

## **INTERNET APPENDIX**

### **Appendix 1: Market Fragmentation and Dark Venues on U.S. Equities Markets**

In the National Market System (NMS) for U.S. equities, price information is provided to the public via the consolidated data of the two Securities Information Processors (SIPs). The SIPs widely disseminate real-time consolidated quotation data on the best-priced quotations, and consolidated trade data on trades as they are executed. The SEC defines dark liquidity as trading interest that is not included in the NMS consolidated quotation data.<sup>1</sup> The fair access rules of Regulation Alternative Trading System (ATS) require ATSS that execute more than 5% trading volume in an NMS security to provide their best-priced quotations for inclusion in the consolidated quotation data and provide traders execution access to those quotations (Rule 301(b)(3)). The SEC Rule 3a1-1 also exempts trading venues from being registered as an exchange if their trading volume is below certain thresholds.<sup>2</sup> As a result, trading on registered exchanges and electronic communication networks (i.e., the lit markets) is subject to both pre-trade and post-trade transparency, while only post-trade information is provided by dark venues. The fair access rules exempt ATSS that execute less than 5% of trading volume in an NMS security from the equal access requirement of Rule 301(b)(5). As registered exchanges or large ATSS, lit markets are not permitted to select or exclude customers.

---

<sup>1</sup> See Securities Exchange Act Release No. 60997, November 13, 2009.

<sup>2</sup> SEC Rule 3a1-1 specifies that a trading venue must be registered as an exchange if its dollar trading volume is: a) 50% or more of the daily average dollar trading volume in any security and 5% or more in any class of securities; or b) 40% or more of the daily average dollar trading volume in any class of securities. By August 2011, the largest dark pool, CrossFinder owned by Credit Suisse has a market share of less than 2.5% of consolidated volume. See Rosenblatt “Let there be light” August 2011.

The SEC identifies two sources of dark liquidity for market centers which are not part of the NMS quotation data: dark pools and broker-dealer internalization (SEC Release No. 60997, 2010).<sup>3</sup> Dark pools are registered ATSS that are not required to provide their best priced orders to the consolidated quotation data. Dark pools operate in a variety of ways (Mittal (2008)). One type of dark pool operates in a similar way to ECNs by accepting limit and market orders. In these markets, customer order flow interacts with those from other customers and potentially with the proprietary trading interest of the dark pool operator and other external liquidity partners that may include high frequency firms. Another type of dark pool operates as continuous crossing networks that cross buy and sell orders as they arrive at a price derived from the NBBO (typically at the midpoint). A third type of dark pool accepts only immediate or cancel (IOC) orders from customers, which are executed against the operator's proprietary flow at the operator's discretion.<sup>4</sup> Internalizers are broker-dealers that execute client trades either as agent or principal (SEC Release No. 60997, 2009) within their own trading system. The two main categories of internalizers are retail market makers, who handle order flow routed by retail brokerage firms and block positioners, who directly negotiate trades with customers or other broker-dealers. Similar to dark pools, internalized trades of broker-dealers represent liquidity that is not included in the consolidated

---

<sup>3</sup> In our study we are unable to further differentiate between individual dark venues due to the limitations of our dataset.

<sup>4</sup> This type of dark pool is typically not a registered ATS and thus do not fall into the SEC's definition of a dark pool. However, these trading centers offer electronic execution services that are analogous to those offered by dark pools (SEC Concept Release No. 34-61358).

quotation data. Many dark pools and broker-dealer systems are linked to each other and an order may directly or indirectly transit many dark venues in search of a counterparty.

Transactions executed on dark venues are reported first to a trade reporting facility (TRF) before the report is submitted to the consolidated trade data. Thus, dark trades are included in consolidated trade data under the TRF identifier as are lit trades from fully transparent, equal access, ECNs.<sup>5</sup> Previous research has relied on this combined data to provide an insight into the trading on dark venues (see O'Hara and Ye (2011)). Currently, all off-exchange trading volumes are reported through the FINRA/Nasdaq and FINRA/NYSE TRFs.

Most dark venues are referred to as dark pools underscoring that they are designed to conceal a trader's trading intentions. Compared with lit markets, the intended purpose of allowing pre-trade opacity on dark venues was to significantly reduce the market impact costs associated with large orders, which is especially attractive to institutional investors who usually have large orders to fill and are more concerned about information leakage. However, as presented in Table 1, dark venues today are no longer large trade facilitators and instead they compete with lit markets for trades of all sizes. Professional liquidity providers who wish to avoid quote competition when quotes are widely accessible can also benefit from the pre-trade opacity offered by dark venues. In addition, the ability to grant market access to selected customers further complicates the relationship between dark trading and market quality.

---

<sup>5</sup> The major ECNs in our sample period are Lava and Bloomberg Tradebook.

Figure A1 depicts the growth in dark trading volume from 1<sup>st</sup> July 2010 to 31<sup>st</sup> March 2011.<sup>6</sup> We calculate the daily trading volume on lit and dark markets, and also the value weighted percentage of dark trading volume. The graph shows that dark market share steadily increases from 26% in July 2010 to above 33% by the end of March 2011, representing an approximate 30% increase over the 9 month period. The rapid growth in dark trading shown in Figure A1 illustrates the importance of understanding the economic consequences of dark trading.

---

<sup>6</sup> The summary of daily trading statistics is provided by NASDAQ.

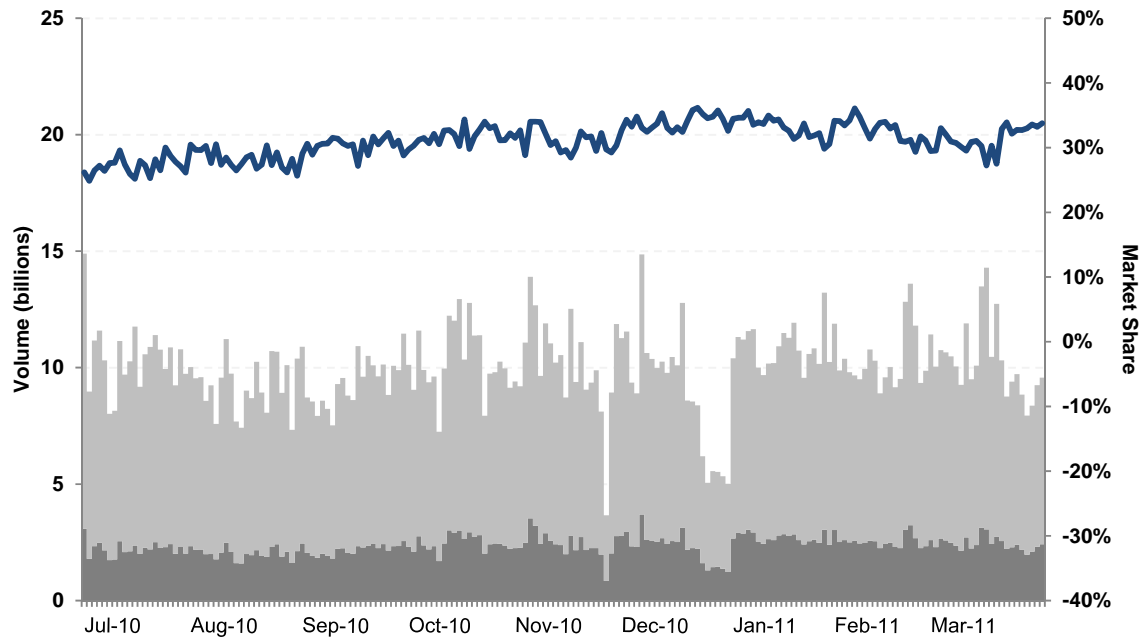


Figure A1. Growth in Dark Market Share

Figure A1 shows trends in dark market share for a sample of 116 stocks listed on NYSE and Nasdaq from July 1 2010 to March 31 2011. The lighter and darker columns represent daily consolidated trading volume on lit and dark markets, respectively. Daily dark market share is calculated by dividing dark volume by consolidated volume. We remove dark volume contributed by Direct Edge, which gained exchange status and stopped reporting to the NMS on July 19 2011.

## Appendix 2: Sample stocks

Table A2. Sample Stocks

List of 116 securities in the final sample and their respective market capitalizations on January 3 2011 (billions).

Ticker Symbol	Market Cap. (billions)	Firm Name	Ticker Symbol	Market Cap. (billions)	Firm Name
AA	16.14	Alcoa Inc.	CKH	2.199	SEACOR Holdings Inc.
AAPL	302.3	Apple Inc.	CMCSA	46.30	Comcast Corp. (Cl A)
ABD	0.485	ACCO Brands Corp.	CNQR	2.768	Concur Technologies Inc.
ADBE	15.92	Adobe Systems Inc.	COO	2.606	Cooper Cos.
AGN	21.36	Allergan Inc.	COST	31.33	Costco Wholesale Corp.
AINV	2.200	Apollo Investment Corp.	CPSI	0.512	Computer Programs & Systems Inc.
AMAT	18.79	Applied Materials Inc.	CPWR	2.571	Compuware Corp.
AMED	1.013	Amedisys Inc.	CR	2.452	Crane Co.
AMGN	52.48	Amgen Inc.	CRI	1.681	Carter's Inc.
AMZN	82.68	Amazon.com Inc.	CRVL	0.584	Corvel Corp.
ANGO	0.390	AngioDynamics Inc.	CSCO	113.6	Cisco Systems Inc.
APOG	0.394	Apogee Enterprises Inc.	CSE	2.318	CapitalSource Inc.
ARCC	3.451	Ares Capital Corp.	CSL	2.469	Carlisle Cos.
AXP	52.24	American Express Co.	CTRN	0.367	Citi Trends Inc.
AYI	2.556	Acuity Brands Inc.	CTSH	22.87	Cognizant Technology Solutions Corp.
AZZ	0.510	AZZ Inc.	DCOM	0.525	Dime Community Bancshares
BAS	0.693	Basic Energy Services Inc.	DELL	26.62	Dell Inc.
BHI	24.71	Baker Hughes Inc.	DIS	71.62	Walt Disney Co.
BIIB	16.01	Biogen Idec Inc.	DK	0.402	Delek US Holdings Inc.
BRCM	20.19	Broadcom Corp.	DOW	40.60	Dow Chemical Co.
BRE	2.822	BRE Properties Inc.	EBAY	37.39	eBay Inc.
BXS	1.378	BancorpSouth Inc.	EBF	0.452	Ennis Inc.
BZ	0.687	Boise Inc.	ERIE	3.383	Erie Indemnity Co. (Cl A)
CB	18.37	Chubb Corp.	ESRX	29.63	Express Scripts Inc.
CBEY	0.478	Cbeyond Inc.	EWBC	2.920	East West Bancorp Inc.
CBT	2.558	Cabot Corp.	FCN	1.700	FTI Consulting Inc.
CBZ	0.313	CBIZ Inc	FFIC	0.451	Flushing Financial Corp.
CCO	0.600	Clear Channel Outdoor Holdings Inc.	FL	3.062	Foot Locker Inc
CDR	0.433	Cedar Shopping Centers Inc.	FMER	2.210	FirstMerit Corp.
CELG	28.25	Celgene Corp.	FPO	0.658	First Potomac Realty Trust
CETV	1.205	Central European Media Enterprises Ltd.	FRED	0.553	Fred's Inc.

Table A2 – *Continued*

Ticker Symbol	Market Cap. (billions)	Firm Name	Ticker Symbol	Market Cap. (billions)	Firm Name
FULT	2.083	Fulton Financial Corp.	MFB	0.574	Maidenform Brands Inc.
GAS	2.294	Nicor Inc.	MIG	0.553	Meadowbrook Insurance Group Inc.
GE	195.4	General Electric Co.	MMM	62.04	3M Co.
GENZ	18.59	Genzyme Corp.	MOD	0.801	Modine Manufacturing Co.
GILD	29.70	Gilead Sciences Inc.	MOS	33.98	Mosaic Co.
GLW	29.99	Corning Inc.	MRTN	0.480	Marten Transport Ltd.
GOOG	150.1	Google Inc. (Cl A)	MXWL	0.509	Maxwell Technologies Inc.
GPS	13.65	Gap Inc.	NC	0.743	NACCO Industries Inc. (Cl A)
HON	42.29	Honeywell International Inc.	NSR	1.945	NeuStar Inc. (Cl A)
HPQ	93.62	Hewlett-Packard Co.	NUS	1.894	Nu Skin Enterprises Inc. (Cl A)
IMGN	0.654	Immunogen Inc.	NXTM	1.238	NxStage Medical Inc.
INTC	116.3	Intel Corp.	PBH	0.593	Prestige Brands Holdings Inc.
IPAR	0.585	Inter Parfums Inc.	PFE	141.6	Pfizer Inc.
ISIL	1.824	Intersil Corp. (Cl A)	PG	183.8	Procter & Gamble Co.
ISRG	10.57	Intuitive Surgical Inc.	PNC	32.32	PNC Financial Services Group Inc.
JKHY	2.534	Jack Henry & Associates Inc.	PNY	2.039	Piedmont Natural Gas Co.
KMB	25.54	Kimberly-Clark Corp.	PPD	0.610	Pre-Paid Legal Services Inc.
KNOL	0.592	Knology Inc.	PTP	1.776	Platinum Underwriters Holdings Ltd.
KR	14.00	Kroger Co.	RIGL	0.402	Rigel Pharmaceuticals Inc.
LANC	1.616	Lancaster Colony Corp.	ROC	3.070	Rockwood Holdings Inc.
LECO	2.820	Lincoln Electric Holdings Inc.	ROCK	0.423	Gibraltar Industries Inc.
LPNT	1.968	Lifepoint Hospitals Inc.	ROG	0.633	Rogers Corp.
LSTR	2.028	Landstar System Inc.	RVI	0.802	Retail Ventures Inc.
MAKO	0.591	MAKO Surgical Corp.	SF	2.217	Stifel Financial Corp.
MANT	0.963	ManTech International Corp. (Cl A)	SFG	2.136	StanCorp Financial Group Inc.
MDCO	0.768	Medicines Co.	SJW	0.489	SJW Corp.
MELI	3.094	MercadoLibre Inc.	SWN	13.20	Southwestern Energy Co.

