Are Market Makers Incentivized to Provide Liquidity? Evidence from the Nasdaq

Version: February 23, 2017

Abstract

Since 2002, Nasdaq has given market participants the discretion over whether to submit limit orders anonymously, or submit their orders non-anonymously along with their market participant identification number (MPID). The vast majority of non-anonymous orders are submitted by market makers, who are mandated by the exchange to submit a certain number of non-anonymous orders per day. However, while market makers are expected to fulfill this quota, the choice of when and under what circumstances to reveal their identities is still at their discretion. Therefore, in this project, we explore the determinants of MPID revelation among market makers on the Nasdaq, and the implications that these non-anonymous order submissions have for market quality.

1 Introduction

While equity market makers were traditionally agents appointed by exchanges to maintain orderly markets and stand ready to provide liquidity, recent years have seen the rise of so-called "endogenous liquidity providers" (ELPs), who are not appointed by exchanges but arise as liquidity providers on their own accounts.[1] The current reliance of many modern U.S. equity trading platforms on ELPs remains somewhat controversial, as exchanges have scrambled to implement programs to properly incentivize and obligate this new type of market maker to help maintain orderly markets.[2] Some questions remain as to whether exchanges have indeed been successful at this task. In a speech before the Economics Club of New York, Securities and Exchange Commission (SEC) chairman Mary L. Shapiro expressed concern "whether the firms that effectively act as market makers during

[1]See, e.g., Bessembinder et al. (2015); Anand and Venkataraman (2016)

[2]For example, the Qualified Market Maker Program on Nasdaq and the Designated Market Market Program on the NYSE.

normal times should have any obligation to support the market in reasonable ways in tough times."[3] In a report written in response to the May 2010 Flash Crash, a joint committee of the SEC and Commodity Futures Trading Commission (CFTC) emphasizes that "...there remain legitimate concerns over the absence of present incentives for market participants to provide liquidity in the present market structure."[4] Recent literature has focused much attention on "market fragility": Without the presence of exchange-mandated obligations, ELPs tend to withdraw liquidity in unison once conditions become suboptimal and liquidity would be most needed such as in times of price uncertainty and order imbalances (see, e.g., Bessembinder et al., 2015; Anand and Venkataraman, 2016).

In addition, another characteristic of modern equity markets that impacts the role of the market maker is that of increasing anonymity of trading. Raman et al. (2014) argue that an increase in market fragility is a direct result of the anonymity of modern electronic markets. In traditional floor-based markets, market makers could use reputation and long-term relationships to identify the information content of trades, and credibly discipline those who trade on private information (Battalio et al., 2007). However, in anonymous markets, market makers have no way of knowing the information levels of traders, which thus increases their exposure to adverse selection and incentivizes the mass withdraw of liquidity in uncertain market conditions. Furthermore, market makers may lack the incentives to reveal their own presence given a choice to trade anonymously, as non-anonymity exposes them to adverse selection in the form of pick-off risk, or serves simply to expose their own strategies and inventories.

Since the introduction of SuperMontage in 2002, Nasdaq has given market participants the discretion over whether to submit limit orders anonymously, or submit their orders non-anonymously along with their market participant identification number (MPID) (Mizrach, 2006). At the same time, the rules governing liquidity provision on Nasdaq require that registered market makers submit a certain amount of non-anonymous, MPID-attributed quotes as part of their role as liquidity providers. Specifically, market makers are expected to stand ready and willing to buy and sell securities in which they are registered as a market maker, and to maintain continuous two-sided quotations that are attributable to the market makers by their MPID.[5] Nasdaq's Qualified Market Maker (QMM) program broadens these obligations somewhat, stipulating that a QMM must, on a daily basis, post displayed quotes during at least 25% of trading hours in at least 1,000 securities.[6] In return, market makers receive additional rebates on liquidity provision, along with other privileges. However, given their fulfillment of this quota of MPID-attributed orders, the choice of when and under what circumstances to reveal their presence is still at the discretion of the individual market maker. This has put modern market makers at a crossroads in terms of fulfilling their

[3]See "Strengthening Our Equity Market Structure", 7 September 2010, available at https://www.sec.gov/news/speech/2010/spch090710mls.htm.

[4]See "Recommendations Regarding Regarding Regulatory Responses to the Market Events of May 6, 2010", 18 Februrary 2011, available at http://www.cftc.gov/idc/groups/public/@aboutcftc/documents/file/jacreport_021811.pdf.

[5]Nasdaq Rule 4613(a)(1): "...the member shall be willing to buy and sell such security for its own account on a continuous basis and shall enter and maintain a two-sided quotation ('Principal Quote'), which is attributed to the market maker by a special maker participant identifier ('MPID')...". See http://Nasdaq.cchwallstreet.com/Nasdaq/pdf/new_listing_rules.pdf.

[6]See http://Nasdaq.cchwallstreet.com/Nasdaq/main/Nasdaq-equityrules/chp_1_1/chp_1_1_4/chp_1_1_4_6/chp_1_1_4_6_5/chp_1_1_4_6_5_4/default.asp.

obligations as market makers, against avoiding adverse selection and exposure in highly anonymous markets.

Therefore, this paper explores the nature of the market maker incentive program's required exposure of market maker MPIDs, and to what extent this rule may in turn constrain market makers more than incentivizing them to adequately stabilize and supply liquidity to uncertain markets. Accordingly, we explores the determinants of non-anonymous intervention by market makers on the Nasdaq, and the implications that this may have for market quality. Specifically, this paper asks: How does the market react to non-anonymous orders posted by market makers? In other words, are MPID-attributed orders by market makers seen by the market as containing information about fundamental value or future order flows? Secondly, are MPID-attributed orders submitted by market makers in line with their obligations to provide liquidity and stabilize markets? In order to address these questions, we use a unique dataset that reconstructs the limit order book for several Nasdaq-traded stocks. Our data contains information on whether or not a limit order is submitted anonymously, and, given non-anonymity, the market participant identifier number (MPID) of the submitter. This allows us to identify the non-anonymous orders submitted by market makers. In answering these questions, we aim to contribute insights into whether liquidity provision incentive schemes put into place by modern exchanges indeed properly incentivized market makers maintain orderly markets and provide liquidity when it is needed.

Our paper is most closely related to that of Comerton-Forde et al. (2011), which examines the determinants and impact of anonymous trading on the Toronto Stock Exchange, using trading book data from 1 May to 31 July 2004. These authors find that traders use anonymity when submitting informed orders, and that anonymity can reduce the executions costs of orders that are large and aggressive. However, it should be noted that the properties of their dataset hint that TSX at this period of time had a very different trading landscape than that of the present study. While 2.86% of orders in our sample are submitted non-anonymously, in the dataset of Comerton-Forde et al. (2011) about 94% are submitted non-anonymously. Furthermore, TSX stocks each have a designated market maker firm who is compensated directly by the issuer for maintaining an orderly market in that stock; such contracts are currently illegal in the U.S. under FINRA Rule 5250.[7] In addition, our paper is related to two working papers. Both Benhami (2006) and Karam (2012) examine the value of the choice of (non-) anonymity on Nasdaq traders; however, both papers use datasets that pre-date the implementation Regulation National Market System (Reg NMS), after which most trading shifted onto ECNs that more easily facilitate anonymous trading.

Our results are consistent with the hypothesis that the market significantly reacts to the information contained in MPID-attributed submissions by market makers. Specifically, MPID-attributed orders tend to be followed by price movements in the same direction as the order (e.g., price increases following MPID-attributed buy orders), and an increase in effective bid-ask spreads. Furthermore, consistent with the idea that market participants avoid trading against an informed order and instead act to trade in the same direction as the order with immediacy, MPID-attributed orders

[7] See FINRA Rule 5250, Payments for Market Making: "No member or person associated with a member shall accept any payment or other consideration, directly or indirectly, from an issuer of a security, or any affiliate or promoter thereof, for publishing a quotation, acting as market maker in a security, or submitting an application in connection therewith." Available at http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=8626.

tend to be followed by a decrease in market order submissions on the same side of the book, and an increase in market orders on the opposite side of the book as the submitted order. Interestingly, the relative tick size appears to play a role in these results, as small tick sizes may effect the informational advantage that the market makers have with respect to competing informed traders.

In terms of the determinants of market maker MPID attribution, our results show that market makers tend to internalize this expected market response to their orders, and choose to submit MPID-attributed orders when they expect the impacts of their orders to be minimized, i.e., when markets are relatively calm. Market makers tend to submit MPID-attributed orders when quoted spreads are low, and are less likely to submit MPID-attributed orders following price changes (both negative and positive). They are also less likely to submit MPID-attributed following increases in opposite-side execution volume; in other words, they avoid contributing additional signals to information flows. Furthermore, MPID-attributed orders also tend to be smaller and less aggressive, to avoid further signals of an informed order. It does not appear that market makers use their informational advantage in order to try to bluff the market, but there is some evidence that market makers are able to use MPID-attributed orders to respond to some adverse market conditions, such as when depth is low. All-in-all, our results show that the mandate for market makers to submit MPID-attributed orders may be counter-productive, as it constrains market makers from using their full capacity in order to stabilize markets, and towards simply using their orders within a context of "damage control" to minimize the de-stabilizing impacts of their orders.

The remainder of this paper is organized as follows. Section 2 describes the economic foundations to our paper and lays out our hypotheses regarding the determinants of and market reactions to non-anonymous market maker intervention. Details on our unique dataset, along with descriptions of the MPID data, are provided Section 3. Section 5 describes the methodology and empirical results for our question regarding market reactions to market makers' non-anonymous interventions, while Section 4 does the same addressing our question on the determinants of market maker intervention. Section 6 explores the robustness of our results. Finally, Section 7 concludes.

2 Economic Foundations

The first aim of our paper is to address the question: Do market makers submit non-anonymous orders according to their obligations to provide liquidity and stabilize markets? In his seminal paper, Demsetz (1968) identifies the non-synchronous arrival of buyers and sellers into a market as one of the key frictions in financial markets. However, he proposes that this problem can be alleviated by market makers who bridge the gaps between these non-synchronous arrivals. In subsequent papers, Garbade and Silber (1979) and Grossman and Miller (1988) show that market makers help to mitigate order imbalances and lower execution risk for other market participants. Therefore, the main role of market makers should be to serve as reliable counterparties to traders when no others are available. A lack of reliable trading partners is likely to occur, for example, when uncertainty is high and market participation rates subsequently drop (see, e.g. Cao et al., 2005). We should therefore see a higher participation rate of market makers (relative to other traders) in times when market participation is low and market uncertainty is high.

Guided by this intuition, if market makers are indeed properly incentivized and/or obligated to act as market makers by the incentive programs on Nasdaq, then submitted quotes that are attributed to a market maker MPID should reflect this role. First, we would expect to see a higher rate of MPID-attributed quotes when market participation rates, i.e., the number of executions and submissions of orders, are low. We might also see a higher rate of MPID-attributed quotes when depth on the books is low, as this represents a low supply of liquidity. Furthermore, market participation has been shown to drop during market downturns (see, e.g., Næs et al., 2011); therefore, we might also expect to see a higher rate of MPID-attributed quotes following large, negative movements in prices.

Secondly, we would expect to see a higher rate of MPID-attributed quotes when uncertainty is high. For our purposes, uncertainty refers to an occasion in which market efficiency fails, and asset prices deviate significantly from fundamental underlying values, and is typically measured using asset volatility (see, e.g., Chung and Chuwonganant, 2014). Furthermore, higher uncertainty is also associated with higher illiquidity as measured by the relative bid-ask spread, as (1) illiquidity prevents information from being quickly incorporated into prices (see, e.g., Chordia et al., 2008), and (2) uncertainty about prices exposes traders to higher adverse selection as in the classic "lemons problem" of Akerlof (1995), and higher adverse selection in turn leads to higher spreads (see, e.g., Kyle, 1985; Huang and Stoll, 1997). Therefore, we should see a higher rate of MPID-attributed quotes when illiquidity is high.

Hypothesis 1: If MPID-attributed orders are submitted by market markets according to incentives to encourage liquidity and stabilize markets, then we should see an increase in MPID submissions in response to higher market uncertainty, such as higher bid-ask spreads and volatility, and lower market participation.

The second aim of our paper is to answer the question: How does the market react to non-anonymous orders posted by market makers? By submitting an MPID-attributed order, a market maker essentially reveals two important signals, to which the market can react. The first signal can be tied to the nature of the order as non-anonymous. Numerous papers on trader anonymity have shown that informed traders prefer to submit their orders anonymously, to prevent other traders from discovering their information or trading strategies (see, e.g., Grammig et al., 2001; Barclay et al., 2003; Comerton-Forde et al., 2011). Therefore, the act of MPID revelation itself may signal to the market that the trade contains a lack of information. This may serve to lower adverse selection (and thus lower relative bid-ask spreads). A reduction in adverse selection may encourage participation from risk averse traders. Following Harris (1997), orders that are more transparent encourage participation from defensive, "reactive" traders, who wait to be presented with valuable trading opportunities rather than initiating a trade and potentially exposing themselves to informed trades. This may encourage market participation in the form of higher submission and execution volumes, and higher depth.

The second signal stems from the fact that, since the MPID reveals the full identity of the submitter, MPID attribution also reveals that the order is submitted by a market maker. Given regulatory concern over low investor confidence in liquidity providers since the 2010 Flash Crash, the revealed presence of a reliable liquidity provider may serve to boost investor confidence and stabilize uncertain markets (see, e.g., Watanabe, 2015; Anand and Venkataraman, 2016). As a result, we might see lower volatility in response to an MPID-attributed order. Same-side order flow may increase, as an MPID-attributed order is viewed as "anchoring" the price on the side of the book (buy or sell) to which it is submitted.

Hypothesis 2: If MPID-attributed orders are submitted by market markets according to incentives to encourage liquidity and stabilize markets, then in response to higher rates of MPID submission we should see an improvement in market quality, such as lower bid-ask spreads and volatility, and an improvement in execution quality.

On the other hand, the literature shows a range of alternative motivations that can drive the determinants of and reactions to market maker order submissions, particularly within the context of market transparency. Simaan et al. (2003) argue that non-anonymity allows market makers to form a regime that enforces an implicit collusion to keep spreads wide, in which defecting market makers can be identified and punished. Furthermore, market makers are occasionally treated in the literature as informed traders, as they can receive information from either their associated clients (Saporta, 1997; Naik et al., 1999) or make inferences from their knowledge of order flow, particularly in the the case of high-frequency market makers (Vayanos, 2001; van Kervel and Menkveld, 2016; Malinova and Park, 2016).[8] As argued by Karam (2012), the fact that market makers are constrained to provide MPID-attributed quotes increases the likelihood that even their non-anonymous quotes will contain some degree of their information. Furthermore, Reiss and Werner (2005) show that, even when an anonymous market is available, market makers prefer to submit their informed trades non-anonymously, in order to avoid a downward spiral in which too many anonymous informed trades increase adverse selection and crowd out liquidity. Lastly, market makers may also use their MPID-attributed orders in attempts to "bluff" the market. Foucault et al. (2013) and Karam (2012) argue that, when market makers are aware that their orders will signal information, they will post non-anonymous quotes at the best bid or ask when spreads are large.

Overall, if market makers are indeed properly incentivized and/or obligated to use their MPID-attributed orders to encourage market participation and to stabilize markets, then these incentives should be reflected in the MPID's strategic choice of when to submit a non-anonymous order, as well as reflected in the market's reaction to an MPID-attributed order. Deviations from Hypothesis 1 and Hypothesis 2 should reflect an alternative motivation for the market maker's submission of MPID-attributed orders, and thus a deviation from the principal goals of exchange-mandated market maker incentive programs.

