An Empirical Analysis of Market Fragmentation on U.S. Equities Markets

Frank Hatheway
The NASDAQ OMX Group, Inc.
Amy Kwan
The University of Sydney
Capital Markets Cooperative Research Center
Hui Zheng*
The University of Sydney
Capital Markets Cooperative Research Center

Disclaimer

- We thank NASDAQ OMX for providing the data for this research. The data employed in this research are equivalent to the TAQ data publicly available through databases such as WRDS. The views expressed herein are not intended to represent the views of NASDAQ OMX, its employees, or directors. The authors are solely responsible for the content, which is provided for informational and educational purposes only. Nothing contained herein should be construed as investment advice, either on behalf of a particular security or an overall investment strategy.

Introduction

- Market fragmentation is an important feature of equity market development in recent time
- An important question is what is the impact of market fragmentation on market quality
- Our study contributes to the answer by examining the impact of market fragmentation in the form of dark pools.
- In particular, we aim to answer the following questions:
  - What is the relationship between dark trading and market quality?
  - Why and how does it happen?
- Our study has important policy implications for the regulation of international equities markets

Institutional Background

- The fair access rules of Reg ATS (Rule 301(b)(3)) exempt ATSs that execute less than 5% trading volume in an NMS security from:
  - Being included in the consolidated quotation data
  - Providing equal access to market participants
- These special exemptions make it possible for small ATSs, who choose to operate as dark pools, to:
  - Maintain a non-transparent market structure
  - Select their customers
  - Allow price increments at sub-penny levels

What We Find

- Dark venues offer sub-penny price improvement which provides an incentive for order flow to migrate
- Orders are significantly less informed on dark venues
- Our results suggest that dark venues rely on their unique features to attract less informed order flow
- This action leaves liquidity providers worse off on lit markets where there is a concentration of informed orders
- It is widely recognized that liquidity providers offset their losses from trading against informed traders with gains from the uninformed
- When informed and uninformed order flow is segmented by venue type, there is little incentive for liquidity providers to stay on the informed market (Admati and Pfleiderer, 1988)

What We Find (Cont.)

- The impact on:
  - Transaction costs:
    - Bolton, Santos and Scheinkman (2011) demonstrate that the worsened condition on lit markets will in term affect the market quality of dark venues,
  - Price efficiency:
    - Madhavan (2011) “fragmented trading can thin out order books in any given venue”, and prices become more volatile
    - Pagano (1989b) show that this increase in volatility will further discourage risk averse investors from entering the market, leaving the market more susceptible to order imbalances and volatility.

What We Find (Cont.)

- Consistent with these predictions, we find that on average dark fragmentation leads to higher transaction costs and lower price efficiency across all markets :
  - a 10% increase in total dark market share will lead to 5.4% increase in transaction costs market wide.
- We also find that the execution of large trades on dark venues does not harm market quality.

Data

TAQ equivalent Trade and Quote Data

1. Stratified sample of **116 stocks**, 57 NASDAQ and 59 NYSE
   1. The 120 sample stocks for HFT studies less four stocks which were delisted by end of sample period.
2. Sample period **January 1 – March 31, 2011**
3. Data is from the Tape Plans and is scrubbed as follows:
   1. Only regular way trades 9:30:30 to 16:00:00
   2. Quotes with one sided quotes and with spreads outside (0.00 , $2.00) deleted
   3. Winsorize trades outside the NBBO to be at the same side of the NBBO
4. Trade and quote matching
   1. Lee & Ready for buy/sell initiation
   2. Lit venues contemporaneous match
   3. Dark venues trades matched with quotes 40ms prior.

Data Summary Stats

Table 1. Sample Descriptive Statistics

The sample consists of trades from 116 stocks occurring between January 3 2011 and March 31 2011. Reported are cross-sectional values averages across the 116 stocks. *Market cap.* is the stock's market capitalization on January 3 2011 in billions. All remaining variables are daily averages. *Quoted spread* is the time-weighted average difference between the bid and ask prices in dollars. *Price* is the average trade price in dollars. *Volume* is the average number of shares traded in millions. *Ntrades* is the average number of transactions in thousands. *Trade size* is the average daily trade size in number of shares. *Market share* is the venue volume (i.e. lit volume or dark volume) of shares traded divided by the total volume of shares traded, expressed as a percentage. Stocks are sorted into terciles based on market capitalization, and Panels B to D report these statistics for each tercile group.

