Why Fragmented Markets Have Better Market Quality: The Flight of Liquidity Order Flows to Off Exchange Venues

Christine Jiang
Fogelman College of Business and Economics, University of Memphis, Memphis, TN 38152
cjiang@memphis.edu, 901-678-5315

Thomas McInish
Fogelman College of Business and Economics, University of Memphis, Memphis, TN 38152
tmcinish@memphis.edu, 901-217-0448

James Upson*
University of Texas, El Paso
College of Business, El Paso, TX 79968
jeupson@utep.edu, 915-747-7758

This version November 2011.

JEL classification: G14, G18, G19

Key Words: Fragmentation, Market Quality, Off-exchange trading

*Contact Author

Abstract:

We analyze and compare the information quality of order flows on the exchange and on off-exchange trading venues reporting to Trade Reporting Facilities. We find that off-exchange order flow has significantly lower information quality, lower effective spreads, and a significantly higher percentage of trades executing inside the quote, compared to exchange order flow. Our results are consistent with the notion that as uninformed liquidity traders segment order flow to off-exchange venues, there is a higher proportion of informed traders on the exchanges, leading to an improvement in the price discovery process and market quality on the exchanges. Use of off-exchange venues is higher with increased market speed and trading intensity, but decreases with higher intraday volatility.

1. Introduction

In the U.S. trades are executed on exchanges or on off-exchange trading venues such as dark pools, Electronic Trading Networks (ECNs), crossed orders within brokerage firms, and numerous others. Exchanges report their own trades, but off-exchange trades are reported to Trade Reporting Facilities (TRFs). O'Hara and Ye (2011) use TRF volumes to measure market fragmentation on a stock by stock basis. These authors conclude that more fragmented markets are more efficient, enjoy better execution quality, and have lower transaction costs, resulting in improved market quality.[1] O'Hara and Ye (2011) conjecture that their results are due to the fact that while trading is spatially fragmented, U.S. equity markets are actually virtually consolidated. However, these authors do not test this conjecture.

We develop and test an alternate explanation based on the observations of Subrahmanyam (1991) and Holden and Subrahmanyam (1992) that as the proportion of informed traders increases in a market, the market becomes more efficient. Specifically, we conjecture that as uninformed traders are able to segment their order flow to off exchange venues, a larger proportion of trades on the exchanges are informed, improving the price discovery and market quality of the overall market. We distinguish between fragmented markets for which the distribution of informed and uninformed traders (Kyle, 1985, and Glosten and Milgrom, 1985) across trading venues is similar and segmented markets that have differing proportions of informed and uninformed traders.

Our analysis focuses on the information and transaction cost differences between trades executed on and off exchanges. We show that off-exchange order flow is significantly less informed, indicating that it is dominated by uninformed liquidity traders. Specifically, the information share (Hasbrouck, 1995) of exchange trades is roughly 10 times higher (0.902 for exchanges versus 0.098 for off exchanges). We define the Information Ratio of the exchange (off-exchange) trades as the ratio of the information share of the exchange (off exchange) trades to the volume share of exchange (off exchange) trades. An information ratio greater than 1.0 indicates that the order flow carries information above and beyond its

[1] O'Hara and Ye (2011) based their analysis on data from 2008. Weaver (2011), using data from October, 2010, finds that market quality decreases with higher market fragmentation.

simple volume share. Our results indicate that information quality of exchange trades is significantly higher than that of off-exchange trades, with an information ratio of 1.125 for the exchange trades compared to an information ratio of 0.495 for off-exchange trades.

We also evaluate the transaction cost differences between exchange and off-exchange trades. Easley and O'Hara (1987) develop a model that shows that informed traders pay higher spreads than uninformed traders and that market makers are able to differentiate between the informed and uninformed by their differing trade sizes. In post Reg NMS markets trade sizes have dropped significantly and the pooling model of Back and Baruch (2007) indicates that informed traders will match the distribution of trade sizes of uninformed liquidity traders to hide their trading. However, the essence of the Easley and O'Hara (1987) conclusion is that if market makers can effectively distinguish between informed and uninformed traders, spreads will be lower for the uninformed. Our results support this conclusion. Effective spreads of off-exchange trades are significantly lower than those of exchange trades, while the realized spreads of off-exchange trades are significantly higher, supporting the information share analysis that off-exchange order flow is dominated by uninformed trading. We confirm this finding using the spread decomposition model of Madhanvan, Richarson, and Roomans (1997), which again shows the effective spreads for off-exchange trades are lower than those of exchange trades. If informed traders split large orders into small trades to pool with small uninformed traders, as the pooling model of Back and Baruch (2007) indicates, we expect a higher serial correlation of exchange order flow as the split trades of the informed continue to execute on one side of the market to complete the full position. We confirm this expectation.

Our empirical research compliments burgeoning theoretical and empirical work on dark pool trading strategies since dark pools account for a substantial share of total off-exchange trading. Examples of this literature are Zhu (2011), Ye (2011), Ready (2010), and Buti, Rindi, and Werner (2011). Although the Daily Trade and Quote (DTAQ) database does not specifically identify dark pool trades, these trades are reported through TRFs. The literature typically characterizes dark pool trades as trades that occur at a price derived from the exchange market at the midpoint of the prevailing quote.

We analyze where exchange trades and off-exchange trades occur on the quote price grid and find that 31.78% of off-exchange trades occur inside the National Best Bid and Offer (NBBO), which is substantially more than the 7.22% of exchange trades that are inside the NBBO. In addition, 14.77% of off-exchange trades are at the NBBO quote midpoint compared to 2.96% for exchange trades. These results indicate that a substantial portion of off-exchange trading is likely from dark pools.

We also find that as markets become faster, traders shift volume to the off-exchange trading venues. [2] Buti, Rindi, and Werner (2011) develop a model that shows that order flow will migrate to dark pools when exchange liquidity is high so that depth is high and spreads are narrow. We do not find support for this conjecture. Also, we find that when intraday volatility is high, uninformed traders send more order flow to exchanges and execute less off exchanges. We believe that this is because during a period when prices are changing rapidly, uninformed traders seek to reduce the opportunity costs created by delayed or non-execution of off-exchange trades. We find that when trading intensity is high, off-exchange volume increases. High relative volume increases the probability of trade execution on off-exchange trading venues, drawing more trades away from the exchanges.

We use regression analysis to evaluate the impact of off-exchange trading on price discovery. Consistent with the primary finding of Zhu (2011), our results indicate that exchange order flow becomes more informed as off-exchange volume increases. In other words, as uninformed traders migrate to off-exchange trading venues, the percentage of informed traders remaining at exchanges increases, improving price discovery on the exchanges. Ye (2011) finds that a proportion of informed traders will migrate to dark pool trading because of lower cost and trading anonymity. We show that Intermarket Sweep Order (ISO) volume reported through TRFs is informed, although our data does not let us identify these trades

[2] On March 10, 2010, the NYSE implemented a significant upgrade to the market computer system. This upgrade dramatically increased the market speed. The liquidity impact of this event is explored in Jiang, McInish, and Upson (2011).

as dark pool trades.[3] This result indicates that a percentage of off-exchange volume is informed even though the information content is focused on smaller firms. Other factors that impact the information quality of exchange trades are market speed, intraday volatility, market liquidity, and trading intensity. As an additional test of the relative importance of exchange- and off-exchange trades, we implement the order imbalance regression technique of Chordia, Roll, and Subrahmanyam (2002) and Chordia, Subrahmanyam (2004). Specifically, we test whether returns are driven by order imbalance on exchanges, off exchanges, or both. Our findings indicate that exchange order imbalance is significant in the return generating process. Contemporaneous exchange order imbalance is significantly positively correlated with the current return, while lagged exchange order imbalance is significantly negatively correlated with the current return. Contemporaneous off-exchange order imbalance is mostly not significant, and only the first lagged off-exchange order imbalance is negative and significant at the 5% level, further supporting our view that off-exchange volume is dominated by uninformed traders. Off-exchange order imbalance results are primarily driven by the small stocks.

Overall, our results indicate that the price discovery on exchanges improves in fragmented markets because uninformed traders are able to self segment their order flow to off-exchange trading venues, leaving a larger proportion of informed traders at the exchange. This consequence was not the result of a conscious regulatory choice, but a byproduct of the introduction of Reg NMS. Particularly, the order protection rule and the mandate to improve intermarket communications create a higher risk for uninformed traders at exchanges. Rather than playing a losing game at the exchanges, uninformed traders choose to migrate to off-exchange trading venues, where the informed seldom trade.

2. Hypotheses Development

We investigate the reason for the increase in market efficiency in more fragmented markets reported by O'Hara and Ye (2011). Our principal hypothesis is based on the following proposition: If the

[3] Details of ISO trades can be found in Chakravarty, Jain, Upson, and Wood (2011). They find that ISO order flow is dominated by informed traders. ISO trades have a higher information share relative to their volume, higher effective spreads, but lower realized spreads, and better execution quality than non-ISO (INSO) trades.

proportion of informed and uninformed traders is constant across trading venues, the proportion of informed traders in the overall market can be defined as $\alpha / (\alpha + \gamma)$, where α is the proportion of informed traders and γ is the proportion of uninformed traders. If a fraction of uninformed traders can credibly segment their trading on a subset of trading venues, then the proportion of informed traders in the remaining trading venues is $\alpha / (\alpha + \gamma - \phi)$, where ϕ is the proportion of uninformed that segment their trading from the primary trading venues.

Uninformed traders might credibly segment their trading in many ways. Off-exchange venues are less attractive to informed traders because of higher execution risk (Zhu, 2011). Similarly, informed traders face prohibitive costs of trading on retail trading venues such as TD Ameritrade and Scotttrade. Additionally, informed traders may not have access to internalized order flow where only unmatched volume (that can be in the opposite direction of the informed trades) is available for trading. Hence, the proportion of informed/uninformed trades might well be different for trades executed on and off exchanges. We are not able to observe directly the proportion of informed and uninformed traders. Therefore, we develop several tests to empirically determine the differences in the information content of the order flows when markets are segmented.

Based on these considerations, we test the following hypotheses:

Hypothesis 1: *Exchange order flow has higher information content than off-exchange order flow.*

If a significant portion of uninformed traders migrate to off-exchange venues, one result will be an increase in the proportion of informed traders on the exchanges. Zhu (2011) develops a model that predicts that dark pools are more attractive to liquidity traders; however informed traders prefer exchanges because of their lower execution risk. In addition, Holden and Subrahmanyam (1992) show that the greater the concentration of informed traders, the faster information is compounded into asset prices. Further, Subrahmanyam (1991) indicates that "price efficiency may be decreasing in the amount of liquidity trading in the market". As uninformed liquidity traders move order flow to off-exchange venues, price discovery at the exchanges is predicted to increase. To examine the relative contribution of exchange and off-exchange trading to price discovery, we use the Information Shares (IS) approach of Hasbrouck (1995).

