, which in 2010 has introduced features that promote HFT, we find that more intense AT reduces commonality in order flow, returns, liquidity, and volatility, and therefore reduce the market's susceptibility to systemic shocks. These declines are more pronounced for algorithmic order flow and for large-cap firms. We attribute our findings to more intense competition among algorithmic than non-algorithmic traders.

* This paper is part of the NSE-NYU Stern School of Business Initiative for the Study of Indian Capital Markets. The authors acknowledge the support of the initiative. The views expressed in this Working Paper are those of the authors and do not necessarily represent those of NSE or NYU. The authors also thank seminar and conference participants at the National University of Singapore and the 2013 Australasian Finance and Banking Conference.

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. The explosive growth in HFT activities has raised 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 flow, returns, and a range of liquidity measures. Examining this question is important as it offers insights on whether faster trading makes markets more susceptible to systemic risk, potentially magnified through the correlation among AT strategies or among its effects on markets.

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. At best, greater commonality in liquidity, returns, or order flow will create externalities that can conceivably magnify the costs of trading. For example, if algorithmic traders often trade contemporaneously into the same direction for multiple securities (i.e., commonality in order flow is high) or trading costs tend to increase at the same time for multiple securities (commonality in liquidity is high), then trading costs are greater compared to markets with less correlated trading patterns.

Conversely, as Chordia et al. (2013) note in their review article, lower latency may encourage faster trading without fundamentally changing either the strategies employed by traders or the underlying economics of financial markets. Under this view, lower latency might not have any impact on commonality in liquidity or order flows. Similarly, the frictionbased theory of comovement elaborated in Barberis et al. (2005) suggests that comovement might actually decrease with more HFT. For example, one of the authors' arguments is based on differential rates of information diffusion. Information about fundamentals might be incorporated faster in some stocks; for instance, in stocks with lower transaction costs or those held by investors with faster access to information. Returns of stocks with similar rates of diffusion would display commonality. Corwin and Lipson (2010) provide evidence on the relation between comovement in order flow and correlated trading decisions within various trader categories. Conceivably, algorithmic traders are efficient monitors of new information so that information diffusion becomes less of an issue when HFT increases.

We provide the first direct evidence on the impact of low-latency trading on comovement in order flows and market conditions. 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' infrastructure, but without their manual intervention. This step is important, because it allows algorithmic traders to access the market without the delay caused by routing through a brokerage. To further reduce external latency, NSE introduced co-location services in January 2010. With colocation, market participants can 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

[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.

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

identify the originator for each incoming message as either AT (algorithmic trader) or non-AT.

We 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. Beyond this difference, the introduction of colocation facilities leads to a significant reduction in order flow comovement for both AT and non-AT that is, however, more pronounced for AT flow.

Order flow is naturally related to prices and liquidity and 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. Taken together, our findings are not consistent with the notion that more AT strategies are correlated in a way that increases systemic risk by accentuating comovement in order flows or market conditions.

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 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 the 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 leave the market. Trading protocols that support low-latency trading could thus lead to crowding out of slow traders. The resulting increase in the share of fast traders, 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 of specialists 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) 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 42 exchanges. They find that, on average, more intense AT activity increases market liquidity for the two-thirds largest firms on each market, but reduces liquidity for the smallest one-third of firms. AT also increases the informational efficiency of prices and volatility. 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 related to 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 the 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.

Chaboud et al. (2009) and Huh (2011) look at commonality in the context of AT. 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 from Huh's in that we use order-level data that more clearly identify trading intention than the NYSE trade-level data. Our data also 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. 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 even before the hybrid market event. As such, the NYSE's 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 research design. In Section 3, we present results for commonality in order flow. We discuss comovement in returns and volatility in section 4 and comovement in liquidity and liquidity imbalance in Section 5. The last section concludes.

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 with much lower market shares. 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.

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. With DMA, institutional clients can directly enter orders in the exchange's trading system. This event supports 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

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

levels of less than ten milliseconds.[4] To further facilitate HFT trading, NSE also provides real-time tick-by-tick market data feeds.