[8] Other prominent examples include Boulatov and George (2013) and Kaniel and Liu (2006), who examine the endogenous choice of informed traders to submit limit or market orders. Kaniel and Liu (2006) shows that informed traders with long-lived information tend to prefer limit orders, and that price impacts following limit orders tend to be higher.

3 Data

Data is obtained from LOBSTER[9] Academic Data, an online data tool that reconstructs the limit order book for the universe of Nasdaq stocks using the Nasdaq TotalView-ITCH direct data feed. The dataset includes order book data on prevailing bid and ask quotes and depths at up to 200 price levels, as well as message files that contain updates to the limit order book. This includes information on the type of event (submissions, partial or total cancellations, and executions of visible or hidden orders), the number of shares, price, direction (buy or sell), and time stamp (to the nanosecond) of the order that the event concerns, as well as a unique order reference number that allows us to track the submission and eventual execution or cancellation of the order. In addition, our data sample uniquely contains information on the Market Participant Identification Number (MPID). This will be described in more detail in Section 3.1.

The main sample in this analysis is composed of eight Nasdaq-listed firms, mostly in the high-tech industry. These firms include: Apple, Inc. (AAPL); Cisco Systems, Inc. (CSCO); eBay, Inc. (EBAY); Facebook, Inc. (FB); Google, Inc. (GOOG); Intel Corporation (INTC); Microsoft Corporation (MSFT), and Yahoo! Inc. (YHOO). The sample time period includes 14 trading days in November 2013, from 4 November to 22 November 2013.[10] Summary statistics for the stock prices, returns, and order flow for these firms is presented in Table 1. From the summary statistics we can see, first, a wide dispersion in share prices between the sample stocks. The two highest-priced stocks (GOOG, at $1025.43; and AAPL, at $521.25) are particularly distinct from the other stocks, as they have stocks prices that are an order of magnitude higher than the next highest-priced stock (EBAY, at $51.94). Furthermore, these stocks also differ widely in terms of order flow. While GOOG has just over 5,000 average daily trades, FB has nearly eleven times as mainly, at nearly 55,000 average daily trades. Lastly, five out of the eight stocks experience, on average, negative daily returns, implying that this time period may be in one in which the market is in particular need of stable liquidity provision by market makers. Also reported for each firms are the percentages of total order submissions that are attributed to an MPID. MPID-attribution rates range from a low of 1.61% for MSFT, to a high of 6.8% for FB.

3.1 Market Maker Identifier Data

As previously mentioned, LOBSTER Academic Data records data directly from the Nasdaq TotalView-ITCH direct data feed. When a market participant submits a limit order to Nasdaq, the information on a number of characteristics of the order, including the order reference number, limit price and size, will be recorded as a submission message. If the market participants has chosen to display their identity to other market participants, than the message will additionally contain their Market Participant Identification (MPID), which is uniquely assigned to each market participant registered on the exchange.

[9]Limit Order Book System – The Efficient Reconstructor; see https://lobsterdata.com/.
[10]It was necessary to exclude 11 November 2013 due to corrupt data files.

Table 1: Sample Stocks Descriptive Statistics

	Mean	Median	Std. Dev.	Min.	Max.
AAPL					
Average Stock Price (USD)	521.25	520.43	3.66	512.38	529.27
Daily Return (bp)	-0.10	-0.21	1.00	-1.61	1.57
Number of Daily Orders	138,681	137,088	23,008	96,796	186,736
Number of Daily Trades	16,368	15,555	2,500	13,614	21,515
% MPID	1.96%				
CSCO					
Average Stock Price (USD)	22.11	21.47	1.03	20.77	24.00
Intradaily Return Volatility (bp)	0.0022	0.0015	0.0028	0.0013	0.0121
Daily Return (bp)	-0.34	0.70	3.31	-10.84	2.13
Number of Daily Orders	249,443	244,267	90,655	152,238	536,612
Number of Daily Trades	22,544	19,668	13,670	12,247	67,736
% MPID	2.14%				
EBAY					
Average Stock Price (USD)	51.94	52.15	1.07	49.87	53.85
Daily Return (bp)	-0.15	-0.30	1.68	-3.34	4.31
Number of Daily Orders	310,645	286,657	113,001	164,888	597,135
Number of Daily Trades	20,620	17,564	7,797	13,560	40,018
% MPID	3.84%				
FB					
Average Stock Price (USD)	47.89	47.89	1.13	45.73	50.45
Daily Return (bp)	-0.29	0.04	2.84	-6.49	4.51
Number of Daily Orders	638,182	649,744	162,879	369,426	852,392
Number of Daily Trades	54,356	59,540	13,418	32,768	78,343
% MPID	6.80%				
GOOG					
Average Stock Price (USD)	1025.43	1025.80	9.58	1005.00	1048.80
Daily Return (bp)	0.05	-0.15	0.88	-1.48	2.05
Number of Daily Orders	75,042	76,744	19,323	49,241	114,180
Number of Daily Trades	5,187	4,710	1,395	3,531	7,935
% MPID	1.81%				
INTC					
Average Stock Price (USD)	24.36	24.34	0.33	23.77	25.28
Daily Return (bp)	-0.11	0.29	1.88	-5.35	2.69
Number of Daily Orders	190,832	172,346	49,940	138,646	326,237
Number of Daily Trades	14,174	12,548	5,644	9,398	31,298
% MPID	1.89%				
MSFT					
Average Stock Price (USD)	37.34	37.45	0.55	35.55	38.22
Daily Return (bp)	0.35	0.43	1.71	-1.77	4.18
Number of Daily Orders	444,870	426,755	105,221	269,153	640,988
Number of Daily Trades	29,026	27,207	10,764	16,448	57,101
% MPID	1.61%				
YHOO					
Average Stock Price (USD)	34.56	34.85	1.32	32.07	36.66
Daily Return (bp)	0.75	0.52	1.94	-2.40	3.15
Number of Daily Orders	378,733	358,077	124,213	208,150	552,838
Number of Daily Trades	21,726	22,149	5,121	15,623	34,732
% MPID	2.85%				

This table shows the average stock price, daily return (using closing transaction prices), daily number of orders (i.e., submitted limit orders), and the number of daily trades (i.e., executed orders or submitted market orders), for eight Nasdaq-traded stocks for 4-22 November 2013. Reported are the mean, median, standard deviation, minimum, and maximum of these variables for each firm. Also reported for each firm are the percentages of order submissions that are attributed to an MPID.

Table 2: Market Participant Identifiers, Types and Relative Submission Contribution

(1) MPID	(2) Firm Name	(3) MPID Type	(4) %Sub.
ATDF	Automated Trading Desk Financial Services, LLC	Market Maker	0.05%
BARD	Robert W. Baird & Co. Incorporated	Market Maker	< 0.00%
DADA	D.A. Davidson & Co.	Market Maker	0.00%
FBCO	Credit Suisse Securities (USA) LLC	Market Maker	0.28%
GSCO	Goldman, Sachs & Co.	Market Maker	1.55%
RHCO	Suntrust Robinson Humphrey, Inc.	Market Maker	< 0.00%
SBSH	Citigroup Global Markets Inc.	Market Maker	17.21%
TMBR	Timber Hill LLC	Market Maker	76.36%
UBSS	UBS Securities LLC	Market Maker	4.44%
WCHV	Wells Fargo Securities, LLC.	Market Maker	< 0.00%
		Total Market Maker	**99.89%**
BOOK	Bloomberg Tradebook LLC	ECN	0.03%
LEHM	Barclays Capital Inc./Le	Nasdaq Participant	0.01%
NITE	Knight Capital Americas LLC	Nasdaq Participant	0.03%
WEMM	Wells Fargo Securities, LLC.	Nasdaq Participant	0.04%
		Total Other	**0.11%**

This table shows the list of Nasdaq market participant identifiers (MPIDs) identified from a sample of eight Nasdaq-traded stocks for 4-22 November 2013. Reported are the MPIDs, the firm name of the market participant, the MPID type, and the percentage of total MPID-attributed submission volume contributed by that particular MPID. The market participants are split according to whether or not Nasdaq registers the market participant as a market maker or not; see ftp://ftp.Nasdaqtrader.com/symboldirectory/mpidlist.txt.

Along with the firm name of the market participant and other identifying information, Nasdaq additionally provides information on the market participant type (i.e., Market Maker, Order Entry Firm, Electronic Crossing Network (ECN), general Nasdaq Market Participant, etc.).[11] This allows us to identify which market participants are registered as market makers. Table 2 shows the list of the Nasdaq MPIDs that are revealed within our sample, along with the corresponding market participant type. Also reported in Column (4) is the relative contribution of each MPID to the total MPID-attributed submission volume. The table confirms that a vast majority (99.89%) of MPID revelation is done by firms that identify as market makers. To remain consistent with our hypotheses on the behavior of market makers, those MPIDs which are not attributed to the market maker MPID type are removed.[12]

Table 3 decomposes the contribution of each market participant to the submission volume of individual stocks. From the table, it becomes clear that only three market makers maintain a presence in all stocks: Citigroup Global Markets LLC (SBSH), Timber Hill LLC (TMBR), and UBS Securities LLC (UBSS). Also reported are the Herfindahl-Hirschman Index (HHI) scores for each stock, calculated as the sum of squares of the market shares (i.e., relative contribution to MPID-attributed submission volume) of each market making firm active within a stock. This measure should capture

[11] For more information on MPID types, see http://www.Nasdaqtrader.com/trader.aspx?id=symboldirdefs and ftp://ftp.Nasdaqtrader.com/symboldirectory/mpidlist.txt.

[12] Note that the MPID "WEMM", formerly belonging to Wells Fargo, Inc., has since been de-listed as an MPID, and thus its MPID type cannot be verified. Therefore, this (relatively small) sample is also removed.

the degree of competition between liquidity providers in each stock, and ranges from a low of 3973 for GOOG (most competitive), to a high of 8943 for EBAY (least competitive).

Table 3: Relative Contribution of Market Participants to Submission Volume, By Stock

	(1) AAPL	(2) CSCO	(3) EBAY	(4) FB	(5) GOOG	(6) INTC	(7) MSFT	(8) YHOO
ATDF	0.21%	0.24%	0.02%	0.02%	0.16%	0.15%	0.07%	0.04%
BARD	–	–	< 0.01%	–	–	–	0.01%	–
BOOK	0.03%	0.11%	0.09%	< 0.01%	0.17%	0.15%	0.02%	0.01%
DADA	–	0.01%	–	–	–	–	–	–
FBCO	0.35%	2.08%	0.05%	0.06%	0.10%	2.13%	0.08%	0.04%
GSCO	–	5.57%	–	–	–	7.52%	10.81%	–
LEHM	–	–	–	–	–	0.11%	0.05%	–
NITE	0.05%	0.07%	–	0.02%	0.96%	0.03%	0.02%	0.01%
RHCO	< 0.01%	–	–	–	–	–	–	–
SBSH	3.72%	18.90%	4.65%	24.11%	51.56%	4.89%	0.87%	16.49%
TMBR	68.08%	57.89%	94.45%	73.06%	33.73%	72.15%	84.95%	81.78%
UBSS	27.23%	14.98%	0.73%	2.71%	13.26%	12.77%	3.09%	1.61%
WCHV	0.01%	–	–	–	–	–	–	–
WEMM	0.32%	0.16%	0.01%	0.02%	0.06%	0.11%	0.03%	0.02%
Herfindahl-Hirschman Index	5390.47	3968.31	8942.97	5926.41	3972.96	5453.76	7343.68	6962.48

This table shows the list of Nasdaq market participant identifiers (MPIDs) from a sample of eight Nasdaq-traded stocks for 4-22 November 2013. Reported are the MPIDs and their relative contribution (in percentage terms) to the total submission volume for each stock. Also reported are each stock's Herfindahl-Hirschman Index (HHI), to capture the each stock's level of competition between liquidity providers.

4 Determinants of Market Maker MPID-Attribuetd Order Submissions

4.1 Methodology: Logistic Regression

In the first step of our analysis, we would like to examine whether market participants submit MPID-attributed orders in their capacity as markets makers, in order to correct order imbalances and stabilize uncertain markets. To that aim, a logistic regression specification is used to regress a binary decision variable regarding the anonymity of a submitted order on the various market characteristics and order submission characteristics that are hypothesized to influence that decision. The output of logistic regression tells the probability of the binary response based on the independent variables. Therefore, we use this specification to say whether the presence of certain market characteristics increases the probability that the market maker will choose to submit at MPID-attributed order.

More precisely, the dependent variable in our model is defined as a dummy variable equal to one if the order has been submitted with an attached MPID, and equal to zero if it has been submitted without an attached MPID (i.e., has been submitted anonymously). For stock i at time t, this is equivalent to:

$$MPID_t^i = \begin{cases} 1 & \text{if attributed to an MPID} \\ 0 & \text{if anonymous} \end{cases} \tag{1}$$

In alternative specifications, the dummy variable in (1) is specified to equal one if the submitted order is a buy order, $MPID_t^{i,B}$, or specified to equal one if the submitted order is a sell order, $MPID_t^{i,S}$.

The logistic regression is estimated for each stock i and submission at time t and can be considered as the estimation of the vector of coefficients γ^i that best fits the logistic regression model:

$$y_t^i = \alpha^i + \mathbf{x}_t^i(\gamma^i)' + \varepsilon_t^i \quad \text{where} \quad MPID_t^i = \begin{cases} 1 & \text{if} \quad y_t^i > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

where $\mathbf{x}_t^i$ is a column vector containing the explanatory variables. These explanatory variables includes market conditions and order characteristics that are hypothesized to influence the decision of the market maker to submit an MPID-attributed order, such as bid-ask spreads and depth. y_t^i can be thought of as an unobservable latent variable that is observed through the binary observable variable $MPID_t^i$. Errors ε_t^i are distributed according to the logistic distribution. This implies that the conditional probability $p(\mathbf{x}_t^i) = \Pr[MPID_t^i = 1|\mathbf{x}_t^i]$ is given by the following *link function*:

$$p(\mathbf{x}_t^i) = \frac{1}{1 - \exp(-(\alpha^i + \gamma^i \mathbf{x}_t))}. \tag{3}$$

The model in (2) is typically estimated using maximum likelihood estimation, where the likelihood function is given by:

$$L(\alpha^i, \gamma^i) = \prod_{t=1}^{T} p(\mathbf{x}_t^i)^{MPID_t^i} \cdot (1 - p(\mathbf{x}_t^i))^{(1-MPID_t^i)}. \tag{4}$$

In order to solve the maximum likelihood problem, typically one would differentiate the log of (4) with respect to the parameters, set to zero and solve for the parameters. However, there is no closed-form solution to maximum likelihood problem for logistic regressions, and numerical approximations must be used. We use the Newton-Raphson iterative procedure, which guesses at initial parameter values, and updates the values based on the gradient and Hessian matrix evaluated at the parameters (for more details, see Cameron and Trivedi, 2005).

Due to the nature of our data as rare event data (only in about 3% of the cases is the variable in equation (1) equal to one), special care has to be taken in estimating the logistic regression using maximum likelihood. Authors such as King and Zeng (2001), Gao and Shen (2007), and others have shown that results from the logistic estimation can be quite biased when used with rare event data. Therefore, this analysis uses the Firth (1993) penalized maximum likelihood estimation, which introduces a penalty term to the likelihood function that adjusts for the presence of bias. Details on the Firth penalized log likelihood methodology are provided in Appendix B.