The migration of uninformed traders to off-exchange venues is also likely to impact the transaction cost of these venues. Normally, highly informative orders contribute to better price discovery, but also tend to worsen adverse selection, resulting in wider spreads and higher price impact. Hence, we test the following hypothesis:

Hypothesis 2: *Adverse selection costs of off-exchange trades are lower than those of exchange trades.*

Our next hypothesis is based on the theoretical work of Easley and O'Hara (1987) which shows a positive relation between trade size and spreads. The intuition of their model is that small liquidity traders are able to credibly signal the market maker that they are uninformed based on the small size of their trade, resulting in a small transaction cost. The key finding of the Easley and O'Hara model is that uninformed traders should have a lower transaction cost if they can creditably signal that they are uninformed. By migrating to off-exchange venues, uninformed traders can signal their lack of information and lower their adverse selection costs. O'Hara and Ye (2011) report that higher fragmentation is associated with faster execution, lower transaction costs, and more efficient prices, results seemingly contradict to our predictions. However, their analysis is at the stock level with stocks of otherwise similar characteristics forming their matched sample. The difference in the level of fragmentation drives the differential trading cost. The more fragmented stocks naturally have a high percentage of liquidity trades off the exchanges, resulting in a lower overall cost. Hence, viewed in this way, their results are entirely consistent with our predictions, which leads to the following hypothesis:

Hypothesis 3: *When execution quality is lower on exchanges, liquidity traders migrate to off-exchange trading venues.*

Execution costs and adverse selection costs are only applicable when trades are executed; indicating that if the probability of off-exchange trade execution is low, volume will not shift to the off-exchange venue. There can also be significant opportunity costs if trade execution is delayed. Therefore, the liquidity provision on exchanges can impact whether orders are attracted to or diverted from off-exchange venues. In general, incentives for shifting to off-exchange venues are higher when the execution quality is lower on the exchanges.[4]

Admai and Pfleiderer (1988) and Foster and Viswanathan (1990) develop models showing that discretionary liquidity traders can move their trading in a temporal fashion to avoid informed traders. These liquidity traders can trade at mid-day rather than at the market open or close, or delay trading until later days of the week to avoid informed traders earlier in the week. By moving away from periods dominated by informed traders, discretionary liquidity traders allow the information of the informed to be compounded into prices. Even though both of these models are based on a single market, the intuition that discretionary liquidity traders can time their trading to avoid trading with informed traders is applicable to our analysis. By segmenting uninformed trading to off-exchange venues, uninformed traders avoid trading with informed traders. There may be an additional benefit to uninformed traders migrating to off-exchange venues–improved price discovery on the exchanges.

Hence, we test the following hypothesis:

Hypothesis 4: *Uniformed traders migrate to off-exchange venues, increasing the information content of exchanges' order flow.*

Chordia, Roll, and Subrahmanyam (2002) focus on the impact of market-wide order imbalance, but our focus is at the stock level as in Chordia and Subrahmanyam (2004). In their model, uninformed discretionary liquidity traders split order flow across trading periods, leading to a temporal dependence in price pressures. Under our paradigm, a portion of discretionary liquidity traders migrate to off-exchange venues, removing price pressure from the market. The predicted empirical result is that contemporaneous and lagged off-exchange order imbalances will have no impact on returns. However, for exchanges contemporaneous order imbalance will be positive and significant and lagged order imbalance will be negative and significant. Hence, we test the following hypotheses:

Hypothesis 5(a): *Exchange order imbalance predicts future stock returns.*

[4] In the case of dark pools, the cost savings will be proportional to the costs from trading on the exchanges as the price in dark pools is often at the midpoint of the NBBO.

Hypothesis 5(b): *Off-exchange order imbalance does not predict future stock returns.*

On March 10, 2008, the NYSE updated its computer systems that process trades and quotes, significantly reducing within system latencies by roughly 800 milliseconds. This upgrade allowed centrally located traders to respond more quickly to changes in market conditions, which increase the ability of low latency traders to profit from trading with high latency traders [5] Hence, the incentive for high latency traders to migrate to off-exchange venues increased.

In addition, Regulation NMS introduced a new order type to the market called an Intermarket Sweep Order. Chakravarty et al. (2011) show that traders using this order type are much more likely to be informed than traders using other order types. Therefore, when appropriate, we condition our analysis on the speed change and whether or not an ISO order was used.

3.0 Sample, Data, and Methods

3.1 Sample

Our sample comprises all common stocks in the DTAQ dataset for the first 6 months of 2008. We also require a minimum of 300 trades each day so that we can implement the information shares approach of Hasbrouck (1995). We classify our sample into quartiles based on market capitalization on the first day of 2008. We obtain our final sample by randomly selecting 50 stocks from each quartile.

Table 1 shows selected statistics for exchange and off-exchange volume and trade size. We report the mean and standard deviation of each measure for the full sample and by firm size. Volume is reported in thousands of shares. About 25% of our sample volume is executed on off-exchange venues.[6] This is comparable to the 27% off-exchange volume found in O'Hara and Ye (2011), and similar to the

[5] McInish and Upson (2011) show that fast traders are able to use their speed advantage to earn arbitrage profits against slow liquidity traders at an estimated level of $281 million per year.
[6] Off-exchange trading is not unique to the U.S. market. Gomber and Pierron (2010)report that the activity on dark pools, crossing networks and OTC is approximately 40% of total traded volume in 2008-2009 for the European equity markets

proportion reported in Weaver (2011) using more recent data.[7] Consistent with the findings of O'Hara and Ye (2011) for NYSE stocks, the largest size quartile off-exchange volume is 26% compared to 22% for the remaining quartiles. The higher off-exchange fragmentation for large stocks on NYSE is a reflection of competition for institutional order flows from alternative trading venues (Conrad, Johnson, and Wahal, 2003). The average trade size of off-exchange trades is larger than that for exchange trades. One possible explanation for the larger trade size of off-exchange trades is that these trades have smaller price impacts than similar sized exchange trades. Hence, uninformed traders can increase trade size, reducing order submission costs per share, but not adversely impacting market prices. We will provide evidence supporting this conjecture.

3.2 Data

We obtain data from the Daily Trade and Quote (DTAQ) dataset, which has time stamps to the millisecond, extensive condition codes, and includes the exchange-calculated NBBO time stamped to the millisecond. For some of our analysis, we condition on whether the trade type is an Intermarket Sweep Order (ISO) or a non-Intermarket Sweep Orders (NISO). ISO trades are identified as condition code F. Chakravarty et al. (2011) show that ISO trades are dominated by informed institutional traders and conditioning on trade type can give added insight into the trading structure of the market.

3.3 Methods

It is important in our analysis to get the best possible trade inference of buyer- and seller-initiated trades. We apply the technique of Lee and Ready (1991) to infer direction, but the quality of this inference will be greatly impacted by our ability to select the quote that is in force when the trade is executed. We need to align the trades and NBBO quotes, but trades and quotes are handled through separate computer systems and are provided in separate files. Matching by time stamps is not sufficient to

[7] Two market centers at the time of our study, BATS and DirectEdge, reported through the TRF facility, but received exchange status under the Weaver (2011) study. If these volumes are included in off-exchange volume for the Weaver analysis, matching the market configuration of our study, off-exchange volume would be 50.83%.

accurately merge the trades and quotes because the differing computer systems for trades and quotes can introduce shift in the time stamps of each entity. We develop and implement the following procedure.

In today's high speed communication systems, smart routers, co-located computer systems, and sophisticated algorithms are all applied to rout trades to the market(s) posting the NBBO. Computers execute most trades without human intervention, matching incoming orders with posted limit orders. Hence, for each stock day in the sample we test time lags from 0 to 1,500 milliseconds, in 25 millisecond increments.[8] For each stock for each day, we select the lag time that maximizes the number of trades executed at the NBBO price.[9]

We apply this approach to exchange and off-exchange trades independently. Exchanges have knowledge of their own trades immediately, but TRFs must receive the report of an off-exchange trade from the trading venue, resulting in a time delay of unknown duration. Further, off-exchange venues have time window within which to report trades.

Since more and more researchers are using the DTAQ dataset, we digress from the main focus of our research to detail the impact of the trade quote alignment process. Recall that on March 10, 2008, the NYSE implemented a significant upgrade to its computer system. Figure 1 shows plots for the percentage of trades that is executed at the NBBO price, as a function of the lag time between the NBBO quote and the trade time stamp, both before and after the speed change, conditioned on exchange trades and off-exchange trades. For exchange trades before the speed change, Figure 1, Panel A, shows that if there is a zero time lag applied to the reported NBBO quote time, 60% of trades would occur at the NBBO, 32% would be between the NBBO quotes, and the remainder would be outside the NBBO quotes. The number of trades at the NBBO is maximized at a lag of about 900 milliseconds, at which point over 80% of trades are executed at NBBO quotes, less than 10% executed inside the NBBO spread, and less than 10% outside of NBBO quotes. Trades outside of the NBBO quote are most likely not trade through violations of the Order Protection Rule, Rule 611, of Reg NMS. Trades can occur outside the NBBO based on the

[8] The use of various time lags to help improve trade inference and subsequent transaction costs estimates is also applied in Bessembinder (2003).
[9] This method was first applied in McInish and Upson (2011).

Flick Quotes Exemption, which defines the reference price for establishing a trade through as the least aggressive NBBO ask and bid prices over the previous one second of trading (McInish and Upson, 2011). In addition, ISO trades are allowed to trade through NBBO prices without violation of the Order Protection Rule (Chakravarty et al., 2011).

Figure 1, Panel B, shows the alignment results for off-exchange trades. While the lag that yields the highest proportion of trades at the NBBO is similar to that for Panel A, the distribution of trade location is significantly different. At the optimal lag, almost 35% of trades are executed inside the NBBO quote, which, we believe, reflects the impact of Dark Pool trading reported through the TRFs.

Figure 1, Panels C and D, show the alignment after the NYSE speed improvement. Clearly, there is a dramatic shift in the time lag that maximizes trades at NBBO quotes. This alignment shift is similar for exchange and off-exchange trades. We estimate that in system latencies at the NYSE decreased by as much at 800 milliseconds due to the system upgrade. This reduction in latency translates to market participants, particularly centrally located market participants, obtaining a clearer, more up-to-date picture of market conditions.

In Figure 2 we plot the time series of the location of prices on the NBBO price grid for exchange and off-exchange trades, using the lag that maximizes the proportion of NBBO trades. Figure 2, Panel A, shows the exchange results and Figure 2, Panel B, shows the off exchange results. Figure 2 shows that the Trade Maximizing Lag Method (TML) generates a relatively consistent level of trading at, inside, and outside of the NBBO quote for both exchange and off-exchange trades. We feel that this method gives a better alignment between trades and NBBO quotes, for the purpose of trade inference and calculations of transaction costs. While we do not claim that the TML method generates a perfect match between trades and NBBO quotes, we do believe that this approach significantly improves the alignment of trades and quotes and recommend its application when using the DTAQ database.

4.0 Results

4.1 Off-exchange time series

We begin our exposition on the impact of off-exchange volume on the information structure and transaction costs of the market by first giving a general sense of the variability of off-exchange use. Figure 3 shows a time series plot of the daily average percent of volume reported through the TRFs conditioned on firm size ranked from smallest (rank 1) to largest (rank 4). Prior to the NYSE speed increase, off-exchange volume is stable at roughly 16%. Figure 3 shows significant variation in the percentage of total volume executed on off-exchange venues after the reduction in latency. Off-exchange volume increases steadily until May of 2008 then subsequently stabilizes at about 23%. Figure 3 shows that exchange latency as well as market conditions might impact traders order routing strategies.

4.2 Trade price grid location

Table 2 shows trade prices relative to the NBBO for exchange and off-exchange trades. Table 2, Panel A, shows the results for all trades, while Table 2, Panels B and C, are conditioned on whether the order was an ISO. For each grid point, we test the null hypothesis of equality of means for exchange and off-exchange trade using a paired t-test. For the full sample, only 7.22% of exchange trades are inside the NBBO quote, while 31.78% of off-exchange trades are inside the quote. Naturally, quotes at the mid-point are inside the NBBO, but we show this particular price point separately in Table 2. Only 2.96% of exchange trades occur at the NBBO quote midpoint compared to 14.77% of off-exchange trades.

Note that it is not necessarily true that only 14.77% of trades are from Dark Pools. As shown in McInish and Upson (2011), the latency of quote transmission will impact execution prices on a venue. For example, if Dark Pool 1 has a quote latency of 500 milliseconds and Dark Pool 2 has a latency of 0 milliseconds, the quote mid-point of Dark Pool 1 will be 500 milliseconds in the past of Dark Pool 2. If this results in the two Dark Pools trading at different prices at the same instant, the trades of Dark Pool 1 will not be at the instantaneous NBBO midpoint. The TML method averages the unobservable latencies for all market venues reporting to the TRF.

We also condition the grid analysis based on trade type, ISO and NISO. ISO trades are more aggressive than NISO trades and have higher transaction costs (Chakravarty et al., 2011). Table 2, Panel B, shows that off-exchange ISO trades have a higher percentage of trades inside the NBBO quotes relative to exchange ISO trades. However, the highest percentage of inside quote trades are observed for off-exchange NISO trades. For example, at the exchanges 6.90% of ISO trades are inside the NBBO compared to 11.24% off-exchange. For NISO trades, 7.52% of exchange trades are inside the NBBO, compared to 36.98% for off-exchange trades. We will show later in the paper that off-exchange ISO trades have high information quality and represent the second highest information quality of the market order flows. These results indicate that traders who rout trades to venues that report through the TRFs have significantly different expectations of execution prices compared to those trading on exchanges. In addition, the price grid results indicate a reasonable distribution of trades based on the TML alignment.

