In: Liquidity, Interest Rates and Banking
Editors: J. Morrey and A. Guyton, pp. 139-167
ISBN: 978-1-60692-775-5

Chapter 7

AN ANALYSIS OF LIQUIDITY ACROSS MARKETS: EXECUTION COSTS ON THE NYSE VERSUS ELECTRONIC MARKETS

Michael A. Goldstein[1,a], Gang Hu[1,b] and J. Ginger Meng[2,c]
[1]Babson College, Babson Park, MA, USA
[2]Stonehill College, Easton, MA, USA

Abstract

We examine liquidity across different types of markets by using execution costs as a proxy for liquidity. We conduct a thorough analysis of execution costs on the NYSE versus a variety of electronic NASD market centers which also trade NYSE-listed stocks ("Electronic Markets"). We adopt a variety of techniques attempting to correct for the selection bias problem. Unlike current literature, we find that the Electronic Markets offer lower execution costs even after controlling for selection biases. In addition to controlling for selection biases at the sample average level of order difficulty, we also carry out our analysis at different levels of order difficulty, measured by a vector of control variables. Our results are robust under different model specifications. Finally, our what-if analysis shows that the Electronic Markets' (*the NYSE's*) orders would have been worse (*better*) off, had they been executed by the NYSE (*Electronic Markets*). Overall, our results highlight the superiority of the Electronic Markets' liquidity and execution quality.

[a] E-mail address: goldstein@babson.edu. Professor of Finance, Babson College, 223 Tomasso Hall, Babson Park, MA 02457. Phone: 781-239-4402. Fax: 781-239-5004.
[b] E-mail address: ghu@babson.edu. Phone: 781-239-4946. Fax: 781-239-5004. Assistant Professor of Finance, Babson College, 121 Tomasso Hall, Babson Park, MA 02457.
[c] E-mail address: gmeng@stonehill.edu. Phone: 508-565-1986, Assistant Professor of Finance, Department of Business Administration, Stonehill College, 320 Washington Street, Easton, MA 02357.
For helpful comments and discussions, we thank Thomas Chemmanur, Ani Chitaley, Arthur Lewbel, Eric Roiter, Erik Sirri, Zhijie Xiao, and seminar participants at Boston College. Meng acknowledges support from a Fidelity Investments research grant. All remaining errors and omissions are our own.

1. Introduction

It is a world-wide trend that the stock exchanges are changing their traditional trading mechanisms, largely attributed to the competition among exchanges. For example, the modernization of European stock markets since the mid-eighties, including the switch to continuous trading and electronic markets, was spurred by the competitive pressure of London. Electronic markets continue to evolve and improve. Pagano and Schwartz (2003) provide a detailed analysis of one of such improvements: the introduction of electronic closing call auctions at Euronext Paris that lowered execution costs for individual participants and sharpened price discovery for the broad market. In the United States, liquidity in NYSE stocks is moving away from the floor of the NYSE towards electronic markets. From the second half of 2003 until the end of 2005, in the 18 months, the percentage of NYSE-listed shares executed electronically increased from 2.4% to almost 10%. As Regulation NMS is implemented, this percentage will likely increase substantially. In the literature, evidence on the relative advantage of the NYSE versus electronic markets is mixed. For example, Kalay and Portniaguina (2001) document the first voluntary switch of a NYSE firm (Aeroflex) to NASDAQ, and find that the switch announcement resulted in a significantly positive abnormal return, subsequent narrowing of the daily bid-ask spread, and significant increase of the daily volume. In contrast, Pruitt, Van Ness, and Van Ness (2002) find that Aeroflex's switch resulted in economically and statistically significant degradations in key trading metrics such as the bid-ask spread and the number of equity trades. Pruitt, Van Ness, and Van Ness (2002) did not find any observed improvements in trading or quoting behavior as a result of the switch. On December 20, 2005, Charles Schwab (SCHW) effectively dropped its dual listing and decided to list only on the NASDAQ. At the same time, the NYSE is taking a series of actions to move fast from flooring trading to electronic trading, for example, its IPO through the merger with the ARCA Exchange (an electronic exchange that used to be an ECN), and the new Hybrid Market system.[1]

Trading costs and listing fees are the main determinants for the exchanges to attract firms interested in listing and for investors. NASDAQ offers lower fees; however, many previous studies find that the NYSE offers lower execution costs over the years. Researchers such as Christie and Schultz (1994) and Chung, Van Ness, and Van Ness (2001) attribute "implicit collusion" among NASDAQ dealers as the cause. On the other hand, there is another strand of literature showing that electronic markets' execution costs have been declining over the years.[2] Van Ness, Van Ness, and Warr (2005) document that NASDAQ spreads steadily declined from 1993 to 2002. Bessembinder (2003) find that the execution costs on electronic markets are actually lower than the NYSE for market orders. However, researchers attribute this difference to order difficulty differences, i.e., the NYSE have been receiving more difficult orders to execute.

In this paper, we conduct a thorough analysis comparing execution costs between the NYSE and electronic NASD market centers which also trade NYSE-listed stocks ("Electronic

[1] See Davis, Pagano, and Schwartz (2006) for a description and discussion of the new NYSE Hybrid Market.

[2] The term "Electronic Markets" or "Electronic Market Center" refers to the combination of NASDAQ book available to brokers and market makers who are members of NASD, the ECN books available to all brokers, market makers and investors sponsored by brokers. All these electronic markets are voluntarily inter-connected. See Goldstein, Shkilko, Van Ness, and Van Ness (2008) for a recent study on competition in these markets during a similar period across market makers and three major ECNs in NASDAQ stocks.

Markets"), adopting a variety of techniques attempting to correct for the selection bias problem. Conventionally, the point comparison of execution costs is conducted at each market's own average, which is widely criticized because of the selection bias problem, i.e., the difficulties of the orders routed to the two markets could be different. Several recent papers attempt to account for the selection bias problem. For example, Lipson (2005) finds that "once we account for the difference in order flow difficulty, the NYSE is no more costly than other exchanges and much less costly than many." Unlike current literature, we find that after controlling for the selection bias problem, the Electronic Markets still offer lower execution costs. We also shed light on what impact different factors have on execution costs. The results are robust under different model specifications.

We make use of a sample of 1,138 NYSE-listed stocks which are traded on both the NYSE and the Electronic Markets. We start with the conventional simple mean comparison as well as share-weighted mean comparison of execution costs, measured by effective spreads. The univariate results show that the effective spreads on the Electronic Markets are lower than those on the NYSE.

We then proceed to OLS regression analysis controlling for order difficulty. We study a wide range of potential explanatory variables and eventually reach a model with a set of significant explanatory variables explaining effective spreads. The model is cross-sectional and is set up with two dummy variables indicating whether the order is executed on the NYSE or the Electronic Markets. The explanatory variables are all constructed as the deviation from the in-sample mean, so that the coefficients of the two dummy indicators may be interpreted as the conditional mean effective spreads at the mean of the explanatory variables vector. The results confirm that execution costs are lower on the Electronic Markets than on the NYSE.

Next we adopted the two-stage procedure advocated by Heckman (1979) and Maddala (1983). This method involves first estimating a Probit model for the choice of venues, generating two new variables from the Probit estimation, and then including these two new variables in an OLS regression model as controls for selection bias. Our results based on this two-stage procedure confirm the superiority of the Electronic Markets' execution.

We conduct several robustness checks of our results. Since the above OLS and two-stage analysis is effectively a point comparison at the sample mean of control variables, it is worthwhile to check whether the results hold at other locations of the sample domain. To do that, we split our sample into two sub-samples according to the fitted values, which is a measure of order difficulty, and repeat our OLS and two-stage selection analysis in both difficulty sub-samples. The results are similar.

As another extension and robustness check, we try to answer the following question: if the NYSE executed the Electronic Markets' orders, will these orders receive better or worse execution? Similarly, what if the Electronic Markets executed the NYSE's orders? In order to answer the former question, we first estimate a model for the NYSE's orders alone. Then we plug in explanatory variables of the Electronic Markets' orders into the NYSE model. The fitted values represent "as-if" NYSE execution costs of the Electronic Markets' orders. The latter question is similarly answered by first estimating the Electronic Markets model, and then plugging in explanatory variables of the NYSE's orders. Results of the above what-if analysis suggest that the Electronic Markets' (*the NYSE's*) orders would have been worse (*better*) off, had they been executed by the NYSE (*Electronic Markets*). Therefore, our results clearly show that the Electronic Markets offer better execution quality than the NYSE.

Given our results of lower execution costs on the Electronic Markets, one might find it puzzling that the NYSE is still dominant in the market place even though it is not fully electronic. We note that although execution costs and listing fees are important considerations in a firm's listing decision, they are by no means the only factors a firm might consider. Chemmanur and Fulghieri (2006) develop a theoretical model of firms' listing decisions in an environment of competition and co-operation among exchanges with endogenous listing standards. In their model, reputation, asymmetric information, and investors' ability to produce information are the main concerns in a firm's listing decision making process. Our results are thus consistent with their theoretical analysis: execution costs alone do not drive a firm's listing decision.

The rest of the paper is organized as follows. Section 2 describes the data and sample selection procedures. Section 3 presents univariate analysis. Section 4 presents OLS and two-stage selection model analysis. Section 5 presents difficulty sub-sample analysis. Section 6 contains an extension answering the what-if question. Section 7 concludes.