Recall from the summary statistics that the stocks in our sample differ widely in terms of prices, which may affect the scaling of coefficients. Therefore, to ease the comparison of coefficients obtained from the regression in (2) across stocks, for each stock i, all variables except binary dummy variables are standardized by subtracting the stock time-series average and dividing by the stock time-series standard deviation. Lastly, consideration is made to distinguish between different sides of the book. Thus, the estimation of the regression in (2) includes only buy-side submissions.

4.2 Determinants of Market Maker Order Submissions: Empirical Results

This section presents our empirical results on the determinants of MPID revelation by market makers by estimating a logistic regression in which the dependent variable is a dummy variable equal to one if an order is submitted with an MPID, and zero otherwise. The dependent variable is regressed on a series of explanatory variables capturing market conditions which are hypothesized to influence the decision on whether to non-anonymously intervene as a market maker. From Hypothesis 1, if MPID-attributed orders are submitted by market markets according to incentives to encourage liquidity and stabilize markets, then we should see an increase in MPID submissions in response to market conditions characterized by high uncertainty and lower market participation.

These measures of market conditions include relative quoted bid-ask spreads ($REL\_QSPR_t^i$); 5-minute volatility (VOL_t^i); total execution volumes (EXE_t^i); total submission volumes (SUB_t^i); and average total depth ($DEPTH_t^i$). To capture the potential effects of downwards markets, also included are (unsigned) relative changes in prices ($\Delta RELPRC_t^i$), a dummy variable equal to one if this relative price change is negative ($DOWN_t^i$), and the interaction between the former two variables to capture the effects of large, negative changes in prices ($\Delta RELPRC_t^i \times DOWN_{t-\Delta t}^i$). Additional control variables include the absolute volume of preceding MPID-attributed orders ($ABSVOL_t^i$), in order to take into account the potential clustering of MPID-attributed orders, and a dummy variable capturing if the order is submitted in the first fifteen minutes of the trading day ($OPEN_t^i$), to capture time-of-day effects. Also included are characteristics of the order itself, including order size ($ORSZ_t^i$) and aggressiveness ($AGGR_t^i$). This is considered the baseline specification. To avoid multicollinearity, alternative specifications replace relative quoted bid-ask spreads with relative effective spreads ($REL\_ESPR_t^i$). Buy-side and sell-side execution and submissions volumes and depth are further considered in different alternative specifications. Lastly, to capture potential effects of upwards markets, an alternative specification replaces ($DOWN_t^i$) with a dummy variable equal to one if the relative price change is positive (UP_t^i). For detailed information on the construction of these variables, see Appendix A.

Note that it is unlikely that market makers respond instantaneously to market conditions; more likely, their strategies depend on market conditions that have persisted for some duration of time. Therefore, rather than taking their instantaneous values, most market characteristic variables are average across a specific interval preceding the order submission. These intervals are determined by event time, rather than calendar time, to ease comparison of the results between stocks that differ in terms of trading timescales. For example, the average time between submissions varies from a low of 40 milliseconds (FB) to a high of 330 milliseconds (GOOG); a calendar time window of, for example, 5 seconds would thus include an average of 125 submissions for FB, but only 15 submissions for GOOG. Therefore, this analysis calculates all variables (with the exception of VOL_t^i, $OPEN_t^i$, $ORSZ_t^i$ and $AGGR_t^i$) over the interval $[t-\Delta, t]$, where t is the time stamp of a submitted order, and Δ is the time that it takes to observe $N = 20$ consecutive order book updates (i.e., message files).

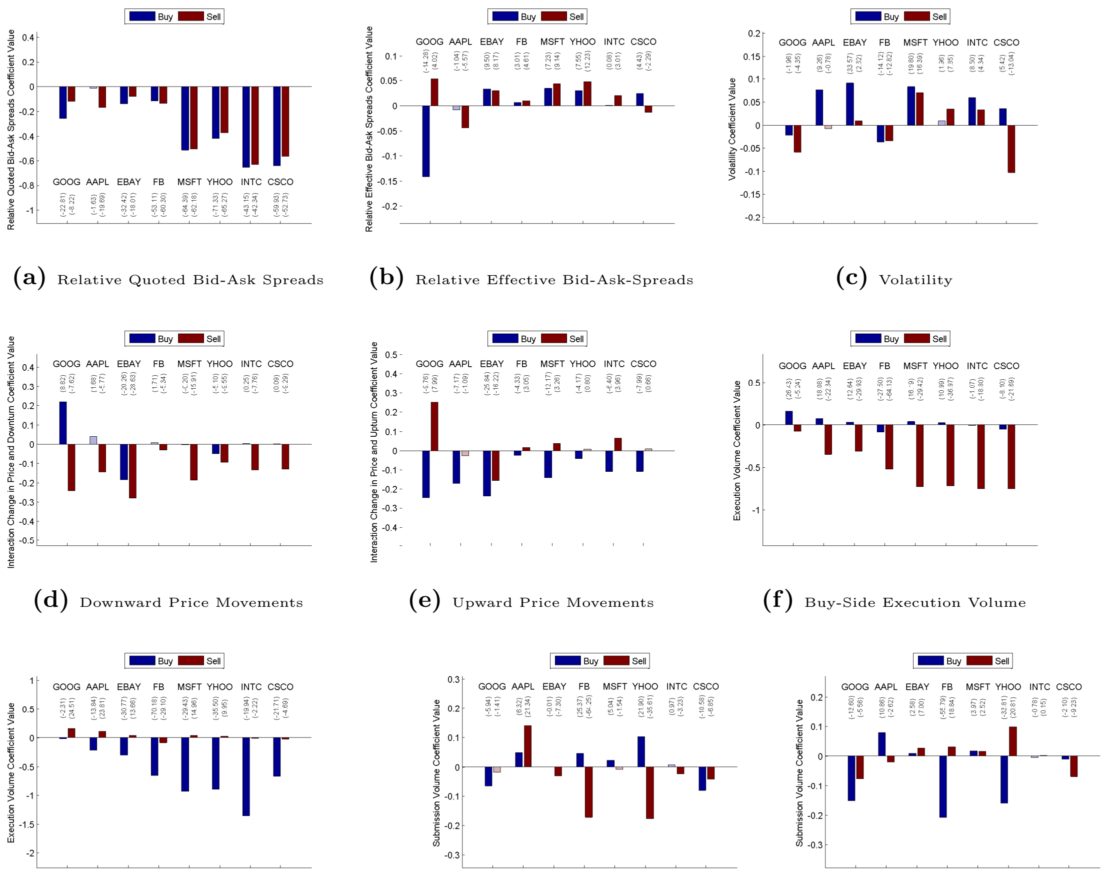
Figure 1 plots the coefficients of interest from the logistic regression, and results are presented separately for each stock. Results are presented separately for specifications in which the dependent variable is either buy-side or sell-side MPID-attributed submissions ($MPID_t^{i,B}$ or $MPID_t^{i,S}$). The coefficients from each stock-level regression are ordered according to the stock's mean share price over the sample period, from highest (GOOG, at \$1025.43) to lowest (CSCO, at \$22.11). Results from the logistic regressions provide several key insights into what determines MPID-attributed order submission by market makers.

In general, we do not find much support for the idea in Hypothesis 1, that market makers use their MPID-attributed orders to encourage liquidity and stabilize markets in conditions of low market participation or high market uncertainty. First, Figure 1a shows that the coefficient of both $MPID_t^{i,B}$ and $MPID_t^{i,S}$ on relative quoted bid-ask spreads are negative for all stocks, implying that MPID-attributed orders are more likely following periods of low relative quoted bid-ask spreads. Since low quoted bid-ask spreads signal a regime in which liquidity supply is plentiful and/or adverse selection is low (see, e.g., Glosten and Milgrom, 1985; Glosten and Harris, 1988; Huang and Stoll, 1997), this is in contrast with the idea in Hypothesis 1 that market makers intervene in times of low available liquidity or high market uncertainty. Instead, we have a first clue that market makers may be submitted MPID-attributed orders in accordance with other strategies.

Secondly, Figure 1d show that market makers are less likely their MPID-attributed sell orders (and are furthermore not necessarily more likely to submit MPID-attributed buy orders) following periods of downward price movements. This is again in contrast to Hypothesis 1, in which we would expect market makers to intervene during market downturns when market participation may be low. Lastly, there does not appear to be a consistent response of MPID-attributed order submissions when volatility is high (Figure 1c).

Likewise, Figures 1f, 1g, 1h, and 1i show results for market makers' MPID-attributed order submissions strategies in response to buy- and sell-side execution and submission volumes. In order to interpret these coefficients, it is helpful to take a more detailed examination of Hypothesis 1.

Figure 1: Coefficients from Logistic Regression



(a) Relative Quoted Bid-Ask Spreads

(b) Relative Effective Bid-Ask-Spreads

(c) Volatility

(d) Downward Price Movements

(e) Upward Price Movements

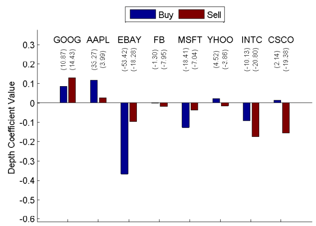
(f) Buy-Side Execution Volume

(g) Sell-Side Execution Volume

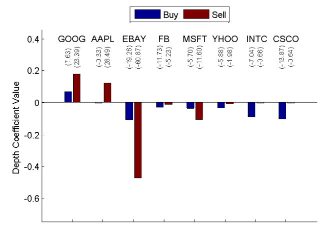
(h) Buy-Side Submission Volume

(i) Sell-Side Submission Volume

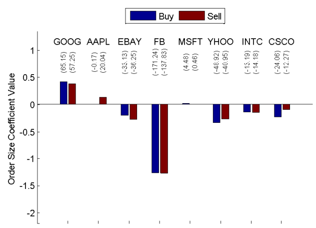
Figure 1: Coefficients from Logistic Regression (Cont.)



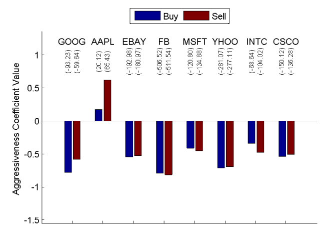
(j) Buy-Side Depth Volume



(k) Sell-Side Depth



(l) Order Size



(m) Aggressiveness

The bar graphs in the figure above correspond to key coefficients from a logistic regression of a dummy variable equal to one if a submitted order is attributed to any MPID, and zero if the submitted order is anonymous, on various variables meant to capture market conditions during the interval $[t - \Delta t, t]$, where Δt is the time it takes to observe $N = 20$ consecutive submissions, prior to the submitted order, as well as additional characteristics of the submitted order. Results are presented separately for buy- and sell-side MPID-attributed orders, and for each stock separately, ordered according to the mean stock price over the sample time period (4-22 November 2013) from highest to lowest. t-statistics are presented in parenthesis below the stock name. Dark blue and red bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light blue and red bars correspond to coefficients that are insignificant at this level.

Clearly, if MPID-attributed orders are indeed submitted in order to fulfill their obligations to provide liquidity when it is otherwise not available, then MPID-attributed orders should be more likely when same-side submissions are low (e.g., a negative coefficient between $MPID_t^{i,B}$ and $SUB_t^{i,B}$). However, the response of MPID-attributed orders to same-side execution volume should be mixed, depending whether the primary task of the market maker is to encourage liquidity, or as a stabilizing response to liquidity demand. A higher buy-side execution volume (i.e., higher execution of sell-side limit orders) signals that there is more impatient trader demand to sell. Therefore, if the market makers wishes to act to stabilize markets, they should increase their buy-side supply of limit orders to meet this demand (i.e., a positive coefficient between $MPID_t^{i,B}$ and $EXE_t^{i,B}$). On the other hand, if market marker's primary task is to encourage liquidity, in response to a lower buy-side execution volume, market makers should increase their buy-side liquidity supply in order to encourage reactive traders (i.e., a negative coefficients between $MPID_t^{i,B}$ and $EXE_t^{i,B}$).

On the other hand, if opposite-side submissions are high, then we might expect to see an increase in MPID-attributed submissions, as market makers try to encourage passive or "reactive" traders to cross the spread. For example, if buy-side submissions are high, then we should see a higher number of MPID-attributed sell-side submissions (i.e., a positive coefficient between $MPID_t^{i,S}$ and $SUB_t^{i,B}$). If market makers view their MPID-attributed buy- and sell-side orders are substitutes, then the response to opposite-side executions should again be mixed. If market makers are responding to high buy-side demand, they should decrease their sell-side limit orders as they substitute these for buy-side limit orders to meet the incoming market orders (a negative coefficient between $MPID_t^{i,S}$ and $EXE_t^{i,B}$). If market makers are encouraging buy-side demand, they should increase their sell-side submissions in response to high buy-side demand, as there is less need to substitute them for buy-side submissions when demand is high (a positive coefficient between $MPID_t^{i,S}$ and $EXE_t^{i,B}$).

In contrast to this idea, Figures 1h and 1i similarly do not show a consistent reaction to either same- or opposite-side submission volume. While the results from Figures 1f and 1g do show mixed responses of MPID-attributed submissions to same-side execution volume, however, they interesting show a clear response of market makers to opposite-side execution volume in terms of their order submission strategies. Figure 2f shows that, for all stocks, the coefficient of $MPID_t^{i,S}$ on $EXE_t^{i,B}$ is strongly negative. This implies that, when buy-side executions (i.e., submissions of sell market orders) are high, market makers will decrease their MPID-attributed submissions to the sell side of the book. The same result holds in Figure 2g for the coefficient of $MPID_t^{i,B}$ on $EXE_t^{i,S}$: when sell-side executions are high, market makers will decrease their MPID-attributed submissions to the buy side of the book. Therefore, it does not seem that market makers are simply using buy- and sell-side MPID-attributed orders as substitutes in accordance to a strategy to meet liquidity demand or to encourage reactive liquidity.

Indeed, the results tend to support the idea that market makers submit in according to another strategy, particularly one in which they internalize the potential information content of their orders. Consider again the result from Figure 1d regarding price movements: this behavior is consistent with a strategy in which the market maker is responding to potential flows of information. First, when prices are moving downwards, market markets avoid submitting a sell order, which may quickly become stale and susceptible to being "picked off" by a more informed trader. Second, they avoid a situation in which they would contribute to information flows by signalling that they have information to sell when prices are already moving downwards. This idea is further strengthened by the symmetric result in Figure 1e, which shows that market makers are less likely to submit MPID-attributed *buy* orders following *upward* price movements.

Secondly, if market makers are acting to reduce the informational impact of their orders, then the impact of execution volume for MPID-attributed orders on the *opposite* side of the book are more clear, and consistent with the results from Figures 1f and 1g. A higher buy-side execution volume signals the potential presence of impatient traders with short-lived information and a need to sell. To avoid contributing to this information flow, market makers should decrease their MPID-attributed presence on the sell side of the book, implying a negative coefficient between $MPID_t^{i,S}$ and $EXE_t^{i,B}$. At the same time, a lower buy-side execution volume signals that there is a lack of short-lived information signalling a need to sell. Therefore, given that they are constrained to submit MPID-attributed orders, market makers may choose these opportunities to reveal their presence on the sell side because it is relatively "safe". This also implies a negative coefficient between $MPID_t^{i,S}$ and $EXE_t^{i,B}$. Consistent with this idea, we indeed see a strongly negative coefficient between $MPID_t^{i,S}$ and $EXE_t^{i,B}$, as well as between $MPID_t^{i,B}$ and $EXE_t^{i,S}$.