4.3 Information Quality

To evaluate the information quality of exchange and off-exchange order flow, we estimate information shares using the method of Hasbrouck (1995). First, we estimate information shares for two price channels–one for exchange and one for off-exchange trades. We use the last trade price of each trade type in each second.[10] Second, we estimate information shares for four price channels–exchange ISO, off-exchange ISO, exchange NISO, and off-exchange NISO. Unless the resulting variance co-variance metric is diagonal, the information share estimate for each trade type is not exactly identified. Therefore, we average the upper and lower bound values.

Hasbrouck (1995) finds that the price discovery process is under represented on regional stock exchanges because the information share of these exchanges is well below the traded volume of shares executed on these exchanges. When evaluating the information quality of order flow, the information

---

[10] Although the DTAQ database has time stamps to the millisecond, the number of observations generated for the method is computationally prohibitive. The use of trade prices follows Hasbrouck (2003), Anand and Chakravarty (2007), and Goldstein, Shkilko, Van Ness, and Van Ness (2008).

share of a given order flow must be conditioned by the proportion of volume attributable to this venue or venue trade type. We define the information ratio (*InfoRatio*) as:

$$InfoRatio_i = \frac{InfoShare_i}{VolumeShare_i} \tag{1}$$

$InfoShare_i$ is the point estimate of the information share for a volume flow (exchange or off-exchange trades) and $VolumeShare_i$ is the percentage of volume over the period, relative to total executed volume. An information ratio greater than 1.0 (below 1.0) indicates that the volume flow carries information greater than (less than) that implied by only the volume share and has higher (lower) information quality.

Our results are shown in Table 3. We report results for the full sample and by firm size but focus our discussion on the full sample results. For this table the information ratio is estimated by first averaging the information share and volume share for each stock, over the sample period, and then calculating the ratio of these two variables. This method controls for the effect of potential outliers. Exchange trades represent 80.2% of the volume, but have an information share of 0.902. The mean information ratio of exchange volume is 1.125, which indicates that exchange volume carries an information level 12.5% greater than its volume indicates. Off-exchange trades represent 19.7% of the volume, but only have an information share of 0.098. The information ratio of off-exchange volume is 0.495, indicating that the information quality of off-exchange volume is much lower than would be expected based on its volume share.

Our four channel results reported in Table 4 are even stronger. Exchange ISO volume is only 22.7% of total volume, but has an information share of 0.359. The information ratio of exchange ISO volume is 1.594, or 59% higher than its volume level would indicate. Exchange NISO volume has an information ratio of 0.990 so that the contribution of NISO volume to price discovery is almost exactly proportional to the volume share.

Off-exchange NISO volume is the least informed, representing 16.5% of total volume, but having an information share of only 0.05. Off-exchange ISO volume is more difficult to interpret. On average off-exchange ISO volume represents 3.2% of the total, with an average information share of 0.032.

However, the average daily information ratio of the sample is 1.178, indicating that off-exchange ISO volume is, at least for some stocks, more informed. The firm size based results indicate that off-exchange ISO order flow is most informed for smaller sized firms. Our expectation is that the positions desired by informed traders for small stocks will be smaller than the positions desired for large stocks. The lower potential liquidity for small stocks off-exchange might be sufficient for informed traders to access, improving the information quality of this flow. The liquidity provision for large stocks off-exchange might be too small to make a significant contribution to the net position required, and, therefore, the majority of trading is on the exchanges.

These results strongly support our first hypothesis, that exchange order flow is more informed than off-exchange order flow. As a result of the migration of uninformed to off-exchange venues, the information quality of off-exchange volume is much lower than the information quality of exchange volume. Although our results indicate that off-exchange volume is dominated by uninformed traders, we also find evidence that, to some degree, informed traders take advantage of the liquidity offered at off-exchange venues. Off-exchange ISO order flow has higher information quality than off-exchange NISO order flow, and, for small stocks, off-exchange ISO order flow represents the highest information quality order flow in the market. This result is consistent with Ye (2011) who predicts that under certain conditions informed traders prefer to trade on off exchange venues.

4.4 Adverse Selection Cost Analysis

4.4.1 Effective Spreads, Realized Spreads, and Execution Quality

Our second hypothesis states that adverse selection costs will be lower for off-exchange trades than for exchange trades. By segmenting order flow to off exchange venues, uninformed traders can credibly signal that they are uninformed. With this credible signal, the model of Easley and O'Hara (1987) indicates that off-exchange trades will have lower execution costs. Table 5 shows results for all trades and conditioned on trade type. The effective spreads for exchange trades are a statistically significant 0.13 cents higher than for off-exchange trades. Further, both ISO and NISO effective spreads are higher for trades on exchanges than on off-exchanges, supporting hypothesis 2.

We also evaluate the realized spreads based on 5 and 30 minute reference points. Considering all trades and the results conditioned on NISO trades, realized spreads of exchange trades are significantly smaller than that of off-exchange. This result supplies additional evidence that off-exchange volume has lower information quality than exchange volume. However, off-exchange ISO trades have significantly lower realized spreads than exchange ISO trades, indicating that off-exchange ISO trades may be more informed than those of the exchanges. This result also is supported by the information ratio analysis on the information quality of off-exchange ISO order flow.

In addition to effective and realized spreads, we also evaluate the Preferencing Measure (PM) proposed by He, Odders-White, and Ready (2006). PM is the ratio of realized spreads to effective spreads. With a significant percentage of trades executing at the quote midpoint (effective spread=0), we first calculate the daily trade weighted effective spread and realized spread and then take the ratio of these values to obtain one observation per stock day of our sample. The lower the PM level, the better the execution quality of the trade.

Again, for all trades and for each trade type, execution quality of off-exchange trades is significantly worse than for exchange trades. The key point is that the execution quality of exchange trades is better because of their substantially lower realized spreads. Uninformed traders cannot simply rout trades to the exchanges to improve execution quality since they are uninformed. Rather, the uninformed reduce effective spread costs by routing trades to off-exchange venues, which is their best cost minimizing strategy. We find that the best execution quality of trades is for off-exchange ISO order flow. Here the PM is -0.80 compared to the PM for exchange ISO order flow at -0.09. For exchange versus off-exchange trades, our spread analysis indicates that transaction costs are significantly lower off-exchange, supporting our second hypothesis.

4.4.2 Spread Decomposition

The preceding spread analysis does not take into consideration potential serial correlation of signed order flow with buys following buys or sells following sells. This serial correlation can lead to inaccurate estimation of the effective spreads as indicated by Madhaven, Richardson, and Roomans (1997) (henceforth, MRR). MRR propose a regression method that decomposes the spread and explicitly takes into consideration correlated signed order flow. We are interested in the following parameters derived from the regression, S, the implied spread, S^E, the implied effective spread, r, the asymmetric information component of the spread, and ρ, the serial correlation of signed order flow. We segment the 125 day sample period into five 25 day sub-periods. To control for intraday spread effects, we also divide each day into six 105 minute segments. The model is estimated for each day for each segment, with 25 days of data. We estimate the model for the exchange and off-exchange venues, and report the results in Table 6. In addition, we estimate the model for each trade type and report the results in Table 7.

Table 6 indicates that exchange transaction costs are significantly higher than for off-exchange trades even after accounting for serial correlations in order flows. The implied spread, S, is 1.96 cents for the exchanges, but only 0.49 cents off-exchange. Also, the implied effective spread, S^E, is much lower off-exchange with savings of close to one cent per share in transaction costs. Our results indicate that the asymmetric component of the spread is not significantly different between the exchanges and off-exchange. Nevertheless, the lower effective spread of off-exchange trades implies that the dollar value of the asymmetric information component is smaller than for the exchanges. The last parameter of Table 6 is ρ, the serial correlation of signed order flow. Our results show that the serial correlation of signed order flow is significantly larger for exchanges compared to off-exchange, 0.53 versus 0.36. We interpret this result as follows. Exchange trades represent small parts of larger, possibly institutional, orders. Assuming that the Back and Baruch (2007) pooling model is correct and informed traders divide larger orders into small trades to pool with uninformed traders, the exchange order flow will have a higher serial correlation because the trades directionally remain the same as the total order size is filled. The smaller ρ for off-exchange trades is consistent with uninformed liquidity traders trading on both sides of the market for purely liquidity needs.

Table 7 shows the MRR regression results by trade type. We compare exchange ISO (NISO) trades against off-exchange ISO (NISO). For ISO trades the implied spread, S, and implied effective spread, S^E, are smaller for the off-exchange trades compared to exchange ISO trades. The asymmetric component of the spread, for both exchange and off-exchange ISO trades is greater than 1.0. This result is consistent with the negative PM evaluation for ISO trades in the previous section and indicates that liquidity suppliers consistently lose to ISO trade initiators, regardless of the venue the trade is executed on. The larger asymmetric component of off-exchange ISO trades indicates that liquidity suppliers lose relatively more to off-exchange ISO trade initiators. We feel that off-exchange ISO trades represent the "Sharks in the Pool" on off exchange venues for smaller firms. The serial correlation parameter of exchange ISO trades is significantly higher than for off-exchange ISO trades. This might be the result of larger orders being worked at the exchanges and smaller orders being worked at the off-exchange venues, or because informed traders initiate a trade series to fill an order at the off-exchange venues to fill what they can and then move to the exchange to fill the greater part of the order because of lower depths on off-exchange venues. Though not reported in the table, we note that the serial correlation of off-exchange ISO trades is significantly larger than off-exchange NISO trades.

The implied spread of off-exchange NISO trades is statistically larger than the implied spread for the exchanges. This result may indicate that the cost of posted liquidity, such as liquidity at ECNs, might be more expensive than on the exchanges. However, the implied effective spreads of off-exchange NISO trades is significantly less than exchange NISO trades, with a difference of 0.42 cents. The asymmetric information component of the spread is less than 1.0 for both ISO and NISO; however the component is significantly higher off-exchange. This might be in response to the "Sharks in the Pool" off-exchange ISO trade initiators. Serial correlation of NISO trades is also higher on the exchanges compared to off-exchange. Overall our results support hypothesis 2, that spreads are lower off-exchange because the venues are credibly dominated by uninformed liquidity traders.

4.5 Determinates of Choice of Venue

Hypothesis 3 states that off-exchange volume is higher when the execution quality is lower on the exchanges. Execution quality can be measured in many dimensions and with various metrics, and we measure quality through liquidity and intraday volatility.[11] In addition, off-exchange volume share will be lower when there is a higher degree of information asymmetry which can lead to higher execution risk on the off-exchange venues. We use idiosyncratic volatility to proxy for information asymmetry. In this section we evaluate the market conditions that lead to an increase in the volume share of off exchange venues. We estimate the following regression:

$$\%TRFvol_{i,t} = \alpha + \beta_1 Spd + \beta_2 Isig_{i,t} + \beta_3 MpVar_{i,t} + \beta_4 Liq_{i,t} + \beta_5 Turn_{i,t} + \varepsilon_{i,t} \qquad (2)$$

where $\%TRFvol_{i,t}$ is the percent of off-exchange volume, *Spd* is a dummy variable that is zero prior to March 10, 2008 and 1 after, *Isig* is the absolute value of the residual from a daily Fama and French 3 factor model regression and proxies for idiosyncratic risk, and $MpVar_{i,t}$ is the NBBO quote mid-point volatility and represents intraday volatility. $Liq_{i,t}$ is a composite liquidity measure at the market level defined as $100*(NBBO_{ask}-NBBO_{bid})/NBBO_{Depth}$, where $NBBO_{Depth}$ is the aggregate quoted depth from all market centers with prices that match the NBBO. $Turn_{i,t}$ is the stock turnover defined as traded volume divided by the number of shares outstanding. This coefficient is multiplied by 100 for reporting. We estimate equation 2 as a fixed effects regression and then, as a robustness check, at the stock level, reporting the average coefficients and testing if they are significantly different from zero. All measures are at the daily level.

