Market Fragmentation and Information Quality: The Role of TRF Trades

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 March 2012

JEL classification: G14, G18, G19

Key Words: Fragmentation, Information quality, TRF trades
*Contact Author

Market Fragmentation and Information Quality: The Role of TRF Trades

Abstract

We analyze and compare the information quality of order flows on the exchange and on off-exchange venues reported to Trade Reporting Facilities. Compared to exchange order flow, we find that off-exchange order flow has significantly lower information quality, including a lower information ratio and a lower price impact, and a significantly higher percentage of trades executing inside the quote. Our results are consistent with the notion that the segmentation of uninformed liquidity traders in off-exchange venues so that there is a higher proportion of informed traders on the exchanges, leads to an improvement in price discovery and information 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

Exchanges report their trades directly to the Consolidated Trade System, but various types of dark pools, including crossing networks, internalizers, and ping destinations, report their traders indirectly to the Consolidated Trade System through Trade Reporting Facilities (TRFs). Using 2008 data and considering all trades and volume (exchange and TRF trades together), O'Hara and Ye (2011) report that a higher proportion of TRF volume to total volume is associated with lower effective spreads.[1] Weaver (2011), using data from October, 2010, finds that exchange quotes are wider when there is more TRF volume. We proceed in the spirit of Subrahmanyam (1991) and Holden and Subrahmanyam (1992) who posit that differences in market quality can be driven by market segmentation. If a proportion of uninformed are able to segment, the level of informed trading on exchanges becomes higher, leading to an improved price discovery on 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 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 as the ratio of a trading venue's information shares to its volume share. An Information Ratio greater than 1.0 indicates that the venue is information dominant. 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.

[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 also evaluate the differences in transaction cost and price impact between exchange and off-exchange trades.[2] 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, making trade size less important. 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 and the price impact of off-exchange trades are significantly lower than those of exchange trades, again leading to the conclusion that off-exchange order flow is dominated by uninformed trading. We confirm this finding using the spread decomposition model of Madhavan, Richarson, and Roomans (1997). In addition, the serial correlation of trades is higher for exchange compared to off exchange order flow. This finding is consistent with the pooling model of Back and Baruch (2007) in which informed traders split large orders into small trades to pool with small uninformed traders, resulting in a high number of small trades on the same side of the market.

Turning to the determinants of TRF order flows, we find that as markets become faster, traders shift volume to the off-exchange trading venues. We do not find support for the positive correlation between spread on the exchanges and TRF volume as many argue that potential cost saving is higher on off-exchange trades when spread is wider (Buti, Rindi and Werner, 2011; Ray, 2011). Also, we find that when intraday volatility is high, order flows are executed 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 off-exchange

[2] TRF order flow is heterogeneous in that multiple trading venues are reported though the TRF facility. We consider the TRF volume as a unit and segment on three characteristics of TRF order flow to help address this issue.

volume increases with trading intensity as 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. Other factors that impact the information quality of exchange trades are market speed, intraday volatility, market liquidity, and trading intensity.

TRF trades are heterogeneous and may have different information content. Although we cannot identify the trades directly, we are able to isolate certain segments of the TRF trades based on the characteristics of the trades. As our conjecture hinges on the routing of uninformed order flows to TRFs, we use various rules to separate out the order flows that are known to contribute little to price discovery (i.e., dark pools, internalized orders) and show even stronger support for our main hypotheses.

Lastly, 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 order imbalance on exchanges is significantly negatively correlated with the current return, but this finding does not hold for off exchange trades.

Overall, our results indicate that the price discovery on exchanges improves in fragmented markets because uninformed traders are able to route 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 face barriers and seldom trade.

.

2. Hypotheses Development

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

Informed traders cannot freely use off-exchange venues to carry out their trades, leaving uninformed traders the major players on these venues. 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 Scottrade. 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 is likely different for trades executed on and off-exchanges. Based on these considerations, we test the following hypothesis:

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

If a significant portion of uninformed trades is executed on 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 choose to route 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 trades of uninformed traders on off-exchange venues are 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. We test the following hypothesis:

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

Hypothesis 2 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 based on the Easley and O'Hara model is that uninformed traders have a lower transaction cost if they can creditably signal that they are uninformed. By routing trades 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 context, their results are entirely consistent with our predictions.

Hypothesis 3: *Off-exchange trading volume increases with potential cost saving and decreases with execution risk.*

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 potential cost saving is high and execution risk is low.[3]

Admati and Pfleiderer (1988) and Foster and Viswanathan (1990) develop models showing that discretionary liquidity traders can move their trading temporally 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. An additional benefit is the improved price discovery on the exchanges when uninformed traders route orders to off-exchange venues. Hence, we test the following hypothesis:

Hypothesis 4: *Uniformed traders have a higher propensity to use 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, discretionary liquidity traders disproportionately use off-exchange venues, removing price pressure from the market. The leads to the prediction 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:

[3] 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(a): *Exchange order imbalance predicts future stock returns.*

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

On 10 March 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, increasing the ability of low latency traders to profit from trading with high latency traders.[4] Hence, the incentive for high latency traders to use off-exchange venues increased.