### Appendix 3: Testing for Trade Time Adjustment

The process of interleaving quote and trade data based on reported time stamps may introduce errors in trade classification when quotes are reported ahead of the trade (Lee and Ready (1991)). Since the adoption of Regulation National Market System (Reg. NMS) in 2005, all market centers must comply with the Order Protection Rule (Reg. NMS Rule 611). Under the Order Protection Rule, trades cannot be executed at prices inferior to protected quotations, i.e., top of book limit orders on other exchanges that are eligible for NBBO Quote participation. Strictly speaking, if the volume of an incoming order exceeds the depth available at the NBBO, the order first executes against the volume at the NBBO, a new NBBO is established, and the remainder of the order will execute against orders displayed at the new NBBO prices. Thus, we should not observe orders executed outside the quotes.

However, there are three reasons why trades can occur at prices outside the NBBO under the current SEC rules and the market data reporting mechanism. Two of these reasons result from exceptions to the Order Protection Rule.<sup>7</sup> The first is the use of Intermarket Sweep Order (ISO). According to the SEC, ‘the intermarket sweep exception enables trading centers that receive sweep orders to execute those orders immediately, without waiting for better-priced quotations in other markets to be updated’ (SEC, Final Rules and Amendments to Joint Industry Plans, p. 35). Chakravarty, Jain, Upson, and Wood (2012) document that 46% of trades in their sample are ISOs, of which 16% are executed outside the quotes (Chakravarty et al. (2012, Table 2)). The second reason is the adoption of the Flickering Quote Exception, which allows market centers to “trade at

---

<sup>7</sup> See SEC, Final Rules and Amendments to Joint Industry Plans, Release No. 34-51808; File No. S7-10-04.



any price equal to or better than the least aggressive best bid or best offer, as applicable, displayed by the other trading center during that one-second window” (SEC, Final Rules and Amendments to Joint Industry Plans, p. 102).<sup>8</sup> Finally, the NBBO and Last Sale data feeds within the Securities Information Processor (SIP, the system that consolidates and disseminates quote and trade data to market participants) are controlled by separate systems and it is possible that a reporting lag exists between trades and quotes (see, Chakravarty et al., 2012, p. 419-420). Reporting delays will result in observations in which the trade price appears to be outside the NBBO even if the trade actually occurred at or within the NBBO.

Thus, trades occurring outside the quotes could be valid, if the trade results from an exception to the Order Protection Rule or invalid, if they are due to reporting delays. Moreover, there are major differences in methods of trade reporting for lit and dark trades. While lit trades are reported directly to the tapes, dark trades are first reported to a trade reporting facility (TRF) before transmission to the trade tapes,<sup>9</sup> resulting in longer reporting delays for dark trades than for lit trades. Lee and Ready (1991) demonstrate that imposing a 5 second delay on trades can greatly reduce trade misclassification. Using more recent data, Ellis, Michaely, and O’Hara (2000) and Bessembinder (2003a) show that trades are best matched to contemporaneous quotes. However, these studies do not address the differences in reporting methods by lit and dark venues.

---

<sup>8</sup> See McNish and Upson (2013) for an academic examination of this rule.

<sup>9</sup> For the majority of dark trades, reporting must be completed within 30 seconds of trade execution ([http://finra.complinet.com/en/display/display\\_viewall.html?rbid=2403&element\\_id=4439&record\\_id=5533](http://finra.complinet.com/en/display/display_viewall.html?rbid=2403&element_id=4439&record_id=5533)).

To minimize the effects of reporting delays, we follow Bessembinder (2003a) and Chakravarty et al. (2012) and test the impact of timing errors on trade and quote matching for the full sample of 116 stocks on 10 randomly selected days. There are a total of 19,896,655 transactions in the sample. The frequency of trades occurring at or inside the best bid price and ask price (INQUOTE%) is used to proxy for the degree of matching accuracy. Table A3-1 below reports INQUOTE% when trades are matched to quotes in effect 1, 5, 10, 50, 100 and 200 milliseconds prior to the trade report time, separated into lit and dark trading venues.<sup>10</sup>

Table A3-1. Trade and Quote Matching Accuracy

Table A3-1 reports INQUOTE% when trades are matched to quotes at various time lags. The sample contains trades on 10 randomly selected trading days between January 3 2011 and March 31 2011 for all 116 stocks in the sample (19,896,655 observations). The indicated time delay is subtracted from the reported time of each trade. INQUOTE% is the percentage of trades that fall within the prevailing best bid and ask price at the adjusted time of the trade.

	Time delay (ms)						
	0	1	5	10	50	100	200
All	96.39	96.21	95.74	95.18	93.01	91.80	90.34
Lit	97.08	96.81	96.14	95.31	92.20	90.78	89.17
Dark	93.99	94.14	94.36	94.71	95.80	95.32	94.37

For the full sample of transactions, there is a monotonic decrease in INQUOTE% as a longer time delay is imposed on trades. Dividing transactions into Lit and Dark indicates that there are systematic differences in the way lit and dark trades are reported. At no delay, INQUOTE% is 97.08% for Lit and 93.99% for Dark. Increasing the time delay reduces INQUOTE% for Lit monotonically. Thus, it is optimal to match lit trades to contemporaneous quotes. For dark trades, however, INQUOTE% improves until a delay of 50 milliseconds. At 100 and 200 millisecond

<sup>10</sup> In results not presented, we also tested time delays up to 10 seconds in 1 second intervals.

delays, INQUOTE% is lower. From these results, the optimal time delay for dark trades is between 10 and 100 milliseconds. Similar tests are repeated for dark trades using 5 millisecond increments occurring between 10 and 100 milliseconds. Table A3-2 reports the results for time delays of 30 to 60 milliseconds in 5 millisecond increments.

Table A3-2. Trade and Quote Matching Accuracy for Dark Venues

Table A3-2 reports INQUOTE% when dark trades are matched to quotes at various time lags. The sample contains dark trades on 10 randomly selected trading days between January 3 2011 and March 31 2011 for all 116 stocks in the sample (4,457,877 observations). The indicated time delay is subtracted from the reported time of each dark trade. INQUOTE% is the percentage of trades that fall within the prevailing best bid and ask price at the adjusted time of the trade.

Time delay (ms)						
30	35	40	45	50	55	60
95.74	95.81	95.84	95.82	95.80	95.76	95.73

Table A3-2 shows that the optimal time delay is 40 milliseconds, as this time delay corresponds to the highest value for INQUOTE%. Based on these results, dark trades are matched to quotes in effect 40 milliseconds before the trade report time. No time delay is applied to lit trades.

Relying on trade classification algorithms to infer buys and sells can be problematic, especially in a high frequency setting (see Easley, Lopez de Prado, and O'Hara (2012) for a discussion). We conduct an additional test to assess the accuracy of the Lee and Ready (1991) trade classification algorithm using transaction data from the NASDAQ TotalView-ITCH database (publicly available). The data contain transaction records with buy and sell identifiers inferred from the original order, which we match with the NASDAQ trades contained in our sample. We then

compare the Lee and Ready (1991) trade classifications with the actual buy and sell indicators recorded in the NASDAQ data.

Table A3-3 shows the results of the comparison. In our sample, 27.61 million transactions are executed on NASDAQ. Within these trades, 27.36 million transactions (99.1%) are matched to those in the NASDAQ ITCH data based on price, time and volume. Of these matched trades, 23.55 million transactions (86.1%) are correctly assigned as buy or sell transactions. Breaking down the sample of stocks into market capitalization terciles, we find that trades from large stocks (86.2%) are classified slightly more accurately than those of small (85.2%) or medium stocks (85.9%).

Table A3-3 Accuracy of the Lee and Ready (1991) Algorithm

Table A3-3 reports the number and percentage of Nasdaq trades correctly classified. Stocks are ranked into terciles based on their market capitalization on January 3 2011.

	Sample	NASDAQ		Correctly Classified	
All Stocks	27,609,942	27,364,337	(99.1%)	23,554,968	(86.1%)
Large Stocks	24,307,996	24,105,016	(99.2%)	20,772,911	(86.2%)
Medium Stocks	2,643,669	2,610,847	(98.8%)	2,225,120	(85.2%)
Small Stocks	658,277	648,474	(98.5%)	556,937	(85.9%)

#### Appendix 4: Effective Spreads, Adverse Selection Costs and Realized Spreads

Table A4-1. Comparison of Relative Effective Spreads between Lit and Dark Markets

Table A4-1 contains a comparison of effective spreads between lit and dark markets. Trade and quote data for 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Stocks are ranked into terciles based on their market capitalization on January 3 2011. For each stock, transactions are ranked into terciles based on the size of the prevailing quoted spread (QSPREAD) at the time of the trade. Daily value weighted effective spreads are calculated for each stock across QSPREAD terciles and venue types. Reported are the mean and median effective spreads in basis points for each venue type. The difference in mean (median) effective spreads between lit and dark venues is tested based on a two-tailed *t*-test (Wilcoxon signed rank test). \*\*\* and \* indicate significance levels of 0.1% and 5%, respectively

QSPREAD	Lit		Dark		Dark-Lit		
	Mean	Median	Mean	Median	Mean	Median	
<i>Panel A: Full Sample</i>							
Small	2.454	1.826	2.272	1.564	-0.182	***	-0.262 ***
Medium	3.598	2.822	3.251	2.337	-0.347	***	-0.485 ***
Large	5.460	3.802	5.852	3.808	0.392	***	0.005 ***
<i>Panel B: Large Stocks</i>							
Small	1.249	0.922	1.214	0.816	-0.034		-0.106 ***
Medium	1.402	1.126	1.344	1.053	-0.058		-0.073 ***
Large	1.502	1.388	1.952	1.746	0.450	***	0.358 ***
<i>Panel C: Medium Stocks</i>							
Small	1.811	1.345	1.672	1.233	-0.139	*	-0.113 ***
Medium	2.471	2.072	2.214	1.854	-0.257	***	-0.218 ***
Large	4.175	3.802	4.521	3.693	0.346	***	-0.109
<i>Panel D: Small Stocks</i>							
Small	3.824	3.400	3.552	2.926	-0.272	*	-0.474 ***
Medium	6.078	5.413	5.478	4.613	-0.600	***	-0.800 ***
Large	10.590	9.637	10.848	9.531	0.259		-0.106

Table A4-2. Comparison of Relative Adverse Selection Costs between Lit and Dark Markets

Table A4-2 contains a comparison of adverse selection costs between lit and dark markets. Trade and quote data for 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Short-term (30 seconds) and long-term (5 minutes) adverse selection costs are calculated for each transaction. Stocks are ranked into terciles based on their market capitalization on January 3 2011. For each stock, transactions are ranked into terciles based on the size of the prevailing quoted spread (QSPREAD) at the time of the trade. Daily value weighted adverse selection costs are calculated for each stock across QSPREAD terciles and venue types. Reported are the mean and median adverse selection costs in basis points for each venue type. The difference in mean (median) adverse selection costs between lit and dark venues is tested based on a two-tailed *t*-test (Wilcoxon signed rank test). \*\*\* indicates a significance level of 0.1%.