Table 8 shows the regression results. In our first regression, only the speed dummy variable is included and the coefficient is 0.039 and significant at the 1% level. This result indicates that the speed increase of the NYSE leads to a 3.9% increase in the proportion of off-exchange volume. Clearly, a

[11] The addition of a direct cost measure, such as effective spreads or the PM in this regression will introduce a significant endogeneity issue in the regression. As uninformed migrate from off-exchange venues to the exchanges their trading will affect the execution costs and execution quality at the exchanges.

significant change in the processing capabilities of a primary exchange like the NYSE has far reaching affects beyond the exchange itself.[12] The second specification, S2, includes only market based conditions. The coefficients of *Liq* and *Isig* are not significant. *MpVar* is significant and negative while *Turn* is significant and positive. We interpret these results as follows. When turnover is high, the higher volume improves the probability of execution off-exchange. Since transaction costs are lower for off-exchange trade execution, uninformed route more volume to off-exchange venues. However, when prices are volatile, missed trade execution or delayed trade execution off-exchange create large opportunity costs. To avoid these costs, the uniformed route relatively more volume to the exchanges to obtain faster execution. Our third specification includes both the speed dummy and measures of market quality and information asymmetry, speed dummy remains highly significant and the results on other variables are similar to those reported in S2.

The stock level regression supports the fixed effects results in that the results are qualitatively similar. The *Isig* variable is significant in this regression. The coefficient is negative and significant at the 1% level. Hence, traders execute fewer off-exchange trades when there is a higher degree of information asymmetry.

4.6 Information Determinates

In section 4.2, we examined the unconditional information quality of exchange and off-exchange trades. In this section, we use regression analysis to better investigate how the migration to off-exchange venues impacts the information quality of exchange order flow. We estimate the following regression:

$$X_{i,t} = \alpha + \beta_1 \%TRFvol_{i,t} + \beta_2 Spd + \beta_3 Isig_{i,t} + \beta_4 MpVar_{i,t} + \beta_5 Liq_{i,t} + \beta_6 Turn_{i,t} + \varepsilon_{i,t} \qquad (3)$$

where $X_{i,t}$ is either the information ratio or information share for stock *i* on day *t* at the exchanges. The other variables are as previously defined. Hypothesis 4 indicates that as uninformed migrate to off-exchange venues; the information quality of exchange trading will improve. This analysis seeks to

[12] Readers interested in a detailed study of this event are referred to Jiang, McInish, and Upson (2011).

discover the factors contributing to the information quality on exchange trading. Our primary variable of interest is *%TRFvol*. The results are show in Table 9. We run a fixed effects regression and for robustness, at the firm level.

Equation 3 contains the endogenous *%TRFvol* variable together with exogenous control variables. In this case OLS can still be used to estimate Equation 3 since all the right hand side variables in Equation 3 are uncorrelated with that equation's error term. In fact, *%TRFVol* is not correlated with the error term because there is no $X_{i,t}$ measure in Equation 2. There is no simultaneity problem because the dependence is not bi-directional, for each equation (equations 2 and 3) it all goes one way (Brooks (2008) pg. 275).

For the information ratio regression, the coefficient for *%TRFvol* is positive and significant at the 1% level in both regression specifications. This indicates that as off-exchange volume increases, the information quality of exchange volume improves, consistent with the prediction of Zhu (2011). However, the same coefficient is negative and significant in the information share regression. As we have show in previous sections of the paper, some quantity of off-exchange volume is informed, so an increase in off-exchange volume will decrease the information share of exchange trading. However, this decrease is smaller than would be expected based on the volume transfer between the two levels. In other words, while off-exchange volume has some information quality, the volume is dominated by uninformed trading.

The coefficient of *Spd* is negative and significant in all regressions. The interpretation of this result is subtle. The speed increase of the NYSE results in the shift of substantial volume to off exchange venues as shown in the previous regression results. Jones, Kaul, and Lipson (1994) show that volatility is driven by the number of trades rather than volume, however as the volume shifts to off-exchange venues, off-exchange venues will also have more trades, increasing the volatility contribution. The information shift of off-exchange volume has two components. The first component increases the information quality of exchange trading because the withdrawal of uninformed traders increases the relative proportion of informed traders at the exchanges. The second component increases the variance contribution off-exchange because there is an increase in the number of trade executions with the shift of volume. The *Spd* dummy is picking up this second component of the information impact of off-exchange volume.

The coefficient of *Isig* and *Liq* are not stable in the regression leaving their importance suspect. Intraday volatility, *MpVar*, is significant and positive in all regressions. *MpVar* can be thought of as the raw value of price discovery. Even though uninformed migrate to the exchanges during days with high intraday volatility; the noise this liquidity trading brings to the price discovery process is unable to mask the price signal of the informed at the exchanges. This positive coefficient indicates that on days with high price discovery, the price discovery process is dominated by exchange trades. The coefficient of *Turn* is negative and significant in the information ratio fixed effects regression, but not significant at the stock level. It is significant and negative in both regressions for the information share. We have previously shown that an increase in turnover leads to an increase in off-exchange volume, but the negative coefficient indicates that the information quality of exchange trades drops. We feel that as turnover increases, a higher level of uninformed trading occurs, perhaps for portfolio rebalancing. Some of this uninformed trading migrates to off-exchange venues, but some of the increase stays on the exchanges, decreasing the information quality of the order flow. Overall, our regression results support Hypothesis 4; an increase in off-exchange volume improves the information quality of exchange trades.

4.7 Order Imbalance

One implication from the Glosten and Milgrom (1985) model is that in the long run, the order imbalance of pure liquidity traders will be zero, ie buys will equal sells in the long run. Although there might be short term order imbalances, market makers need not change prices because at some point the imbalance will reverse.[13] If off-exchange order flow is dominated by uninformed traders, then order imbalances should not have an impact on market returns. We estimate the following regression based on Chorida and Subrahmanyam (2004):

$$R_{i,t} - R_{m,t} = \alpha + \sum_{k=0}^{4} \beta_k ExOB_{i,t-k} + \sum_{k=0}^{4} \beta_{k+5} TrfOB_{i,t-k} + \varepsilon_{i,t} \qquad (4)$$

where $R_{i,t}$ is the return for stock *i* on day *t* and $R_{m,t}$ is the equally weighted return for the market on day *t*. $ExOB_{i,t-k}$ is the volume order imbalance on the exchanges, (BuyVol-SelVol)/(BuyVol+SelVol) and $TrfOB_{i,t-k}$ is the volume imbalance off-exchange. The regression includes the contemporaneous order imbalance and four lagged values for exchange and off-exchange trades. Regressions are run for the full sample and conditioned on firm size. The results are shown in Table 10. Consistent with our Hypotheses 5a and 5b, off-exchange order imbalance has limited significance in the regressions. Only for the small firms does the order imbalance have significance and the predicted sign. This is consistent with the findings that off-exchange ISO order flow is informed for these small firms. On the other hand, both the contemporaneous order imbalance and its first lag have statistically significant impact on returns. Overall our findings support the main position of our paper, that off-exchange order flow is dominated by uninformed traders and that the shift of volume to off-exchange venues improves the information quality at the exchanges.

[13] This statement assumes that the market maker has infinite liquidity and can always outlast any order imbalance, thus avoiding the problem of the gamblers ruin (Garman, 1976).

5.0 Conclusion

O'Hara and Ye (2011) show that efficiency is improved for fragmented markets. We investigate an explanation for these findings. If markets are purely fragmented then the level of informed trading can be defined as $\alpha/(\alpha+\gamma)$ where α is the proportion of informed traders and γ is the proportion of uninformed traders. If some uninformed traders are able to credibly segment their trading onto off-exchange venues, the level of informed trading on exchanges becomes $\alpha/(\alpha+\gamma-\varphi)$, where φ is the proportion of uninformed that are able to segment. Thus, our hypothesis is that price discovery will improve on the exchanges. Zhu (2011) argues that execution risk is higher for informed investors because their trades tend to be on the same side of the order book. Consequently, off-exchange venues such as dark pools, ECNs, and crossing networks attract mostly uninformed traders, leaving the informed trades on exchanges. Our hypothesis is consistent with these observations of Zhu (2011)

We investigate this explanation with a sample of NYSE firms for the first six months of 2008. We focus on the information quality of exchange order flow compared to off-exchange order flow. We show that off-exchange order flow is significantly less informed than exchange order flow. We also show that the information quality of exchange order flow is increasing in the percentage of off-exchange volume reported through the TRFs. In other words as uninformed traders migrate to off-exchange venues, the concentration of informed traders on exchanges increases, improving price discovery.

The model of Easley and O'Hara (1987) indicates that if uninformed traders can credibly signal that they are uninformed, adverse selection costs of their trades will be lower. Our results indicate that off-exchange trades have significantly lower effective spreads than exchange trades. However, execution quality is better for exchange trades. As a robustness test we estimates spreads using the regression model of Madhaven, Richardson, and Roomans (1997) and confirm that off-exchange trades have lower execution costs.

We also investigate the conditions that prompt traders to route order flow to off-exchange venues. We find that as markets become faster, uninformed traders migrate to off-exchange venues. Faster markets give a distinct advantage to informed traders, and uninformed traders move to off-exchange venues to avoid losses to the informed in faster markets. However, when trading intensity is high and prices are volatile, the volume share of off-exchange venues decreases.

We also compare the impact of exchange and off-exchange order imbalance on stocks returns. If off-exchange order flow is dominated by uninformed liquidity traders, then off-exchange order imbalance should not impact returns. If they know that a given order is uninformed, liquidity suppliers should not change prices. Our regression results support this assertion. While contemporaneous and lagged order imbalances on exchanges significantly impact stock returns, contemporaneous and lagged order imbalances at off-exchange venues are mostly insignificant.

Our results indicate that the reason for the observed improvement in market quality, price discovery, and market efficiency in fragmented markets is the ability of uninformed liquidity traders to credibly segment their trading on off-exchange venues. When uninformed traders migrate to off-exchange venues, higher concentrations of informed remain at the exchanges. With fewer uninformed at the exchanges, competitive informed traders are less able to hide demand, then, therefore, trade more aggressively, improving the price discovery process at the exchanges.

References

Admati, A, Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3 – 40.

Anand, A., Chakravarty, S., 2007. Stealth trading in options markets. *Journal of Financial and Quantitative Analysis* 42, 167 – 188.

Back, K., Baruch, S., 2007. Working orders in limit order markets and floor exchanges. Journal of Finance 63, 1589 – 1621.

Bessembinder, H., 2003. Issues in Assessing Trade Execution Costs. *Journal of Financial Markets* 6, 233–257.

Brooks, C., 2008, Introductory Economics for Finance, *Cambridge University Press.*

Buti, S., Rindi, B., Werner, I., 2011. Dark Pool Trading Strategies. Working paper, Ohio State University.

Chakravarty, S., Jain, P., Upson, J., Wood, R., 2011. Clean Sweep: Informed trading through intermarket sweep orders, Journal of Financial and Quantitative Analysis, forthcoming.

Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity, and market returns. Journal of Financial Economics 65, 111 – 130.

Chordia, T., Subrahmanyam, A., 2004. Order imbalance and individual stock returns: Theory and evidence. Journal of Financial Economics 72, 485 – 518.

Conrad, J, Johnson, K., Wahal, S., 2003. Institutional trading and alternative trading system. Journal of Financial Economics 70, 99-134.

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

Foster, F., and S. Viswanathan. 1990, A theory of the intraday variations in volume, variance and trading costs in securities markets. *Review of Financial Studies,* 3), 583 – 624.

Garman, M., 1976. Market Microstructure. Journal of Financial Economics, 3, 257 – 275.

Glosten, L., Milgrom, P., 1985. Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 13, 71 – 100.

Goldstein, M., Shkilko, A., Van Ness, B., Van Ness, R., 2008. Competition in the market for NASDAQ securities. *Journal of Financial Markets* 11, 113 – 143.

Gomber, P., Pierron, A., 2010. MiFID Spirit and Reality of a European Financial Markets Directive. Report Published by Celent.