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 introduction of co-location facilities. With systems to handle algo trading already in place, participants likely do not require long acclimatization periods after the adoption of colocation. Hence, this event is 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 before the introduction of colocation. For our post-event sample, we obtain data for two weeks from two distinct time periods: two months and four months after the event.

2.2 Data and sample construction

Our proprietary NSE data contain complete order book and trade data for the periods mentioned earlier. The 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 a clean 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

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

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 at the order level.

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, 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 “Index stocks” form the first group in our analysis. 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.

2.3 Methodology

If colocation facilitates more AT, and algo 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.

[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).

We follow Lee et al. (2004) and formally define order flow or order imbalance as the difference between marketable buy and sell orders.[6] We reconstruct the entire limit order book for every book event and record the best bid and ask at each time stamp, and then use these inside quotes to classify order flow. We normalize net 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 flow 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 the introduction of colocation facilities. Specifically, we estimate the following market model for each firm and each period:

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

where ΔOF_{it} refers to change in net order flow for firm i and $\Delta OF_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide net order flow. The latter is computed as the average $OF_{\mathrm{m,it}}$ of all individual securities except firm i's order flow. 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). We estimate all regressions with these controls and also without controls. Because results are qualitatively identical, we report only the models without controls unless

[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.

indicated otherwise. 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.

We use standard measures of liquidity. 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

$$RQS_{it} = \frac{\sum_{j=1}^{N} \tau_{it_j} \frac{Ask_{it_j} - Bid_{it_j}}{M_{it_j}}}{\sum_{j=1}^{N} \tau_{it_j}} \qquad (2)$$

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

$$RES_{it_j} = q_{it_j} \frac{P_{it_j} - M_{it_j}}{M_{it_j}} \qquad (3)$$

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.

3. Commonality in order flow

Table II presents results of our analysis for order flows. The market model is estimated separately for each stock and for each period without controls. For sake of brevity, we report only the cross-sectional median of β_i for the pre-event period and the median change in $\beta_i$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 the market model R^2 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 net order flow beta for the full sample is 0.419 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 R^2, is 17.6%. These results suggest the presence of significant - both economically and statistically - commonality in order flow. Aggressiveness in Buys (sells) in a stock tends to be correlated with the aggressiveness of buys (sells) in other stocks. Examining the size sub-samples, we find that that commonality is similarly prevalent in large, mid cap, and small firms.

To examine the impact of colocation on the commonality of order flow, we estimate the market model separately for each of the two post-event periods. For each stock, we then compute the post minus pre difference around colocation. These results are reported in the two right-most columns 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 decreases substantially after introduction of colocation services. These declines are more pronounced for AT in that the proportional decline in commonality is much larger in the AT flow than in the non-AT flow. This makes intuitive sense, because as colocation allows AT strategies to be more effective, and AT reduces commonality, we expect colocation to reduce commonality more for AT flow. Results are very similar and qualitatively identical when we control for lagged returns in each of the cross-sectional firm-level regressions (see Table III).

To more directly assess the difference between AT and non-AT flow regarding their differential effects on comovement in net order flow, we estimate firm-level regressions separately for AT and NAT flow. The results in Table IV show that NAT flow is associated with significantly greater commonality in each of the three periods. This suggests that it is NAT flow, rather than AT flow, that drives commonality in order flow.

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 use a panel regression that allows 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. We estimate the following model:

$$\Delta OF_{it} = \alpha_i + \beta \Delta OF_{m,it} + \lambda \delta_t \Delta OF_{m,it} + \gamma X_{it} + \varepsilon_{it} \quad (4)$$

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

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.

We use t-statistics based on Rogers standard errors that account for heteroskedasticity and autocorrelation for statistical inference (Petersen, 2009).

The results in Table V are economically and statistically significant and similar to those obtained from the firm-level regressions. For instance, β and λ are 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 III). Thus, our inferences are not unduly influenced by any cross correlation in errors.

4. Commonality in returns and volatility

An extensive body of literature 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.

4.1 Return commonality

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

$$y_{it} = \alpha_i + \beta_i y_{\mathrm{m,it}} + \varepsilon_{it} \qquad (5)$$

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

It is evident from Table VI that returns display high comovement. Cross-sectional median R^2 is 30.6% for index stocks; median beta is 0.553 and is statistically significant. There is no discernible size effect in return betas. To examine the impact of latency on commonality, we compare the estimate of return beta (and 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 VI. Commonality declines significantly during the both post-event periods. Again, these results are similar across firm size categories.