QSPREAD		Lit		Dark		Dark - Lit			
		Mean	Median	Mean	Median	Mean	Median		
<i>Panel A: Full Sample</i>									
Short-term	Small	3.239	2.488	0.642	0.390	-2.596	***	-2.098	***
	Medium	3.720	2.77	0.814	0.481	-2.906	***	-2.289	***
	Large	3.830	2.875	1.293	0.710	-2.537	***	-2.165	***
Long-term	Small	3.564	2.418	1.046	0.571	-2.518	***	-1.847	***
	Medium	4.394	2.947	1.068	0.649	-3.326	***	-2.298	***
	Large	4.565	3.254	1.496	0.991	-3.070	***	-2.263	***
<i>Panel B: Large Stocks</i>									
Short-term	Small	1.544	1.241	0.393	0.279	-1.152	***	-0.963	***
	Medium	1.787	1.388	0.424	0.330	-1.364	***	-1.058	***
	Large	1.569	1.372	0.454	0.411	-1.115	***	-0.960	***
Long-term	Small	1.650	1.222	0.485	0.362	-1.165	***	-0.860	***
	Medium	2.007	1.369	0.620	0.422	-1.387	***	-0.948	***
	Large	1.727	1.287	0.336	0.469	-1.392	***	-0.818	***
<i>Panel C: Medium Stocks</i>									
Short-term	Small	2.564	2.222	0.544	0.434	-2.020	***	-1.788	***
	Medium	2.819	2.546	0.588	0.485	-2.232	***	-2.062	***
	Large	3.221	2.946	1.052	0.716	-2.169	***	-2.230	***
Long-term	Small	2.803	2.304	0.980	0.673	-1.823	***	-1.631	***
	Medium	3.140	2.719	0.734	0.642	-2.406	***	-2.077	***
	Large	3.787	3.433	1.157	1.133	-2.630	***	-2.299	***
<i>Panel D: Small Stocks</i>									
Short-term	Small	4.961	4.379	0.904	0.677	-4.057	***	-3.703	***
	Medium	5.820	4.772	1.284	0.855	-4.536	***	-3.917	***
	Large	6.637	5.745	2.324	1.547	-4.312	***	-4.198	***
Long-term	Small	5.507	4.932	1.497	0.979	-4.009	***	-3.953	***
	Medium	7.118	5.791	1.677	1.247	-5.441	***	-4.544	***
	Large	8.102	6.746	2.924	2.215	-5.178	***	-4.531	***

Table A4-3. Comparison of Relative Realized Spread between Lit and Dark Markets

Table A4-3 contains a comparison of realized spread between lit and dark markets. Trade and quote data for 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Short-term (30 seconds) and long-term (5 minutes) realized spreads are calculated for each transaction. Stocks are ranked into terciles based on their market capitalization on January 3 2011. For each stock, transactions are ranked into terciles based on the size of the prevailing quoted spread (QSPREAD) at the time of the trade. Daily value weighted realized spreads are calculated for each stock across QSPREAD terciles and venue types. Realized spreads are also reported after adjustment for the amount of the liquidity rebate or charge imposed by the market center where the trade is executed, which are outlined in Internet Appendix 5. Reported are the mean and median realized spreads in basis points for each venue type. The difference in means (medians) between the realized spread on dark venues and the rebate-adjusted realized spread on lit markets is tested based on a two-tailed *t*-test (Wilcoxon signed rank test). \*\*\* indicates a significance level of 0.1%.

		Lit				Dark		Dark - Lit		
QSPREAD		No Rebate Adj.		Rebate Adj.						
		Mean	Med.	Mean	Med.	Mean	Med.	Mean	Median	
<i>Panel A: Full Sample</i>										
Short-term	Small	-0.786	-0.501	0.143	0.145	1.623	1.060	1.481	***	0.914 ***
	Medium	-0.134	-0.174	0.823	0.479	2.432	1.565	1.610	***	1.086 ***
	Large	1.610	0.673	2.621	1.427	4.559	2.922	1.938	***	1.495 ***
Long-term	Small	-1.100	-0.538	-0.167	0.135	1.205	0.896	1.371	***	0.762 ***
	Medium	-0.798	-0.307	0.161	0.343	2.166	1.397	2.005	***	1.054 ***
	Large	0.891	0.489	1.903	1.180	4.364	2.647	2.461	***	1.467 ***
<i>Panel B: Large Stocks</i>										
Short-term	Small	-0.296	-0.253	0.198	0.196	0.821	0.657	0.623	***	0.461 ***
	Medium	-0.386	-0.246	0.139	0.231	0.920	0.781	0.781	***	0.550 ***
	Large	-0.101	0.031	0.522	0.497	1.496	1.274	0.974	***	0.778 ***
Long-term	Small	-0.400	-0.193	0.098	0.244	0.734	0.648	0.636	***	0.404 ***
	Medium	-0.603	-0.158	-0.076	0.289	0.726	0.750	0.802	***	0.461 ***
	Large	-0.253	0.087	0.370	0.560	1.617	1.212	1.247	***	0.652 ***
<i>Panel C: Medium Stocks</i>										
Short-term	Small	-0.755	-0.693	-0.065	-0.104	1.130	0.970	1.195	***	1.074 ***
	Medium	-0.356	-0.297	0.360	0.324	1.625	1.405	1.265	***	1.081 ***
	Large	0.930	0.720	1.767	1.451	3.465	2.874	1.698	***	1.424 ***
Long-term	Small	-0.984	-0.711	-0.289	-0.065	0.696	0.845	0.985	***	0.910 ***
	Medium	-0.668	-0.465	0.052	0.157	1.475	1.274	1.423	***	1.117 ***
	Large	0.378	0.347	1.216	1.084	3.363	2.645	2.147	***	1.561 ***

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

Table A4-3 – *Continued*

QSPREAD		Lit				Dark		Dark - Lit			
		No Rebate Adj.		Rebate Adj.				Dark - Lit			
		Mean	Med.	Mean	Med.	Mean	Med.	Mean	Median		
Panel D: Small Stocks											
Short-term	Small	-1.136	-0.853	0.292	0.503	2.630	2.145	2.339	***	1.642	***
	Medium	0.234	0.351	1.696	1.782	4.184	3.472	2.488	***	1.690	***
	Large	3.953	3.438	5.516	4.995	8.531	7.324	3.016	***	2.330	***
Long-term	Small	-1.664	-1.130	-0.231	0.222	2.005	1.794	2.237	***	1.572	***
	Medium	-1.046	-0.406	0.417	1.002	3.768	3.216	3.352	***	2.214	***
	Large	2.515	2.501	4.079	4.016	7.944	6.880	3.865	***	2.863	***



## **Appendix 5: Rebate Rates**

Table A5. Rebate Rates for Selected Market Centers

Table A5 shows the maker-taker rebates that applied to selected market centers during January 3 2011 to March 31 2011. Maker (taker) pricing schedules are used in calculations of adjusted realized spreads (Table 2 and Table A4-3) and adjusted effective spreads (Table A6) respectively. Positive values represent rebates while negative values represent costs incurred, from the perspective of the maker or taker as indicated in the column headings.

Venue	Maker (cents per share)	Taker (cents per share)
<i>Maker-taker</i>		
NYSE	0.15	-0.23
ARCA	0.30	-0.30
NASDAQ OMX	0.295	-0.30
PSX <sup>1</sup>	0.24/0.26	-0.27
BATS BZX	0.27	-0.28
EDGX <sup>2</sup>	0.26/0.23	-0.30
<i>Reverse maker-taker</i>		
BX	-0.15	0.14
BATS BYX	0	0.03
EDGA	-0.025	0.015

<sup>1</sup> PSX provides a 0.24 (0.26) cents per share rebate for orders less than (equal to or more than) 2,000 shares.

<sup>2</sup> EDGX changed the rebate for adding liquidity on March 1 2011 to 0.23 cents per share. Previously, the rebate for adding liquidity was 0.26 cents per share.

## Appendix 6: Effective Spreads Adjusted for Maker-Taker Pricing

Section 3.1 compares effective spreads between lit and dark markets without taking into account maker-taker pricing schemes. We adjust the effective spread for each transaction for the taker fee or rebate based on the market center where execution occurs. Dark venues may also impose maker-taker pricing schedules. However, in contrast to lit markets, they are not required to publicly disclose their pricing schemes and for this reason, we make no adjustment to the effective spreads on dark venues.

Effective spreads for lit market are adjusted as:

$$\text{EFFECTIVE\_SPREAD} = q_t \frac{(p_t - m_t)}{m_t} + \frac{\text{REBATE}}{m_t},$$

where  $m_t$  is the bid ask midpoint at the time when the current trade takes place,  $p_t$  is the trade price and  $q_t$  is a buy-sell indicator, which equals to 1 (-1) if the trade is buyer- (seller-) initiated. The results are reported in Table A6.

Compared to the results in Table 2, the magnitude of effective spreads on lit markets are greater after the rebate adjustment. This is because most lit markets impose a cost to traders for taking liquidity (see Internet Appendix 5). Consistent with Table 2, effective spreads are significantly larger on lit markets than dark venues across all stock sizes and QSPREAD terciles.

Table A6. Comparison of Relative Effective Spreads Adjusted for Maker-Taker Pricing between Lit and Dark Markets

Table A6 contains a comparison of effective spreads between lit and dark markets. Trade and quote data of 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Stocks are ranked into terciles based on their market capitalization on January 3 2011. For each stock, transactions are ranked into terciles based on the size of the prevailing quoted spread (QSPREAD) at the time of the trade. Daily value weighted effective spreads are calculated for each stock across QSPREAD terciles and venue types. Reported are the mean and median effective spreads in basis points for each venue type. Effective spreads are adjusted for maker-taker costs and rebates outlined in Appendix 3. The difference of effective spreads cost between the lit and dark markets is tested, and \* and \*\*\* indicates a significance level of 5% and 0.1% respectively based on a two-tailed t-test (Wilcoxon signed rank test) of the differences in means (medians).

QSPREAD	Lit		Dark		Dark - Lit			
	Mean	Median	Mean	Median	Mean		Median	
<i>Panel A: Full Sample</i>								
Small	3.272	2.438	1.819	1.277	-1.453	***	-1.161	***
Medium	4.449	3.414	2.872	2.108	-1.577	***	-1.307	***
Large	6.338	4.525	5.417	3.510	-0.921	***	-1.015	***
<i>Panel B: Large Stocks</i>								
Small	1.666	1.240	0.896	0.601	-0.770	***	-0.639	***
Medium	1.863	1.362	1.053	0.841	-0.810	***	-0.521	***
Large	2.060	1.790	1.636	1.503	-0.423	***	-0.287	***
<i>Panel C: Medium Stocks</i>								
Small	2.396	1.737	1.309	1.004	-1.088	***	-0.732	***
Medium	3.130	2.652	1.927	1.647	-1.203	***	-1.005	***
Large	4.917	4.400	4.072	3.472	-0.845	***	-0.928	***
<i>Panel D: Small Stocks</i>								
Small	5.116	4.872	2.925	2.621	-2.191	***	-2.251	***
Medium	7.363	6.802	4.957	4.294	-2.406	***	-2.508	***
Large	11.918	11.122	10.317	9.159	-1.600	***	-1.963	***

## Appendix 7: Cream Skimming by Dark Venues

In this appendix we analyze how dark venues could engage in cream-skimming activity. Previous research find evidence consistent with the cream-skimming of order flow by competing exchanges (Easley, Kiefer, and O'Hara (1996), Chordia and Subrahmanyam (1995), and Lin, Sanger, and Booth, (1995)) and by dealers on the upstairs markets (Madhavan and Cheng (1997)).<sup>11</sup> In the following analysis, we provide evidence on the use of sub-penny price improvement by dark venues to attract order flow and end with a discussion on current market regulations that allow dark venues to screen customers based on their information and thus, cream-skim the order flow.

### 1. Sub-penny pricing and price improvement

The 'sub-penny' rule of Reg NMS creates a minimum tick size of 1 cent for all stocks in our sample.<sup>12</sup> However, we observe a significant number of transactions taking place at sub-penny

---

<sup>11</sup> Previous literature has also identified another type of cream-skimming activity, i.e., order preferencing. For example, Chung, Chuwonganant and McCormick (2004) find that less informed orders are more likely to be preferenced and this order preferencing increases transaction costs. Since market makers use quotations to compete for liquidity (Bessembinder (2003b)), the practice of order preferencing creates disincentives for posting competitive quotes (Huang and Stoll (1997)). In our study, price improvement offered by dark venues serve as a partial payment to brokers for routing uninformed liquidity to these markets and our findings are consistent with the conclusions of Chung et al. (2004).

<sup>12</sup> Reg NMS Rule 612 specifies that "No national securities exchange, national securities association, alternative trading system, vendor, or broker or dealer shall display, rank, or accept from any person a bid or offer, an order, or an indication of interest in any NMS stock priced in an increment smaller than \$0.01 if that bid or offer, order, or indication of interest is priced equal to or greater than \$1.00 per share".

price increments on dark venues. One benefit of sub-penny pricing is that dark venues can attract liquidity by offering price improvements.

We calculate price improvement by comparing the trade price to the NBBO price. A buy (sell) trade receives price improvement if the trade price is smaller (greater) than the NBBO ask (bid) price. A total of 13 price improvement levels are classified, ranging from trades that receive no price improvement (Level 1) to trades that receive more than 1 cent of price improvement (Level 13). The upper and lower bounds for each price improvement level are outlined in Table A7-1. For each stock, the frequency of transactions falling into each level is calculated for the lit and dark markets on each day before averaging across all trading days. Table A7-1 reports the cross-sectional means and medians of the frequencies for each price improvement level.