Hasbrouck, J., 1995. One security, Many Markets: Determining the Contributions to Price Discovery. *Journal of Finance* 50, 1175–1199.

Hasbrouck, J., 2003. Intraday Price Formation in U.S. Equity Index Markets. *Journal of Finance* 58, 2375–2399.

He. C., E. Odders-White, M. Ready. 2006, The Impact of Preferencing on Execution Quality. *Journal of Financial Markets* 9 , 246–273.

Holden, C., Subrahmanyam, A., 1992. Long-lived private information and imperfect competition. Journal of Finance 47, 247 – 270.

Jiang, C., McInish, T., Upson, J., 2011. Liquidity and market stability: The impact of market speed. Working paper, University of Texas at El Paso.

Jones, C., Kaul, G., Lipson, M., 1994. Information, trading, and volatility. Review of Financial Studies 36, 127 – 154.

Kyle, A., 1985. Continuous Auctions and Insider Trading. *Econometrica* 53, 1315–1336.

Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. Journal of Finance 46, 733 – 747.

Madhavan A., M. Richardson and M. Roomans, 1997. Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks. *The Review of Financial Studies* 10, 1035-1064.

McInish, T., Upson, J., 2011. Strategic liquidity supply in a market with fast and slow traders. Working paper, University of Texas at El Paso.

O'Hara, M., Ye, M., 2011. Is market fragmentation harming market quality? Journal of Financial Economics 100, 459 – 474.

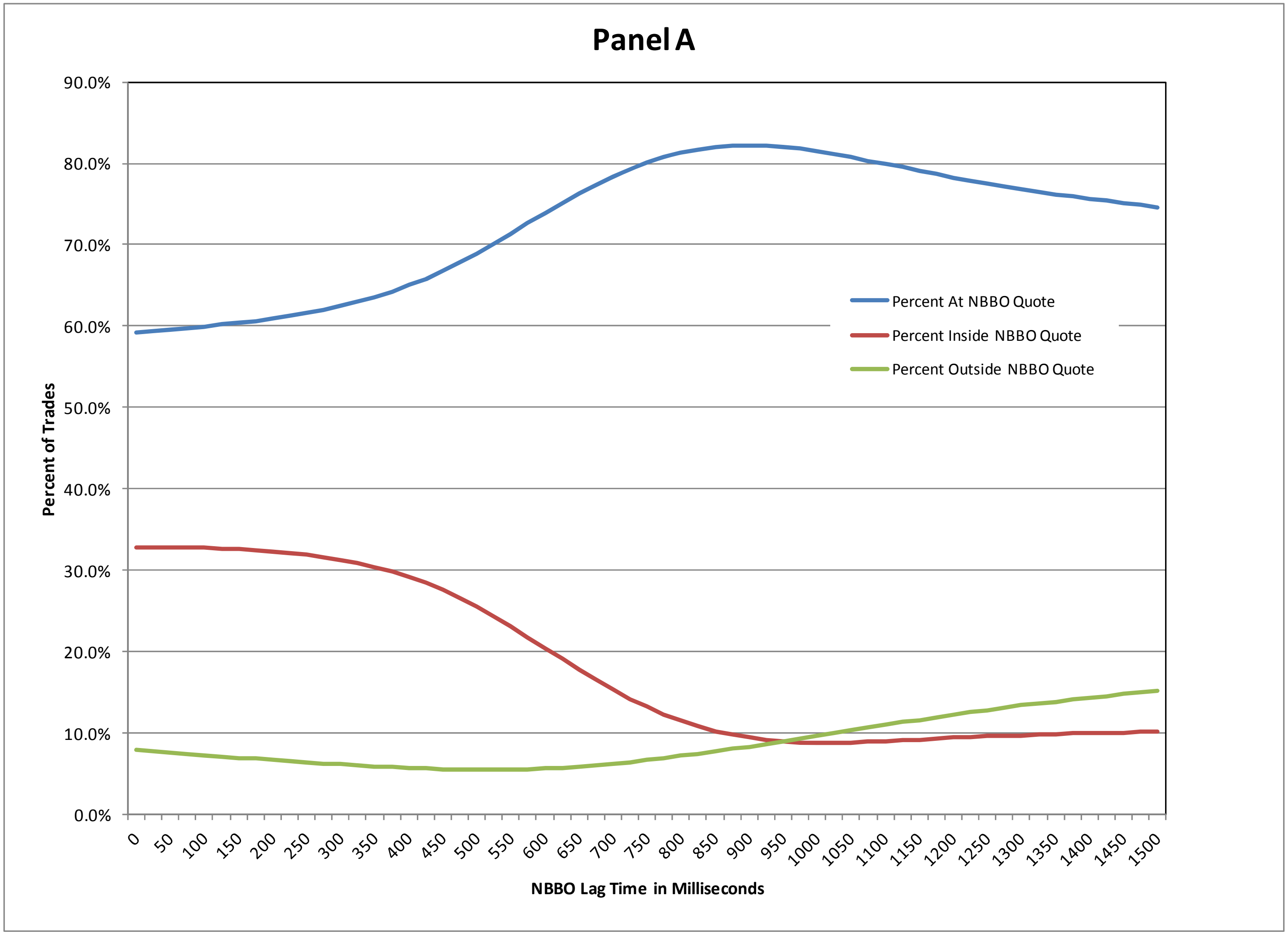
Ready, M., 2010. Determinants of volume in dark pools. Working paper, University of Wisconsin.

Subrahmanyam, A., 1991. Risk aversion, market liquidity, and price efficiency. Review of Financial Studies 4, 417 – 442.

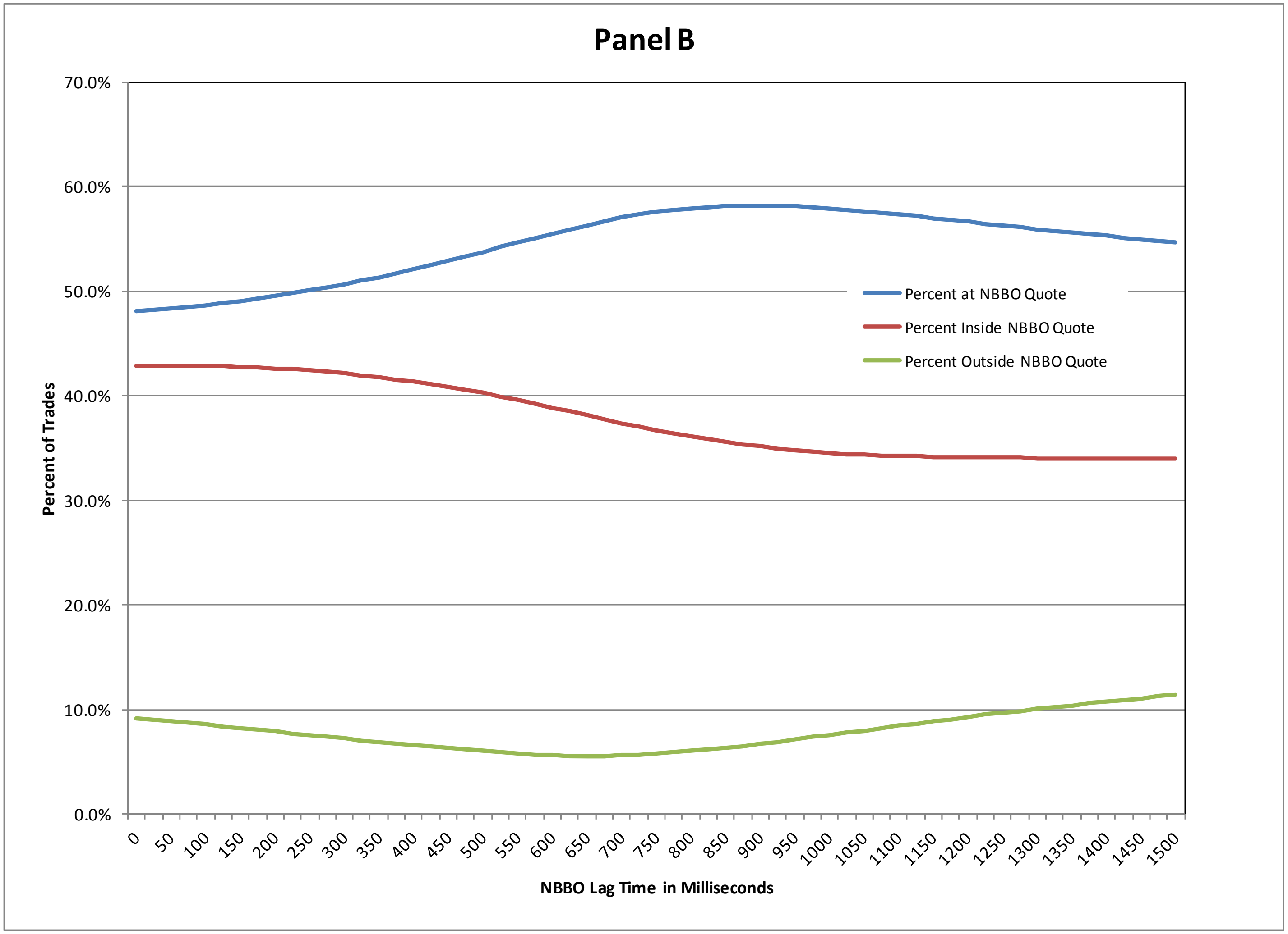
Weaver, D., 2011. Internalization and market quality in a fragmented market structure. Working paper, Rutgers University.

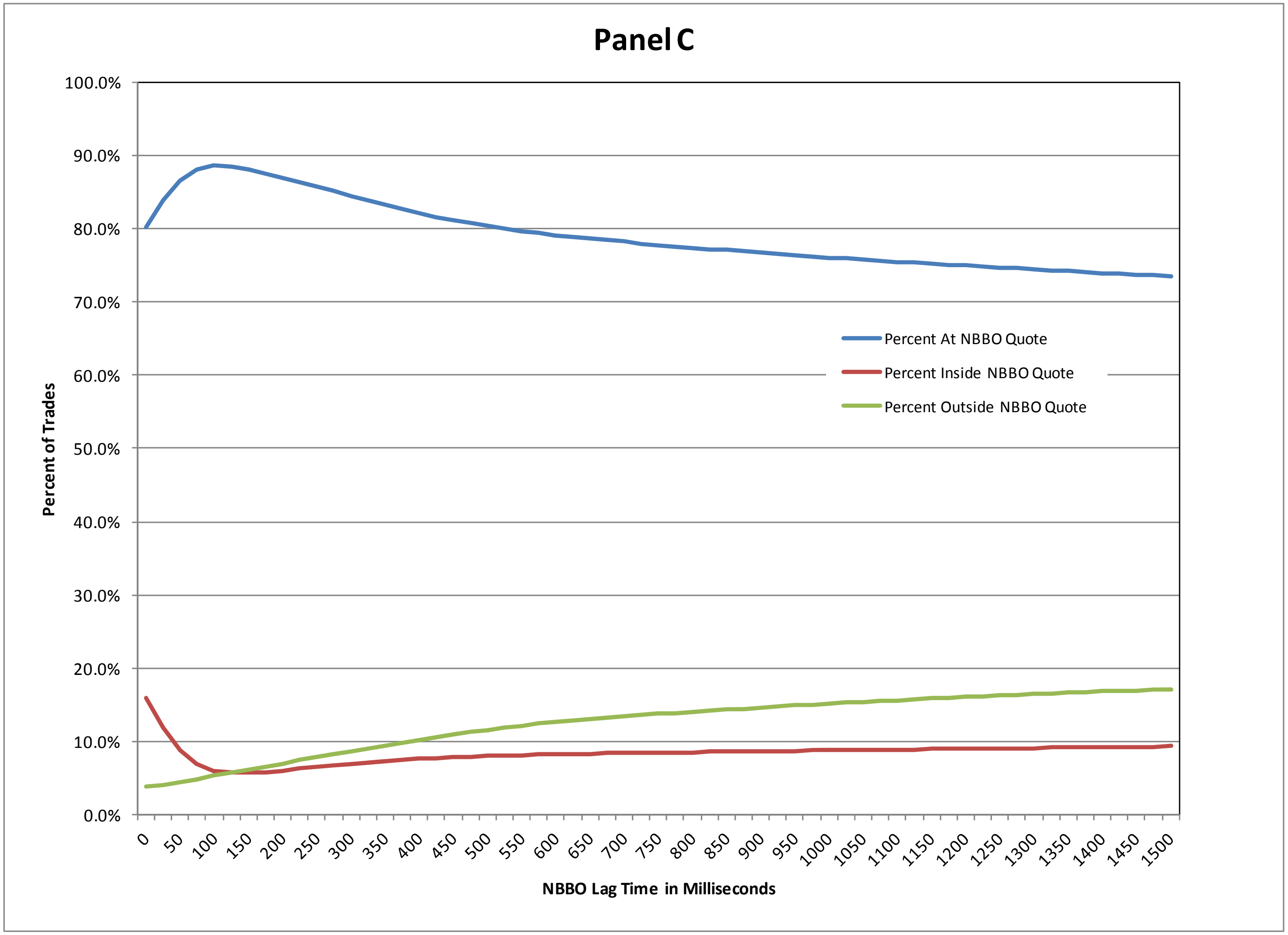
Ye, M., 2011. A glimpse into the dark: Price formation, transaction costs, and market share in the crossing network. Working paper, University of Illinois at Urbana-Champaign.