	Lit					Dark				
	Mean	**Stddev**	**Q1**	**Median**	**Q3**	**Mean**	**Stddev**	**Q1**	**Median**	**Q3**
Panel A: Full sample										
Market cap.	20.72	44.95	0.644	2.256	20.77					
Quoted spread	0.038	0.071	0.011	0.020	0.035					
Price	46.44	70.04	16.26	30.79	54.35	46.44	70.04	16.26	30.80	54.35
Volume	3.132	7.123	0.144	0.339	2.668	1.426	3.747	0.056	0.116	1.104
Ntrades	13.57	21.10	1.068	2.833	18.06	3.903	7.793	0.239	0.505	3.734
Trade size	153.4	59.5	123.9	134.0	151.6	256.1	77.88	208.1	235.2	280.5
Market share	73.80	5.231	70.78	73.91	77.04	26.20	5.231	22.96	26.09	29.22

Transaction Costs

$$Effective\ spread_t = q_t \frac{(p_t - m_t)}{m_t}$$

Total transaction costs paid by liquidity takers, or equivalently total revenue received by liquidity providers

$$Adverse\ selection_t = q_t \frac{(m_{t+30} - m_t)}{m_t}$$

Costs of liquidity provision due to adverse selection

$$Realized\ spread_t = q_t \frac{(p_t - m_{t+30})}{m_t} + \frac{rebate}{m_t}$$

Costs of liquidity provision due to adverse selection

Results: Comparison of Spread Decompositions Adjusted for Lit Market Fees

Qspread	Lit			Dark			Dark – Lit		
	Effective Spreads	Adverse Selection	Realized Spreads	Effective Spreads	Adverse Selection	Realized Spreads	Effective Spreads	Adverse Selection	Realized Spreads
Panel A: Full sample									
Small	2.253	3.239	-0.059	1.819	0.642	1.176	-0.434 ***	-2.596 ***	1.2345 ***
Medium	3.397	3.720	0.625	2.872	0.814	2.054	-0.525 ***	-2.906 ***	1.4285 ***
Large	5.227	3.830	2.391	5.417	1.293	4.124	0.190 *	-2.537 ***	1.7331 ***

Results: Information Contribution to Market Price

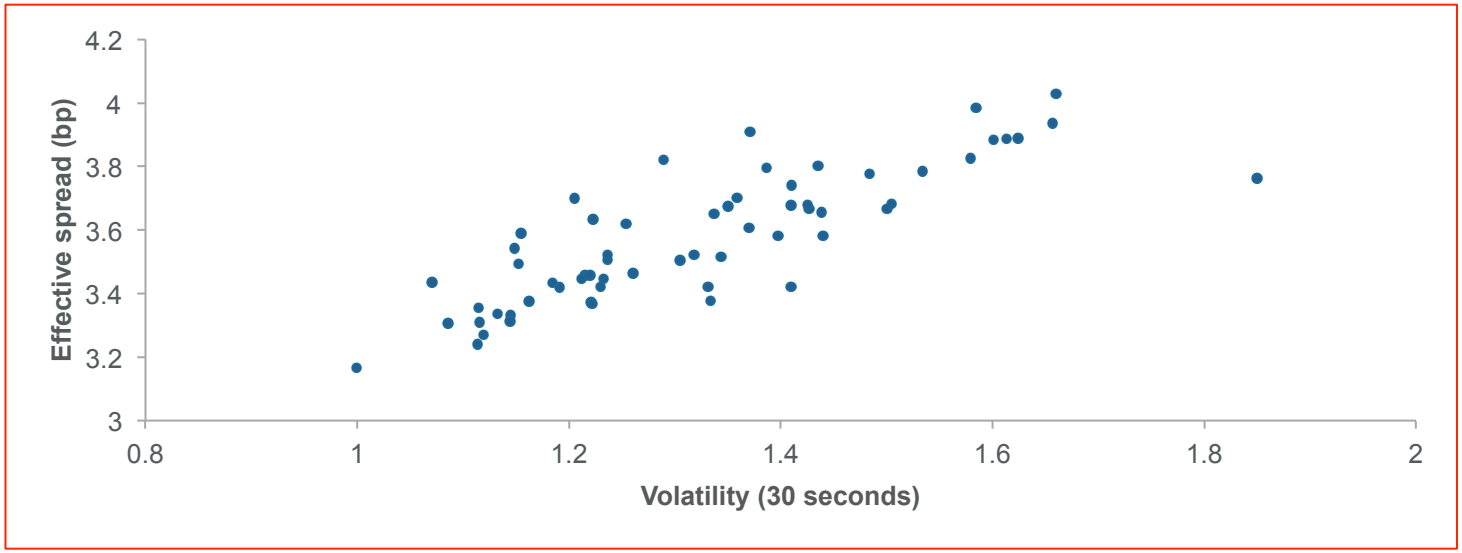
Hasbrouck (1995) 1 Minute Sampling interval