Table A7-1. Price Improvement on Lit and Dark

Table A7-1 contains a comparison of price improvement between Lit and Dark venues. Trade and quote data of 116 stocks listed on NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Stocks are ranked into terciles based on their market capitalization on January 3 2011. Price improvement is calculated by comparing the transaction prices with the prevailing NBBO and classified into 13 levels based on the magnitude of improvement. For each stock, the frequency of transactions with price improvement falling into each level is calculated for the lit and dark venues. The mean and median of the frequencies for each price level across stocks are reported.

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

Level	Price improvement (cents)	Lit		Dark	
		Mean	Median	Mean	Median
Panel A: Full Sample					
1	0	83.61	86.32	50.48	49.58
2	$0 < x \leq 0.10$	0.00	0.00	11.73	11.45
3	$0.10 < x \leq 0.20$	0.00	0.00	2.68	2.29
4	$0.20 < x \leq 0.30$	0.00	0.00	2.83	2.78
5	$0.30 < x \leq 0.40$	0.00	0.00	1.18	1.00
6	$0.40 < x < 0.50$	0.00	0.00	0.51	0.37
7	0.5	1.66	1.43	12.38	11.65
8	$0.50 < x < 0.60$	0.00	0.00	0.13	0.08
9	$0.60 \leq x < 0.70$	0.00	0.00	0.36	0.27
10	$0.70 \leq x < 0.80$	0.00	0.00	0.45	0.46
11	$0.80 \leq x < 0.90$	0.00	0.00	0.45	0.39
12	$0.90 \leq x < 1.00$	0.00	0.00	0.16	0.12
13	$1.00 \leq x$	14.73	12.14	16.64	17.49
Panel B: Large Stocks					
1	0	89.92	93.76	51.55	51.60
2	$0 < x \leq 0.10$	0.00	0.00	12.34	11.87
3	$0.10 < x \leq 0.20$	0.00	0.00	3.74	3.57
4	$0.20 < x \leq 0.30$	0.00	0.00	3.83	3.99
5	$0.30 < x \leq 0.40$	0.00	0.00	1.55	1.44
6	$0.40 < x < 0.50$	0.00	0.00	0.90	0.71
7	0.5	2.57	1.94	18.52	20.24
8	$0.50 < x < 0.60$	0.00	0.00	0.12	0.02
9	$0.60 \leq x < 0.70$	0.00	0.00	0.30	0.05
10	$0.70 \leq x < 0.80$	0.00	0.00	0.27	0.09
11	$0.80 \leq x < 0.90$	0.00	0.00	0.33	0.09
12	$0.90 \leq x < 1.00$	0.00	0.00	0.15	0.04
13	$1.00 \leq x$	7.50	2.46	6.41	1.56
Panel C: Medium Stocks					
1	0	82.58	84.66	50.76	49.19
2	$0 < x \leq 0.10$	0.00	0.00	10.83	9.47
3	$0.10 < x \leq 0.20$	0.00	0.00	2.24	1.92
4	$0.20 < x \leq 0.30$	0.00	0.00	2.62	2.61
5	$0.30 < x \leq 0.40$	0.00	0.00	1.03	0.76
6	$0.40 < x < 0.50$	0.00	0.00	0.33	0.24
7	0.5	1.34	1.17	11.09	10.42
8	$0.50 < x < 0.60$	0.00	0.00	0.13	0.10
9	$0.60 \leq x < 0.70$	0.00	0.00	0.37	0.25
10	$0.70 \leq x < 0.80$	0.00	0.00	0.51	0.51
11	$0.80 \leq x < 0.90$	0.00	0.00	0.45	0.37
12	$0.90 \leq x < 1.00$	0.00	0.00	0.17	0.14
13	$1.00 \leq x$	16.07	14.03	19.47	22.79

Table A7-1 – *Continued*

Level	Price improvement (cents)	Lit		Dark	
		Mean	Median	Mean	Median
Panel D: Small stocks					
1	0	78.49	78.63	49.17	49.14
2	$0 < x \leq 0.10$	0.00	0.00	12.03	11.66
3	$0.10 < x \leq 0.20$	0.00	0.00	2.11	2.12
4	$0.20 < x \leq 0.30$	0.00	0.00	2.07	2.02
5	$0.30 < x \leq 0.40$	0.00	0.00	0.97	0.92
6	$0.40 < x < 0.50$	0.00	0.00	0.31	0.26
7	0.5	1.09	0.88	7.70	6.55
8	$0.50 < x < 0.60$	0.00	0.00	0.15	0.12
9	$0.60 \leq x < 0.70$	0.00	0.00	0.42	0.39
10	$0.70 \leq x < 0.80$	0.00	0.00	0.58	0.53
11	$0.80 \leq x < 0.90$	0.00	0.00	0.56	0.51
12	$0.90 \leq x < 1.00$	0.00	0.00	0.17	0.14
13	$1.00 \leq x$	20.42	21.05	23.77	24.00

Panel A of Table A7-1 presents the results for the full sample of stocks. Over 80% of transactions on the lit markets are executed at the NBBO and do not receive price improvement. In contrast, about 50% of dark trades receive some level of price improvement. Although most of this benefit is less than 1 cent, these results provide strong evidence that dark venues provide price improvement more frequently than lit markets.

A closer examination reveals differences in how lit and dark markets provide price improvement. First, more than 20% of transactions on dark venues receive sub-penny price improvement after excluding trades with a price increment of 0.5 cents (Level 7). In contrast, no transactions take place at the sub-penny levels on lit markets. Second, the difference between lit and dark markets for price improvement over a penny (Level 13) is less apparent. We find that 14.73% of lit transactions receive price improvement of at least 1 cent, which is only slightly below the 16.64%

on dark venues.<sup>13</sup> Given that price improvement of at least 1 cent can only occur when the bid-ask spread is greater than 1 cent, our results suggest that lit venues are better positioned to compete with dark venues when minimum tick sizes are not a constraining factor on spreads.

Panels B – D report the price improvement frequencies for the market capitalization subsamples. For the lit markets, the frequency of trades receiving price improvement is negatively related to stock sizes, while the percentage of trades receiving price improvement on dark venues remains relatively constant across stocks of different sizes. The difference between the percentage of trades receiving price improvement in lit and dark markets decreases with stock size. Since spreads are more likely to be wider for small stocks, these results are consistent with our findings in Panel A and suggest that sub-penny pricing is less valuable to dark venues when spreads are wide.<sup>14</sup> Our results are consistent with Buti, Consonni, Rindi, Wen, and Werner (2014) who find that the use of sub-penny pricing is more prevalent in low priced stocks that are tick size constrained.

## 2. Regulation on fair access

The results in Table A7-1 suggest that compared with lit markets dark venues offer a price

---

<sup>13</sup> When price improvement is below 1 cent, a small portion of lit transactions (1.66%) receive price improvement of 0.5 cent through mid-point peg orders. Limit orders submitted to the lit markets can peg to the midpoint of the spread as a hidden order. See, SEC Release No. 34-57537, File No. SR-NASDAQ-2008-021.

<sup>14</sup> Acting as broker-dealers, dark venues can accept orders with terms that allow them to derive an explicit price at sub-penny increments. However, the broker-dealer cannot accept an order that is explicitly priced at a sub-penny increment. See “Responses to Frequently Asked Questions Concerning Rule 612 (Minimum Pricing Increment) of Regulation NMS”, available at: <http://www.sec.gov/divisions/marketreg/subpenny612faq.htm>.



advantage to liquidity takers through the widespread use of sub-penny pricing. Hence all liquidity demanders are attracted naturally to the lower explicit transaction costs in dark venues. However, it is possible that not all can trade on dark venues, because under the current regulatory environment dark venues are not subject to the fair access rule and thus can prohibit or limit access to their services (see Reg ATS Rule 301(b)(5); SEC Concept Release 2010). Our main results show that the adverse selection risk is significantly lower on dark venues, suggesting that dark venues preferentially screen for less informed orders. Taken together, our findings support the notion that dark venues attract specifically uninformed order flow from lit markets by bettering the displayed prices on lit markets through sub-penny price improvement.

This order segmentation activity has been documented for retail market makers, who attract uninformed orders through payment for order flow agreements with retail brokers (Easley et al. (1996), Bessembinder and Kaufman (1997), and Chung et al. (2004)). In Seppi's (1990) model, upstairs brokers offer superior pricing because they can screen block traders based on their information. Likewise, the evidence we document on dark venues is essentially a variation of these practices; price improvement is analogous to payments for order flow and the ability to identify the uninformed is similar to the screening mechanisms modeled by Seppi (1990).<sup>15</sup>

Overall, our results suggest that by offering sub-penny price improvement dark venues attract uninformed order flow to the adverse selection risk on their markets. Liquidity providers on dark

---

<sup>15</sup> Battalio, Hatch and Jennings (2003) argue that the payment for order flow may partially flow through brokers to investors in the form of lower commissions and/or better services. Therefore, there are also incentives for investors to participate in the payment for order flow scheme.

venues are able to offer discounts to uninformed orders while still receiving a large reward for their provision of liquidity. The effect is a segmentation of uninformed and informed order flow.

### 3. Price improvement by QSPREAD

Our results in Table 2 show that effective spreads are lower on the dark venues than on the lit markets. However, the opposite result holds when quoted spreads are wide. Specifically, Panel A of Table 2 shows that quoted spreads are 23.9% and 18.3% higher on lit markets than dark venues in the two lowest QSPREAD categories. In contrast, traders pay 3.5% more to trade on dark venues when QSPREAD is wide. One reason for this finding is that lit markets are able to compete for order flow by offering price improvement through mid-point peg orders when quoted spreads are wide.

More specifically, the amount of price improvement offered by lit markets is compared to that of dark venues based on the prevailing QSPREAD at the time of the trade. To calculate price improvement, we compute the distance between the trade price and the prevailing NBBO price depending on whether the trade is buyer or seller initiated. We define 13 price improvement levels and the cross-sectional frequencies of trades occurring at each level are reported in Table A7-2. When spreads are narrow, the results in Panel A show that dark liquidity providers are able to lower transaction costs by offering sub-penny price improvement; 43% of transactions in dark venues are price improved while only 4.8% of transactions in lit markets are price improved. When quoted spreads are wide, the difference in the frequency of trades receiving price improvement in lit and dark venues is much smaller. Panel C shows that 45% of lit trades and 59% of dark trades are price improved. Furthermore, we find that the amount of price improvement offered by lit

markets is larger than the amount offered by dark venues. Approximately all of transactions receiving price improvement in lit markets are improved by 1cent or more while only 58% of trades that are price improved receive 1 cent or more in dark venues. These results are consistent with the observation that transaction costs are lower in lit markets relative to dark venues when quoted spreads are wide. In wide spread environments, sub-penny pricing is less valuable to dark venues, allowing lit markets to more successfully compete for order flow.

Table A7-2. Price improvement by QSPREAD

Table A7-2 contains a comparison of price improvement between lit and dark markets. Trade and quote data of 116 stocks listed on NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. All transactions reported within the sample period are ranked into terciles based on the size of the prevailing quoted spread (QSPREAD) at the time of the trade for every stock. Price improvement is calculated by comparing the transaction prices with the prevailing NBBO and classified into 13 levels based on the magnitude of improvement. For each stock, the frequency of transactions with price improvement falling into each level is calculated for the lit and dark markets. The mean and median of the frequencies for each price level across stocks are reported.