Lastly, Figures 1l and 1m shows results for the order characteristics of MPID-attributed market maker orders. The mostly negative coefficients of both $MPID_t^{i,B}$ and $MPID_t^{i,S}$ on order size and aggressiveness, $ORSZ_t^i$ and $AGGR_t^i$, show that MPID-attributed orders by market makers tend to be smaller and less aggressive. As large, aggressive orders are typically associated with the highest price impact and risk of informed trading (see, e.g., Easley and O'Hara, 1987; Biais et al., 1995; Kaniel and Liu, 2006), market makers may avoid the use of such orders to avoid any further destabilizing price effects from a MPID-attributed submission. Therefore, market makers may be more likely to submit orders that are smaller and less aggressive when they are wary that their orders may signal information.

It should be noted that some results are consistent with Hypothesis 1. Interestingly, Figure 1b shows a tendency towards MPID revelation when effective bid-ask spreads are high. This is somewhat inconsistent with the idea that market makers submit MPID-attributed orders when their exposure to adverse selection is low, and instead reflects an intervention by market makers when adverse selection is high. Secondly, Figures 1j and 1k show that MPID-attribution is higher when depth is low, on both sides of the book. This is again more consistent with Hypothesis 1, reflecting that market makers act to encourage liquidity. Thus, it appears that market makers to a certain extent may still use MPID-attributed orders according to an obligations to stabilize markets; however, the remaining results reflect that the constraints put on market makers by requiring them to disclose their identity causes them to avoid certain circumstances in which a calming presence would otherwise be warranted.

Therefore, while not inconsistent with the idea that market makers may still use MPID-attributed orders in order to either stabilize markets or encourage liquidity in certain circumstances, the results from the logistic regression are more in support of an alternative idea that market makers internalize the market's reaction to the information content of their MPID-attributed orders. Therefore, the fact that the liquidity provision incentive programs on Nasdaq mandate that market makers submit

MPID-attributed orders may be counterproductive, as it may constrain market makers from using their full capacity in order to stabilize markets. Whether or not the market indeed views MPID-attributed orders by market makers as informed, and whether MPID-attributed orders can therefore have a de-stabilizing, rather than stabilizing, effect on markets, remains to be explored in the following Section.

5 Effects of of Market Maker MPID-Attributed Order Submissions

5.1 Methodology: Heckman Correction Regression

In this section, we aim to explore the market's reaction to MPID-attributed orders posted by market makers, and the implications that this has for market quality. This takes the form of a simple regression of ex-post market quality measures (i.e., measures of market conditions following the submissions of limit orders) on a dummy variable equal to one if a submitted order is attributed to an MPID. The estimated coefficient on the MPID dummy thus measures the marginal effect that the submission of an MPID-attributed order has on market quality and market conditions.

However, in estimating this regression, we must take into account that, although faced with a quota of non-anonymous order that they must submit, market makers still have discretion over the timing of these orders, and their selection of when to reveal likely depends on the expected costs of doing so. By submitting an MPID-attributed order, for example, when they expect execution volumes to increase, market makers reduce their own exposure to execution risk. Therefore, it would be difficult for an observer to determine whether executions increase as a result of the market maker's submission, or the submission was a result of an expected increase in executions. This introduces a likely endogeneity problem between the market maker's submission decision, and ex-post market conditions. Therefore, in to correct for this potential endogeneity problem, the Heckman correction methodology is used. The idea behind the use of this methodology is that order characteristics (such as MPID attribution) and market conditions are endogenously determined, and thus this procedure includes a "correction" term into the regression to account for this endogeneity. The exact procedure can be described as follows.

Motivate why exactly we need this specification: "Marginal effect of MPID on market conditions"

Consider the following model:

$$
\begin{aligned}
w_{t+1}^i &= \mu^i + \mathbf{v}_t^i(\beta^i)' + \theta^i MPID_t^i + u_{t+1}^i, && (5)\\
y_t^i &= \alpha^i + \mathbf{x}_t^i(\gamma^i)' + \varepsilon_t^i && (6)\\
MPID_t^i &= \begin{cases} 1 & \text{if} \quad y_t^i > 0 \\ 0 & \text{otherwise,} \end{cases} && (7)
\end{aligned}
$$

where in our case w_{t+1}^i in (5) can take the form of several different variables capturing ex-post market conditions following an order submission at time t, such as bid-ask spreads and volatility. $\mathbf{v}_t^i$ is a column vector of explanatory variables that influence market conditions; considering the strong autocorrelation between market conditions, these explanatory variables include, e.g., lagged bid-ask spreads and volatility. Note that, in our case, of particular interest will be the coefficient θ^i on the dummy variable $MPID_t^i$, which should given the marginal effect of an MPID-attributed submission on ex-post market conditions.

Equation (6) above is defined similarly to the logistic model in (2), in which y_t^i is an unobserved latent variable. Again, this unobserved latent variable is captured by the observed dummy variable in (7), equal to one if the order is attributed to an MPID, and zero otherwise. However, one key difference is that, in (6), the model is defined according to a *probit* specification, in which the error terms are assumed to follow the standard normal distribution, $\varepsilon_t^i \sim N(0,1)$. More specifically, the Heckman correction procedure relies on the assumption that the error terms $u_{t+1}^i, \varepsilon_t^i$ have the *bivariate normal distribution*:

$$
\begin{bmatrix} u_{t+1}^i \\ \varepsilon_t^i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \right)
$$

Then, the condition expectation of w_{t+1}^i given a positive value of y_t^i follows from the properties of the conditional expectation of truncated normal distribution:

$$
\begin{aligned}
E[w_{t+1}^i | \mathbf{v}_t^i, MPID_t^i, y_t^i > 0] &= \mu^i + \mathbf{v}_t^i(\beta^i)' + \theta^i MPID_t^i + E[u_{t+1}^i | \varepsilon_t^i > -(\alpha^i + \mathbf{x}_t^i(\gamma^i)')] \\
&= \mu^i + \mathbf{v}_t^i(\beta^i)' + \theta^i MPID_t^i + \frac{\sigma_{12}}{\sigma_2} \cdot \frac{\phi(-(\alpha^i + \mathbf{x}_t^i(\gamma^i)'))}{1 - \Phi(-(\alpha^i + \mathbf{x}_t^i(\gamma^i)'))}
\end{aligned}
$$

where ϕ and Φ are, respectively, the standard normal density and cumulative distribution functions. Note that $E[w_{t+1}^i | \mathbf{v}_t^i, MPID_t^i, y_t^i \leq 0]$ can be found analogously as:

$$
E[w_{t+1}^i | \mathbf{v}_t^i, MPID_t^i, y_t^i \leq 0] = \mu^i + \mathbf{v}_t^i(\beta^i)' + \theta^i MPID_t^i + \frac{\sigma_{12}}{\sigma_2} \cdot - \frac{\phi(-(\alpha^i + \mathbf{x}_t^i(\gamma^i)'))}{\Phi(-(\alpha^i + \mathbf{x}_t^i(\gamma^i)'))}
$$

The second terms in the above equations are also known as the inverse Mills ratio. Heckman (1979) thus considers the selection problem as the case of an "omitted variable" problem, with the inverse Mills ratio acting as the omitted variable. Therefore, his procedure consists of two steps. In the first step, a probit model as represented by (6) and (7) is estimated, and the predicted values $\hat{y}_t^i = \hat{\alpha}^i + \mathbf{x}_t^i(\hat{\gamma}^i)'$ are obtained using the estimated intercepts and parameters $\hat{\gamma}^i$. Then, the inverse Mills ratio, or the *Heckman selectivity correction terms* is obtained as:

$$\lambda_t^i = MPID_t^i \cdot \frac{\phi(-\hat{y}_t^i)}{1-\Phi(-\hat{y}_t^i)} + (1-MPID_t^i) \cdot -\frac{\phi(-\hat{y}_t^i)}{\Phi(-\hat{y}_t^i)} \tag{8}$$

In the second step, these terms are then included in an OLS regression of market conditions and order characteristics, including MPID attribution and the Heckman selectivity correction term as above. In this way, the relationship between MPID revelation and ex-post market quality can be explored, while controlling for the endogeneity of these market quality variables in the decision to reveal. The second-stage regression thus takes the following form:

$$w_{t+1}^i = \alpha^i + \mathbf{v}_t^i(\beta^i)' + \theta^i MPID_t^i + \delta^i \lambda_t^i + u_t^i, \tag{9}$$

Similarly to in Section 4.1, to ease the comparison of coefficients obtained from the regression in (9) across stocks, for each stock i, all variables except binary dummy variables are standardized by subtracting the stock time-series average and dividing by the stock time-series standard deviation. The estimation of the regression in (9) includes only buy-side submissions; an analysis of the sell side of the book is performed as a robustness check.

Recall from Section 4.1 that, due to the rare event nature of our data, we ideally would like to estimate the first-stage regression given by (6) and (7) using the Firth penalized maximum likelihood estimation. However, the selection model as proposed by Heckman (1979) is meant to be paired with a probit analysis, which uses a different (less fat-tailed) probability link function. Therefore, we use the procedure from Lee (1983), based on the inverse cumulative distribution function of the normal distribution, to calculate "quasi probit" scores from those of the logistic model. In this procedure, the predicted values $\hat{\tilde{y}}_t^i$ are obtained from the logistic regression from Section 4.1, and the predicted probabilities implied by the model are calculated using the link function that was given in (3):

$$p(\mathbf{x}_t^i) = \frac{1}{1+\exp(-\hat{\tilde{y}}_t^i)}$$

These predicted probabilities are then transformed into "quasi probit" scores using the inverse cumulative distribution function of the normal distribution: $\hat{y}_t^i = \Phi^{-1}(p(\mathbf{x}_t^i))$. These are then used in the calculation of the Heckman selectivity correction term in (8).

5.2 Effects of of Market Maker Order Submissions: Empirical Results

This section presents our empirical results on the effects of market maker submissions on market quality from the estimation of the Heckman correction regression as in Section 5.1. The results from the first stage of the Heckman procedure are identical to those presented in Section 4. The second stage involves a regression, given in (9), of market conditions on explanatory variables that are hypothesized to influence market conditions, along with the Heckman selectivity correction term. The key regressor in this specification is $MPID_t^i$, given in (1), as the coefficient on this variable should give the additional effect that an MPID-attributed submission by a market maker has an various market conditions. According to Hypothesis 2, if market makers are indeed submitting MPID-attributed orders in their capacity to correct order imbalances and stabilize markets, then we should see an improvement in market quality, including lower volatility, lower relative bid-ask spreads, and higher market participation rates. The dependent variables in the second-stage regression thus takes the form of a variety of different ex-post market quality measures.

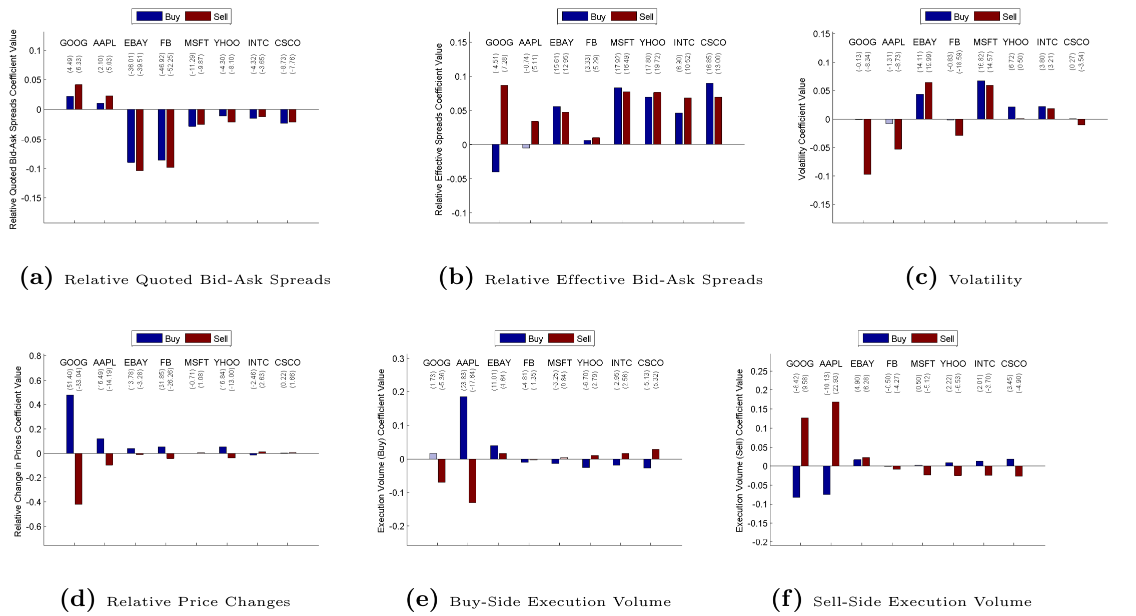
The ex-post market quality measures included as dependent variables in the second-stage regression include volatility (VOL_{t+1}^i); relative quoted and relative effective bid-ask spreads ($REL\_QSPR_{t+1}^i$ and $REL\_ESPR_{t+1}^i$); buy- and sell-side execution volume ($EXE_{t+1}^{i,B}$ and $EXE_{t+1}^{i,S}$); buy- and sell-side submission volume ($SUB_{t+1}^{i,B}$ and $SUB_{t+1}^{i,S}$); relative changes in price ($RELPRC_{t+1}^i$), and buy- and sell-side depth depth ($DEPTH_{t+1}^{i,B}$ and $DEPTH_{t+1}^{i,S}$). All variables are constructed as described in Appendix A, and, similarly to in Section 4.2, variables are calculated over an interval in order to capture potentially slower reaction times. Specifically, ex-post market quality measures are calculated over the interval $[t, t+\Delta]$, where t is the time stamp of a submitted order, and Δ is the time that it takes to observe $N = 20$ consecutive order book updates (i.e., message files).

Figure 2 plots the coefficients on $MPID_t^i$ for each stock-level regression, for the different dependent variables, where the coefficients from each stock-level regression are ordered according to the stock's mean share price over the sample period, from highest (GOOG, at \$1025.43) to lowest (CSCO, at \$22.11).[13] Results are presented separately for specifications in which either the either buy-side or sell-side MPID-attributed submissions ($MPID_t^{i,B}$ or $MPID_t^{i,S}$) serves as the regressor of interest. There are several implications that can be drawn from the results.

Generally, the results do not fully support the idea from Hypothesis 2, that market quality improves following MPID-attributed market maker orders. Figure 2a shows mostly negative coefficients from the regression of relative quoted bid-ask spreads on $MPID_t^{i,B}$ and $MPID_t^{i,S}$, implying an improvement in market quality in terms of lower *quoted* spreads following both buy- and sell-side MPID-attributed orders by market makers. However, Figure 2b shows that relative *effective* spreads largely increase following MPID-attributed submissions on both the buy and sell side. Foucault et al. (2013) argue that effective spreads are a better measure of realized trading costs, while quoted spreads capture available liquidity. Thus, it appears that, while the availability of liquidity may improve following MPID-attributed submissions, trading costs actually obtained by traders tend to go up – i.e., price impacts have a tendency to increase. This results points to an increase in adverse

[13] Note that we do not attempt to aggregate the regression results using, e.g., DuMouchel framework as in Bessembinder et al. (2009), due to the low number of stocks used in this analysis.

Figure 2: Coefficients from Heckman Correction Regression



(a) Relative Quoted Bid-Ask Spreads **(b)** Relative Effective Bid-Ask Spreads **(c)** Volatility

(d) Relative Price Changes **(e)** Buy-Side Execution Volume **(f)** Sell-Side Execution Volume

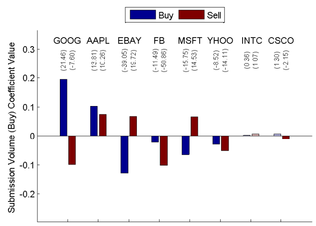
selection following MPID-attributed submissions, which contradicts the predictions of Hypothesis 2.