4.2 Volatility commonality

We estimate Model (5) with our proxies for volatility as dependent variables. Panel A of Table VII presents 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 Kelly 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 0.287 for the entire sample; the median explanatory power is 8.2%. The stocks in our sample display significant commonality in volatility. Most importantly, as with returns, we find a significant decline in volatility comovement after colocation. We obtain similar results based on the intra-interval price range, reported in Panel B.

5. Commonality in liquidity and liquidity imbalance

5.1. Liquidity commonality

We estimate Chordia et al.'s (2000) market model of liquidity that models each security's liquidity 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:

$$\Delta y_{it} = \alpha_i + \beta_i \Delta y_{m,it} + \gamma_i X_{it} + \varepsilon_{it} \quad (6)$$

where Δy_{it} refers to change in liquidity measure for stock *i* and $\Delta y_{m,it}$ 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). As before, we obtain qualitatively identical results whether we include these controls or not, and report only estimates without controls. 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 VIII presents commonality in Relative Effective Spreads (RES). The median liquidity beta for the full sample is estimated to be 0.066 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 R^2, is 0.6%. We find similar results for RQS and quoted depth up to five ticks – each of these measures exhibits a significant commonality.

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 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 those earlier results.

The well-documented size effect in comovement is also evident in our results. Large cap stocks have a median liquidity beta of 0.112, while the smallest stocks have a beta of only 0.054. This suggests larger commonality among larger stocks. While these stocks tend to have the lowest spreads, their spreads react most to contemporaneous changes in market-wide spreads.

After colocation, liquidity commonality declines across the board, in particular for RQS and, for the full sample, for Depth. The declines in RES commonality are only significant during the first post-event period. Examining changes in R^2 offers a similar interpretation. Most of reduction in commonality appears to be driven by the large cap stocks.

5.2. Commonality in liquidity imbalance

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 colocation on comovement in liquidity imbalance and examine if it can explain changes in liquidity commonality.

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 in each period:

$$\Delta LiqImb_{it} = \alpha_i + \beta_i \Delta LiqImb_{\mathrm{m,it}} + \gamma_i X + \varepsilon_{it} \qquad (7)$$

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. We report only result without controls as they are qualitatively identical to the ones with these two control variables. 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 IX 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 betas is positive and statistically significant for all groups. Comovement in liquidity imbalance tends to increase with market cap. As with liquidity measures, we find a significant decline in commonality of liquidity imbalance after colocation.

In Panels B and C, we present results separately for AT and non-AT flow. Commonality is somewhat higher in the AT flow except in the small cap group. But the AT flow also shows the largest decline in commonality after colocation. In fact, comparing the estimates in Panels B and C suggests that the post-event commonality is lower for AT than NAT flow. Table X provides a more direct look at the difference between AT and NAT flow. Before colocation, the commonality associated with NAT flow is smaller than that associated with AT flow in all but the smallest firms. In both post-colocation periods, the difference is no longer distinguishable from zero. This suggests that overall, the effect of AT flow on commonality of liquidity imbalance is not different from the effect that NAT flow has on this commonality.

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, returns, volatility, and liquidity. We provide the first direct evidence on this issue using anew data set around a natural experiment at the National Stock Exchange of India. We can cleanly identify algorithmic order flow at the order level and exploit a colocation event that makes algorithmic trading strategies more effective. Contrary to Chabould et al. (2009) and Huh (2011), we find that order flow from algo traders is less correlated than the order flow of other traders. Correspondingly, a reduction in latency around the introduction of colocation facilities leads to a significant reduction in order flow commonality for both trader categories. The comovement in other relevant firm-specific attributes such as returns, volatility and liquidity also show a significant decline around this event.

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, we show that more intense algo trading is associated with lower commonality. These results 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.