In addition, Regulation NMS introduced a new order type – Intermarket Sweep Order (ISO). 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 (in thousand shares) and trade size for the full sample and by firm size. About 25% of our sample volume is executed on off-exchange venues.[5] This is comparable to the 27% off-exchange volume found in O'Hara and Ye (2011),

[4] 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.

[5] 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.

and similar to the results reported in Weaver (2011) using more recent data.[6] Consistent with the findings of O'Hara and Ye (2011) for NYSE stocks, the largest size quartile's off-exchange volume is 26% compared to 22% for the remaining quartiles. The higher off-exchange fragmentation for large stocks on the 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 off-exchange is that these trades have smaller price impact 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. 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

We use the technique of Lee and Ready (1991) to infer trade direction, but the quality of the inference depends on whether trades and quotes are properly aligned. Trades and quotes are handled through separate computer systems and are provided in separate files in the DTAQ dataset. Various ways of dealing with the differing latencies of the two computers have been proposed, including arbitrarily adding 1 second, 2 seconds, and so forth to the quote times. Today computers execute most trades without human intervention, matching incoming orders with posted limit orders. Hence, following McInish and Upson (2011), we believe that the best match lag is the one that maximizes the number of trades executed

[6] Two market centers at the time of our study, BATS and DirectEdge, reported through the TRF facility, but became exchanges by the time of the Weaver (2011) study. Weaver's study reports a combined volume for these two venues and off-exchange venues of 50.83%.

at the NBBO. We identify this lag as follows. For each stock day in the sample, we test time lags from 0 to 1,500 milliseconds, in 25 millisecond increments and select the lag time that maximizes the number of trades executed at the NBBO.

We apply this approach to exchange and off-exchange trades separately. 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.

Because of the increasing use of the DTAQ dataset, we present a more detailed discussion of the impact of the trade quote alignment procedure in the Appendix.

4.0 Results

4.1 Off-exchange time series

Exchange latency might impact traders order routing strategies. Figure 1 shows a time series plot of the daily average percent of volume reported through the TRFs 50 randomly selected firms for each quartile based on market capitalization (rank 1 = smallest). Prior to the NYSE speed increase, off-exchange volume is stable at roughly 16%. Figure 1 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 stabilizes at about 23%.

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 impacts execution prices on a particular 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. Our method detailed in section 3.3 and the appendix 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 relative to exchange ISO trades. However, the highest percentage of inside NBBO trades is for off-exchange NISO trades. For example, 6.90% of ISO trades are inside the NBBO on exchanges 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.[7] 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.

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 channels–exchange and off-exchange. We use the last trade price of each trade type in each second.[8] Second, we estimate information shares for four channels–exchange ISO, off-exchange ISO, exchange NISO, and off-exchange NISO. Unless the resulting variance co-variance matrix is diagonal, the

[7] We show below that off-exchange ISO trades have high information quality and represent the second highest information quality of the market order flows.
[8] Although the DTAQ database has time stamps to the millisecond, using all the trades is computationally prohibitive. Our use of trade prices follows Hasbrouck (2003), Anand and Chakravarty (2007), and Goldstein, Shkilko, Van Ness, and Van Ness (2008).

information share estimate for each trade type is not exactly identified. Therefore, we average the upper and lower bound values to obtain a point estimate of the information share.

Hasbrouck (1995) finds that the regional stock exchanges are not information dominant because their information share is lower than their share of traded volume. We define the information ratio (*InfoRatio*) as:

$$InfoRatio_i = \frac{InfoShare_i}{VolumeShare_i} \quad (1)$$

$InfoShare_i$ 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 for a channel over the period, relative to total executed volume. If $InfoShare_i > 1.0$, the trading venue is information dominant and has higher information quality.

Table 3 reports results for the full sample and by firm size. We report results obtained 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. We report results for the full sample and by quartiles of firm size. For exchange volume share is 80.3% and information share is 0.902, giving an information ratio of exchange volume is 1.125. The off-exchange information ratio is 0.495, indicating lower information quality. Our results by quartile are similar.

Table 4 reports even stronger results for our analysis of our four ISO and NISO channels. Exchange ISO volume is only 22.7% of total volume, but has an information share of 0.359, giving an average information ratio of 1.594. Exchange NISO volume has an information ratio of 0.990. 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 information ratio 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 is 1.178, indicating that off-exchange ISO volume is more informed, at least for some stocks.