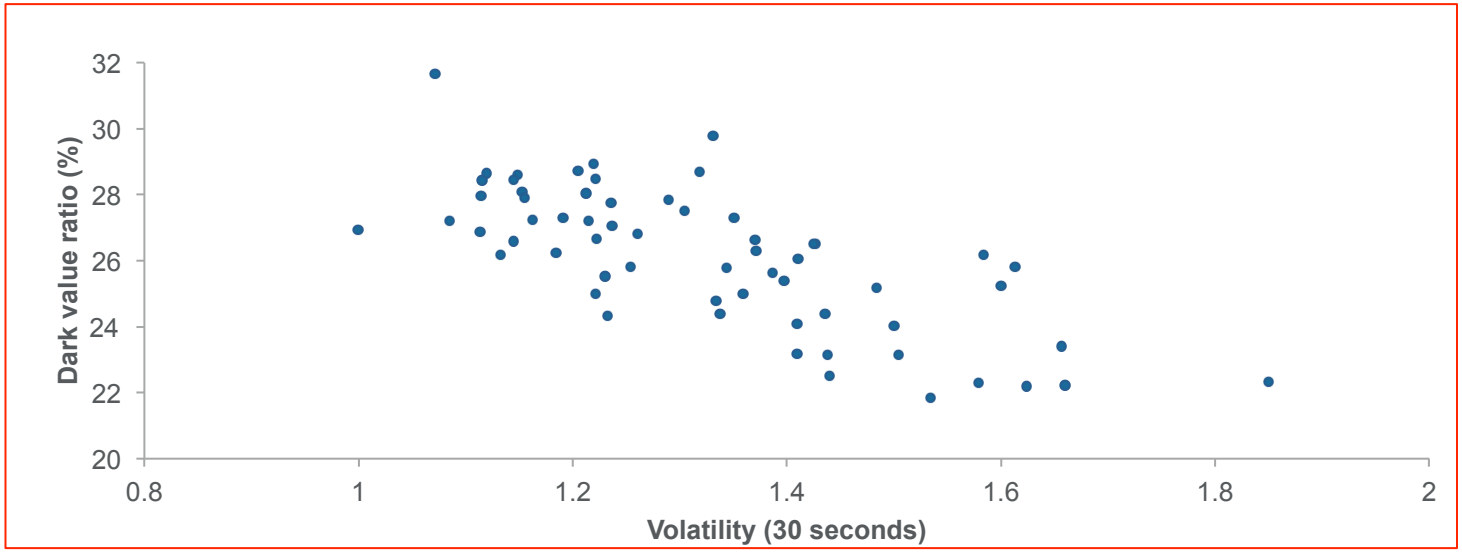
	Lit			Dark		
	Maximum	**Minimum**	**Midpoint**	**Maximum**	**Minimum**	**Midpoint**
Panel B: 1 minute						
All	0.903	0.507	0.705	0.493	0.097	0.295
	(0.064)	(0.257)	(0.118)	(0.257)	(0.064)	(0.118)
Large stocks	0.947	0.182	0.564	0.818	0.053	0.436
	(0.017)	(0.087)	(0.044)	(0.087)	(0.017)	(0.044)
Medium stocks	0.914	0.636	0.775	0.364	0.086	0.225
	(0.042)	(0.172)	(0.086)	(0.172)	(0.042)	(0.086)
Small stocks	0.849	0.694	0.771	0.306	0.151	0.229
	(0.074)	(0.068)	(0.057)	(0.068)	(0.074)	(0.057)