2. Data and Sample Selection

2.1. Data

Execution costs are often computed based on trade level data such as Trade and Quote database (TAQ), disseminated by the NYSE. The drawback is that order direction, order size, and order arrival time are not observable and must be estimated using approximation methods. On November 17, 2000, the SEC adopted Exchange Act Rule 11Ac1-5 (the "Dash-5 reports"). Regarding the purpose of compiling Dash-5 reports, the SEC states "*one of the primary objectives of the Rule is to generate statistical measures of execution quality that provides a fair and useful basis for comparisons among different market centers*." The main advantages of the Dash-5 reports are that the order direction is known, the benchmark price is the best quote at order receipt time, the time between order receipt and execution is reported, and they provide order volume and execution quality for all market centers individually. Dash-5 reports have drawbacks as well. For example, they only report aggregate monthly averages. Due to the advantages of Dash-5 reports, several recent academic studies use Dash-5 data to answer related research questions (see, e.g., Bessembinder (2003), Lipson (2005), and Nguyen, Van Ness, and Van Ness (2005)). The Dash-5 reports provide the most relevant publicly available data for our analysis. Our Dash-5 data are from Transaction Audit Group, Inc (www.tagaudit.com).

Exchange Act Rule 11Ac1-5 mandates that all markets in the United States report order data and regular-way execution data received for all stock orders of less than 10,000 shares from both individual and institutional investors. Dash-5 reports do not include any order for which the customer requests special handling, such as orders to be executed at the market opening price or closing price, orders submitted with stop prices, orders to be executed only at their full size, orders to be executed on a particular type of tick or bid, orders submitted on a "not held" basis, orders for other than regular settlement, and orders to be executed at prices unrelated to the market price of the security at the time of execution. Dash-5 reports provide data by order size/order type/security/market center/month/participant. The orders are divided into four size categories: 100 to 499 shares, 500 to 1999 shares, 2000 to 4999 shares, and 5000 or greater shares. Order size can be viewed as a measure of order difficulty. Executions for large orders are generally expected to be more costly than for small orders.

Dash 5 reports include market orders and marketable limit orders. We follow Boehmer (2005) and focus on market orders only. As pointed out by Boehmer (2005), results for marketable limit orders based on Dash-5 reports are hard to interpret for at least the following four reasons. "First, because Dash 5 reports do not include information on the opportunity cost of non-execution, ex post execution costs for marketable limits understate their true cost. Consequently, estimates for marketable limits would not be comparable to those in SEC (2001), which uses an ex post adjustment for unfilled marketable limits. This analysis cannot be replicated using Dash 5 data, because they include only monthly aggregates. Second, the time-to-execution for this order type is censored, because cancelled and expired orders are not considered in the computation. Third, summary statistics on speed are dominated by orders that happen to be submitted as the market moves away and, therefore, do not execute immediately. Finally, usage of marketable limit orders differs systematically across markets. All NYSE specialists accept market orders, but some Nasdaq market centers do not. For example, some marketable limits reported by Island, which does not accept market orders, are probably functionally equivalent to market orders." See Peterson and Sirri (2002) for an analysis of the two order types.

2.2. Sample Selection

The Center for Research in Security Prices (CRSP) database is used to construct the sample of stocks and several control variables based on stock characteristics, such as price, volume, shares outstanding, and return. We select NYSE-listed stocks from the CRSP database. The sample selection criteria are similar to those used in SEC (2001). Our Dash-5 data are from January to December 2003. We start with all the 2,557 NYSE listed securities as of December 31, 2002. From this list, we eliminated dual classes, foreign-incorporated securities, ADRS, REITS, Certificates, SNIs, Units, Closed End Funds, etc., leaving us with 1,329 NYSE common stocks. This list was further reduced to 1,138 NYSE securities after removing securities whose daily trading volume was less than \$20,000, whose average closing price was less than \$3, which switched exchanges, or which had missing data, or for which data was not available in Dash-5 reports. The detailed sample selection procedure is shown in the Appendix, Table A1. The sample was then merged with Dash-5 reports data. We further apply a filter to exclude outliers: we exclude orders where the effective spread or the quoted spread is equal to or larger than half of the trading price. In our data, there are 7 NYSE specialist firms and 32 electronic market centers. The list of market centers are in the Appendix, Table A2.

3. Univariate Analysis

3.1. Measures of Execution Costs

The *quoted spread* is defined as the bid ask difference, which reflects market and order flow conditions at the time of order arrival. The *effective spread,* first developed by Blume and

Goldstein (1992) and Petersen and Fialkowski (1994), is defined, for buy orders, as double the amount of the difference between the execution price and the midpoint of the consolidated best bid and offer at the time of order receipt and, for sell orders, as double the amount of difference between the midpoint of the consolidated best bid and offer at the time of order receipt and the execution price. Dash-5 data calculate at a record level, the share-weighted average of effective spreads for order executions in the month. The *realized spread* is defined, for buy orders, as double the amount of difference between the execution price and the midpoint of the consolidated best bid and offer five minutes after the time of order execution and, for sell orders, and double the amount of difference between the midpoint of the consolidated best bid and offer five minutes after the time of order execution and the execution price.

As noted in Blume and Goldstein (1992) and Petersen and Fialkowski (1994), effective spreads are a better measure of execution costs than quoted spreads, because orders do not always execute exactly at the bid or offer price. The effective spread takes this into account by incorporating any price improvement or dis-improvement that an order may receive. The effective spread calculates how much above the midpoint price you paid on a buy order and how much below the midpoint price you received on a sell order. While price improvement is a good tool for measuring execution quality, effective spread captures both how often, and also by how much, a broker-dealer price improves trades. Therefore, the effective spread can be interpreted as the total price impact of the trade, a measure of the non-commission, out-of-pocket cost of a trader.

Effective spreads can be decomposed into two parts: realized spread and the information component or price impact, which is the difference of the bid-ask midpoint five minutes later and that at the time of order receipt. The information component can measure the extent to which “informed” and “uninformed” orders are routed to different market centers. Informed orders are those submitted by persons with better information than is generally available in the market. They therefore represent a substantial risk to liquidity providers that take the other side of these informed trades. In contrast, order submitted by persons without an information advantage (often small orders) present less risk to liquidity providers and in theory should receive the most favorable effective spreads available in the market.

The smaller the average realized spread, the more market prices have moved adversely to the market cent’s liquidity providers after the order was executed, which shrinks the spread “realized” by the liquidity providers. In other words, a low average realized spread indicates that the market center was providing liquidity even though prices where moving against it for reasons such as news or market volatility. Spreads are not the perfect measure of trading costs. However, they are simple to measure, readily available, and are usually reasonable indicators of actual trading costs.

3.2. Univariate Results

Table 1 presents summary statistics of execution quality measures in the two markets, the NYSE and the Electronic Markets. It reports the median, simple average, as well as share-weighted average (aggregated across all stocks traded at each market over the 12 month period) of the effective spread, quoted spread, realized spread, information component (effective spread less realized spread), and execution speed.

Table 1. Summary Statistics

This table presents summary statistics from Dash-5 database for the 1,138 selected securities. Summary statistics are computed separately for the NYSE and the Electronic Markets (EM), and are further divided into four size categories: very small (100~499 shares), small (500~1,999 shares), medium (2,000~4,999 shares), and large (5,000~9,999 shares). Monthly shares ordered, monthly shares executed, and monthly number of orders are reported in millions. Average effective spread (AES), average quoted spread (AQS), average realized spread (ARS), and average information component (INFO, the difference of AES and ARS) are reported in cents. Average speed (SPEED) is also reported, in seconds. Median, mean, and share-weighted mean are reported for the above variables.

		NYSE					Electronic Markets (EM)				
		All	**Very Small**	**Small**	**Medium**	**Large**	**All**	**Very Small**	**Small**	**Medium**	**Large**
Shares Ordered (M)		5,517.10	1,620.73	2,188.51	1,150.60	557.27	769.94	185.81	332.42	171.74	79.97
Shares Executed(M)		5,455.13	1,602.38	2,161.68	1,139.94	551.13	743.28	181.71	324.45	163.37	73.74
Number of Orders (M)		11.98	8.83	2.63	0.42	0.09	1.43	0.95	0.40	0.07	0.01
	Median	3.53	2.39	3.36	6.12	8.76	2.33	1.84	2.33	3.41	2.39
AES (cents)	Mean	5.95	2.80	4.70	8.31	11.63	3.94	2.30	3.36	6.13	2.80
	Wtd. Mean	3.03	2.38	2.83	3.58	4.59	2.16	1.47	1.85	2.80	2.38
	Median	8.04	6.62	7.57	10.33	13.03	5.04	4.59	5.00	5.98	6.20
AQS (cents)	Mean	10.81	8.02	9.51	12.96	16.18	9.47	8.31	8.76	11.54	12.39
	Wtd. Mean	7.28	6.70	7.16	7.65	8.65	5.88	4.78	5.63	6.74	7.81
	Median	1.00	0.40	0.79	2.04	3.49	1.27	1.20	1.15	1.56	1.95
ARS (cents)	Mean	2.52	1.09	2.52	2.84	5.11	1.69	1.23	1.25	2.59	3.18
	Wtd. Mean	10.40	6.53	20.40	1.07	1.76	1.07	0.97	0.85	1.33	1.73
	Median	2.71	1.97	2.68	3.92	4.57	1.00	0.51	1.11	2.00	2.18
INFO (cents)	Mean	3.43	1.72	2.18	5.46	6.51	2.25	1.07	2.11	3.54	4.72
	Wtd. Mean	-7.37	-4.14	-17.58	2.51	2.82	1.09	0.50	1.00	1.47	2.07
	Median	17.05	14.33	15.51	19.33	23.80	13.21	7.91	12.00	19.70	28.80
SPEED	Mean	19.43	16.40	17.33	21.41	27.26	31.07	21.73	24.26	38.55	76.50
	Wtd. Mean	17.14	16.03	16.61	18.02	20.66	14.37	5.51	9.92	22.88	36.96

Statistics are averaged across all categories and aggregated up to the markets level (either the NYSE or the Electronic Markets (*EM*)). This method, while simple, may be distorted by variations in executed volume among market centers. Share-weighted average statistics are also provide to account for share volume differences. We also report monthly shares ordered, shares executed, and number of orders. We further examine the above variables in four order size categories: very small (100~499 shares), small (500~1,999 shares), medium (2,000~4,999 shares) and large (5,000~9,999 shares).