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 are 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 use of off-exchange venues by the uninformed, 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, Price Impact, and Execution Quality

Our second hypothesis states that adverse selection costs are lower for off-exchange trades than for exchange trades. By using off exchange venues, uninformed traders credibly signal that they are uninformed so that under these circumstances the model of Easley and O'Hara (1987) predicts that off-exchange trades have lower execution costs. Table 5 shows results for all trades and trades 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 price impact based on 5 and 30 minute reference points. Our measure of price impact is based on Bessembinder (2003). Many recent studies use a five minute window to compute price impact. Following the literature, we also use 5-minute along with a longer window of 30 minutes after the completion of a trade. The direction of price impact depends on the underlying information. Price impact can be positive (negative) when a trade is associated with a move in the price in the (opposite) direction of the trade. Considering all trades and the results conditioned on ISO/NISO trades reported in Table 5, Panel A, the price impact of exchange trades is significantly than larger off-exchange trades. This result supplies additional evidence that off-exchange volume has lower information quality than exchange volume.

Taking the analysis of effective spread and price impact together, the improvement in information quality promoted by uninformed order flows to TRF venues need not coincide with higher liquidity or welfare. As Zhu (2011) points out "highly informative orders correspond to better price discovery, but also tend to worsen the adverse selection on the exchange, resulting in wider spreads and higher price impacts."

In addition to effective spreads and price impact, 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, as reported in Table 5, information quality of off-exchange trades is significantly worse than for exchange trades. 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. Results in Panel A of Table 5 are aggregated by taking the average of all individual trades on the exchange and off-exchanges. To rule out the possibility that differences in transaction sizes of trades on the exchange versus off-exchanges may systematically affect our results, we also compute volume-weighted measures of effective spread, price impact, and PM and report them in Table 5, Panel B.

As large trades can have higher transaction costs and price impact, the volume weighted measures are all higher in Panel B than the simple averages reported in Panel A. However, exchange trades continue to show higher effective spreads, price impact, and PM. The only exception is for volume-weighted measures for ISO trades. Although ISO trades are only a very small portion (3.2%) of the overall distribution, their trade size is larger, resulting in significantly higher effective spread on off-exchange ISO than that of exchange ISOs.

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 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 MRR model is estimated for each day for each segment, with 25 days of data. We obtain estimates for 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 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 off-exchange 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. Examining ρ, the serial correlation of signed order flow, in Table 6 we see 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 has 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 results by trade type. We compare exchange against off-exchange trades by type of trade. 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 venue. 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, 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.

4.5 Determinates of Choice of Venue

4.5.1 Determinates of total TRF volume

Hypothesis 3 states that off-exchange volume increases with potential cost saving and decreases with execution risk. Cost saving is directly related to spread. Execution risk can be measured in many dimensions and with various metrics, and we use intraday volatility, speed of the market and degree of information asymmetry.[9] Off-exchange volume share will be lower when there is a higher degree of information asymmetry that can lead to higher execution risk. Similarly, execution risk will be higher in a slower and/or highly volatile market. In this section we evaluate the market conditions that lead to a change 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 10 March 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. Generally, the larger spread is an indication of higher potential savings on off-exchanges, which leads to an increase in TRF volume. Ray (2010) raises the concern that quoting exchange prices are more easily manipulated with a larger spread. Thus, a negative coefficient indicates that traders avoid using TRFs in fear of gaming. The relationship between TRF volume and spreads is more complex, depending on which effect is dominant. $Turn_{i,t}$ is the stock turnover defined as

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

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 significant change in the processing capabilities of a primary exchange like the NYSE has far reaching affects beyond the exchange itself. 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.

4.5.2 Determinates of sub-categories of TRF volume

We recognize that TRF order flow is heterogeneous, coming from venues that adopt different trading strategies for clients with different levels of information quality. We attempt to break down TRF order flow in the following manner to isolate, as best as possible given the data, the various sources of trading. Our first cut of the order flow is to remove ISO volume and only use NISO TRF order flow. As shown in this paper and other research, ISO order flow tends to be more informed than NISO order flow on the off-exchange venues, consists of only a small portion of the volume. The remaining NISO order flow is further partitioned in two additional ways. First, we include only NISO trades that are executed at the quote midpoint of the prevailing NBBO. This order flow proxies for crossing networks, which derive the trading price from the NBBO quote in lit markets. This corresponds most closely to the dark pools acting as agents, relying on lit venues to determine prices, and, hence, do not provide price discovery (Zhu, 2011; Ready, 2010). Second, we consider all NISO trades that occur inside the quote. This grouping includes internalized order flow with the crossing networks. Together, internalized order flow and crossing network trading are expected to be the least informed of the market.