Secondly, Figures 2e and 2f show that execution volumes tend to decrease on the same side of the book as the MPID-attributed order. Following a MPID-attribute buy order, buy-side executions (i.e., the execution of buy-side limit orders by sell market or marketable limit orders) decrease; likewise, sell-side executions tend to decrease following MPID-attributed sell orders. This is in further contradiction to Hypothesis 1, as MPID-attributed orders by market makers do not appear to be signalling the presence of a reliable counterparty, and instead seem to repel rather than attract reactive traders to trade against their orders.

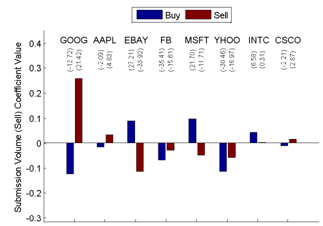
Lastly, there does not appear to be a clear and consistent impact on volatility (Figure 2c), which is inconsistent with our Hypothesis 2, but somewhat consistent with the literature. On the one hand, higher adverse selection following informed trades may reduce liquidity, thereby increasing volatility (Foucault et al., 2007). On the one hand, volatility might decrease as new information reduces the uncertainty about the value of the asset (Karam, 2012). More generally, however, a lack of clear decrease in volatility following MPID-attributed orders perhaps signals that the market does not view these MPID-attributed orders as part of market-stabilizing liquidity provision.

Instead, consistent with the results from Section 4.2, the results once again support the idea that the market views MPID-attributed orders as containing information. First, the above results from Figures 2e and 2f are consistent with the idea that, if the market expects MPID-attributed orders

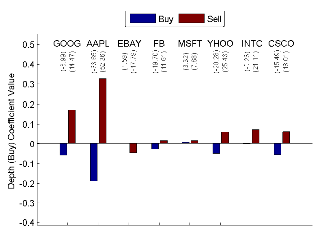
Figure 2: Coefficients from Heckman Correction Regression (Cont.)



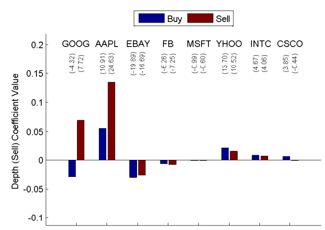
(g) Buy-Side Submission Volume



(h) Sell-Side Submission Volume



(i) Buy-Side Depth



(j) Sell-Side Depth

The bar graphs in the figure above correspond to key coefficients from a regression of measures capturing various dimensions of market quality, on a dummy variable equal to one if a submitted order is attributed to an MPID, and zero otherwise. The dependent variables are calculated during the interval $[t, t + \Delta t]$, where Δt is the time it takes to observe $N = 20$ consecutive submissions, proceeding the submitted order. Results are presented separately for buy- and sell-side MPID-attributed orders, and separately for each firm in the sample, ordered according to the stock's mean stock price during the sample time period, 4-22 November, from highest to lowest. t-statistics are presented in parenthesis above the bar. Dark blue and red bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light blue and red bars correspond to coefficients that are insignificant at this level.

to contain information, then traders will avoid trading against an order that they view as informed. Furthermore, Figures 2e and 2f show that execution volumes have a tendency to increase on the *opposite* side of the book. In other words, following an MPID-attributed buy order, the amount of buy-side market orders (execution of sell-side limit orders) tends to increase. This is also consistent with the idea that MPID-attributed orders contain information. If the market views a market market-attributed buy order as containing positive information about the stock, then traders may move to act on that information with immediacy, thereby submitting buy market orders. Figures 2g and 2h do not show much consistency in the results for submission volume, although there is a slightly tendency for submission volume to decrease on the same side of the book as the MPID-attributed order. This is again consistent with the idea that the market views MPID-attributed orders as informed, and choose to stay away from the market in the sprit of Foucault et al. (2007). It also could reflect a choice of market participants to substitute of limit orders with market orders when trading on the information reflected in MPID-attributed orders.

Secondly, Figure 2d shows that, from the positive coefficients of relative prices changes $RELDPR_t^i$ on $MPID_t^{i,B}$ and its negative coefficients on $MPID_t^{i,S}$, prices have a strong tendency to move in the same direction as an MPID-attributed market maker order. Following an MPID-attributed buy order, prices increase, while prices decrease following an MPID-attributed sell order. This contributes clear evidence that the market views MPID-attributed orders as containing information, and thus prices move in the same direction as the order.

Note that, while the above results for order flow variables hold for the majority of stocks, some differences remain between the cross-section of stocks. In particular, the opposite results tend to hold for the stocks with the highest prices. This is an interesting finding that requires further exploration. Following the implementation of Regulation National Market System, tick sizes have been set at one penny ($0.01) for all U.S. equities above $1; therefore, high-priced stocks will have much smaller *relative* tick sizes than medium- and low-priced stocks. Small relative tick sizes reduce the profitability of market making by lowering spreads, and increases competition between market makers by making it easier for competing liquidity providers to undercut their orders (Harris, 1996; O'Hara et al., 2015). Indeed, Table 3 shows that the two highest-priced stocks, AAPL and GOOG, have the highest and third-highest HHI scores, respectively. On the other hand, information-based trading has been shown to be more profitable in stocks with low relative tick sizes (Zhao and Chung, 2006), as the profitability of information outweighs trading costs. Following Boulatov and George (2013), when market makers lose their information advantage compared to other informed traders, informed traders will prefer market orders. Therefore, the order flow results for low-tick stocks may reflect a more intense competition environment among informed traders within low-tick stocks, in which more informed traders use market orders to "pick off" the less informed orders by market makers.

The idea that the market may view MPID-attributed orders by market makers as informed is also supported by the literature. Several studies have shown that market participants react significantly to the information contained in trader IDs, and use this information to gauge whether or not the non-anonymous trader is informed (Frino et al., 2010; Linnainmaa and Saar, 2012). For example, Linnainmaa and Saar (2012) show that price impacts are larger following trades by brokers with large institutional clients, implying that the market associates these trades with a higher probability of being informed. Furthermore, as mentioned in Section 2 many papers have shown that market makers may have access to private information, as they can receive information from either their associated clients (Saporta, 1997; Naik et al., 1999) or order flow (Vayanos, 2001; van Kervel and Menkveld, 2016; Malinova and Park, 2016). Other authors, such as Boulatov and George (2013) and Kaniel and Liu (2006), have shows that informed traders with long-lived information tend to prefer limit orders, and that price impacts following limit orders tend to be higher.

All-in-all, the results from this analysis are largely inconsistent with Hypothesis 2, and do not support the idea that MPID-attributed orders by market makers act to encourage liquidity or to stabilize uncertain markets. Instead, the results show that the market largely views MPID-attributed orders by market makers as informed, and reacts accordingly. Prices tend to move in the same direction as market maker MPID-attributed orders, and effective spreads following MPID-attributed orders tend to increase. Furthermore, MPID-attributed orders appear to repel, rather than attract, counterparties on the same side of the book. Interestingly, the relative tick size appears to play a role, as it may alter the informational advantage of market makers. From Section 4.2, market makers in turn internalize the expected reaction of the market, and choose to submit orders such that they minimize the impacts of their MPID-attributed orders. Once again, the results point to the counter-productivity of mandating MPID disclosure by market makers: on the one hand, rather than stabilizing markets, MPID-attributed orders have a tendency to disrupt markets by signalling the presence of informed trading; on the other hand, market makers who internalize the expected reaction by the market are unable to use MPID-attributed orders in accordance with their obligations to provide and encourage liquidity.

6 Robustness

6.1 Different Event Time Windows

In the first set of robustness checks, we want to check whether our results depend on our choice of the event time intervals $[t-\Delta, t]$ and $[t, t+\Delta]$, over which we average our ex ante and ex-post measures of market conditions. In the baseline specification, we consider Δ to be the time that it takes to observe $N = 20$ consecutive order book updates. However, this selection is somewhat arbitrary, as the exact speed at which market makers will reaction to market conditions, or the speed at which markets react to market maker MPID-attributed orders, is unobservable.

Therefore, Figures 3 and 4 in Appendix C shows results from a Heckman Correction procedure and a logistic model, in which the event time is allowed to vary. Specifically, we consider results for $N = 10$, $N = 20$, and $N = 100$ consecutive order book updates. For simplicity, only the results including buy-side submissions are reported, although results are qualitatively similarly for sell-side submissions.

Figure 3 shows results for the logistic regression as described in Section 4.1. Again, these figures show the coefficients (separately for each stock) from a regression of the dummy variable $MPID_t^{i,B}$, equal to one if a submitted buy order is MPID-attributed, on various measures that are hypothesized to influence the decision of the market maker to submit a non-anonymous order. Note that we see the coefficient on $MPID_t^{i,B}$ monotonically increase for many variables as the time window is allowed to increase, most notably in the case of relative quoted bid-ask spreads in Figure 3a; this shows that the market may react more slowly to the information revealed in MPID-attributed orders than is captured by out $N = 20$ baseline specification. In a few other cases, such as with sell-side execution volume in Figure 3g, coefficients decrease as the time window increases; this implies that these measures of market conditions may react quite quickly to the information revelation. However, is most cases, coefficients either monotonically increase or decrease as the time window increases, implying a consistency of our results across event time windows.

Only rarely do we see signs flip (and remain significant) as the time window varies. The most notable cases include the coefficients on execution and submission volume. However, our main result strongly holds, that market makers are more likely to submit MPID-attributed buy orders when sell-side execution volume is low (Figure 3g, when relative bid-ask spreads are low (Figure 3a), and effective bid ask-spreads are high (Figure 3b), and following upward movements in price (Figures 3d and Figures 3e).

Figure 4 shows results for the Heckman Correction procedure as described in Section 4.1. As in the previous analyses, presented are the coefficients on the various measures of market quality, when regressed on $MPID_t^{i,B}$, separately for each stock. Again, the results show that the coefficients have a tendency to either monotonically increase or decrease as the time window increases, and sign flips are only occasionally observed. One notable exception is for sell-side depth in Figure 4j, which shows a sign flip for the lower-priced (i.e., higher-tick) stocks when $N = 100$. These longer time-window results imply that, in the longer run, sell-side depth decreases following MPID-attributed buy orders. Note also that, from Figure 4i, the longer time windows show a stronger reaction of buy-side depth to MPID-attributed buy orders. Given that the presence of informed trading would predict a drop in submissions and a potential drop in depth on both sides of the market, this affirms our results that the market views MPID-attributed orders as containing information.

6.2 Multicollinearity

Another concern may be the effects of multicollineary on our model estimation, which may effect the validity of our individual predictors. For example, one concern might be the multicollinearity between bid-ask spreads and volatility – variables that have been shown in the literature to have very strong interdependencies, particularly on a high-frequency level (see, e.g., Chordia et al., 2001; Vayanos, 2004; Hautsch and Jeleskovic, 2008). Therefore, in a next set of robustness tests, we consider the effects of multicollinearity on our estimated coefficients.

Figure 3 shows results for the logistic regression as described in Section 4.1. Again, these figures show the coefficients from a regression of the dummy variable $MPID_t^{i,B}$, equal to one if a submitted buy order is MPID-attributed, on (1) a reduced model in which only the respective variable is included, and (2) the full baseline model as described in Section 4.2. Coefficients from the reduced model tend to be larger and more significance, but for the most part do not switch signs in the reduced model. Interestingly, the reduced model shows a much clearer picture for the reaction of market makers to submission volume than is observed from the full model. Specifically, MPID-attributed buy order attribution tends to increase when buy-side submission volume is low (Figure 3h), and sell-side submission volume is high (Figure 3i). These results are more consistent with Hypothesis 1, as they reflect a submission strategy that attempts to attract liquidity to the buy side of the book. However, the results hold that MPID-attributed order submissions are lower when quoted bid-ask spreads are high, and following upward price movements. Again, as was discussed in Section 4.2, to a certain extent market makers may still use MPID-attributed orders according to an obligations to stabilize markets. However, the results are most consistent with a strategy in which market makers internalize the information content of their orders, and attempt to avoid the de-stabilizing impacts of informed orders.

Figure 6 shows results for the Heckman Correction procedure as described in Section 5.1, in which results are presented for a regression of the respective variable on (1) a reduced model in which only the $MPID_t^{i;B}$ is included along with the Heckman correction term $\lambda_t^{i,}$, and (2) the full baseline model as described in Section 5.2. As with the above, coefficients tend to increase in the reduced model, but for the most part do not switch signs. One notable exception is found in Figure 6a: while the full model shows a tendency of relative quoted bid-ask spreads to increase following MPID-attributed buy orders, the reduced model shows the opposite. This is consistent with our main result that the market largely views MPId-attributed orders as informed.

6.3 Individual Market Markers

The next robustness check considers that different market makers may have different strategies in terms of MPID attribution, and the market may additionally view different market makers as more or less informed. Recall from Table 2 that Timber Hill, LLC (MPID: TMBR) accounts for 76.36% of total MPID submission volume across stocks. This leaves us with a sufficient sample of MPID-attributed orders, in order to consider the market's reaction to TMBR-attributed orders, and Timber Hill, LLC's subsequent strategy in placing TMBR-attributed orders. Therefore, to decrease the potential noise created by aggregating all market makers into one analysis, the next set of robustness tests considers analyses in which only order submissions by Timber Hill, LLC are included.

Figure 3 shows results for the logistic regression as described in Section 4.1, in which $MPID_t^{i,TMBR}$ is used as the dependent variable. Again, the results are largely consistent with those from the main analysis in Section 4.2. Figure 7c shows a stronger tendency towards higher MPID attribution when volatility is high. There is also a clearer picture of the relationship between MPID attribution and depth, as Timber Hill, LLC is more likely to submit a TMBR-attributed order when same-side depth is low. Therefore, TMBR-attributed orders may be seen as less informed in lower-priced stocks, and Timber Hill, LLC may still be able to use MPID-attributed orders in order to encourage liquidity.

Figure 8 shows results for the Heckman Correction procedure as described in Section 4.1, in which the dummy variable $MPID_t^{i,TMBR}$ is defined as one if the submitted order is attributed to TMBR, and zero otherwise. Results are presented separately for buy- and sell-side submissions. The results are largely consistent with those from Section 5.2. However, it should be noted that the results for execution volume tend to lose their significance for the lower-prices stocks. Therefore, especially considering the above results, it is less likely that, in these stocks, the market views TMBR-attributed orders as informed.

7 Conclusion

Since the mid-2000s, technological advances and regulatory changes have largely pushed out the role of traditional market makers, mandated to maintain orderly markets, and led to the rise of endogenous liquidity providers, who provide liquidity on their own accounts without direct obligations. To ensure a smooth and continuous provision of liquidity, Nasdaq and other exchanges have implemented a number of programs, to encourage market participants who register as market makers to fulfill the roles that make market makers invaluable to financial markets: namely, to step in as counterparties when they are scarce, and to stabilize markets when uncertainty is high. On the other hand, such incentive programs mandate that market makers disclose their identities for a certain quota of submissions throughout the trading day. This leaves market makers at a crossroads between fulfilling their obligations as liquidity providers, and avoiding the adverse selection that comes along with order disclosure within highly anonymized electronic markets. Thus, with this paper, we aim to explore whether MPID-attributed orders by market makers are indeed aligned with incentives to act according to these obligations, and the potential implications that this has for market quality in modern-day equity markets.