Heckman's Two-step Procedure

- The dependent variable is *Dark_value_ratio*, which is the proportion of total trading value on dark venues.
- *Stage 1:* $Dark\_value\_ratio_{it} = \phi(Z_{it}\gamma + \mu_{it})$
- *Stage 2:* $Eff_{spread_{it}} = Dark\_value\_ratio_{it}\beta_1 + X_{it}\beta_2 + \theta\hat{\lambda}_{it} + \varepsilon_{it}$
- Where $\phi(\cdot)$ is the standard normal cumulative distribution function and $\hat{\lambda}$ is the inverse Mills ratio.
- The control variables for Stage 1 regression are:
  - *Mcap*: market capitalization
  - *Price*: daily value weighted average trade price
  - *Trade_size_ratio*: the ratio of the average trade size on day *t* and the average trade size for the whole sample period for each stock
  - *Total_value*: the daily dollar turnover
  - *Value_imbalance*: daily buy value minus sell value
- The control variables for the second stage regression include *Price*, *Trade_size_ratio* and *Total_value*.
- Except for *Dark_value_ratio* and *Trade_size_ratio*, all variables are log transformed.

Trading Risk, Spreads and Dark Market Share





Additional controls

- *Post-trade volatility:* the standard deviation of one second midpoint returns in the 30 seconds after the transaction averaged over the trading day
- The probability of informed trading (PIN) for each stock day
- We also control for the execution of “blocks” on dark venues. Block trades are defined as the largest 1% trades during sample period.

Table 5: Transaction Costs for Consolidated Markets

Controlling for the Informational Environment and BLOCKs

	Stage 1			Stage 2		
	Coefficient	StdErr		Coefficient	StdErr.	
Dark_value_ratio				**1.8593**	**0.6842**	**
Dark_block_ratio				**-1.5937**	**0.2728**	***
Lambda				-2.4991	1.4875	*
PIN				0.0575	0.0236	**
Volatility	-452.45	52.326	***	1.6057	0.1648	***
Price	-0.1551	0.0053	***	-0.2888	0.2393	
Trade_size_ratio	1,721.0	291.26	***	5.9222	1.4338	***
Total_value				-0.3188	0.0848	***
Value_imbalance	0.5484	0.0418	***			
Mcap	0.0642	0.0028	***			
Intercept	-1.5646	0.0582	***	10.400	2.6469	***
Adj-R^2	0.1431			0.7785		

Table 7: Role of Market Cap

For large and mid-cap stocks effect on dark pool market is statistically significant. Small caps are indeterminate

	All			Lit			Dark		
	Coeff.	**StdErr.**		**Coeff.**	**StdErr.**		**Coeff.**	**StdErr.**	
Panel I: Large stocks									
Dark_value_ratio	1.865	0.487	***	2.112	0.500	***	1.990	0.534	***
Dark_block_ratio	-1.337	0.353	***	-1.384	0.362	***	-1.264	0.356	***
Panel II: Medium stocks									
Dark_value_ratio	2.393	0.765	**	2.815	0.731	***	1.674	0.900	*
Dark_block_ratio	-1.232	0.344	***	-1.143	0.341	***	-1.455	0.392	***
Panel III: Small stocks									
Dark_value_ratio	-0.833	0.799		0.565	0.661		-4.799	1.229	***
Dark_block_ratio	-1.208	0.416	**	-0.909	0.413	*	-0.808	0.604	

Other market quality measures
Price efficiency

Use the Variance Ratio test for Price Efficiency, measure as absolute value of deviation from 1.

Follow O'Hara and Ye (2011) and calculate the variance ratio as:

$$Variance\_ratio = \left| 1 - \frac{\sigma^2_{short}}{\frac{1}{n}\sigma^2_{long}} \right|$$

Where σ_{short} and σ_{long} are variances of returns measured over short and long intervals, respectively and *n* is the ratio of the intervals.