Table 9 reports three additional regressions based on Equation 2, where the dependent variable is off-exchange NISO volume, off-exchange NISO executed inside the quote, and off-exchange NISO at the midpoint, respectively. We find results quite similar to those when total TRF volume (both ISO and NISO are included). The percentage of uninformed TRF volume is higher in faster market, and in stocks with higher turnover, but declines when volatility is higher. The coefficient of *Liq* remains insignificant.

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 use of 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 the use of off exchange venues by the uninformed improves the information quality of exchanges. This analysis seeks to 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) the direction is 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 differences in the volumes.

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 by 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 route orders 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. The coefficient of *Turn* is significant and negative in both regressions for the information share. Overall, our regression results support Hypothesis 4; an increase in off-exchange volume improves the information quality of exchange trades.

O'Hara and Ye (2011) caution that TRF order flows are not homogeneous and may not be appropriate to aggregate. This is important to note as our hypothesis hinges on the credible segmentation of order flows. Thus, as explained in the earlier section, we attempt to single out the off-exchange order flows that are most likely to be uninformed by using three additional classifications.

Table 11 reports results on the regression of information shares and ratios using three alternative measures of TRF order flows. Results are qualitatively similar to those reported in Table 10, with the exception of the negative coefficients for the variable *Liq*.

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 so that buys 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.[10] 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):

[10] 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).

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

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 12. 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 on the exchanges 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.

The large number of NISO TRF trades occurring at and inside the quote creates a potential trade inference issue with the Lee and Ready algorithm. Specifically, for trades at the quote midpoint executed off exchange, the tick test of the LR method is problematic. If the trade is truly from a crossing network, both parties initiate the trade and should be removed from the classification of trade directions. In addition, the price improvements offered by internalizers of retail order flow introduce a price shift that can lead to an incorrect trade inference. Therefore, we modify the order imbalance estimation for TRF volume in two ways. First, all midpoint TRF trades are dropped from the analysis. The remaining trades are signed using Lee and Ready for the order imbalance calculation. Second, all trades that are inside the quote are dropped from the analysis, with only trades at or outside of the prevailing quote signed for the imbalance calculation. As each level of TRF volume is removed from the imbalance calculation, we believe that the information quality and more importantly, the accuracy of signed trading will also improve.

Table 13 reports the estimation results of Equation 4 using the two more restrictive measures of order imbalances. The significance of off-exchange order flows does improve as expected. However, the main driving factors for explaining excess returns are still from the contemporaneous and the first-lag order imbalances on the exchanges.

5.0 Conclusion

O'Hara and Ye (2011) show that efficiency is improved for fragmented markets. 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 φ proportion of uninformed that are able to segment, the level of informed trading on exchanges becomes $\alpha/(\alpha+\gamma-\varphi)$,. Higher φ leads to improved price discovery on 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 a high concentration of uninformed traders in off-exchange venues is associated with a higher concentration 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 and lower price impacts than exchange trades. As a robustness test we estimate spreads using the regression model of Madhaven, Richardson, and Roomans (1997) and confirm that off-exchange trades have lower execution costs and lower adverse selection 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 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.

Appendix

Trade Alignment

On March 10, 2008, the NYSE implemented a significant upgrade to its computer system. Trades and quotes are routed through different computers with differing latencies. Hence, to order trades and quotes in the DTAQ dataset we subtract X milliseconds from the reported trades times. We explore how the NYSE computer upgrade affected X.

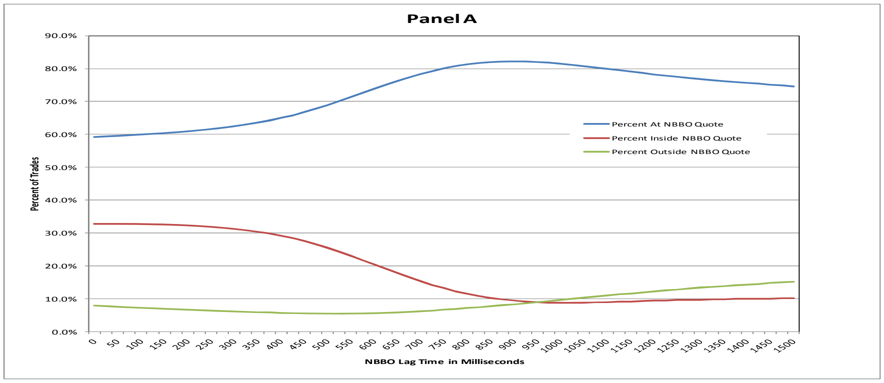
We begin with the clock times as given in the DTAQ trade and quote files so that X = 0. For each stock for each day before the speed change, we compute the percentage of trades executed at the NBBO, between the NBBO and outside the NBBO. We repeat the procedure for each 25 millisecond increament to X = 1,500 milliseconds. We repeat the procedure for all stocks and for the period after the speed change report the results in Figure A1.