NYSE orders seem to have higher average effective and quoted spreads. The overall effective spread reported by the Electronic Markets has a simple mean of 3.94 cents versus 5.95 cents for the NYSE, and a share-weighted mean of 2.16 cents versus 3.03 cents. In addition, the Electronic Markets' effective spread is lower across all four size categories.

One explanation for why the Electronic Markets can offer lower effective spreads than the NYSE for the same NYSE stocks is that the Electronic Markets compete and attract "easy orders." This selection bias could cause the difference in effective spreads. Before formally accounting for this selection bias in a multivariate selection model framework, we first sort the data on quoted spread, an important variable since it reflects the market condition at the time of the order. We segment the data into 8 ranges of quoted spread. Then we calculate the share-weighted effective spread for all the NYSE records in each of the 8 quoted spread ranges. We do the same for all records in the Electronic Markets. The results are shown in Table 2. The Electronic Markets offer lower average effective spreads in 7 out of the 8 ranges of quoted spreads. The NYSE only offers marginally lower effective spreads in the lowest quoted spread range (0~4 cents), 1.17 cent for the NYSE versus 1.23 for the Electronic Markets. Note that this range also contains relatively less number of symbols and shares executed. In summary, results in Table 2 suggest that the Electronic Markets seem to be able to offer lower effective spreads than the NYSE even after controlling for the differences in quoted spreads across these two markets.

Table 2. Quoted Spread Bins

This table presents results by splitting the data according to quoted spread bins. The number of symbols, shares executed in millions, and share-weighted effective spreads in each range for the NYSE versus the Electronic Markets (EM) are reported.

	Number of Symbols		**Shares Executed (M)**		**AES (cents)**	
Quoted Spread	**NYSE**	**EM**	**NYSE**	**EM**	**NYSE**	**EM**
0~4 cents	16	39	4,800	2,695	1.17	1.23
4~6 cents	130	212	17,053	3,318	1.94	1.74
6~8 cents	276	228	23,871	1,556	2.82	2.53
8~10 cents	260	204	10,850	701	3.90	3.69
10~12 cents	181	126	5,642	220	4.91	4.73
12~14 cents	114	108	2,221	180	6.14	5.09
14~16 cents	59	80	614	130	7.29	5.54
>= 16 cents	102	141	423	132	10.07	7.72

4. Regression and Selection Model Analysis

4.1. Factors Affecting Execution Costs

We consider a wide range of variables explaining execution costs based on related microstructure literature. Many of these variables have been used in one or more of previous studies, such as Bessembinder (2003), Boehmer (2005), Lipson (2005), and Nguyen, Van Ness, and Van Ness (2005). Some of these variables are cost-based, while others are only reflective. For example, the quoted spread directly represents the financial loss a trader incurs from a particular transaction. On the other hand, though the trading volume indicates whether a particular stock is liquid or not, it does not show us how costly it is to actually trade the stock. The explanatory variables can be classified into two groups: order-specific measures and stock-specific measures. Order-specific measures capture the nature and difficulty of different orders, while stock-specific measures capture the characteristics of stocks traded and are the same across different orders within the same stock. We further divide the stock-specific measures into liquidity measures and volatility measures. The detailed definitions of these variables are as follows:

Order-specific Measures:

Log(num.ord): The natural logarithm of the number of orders.

AQS: Average quoted spread. This variable is included as a measure of market conditions at order time. It can be viewed as a cost-based liquidity measure because it examines the financial loss a trader incurs from a particular transaction. It is highly correlated with the average effective spread.

INFO: It is defined as the difference between the effective spread and the realized spread. Therefore it is the mirror image of realized spread. The smaller the average realized spread, the more market prices have moved adversely to the market center's liquidity providers after the order was executed, which shrinks the spread "realized" by the liquidity providers. In other words, a low average realized spread indicates that the market center was providing liquidity even though prices were moving against it for reasons such as news or market volatility.

SPEED: Another dimension of execution quality measures beyond trading costs. There is a trade-off between the urgency and the absolute cost. Therefore one would expect that the faster the speed, the higher the cost. SEC (2001) used a five-day period in June and found that for smaller orders (orders below 5,000 shares), the NYSE execution costs are below NASDAQ costs, but NASDAQ orders generally execute faster. Boehmer (2005) extended this part of study and found that small orders (below 2,000 shares) execute at lower cost on the NYSE, but substantially faster on NASDAQ. However this results reverses for larger orders (between 2,000 and 9,999 shares). These execute more cheaply on NASDAQ, but faster on the NYSE.

Log(Ord.Sz) and Ord.Sz/Vol: Ord.Sz/Vol is the standardized order size, calculated as the order size for the security at the specific markets, versus its total trading volume in the last month of 2002. It is intuitive that larger orders should be more difficult to execute (due to pure liquidity reasons, regardless of information content) and also should contain more information. We expect both reasons to cause a positive relationship between order size and the effective spread. Boehmer (2005) describes a specific example for why larger NYSE orders should contain more information. He reasons that traders who have either no private information or whose information is sufficiently long-lived often use floor brokers to work large orders. This involves delegating control over the actual trading decisions to a floor broker, who then seeks favorable (partial) executions until the order is filled. An informed trader with short-lived information cannot afford to use this option because it is slow, and the trader risks that others discover the same information before the orders are filled. For the same reason, small NYSE orders are not useful for informed traders because they are executed sequentially. One would thus expect informed traders (or their agents) to submit large orders directly to the specialist.

Stock-specific Measures:

Liquidity measures: Illiquidity reflects the impact of order flow on price: the discount that a seller concedes or the premium that a buyer pays when executing a market order, which results from adverse selection costs and inventory costs.

MCAP and MCAP Rank: MCAP is the market capitalization, calculated as the product of price and shares outstanding. MCAP Rank is the market capitalization rank (1~20) based on Fama-French NYSE Breakpoints. They are common proxies for liquidity since a larger stock issue has smaller price impact for a given order flow and a smaller bid–ask spread: large firms are more liquid.

1/PRC: The inverse of price. The higher this factor, the more liquid the order.

Turnover: The volume in the stock divided by the number of shares outstanding.

ADV: Average dollar volume (price times share volume) in the fourth quarter of 2002. High volume levels may indicate that a particular security is very liquid. However, it does not tell us how costly it is to actually trade the security.

CBMA: Gibbs estimate of transaction cost, c, from Basic Market-Adjusted Model from Hasbrouck (2006). It is a daily liquidity proxy.

$$CBMA = \begin{cases} \sqrt{-\operatorname{cov}(r_t, r_{t-1})} & \text{if } \operatorname{cov}(r_t, r_{t-1}) < 0 \\ 0 & \text{otherwise} \end{cases} . \tag{1}$$

I1: Amihud's (2002) illiquidity measure, calculated as the average daily ratio over year of the absolute value of daily return divided by the daily trading volume in millions of dollars.

$$I1 = 1{,}000{,}000 \times \frac{abs(ret)}{prc \times vol} . \tag{2}$$

As can be seen from the above definition, Amihud's illiquidity measure can be interpreted as the daily price response (sensitivity) associated with one dollar of trading volume, thus serving as a rough measure of price impact. Amihud (2002) shows that illiquid stocks are more difficult to trade. We expect orders in more illiquid stocks more likely to be submitted to NYSE rather than the Electronic markets.

PSGAMMA: Pastor-Stambaugh (2003) gamma. However, the authors caution against its use as a liquidity measure for individual securities, noting the large sample error in the individual estimates.

Volatility Measures:

VOLA: The standard deviation of daily returns in the fourth quarter of 2002. The specification of this variable is slightly different in other papers. For example, Lipson (2005) defines it as the standard deviation of daily trade-weighted prices.

RR: The average daily relative price range during the fourth quarter of 2002. Daily relative price range is defined as the daily range divided by the closing price. This is an intra-day measure. Compare to the other volatility measure *VOLA, RR* does not rely on stationary assumptions over the period of time needed to calculation *VOLA*.