Using a unique dataset that contains information on the identities of the market makers who submit orders, our results show that the market largely views MPID-attributed orders submissions by market markers as informed. MPID-attributed orders tend to be followed by price movements in the same direction as the order, and order flow shifts to reflect market participants' reactions to the information revealed by the MPID-attributed order. We see a general deterioration of certain measures of market quality, such as an increase in effective bid-ask spreads.

Our results further show that, internalizing the market's response to their MPID-attributed orders, market makers choose to submit MPID-attributed orders when they expect the impacts of their orders to be minimized. Market makers tend to submit MPID-attributed orders when quoted spreads are low, and are less likely to submit MPID-attributed orders following price changes (both negative and positive). They are also less likely to submit MPID-attributed following increases in opposite-side execution volume; in other words, they avoid contributing additional signals to information flows. Furthermore, MPID-attributed orders also tend to be smaller and less aggressive, to avoid signalling further that the order is informed. These results show that mandates for identity disclosure may constrain market makers more than incentivize them, as they must act to minimize the de-stabilizing impacts of their non-anonymous orders, rather than using them in their capacity to stabilize markets.

A Measures of Market Conditions and Order Characteristics: Details

- **Absolute Volume of MPID Orders**: $ABSVOL_t^i$, The submission volume during the interval $[t-\Delta, t]$ that is attributed to any MPID; letting v_t^i represent the dollar volume of the submitted order and letting $\mathbb{I}_t^i$ represent a dummy variable equal to one if a submission is attributed to an MPID, for stock i for an order submission at time t the measure is equal to:

$$ABSVOL_t^i = \sum_{s=t-\Delta t}^{t} (\mathbb{I}_s^i \cdot v_s^i)$$

- **Relative Quoted Bid-Ask Spread**: $REL\_QSPR_{t-\Delta t:t}^i$, the average difference between the observed bid and ask quotes during the interval $[t-\Delta, t]$, standardized by the average midquote (i.e., the average of the bid and ask price) during the respective interval; for stock i for an order submission at time t, denoting by $\#(\cdot)$ the number of quotes during that interval:

$$REL\_QSPR_t^i = \frac{\sum_{s=t-\Delta t}^{t}(ask_s^i - bid_s^i)/\#(t-\Delta t:t)}{\sum_{s=t-\Delta t}^{t}\frac{1}{2}(ask_s^i + bid_s^i)/\#(t-\Delta t:t)} \tag{10}$$

- **Relative Effective Bid-Ask Spread**: $REL\_ESPR_t^i$, the average difference between the prices actually obtained by investors upon order execution and the midquote during the interval $[t-\Delta, t]$, standardized by the midquote. Let $\mathbb{I}_t^{EXE}$ be an indicator variable equal to 1 if the order book message at time t is represents an order execution, and 0 otherwise. Denote by p_t^i and by d_t^i the price and direction of an order executed at time t, and by m_t^i the prevailing midquote at time t (as defined about and in equations 10 and **??**); for firm i for an order submission at time t, this is equal to:

$$ESPR_{t-\Delta t:t}^i = \sum_{s=t-\Delta t:t}^{t} d_s^i \cdot (p_s^i - m_s^i) \cdot \mathbb{I}_s^{EXE} / \sum_{s=t-\Delta t:t}^{t} \mathbb{I}_s^{EXE} \tag{11}$$

$$REL\_ESPR_{t-\Delta t:t}^i = \frac{ESPR_{t-\Delta t:t}^i}{m_{t-\Delta t:t}^i} \tag{12}$$

- **Volatility**: VOL_t^i, for a given order, is defined as the sum of squared 1-minute midquote returns for the five minutes prior to the that order, calculated using sub-sampling over 10-second grids. This definition follows that of Hautsch (2011). A more detailed description of the calculation steps is given below.

  In a first step, log midquotes are calculated as the average bid and ask price; for firm i at time t this is given by:

$$m_t^i := \log\left(\frac{1}{2}(ask_t^i + bid_t^i)\right). \tag{13}$$

The interval of interest is taken as the x-minute interval prior to an order submission, denoted by $[t-x, t]$. For simplicity, in the following notation the interval is normalized to $[0, t]$. Midquote returns of length $\Delta_n = t/n$ are then calculated. The realized variance is then calculated as the sum of the n squared returns during the x-minute interval prior to a given submission time t:

$$RV^i_{0:t,n} := \sum_{s=1}^{n} (m^i_{s\Delta_n} - m^i_{(s-1)\Delta_n})^2 =: \sum_{s=1}^{n} (r^i_{j,n})^2$$

Next, consider K sub-intervals of midquotes given by:

$$\begin{gathered}\{m_{1\Delta_n}, m_{(K+1)\Delta_n}, m_{(2K+1)\Delta_n}, ..., m_{(n_1 K+1)\Delta_n}\}\\ \{m_{2\Delta_n}, m_{(K+2)\Delta_n}, m_{(2K+2)\Delta_n}, ..., m_{(n_2 K+2)\Delta_n}\}\\ \vdots \\ \{m_{K\Delta_n}, m_{2K\Delta_n}, m_{3K\Delta_n}, ..., m_{(n_k+1)K\Delta_n}\}\end{gathered}$$

The realized variance specific to each sub-interval is thus given by (for $k = 1, ..., K$):

$$RV^i_{0:t,n_k} := \sum_{s=1}^{n_k} (m^i_{(sK+k)\Delta_n} - m^i_{((s-1)K+k)\Delta_n})^2 =: \sum_{s=1}^{n_k} (r^i_{j,n_k})^2$$

The volatility measure is then calculated as the average realized variances over the sub-intervals. In this study, to ensure an equally-spaced grid the interval of interest is the $x = 5$-minute interval prior to a given order submission, with returns calculated over interval length 1 minute such that $\Delta_n = 1/5$. For the sub-sampling, $K = 10$ intervals are chosen over a 10-second grid. Therefore, for stock i and order submission time t, this is given by:

$$VOL^i_t := \frac{1}{10} \sum_{k=1}^{10} RV^i_{t-5:t,n_k}$$

- **Absolute Volume of Submission Orders**: SUB^i_t, the dollar volume of submissions during the interval $[t-\Delta, t]$ preceding a given submission; denoting by $\mathbb{I}^{SUB,i}_t$ an indicator variable equal to 1 if the order book message for stock i at time t represents an order submission and 0 otherwise and by v^i_t the dollar volume of an order, for stock i for an order submission at time t the measure is equal to:

$$SUB^i_t = \sum_{s=t-\Delta t}^{t} \mathbb{I}^{SUB,i}_s \cdot v^i_s$$

- **Absolute Volume of Execution Orders**: $EXE^i_{t-\Delta t:t}$, the dollar volume of executions during the interval $[t-\Delta, t]$ preceding a given submission; $\mathbb{I}^{EXE}_t$ be an indicator variable equal to 1 if the order book message at for stock i at time t is represents an order execution and 0 otherwise and by v^i_t the dollar volume of an order, for stock i for an order submission at time t the measure is equal to:

$$EXE_t^i = \sum_{s=t-\Delta t}^{t} \mathbb{I}_s^{EXE,i} \cdot v_s^i$$

- **Total Depth**: $DEPTH_{t-\Delta t:t}^i$, the average total depth available on both sides of the book during the interval $[t-\Delta, t]$. Denoting by bid_t^i the prevailing best bid quote, $qbid_t^i$ the quantity of shares available at the prevailing best bid quote, ask_t^i the prevailing best ask quote, and $qask_t^i$ the quantity of shares available at the prevailing best ask quote; for stock i for an order submission at time t, denoting by $\#(\cdot)$ the number of quotes during that interval:

$$DEPTH_t^i := \sum_{s=t-\Delta t}^{t} bid_s^i \times qbid_s^i + ask_s^i \times qask_s^i / \#(t-\Delta t : t)$$

- **Relative Changes in Price**: $RELPRC_{t-\Delta t:t}^i$, the relative change in the prices over during the interval $[t-\Delta, t]$ preceding an order submission. The absolute value is taken such this variable captures the magnitude but not the direction of the change. Denoting by p_t^i the price at time t, for stock i for an order submission at time t this is equal to:

$$\Delta RELPRC_t^i = |(p_t - p_{t-\Delta t})/p_{t-\Delta t}| \tag{14}$$

- **Aggressiveness**: $AGGR_t^i$, the aggressiveness of an order submission is defined as the distance of an submitted order's price to the best same-side quote divided by the midquote. The midquote for firm i at time t is defined as the average of the prevailing and bid and ask quote as in (13).

  Denoting by p_t^i the price of the submitted order and d_t^i the direction of the order (where $d_t^i = 1$ represents a buy order and $d_t^i = -1$ represents a sell order), the aggressiveness $AGGR_t^i$ of a submitted buy order for stockm i at submission time t is thus given by:

$$ABS\_AGGR_t^i = \begin{cases} p_t^i - bid_t^i & \text{if } d_t^i = 1 \\ ask_t^i - p_t^i & \text{if } d_t^i = -1. \end{cases}$$
$$AGGR_t^i = \frac{ABS\_AGGR_t^i}{m_t^i}$$

- **Order Size**: $ORSZ_t^i$, the size of an order submitted at time t in terms of U.S. dollars. Denote by p_t^i and s_t^i, respectively, the price and size (in terms of number of shares) of an order submitted at time t; for stock i for an order submission at time t the measure is equal to:

$$ORSZ_t^i = p_t^i \cdot s_t^i$$

B Firth Penalized Maximum Likelihood Estimation

Traditionally, the solution to maximum likelihood estimation (MLE) problems is simply derived as the solution to the first derivative of the log-likelihood function, also known as the *score function*:

$$U(\boldsymbol{\theta}) \equiv \frac{\partial \text{log} L(\boldsymbol{\theta}; \mathbf{y})}{\partial \boldsymbol{\theta}} = 0,$$

where the data is given by $\mathbf{y}$ and vector-valued parameters by $\boldsymbol{\theta}$, and $L(\boldsymbol{\theta}; \mathbf{y})$ is the likelihood function. The MLE parameter estimates $\hat{\boldsymbol{\theta}}$ contain a bias $b(\boldsymbol{\theta})$, which asymptotically can be expressed as:

$$b(\boldsymbol{\theta}) = \frac{b_1(\boldsymbol{\theta})}{n} + \frac{b_2(\boldsymbol{\theta})}{n^2} + \dots \quad . \tag{15}$$

Note that the first term ("first order bias") is $O(n^{-1})$, i.e., it decreases in the sample size at the rate $1/n$. Earlier attempts to reduce the first order bias in (15) focused on computationally expensive, ex-post "jackknife" methods that attempted to correct the bias after the parameters have been estimated (see, e.g., Cox and Hinkley, 1979). Instead, Firth (1993) suggests a modification to the score function to ex ante reduce the score function proportional to the bias. His modified score function takes the form of

$$U^*(\boldsymbol{\theta}) \equiv U(\boldsymbol{\theta}) + A(\boldsymbol{\theta}),$$

where the modification term A is $O(1)$ as $n \to \infty$ and is allowed to depend on the characteristics of the data. Firth (1993) shows that, generally, when $A(\boldsymbol{\theta}) = -i(\boldsymbol{\theta})b_1(\boldsymbol{\theta})$, where $-i(\boldsymbol{\theta}) = U'(\boldsymbol{\theta})$ is the negative of the local gradient of the score function, also known as the Fisher information matrix, then the solution to the score function will be free from its first-order bias. From the properties of the joint null cumulants of the score function $U(\boldsymbol{\theta})$, for estimating $\boldsymbol{\theta}$ for the exponential family of models, as in a logistic regression, the penalty term $A(\boldsymbol{\theta})$ is a matrix in which the r-th term (where $r \in (1, ..., R)$ is the number of parameters) is given by:

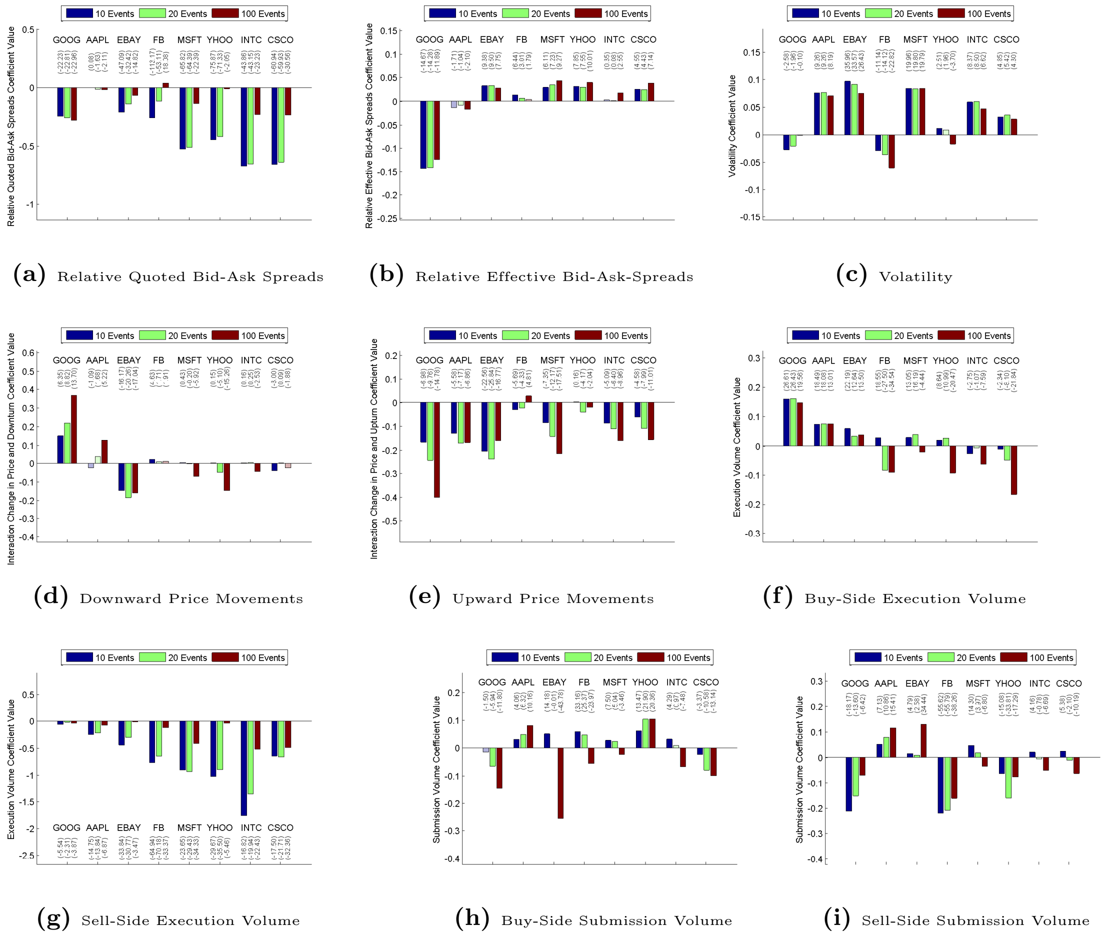
$$a_r = \frac{1}{2}\text{trace}\left\{ i(\boldsymbol{\theta})^{-1} \left(\frac{\partial i(\boldsymbol{\theta})}{\partial \theta_r} \right) \right\}. \tag{16}$$

Thus, the solution $\boldsymbol{\theta}$ to the Firth penalized maximum likelihood estimation is given by the solution to $U_r^*(\boldsymbol{\theta}) \equiv U_r(\boldsymbol{\theta}) + a_r(\boldsymbol{\theta}) = 0$. One additional benefit of the Firth bias reduction technique is that it also corrects for "separation", in which parameter estimates diverge towards infinity, while other bias reduction techniques (such as the jackknife techniques) require finite parameter estimates.[14]

[14]We use the R package **logistf** in order to calculate the firth penalized likelihood logistic regression, available at https://cran.r-project.org/web/packages/logistf/index.html. Please see the package documentation for more estimation details.