For each stock and trading day, we calculate the variance of midpoint returns at 60-second, 600-second, 900-second and 1800-second intervals. The variance ratio is then calculated over four frequencies: 60/600, 60/1800, 300/900 and 300/1800.

Table 8: Price Efficiency

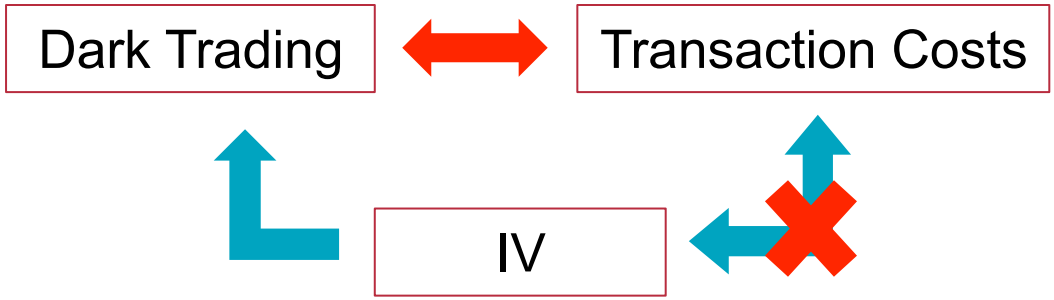
Price less efficient with more dark volume

	60/600			60/1800		
	Coefficient	Stderr		Coefficient	Stderr	
Dark_value_ratio	0.2259	0.1129	*	0.4020	0.2139	*
Dark_block_ratio	-0.1152	0.0539	*	-0.1708	0.1198	
PIN	0.0025	0.0058		0.0099	0.0148	
Volatility	33.31	149.53		-34.10	369.91	
Price	0.0189	0.0111	*	0.0367	0.0249	
Trade_size_ratio	5624	1373	***	9920	2858	***
Total_value	-0.0136	0.0059	*	-0.0188	0.0112	*
Intercept	0.4846	0.0997	***	0.8074	0.2313	***
Adj-R^2	0.0449			0.0228		

Robustness tests

Endogeneity and by-directional causality

- We show that an increase in dark trading is associated with an increase in transaction costs.
- However, the effect can be bi-directional.
- A major difficulty is to isolate the effect of dark trading on transaction costs.
- One solution is to find another variable, called an Instrumental Variable (IV), so that:



- We suggest the following instrumental variable:

$$Dark\_block\_size\_ratio = \frac{\text{Dark\_block\_size}}{\text{Average\_trade\_size}}$$

Table 9: 2SLS with instrumental Variable

$$Dark\_block\_size\_ratio = \frac{\text{Dark\_block\_size}}{\text{Average\_trade\_size}}$$

2nd stage results, previous findings hold

	Coefficient	StdErr	
Dark_block_size_ratio	2.4826	1.1286	*
Dark_block_ratio	-1.4883	0.2660	***
Lambda	-2.3964	1.5277	
Volatility	1.5848	0.1683	***
PIN	0.0631	0.0241	**
Price	-0.3177	0.2361	
Trade_size_ratio	6.2432	1.4039	***
Total_value	-0.3137	0.0834	***
Intercept	10.1000	2.6831	***
Adj-R^2	0.77		

Conclusions

- Using a representative sample of 116 stocks we try and develop a methodology to address a number of issues related to measuring the impact on market quality from a structural change in the market.
- The example taken in this paper echoes the increase in the level of 'dark' trading between 2009 and 2011.
- After controlling for trading risk, we find that trading in dark markets:
  - Increases transaction costs;
  - Reduces market price efficiency
  - The causality is from dark trading to transaction costs
  - Relationship is not driven by the level of High Frequency Trading in the market (second robustness tests, not presented)

Policy recommendations

- Our results indicate that behaviors rational from the perspective of an individual actor can be in aggregate harmful to the quality of the overall equity markets
- We provide three policy recommendations for regulators to reduce the negative impact of dark pools:
  - Expand the fair access requirement to market centers with less than 5% market share
  - Adoption of a tick size table to allow for sub-penny pricing for tick constrained stocks on all market centres
  - The implementation of a “trade-at” rule

Thank you