Level	Price improvement (cents)	Lit		Dark	
		Mean	Median	Mean	Median
<i>Panel A: Low</i>					
1	0	95.22	97.59	57.06	56.82
2	$0 < x \leq 0.10$	0.00	0.00	11.50	11.36
3	$0.10 < x \leq 0.20$	0.00	0.00	2.63	2.31
4	$0.20 < x \leq 0.30$	0.00	0.00	4.04	3.98
5	$0.30 < x \leq 0.40$	0.00	0.00	1.16	0.95
6	$0.40 < x < 0.50$	0.00	0.00	0.49	0.34
7	0.5	2.05	1.97	18.51	19.26
8	$0.50 < x < 0.60$	0.00	0.00	0.03	0.00
9	$0.60 \leq x < 0.70$	0.00	0.00	0.08	0.00
10	$0.70 \leq x < 0.80$	0.00	0.00	0.11	0.00
11	$0.80 \leq x < 0.90$	0.00	0.00	0.11	0.00
12	$0.90 \leq x < 1.00$	0.00	0.00	0.02	0.00
13	$1.00 \leq x$	2.73	0.00	4.25	0.00

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

Table A7-2 – *Continued*

Level	Price improvement (cents)	Lit		Dark	
Panel B: Medium					
1	0	84.30	85.90	50.55	49.76
2	0 < x ≤ 0.10	0.00	0.00	11.28	10.94
3	0.10 < x ≤ 0.20	0.00	0.00	2.63	2.22
4	0.20 < x ≤ 0.30	0.00	0.00	2.80	2.53
5	0.30 < x ≤ 0.40	0.00	0.00	1.12	0.98
6	0.40 < x < 0.50	0.00	0.00	0.50	0.39
7	0.5	1.47	0.97	13.18	12.11
8	0.50 < x < 0.60	0.00	0.00	0.14	0.00
9	0.60 ≤ x < 0.70	0.00	0.00	0.33	0.06
10	0.70 ≤ x < 0.80	0.00	0.00	0.49	0.12
11	0.80 ≤ x < 0.90	0.00	0.00	0.35	0.16
12	0.90 ≤ x < 1.00	0.00	0.00	0.18	0.00
13	1.00 ≤ x	14.23	12.61	16.46	13.10
Panel C: High					
1	0	54.83	56.20	41.13	41.66
2	0 < x ≤ 0.10	0.00	0.00	10.83	10.19
3	0.10 < x ≤ 0.20	0.00	0.00	2.12	2.07
4	0.20 < x ≤ 0.30	0.00	0.00	1.41	1.39
5	0.30 < x ≤ 0.40	0.00	0.00	1.12	0.95
6	0.40 < x < 0.50	0.00	0.00	0.38	0.27
7	0.5	0.54	0.23	5.44	3.66
8	0.50 < x < 0.60	0.00	0.00	0.24	0.15
9	0.60 ≤ x < 0.70	0.00	0.00	0.67	0.54
10	0.70 ≤ x < 0.80	0.00	0.00	0.94	0.87
11	0.80 ≤ x < 0.90	0.00	0.00	1.03	0.88
12	0.90 ≤ x < 1.00	0.00	0.00	0.46	0.27
13	1.00 ≤ x	44.63	43.13	34.24	35.50

## Appendix 8: Hasbrouck (1995) Information Shares

Table A8-1. Hasbrouck Information Shares (10 second)

Information shares are estimated daily using 10 lags at 10 second sampling intervals. Lit and Dark represents the average of the maximum and minimum market contributions based on the ordering of the variables in the Cholesky factorization.

Symbol	Lit	Dark	Symbol	Lit	Dark	Symbol	Lit	Dark	Symbol	Lit	Dark
AA	0.593	0.407	CELG	0.778	0.222	FL	0.839	0.161	MDCO	0.908	0.092
AAPL	0.580	0.420	CETV	0.886	0.114	FMER	0.895	0.105	MELI	0.866	0.134
ABD	0.908	0.092	CKH	0.936	0.064	FPO	0.853	0.147	MFB	0.911	0.089
ADBE	0.728	0.272	CMCSA	0.628	0.372	FRED	0.855	0.145	MIG	0.810	0.190
AGN	0.887	0.113	CNQR	0.947	0.053	FULT	0.826	0.174	MMM	0.824	0.176
AINV	0.791	0.209	COO	0.930	0.070	GAS	0.911	0.089	MOD	0.901	0.099
AMAT	0.617	0.383	COST	0.855	0.145	GE	0.578	0.422	MOS	0.686	0.314
AMED	0.881	0.119	CPSI	0.807	0.193	GENZ	0.760	0.240	MRTN	0.775	0.225
AMGN	0.746	0.254	CPWR	0.731	0.269	GILD	0.714	0.286	MXWL	0.882	0.118
AMZN	0.683	0.317	CR	0.954	0.046	GLW	0.605	0.395	NC	0.745	0.255
ANGO	0.761	0.239	CRI	0.911	0.089	GOOG	0.784	0.216	NSR	0.907	0.093
APOG	0.866	0.134	CRVL	0.619	0.381	GPS	0.667	0.333	NUS	0.901	0.099
ARCC	0.777	0.223	CSCO	0.660	0.340	HON	0.831	0.169	NXTM	0.895	0.105
AXP	0.697	0.303	CSE	0.718	0.282	HPQ	0.596	0.404	PBH	0.851	0.149
AYI	0.943	0.057	CSL	0.945	0.055	IMGN	0.878	0.122	PFE	0.596	0.404
AZZ	0.771	0.229	CTRN	0.876	0.124	INTC	0.608	0.392	PG	0.652	0.348
BAS	0.923	0.077	CTSH	0.864	0.136	IPAR	0.769	0.231	PNC	0.815	0.185
BHI	0.789	0.211	DCOM	0.782	0.218	ISIL	0.773	0.227	PNY	0.910	0.090
BIIB	0.873	0.127	DELL	0.600	0.400	ISRG	0.863	0.137	PPD	0.625	0.375
BRCM	0.675	0.325	DIS	0.672	0.328	JKHY	0.913	0.087	PTP	0.893	0.107
BRE	0.940	0.060	DK	0.775	0.225	KMB	0.819	0.181	RIGL	0.855	0.145
BXS	0.888	0.112	DOW	0.707	0.293	KNOL	0.846	0.154	ROC	0.935	0.065
BZ	0.764	0.236	EBAY	0.658	0.342	KR	0.642	0.358	ROCK	0.854	0.146
CB	0.882	0.118	EBF	0.828	0.172	LANC	0.851	0.149	ROG	0.828	0.172
CBEY	0.825	0.175	ERIE	0.672	0.328	LECO	0.858	0.142	RVI	0.897	0.103
CBT	0.923	0.077	ESRX	0.799	0.201	LPNT	0.912	0.088	SF	0.940	0.060
CBZ	0.738	0.262	EWBC	0.873	0.127	LSTR	0.913	0.087	SFG	0.904	0.096
CCO	0.870	0.130	FCN	0.934	0.066	MAKO	0.837	0.163	SJW	0.678	0.322
CDR	0.816	0.184	FFIC	0.748	0.252	MANT	0.871	0.129	SWN	0.763	0.237

Table A8-2. Hasbrouck Information Shares (1 minute)

Information shares are estimated daily using 10 lags at 1 minute sampling intervals. Lit and Dark represent the average between the maximum and minimum market contributions based on the ordering of the variables in the Cholesky factorization.

Symbol	Lit	Dark	Symbol	Lit	Dark	Symbol	Lit	Dark	Symbol	Lit	Dark
AA	0.521	0.479	CELG	0.574	0.426	FL	0.633	0.367	MDCO	0.822	0.178
AAPL	0.513	0.487	CETV	0.817	0.183	FMER	0.783	0.217	MELI	0.764	0.236
ABD	0.849	0.151	CKH	0.864	0.136	FPO	0.822	0.178	MFB	0.849	0.151
ADBE	0.540	0.460	CMCSA	0.537	0.463	FRED	0.785	0.215	MIG	0.786	0.214
AGN	0.674	0.326	CNQR	0.831	0.169	FULT	0.709	0.291	MMM	0.599	0.401
AINV	0.679	0.321	COO	0.839	0.161	GAS	0.831	0.169	MOD	0.815	0.185
AMAT	0.516	0.484	COST	0.616	0.384	GE	0.517	0.483	MOS	0.539	0.461
AMED	0.805	0.195	CPSI	0.783	0.217	GENZ	0.675	0.325	MRTN	0.768	0.232
AMGN	0.556	0.444	CPWR	0.620	0.380	GILD	0.567	0.433	MXWL	0.810	0.190
AMZN	0.535	0.465	CR	0.866	0.134	GLW	0.537	0.463	NC	0.709	0.291
ANGO	0.667	0.333	CRI	0.812	0.188	GOOG	0.592	0.408	NSR	0.830	0.170
APOG	0.787	0.213	CRVL	0.650	0.350	GPS	0.559	0.441	NUS	0.798	0.202
ARCC	0.616	0.384	CSCO	0.574	0.426	HON	0.596	0.404	NXTM	0.796	0.204
AXP	0.532	0.468	CSE	0.615	0.385	HPQ	0.520	0.480	PBH	0.774	0.226
AYI	0.850	0.150	CSL	0.861	0.139	IMGN	0.798	0.202	PFE	0.551	0.449
AZZ	0.715	0.285	CTRN	0.839	0.161	INTC	0.539	0.461	PG	0.522	0.478
BAS	0.829	0.171	CTSH	0.633	0.367	IPAR	0.674	0.326	PNC	0.580	0.420
BHI	0.568	0.432	DCOM	0.759	0.241	ISIL	0.621	0.379	PNY	0.857	0.143
BIIB	0.628	0.372	DELL	0.551	0.449	ISRG	0.764	0.236	PPD	0.661	0.339
BRCM	0.534	0.466	DIS	0.538	0.462	JKHY	0.816	0.184	PTP	0.832	0.168
BRE	0.856	0.144	DK	0.731	0.269	KMB	0.583	0.417	RIGL	0.772	0.228
BXS	0.769	0.231	DOW	0.540	0.460	KNOL	0.826	0.174	ROC	0.799	0.201
BZ	0.704	0.296	EBAY	0.523	0.477	KR	0.530	0.470	ROCK	0.783	0.217
CB	0.648	0.352	EBF	0.809	0.191	LANC	0.783	0.217	ROG	0.818	0.182
CBEY	0.835	0.165	ERIE	0.680	0.320	LECO	0.810	0.190	RVI	0.797	0.203
CBT	0.811	0.189	ESRX	0.587	0.413	LPNT	0.820	0.180	SF	0.869	0.131
CBZ	0.717	0.283	EWBC	0.702	0.298	LSTR	0.838	0.162	SFG	0.819	0.181
CCO	0.833	0.167	FCN	0.806	0.194	MAKO	0.729	0.271	SJW	0.659	0.341
CDR	0.768	0.232	FFIC	0.747	0.253	MANT	0.795	0.205	SWN	0.562	0.438

## Appendix 9: O'Hara and Ye (2011) replication

We replicate the tests by O'Hara and Ye (2011) with our dataset. Following their study, we calculate the daily TRF\_RATIO as the ratio of dark volume on the overall market trading volume. The remaining variables are defined in the main paper. Table A9 reports the results of the two-stage Heckman correction model.

Table A9. O'Hara and Ye (2011) Replication

Table A9 reports the estimates for the two-stage Heckman correction model based on the model specification in O'Hara and Ye (2011). Trade and quote data of 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. The dependent variable for the first stage probit is TRF\_RATIO, which is calculated as the proportion of total trading volume on dark venues. The dependent variable for the second stage OLS is EFFSPREAD, which is the daily value weighted relative effective spread. LAMBDA is the inverse Mills ratio obtained from the first stage probit model. PRICE is the log of the daily value-weighted average price. TRADE\_SIZE\_RATIO is the ratio of the average trade size on day t and the average trade size for the whole sample period for each sample stock. TOTAL\_TRADES is the log of the daily total number of trades. TOTAL\_VOLUME is the log of the daily total trading volume. MCAP is the log of market capitalization on 3 January 2011. In the second-stage OLS, all coefficients except for TRADE\_SIZE\_RATIO are scaled by a factor of 10,000. Standard errors reported in second stage OLS are corrected for double clustering by date and stock (Thompson (2011)). \*\*\*, \*\* and \* indicate significance levels of 0.1%, 1% and 5% respectively

	First-stage OLS			Second-stage OLS		
	Coefficient	Std. Err.		Coefficient	Std. Err.	
TRF_RATIO				-3.0628	1.1994	**
LAMBDA				-3.2307	1.5036	*
PRICE	0.0448	0.0067	***	-0.9843	0.2919	***
TRADE_SIZE_RATIO	6,312	265.3	***	9.4459	2.1731	***
TOTAL_TRADES				-0.5381	0.1219	***
TOTAL_VOLUME	0.1015	0.0024	***			
MCAP	-0.1491	0.0055	***			
NASDAQ				0.1015	0.2374	
Intercept	-0.1649	0.0591	**	15.4000	2.3101	***
Adj-R	0.2924			0.6565		

Consistent with O'Hara and Ye (2011, p. 468, Table 6), our results show that TRF\_RATIO is negative and significant, indicating that market transaction costs decreases with dark trading. The signs of all other variables are also the same as those reported by O'Hara and Ye (2011). Specifically, in the first-stage model, the TRF\_RATIO is negatively correlated with PRICE and MCAP, and positively related to TRADE\_SIZE\_RATIO; in the second-stage model, EFFSPREAD is negatively related to PRICE and TOTAL\_TRADES, and positively related to TRADE\_SIZE\_RATIO. However, we do find that LAMBDA is negative and significant at the 5% level, which means that we cannot reject the hypothesis of no selection bias in our data. In addition, EFFSPREAD is positively related to TOTAL\_VOLUME while the relationship between EFFSPREAD and NASDAQ is insignificant. In summary, our results indicate that it is critical to control for the level of trading risk in the market when examining the impact of dark fragmentation.