4.2. OLS Regressions

The OLS regression method is specified in the following model, across different stocks:

$$AES_{ip} = \alpha_1 NYSE_{ip} + \alpha_2 NON.NYSE_{ip} + \alpha_3 X_{ip} + e_{ip}, \tag{3}$$

where AES_{ip} is the mean effective spread for stocks *i* at market *p*, $p \in (N, EM)$. $NYSE_{ip}$ is a dummy variable which equals one if the market is *N*, zero if the market is *EM*. $NON.NYSE_{ip}$ is another dummy variable which equals zero if the market is *N*, and one if the market is *EM*. X_{ip} is a vector of explanatory variables selected from the list of variables discussed previously. All X_{ip} are measured as deviations from their own sample cross-sectional mean. This way, when the X_{ip} are excluded from the regression, the coefficients estimates α_1 and α_2 produce the simple cross-sectional mean effective spreads of the two markets. When the control variables X_{ip} are included in the regression, the coefficient estimates on the two dummy variables reveal conditional mean execution costs on the NYSE versus the Electronic Markets, evaluated at the mean of the variables that comprise the X_{ip} vector.

We start our OLS regression analysis with specifications similar to Bessembinder (2003), and then add other potentially related factors discussed previously. The results are shown in Table 3. We run different specifications of the following regression model:

$$\begin{aligned} &AES_{ip} = \alpha_1 NYSE_{ip} + \alpha_2 EM_{ip} + \alpha_3 \log(NUM.ORD_i) + \alpha_4 INFO_{ip} \\ &+\alpha_5 \frac{1}{PRC_i} + \alpha_6 \frac{ORD.SZ_{ip}}{VOL_i} + \alpha_7 SPEED_{ip} + \alpha_8 ADV_i + \alpha_9 TURNOVER_i \ . \\ &+\alpha_{10} MCAP.RANK_i + \alpha_{11} VOLA_i + \alpha_{12} RR_i + e_{ip} \end{aligned} \tag{4}$$

We start by running an OLS regression on the two dummy variables: one for the NYSE and one for the Electronic Markets (therefore there is no intercept term). The coefficients on the two dummy variables represent the unconditional cross-sectional simple averages for these two markets. Specifically, the effective spread is 7.65 cents for the NYSE and 4.92 cents for the Electronic Markets. In other words, the effective spread for the Electronic Markets is 2.73 cents (or 36%) lower than the NYSE.

Table 3. OLS Regressions

This table presents OLS regression results. The dependent variable is average effective spread (AES). NYSE is a dummy variable that equals 1 for orders executed by the NYSE and 0 otherwise. EM is a dummy variable that equals 1 for orders executed by the Electronic Markets and 0 otherwise. There is no intercept term since both NYSE and EM are included in the regression.

	NYSE	EM	Log(NUM.ORD)	INFO	1/PRC	ORD.SZ/VOL	SPEED	ADV	TURNOVER	MCAP.RANK	VOLA	RR	R^2
	α_4	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	
Model 1	7.65	4.92											70.45%
(t-stat)	59.20***	37.26***											
Model 2	7.68	5.66	-0.55	76.19	-15.04	7.86							95.15%
(t-stat)	99.61***	66.39***	-14.77***	52.23***	-14.72***	13.06***							
Model 3	7.70	5.65	-0.54	76.03	-15.05	7.85	0.01						95.12%
(t-stat)	96.47***	65.92***	-14.60***	51.80***	-14.73***	13.04***	0.97						
Model 4	7.24	6.03	-0.20	72.75	-21.69	7.67	0.01	0.01	-21.17	-0.14			95.26%
(t-stat)	68.86***	57.08***	-3.23***	47.61***	-16.28***	12.49***	0.48	-0.09	-2.75***	-7.87***			
Model 5	7.26	6.01	-0.22	72.69	-22.09	7.78	0.01	0.01	-19.28	-0.13	-10.38	10.03	95.27%
(t-stat)	67.68***	55.7.9***	-3.37***	47.18***	-12.92***	12.10***	0.51	0.04	-2.14**	-7.73***	-1.25	1.17	
Model 6	7.23	6.04	-0.20	72.89	-21.30	7.63			-18.83	-0.14	-2.21		95.26%
(t-stat)	76.27***	62.26***	-3.60***	47.61***	-13.69***	12.39***			-2.10**	-7.92***	-0.49		

In the second model, we add four control variables, as in Bessembinder (2003): logarithm of number of orders, information component, inverse price, and standardized order size. We note three things here: first, after adding these control variables, the effective spread for the NYSE almost remain unchanged, at 7.68 cents. On the other hand, the effective spread for the Electronic Markets increases dramatically, from 4.92 cents to 5.66 cents, causing the difference between the two markets to narrow to 2.02 cents (or 26%). This means that the Electronic Markets received relatively easier orders than did the NYSE, hence controlling for difficulty reduces the cost advantage shown by unconditional results. These results highlight the importance of controlling for the relative difficulty of the orders received by different markets before drawing inferences. Second, the slope coefficients are generally consistent with those reported in prior research (see, e.g., Bessembinder (2003)). The average effective spread decreases with the trading activity as measured by total orders in the stock, increases with average information component, decreases with the inverse share price (or increase with share price), and increases with average order size. Each coefficient estimate is highly significant. Third, the R^2 increases to 95.15% from the first regression's 70.45%.

In the third model, we further add *SPEED* to the regression. Though it is considered to be the other important dimension of execution quality, it does not seem to have marginal explanatory power for effective spread. The coefficient for *SPEED* is not significantly different from zero. The coefficients for the two dummy variables and the R^2 remain almost unchanged.

In the fourth model we further add *ADV, TURNOVER*, and *MCAP.RANK*, trying to capture effects of illiquidity on effective spreads. Coefficients on *TURNOVER* and *MCAP.RANK* are significant and negative, which is intuitive. High turnover and large market cap securities are more liquid, and therefore should have lower execution costs. The coefficient on *ADV* is not significant. The conditional mean effective spread for the NYSE decreases to 7.24 cents, while that for the Electronic Markets increases to 6.03 cents.

The fifth model is the full specification. We further add two volatility variables *VOLA* and *RR*. Neither of them turns out to be significant. The conditional effective spreads for the NYSE and the Electronic Markets are almost unchanged.

The sixth model is our selected model, chosen mainly based on the statistical significance of different factors in previous models. It includes market condition factors such as number of orders, order size; and information component, security illiquidity proxies such as price, turnover, and market cap rank, and a volatility measure, VOLA (though it is the only insignificant factor). The conditional mean effective spread for the NYSE is 7.23 cents, versus 6.04 cents for the Electronic Markets, which represents a difference of 1.19 cents or 16%. Overall, our OLS regression results show that controlling for the above factors narrows the difference in conditional mean effective spread for the NYSE versus the Electronic Markets. However, the Electronic Markets still outperform the NYSE.

4.3. Two-Stage Selection Model

The two-stage procedure to control for selection bias follows the work by Heckman (1979) and Maddala (1983). Effective spreads are modeled for the NYSE (*N*) and the Electronic Markets (*EM*) as:

$$\begin{aligned} ES_{i,N} &= \beta_N ' X_i + \varepsilon_{i,N}, \\ ES_{i,EM} &= \beta_{EM} ' X_i + \varepsilon_{i,EM}, \end{aligned} \tag{5}$$

where X_i is a vector of conditioning variables for each security i, β is a vector of parameters to be estimated, and the ε's are error terms. We assume the difference in effective spreads across the two markets is a factor that determines the market selected by a trader for that stock. The difference in expected effective spreads is

$$\begin{aligned} y_i^* &= E\left[ES_{i,N} - ES_{i,EM} \mid X_i \right] \\ &= \left(\beta_N ' - \beta_{EM} ' \right) X_i + \varepsilon_{i,N} - \varepsilon_{i,EM} \,. \\ &= \gamma ' X_i + \zeta_i \end{aligned} \tag{6}$$

In this model, a trader of stock i chooses to trade the order in the NYSE if $y_i^* \le 0$, i.e., the NYSE trade is expected to be less costly. The order submission rule for a security a traders wishes to trade is

$$\begin{aligned} y_i &= 1, \quad if \quad y_i^* \le 0; \\ y_i &= 0, \quad if \quad y_i^* > 0; \end{aligned} \tag{7}$$

with $y_i = 1$ indicates the NYSE and $y_i = 0$ indicates the Electronic Markets. We will use several independent variables discussed previously to model the choice which market to choose from in a Probit framework. Then, the probability a trader chooses the NYSE is estimated, which is,

$$\Pr(y_i = 1) = \Phi(\gamma ' X_i), \tag{8}$$

where Φ is the cumulative distribution function of the standard normal. The next step is to multiply the parameter estimates from the Probit, γ, with the complementary set of observations, X, to estimate the probability of choosing the NYSE ($\Phi(\gamma' X)$).