The important open question is why and through which channels algorithmic trading strategies can reduce commonalities. Clearly, if algo traders followed strategies that merely mimic each other or are derived from a common model, we should observe more co-movement in order flow, not less. We suggest that, instead, that the results are driven by fiercely competitive trading. For example, if alpha is limited in scale, then correlated order flow signals a reduced scale limit. Traders observing that their strategy is correlated with that of other traders should modify their strategy to increase the scale of potential alpha, and thus maximize the potential size of assets in this strategy. If traders are competitive in this way and modify strategies in response to comovement, we should expect declines in commonalities as algorithmic trading becomes more intense.

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.

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 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 are percentage liquidity demanded and supplied by algorithmic traders (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 period I. For each post-event period, we report the cross-sectional mean of changes relative to the corresponding variables in period I. We report t-tests where *** and ** denote significance at 99% and 95% confidence levels.

	Period I: Pre co-location			Event I: Change two months after co-location			Event II: Change four months after co-location		
	Big	Med	Small	Big	Med	Small	Big	Med	Small
Return (bp)	1.4***	0.8	1.4***	-1.1***	-0.1	-1.5***	-2.6***	-1.6***	-2.8***
RQS (bp)	4.9***	9.7***	10.7***	-1.3***	-2.5***	-2.0***	-0.8***	-1.7***	-0.6
RES (bp)	3.4***	6.1***	6.8***	-0.9***	-1.7***	-1.5***	-0.6***	-1.3***	-0.8***
% Liq Dem	12***	9.2***	9.5***	2.6***	0.7	-0.7	1.8***	2.6***	-0.5
%LiqSupp	64.1***	56.5***	57.3***	0.6	3.6	6.0	8.0***	9.8***	7.1**

Table II: Commonality in net 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: $\Delta OrdImb_{it} = \alpha_i + \beta_i \Delta OrdImb_{\mathrm{m,it}} + \varepsilon_{it}$, 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 order flow, excluding firm i's imbalance. 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. R2 refers to cross-sectional median of R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i or R^2 along with a signed-rank test where *** and ** denote 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	
	β_i	$R2$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$
Panel A: All orders						
Full	0.419***	17.6	-0.167***	-10.2***	-0.105***	-7.2***
Big	0.457***	20.9	-0.206***	-13.0***	-0.113***	-9.8***
Med	0.357***	12.7	-0.153***	-9.6***	-0.095***	-5.9***
Small	0.413***	17.0	-0.160***	-8.8***	-0.098***	-7.2***
Panel B: AT orders						
Full	0.199***	4.0	-0.141***	-2.5***	-0.071***	-1.6***
Big	0.300***	9.0	-0.185***	-7.1***	-0.136***	-5.5***
Med	0.151***	2.3	-0.090***	-1.5***	-0.040***	-0.6
Small	0.150***	2.3	-0.120***	-1.9***	-0.059***	-1.1***
Panel C: Non-AT orders						
Full	0.383***	14.8	-0.141***	-7.9***	-0.099***	-5.6***
Big	0.414***	17.3	-0.154***	-8.6***	-0.100***	-7.4***
Med	0.337***	11.4	-0.117***	-6.0***	-0.091***	-4.0***
Small	0.391***	15.3	-0.143***	-8.8***	-0.101***	-7.2***

Table III: Commonality in order flow with controls for lagged returns

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 order flow, excluding firm i's imbalance. We control for lagged market return and lagged firm return. 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. R2 refers to cross-sectional median of R^2 (in %). For each post-event period, we report the median of difference between pre- and post- β_i or R^2 along with a signed-rank test where *** and ** denote 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	
	β_i	$R2$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$
Panel A: All orders						
Full	0.380***	27.3	-0.181***	-9.2***	-0.110***	-7.4***
Big	0.444***	30.5	-0.214***	-12.8***	-0.111***	-8.1***
Med	0.340***	24.8	-0.158***	-8.6***	-0.101***	-6.3***
Small	0.378***	27.0	-0.182***	-8.5	-0.116***	-7.8
Panel B: AT orders						
Full	0.191***	7.2	-0.126***	-3.6***	-0.075***	-3.1
Big	0.275***	11.3	-0.189***	-6.5***	-0.132***	-5.0
Med	0.154***	5.6	-0.081***	-3.1***	-0.013	-0.8**
Small	0.15***	5.6	-0.089***	-2.6***	-0.075***	-2.9***
Panel C: Non-AT orders						
Full	0.351***	23.8	-0.142***	-6.9***	-0.087***	-6.3***
Big	0.392***	26.1	-0.163***	-8.3***	-0.075***	-6.5***
Med	0.315***	22.0	-0.117***	-5.2***	-0.072***	-5.6***
Small	0.356***	24.5	-0.149***	-7.1***	-0.112***	-7.3***