## Appendix 10: Additional Robustness Tests

We provide three additional robustness tests to support the results reported in the main body of the paper. First, we include firm fixed effects in the second stage regression in Table 4 to better control for unobserved heterogeneity between firms. Second, we use alternative measures for market quality and examine the impact of dark trading on these variables. Finally, we examine the impact of high frequency trading on the relation between dark trading and market quality. To improve the flow of the discussion, we have relegated all tables to the end of this appendix.

### 1. Fixed effects

For our main results, we estimate the regressions using standard errors double clustered by firm and day, as described in Petersen (2009) and Thompson (2011). Gormley and Matsa (2014) argue that the fixed effects estimator should be used when there is unobserved heterogeneity among groups of observations. The results in Table A10-1 incorporate the controls for firm fixed effects into the Heckman procedure with standard errors double clustered by firm and day. Consistent with our main results, we find that DARK\_VALUE\_RATIO is positive and significant. While DARK\_BLOCK\_RATIO remains negative, the coefficient is statistically insignificant, which based on additional analysis appears to be due to a different impact of DARK\_BLOCK\_RATIO among firms of different market capitalizations.

[Insert Table A10-1]

In Table 5 of the main text, we estimate the regressions separately for large, medium and small stocks. We repeat these regressions after controlling for firm fixed effects and report the results in Table A10-2. Table A10-2 shows that DARK\_BLOCK\_RATIO is negative for all three groups and is significant for the large and medium sized firms. We also find that DARK\_VALUE\_RATIO

is positive and significant for large and medium stocks. These results are consistent with those reported in Table 5.

[Insert Table A10-2]

## 2. Alternative market quality measures

To ensure the robustness of our main conclusions in Table 4 of the main text, we test the relationship between DARK\_VALUE\_RATIO and alternate measures of market quality. Buti, Rindi, and Werner (2011) measure market quality using several proxies including quoted spreads, effective spreads and various measures of market depth. Similarly, Weaver (2014) uses dollar and percentage effective and realized spreads. Following these studies, we proxy for market quality using daily value weighted effective spreads (cents), time weighted quoted spreads (percentage and cents), time weighted total depth, bid depth and ask depth. Bid and ask depths (total depth) are the sum of the dollar volume quoted at the best bid or ask (best bid and ask) prices across all exchange venues, respectively.

Table A10-3 presents the results from the second stage regression to these alternate measures of market quality. For ease of comparison, we also reproduce our original results using effective spreads (percentage). Consistent with our main conclusions, the results show that higher levels of dark trading are correlated with lower market quality. Specifically, we find that dark trading is associated with higher quoted spreads (percentage) and lower market depth. For our results using dollar effective and quoted spreads, while DARK\_VALUE\_RATIO remains positive, the coefficient is insignificant for the full sample. To investigate further, we split the full sample into subsamples based on market capitalization and find that DARK\_VALUE\_RATIO is associated with higher dollar effective and quoted spreads, which is significant for large and medium stocks.

<sup>16</sup> This result is similar to our subsample results in Table 5 of the main text, which is based on percentage effective spreads.

[Insert Table A10-3]

In Panel A of Table A10-3, the results for dark block trading are also consistent with our main findings using the percentage effective spread. In particular, we find that higher levels of dark block trading are associated with lower quoted and effective spreads.

Turning to the results for market depth in Panel B of Table A10-3, we find that both DARK\_BLOCK\_RATIO and DARK\_VALUE\_RATIO are negatively correlated with total market depth at the best bid and ask prices as well as the depth at the best bid and ask prices separately. This result indicates that dark venues impose a crowding out effect on market liquidity. As order flow is directed towards the dark venues, market depth on the lit venues decreases.

Taken together, these additional robustness tests provide strong evidence that higher levels of dark trading are associated with lower market quality, while dark block trades tend to have the opposite effect.

### 3. High frequency trading

The recent growth in high frequency trading (HFT) has raised concerns about its impact on market quality. It is widely believed that dark venues screen out high frequency traders (HFTs). Therefore, if dark venues are selectively excluding HFTs from trading, then it is possible that the relationship between dark trading and transaction costs we observe is driven by HFT activity.

---

<sup>16</sup> The results based on market capitalization subsamples are available from the authors.

To test this hypothesis, we need to identify HFT activity, which is not available in the SIP data. We obtain data from NASDAQ that identify 26 of the most active HFT firms on the NASDAQ market over the same sample period. The total trading value of these HFT firms is \$445 billion, which represents 60.13% of the total value traded on NASDAQ. For each observation, there is information on the stock symbol, transaction date and time, price, volume, a buy (sell) indicator and indicators for whether the buyer and seller are a HFT firm. We exclude from our SIP data dark transactions that are not executed on NASDAQ, and construct the following variables to analyse the impact of HFT:

$$(8) \quad HFT_{ALL} = \frac{TOTAL\ VALUE_{HFT}}{TOTAL\ VALUE_{\times 2}},$$

$$(9) \quad HFT_{MAKE} = \frac{TOTAL\ VALUE_{HFT\_MAKE}}{TOTAL\ VALUE_{MAKE}},$$

$$(10) \quad HFT_{TAKE} = \frac{TOTAL\ VALUE_{HFT\_TAKE}}{TOTAL\ VALUE_{TAKE}}.$$

Specifically,  $TOTALVALUE_{HFT}$  is the sum of the value of transactions in which an HFT provides liquidity and the value of transactions in which an HFT takes liquidity,  $TOTALVALUE_{HFT\_MAKE}$  is the value of transactions in which an HFT provides liquidity, and  $TOTALVALUE_{HFT\_TAKE}$  is the value of transactions in which an HFT takes liquidity.  $TOTALVALUE$  is the total trading value on NASDAQ. Therefore,  $HFT_{ALL}$  measures the total level of HFT activity, while  $HFT_{MAKE}$  and  $HFT_{TAKE}$  measure the level of HFT liquidity provision and consumption, respectively. We estimate these values daily and incorporate the variables into the Heckman two-stage model. The results are reported in Table A10-4

[Insert Table A10-4]

Model 1 of Table A10-4 re-estimates Model 2 of Table 4 with the new dataset. The results are largely consistent:  $DARK\_VALUE\_RATIO$  is significantly positive while

DARK\_BLOCK\_RATIO is significantly negative. In Model 2 of Table A10-4 we include HFT<sub>ALL</sub>, our proxy for HFT activity. HFT<sub>ALL</sub> is negative and significant, indicating that the level of HFT trading activity is negatively related to effective spreads. In Model 3 of Table A10-4, we decompose total HFT activity into maker and taker components. HFT<sub>MAKE</sub> is insignificant while HFT<sub>TAKE</sub> is significantly negative, indicating that the negative correlation between HFT activity and effective spreads is associated with the trading of HFT liquidity takers. This result is consistent with Hendershott and Riordan (2013), who show that HFTs tend to take liquidity when it is cheap to do so. Hagstromer and Norden (2013) also find that HFTs engaging in arbitrage and directional strategies tend to take more liquidity following a decrease in the minimum tick size. Across all model specifications, the coefficient of DARK\_VALUE\_RATIO (DARK\_BLOCK\_RATIO) remains positive (negative) and significant, indicating that HFT activity does not drive our results.

Table A10-1. Relationship between dark fragmentation and effective spreads (replication of Table 4 after including firm fixed effects)

Table A10-1 reports the estimates for the two-stage Heckman correction model. Trade and quote data for 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Model 1 reports the results for the first stage probit model. The dependent variable, DARK\_VALUE\_RATIO, is calculated as the daily value of trades on dark venues divided by the daily total value of all trading for stock  $i$  and trading day  $t$ .

$$DARK\_VALUE\_RATIO_{it} = \Phi(\beta_0 + \beta_1 VOLATILITY_{it} + \beta_2 PRICE_{it} + \beta_3 TRADE\_SIZE\_RATIO_{it} + \beta_4 LIT\_VALUE\_IMBALANCE\_RESIDUAL_{it} + \beta_5 MCAP_i + \varepsilon_{it}),$$

$\Phi(\cdot)$  is the standard normal cumulative distribution function. LIT\_VALUE\_IMBALANCE\_RESIDUAL is the residual term from the regression:

$$LIT\_VALUE\_IMBALANCE_{it} = \alpha + \beta INF_{it} + \varepsilon_{it},$$

where LIT\_VALUE\_IMBALANCE is the absolute difference between the value of buy and sell transactions on lit markets, divided by the total trading value of all markets and INF is the absolute value of the daily returns,  $|P_{i,t}/P_{i,t-1} - 1|$ . Model 2 reports the results for the second stage OLS regression.

$$EFFSPREAD_{it} = \beta_0 + \beta_1 DARK\_VALUE\_RATIO_{it} + \beta_2 DARK\_BLOCK\_RATIO_{it} + \beta_3 LAMBDA_{it} + \beta_4 PIN_{it} + \beta_5 VOLATILITY_{it} + \beta_6 PRICE_{it} + \beta_7 TRADE\_SIZE\_RATIO_{it} + \beta_8 TOTAL\_VALUE_{it} + \varepsilon_{it}.$$

The dependent variable EFFSPREAD is the daily value weighted relative effective spread. DARK\_BLOCK\_RATIO is the daily value of the top 1% of trades on dark venues divided by the daily total value of dark trading. LAMBDA is the inverse Mills ratio obtained from the first stage probit model. PIN is estimated following the procedure in Easley, Kiefer, and O'Hara (1997) for each trading day. VOLATILITY is the standard deviation of one second midpoint returns in the 30 seconds after the transaction averaged over the trading day. Price is the log of the daily value-weighted average price. TRADE\_SIZE\_RATIO is the ratio of the average trade size on day  $t$  to the average trade size for the whole sample period for each sample stock. TOTAL\_VALUE is the log of the daily total trading value. MCAP is the log of the stock's market capitalization on January 3 2011. All coefficients except for VOLATILITY and TRADE\_SIZE\_RATIO are enlarged by a factor of 10,000. Dummy variables (not reported) are used to control for firm fixed effects. Standard errors reported in Models 2 are corrected for double clustering by date and stock (Thompson (2011)). \*\*\* indicates significance levels of 1%.

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

	Model 1			Model 2		
	Coefficient	Std. Err.		Coefficient	Std. Err.	
DARK_VALUE_RATIO				2.789	0.989	***
DARK_BLOCK_RATIO				-0.099	0.316	
LAMBDA				1.813	0.538	***
PIN				0.021	0.020	
VOLATILITY	-637.74	53.633	***	1.659	0.249	***
PRICE	-0.1672	0.0055	***	-1.237	0.168	***
TRADE_SIZE_RATIO	4,443.4	295.92	***	4.694	0.667	***
TOTAL_VALUE				-0.265	0.038	***
LIT_VALUE_IMBALANCE_RESIDUAL	-1.0722	0.0562	***			
MCAP	0.0477	0.0029	***			
Intercept	-1.1091	0.0589	***	6.816	1.037	***
Adj- $R^2$	0.1666			0.7018		

Table A10-2. Relationship between dark fragmentation and effective spreads by market capitalization

Table A10-2 reports the estimates for the second stage OLS regression of the Heckman correction model for different market capitalization subsamples. Trade and quote data for 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. Stocks are ranked into market capitalization terciles on January 3 2011. The dependent variable, EFFSPREAD, is the daily value weighted relative effective half spread for all, lit or dark markets as indicated. DARK\_VALUE\_RATIO is the proportion of total trading value on dark venues. DARK\_BLOCK\_RATIO is the daily value of the top 1% of trades on dark venues divided by the daily total value of dark trades. LAMBDA is the inverse Mills ratio obtained from the first stage probit model described in Table R1. PIN is estimated following the procedure in Easley, Kiefer and O'Hara (1997) for each trading day. VOLATILITY is the standard deviation of one second midpoint returns in the 30 seconds after the transaction averaged over the trading day. Price is the log of the daily value-weighted average price. TRADE\_SIZE\_RATIO is the ratio of the average trade size on day t to the average trade size for the whole sample period for each sample stock. TOTAL\_VALUE is the log of the daily total trading value. All coefficients except for VOLATILITY and TRADE\_SIZE\_RATIO are enlarged by a factor of 10,000. Dummy variables (not reported) are used to control for firm fixed effects. Standard errors are corrected for double clustering by date and stock (Thompson (2011)). \*\*\*, \*\* and \* indicate significance levels of 1%, 5% and 10%, respectively.