In the second stage of this method, the Probit probability estimates are used to control for the selection bias. Because an stock i is traded on the NYSE only when $y_i^* \le 0$, the error term $\varepsilon_{i,N}$ does not have a zero mean, conditional on being NYSE. The conditional expected execution costs for the NYSE and the Electronic Markets are denoted as:

$$\begin{aligned} E\left[ES_{i,N} \mid y^* \le 0 \right] &= \beta_N ' X_i + \alpha_N \frac{\phi(\gamma ' X_i)}{\Phi(\gamma ' X_i)}, \\ E\left[ES_{i,EM} \mid y^* > 0 \right] &= \beta_{EM} ' X_i + \alpha_{EM} \frac{-\phi(\gamma ' X_i)}{1 - \Phi(\gamma ' X_i)}; \end{aligned} \tag{9}$$

where $\phi(\gamma ' X_i)$ and $\Phi(\gamma ' X_i)$ are the density and cumulative distribution function of the standard normal evaluated at $\gamma ' X_i$, respectively. $\alpha = \mathrm{cov}(\varepsilon_i, \zeta_i)$. In this methodology, estimating the second stage equation by OLS provides consistent estimates of the parameters.

$\lambda_1 = \dfrac{\phi(\gamma' X_i)}{\Phi(\gamma' X_i)}$ is the "Inverse Mills Ratio". It is monotonically decreasing in the probability that an order will be routed to the NYSE. $\lambda_2 = -\dfrac{\phi(\gamma' X_i)}{1-\Phi(\gamma' X_i)}$. The Heckman method can detect the selection bias in a rather straightforward fashion. Potentially, the parameters from the Heckman method can also be used to examine the trade-offs in the market selection strategies.

4.4. Selection Model Results

The full specification for the first stage Probit regression is as follows:

$$\begin{aligned}\Pr(NYSE_{ip}=1) &= g(\beta_0+\beta_1\log(MCAP_i)+\beta_2\log(VOL_i)+\beta_3\log(ORD.SZ_{ip}) \\ &+\beta_4 Info_{ip}+\beta_5 AQS_{ip}+\beta_6 SPEED_{ip}+\beta_7 RR_i+\beta_8 VOLA_i+\beta_9 CBMA_i+\beta_{10} I1 \\ &+\beta_{11} PSGAMMA+\varepsilon_{ip})\end{aligned} \tag{10}$$

In the second stage OLS regression correcting for selection bias, we use the full model and the model selected previously. The specification of the full model is as follows:

$$\begin{aligned}AES_{ip} &= \alpha_1 NYSE_{ip}+\alpha_2 EM_{ip}+\alpha_3(\lambda_1 NYSE_{ip})+\alpha_4(\lambda_2 EM_{ip}) \\ &+\alpha_5\log(NUM.ORD_i)+\alpha_6 INFO_{ip}+\alpha_7\frac{1}{PRC_i}+\alpha_8\frac{ORD.SZ_{ip}}{VOL_i}+\alpha_9 SPEED_{ip} \\ &+\alpha_{10}ADV_i+\alpha_{11}TURNOVER_i+\alpha_{12}MCAP.RANK_i+\alpha_{13}VOLA_i+\alpha_{14}RR_i+e_{ip}\end{aligned} \tag{11}$$

In the literature there appears to be some inconsistency as to whether only λ_1, or both λ_1 and λ_2 should be included in the second stage regression. We did both. To adapt to our specific model where the two parts of sample (the NYSE and the Electronic Markets) are combined into one model, indicated by two dummy variables, we create two new variables $\lambda_1 NYSE$ and $\lambda_2 EM$. Note that these two variables are highly correlated with each other.

Table 4 Panel A presents first stage Probit regression results. We start with a model similar to that in Bessembinder (2003). The independent variables include market capitalization, order volume, order size, information component, and quoted spread. All coefficients are significant at least at the 95% level, confirming the presence of systematic selection biases in order routing. The coefficient on order volume is significantly negative, indicating that actively traded stocks are more likely to be executed in the Electronic Markets. Coefficients on market capitalization, order size, information component, and quoted spread are all significantly positive, suggesting that orders for larger stocks, with larger order sizes, containing more information, and with worse market condition, tend to be executed on the NYSE. We then further add variables including *SPEED, RR, VOLA, CBMA, I1* and *PSGAMMA* to the Probit model. These variables comprise volatility measures and illiquidity measures.

Table 4. Two-Stage Selection Model

This table presents selection model results. Panel A presents the first stage Probit regressions for the likelihood of an order being executed by the NYSE. Panel B presents the second stage OLS regression of conditional average effective spreads.

Panel A. First Stage Probit Regressions for the Likelihood of an Order Being Executed by the NYSE

	Log (MCAP)	Log (VOL)	Log (ORD.SZ)	INFO	AQS	SPEED	RR	VOLA	CBMA	I1	PSGAMMA		
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	Pseudo R^2	Prob > χ^2
Model 1	1.04	-2.29	16.99	28.56	6.67							89.57%	0.00
(z-stat)	6.87***	-10.17***	17.45***	8.57***	2.34**								
Model 2	0.90	-2.32	18.94	32.03	9.82	-0.10	-49.07	54.19	-117.63	3.58	-56665.03	93.10%	0.00
(z-stat)	3.14***	-6.07***	12.79***	7.95***	2.62***	-8.23***	-2.24***	2.59***	-2.39***	1.59	-2.01**		

Panel B. Second Stage OLS Regressions of Conditional Average Effective Spreads

	NYSE	EM	λ_1NYSE	λ_2EM	Log(NUM. ORD)	INFO	1/PRC	ORD.SZ /VOL	SPEED	ADV	TURNOVER	MCAP. RANK	VOLA	RR	R^2
	α_4	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	
Model 1	7.14	4.92	-0.74	-0.02											72.42%
(t-stat)	57.42***	39.68***	-1.72*	-0.05											
Model 2	7.75	4.92	-1.32												70.56%
(t-stat)	57.86***	37.32***	-2.78***												
Model 3	7.08	4.92		-0.02											72.38%
(t-stat)	59.22***	39.66***		-0.05											
Model 4	7.25	6.02	0.11		-0.22	72.82	-22.09	7.74	0.01	0.01	-19.66	-0.13	-10.42	10.01	95.27%
(t-stat)	65.80***	55.71***	0.54		-3.31***	46.75***	-12.92***	11.94***	0.53	0.01	-2.18**	-7.74***	-1.26	1.17	
Model 5	7.22	6.04	0.10		-0.20	73.01	-21.30	7.59			-19.18	-0.14	-2.27		95.26%
(t-stat)	74.01***	62.21***	0.52		-3.56***	47.16***	-13.69***	12.20***			-2.13**	-7.93***	-0.50		

Table 5. OLS Regressions for Two Order Difficulty Sub-Samples

This table presents OLS regression results for the two order difficulty sub-samples. Panels A and B present results for easy and difficult orders respectively, where order difficulty is defined by the fitted value using the model for the whole sample.

Panel A. Sub-Sample of Easy Orders

	NYSE	EM	Log(NUM.ORD)	INFO	1/PRC	ORD.SZ/VOL	TURNOVER	MCAP.RANK	VOLA	R^2
	α_4	α_2	α_3	α_4	α_5	α_6	α_9	α_{10}	α_{11}	
Model 1	3.70	3.11								82.74%
(t-stat)	51.08***	48.01***								
Model 2	3.82	3.01	-0.18	46.87	-16.35	29.34	23.96	-0.04	-2.60	90.84%
(t-stat)	41.56***	39.23***	-3.37***	14.55***	-10.76***	10.49***	2.83**	-2.00**	-0.63	

Panel B. Sub-Sample of Difficult Orders

	NYSE	EM	Log(NUM.ORD)	INFO	1/PRC	ORD.SZ/VOL	TURNOVER	MCAP.RANK	VOLA	R^2
	α_4	α_2	α_3	A_4	α_5	α_6	α_9	α_{10}	α_{11}	
Model 1	10.70	7.29								84.63%
(t-stat)	64.88***	39.87***								
Model 2	6.59	4.52	-0.33	70.49	-28.62	6.04	-56.98	-0.15	-4.51	96.18%
(t-stat)	30.02***	19.50***	-2.94***	31.59***	-9.52***	6.89***	-3.30***	-5.40***	-0.54	

Table 6. Two-Stage Selection Model for Two Order Difficulty Sub-Samples

This table presents selection model results for the two order difficulty sub-samples. Panels A and B present results for easy and difficult orders respectively, where order difficulty is defined by the fitted value using the model for the whole sample.