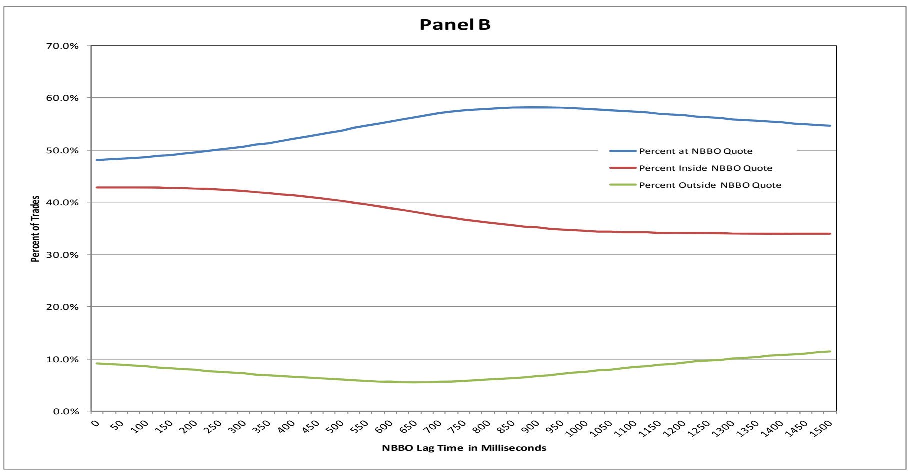
Figure A1, Panel A, shows that if there is a zero time lag applied to the reported NBBO quote time, 60% of trades would occur at the NBBO, 32% would be between the NBBO quotes, and the remainder would be outside the NBBO quotes. The number of trades at the NBBO is maximized at a lag of about 900 milliseconds, 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 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).

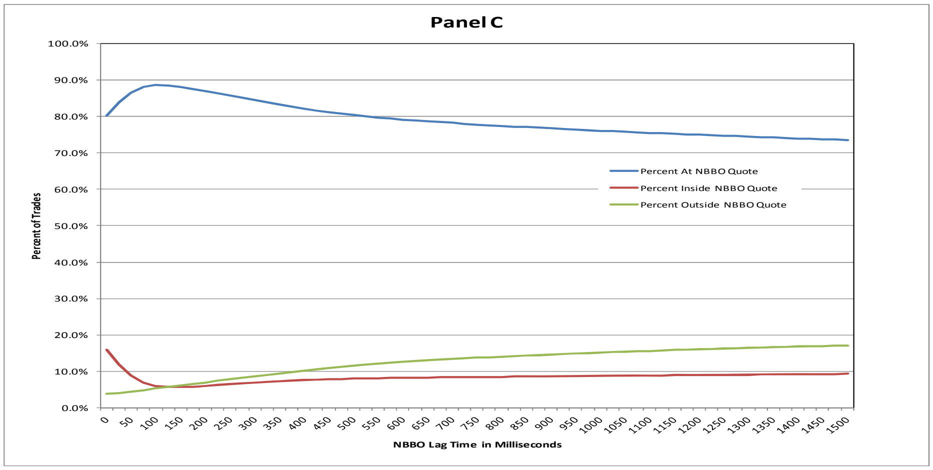
Figure A1, Panel B, shows the alignment results for off-exchange trades. While the lag that yields the highest proportion of trades at the NBBO is similar to that for Panel A, the distribution of trade location is significantly different. At the optimal lag, almost 35% of trades are executed inside the NBBO quote, which, we believe, reflects the impact of Dark Pool trading reported through the TRFs.

Figure A1, 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 A2 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 A2, Panel A, shows the exchange results and Figure A2, Panel B, shows the off exchange results. Figure A2 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.


Panel A
Percent of Trades
NBBO Lag Time in Milliseconds
Percent At NBBO Quote
Percent Inside NBBO Quote
Percent Outside NBBO Quote


Panel B
Percent of Trades
NBBO Lag Time in Milliseconds
Percent at NBBO Quote
Percent Inside NBBO Quote
Percent Outside NBBO Quote