C Figures: Robustness Checks

Figure 3: Coefficients from Logistic Regression: Different Event Time Windows

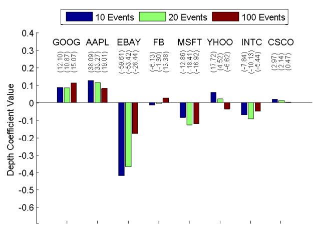


(a) Relative Quoted Bid-Ask Spreads **(b)** Relative Effective Bid-Ask-Spreads **(c)** Volatility

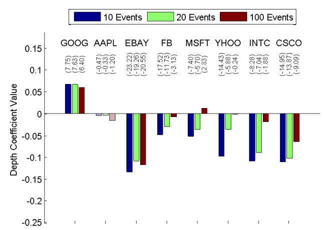
(d) Downward Price Movements **(e)** Upward Price Movements **(f)** Buy-Side Execution Volume

(g) Sell-Side Execution Volume **(h)** Buy-Side Submission Volume **(i)** Sell-Side Submission Volume

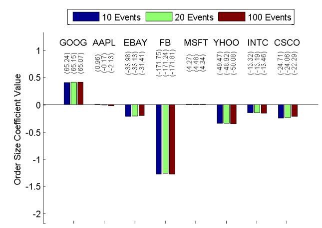
Figure 3: Coefficients from Logistic Regression: Different Event Time Windows (Cont.)



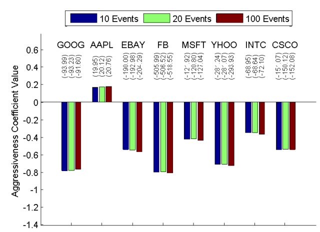
(j) Buy-Side Depth Volume



(k) Sell-Side Depth



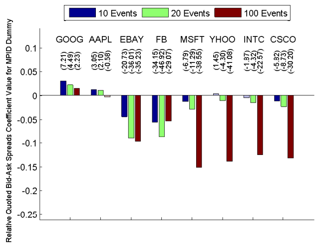
(l) Order Size



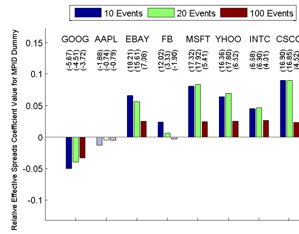
(m) Aggressiveness

The bar graphs in the figure above correspond to key coefficients from a logistic regression of a dummy variable equal to one if a submitted order is attributed to any MPID, and zero if the submitted order is anonymous, on various variables meant to capture market conditions during the interval $[t - \Delta t, t]$. Different time windows are considered, such that Δt is the time it takes to observe $N = 10$, $N = 20$, or $N = 100$ consecutive submissions, prior to the submitted order. Also included are additional characteristics of the submitted order. Results are presented for buy-side MPID-attributed orders only, and for each stock separately, ordered according to the mean stock price over the sample time period (4-22 November 2013) from highest to lowest. t-statistics are presented in parenthesis below the stock name. Dark-colored bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light-colored bars correspond to coefficients that are insignificant at this level.

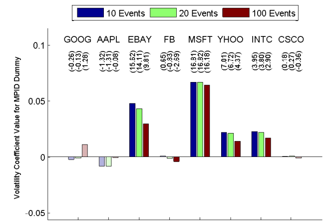
Figure 4: Coefficients from Heckman Correction Regression: Difference Event Time Windows



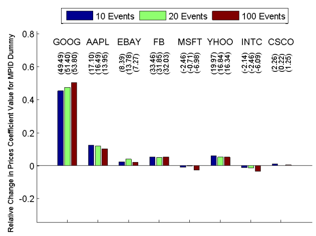
(a) Relative Quoted Bid-Ask Spreads



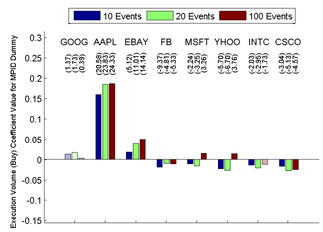
(b) Relative Effective Bid-Ask Spreads



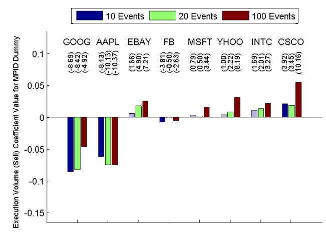
(c) Volatility



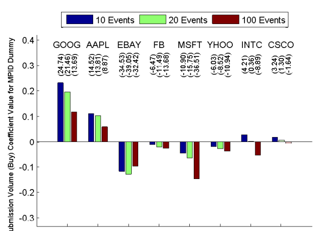
(d) Relative Price Changes



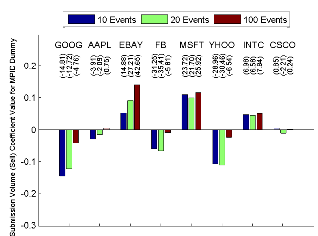
(e) Buy-Side Execution Volume



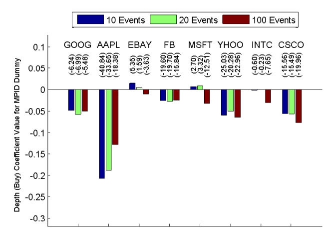
(f) Sell-Side Execution Volume



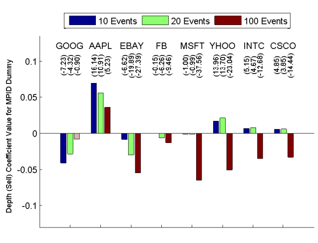
(g) Buy-Side Submission Volume



(h) Sell-Side Submission Volume



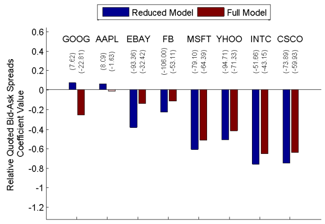
(i) Buy-Side Depth



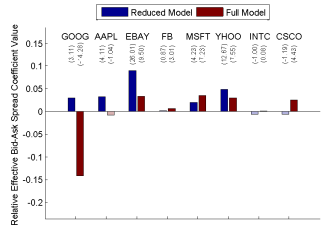
(j) Sell-Side Depth

The bar graphs in the figure above correspond to key coefficients from a regression of measures capturing various dimensions of market quality, on a dummy variable equal to one if a submitted order is attributed to Timber Hill LLC (TMBR), and zero otherwise. The dependent variables are calculated during the interval $[t - \Delta t, t]$. Different time windows are considered, such that Δt is the time it takes to observe $N = 10$, $N = 20$, or $N = 100$ consecutive submissions, prior to the submitted order. Results are presented for buy-side MPID-attributed orders only, and separately for each firm in the sample, ordered according to the stock's mean stock price during the sample time period, 4-22 November, from highest to lowest. t-statistics are presented in parenthesis above the bar. Dark-colored bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light-colored bars correspond to coefficients that are insignificant at this level.

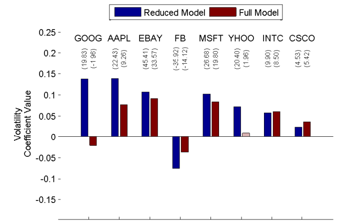
Figure 5: Coefficients from Logistic Regression: Multicollinearity



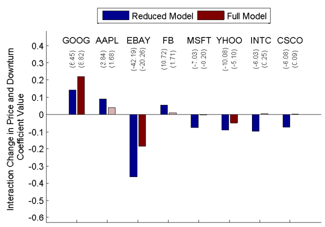
(a) Relative Quoted Bid-Ask Spreads



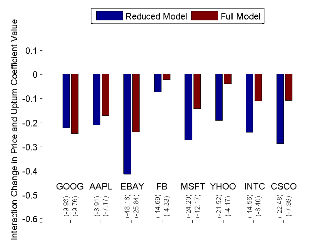
(b) Relative Effective Bid-Ask-Spreads



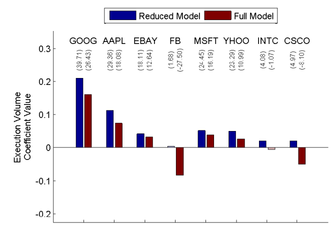
(c) Volatility



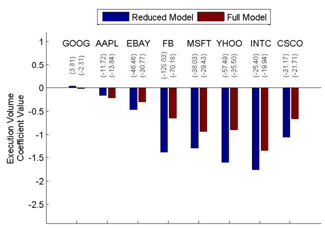
(d) Downward Price Movements



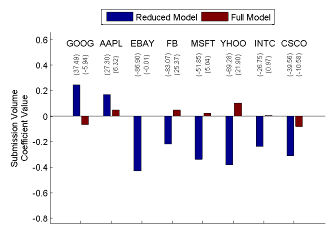
(e) Upward Price Movements



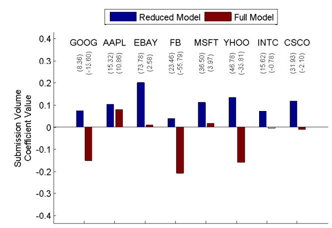
(f) Buy-Side Execution Volume



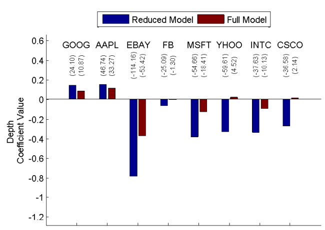
(g) Sell-Side Execution Volume



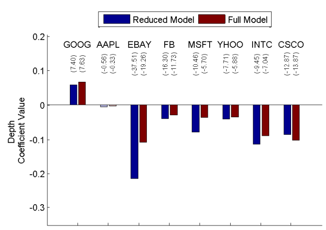
(h) Buy-Side Submission Volume



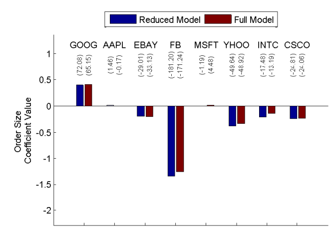
(i) Sell-Side Submission Volume



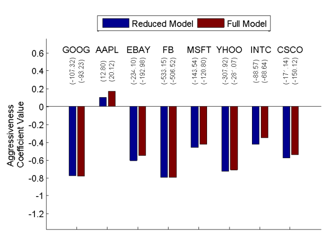
(j) Buy-Side Depth Volume



(k) Sell-Side Depth



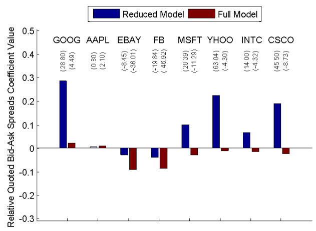
(l) Order Size



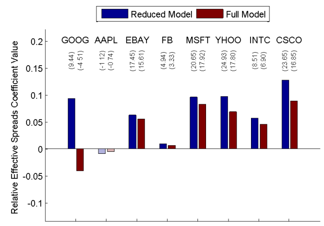
(m) Aggressiveness

The bar graphs in the figure above correspond to key coefficients from a logistic regression of a dummy variable equal to one if a submitted order is attributed to any MPID, and zero if the submitted order is anonymous, on various variables meant to capture market conditions during the interval $[t - \Delta t, t]$, where Δt is the time it takes to observe $N = 20$ consecutive submissions, prior to the submitted order, as well as additional characteristics of the submitted order. Results are presented for buy-side MPID-attributed orders only, and for each stock separately, ordered according to the mean stock price over the sample time period (4-22 November 2013) from highest to lowest. t-statistics are presented in parenthesis below the stock name. Dark blue and red bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light blue and red bars correspond to coefficients that are insignificant at this level. Blue bars corresponds to coefficients of a regression in which only the respective variable is included as a regressor ("Reduced Model"); red bars correspond to coefficients on the variable in which the baseline set of regressors as described in Section 5.2 are included ("Full Model").

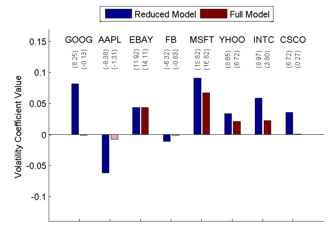
Figure 6: Coefficients from Heckman Correction Regression: Multicollinearity



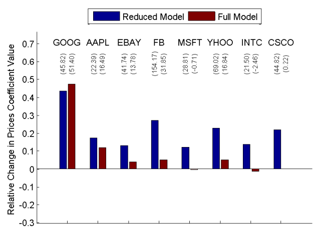
(a) Relative Quoted Bid-Ask Spreads



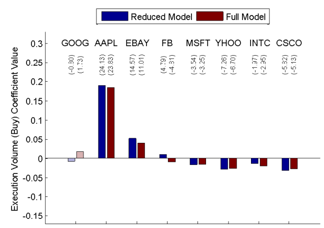
(b) Relative Effective Bid-Ask Spreads



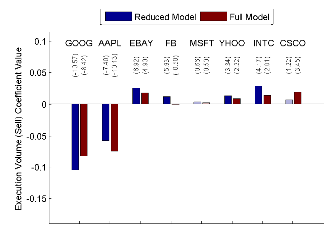
(c) Volatility



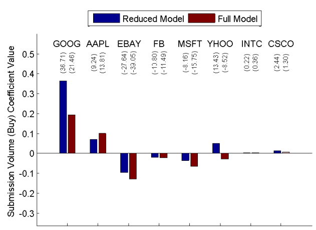
(d) Relative Price Changes



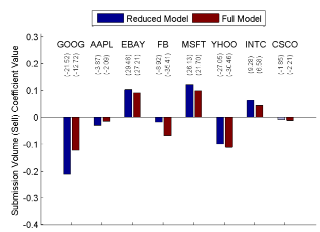
(e) Buy-Side Execution Volume



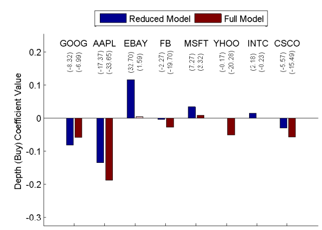
(f) Sell-Side Execution Volume



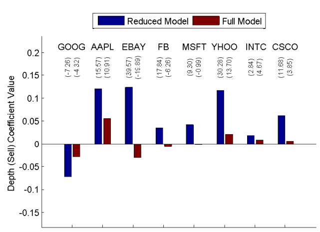
(g) Buy-Side Submission Volume



(h) Sell-Side Submission Volume



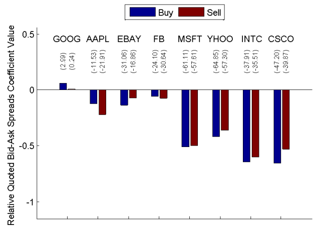
(i) Buy-Side Depth



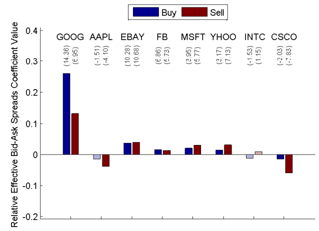
(j) Sell-Side Depth

The bar graphs in the figure above correspond to key coefficients from a regression of measures capturing various dimensions of market quality, on a dummy variable equal to one if a submitted order is attributed to any MPID, and zero otherwise. The dependent variables are calculated during the interval $[t, t + \Delta t]$, where Δt is the time it takes to observe $N = 20$ consecutive submissions, proceeding the submitted order. Results are presented separately for buy-side MPID-attributed orders only, and separately for each firm in the sample, ordered according to the stock's mean stock price during the sample time period, 4-22 November, from highest to lowest. t-statistics are presented in parenthesis above the bar. Dark blue and red bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light blue and red bars correspond to coefficients that are insignificant at this level. Light blue and red bars correspond to coefficients that are insignificant at this level. Blue bars corresponds to coefficients of a regression in which only the respective variable is included as a regressor ("Reduced Model"); red bars correspond to coefficients on the variable in which the baseline set of regressors as described in Section 5.2 are included ("Full Model").