Zhu. H., 2011. Do dark pools harm price discovery? Working paper, Stanford University.



Panel A
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Percent of Trades
Percent At NBBO Quote
Percent Inside NBBO Quote
Percent Outside NBBO Quote
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
1050
1100
1150
1200
1250
1300
1350
1400
1450
1500
NBBO Lag Time in Milliseconds


Panel B
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Percent of Trades
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
1050
1100
1150
1200
1250
1300
1350
1400
1450
1500
NBBO Lag Time in Milliseconds
Percent at NBBO Quote
Percent Inside NBBO Quote
Percent Outside NBBO Quote


Panel C
Percent of Trades
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Percent At NBBO Quote
Percent Inside NBBO Quote
Percent Outside NBBO Quote
0 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250 1300 1350 1400 1450 1500
NBBO Lag Time in Milliseconds

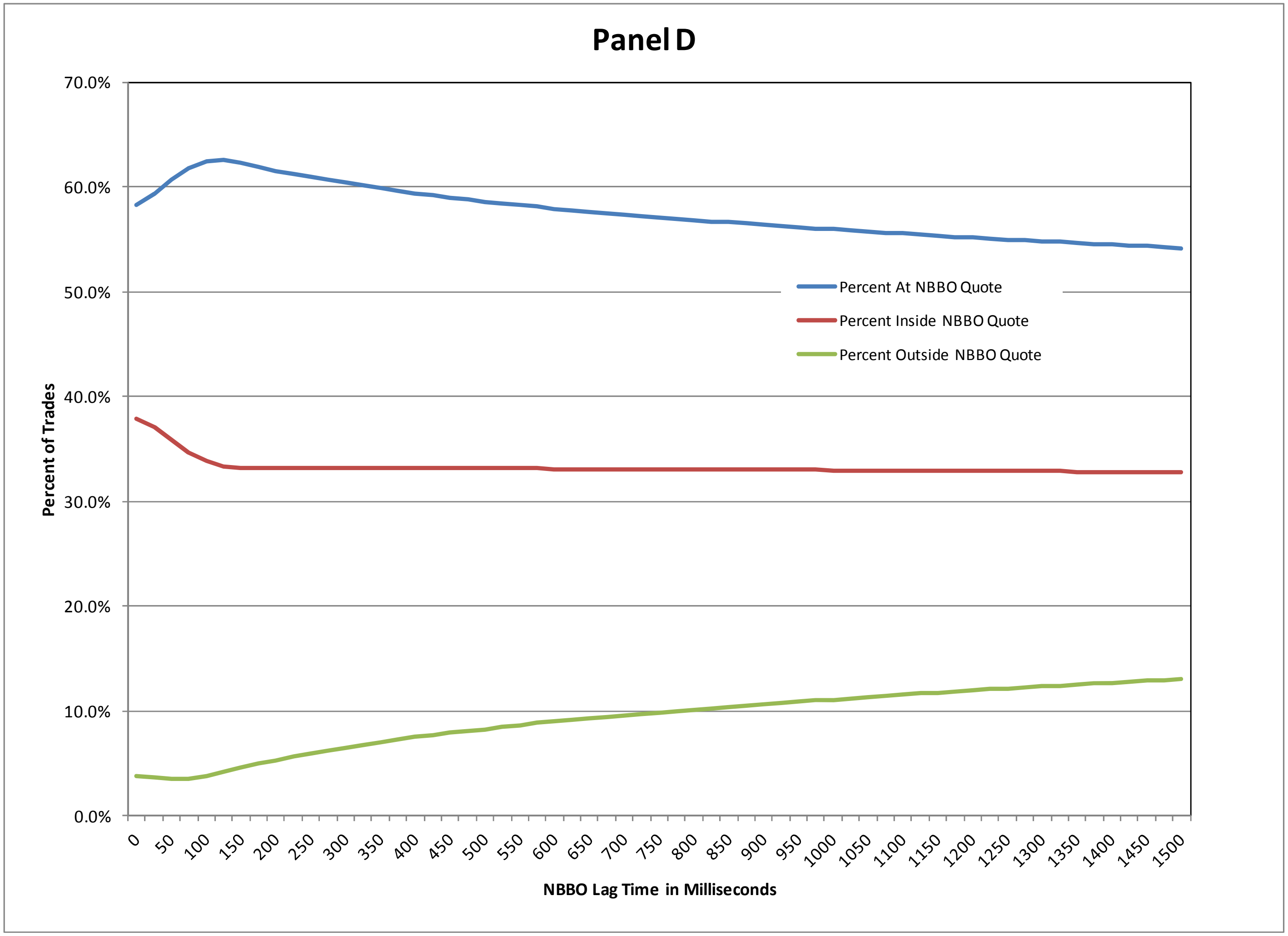
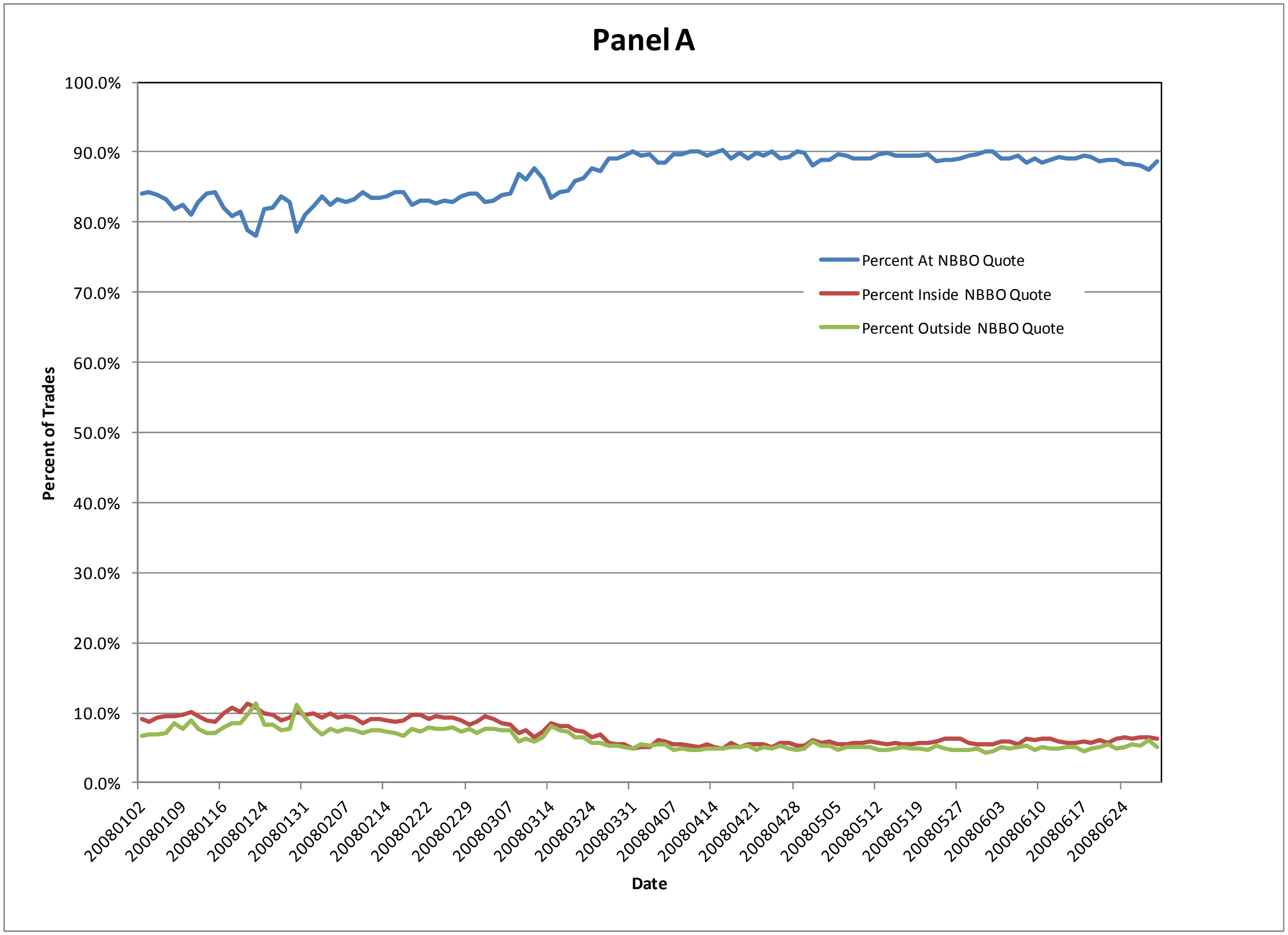


Figure 1: Trade price grid location as a function of NBBO quote lag time. For each stock for each day, we align trades and quotes using the time lag that maximizes the number of trades at the NBBO. We plot the percent of trades that occur at the NBBO quote, inside the NBBO quote, and outside the NBBO quote as a function of the quote lag time. Exchange trades and off-exchange trades are evaluated independently. On 10 March 2008, the NYSE significantly upgraded its computer systems. Panel A shows the alignment results for exchange executed trades prior to this upgrade while Panel B shows the alignment for off-exchange trades pre upgrade. Panels C and D show the alignment for the Post period for exchange and off-exchange trades, respectively


Panel A
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Percent of Trades
Percent At NBBO Quote
Percent Inside NBBO Quote
Percent Outside NBBO Quote
20080102
20080109
20080116
20080124
20080131
20080207
20080214
20080222
20080229
20080307
20080314
20080324
20080331
20080407
20080414
20080421
20080428
20080505
20080512
20080519
20080527
20080603
20080610
20080617
20080624
Date