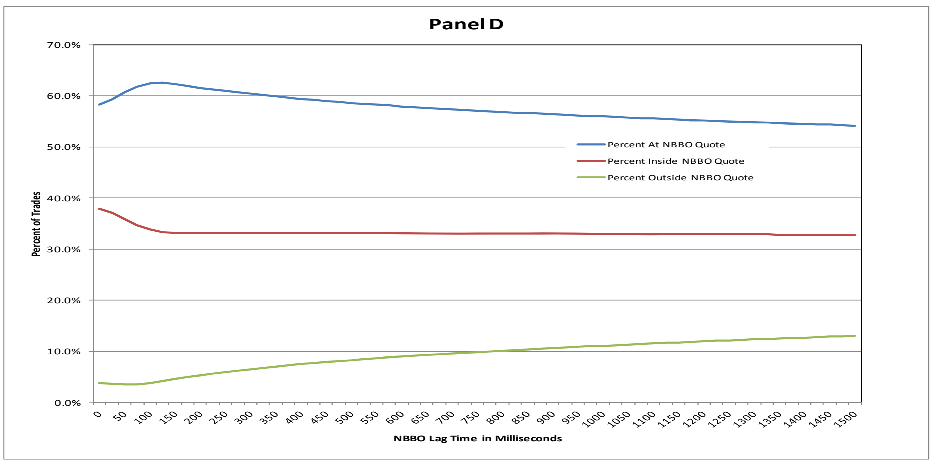
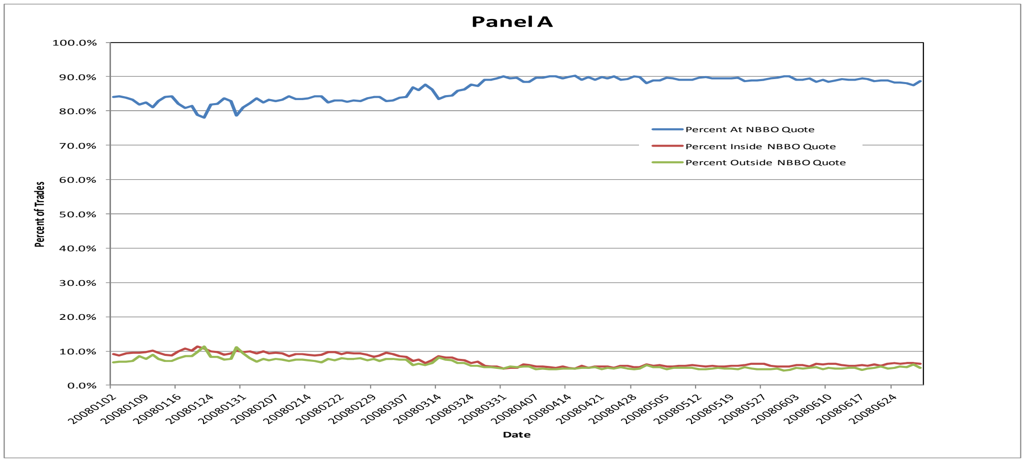


Figure A1: Trade price grid location as a function of NBBO quote lag time. For each stock for each day, we align trades and quotes using the time lag that maximizes the number of trades at the NBBO. We plot the percent of trades that occur at the NBBO quote, inside the NBBO quote, and outside the NBBO quote as a function of the quote lag time. Exchange trades and off-exchange trades are evaluated separately. On 10 March 2008, the NYSE significantly upgraded its computer systems. Panels A and B show the alignment results for exchange and off-exchange trades, respectively, prior to the upgrade. Panels C and D show the alignment for the post period for exchange and off-exchange trades, respectively.



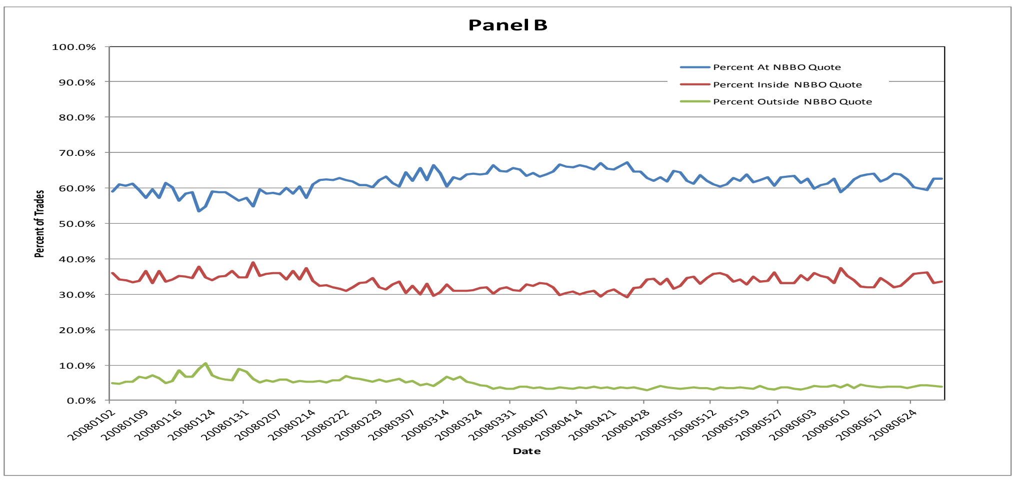


Figure A2: Trade/Quote Alignment for Exchange and Off-exchange trades. Panel A shows the time series of the trade location based on the optimum alignment lag time between the exchange calculated NBBO and exchange executed trades. Panel B shows the times series of off-exchange trades based on the optimal 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.