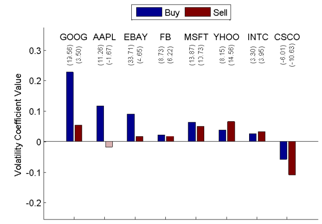
Figure 7: Coefficients from Logistic Regression: Timber Hill, LLC



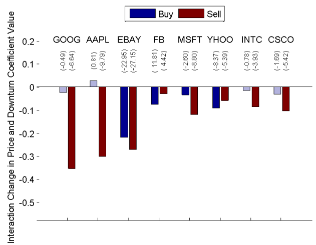
(a) Relative Quoted Bid-Ask Spreads



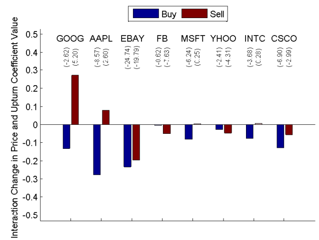
(b) Relative Effective Bid-Ask-Spreads



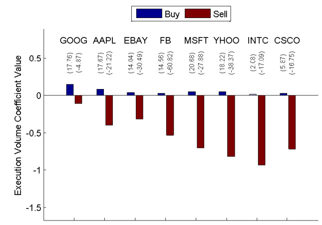
(c) Volatility



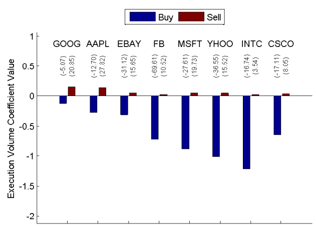
(d) Downward Price Movements



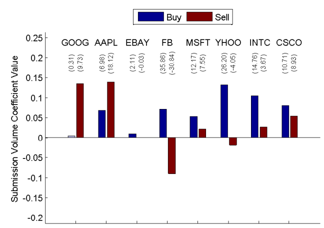
(e) Upward Price Movements



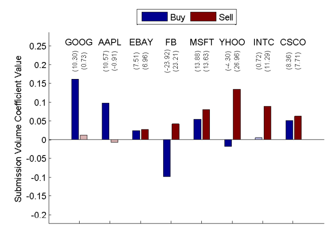
(f) Buy-Side Execution Volume



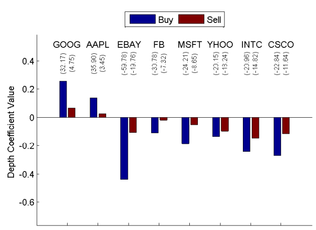
(g) Sell-Side Execution Volume



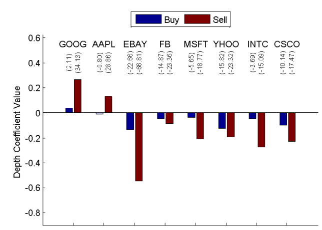
(h) Buy-Side Submission Volume



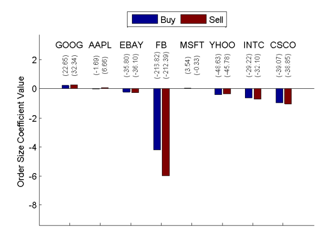
(i) Sell-Side Submission Volume



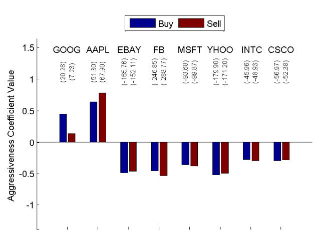
(j) Buy-Side Depth Volume



(k) Sell-Side Depth



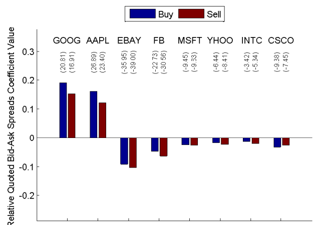
(l) Order Size



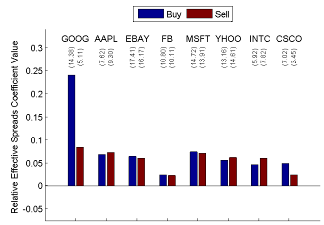
(m) Aggressiveness

The bar graphs in the figure above correspond to key coefficients from a logistic regression of a dummy variable equal to one if a submitted order is attributed to Timber Hill LLC (TMBR), and zero otherwise, on various variables meant to capture market conditions during the interval $[t - \Delta t, t]$, where Δt is the time it takes to observe $N = 20$ consecutive submissions, prior to the submitted order, as well as additional characteristics of the submitted order. Results are presented separately for buy- and sell-side MPID-attributed orders, and for each stock separately, ordered according to the mean stock price over the sample time period (4-22 November 2013) from highest to lowest. t-statistics are presented in parenthesis below the stock name. Dark blue and red bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light blue and red bars correspond to coefficients that are insignificant at this level.

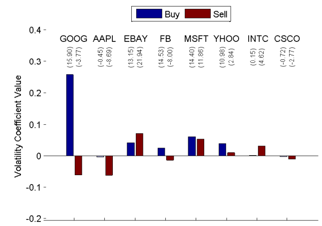
Figure 8: Coefficients from Heckman Correction Regression: Timber Hill, LLC



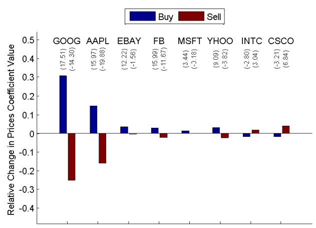
(a) Relative Quoted Bid-Ask Spreads



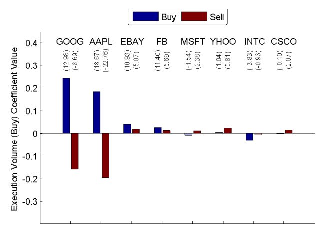
(b) Relative Effective Bid-Ask Spreads



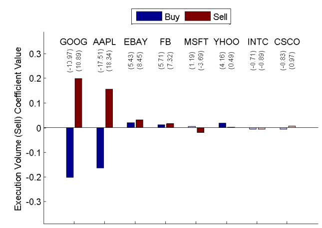
(c) Volatility



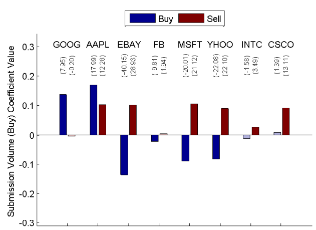
(d) Relative Price Changes



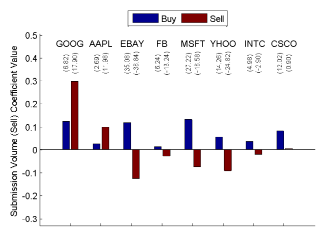
(e) Buy-Side Execution Volume



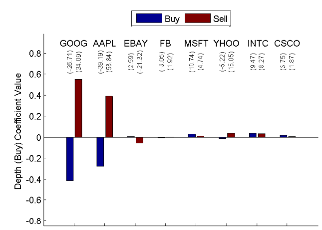
(f) Sell-Side Execution Volume



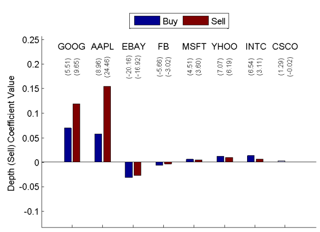
(g) Buy-Side Submission Volume



(h) Sell-Side Submission Volume



(i) Buy-Side Depth



(j) Sell-Side Depth

The bar graphs in the figure above correspond to key coefficients from a regression of measures capturing various dimensions of market quality, on a dummy variable equal to one if a submitted order is attributed to Timber Hill LLC (TMBR), and zero otherwise. The dependent variables are calculated during the interval $[t, t + \Delta t]$, where Δt is the time it takes to observe $N = 20$ consecutive submissions, proceeding the submitted order. Results are presented separately for buy- and sell-side MPID-attributed orders, and separately for each firm in the sample, ordered according to the stock's mean stock price during the sample time period, 4-22 November, from highest to lowest. t-statistics are presented in parenthesis above the bar. Dark blue and red bars correspond to coefficients that are significant at a significance level of $\alpha = 5\%$. Light blue and red bars correspond to coefficients that are insignificant at this level.

References

Akerlof, G. (1995). The market for lemons: Quality uncertainty and the market mechanism. In *Essential Readings in Economics*, pages 175–188. Springer.

Anand, A. and Venkataraman, K. (2016). Market conditions, fragility and the economics of market making. *Journal of Financial Economics (JFE), Forthcoming.*

Barclay, M. J., Hendershott, T., and McCormick, D. T. (2003). Competition among trading venues: Information and trading on electronic communications networks. *The Journal of Finance*, 58(6):2637–2666.

Battalio, R., Ellul, A., and Jennings, R. (2007). Reputation effects in trading on the new york stock exchange. *The Journal of Finance*, 62(3):1243–1271.

Benhami, K. (2006). Liquidity providers valuation of anonymity: The nasdaq market makers evidence. Technical report, Working Paper, University of Toulouse.

Bessembinder, H., Hao, J., and Zheng, K. (2015). Market making contracts, firm value, and the ipo decision. *The Journal of Finance*, 70(5):1997–2028.

Bessembinder, H., Panayides, M., and Venkataraman, K. (2009). Hidden liquidity: an analysis of order exposure strategies in electronic stock markets. *Journal of Financial Economics*, 94(3):361–383.

Biais, B., Hillion, P., and Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the paris bourse. *The Journal of Finance*, 50(5):1655–1689.

Boulatov, A. and George, T. J. (2013). Hidden and displayed liquidity in securities markets with informed liquidity providers. *Review of Financial Studies*, 26(8):2096–2137.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Cao, H. H., Wang, T., and Zhang, H. H. (2005). Model uncertainty, limited market participation, and asset prices. *Review of Financial Studies*, 18(4):1219–1251.

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, 56(2):501–530.

Chordia, T., Roll, R., and Subrahmanyam, A. (2008). Liquidity and market efficiency. *Journal of Financial Economics*, 87(2):249–268.

Chung, K. H. and Chuwonganant, C. (2014). Uncertainty, market structure, and liquidity. *Journal of Financial Economics*, 113(3):476–499.

Comerton-Forde, C., Putniņš, T. J., and Tang, K. M. (2011). Why do traders choose to trade anonymously? *Journal of Financial and Quantitative Analysis*, 46(04):1025–1049.

Cox, D. R. and Hinkley, D. V. (1979). *Theoretical statistics*. CRC Press.

Demsetz, H. (1968). The cost of transacting. *The quarterly journal of economics*, pages 33–53.

Easley, D. and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial economics*, 19(1):69–90.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):27–38.

Foucault, T., Moinas, S., and Theissen, E. (2007). Does anonymity matter in electronic limit order markets? *Review of financial studies*, 20(5):1707–1747.

Foucault, T., Pagano, M., and Röell, A. (2013). *Market liquidity: theory, evidence, and policy.* Oxford University Press.

Frino, A., Johnstone, D., and Zheng, H. (2010). Anonymity, stealth trading, and the information content of broker identity. *Financial Review*, 45(3):501–522.

Gao, S. and Shen, J. (2007). Asymptotic properties of a double penalized maximum likelihood estimator in logistic regression. *Statistics & probability letters*, 77(9):925–930.

Garbade, K. D. and Silber, W. L. (1979). Structural organization of secondary markets: Clearing frequency, dealer activity and liquidity risk. *The Journal of Finance*, 34(3):577–593.

Glosten, L. R. and Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of financial Economics*, 21(1):123–142.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1):71–100.

Grammig, J., Schiereck, D., and Theissen, E. (2001). Knowing me, knowing you:: Trader anonymity and informed trading in parallel markets. *Journal of Financial Markets*, 4(4):385–412.

Grossman, S. J. and Miller, M. H. (1988). Liquidity and market structure. *the Journal of Finance*, 43(3):617–633.

Harris, L. (1996). *Does a large minimum price variation encourage order exposure?*, volume 96. New York Stock Exchange.

Harris, L. (1997). Order exposure and parasitic traders. *University of Southern California working paper.*

Hautsch, N. (2011). *Econometrics of financial high-frequency data.* Springer Science & Business Media.

Hautsch, N. and Jeleskovic, V. (2008). Modelling high-frequency volatility and liquidity using multiplicative error models. *Available at SSRN 1292493.*

Heckman, J. (1979). Sample specification bias as a selection error. *Econometrica*, 47(1):153–162.

Huang, R. D. and Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies*, 10(4):995–1034.

Kaniel, R. and Liu, H. (2006). So what orders do informed traders use?*. *The Journal of Business*, 79(4):1867–1913.

Karam, A. (2012). The value of non-anonymity on the nasdaq. *Working Paper.*

King, G. and Zeng, L. (2001). Logistic regression in rare events data. *Political analysis*, 9(2):137–163.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica: Journal of the Econometric Society*, pages 507–512.

Linnainmaa, J. T. and Saar, G. (2012). Lack of anonymity and the inference from order flow. *Review of Financial Studies*, 25(5):1414–1456.

Malinova, K. and Park, A. (2016). modernmarket makers. *Unpublished Manuscript, University of Toronto.*

Mizrach, B. (2006). Does size mater-liquidity provision by the nasdaq anonymous trading facility. *Competition & Reg. Network Indus.*, 7:471.

Næs, R., Skjeltorp, J. A., and Ødegaard, B. A. (2011). Stock market liquidity and the business cycle. *The Journal of Finance*, 66(1):139–176.

Naik, N. Y., Neuberger, A., and Viswanathan, S. (1999). Trade disclosure regulation in markets with negotiated trades. *Review of Financial Studies*, 12(4):873–900.

O'Hara, M., Saar, G., and Zhong, Z. (2015). Relative tick size and the trading environment. *Available at SSRN 2463360.*

Raman, V., Robe, M. A., and Yadav, P. K. (2014). Electronic market makers, trader anonymity and market fragility. *Trader Anonymity and Market Fragility (May 29, 2014).*

Reiss, P. C. and Werner, I. M. (2005). Anonymity, adverse selection, and the sorting of interdealer trades. *Review of Financial Studies*, 18(2):599–636.

Saporta, V. (1997). Which inter-dealer market prevails? an analysis of inter-dealer trading in opaque markets.

Simaan, Y., Weaver, D. G., and Whitcomb, D. K. (2003). Market maker quotation behavior and pretrade transparency. *The Journal of Finance*, 58(3):1247–1267.

van Kervel, V. and Menkveld, A. J. (2016). High-frequency trading around large institutional orders.

Vayanos, D. (2001). Strategic trading in a dynamic noisy market. *The Journal of Finance*, 56(1):131–171.

Vayanos, D. (2004). Flight to quality, flight to liquidity, and the pricing of risk. Technical report, National Bureau of Economic Research.

Watanabe, M. (2015). Supply ambiguity and market fragility. *Working Paper.*

Zhao, X. and Chung, K. H. (2006). Decimal pricing and information-based trading: Tick size and informational efficiency of asset price. *Journal of Business Finance & Accounting*, 33(5-6):753–766.