Table IV: Commonality in order flow: AT vs Non-AT

We compare commonality in order flow of AT and non-AT traders. 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 order flow, excluding firm i's imbalance. We control for lagged market return and lagged firm return. We estimate the model separately for AT and non-AT order flow. In Panel A, we compute the difference between β_i^{NAT} and β_i^{AT} for each firm and report the cross-sectional median of this difference. We also report the median of difference between R2 of non-AT and AT order flow. In Panel B, for each post-event period, we report the median of difference between pre- and post- (β_i^{NAT} - β_i^{AT}) and ($R2^{\mathrm{NAT}}$ - $R2^{\mathrm{AT}}$) along with results of inference tests based on a signed-rank test. R2 refers to R^2 (in %). Inferences are based on non-parametric signed-rank test, where *** and ** denote significance at 99% and 95% confidence levels.

	Period I: Pre co-location		Period II: Two months after co-location		Period III: Four months after co-location	
	β_i^{NAT}-β_i^{AT}	$R2^{\mathrm{NAT}} - R2^{\mathrm{AT}}$	β_i^{NAT}-β_i^{AT}	$R2^{\mathrm{NAT}} - R2^{\mathrm{AT}}$	β_i^{NAT}-β_i^{AT}	$R2^{\mathrm{NAT}} - R2^{\mathrm{AT}}$
Full	0.186***	10.2***	0.161***	4.4***	0.157***	4.9***
Big	0.148***	10.4***	0.184***	6.6***	0.135***	6.0***
Med	0.193***	7.9***	0.108***	2.1***	0.165***	4.3***
Small	0.217***	11.3***	0.186***	5.3***	0.167***	4.3***

Table V: Commonality in order flow: Panel regression

We test how colocation impacts commonality in net order flow, defined as the difference between marketable buy and sell orders normalized by all marketable orders. We estimate the panel regression model in (4) with firm fixed effects for each group in our sample: $\Delta OF_{it} = \alpha_i + \phi_i\delta_t + \beta\Delta OF_{m,it} + \lambda\delta_t\Delta OF_{m,it} + \gamma X_{it} + \varsigma\delta_t X_{it} + \varepsilon_{it}$, $\varepsilon_{it} = \vartheta_t + \zeta_{it}$, where 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 are based on Rogers standard errors and *** and ** denote significance at 99% and 95% confidence levels.

	Event I: Two months after co-location			Event II: Four months after co-location		
	β	λ	R^2	β	λ	R^2
Panel A: All orders						
Full	0.386***	-0.186***	21.9	0.386***	-0.100***	22.6
Big	0.444***	-0.220***	25.0	0.444***	-0.111***	26.4
Med	0.339***	-0.160***	19.8	0.339***	-0.077***	20.5
Small	0.373***	-0.178***	21.3	0.373***	-0.113***	21.1
Panel B: AT orders						
Full	0.188***	-0.121***	4.0	0.188***	-0.069***	4.2
Big	0.271***	-0.187***	6.3	0.271***	-0.107***	6.9
Med	0.142***	-0.076***	3.0	0.142***	-0.035***	3.5
Small	0.150***	-0.102***	3.3	0.150***	-0.065***	3.1
Panel C: Non-AT orders						
Full	0.350***	-0.141***	19.5	0.350***	-0.094***	19.6
Big	0.396***	-0.160***	21.6	0.396***	-0.092***	22.4
Med	0.310***	-0.123***	18.0	0.310***	-0.143***	18.0
Small	0.345***	-0.140***	19.3	0.345***	-0.203***	18.7