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.

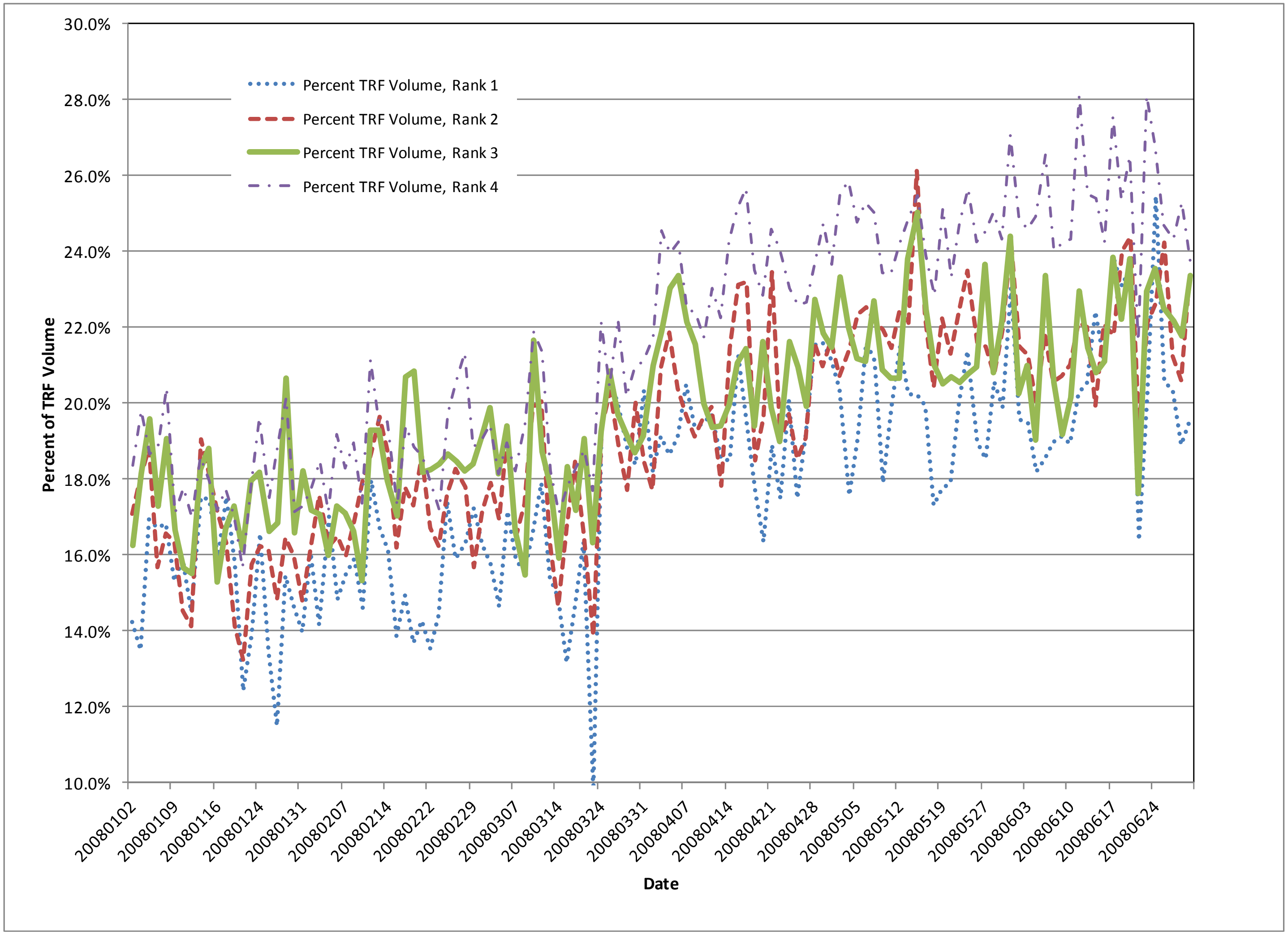


Figure 1: Percent of Off-exchange Volume. For the period 2 January 2008 through 30 June 2008, we present the percentage of off-exchange volume for 50 randomly selected NYSE stocks for each quartile of firms based on market capitalization (Rank 1 = smallest).