Panel A. Sub-Sample of Easy Orders

First Stage Probit Regressions for the Likelihood of an Order Being Executed by the NYSE

	Log (MCAP)	**Log (VOL)**	**Log (ORD.SZ)**	**INFO**	**AQS**	**SPEED**	**RR**	**VOLA**	**CBMA**	**I1**	**PSGAMMA**		
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	**Pseudo R^2**	**Prob > χ^2**
	2.29	-5.04	45.60	32.85	23.98	-0.11	-68.14	75.07	50.94	10.59	-112142	95.82%	0.00
(z-stat)	3.32***	-4.95***	5.70***	2.46***	1.82*	-3.78***	-1.35	1.64*	0.46	1.25	-1.01		

Second Stage OLS Regressions of Conditional Average Effective Spreads

	NYSE	**EM**	**λ_1NYSE**	**Log(NUM. ORD)**	**INFO**	**1/PRC**	**ORD.SZ /VOL**	**TURNOVER**	**MCAP. RANK**	**VOLA**	**R^2**
	α_1	α_2	α_3	α_5	α_6	α_7	α_8	α_{11}	α_{12}	α_{13}	
Model 1	3.69	3.10	0.02								82.73%
(t-stat)	49.48***	47.83***	0.04								
Model 2	3.79	3.01	0.51	-0.18	0.48	-16.21	29.68	24.50	-0.04	-2.70	90.85%
(t-stat)	40.75***	39.28***	1.48	-3.29***	14.70***	-10.67***	10.60***	2.90***	-1.83*	-0.66	

Table 6. Continued

Panel B. Sub-Sample of Difficult Orders

First Stage Probit Regressions for the Likelihood of an Order Being Executed by the NYSE

	Log (MCAP)	**Log (VOL)**	**Log (ORD.SZ)**	**INFO**	**AQS**	**SPEED**	**RR**	**VOLA**	**CBMA**	**I1**	**PSGAMMA**		
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	**Pseudo R^2**	**Prob > χ^2**
	0.67%	-1.87	14.12	26.46	4.39	-0.06	-104.25	91.97	-82.22	0.49	-37041.59	94.40%	0.00
(z-stat)	1.06	-2.14^{**}	7.05^{***}	4.22^{***}	0.68	-4.36^{***}	-2.38^{**}	2.20^{**}	-1.30	0.18	-1.24		

Second Stage OLS Regressions of Conditional Average Effective Spreads

	NYSE	**EM**	**λ_1NYSE**	**Log(NUM. ORD)**	**INFO**	**1/PRC**	**ORD.SZ /VOL**	**TURNOVER**	**MCAP. RANK**	**VOLA**	**R^2**
	α_1	α_2	α_3	α_5	α_6	α_7	A_8	α_{11}	α_{12}	α_{13}	
Model 1	10.71	7.29	-0.31								84.63%
(t-stat)	63.32^{***}	37.85^{***}	-0.30								
Model 2	6.63	4.50	-0.47	-0.35	70.36	-28.56	6.04	-55.27	-0.15	-4.82	96.18%
(t-stat)	29.58^{***}	19.37^{***}	-0.90	-3.03^{***}	31.45^{***}	-9.50^{***}	6.88^{***}	-3.18^{***}	-5.41^{***}	-0.58	

Out of those new variables, the coefficients on *SPEED, RR, VOLA,* and *CBMA* are all significant at the 99% level, and *PSGAMMA* is significant at the 95% level. Interestingly, many of these same variables have shown little explanatory power in previous OLS regressions on effective spreads. One explanation is that these factors affect order routing decisions rather than execution costs directly.

Table 4 Panel B presents the second stage OLS regressions. When both $\lambda_1 NYSE$ and $\lambda_2 EM$ are added to the model in addition to the two dummy variables, the coefficient on $\lambda_1 NYSE$ is significant only at the 90% level, while the coefficient on $\lambda_2 EM$ is not significant. This could be because the two constructed variables are highly correlated. So we also include $\lambda_1 NYSE$ and $\lambda_2 EM$ separately in second and third model. $\lambda_1 NYSE$ becomes more significant, but $\lambda_2 EM$ is still not significant. In the fourth and fifth models, we include $\lambda_1 NYSE$ (the Probit factor) and other control variables. The Probit factor becomes insignificant. It is probably because that the constructed variable is closely linked to variables that are known to affect execution costs, therefore contain similar information. This is consistent with Bessembinder's (2003) finding that OLS regressions seem to do a good job at controlling for order difficulty, and selection models do not seem add incremental explanatory power.

To summarize results obtained so far, unconditional mean effective spread is higher for the NYSE than the Electronic Markets. Controlling for stock and order characteristics in multivariate OLS regressions decreases this difference, though the NYSE still has significantly higher conditional effective spread than the Electronic Markets. Selection models do not significantly change OLS regression results.

5. Difficulty Sub-Sample Analysis

Since the above analysis only examines the execution cost at the in-sample mean points, it is worthwhile to conduct analysis for the two markets at other difficulty levels. One might think that the impact of different dimensions of difficulty may change depending on the levels of difficulties. One might also expect that the selection effect could be stronger among more difficult orders. In order to do this, we calculate the fitted value based on our selected model $\alpha'X$ without intercept. These fitted values are estimates of order difficulty. We then sort our sample according to this order difficulty measure and split our sample to two sub-samples. The OLS analysis and selection model analysis are reproduced for each difficulty sub-sample.

Table 5 presents OLS results for the two difficulty sub-samples. Panel A is for the sub-sample of easy orders, while Panel B is for the sub-sample of difficult orders, according to our difficulty measure $\alpha'X$. We note some interesting results here: (1) More difficult orders show a greater cost decrease in mean effective spreads conditionally (after including various factors in the regression) versus unconditionally. The NYSE conditional mean effective spread dropped 38.41%, from unconditional mean of 10.70 cents to conditional mean of 6.59 cents. The Electronic Markets mean effective spread dropped a similar 38.00%, from unconditional mean of 7.29 cents to conditional mean of 4.52 cents. On the other hand, the conditional means of effective spreads for easy orders do not change as much. (2) For easy orders, the conditional effective spread difference on and off the NYSE is wider than the unconditional means'. It means that including demeaned explanatory variables increases divergence. This result is in contrast to the common conception that easier orders are more often sent to the Electronic Markets. (3) For difficult orders, the conditional effective spread difference on and off the NYSE is narrower, which confirms the NYSE's assertion that they receive more difficult orders. (4) For difficult order sub-sample, turnover has a negative correlation with the effective spread, meaning that high volume orders are charged with lower costs. However, for easy orders, the relationship between turnover and effective spread is significantly reversed. High volume easy orders get higher execution costs. (5) The R^2 increases more for the difficult order sub-sample, from unconditional model's 84.63% to the conditional model's 96.18%. The R^2 change in the easy order sub-sample is not as big: it increases from 82.74% to 90.84%.

Table 6 presents the Heckman (1979) two stage selection model, correcting for selection bias. Again, Panel A is for the sub-sample of easy orders, Panel B is for the difficult order sub-sample, according to our difficulty measure $\alpha'X$. The second stage OLS regression results are similar to the results of simple OLS regression results shown in Table 5. The coefficients of the Probit regression factor are not significant in both sub-samples.

6. What if the NYSE and the Electronic Markets Executed Each Other's Orders?

Like most other studies, our analysis so far has focused on the same model for orders executed both on the NYSE and the Electronic Markets. Implicitly, the underlying assumption is that the "definition" of order difficulty is the same across the two markets. Specifically, in the regression framework, orders executed on the two markets are "pooled" together as observations, and the regression coefficients on independent variables are constrained to be the same for the two markets, even though the intercepts were allowed to differ by using two dummy variables for the two markets. It might be reasonable to think that the same factors may have different impact on execution costs for different market mechanisms. One could interact the dummy variables with the independent variables to allow the coefficients to differ across the two markets, though this has not been done by previous studies, perhaps because this may make the model cumbersome.

We adopt a simple methodology: we model the two markets separately. This allows for maximum freedom to capture the differences across the two markets. This simple method also enables us to answer very interesting questions: what if the NYSE executed the Electronic Markets' orders? And what if the Electronic Markets executed the NYSE' orders? To answer the former question, after we estimate different models for the two markets, we then plug the observations from the Electronic Markets (using the Electronic Markets' independent variable values) into the estimated model for the NYSE (using the NYSE model coefficients). These fitted values ($\alpha_N + \beta_N' X_{EM}$) have a very nice interpretation: these are the expected effective spreads that the Electronic Markets' orders would have received, had they been executed on the NYSE. We can then compare these fitted effective spreads using the NYSE coefficients with the observed (realized) effective spreads of the Electronic Markets' orders.