Table VI: Commonality in returns

We examine how colocation facilities impact commonality in returns. We estimate the following model for each firm for each period: $y_{it} = \alpha_i + \beta_i y_{\mathrm{m,it}} + \varepsilon_{it}$, where y_{it} refers to the return for firm i, $y_{\mathrm{m,it}}$ refers to the contemporaneous market-wide measure excluding firm i's value. Returns are computed from bid-ask midpoints. 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. R2 refers to the cross-sectional median of R^2 (in %). For each post-event period, we report the median difference between pre- and post- β_i or R^2 along with a signed-rank test, where *** and ** denote 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	
	β_i	$R2$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$
Full	0.553***	30.6	-0.147***	-14.7***	-0.049***	-4.9***
Big	0.590***	34.8	-0.151***	-13.7***	-0.078***	-8.0***
Med	0.485***	23.5	-0.142***	-10.7***	-0.021	-1.6
Small	0.563***	31.7	-0.154***	-13.9***	-0.063***	-5.7***

Table VII: Commonality in volatility

We examine how colocation facilities impact commonality in volatility. 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 the volatility for firm i, $y_{m,it}$ refers to the contemporaneous market-wide measure excluding firm i's value. We use two different proxies for volatility: absolute returns and intra-day range (high – low scaled by average price for that interval). Returns are computed from bid-ask midpoints. 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. R2 refers to the cross-sectional median of R^2 (in %). For each post-event period, we report the median difference between pre- and post- β_i or R^2 along with a signed-rank test, where *** and ** denote 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	
	β_i	$R2$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$
Panel A: Absolute returns						
Full	0.287***	8.2	-0.139***	-6.0***	-0.039***	-1.4***
Big	0.349***	12.2	-0.164***	-7.4***	-0.051***	-3.9***
Med	0.236***	5.6	-0.111***	-3.7***	-0.010	-0.2
Small	0.279***	7.8	-0.153***	-6.1***	-0.038**	-1.2**
Panel B: High-low scaled by average price						
Full	0.397***	15.8	-0.175***	-11.1***	-0.034***	-2.4***
Big	0.472***	22.3	-0.198***	-15.5***	-0.067**	-6.5***
Med	0.333***	11.1	-0.120***	-6.1***	-0.005	-0.4
Small	0.391***	15.3	-0.177***	-11.7***	-0.041	-2.7

Table VIII: Commonality in liquidity

We examine how colocation facilities impact commonality in relative effective spreads (RES), relative quoted spreads (RQS), and depth. 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 one of the three liquidity measures for firm i, $y_{m,it}$ refers to the corresponding, contemporaneous market-wide measure excluding firm i's value. Data is sampled at 15-minute intervals. RES for an interval is computed as the value-weighted average of RES during that interval. Depth and RQS for any interval are again computed as the time-weighted average of intraday values. 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. R2 refers to the cross-sectional median of R^2 (in %). For each post-event period, we report the median difference between pre- and post- β_i or R^2 along with a signed-rank test, where *** and ** denote 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	
	β_i	$R2$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$
Panel A: Relative effective spreads (in bp)						
Full	0.066***	0.6	-0.032***	-0.1***	-0.006	-0.1
Big	0.112***	1.3	-0.064***	-0.8***	-0.011	-0.2
Med	0.048***	0.4	-0.016	-0.0	0.009	0.0
Small	0.054***	0.5	-0.004	-0.0	-0.004	-0.2
Panel B: RQS						
Full	0.117***	1.4	-0.079***	-0.9***	-0.045***	-0.5***
Big	0.177***	3.1	-0.123***	-2.6***	-0.046	-0.6**
Med	0.116***	1.3	-0.079***	-0.9***	-0.047***	-0.8***
Small	0.099***	1.0	-0.053***	-0.4***	-0.040***	-0.1
Panel C: Depth at five ticks						
Full	0.036***	0.4	-0.030***	-0.1**	-0.035***	-0.1**
Big	0.032***	0.4	-0.037	0.0	-0.039	0.1
Med	0.037***	0.5	-0.009	-0.1	-0.039	-0.1**
Small	0.042***	0.4	-0.056***	-0.1	-0.021**	-0.1