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 and Price Impact Analysis of Exchange and Off-Exchange Trades

We calculate the trade weighted and volume weighted effective spreads and price impacts each stock day. Price impacts are calculated using NBBO quotes 5 minutes and 30 minutes, in turn, after the trade. The Preference Measure (PM) is the ratio of realized spreads to effective spreads based on the NBBO quote 5 minutes after the trade. 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 equal using a paired t-test. To control for outliers, the PM measure is winsorized at the 1% and 99% extremes.

	Effective Spread			Price Impact (5 Min)			Price Impact (30 Min)			PM (5 Min)		
	Exch	Off	Diff	Exch	Off	Diff	Exch	Off	Diff	Exch	Off	Diff
Panel A: Trade Weighted												
All Trades	1.35	1.22	0.13**	1.42	0.88	0.54**	1.36	0.86	0.51**	-0.07	0.26	-0.33**
ISO Trades	1.50	1.48	0.02**	1.58	1.70	-0.11**	1.53	1.68	-0.15**	-0.10	-0.18	0.08**
NISO Trades	1.23	1.16	0.07**	1.28	0.67	0.61**	1.23	0.64	0.59**	-0.04	0.40	-0.44**
Panel B: Volume Weighted												
All Trades	1.67	1.40	0.27**	1.69	0.69	1.01**	1.60	0.70	0.90**	-0.05	0.47	-0.52**
ISO Trades	1.64	1.78	-0.14**	1.69	1.66	0.04	1.62	1.64	-0.02	-0.09	-0.08	0.00
NISO Trades	1.67	1.27	0.40**	1.67	0.51	1.17**	1.57	0.52	1.05**	-0.05	0.56	-0.61**

* 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.

	Trades		
Estimate	Exchange	Off-exchange	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	Stock 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
Robustness Check of TRF Determinates
TRF order flow is heterogeneous in that trades can be executed on crossing networks, on ECN's, internalized, or another type of execution venue. We attempt to isolate these potential effects by segmenting the TRF order flow as follows. First type ISO volume is dropped. Then, for the remaining NISO volume we evaluate %TRFvol as 1) all NISO TRF volume, 2) NISO volume that executes inside the NBBO quote, and 3) NISO volume that executes only at the NBBO quote midpoint. 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 adjusted 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.

Variable	NISO TRF	NISO In Quote	NISO At Mid
Intercept	0.140**	0.044**	0.023**
Spd	0.038**	0.013**	0.009**
Isig	-0.001	0.000	0.000
MpVar	-0.021**	-0.009**	-0.007**
Liq	-0.001	0.000	-0.001
Turnx100	0.109**	0.034**	0.032**
N	24,957	24,957	24,957
Adj R2	0.244	0.130	0.103

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

Table 10

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 11

Robustness Test of Information Share Regression

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. Since TRF volume is heterogeneous we break out TRF volume three ways. First, using only NISO volume, second NISO volume executed inside the quote, and finally as NISO volume executed at the quote midpoint. %*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.

	Information Ratio			Information Share		
Variable	NISO	In Quote	At Mid	NISO	In Quote	At Mid
Intercept	0.871**	1.022**	1.056**	0.923**	0.908**	0.906**
%TRFvol	1.675**	1.886**	2.091**	-0.135**	-0.078**	-0.073**
Spd	-0.048**	-0.009**	-0.002	-0.025**	-0.029**	-0.029
Isig	0.001	-0.001	0.000	0.001*	0.001*	0.001
MpVar	0.045**	0.027**	0.024**	0.039**	0.041**	0.041**
Liq	-0.003*	-0.003*	-0.003	-0.003**	-0.003**	-0.003
Turnx100	-0.028	0.090**	0.001**	-0.060**	-0.072**	-0.072**
N	24,957	24,957	24,957	24,957	24,957	24,957
Adj R2	0.477	0.350	0.280	0.258	0.249	0.248

* significant at the 5% level

** significant at the 1% level

Table 12
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

Table 13

Robustness test of Order 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. For the No Mid column, TRF trades at the quote midpoint are dropped. For the No Inside column, TRF trades inside of the quote are dropped. All TRF ISO volume is dropped from this analysis. 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.

	No Mid	No Inside
$ExOb_t$	7.13**	7.13**
$ExOb_{t-1}$	-1.27**	-1.25**
$ExOb_{t-2}$	-0.46*	-0.42*
$ExOb_{t-3}$	-0.10	-0.13
$ExOb_{t-4}$	-0.18	-0.17
$TrfOb_t$	-0.25*	-0.35**
$TrfOb_{t-1}$	-0.34**	-0.09
$TrfOb_{t-2}$	-0.07	-0.01
$TrfOb_{t-3}$	0.00	0.10
$TrfOb_{t-4}$	0.17	0.11

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