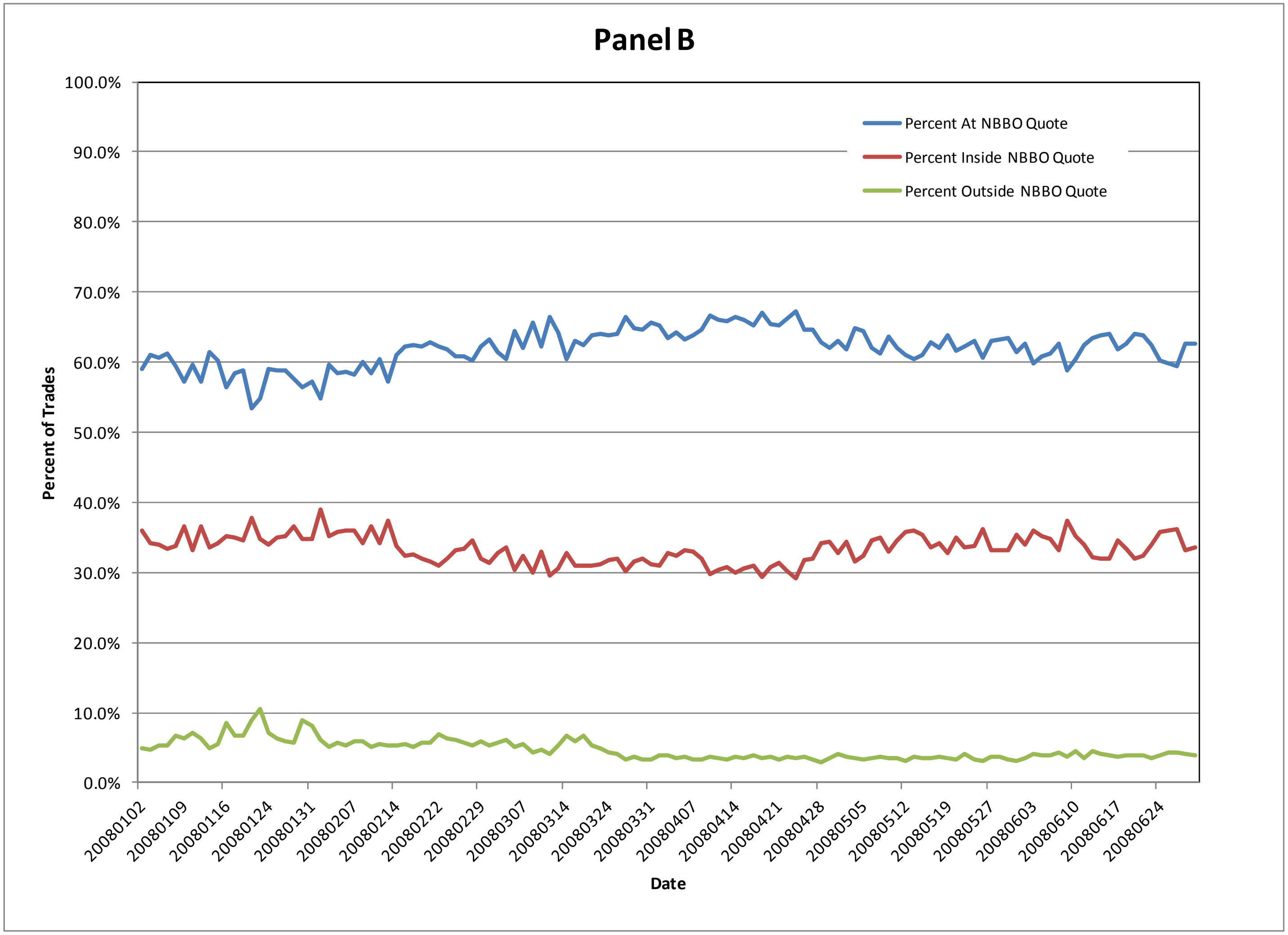


Figure 2: Trade/Quote Alignment for Exchange and Off-exchange trades. Panel A shows the time series of the trade location based on the optimum alignment lag time between the exchange calculated NBBO and exchange executed trades. Panel B shows the times series of off-exchange trades based on the optimal alignment. For each stock for each day, we calculate the number of trades executing at the NBBO for exchange trades and off-exchange trades. Each stock day has one alignment time for exchange and one for off-exchange trades. We report the percentage of trades that execute at the NBBO quote, inside the NBBO quote, and outside of the NBBO quote.

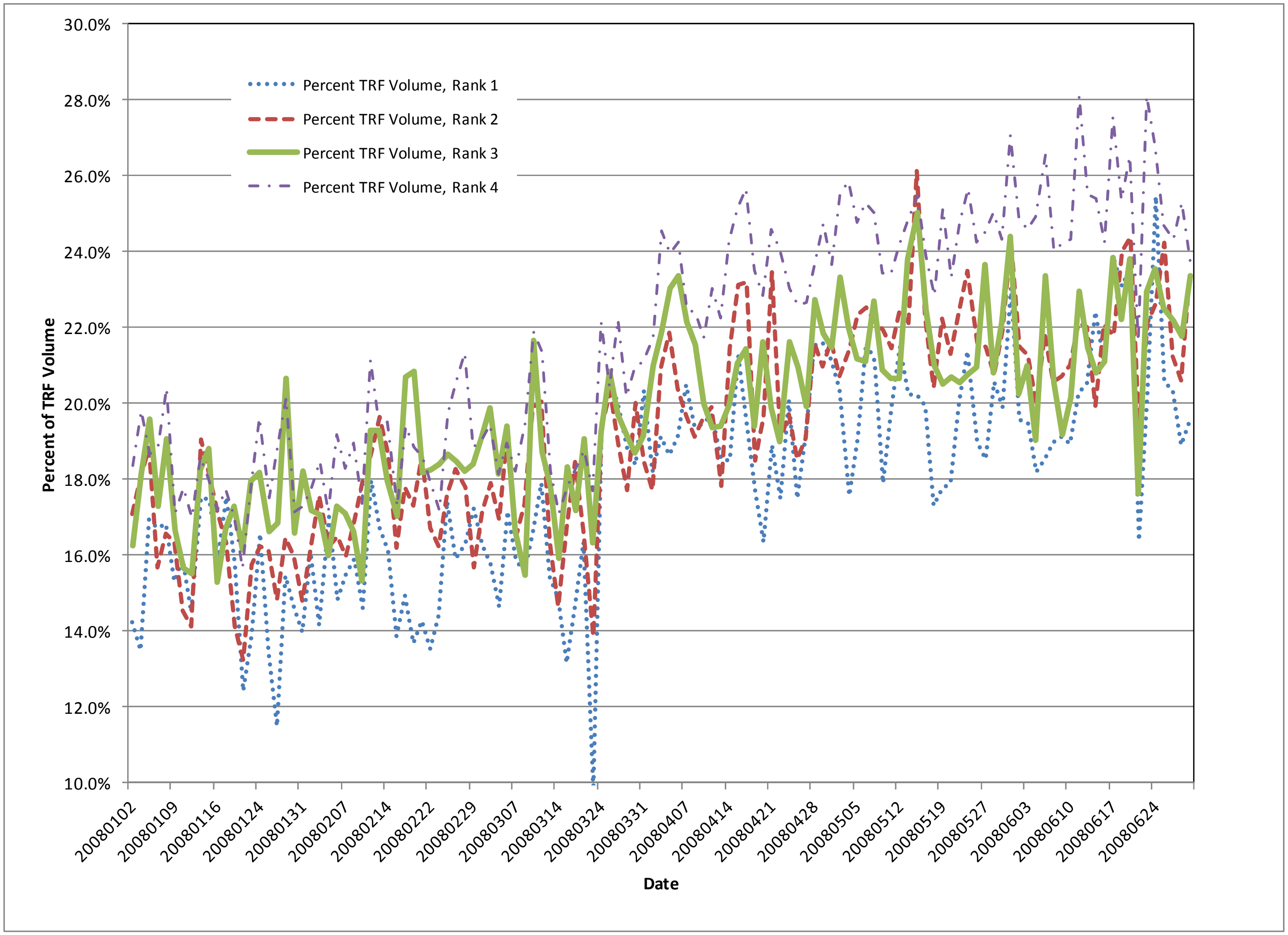


Figure 3: Percent of Off-exchange Volume. We present the percentage of off-exchange volume for each firm size group considered in our analysis. We group all NYSE common stocks into four quintiles. We then randomly select 50 stocks from each of the remaining quintiles. Our sample period is from January 2, 2008 through June 30, 2008. Rank 1(4) represents the lowest (highest) market capitalization group in the analysis.

Table 1
Sample Descriptive Statistics
We report the mean and standard deviation (STD) of Volume (thousands of shares per day) and number of shares per trade for exchange and off-exchange trades. We report statistics for the full sample and by quartiles of firm size.

	Exchange				Off-exchange			
	Volume (1,000s)		Trade Size		Volume (1,000s)		Trade Size	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Full Sample	2,449	7,177	174.8	67.0	800	2,871	272.1	318.4
Firm Size (Quartile)								
Small	396	628	161.2	39.4	109	290	297.4	404.1
2	865	945	167.1	42.2	239	361	260.1	216.8
3	1,305	1,357	172.4	51.9	371	519	269.1	403.2
Large	7,231	13,118	198.4	105.4	2,483	5,357	261.9	178.0

Table 2

Trade Price Grid Location for Exchange and Off-Exchange Trades

We report the percentage of trades at prices relative to the NBBO for exchange and off-exchange trades. For each case the trade price (TP) relative to the NBBO is: Over Ask, TP > NBBO ask; At Ask, TP = NBBO Ask; Inside Quote, NBBO Ask > TP > NBBO Bid; At Midpoint, TP = ((NBBO Ask + NBBO Bid)/2); At Bid, TP = NBBO Bid, Under Bid, TP < NBBO Bid. The five categories, excluding At Midpoint, are exhaustive and mutually exclusive. All trades, ISO trades and NISO trades are reported in Panels A, B, and C, respectively. We report the difference in the exchange and off-exchange means and the significance level for a paired t test.

	Trade Location on Price Grid					
	Over Ask	At Ask	Inside Quote	At Midpoint	At Bid	Under Bid
Panel A: All Trades						
Exchange	3.32%	43.87%	7.22%	2.96%	42.50%	3.09%
Off-Exchange	2.42%	32.39%	31.78%	14.77%	31.07%	2.34%
Difference	0.91**	11.47**	-24.56**	-11.81**	11.42**	0.75**
Panel B: ISO Trades						
Exchange	5.08%	42.64%	6.90%	2.65%	40.62%	4.75%
Off-Exchange	4.24%	41.32%	11.24%	4.64%	39.26%	3.94%
Difference	0.84**	1.32**	-4.34**	-2.00**	1.37**	0.81**
Panel C: NISO Trades						
Exchange	1.97%	44.81%	7.52%	3.23%	43.89%	1.81%
Off-Exchange	1.86%	30.23%	36.98%	17.25%	29.09%	1.85%
Difference	0.11**	14.59**	-29.46**	-14.02**	14.80**	-0.04**

* significant at the 5% level

** significant at the 1% level

Table 3

Information Share of Exchange and Off-exchange Trades

For each firm for exchange and off-exchange trades, we report the Volume Share, Information Share, and the Information Ratio, which is the ratio of the first two variables. We report the mean of each of these variables for the full sample and by quartile of firm size. The Information Share is for the two channels–exchange and off-exchange trades. An Information Ratio greater than 1.0 indicates that the trade channel carries more information than would be expected based on the volume would indicate. We conduct a paired difference t-test to test the null hypothesis of equality of means for the Information Ratio and report the results as Info Ratio Diff.

	Full		Rank 1 (Small Firms)		Rank 2		Rank 3		Rank 4 (Large Firms)	
	Exch	Off	Exch	Off	Exch	Off	Exch	Off	Exch	Off
Volume Share	0.803	0.197	0.822	0.178	0.807	0.193	0.802	0.198	0.783	0.217
Information Share	0.902	0.098	0.923	0.077	0.908	0.092	0.903	0.097	0.876	0.124
Information Ratio	1.125	0.495	1.126	0.441	1.126	0.481	1.127	0.495	1.120	0.565
Info Ratio Diff	0.629**		0.685**		0.645**		0.632**		0.555**	

* significant at the 5% level

** significant at the 1% level

Table 4

Information Share of Exchange and Off-exchange Trades, by Trade Type

We evaluate the information content of exchange and off-exchange trades, segmented by whether the trade type was ISO or NISO. The analysis is based on four price channels–Exchange, ISO, Exchange; NISO; Off-exchange, ISO, Off-exchange, NISO. We report the mean of volume share, information share, and the information ratio. Information ratio is the ratio of information share to volume share. An information share greater than 1.0 indicates that the trade channel carries more information than would be expected based on its volume. We conduct a paired difference t-test of the information ratio for the price channel and trade type and report the results in the column labeled Diff.