Table IX: Commonality in liquidity imbalance

We examine how colocation facilities impact commonality in liquidity imbalance, 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: $\Delta LiqImb_{it} = \alpha_i + \beta_i \Delta LiqImb_{\mathrm{m,it}} + \varepsilon_{it}$, where $\Delta LiqImb_{it}$ refers to the change in liquidity imbalance flow for firm i, $\Delta LiqImb_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide liquidity imbalance of orders, excluding firm i's imbalance. 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. R2 refers to the cross-sectional median of R^2 (in %). For each post-event period, we report the median difference between pre- and post- β_i or R^2 along with a signed-rank test, where *** and ** denote 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	
	β_i	$R2$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$	β_i-β_i^{pre}	$R2$-$R2_i^{pre}$
Panel A: All orders						
Full	0.110***	1.2	-0.069***	-0.7***	-0.058***	-0.8***
Big	0.167***	2.8	-0.100***	-1.6***	-0.075***	-1.4***
Med	0.095***	0.9	-0.073***	-0.4***	-0.051**	-0.4**
Small	0.102***	0.1	-0.026***	-0.4***	-0.068***	-0.7**
Panel B: AT orders						
Full	0.084***	1.2	-0.062***	-0.5***	-0.019**	-0.4***
Big	0.141***	1.8	-0.086***	-1.4***	-0.029**	-0.8**
Med	0.054***	0.6	0.002	-0.1**	-0.001	-0.5**
Small	0.051***	0.3	-0.071***	-0.0	-0.046	-0.1
Panel C: Non-AT orders						
Full	0.068***	0.8	-0.024**	-0.3***	0.008	-0.1***
Big	0.097***	1.0	-0.043**	-0.7***	0.008	-0.0
Med	0.019***	0.7	-0.021	-0.3**	0.030	-0.1
Small	0.066***	0.7	-0.024	-0.1	-0.029**	-0.2***

Table X: Commonality in liquidity imbalance: AT vs Non-AT

We compare commonality in liquidity imbalance of AT and non-AT traders. 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: $\Delta LiqImb_{it} = \alpha_i + \beta_i \Delta LiqImb_{\mathrm{m,it}} + \varepsilon_{it}$, where $\Delta LiqImb_{it}$ refers to the change in liquidity imbalance flow for firm i, $\Delta LiqImb_{\mathrm{m,it}}$ refers to the contemporaneous change in market-wide liquidity imbalance of orders, excluding firm i's imbalance. We estimate the model separately for AT and non-AT order flow. In Panel A, we compute the difference between ${\beta_i}^{\mathrm{NAT}}$ and ${\beta_i}^{\mathrm{AT}}$ for each firm and report the cross-sectional median of this difference. We also report the median of difference between R2 of non-AT and AT order flow. In Panel B, for each post-event period, we report the median of difference between pre- and post- (${\beta_i}^{\mathrm{NAT}}$ - ${\beta_i}^{\mathrm{AT}}$) and ($R2^{\mathrm{NAT}}$ - $R2^{\mathrm{AT}}$) along with results of inference tests based on a signed-rank test. R2 refers to R^2 (in %). Inferences are based on non-parametric signed-rank test, where *** and ** denote significance at 99% and 95% confidence levels.

	Period I: Pre co-location		Period II: Two months after co-location		Period III: Four months after co-location	
	${\beta_i}^{\mathrm{NAT}}$-${\beta_i}^{\mathrm{AT}}$	$R2^{\mathrm{NAT}} - R2^{\mathrm{AT}}$	${\beta_i}^{\mathrm{NAT}}$-${\beta_i}^{\mathrm{AT}}$	$R2^{\mathrm{NAT}} - R2^{\mathrm{AT}}$	${\beta_i}^{\mathrm{NAT}}$-${\beta_i}^{\mathrm{AT}}$	$R2^{\mathrm{NAT}} - R2^{\mathrm{AT}}$
Full	-0.029***	-0.1**	-0.005	0.0	-0.005	0.0
Big	-0.035***	-0.5***	-0.001	0.0	-0.005	-0.0
Med	-0.040***	-0.1	-0.013	-0.1	-0.004	-0.0
Small	0.017	0.1	-0.001	0.0	-0.006	0.0