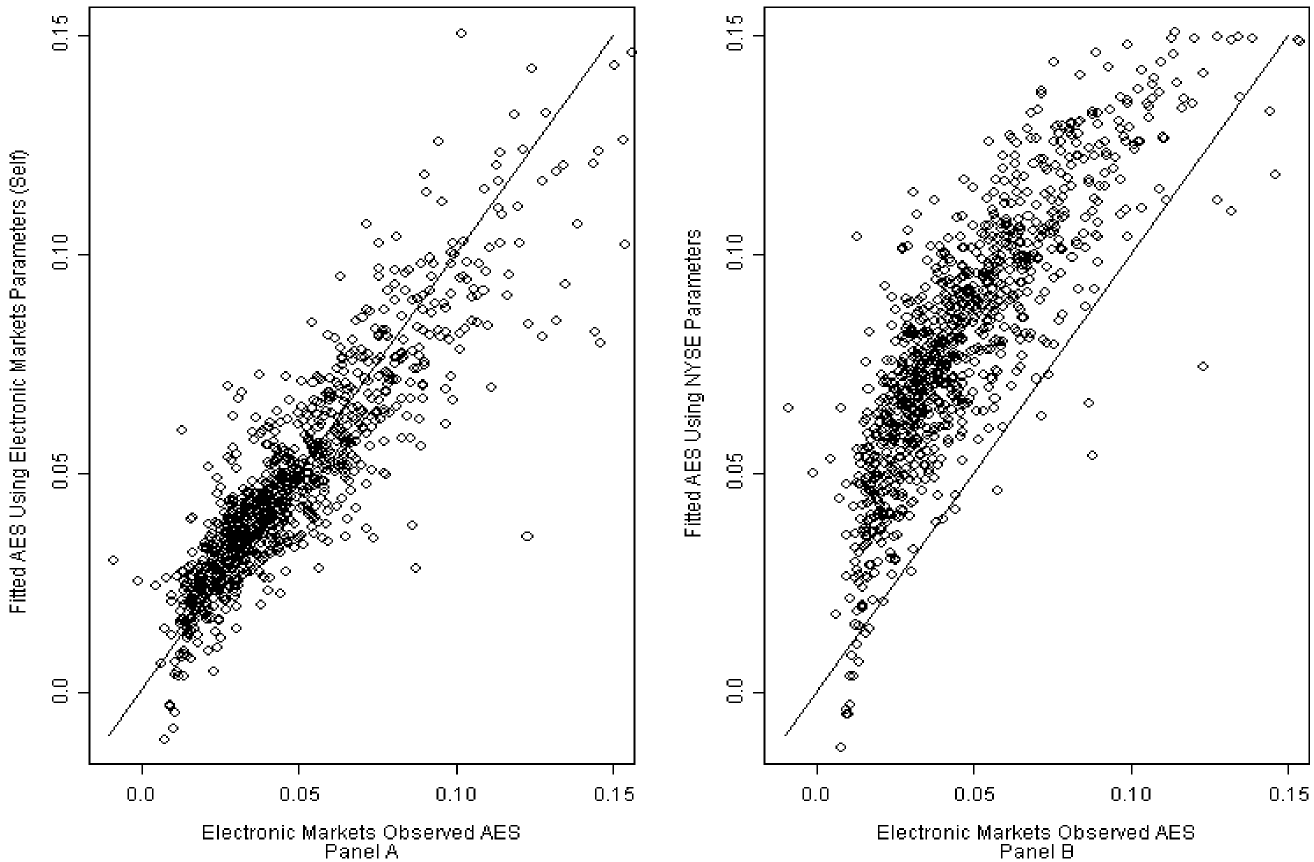

Figure 1. What If the NYSE Executed the Electronic Markets' Orders?

Table 7. What If the NYSE and the Electronic Markets Executed Each Other's Orders?

This table presents results of the what-if analysis. Panel A reports observed AES and fitted AES for the Electronic Markets' orders, and fitted AES assuming these orders are executed according the NYSE's model parameters (and its difference from observed AES). Panel B reports observed AES and fitted AES for the NYSE' orders, and fitted AES assuming these orders are executed according the Electronic Markets' model parameters (and its difference from observed AES).

Panel A. What If the NYSE Executed the Electronic Markets' (EM) Orders?

Decile	EM Observed AES (1)	Fitted AES Using EM Parameters (Self) (2)	Fitted AES using NYSE Parameters (3)	Diff (3) – (1)	t-stat for Diff
1	1.20	1.83	3.93	2.73	10.12***
2	2.24	2.68	5.45	3.21	23.78***
3	2.86	3.41	6.56	3.70	24.19***
4	3.37	3.82	7.23	3.86	30.88***
5	3.97	4.16	7.77	3.80	27.24***
6	4.61	4.69	8.36	3.75	26.94***
7	5.49	5.25	9.04	3.55	26.34***
8	6.43	6.13	10.18	3.76	24.96***
9	7.89	7.30	11.21	3.32	19.59***
10	11.73	10.52	14.28	2.55	10.77***
All	4.98	4.98	8.40	3.42	61.31***

Table 7. Continued.

Panel B. What If the Electronic Markets (EM) Executed the NYSE's Orders?

Decile	NYSE Observed AES (1)	Fitted AES Using NYSE Parameters (Self) (2)	Fitted AES using EM Parameters (3)	Diff (3) – (1)	t-stat for Diff
1	1.82	1.98	1.77	-0.05	-0.74
2	2.89	3.54	2.82	-0.07	-1.12
3	3.90	4.69	3.76	-0.14	-0.58
4	4.78	5.23	4.20	-0.58	-3.56***
5	5.80	6.09	5.00	-0.80	-6.15***
6	7.12	7.31	6.07	-1.05	-10.54***
7	8.53	8.26	7.00	-1.53	-11.55***
8	10.32	9.97	8.66	-1.66	-11.59***
9	13.02	12.59	11.69	-1.33	-5.93***
10	18.32	16.85	16.18	-2.14	-6.44***
All	7.65	7.65	6.71	-0.94	-15.54***

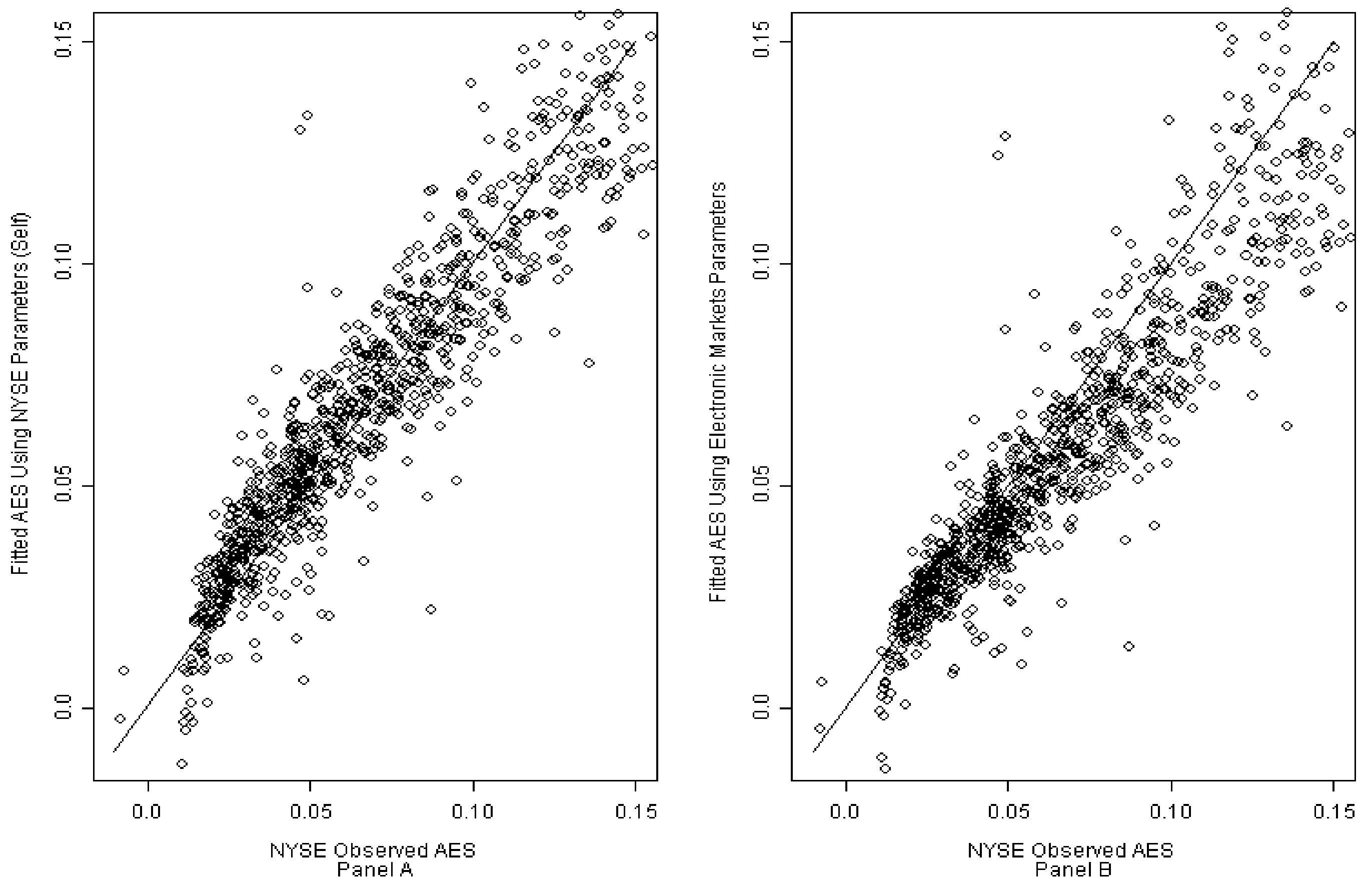

Figure 2. What If the Electronic Markets Executed the NYSE's Orders?

Figure 1 plots the Electronic Markets' observed effective spreads on the horizontal axis. Figure 1 Panel A plots fitted effective spreads using parameters of the model for the Electronic Markets on the vertical axis ($\alpha_{EM} + \beta_{EM}'X_{EM}$). The sample points should be scattered evenly around the 45 degree line, since this is using the Electronic Markets' "self" model. Figure 1 Panel A confirms this. Figure 1 Panel B plots fitted effective spreads using the NYSE model parameters ($\alpha_N + \beta_N'X_{EM}$) on the vertical axis. Figure 1 Panel B clearly shows that the NYSE executions would have been more costly for the Electronic Markets' orders, because most sample points are above the 45 degree line. These results are also presented in Table 7 Panel A. It first sorts the observed effective spreads of the Electronic Markets' orders into 10 deciles. We then calculate the fitted effective spreads using its "self" model, as well as using the NYSE model in each decile. Fitted effective spreads using the NYSE model parameters are significantly larger than the corresponding observed Electronic Markets effective spreads in all 10 deciles. These results confirm that if these the Electronic Markets' orders had been executed on the NYSE, they would have received higher execution costs.