	Large stocks			Medium stocks			Small stocks		
	Coefficient	Std. Err.		Coefficient	Std. Err.		Coefficient	Std. Err.	
DARK_VALUE_RATIO	1.7906	0.4283	***	3.1290	0.8950	***	1.9717	1.3855	
DARK_BLOCK_RATIO	-0.2916	0.1738	*	-0.6684	0.3068	**	-0.1701	0.5096	
LAMBDA	-1.3758	1.2085		3.5762	1.7362	**	14.1833	3.7735	***
PIN	0.0182	0.0092	**	0.0168	0.0285		0.0610	0.0329	*
VOLATILITY	0.6393	0.0844	***	1.2520	0.2681	***	2.9901	0.4722	***
PRICE	-0.6430	0.1000	***	0.1944	0.3363		1.6081	0.8112	**
TRADE_SIZE_RATIO	22.3807	12.9160	*	12.9281	3.2167	***	-1.7685	1.8767	
TOTAL_VALUE	0.1101	0.0370	***	-0.1027	0.1364		-0.4841	0.1675	***
Intercept	2.383	1.5264		6.038	1.8568	***	20.974	3.4216	***
Adj-R <sup>2</sup>	0.7296			0.3405			0.4949		



Table A10-3. Relationship between dark fragmentation and alternate measures of market quality (replication of Table 4, Model 2)

Table A10-3 reports the estimates for the second stage of the two-stage Heckman correction model, where market quality is proxied by alternative measures. Trade and quote data for 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. The dependent variable for the second stage regression is MQPROXY, which is measured as the daily value weighted effective spread, time weighted quoted spread, time weighted total depth, bid depth or ask depth, as indicated in the column heading. Bid and ask depths (total depth) are the sum of the dollar volume quoted at the best bid or ask (best bid and ask) prices across all exchange venues.

$$MQPROXY_{it} = \beta_0 + \beta_1 DARK\_VALUE\_RATIO_{it} + \beta_2 DARK\_BLOCK\_RATIO_{it} + \beta_3 LAMBDA_{it} + \beta_4 PIN_{it} + \beta_5 VOLATILITY_{it} + \beta_6 PRICE_{it} \\ + \beta_7 TRADE\_SIZE\_RATIO_{it} + \beta_8 TOTAL\_VALUE_{it} + \varepsilon_{it}$$

DARK\_VALUE\_RATIO is calculated as the daily value of trades on dark venues divided by the daily total value of all trading for stock *i* and trading day *t*. DARK\_BLOCK\_RATIO is the daily value of the top 1% of trades on dark venues divided by the daily total value of dark trading. LAMBDA is the inverse Mills ratio obtained from the first stage probit model described in Table R3. PIN is estimated following the procedure in Easley et al. (1997) for each trading day. VOLATILITY is the standard deviation of one second midpoint returns in the 30 seconds after the transaction averaged over the trading day. PRICE is the log of the daily value-weighted average price. TRADE\_SIZE\_RATIO is the ratio of the average trade size on day *t* to the average trade size for the whole sample period for each sample stock. TOTAL\_VALUE is the log of the daily total trading value. Standard errors reported are corrected for double clustering by date and stock (Thompson (2011)). \*\*\*, \*\* and \* indicate significance levels of 1%, 5% and 10%, respectively.

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

	Effective spread (%)			Effective spread (cents)			Quoted spread (%)			Quoted spread (cents)		
	Coefficient	Std. Err.		Coefficient	Std. Err.		Coefficient	Std. Err.		Coefficient	Std. Err.	
Panel A: Effective And Quoted Spreads												
DARK_VALUE_RATIO	3.568	0.988	***	152.8	110.6		45.33	25.43	*	491.4	312.2	
DARK_BLOCK_RATIO	-0.644	0.347	*	-106.8	41.61	**	-42.59	8.414	***	-307.0	130.4	**
LAMBDA	1.589	0.856	*	201.3	80.72	**	-16.70	19.25		208.0	148.6	
PIN	0.069	0.023	***	11.08	3.739	***	1.275	0.729	*	33.51	12.18	***
VOLATILITY	1.921	0.241	***	66.83	22.91	***	43.36	8.318	***	194.1	69.42	***
PRICE	-0.670	0.209	***	126.1	32.72	***	-4.443	5.478		456.1	126.7	***
TRADE_SIZE_RATIO	6.227	1.233	***	425.3	212.6	**	215.1	50.37	***	1,096	732.7	
TOTAL_VALUE	-0.194	0.048	***	0.958	7.286		-10.00	3.889	**	-34.76	32.07	
Intercept	3.572	1.250	***	-694.6	204.1	***	233.6	56.03	***	-1,219	722.9	*
Adj-R <sup>2</sup>	0.6341			0.4753			0.2243			0.4706		

	Total Depth			Bid Depth			Ask Depth		
	Coefficient	Std. Err.		Coefficient	Std. Err.		Coefficient	Std. Err.	
Panel B: Market Depth									
DARK_VALUE_RATIO	-9.630	3.969	**	-9.139	4.220	**	-9.202	4.027	**
DARK_BLOCK_RATIO	-6.938	2.267	***	-5.546	2.048	***	-5.483	2.018	***
LAMBDA	-3.319	1.738	*	-3.666	4.609		-4.051	4.192	
PIN	0.743	0.145	***	0.479	0.130	***	0.468	0.136	***
VOLATILITY	-3.806	0.735	***	-2.819	0.782	***	-2.819	0.728	***
PRICE	-17.40	1.171	***	-4.796	1.195	***	-4.593	1.211	***
TRADE_SIZE_RATIO	20.91	4.924	***	20.83	6.949	***	20.25	6.919	***
TOTAL_VALUE	7.678	0.377	***	6.243	0.427	***	6.152	0.425	***
Intercept	12.87	5.777	**	17.64	7.856	**	18.96	7.222	***
Adj-R <sup>2</sup>	0.8488			0.8086			0.8087		

Table A10-4. HFT activity and the relationship between dark fragmentation and effective spreads

Table A10-4 reports the estimates for the two stage Heckman correction model as in Model 2 of Table 4. Additional variables representing trading activities of HFTs on NASDAQ are incorporated. Transactions for the 26 most active HFT firms on the NASDAQ market are identified.  $HFT_{MAKE}$  ( $HFT_{TAKE}$ ) is the ratio of the value of transactions in which an HFT provides (supplies) liquidity to the total trading value on NASDAQ.  $HFT_{ALL}$  is the ratio of the value of transactions in which an HFT provides or supplies liquidity to two times the total market trading value on NASDAQ.  $DARK\_VALUE\_RATIO$  is the proportion of total trading value on dark venues.  $DARK\_BLOCK\_RATIO$  is the daily value of the top 1% of trades on dark venues divided by the daily total value of dark trading.  $LAMBDA$  is the inverse Mills ratio obtained from the first stage probit model.  $PIN$  is estimated following the procedure in Easley et al. (1997) for each trading day.  $VOLATILITY$  is the standard deviation of one second midpoint returns in the 30 seconds after the transaction averaged over the trading day.  $Price$  is the log of the daily value-weighted average price.  $TRADE\_SIZE\_RATIO$  is the ratio of the average trade size on day  $t$  to the average trade size for the whole sample period for each sample stock.  $TOTAL\_VALUE$  is the log of the daily total trading value.  $LIT\_VALUE\_IMBALANCE$  is the absolute difference between the value of buy and sell transactions on lit markets, divided by the total trading value of all markets for the trading day.  $MCAP$  is the log of the stock's market capitalization on January 3 2011. All coefficients except for  $VOLATILITY$  and  $TRADE\_SIZE\_RATIO$  are enlarged by a factor of 10,000. Standard errors reported in are corrected for double clustering by date and stock (Thompson (2011)). \*\*\*, \*\* and \* indicate significance levels of 1%, 5% and 10%, respectively.

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

	Model 1			Model 2			Model 3		
	Coefficient	Std. Err.		Coefficient	Std. Err.		Coefficient	Std. Err.	
HFT <sub>ALL</sub>				-1.7205	0.6680	*			
HFT <sub>MAKE</sub>							0.7868	0.6737	
HFT <sub>TAKE</sub>							-1.9083	0.4777	***
DARK_VALUE_RATIO	1.1922	0.5282	*	1.1525	0.5159	*	1.1662	0.5237	*
DARK_BLOCK_RATIO	-0.9699	0.3137	**	-1.1332	0.3134	***	-1.0491	0.2922	***
LAMBDA	3.3250	1.5609	*	2.7188	1.6299		2.7502	1.5967	
PIN	1.5440	0.5124	**	1.4073	0.5042	**	1.1958	0.4935	*
VOLATILITY	1.5904	0.2673	***	1.6490	0.2700	***	1.6588	0.2681	***
PRICE	-1.0765	0.3049	***	-0.9698	0.3055	**	-0.7675	0.3461	*
TRADE_SIZE_RATIO	2.6381	0.6769	***	2.4759	0.6699	***	2.4129	0.6495	***
TOTAL_VALUE	-0.0973	0.0716		-0.0768	0.0664		-0.1508	0.0723	*
Intercept	3.0738	1.5235	*	3.4348	1.5850	*	3.9795	1.5369	**
Adj-R2	0.7762			0.7788			0.7826		

## Appendix 11: Comparison of market impact costs between block and non-block transactions on lit and dark markets

We compare market impact costs of block and non-block trades executed on lit and dark markets. Comparing between markets, we find that both block and non-block trades executed on lit markets have greater market impact costs than those executed on dark venues. Within individual markets, the results show that block trades are more informed than non-block trades on lit markets while the opposite holds for dark venues. Across all categories, block trades on dark venues have the lowest market impact costs. The results are statistically significant and consistent for both short-term and long-term measures of market impact costs.

Table A11. Comparison of Market Impact Costs

Table A11 compares market impact costs of block and non-block trades on lit and dark markets. For each stock, a block trade is defined as the largest 1% of trades for each stock over the sample period. Short-term (long-term) market impact costs are calculated as the bid-ask midpoint return over 30 seconds (5 minutes) after trade execution respectively, and the reported are the average short-term and long-term market impact costs across all sample transactions. \*\*\* indicates significance at the 0.1% level.

	Lit		Dark		Lit – Dark	
	Short-term (bps)	Long-term (bps)	Short-term (bps)	Long-term (bps)	Short-term (bps)	Long-term (bps)
Non-block trades	3.31	3.81	1.07	1.55	2.25 ***	2.25 ***
Block trades	4.47	5.05	0.71	0.85	3.76 ***	4.21 ***
Non-block - Block	-1.2 ***	-1.2 ***	0.36 ***	0.71 ***		

## Appendix 12: Alternative Instrumental Variable

We follow Roll, Schwartz, and Subrahmanyam (2009) and use one day lagged values of DARK\_VALUE\_RATIO as an alternative instrumental variable. The average autocorrelation of DARK\_VALUE\_RATIO is 0.44 (significant at the 0.001 level) over our sample period. Therefore, a day of high (low) dark market share is more likely to be followed by another day of high (low) dark market share. On the other hand, it is difficult to establish a reverse causality between lagged DARK\_VALUE\_RATIO and current transaction costs. Therefore, lagged DARK\_VALUE\_RATIO satisfies the requirement for an instrumental variable.

Table A12. Dark Fragmentation and Effective Spreads Estimated in a 2SLS Model with Lagged Value of DARK\_VALUE\_RATIO as the Instrumental Variable

Table A12 reports the results based on a 2SLS instrumental variable framework to examine the relationship between Dark fragmentation and effective spreads. Trade and quote data of 116 stocks listed on the NASDAQ and NYSE are examined over the period January 3 2011 to March 31 2011. We use the lagged value of DARK\_VALUE\_RATIO to instrument for DARK\_VALUE\_RATIO. The instrumented variable is then used in the second-stage regression reported here. The dependent variable is EFFSPREAD, which is the daily value weighted relative effective spread. DARK\_BLOCK\_RATIO is the daily value of the top 1% of trades on dark venues divided by the daily total value of dark trading. LAMDA is the inverse Mills ratio obtained from the first stage probit model. VOLATILITY is the standard deviation of one second midpoint returns in the 30 seconds after the transaction averaged over the trading day. PIN is the probability of informed trading calculated for each stock day. PRICE is the log of the daily value-weighted average price. TRADE\_SIZE\_RATIO is 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 is the log of the daily total trading value. All coefficients except for VOLATILITY and TRADE\_SIZE\_RATIO are enlarged by a factor of 10,000. Standard errors reported in are corrected for double clustering by date and stock (Thompson (2011)). \*\*\*, \*\* and \* indicate significance levels of 0.1%, 1% and 5% respectively.

*An Empirical Analysis of Market Segmentation on U.S. Equities Markets*

	Coefficient	Std. Err.	
DARK_BLOCK_SIZE_RATIO	2.3100	1.0095	*
DARK_BLOCK_RATIO	0.3236	0.4227	
LAMBDA	2.0906	0.6891	***
VOLATILITY	1.8406	0.2403	***
PIN	0.0778	0.0229	***
PRICE	-0.8535	0.1969	***
TRADE_SIZE_RATIO	6.9501	1.2784	***
TOTAL_VALUE	-0.1343	0.0512	**
Intercept	2.6577	1.3198	*
Adj-R <sup>2</sup>	0.6316		

The results presented in Table A12 show that lagged DARK\_VALUE\_RATIO remains positive and significant, providing further support for our main findings. DARK\_BLOCK\_RATIO is insignificant, which is consistent with our general conclusion that the execution of dark block trades does not harm market quality.