	ISO			NISO		
	Exch	OFF	Diff	Exch	OFF	Diff
Full Sample						
Volume Share	0.227	0.032		0.576	0.165	
Information Share	0.359	0.032		0.569	0.050	
Information Ratio	1.594	1.178	0.416**	0.990	0.306	0.684**
Rank 1 Small Firm						
Volume Share	0.205	0.022		0.616	0.157	
Information Share	0.328	0.036		0.589	0.049	
Information Ratio	1.610	1.948	-0.338	0.959	0.326	0.633**
Rank 2						
Volume Share	0.225	0.028		0.582	0.165	
Information Share	0.339	0.029		0.591	0.047	
Information Ratio	1.524	1.131	0.393**	1.022	0.283	0.739**
Rank 3						
Volume Share	0.223	0.036		0.579	0.162	
	0.359	0.029				
Information Share				0.574	0.045	
Information Ratio	1.622	0.870	0.753**	0.993	0.281	0.712**
Rank 4 Large Firm						
Volume Share	0.254	0.044		0.529	0.174	
Information Share	0.411	0.032		0.522	0.060	
Information Ratio	1.619	0.763	0.856**	0.988	0.335	0.653**

* significant at the 5% level

** significant at the 1% level

Table 5

Spread Analysis of Exchange and Off-exchange Trades

We calculate the trade weighted effective and realized half spreads for each stock day. Realized half spreads are calculated using prices 5 minutes and 30 minutes, in turn, after the trade. PM is the ration of realized spreads to effective spreads. Exch represents exchange and OFF represents off-exchange trades. Results are presented for All Trades, ISO trades, and NISO trades. We test the null hypothesis that the means for the exchange and off-exchange values are the same using a paired t-test.

	Effective Spread			Realized Spread (5 Min)			Realized Spread (30 Min)			Preference Measure (5 Min)		
	Exch	OFF	Diff	Exch	OFF	Diff	Exch	OFF	Diff	Exch	OFF	Diff
All Trades	1.35	1.23	0.13**	-0.06	0.35	-0.41**	-0.01	0.37	-0.38**	-0.07	0.26	-0.33**
ISO Trades	1.50	1.48	0.02**	-0.08	-0.21	0.13**	-0.02	-0.20	0.17**	-0.09	-0.80	0.71
NSO Trades	1.24	1.16	0.07**	-0.05	0.49	-0.54**	0.00	0.52	-0.52**	-0.04	0.40	-0.44**

* significant at the 5% level

** significant at the 1% level

Table 6
MRR Analysis
We estimate the implied spreads of ISO and NISO trades using the method of Madhavan, Richardson, and Roomans (1997) who propose estimating the following regression:

$$p_t - p_{t-1} = (\phi + \theta)x_t - (\phi + \rho\theta)x_{t-1} + \varepsilon_t + \xi_t - \xi_{t-1}.$$

subject to the following moment constraints:

$$E\left(x_t x_{t-1} - x_t^2 \rho, |x_t| - (1-\lambda), u_t - \alpha, (u_t - \alpha)x_t, (u_t - \alpha)x_{t-1} \right) = 0$$

where p_t is the trade price, θ is the asymmetric information parameter, ϕ is the cost of supplying liquidity, λ is the probability a trade occurs inside the quote, ρ is the autocorrelation of order flow, $u_t = p_t - p_{t-1} - (\phi + \theta)x_t + (\phi + \rho\theta)x_{t-1}$, α is a constant (drift) parameter, and x_t is a trade direction indicator. In particular, x_t is 1 if the trade is buyer initiated (at or above the NBBO ask), -1 if the trade is seller initiated (at or below the NBBO bid), and 0 if the trade is inside the NBBO quote. The implied spread, S, can this be consistently estimated as $S = 2(\phi + \theta)$, the effective spread, S^E, can be estimated as $S^E = (1-\lambda)(2\phi + \theta)$, and the fraction of implied spread attributed to asymmetric information, r, can be estimated as $r = \theta / (\phi + \theta)$. We estimate the equation separately for exchange and off-exchange trades. To calculate the price change, p_t - p_{t-1}, p_t is always the trade price from the trade type that we are estimating, but p_{t-1} is simply the last trade price and can be either exchange or off-exchange trade prices. We estimate the model for each 25 trading days in the sample for each stock. The trading day is divided into 6 equal sections and an estimate is conducted for each section. Tests are based on paired differences. Spread results are in cents.

Estimate	Exchange Trades	Off-exchange Trades	Difference
S	1.96	0.49	1.47**
S^E	1.19	0.20	0.99**
r	0.61	0.68	-0.07
ρ	0.53	0.36	0.16**

* significantly different at the 5% level
** significantly different at the 1% level

Table 7

MRR Regression Results by Trade Type and Venue

We estimate the implied spreads of ISO and NISO trades using the method of Madhavan, Richardson, and Roomans (1997) who propose estimating the following regression:

$$p_t - p_{t-1} = (\phi + \theta)x_t - (\phi + \rho\theta)x_{t-1} + \varepsilon_t + \xi_t - \xi_{t-1}.$$

subject to the following moment constraints:

$$E\left(x_t x_{t-1} - x_t^2\rho, |x_t| - (1-\lambda), u_t - \alpha, (u_t - \alpha)x_t, (u_t - \alpha)x_{t-1}\right) = 0$$

where p_t is the trade price, θ is the asymmetric information parameter, ϕ is the cost of supplying liquidity, λ is the probability a trade occurs inside the quote, ρ is the autocorrelation of order flow, $u_t = p_t - p_{t-1} - (\phi + \theta)x_t + (\phi + \rho\theta)x_{t-1}$, α is a constant (drift) parameter, and x_t is a trade direction indicator. In particular, x_t is 1 if the trade is buyer initiated (at or above the NBBO ask), -1 if the trade is seller initiated (at or below the NBBO bid), and 0 if the trade is inside the NBBO quote. The implied spread, S, can this be consistently estimated as $S = 2(\phi + \theta)$, the effective spread, S^E, can be estimated as $S^E = (1-\lambda)(2\phi + \theta)$, and the fraction of implied spread attributed to asymmetric information, r, can be estimated as $r = \theta / (\phi + \theta)$. We estimate the equation separately for ISO and NISO trades for exchange and off-exchange trades. To calculate the price change, p_t - p_{t-1}, p_t is always the trade price from the trade type that we are estimating, but p_{t-1} is simply the last trade price and can be either ISO or NISO trades. We estimate the model for each 25 trading days in the sample for each stock. The trading day is divided into 6 equal sections and an estimate is conducted for each section. Tests are based on paired differences. Spread results are in cents.

	ISO Trades			NISO Trades		
Estimate	Exchange	Off-exchange	Diff	Exchange	Off-exchange	Diff
S	1.64	1.36	0.28**	1.68	1.87	-0.19**
S^E	0.62	0.47	0.15**	0.94	0.52	0.42**
r	1.16	1.35	-0.19**	0.73	0.93	-0.20**
ρ	0.56	0.39	0.17**	0.42	0.33	0.09**

* significantly different at the 5% level

** significantly different at the 1% level

Table 8

Determinates of Off-exchange Trading

We investigate how structural changes to the market and market quality on the exchange impact the choice of order flows routed to off-exchange venues. The following equation is estimated:

$$\%TRFvol_{i,t} = \alpha + \beta_1 Spd + \beta_2 Isig_{i,t} + \beta_3 MpVar_{i,t} + \beta_4 Liq_{i,t} + \beta_5 Turn_{i,t} + \varepsilon_{i,t}$$

where %*TRFvol* is the percentage of off-exchange volume for stock *i* on day *t*. *Spd* is a dummy variable that is 1 after March 10, 2008 and zero otherwise. *Isig* is the absolute value of residual from a Fama and French 3 factor regression and proxies idiosyncratic risk, *MpVar* is the NBBO quote midpoint volatility for the day, *Liq* is the time weighted daily average of the NBBO spread, in cents, divided by the total quoted depth at NBBO prices, in round lots, and *Turn* is the turnover of the stock defined as total traded volume divided by shares outstanding. The coefficient is multiplied by 100 for reporting. We estimate three regression specifications as fixed effects. As a robustness check, we estimate the regression at the stock level and report the average coefficient. We test if the average coefficient is statistically different from zero. Regression standard errors are adjusted for heteroscedasticity.

	Fixed Effects			
Variable	S1	S2	S3	Stk Level
Intercept	0.177**	0.205**	0.179**	0.159**
Spd	0.039**		0.038**	0.036**
Isig		-0.001	-0.001	-0.005**
MpVar		-0.032**	-0.024**	-0.091**
Liq		-0.003	-0.001	-0.230
Turnx100		0.132**	0.129**	0.576**
N	24,957	24,957	24,957	
Adj R2	0.235	0.219	0.259	

* significant at the 5% level

** significant at the 1% level

Table 9

Information Share Regression Results

We investigate how the information flow at the exchange is impacted by market conditions and the level of volume executed at off-exchange venues. We estimate the following equation:

$$X_{i,t} = \alpha + \beta_1 \%TRFvol_{i,t} + \beta_2 Spd + \beta_3 Isig_{i,t} + \beta_4 MpVar_{i,t} + \beta_5 Liq_{i,t} + \beta_6 Turn_{i,t} + \varepsilon_{i,t}$$

where *X* represents the Information Ratio or the Information Share for stock *i* on day *t*, in turn. *%TRFvol* is the percentage of total volume for the stock executed off-exchange. *Spd* is a dummy variable to control for the system upgrade of the NYSE on 10 March 2008, which is 0 prior to this date and one after. *Isig* is the absolute value of the residual from a Fama and French 3 factor regression and proxies the idiosyncratic risk,, *MpVar* is the NBBO quote midpoint volatility for the day, *Liq* is the time weighted market liquidity proxy of NBBO spread divided by total NBBO depth, and *Turn* is the turnover of the stock defined as total traded volume divided by shares outstanding. The coefficient is multiplied by 100 for reporting. We estimate the equation as a fixed effects regression and at the stock level. For the stock level, we report the average coefficient and test if it is statistically different from zero.

	Information Ratio		Information Share	
Variable	Fixed	Stk Lev	Fixed	Stk Lev
Intercept	0.791**	0.839**	0.929**	0.939**
%TRFvol	1.752**	1.485**	-0.139**	-0.177**
Spd	-0.050**	-0.032**	-0.025**	-0.018**
Isig	0.001	-0.004**	0.001*	-0.001
MpVar	0.052**	0.143**	0.038**	0.129**
Liq	-0.002	0.550	-0.003**	0.396
Turnx100	-0.072**	0.002	-0.056**	-0.108**
N	24,957		24,957	
Adj R2	0.565		0.260	

* significant at the 5% level

** significant at the 1% level

Table 10
Volume Imbalance
This table reports the cross sectional average coefficients of contemporaneous and lagged values of volume imbalance for exchange and off-exchange order flow. We estimate the following regression based on Chordia and Subrahmanyam (2004):

$$R_{i,t} - R_{m,t} = \alpha + \sum_{k=0}^{4} \beta_k ExOB_{i,t-k} + \sum_{k=0}^{4} \beta_{k+5} TrfOB_{i,t-k} + \varepsilon_{i,t}$$

where $R_{i,t}$ is the return for stock i on day t, $R_{m,t}$ is the market return on day t, *ExOB* is the volume imbalance for exchange order flow, and *TrfOB* is the volume imbalance for the order flow reported through off-exchange venues. We estimate the regression for each stock in the sample and then report the average coefficient. We test whether the average coefficient is statistically different from zero.

		Firm Size			
	Full Sample	1	2	3	4
$ExOB_{t}$	7.11**	6.78**	6.67**	5.69**	9.30**
$ExOB_{t-1}$	-1.32**	-1.53**	-1.17*	-0.81*	-1.79**
$ExOB_{t-2}$	-0.45*	-0.49	-0.55	-0.59*	-0.18
$ExOB_{t-3}$	-0.10	-0.48	-0.44	0.36	0.16
$ExOB_{t-4}$	-0.19	-0.01	-0.15	-0.62	0.04
$TrfOB_{t}$	0.04	0.49*	0.04	0.11	-0.46
$TrfOB_{t-1}$	-0.30*	-0.59**	-0.28	0.02	-0.35
$TrfOB_{t-2}$	-0.18	0.19	-0.41*	-0.15	-0.35
$TrfOB_{t-3}$	0.03	-0.21	0.19	0.15	-0.02
$TrfOB_{t-4}$	0.25	-0.10	0.34	0.42*	0.33

* significant at the 5% level
** significant at the 1% level