Figure 2 and Table 7 Panel B present results for the "reverse test": what if the Electronic Markets executed the NYSE's orders. In Figure 2 Panel B, most sample points are below the 45 degree line, suggesting that the NYSE's orders would have received lower effective spreads, had they been executed by the Electronic Markets. Similarly in Table 7 Panel B, we see that the fitted effective spreads for the NYSE's orders using the Electronic Markets' model parameters are significantly lower than the observed NYSE effective spreads overall. They are also lower in each of the 10 deciles (statistically significant in 7 deciles). Overall, our what-if analysis in this section suggests that the Electronic Markets' (NYSE's) orders would have been worse (better) off, had they been executed by the NYSE (Electronic Markets).

7. Conclusion

In this paper, we conduct a thorough analysis comparing execution costs between the NYSE and the Electronic Markets, adopting a variety of techniques attempting to correct for the selection bias problem. Conventionally, the point comparison of execution costs is conducted at each market's own average, which is widely criticized because of the selection bias problem, i.e., the difficulties of the orders routed to the two markets could be different. Unlike current literature, we find that after controlling for the selection bias problem, the Electronic Markets still offer lower execution costs than the NYSE. We also carry out our analysis at different levels of order difficulty, measured by a vector of control variables, instead of just controlling for selection biases at sample mean level of order difficulty. Our results are robust under different model specifications. In addition, the results of our what-if analysis show that the Electronic Markets' (the NYSE's) orders would have been worse (better) off, had they been executed by the NYSE (Electronic Markets). Overall, our results highlight the superiority of the Electronic Markets' execution quality. In terms of a firm's exchange listing decision, our results are consistent with Chemmanur and Fulghieri's (2006) theoretical analysis: execution costs alone do not drive a firm's listing decision.

Appendix: Sample Selection and List of Market Centers

Table A1. Sample Selection

This table describes the selection of the final sample from all securities in the CRSP database. The filters are not mutually exclusive; therefore their ordering is important.

General Filters

All securities on 12/31/2002	2557
+ Single class	2355
+ Ordinary common stock which need not be further defined	1329
+ not "no price on 12/31/2002"	1329
+ not "no SIC code on 12/31/2002"	1328
+ no "missing daily price during 01/01/2001 and 12/31/2003"	1218
+ no switch	1204
+ mean daily trading volume >=$20,000	1192
+ no missing daily volume, any day during the fourth quarter of 2002	1192
+ no daily price during the fourth quarter of 2002<$3.00	1111
+ no change exchange	1108
Total symbols	1116

Further Modifications

Top 10% of market capitalization on 2002/12/31	19 (out of 343)
Top 10% of average daily volume during the fourth quarter of 2002	18 (out of 135)
Top 10% of average daily dollar volume during the fourth quarter of 2002	8 (out of 135)
Total final symbols	1146
For dash 5 data	1142 (no records for NMG, CCR, PZL, KM)
Exclude those without full year's data	1138 (HI, MIR, PHA, UAL)

Table A2. List of Market Centers

This table lists market centers for the NYSE and electronic markets, respectively.

Panel A. NYSE Specialist firms

N0003	WAGNER STOTT BEAR SPEC.
N0034	LA BRANCHE CO.
N0041	FLEET MEEHAN SPECIALIST
N0050	SUSQUEHANNA SPECIALISTS
N0055	SPEAR LEEDS AND KELLOGG
N0061	VAN DER MOOLEN SPECIALISTS USA
N0070	PERFORMANCE SPECIALIST GROUP LLC
N9999	NYSE ITS

Table A2. Continued

Panel B. Electronic (NASD) Market Centers

SCHB	SCHB(US) SCHWAB CAPITAL MARKETS L.P.
TACT	TACT(US) AUTOMATED CONFIRMATION TRANSACTION SERVICE
TARCA	ARCA(US) ARCHIPELAGO SECURITIES L.L.C.
TAUTO	AUTO(US) AUTOMATED TRADING DESK FINANCIAL SERVICES, LLC
TBBNT	BBNT (US) SCOTT AND STRINGFELLOW INC.
TBRGE	BRGE (US) NEWBRIDGE SECURITIES CORPORATION
TBRUT	BRUT(US) BRUT, LLC
TCAES	CAES(US) COMPUTER ASSISTED EXECUTED SYSTEM
TDIRA	DIRA (US) DIRECT ACCESS BROKERAGE SERVICES
TFAHN	FAHN (US) OPPENHEIMER & CO. INC.
TFPKI	FPKI (US) FOX-PITT KELTON INC.
TFRGP	FRGP (US) FORGE FINANCIAL GROUP, INC.
TINET	INET(US) INET ATS, INC.
TISLD	ISLD(US) ISLAND CORPORATION
TJBOC	JBOC (US) NATIONAL CLEARING CORP.
TLQNT	LQNT(US) LIQUID NET INC.
TLYON	
TMADF	MADF(US) BERNARD L. MADOFF
TMAYF	MAYF(US) MAY FINANCIAL CORP
TMCBT	MCBT(US) MOORS AND CABOT INC.
TMONT	MONT(US) BANC OF AMERICA SECURITIES LLC
TNYFX	NYFX(US) NYFIX MILLENIUM, L.L.C.
TPRMX	PRMX(US) PRIMEX PRIME ELECTRONIC EXECUTION INC.
TSBSH	SBSH (US) CITIGROUP GLOBAL MARKETS INC.
TSOAR	
TSSBS	SSBS(US) STATE STREET GLOBAL MARKETS, LLC
TSWST	SWST(US) SOUTHWEST SECURITIES, INC.
TTDCM	TDCM(US) TD WATERHOUSE CAPITAL MARKETS, INC.
TTHRD	THRD(US) THE THIRD MARKET CORP.
TTRIM	TRIM(US) KNIGHT CAPITAL MARKETS, INC.
TUBSW	UBS SECURITIES LLC
TVFIN	VFIN (US) VFINANCE INVESTMENTS INC.
TWRHC	WRHC (US) WILLIAM R. HOUGH & CO.
WATHWATH	WATH(US) TD WATERHOUSE INVESTOR SERVICES, INC.

References

Amihud, Y., 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* **5,** 31-56.

Bessembinder, H., 2003, Selection Biases and cross-market trading cost comparisons, working paper.

Blume, M.E., and Goldstein, M.A., 1992, Displayed and effective spreads by market, *Rodney L. White Center for Financial Research* Working Paper 27-92, The Wharton School, December.

Boehmer, E., 2005, Dimensions of execution quality: Recent evidence for U.S. equity markets, *Journal of Financial Economics* **78**, 553-582.

Chemmanur, T. J., and Fulghieri, P., 2006, Competition and co-operation among exchanges: a theory of cross listing and endogenous listing standards, *Journal of Financial Economics* **82**, 455-489.

Christie, W., and Schultz, P., 1994, Why do Nasdaq market makers avoid odd-eighth quotes? *Journal of Finance* **49**, 1813-1840.

Chung, K., Van Ness, B., and Van Ness, R., 2001, Can the treatment of limit orders reconcile the differences in trading costs between NYSE and Nasdaq issues, *Journal of Financial and Quantitative Analysis* **36**, 267-286.

Davis, P.L., Pagano, M.S., and Schwartz, R.A., 2006, Life after the Big Board goes electronic, *Financial Analysts Journal* **62** (5), 14-20.

Goldstein, M., Shkilko, A., Van Ness, B., and Van Ness, R., 2008, Competition in the market for NASDAQ securities, *Journal of Financial Markets* **11** (2), 113-143.

Hasbrouck, J., 2006, Trading costs and returns for US equities: Estimating effective costs from daily data, *Journal of Finance* **64** (3), 1445 - 1477.

Heckman, J., 1979, Sample selection biases as a specification error, *Econometrica* **47**, 153-162.

Kalay, A., and Portniaguina, E., 2001, Swimming against the tides: the case of Aeroflex move from NYSE to Nasdaq, *Journal of Financial Markets* **4**, 261-267.

Lipson, M., 2005, *Competition among market centers*, working paper.

Maddala, G., 1983, *Limited dependent and qualitative variables in econometrics*, Cambridge University Press.

Nguyen, V., Van Ness, B., and Van Ness, R., 2005, Archipelago's move towards exchange status: An analysis of Archipelago trading in NYSE and Nasdaq stocks, *Journal of Economics and Business* **57**, 541-554.

Pagano, M.S., and Schwartz, R.A., 2003, A closing call's impact on market quality at Euronext Paris, *Journal of Financial Economics* **68**, 439-484.

Pastor, L., and Stambaugh, R.F., 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* **111**, 642-685.

Petersen, M.A., and Fialkowski, D., 1994, Posted verses effective spreads: good prices or bad quotes?, *Journal of Financial Economics* **35**, 269-292.

Peterson, M., and Sirri, E., 2002, Order submission strategy and the curious case of marketable limit orders, *Journal of Financial and Quantitative Analysis* **37**, 221–241.

Pruitt, S., Van Ness, B., and Van Ness, R., 2002, The first of many? The microstructure effects of Aeroflex Corporation's move from the NYSE to the Nasdaq, *Journal of Applied Finance* **12**, 46-54.

SEC, 2001, Report on the comparison of order execution across equity market structures, U.S. Securities and Exchange Commission, Washington.

Van Ness, B., Van Ness, R., and Warr, R., 2005, Nasdaq trading and trading costs: 1993-2002, *Financial Review* **40**, 281-304.