### Appendix 13: Policy Recommendations

In this paper we study the effects of the growing level of trading by market centers which are not full participants in the National Market System (NMS). These dark venues differ from lit markets in a number of dimensions most notably in whether they must provide fair-access and pre-trade transparency and restrict sub-penny trading increments. Our results have important implications for the integrity of the current US equities markets. We find that transaction costs increases and price efficiency decreases as the fraction of trading on dark venues increases. Our results also show that dark venue trading adversely affects investors trading on both lit and dark markets.

The increase in transaction costs and reduction in price efficiency can have long-term consequences on investor confidence. Among all investors, the negative effects are stronger for liquidity-motivated and long-term investors since these investors trade infrequently and hence, rely more on market efficiency. Traders pursue statistical arbitrage or market making strategies exploit short run market inefficiencies and consequently are less likely to be disadvantaged. In fact, short-term market inefficiencies and higher transaction costs may represent potential profit-generating opportunities for professional traders whose activities are generally viewed with skepticism by long term investors. Losing the confidence of long-term investors can have significant negative consequence including increases in cost of capital and destabilisation of financial markets.

Recent studies also find that trading on dark venues reduces market liquidity (Sarkar, Schwartz, and Klagge (2008), and Degryse, De Jong, and Van Kervel (2015)), which support our finding that dark trading is associated with lower price efficiency. The increase in excessive short-term



volatility may lead to a more profound consequence during volatile periods when liquidity is most needed. Since there is no time priority currently across different market centers, rational liquidity providers will refrain from trading while they wait for the volatility to settle without losing the opportunity to trade. The reduced liquidity will in term reduce price efficiency further, causing a cascading effect (see Pagano (1989)). The observation of reduction in dark market trading during the 2010 Flash Crash and market volatility of August 2011 is consistent with this analysis.

Overall, our results support the theoretical predictions of Bolton, Santos and Scheinkman (2015). Specifically, trading in dark venues in the current form is detrimental to the market's central role of allocating efficiently scarce resources in the economy. Accordingly, we recommend the following policy changes.

1. There is an urgency to reduce the negative impact of dark venues.

From the previous analysis, the detrimental impact of dark trading is mainly driven by their ability to segment market participants through their exemptions from certain provisions of Regulation NMS. Therefore, it is important to allow equal market access, eliminate existing discrepancies in the minimum price increment, and reward for price discovery.

2. Expand the fair access requirements

Equal market access on dark venues can not only prevent order flow segmentation but also lead to a general improvement of investor welfare. Results in Table 2 show that about 70% of transaction costs on dark venues are charged by liquidity providers as market-making profits; while this figure for lit markets is about 21% on average. There is a significant lack of competition in the supply of

liquidity on dark venues. Unequal market access affects the entry of all willing liquidity providers, which intensifies liquidity competition and hence reduces total transaction costs for investors on dark venues.

Currently, alternative trading systems that execute less than 5% of trading volume in a lit security are exempted from several requirements of lit market centers. For example, these venues are not required to disclose their best priced orders for inclusion in the public quotation data (Rule 301(b)(3)(ii)) and are exempted from fair access regulations (Rule 301(b)(5)(ii)(B)). While no individual market center exceeds the 5% threshold, dark venues as a group represent over 25% of the total market volume on average. We suggest a reduction in the 5% threshold in order to prevent the collectively significant negative impact from dark venues.

### 3. Adoption of a tick size table to allow for sub-penny pricing

The sub-penny rule affects a minimum tick size of \$0.01 for stocks priced equal to or greater than \$1.00. As documented in Internet Appendix 12, operating as broker-dealers dark venues are able to offer sub-penny executions, which allow them to attract uninformed orders by offering minimal price improvement. We propose a harmonization in minimum tick size, which is particularly important for tick constrained stocks; we find sub-penny pricing is most beneficial to dark venues when stocks are trading at low quoted spreads. Consistent with this proposal, Multilateral Trading Facilities (MTFs) and exchanges under MiFID agreed to a harmonized tick size regime for the

most liquid stocks on European markets in June 2009.<sup>17</sup> We believe a similar tick size regime in the U.S. equities markets will increase the market efficiency by enhancing the competitiveness of lit markets.<sup>18</sup>

#### 4. Implementation of a “trade-at” rule

Our analysis indicates that there are significant costs to providing liquidity in lit markets and benefits to providing liquidity on dark venues. Since lit markets set the benchmark used for trading on both lit and dark markets more market centers should be encouraged to display liquidity and operate in a way which gives priority to displayed liquidity. Recently the SEC proposed a “trade-at” rule.<sup>19</sup> Orders received by trading centers not displaying the NBBO at the time the order is received must be executed with *significant* price improvement or routed to a market center displaying the NBBO.

---

<sup>17</sup> See “European exchanges introduce harmonized tick size regimes in Europe”, FESE. Available from: [http://www.fese.be/\\_mdb/pressdocs/European%20exchanges%20introduce%20harmonised%20tick%20size%20regimes%20in%20Europe.pdf](http://www.fese.be/_mdb/pressdocs/European%20exchanges%20introduce%20harmonised%20tick%20size%20regimes%20in%20Europe.pdf), accessed on October 23 2011. For the remaining stocks, exchanges choose one of four tick size tables with up to 17 tick size bands based on stock price. MTFs apply the tick size table adopted by the listing venue for the underlying shares.

<sup>18</sup> In his response to the Pan Europe Tick Size Harmonization Plans of the FESE, Tamas Madlena, CEO of Quote MTF, commented that “the commoditization of tick sizes across the European equities market is to the benefit of all participants and investors by encouraging efficiency, consistency and clarity.” (Quote MTF, Press Release on 28 September, 2010).

<sup>19</sup> SEC Concept Release No. 34-61358, p. 70.

There are two elements to the proposal. The first element is based on our demonstrating that price improvement allows dark venues to cream-skim uninformed order flow at minimal cost; in Appendix 11 we document a large number of transactions receiving minimal price improvement (i.e.,  $< 0.10$  cent). By requiring significant price improvement, the trade-at rule increases the cost of attracting orders and places lit markets in a better position to compete for orders.

The second element rewards an order displaying at the NBBO. While lit markets provide priority to displayed order over non-displayed orders there is currently no display priority across market centers. As a result, an order can be executed at the NBBO price on the dark venues ahead of orders on the lit markets even when the lit orders are displayed. Pre-trade opacity in dark venues allows traders to free ride on prices discovered on lit markets. Since liquidity provision is costly due to adverse selection risk, a display priority in place is important to encourage liquidity providers to compete for execution and therefore, enhance market liquidity and price discovery.

These policy recommendations are essential to improving the competitive landscape of U.S. equity markets. Without these policies in place, it is difficult for lit exchanges to compete with the dark without compromising, to some extent, their public obligations. For example, the NYSE recently proposed a Retail Liquidity Program where potential price improvement is given to orders deemed to originate from retail investors. This may contradict the requirement for a national security exchange as “the rules of the exchange ... are not designed to permit unfair discrimination between customers, issuers, brokers, or dealers ...” (Securities Exchange Act of 1934, Section 6(b)(5)).<sup>20</sup>

---

<sup>20</sup> In its proposal, the NYSE particularly defended its position. See SEC Release No. 34-65672.

In 1975, Congress established the framework for a National Market System to connect together a growing number of different markets. Our policy recommendations promote the interaction of informed and uninformed order flow among these markets in working towards ‘efficient, competitive, fair and orderly markets’ (SEC, Reg NMS, Release No. 34-51808). Today such market centers are especially important for the restoration of confidence in the financial markets.

## References

- Battalio, R.; B. Hatch; and R. Jennings. "All Else Equal? A Multidimensional Analysis of Retail, Market Order Execution Quality." *Journal of Financial Markets*, 6 (2003), 143-162.
- Bessembinder, H. "Issues in Assessing Trade Execution Costs." *Journal of Financial Markets*, 6 (2003a), 233-257.
- Bessembinder, H. "Selection Biases and Cross-Market Trading Cost Comparisons." Working Paper, University of Utah (2003b).
- Bessembinder, H., and H. M. Kaufman. "A Comparison of Trade Execution Costs for NYSE and Nasdaq-Listed Stocks." *Journal of Financial and Quantitative Analysis*, 32 (1997), 287-310.
- Bolton, P.; T. Santos; and J. A. Scheinkman. "Cream-Skimming in Financial Markets." *The Journal of Finance*, (Forthcoming).
- Buti, S.; F. Consonni; B. Rindi; Y. Wen; and I. M. Werner. "Sub-Penny and Queue-Jumping." *Charles A. Dice Center Working Paper No. 2013-18* (2014).
- Buti, S.; B. Rindi; and I. M. Werner. "Dark Pool Trading Strategies." In *Innocenzo Gasparini Institute for Economic Research*: Bocconi University (2011).

Chakravarty, S.; P. Jain; J. Upson; and R. Wood. "Clean Sweep: Informed Trading through Intermarket Sweep Orders." *Journal of Financial and Quantitative Analysis*, 47 (2012), 415-435.

Chordia, T., and A. Subrahmanyam. "Market Making, the Tick Size, and Payment-for-Order Flow: Theory and Evidence." *The Journal of Business*, 68 (1995), 543-575.

Chung, K. H.; C. Chuwonganant; and D. T. McCormick. "Order Preferencing and Market Quality on Nasdaq before and after Decimalization." *Journal of Financial Economics*, 71 (2004), 581-612.

Degryse, H.; F. de Jong; and V. van Kervel. "The Impact of Dark Trading and Visible Fragmentation on Market Quality." *Review of Finance*, 19 (2015), 1587-1622.

Easley, D.; N. M. Kiefer; and M. O'Hara. "Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow." *The Journal of Finance*, 51 (1996), 811-833.

Easley, D.; M. M. López de Prado; and M. O'Hara. "Flow Toxicity and Liquidity in a High-Frequency World." *Review of Financial Studies*, 25 (2012), 1457—1493.

Ellis, K.; R. Michaely; and M. O'Hara. "The Accuracy of Trade Classification Rules: Evidence from Nasdaq." *Journal of Financial and Quantitative Analysis*, 35 (2000), 529-551.

Gormley, T. A., and D. A. Matsa. "Common Errors: How to (and Not to) Control for Unobserved Heterogeneity." *Review of Financial Studies*, 27 (2014), 617-661.

Hagströmer, B., and L. Nordén. "The Diversity of High-Frequency Traders." *Journal of Financial Markets*, 16 (2013), 741-770.

Hasbrouck, J. "One Security, Many Markets: Determining the Contributions to Price Discovery." *The Journal of Finance*, 50 (1995), 1175—1199.

Heckman, J. J. "Sample Selection Bias as a Specification Error." *Econometrica*, 47 (1979), 153—161.

Hendershott, T., and R. Riordan. "Algorithmic Trading and the Market for Liquidity." *Journal of Financial and Quantitative Analysis*, 48 (2013), 1001-1024.

Huang, R., and H. Stoll. "The Components of the Bid-Ask Spread: A General Approach." *Review of Financial Studies*, 10 (1997), 995-1034.

Lee, C. M. C., and M. J. Ready. "Inferring Trade Direction from Intraday." *The Journal of Finance*, 46 (1991), 733-746.

Lin, J.-C.; G. Sanger; and G. Booth. "Trade Size and Components of the Bid-Ask Spread." *Review of Financial Studies*, 8 (1995), 1153-1183.



Madhavan, A., and M. Cheng. "In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets." *Review of Financial Studies*, 10 (1997), 175—203.

McInish, T. H., and J. Upson. "The Quote Exception Rule: Giving High Frequency Traders an Unintended Advantage." *Financial Management*, 42 (2013), 481-501.

Mittal, H. "Are You Playing in a Toxic Dark Pool? A Guide to Preventing Information Leakage." *Journal of Trading*, 3 (2008), 20-33.

O'Hara, M., and M. Ye. "Is Market Fragmentation Harming Market Quality?" *Journal of Financial Economics*, 100 (2011), 459-474.

Pagano, M. "Endogenous Market Thinness and Stock Price Volatility." *The Review of Economic Studies*, 56 (1989), 269-287.

Petersen, M. A. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies*, 22 (2009), 435-480.

Roll, R.; E. Schwartz; and A. Subrahmanyam. "Options Trading Activity and Firm Valuation." *Journal of Financial Economics*, 94 (2009), 345-360.

Sarkar, A.; R. Schwartz; and N. Klagge. "Liquidity Begets Liquidity Implications for a Dark Pool Environment." *Institutional Investor's Guide to Global Liquidity*, 1 (2009), 15-20.

Seppi, D. J. "Equilibrium Block Trading and Asymmetric Information." *The Journal of Finance*, 45 (1990), 73-94.

Thompson, S. B. "Simple Formulas for Standard Errors That Cluster by Both Firm and Time." *Journal of Financial Economics*, 99 (2011), 1-10.

Weaver, D. G. "The Trade-at Rule, Internalization, and Market Quality." Working Paper, Rutgers Business School (2014).