# The Effect of Payment for Order Flow on the Market Share of Trading Centers

Saketh Aleti

February 26, 2018

## 1 Introduction

One of the most significant changes in US equity markets in the last decade is the emergence of widespread competition for retail order flow[1]. About 20 years ago, stocks listed on an exchange such as NASDAQ were only traded on that exchange. Today, when a broker receives an order, they can redirect the order to one of hundreds of locations or internalize[2] it. If they choose to redirect the order, it may go to an exchange, a market maker, or an electronic communications network. Some venues offer brokers some rebate, usually less than a penny per share, in order to attract order flow. The SEC regulates these venues and refers to them using the term *market center* and are also sometimes referred to as trading centers; the SEC defines a market center as "any exchange market maker, OTC market maker, alternative trading system, national securities exchange, or national securities association" as per the Securities Exchange Act of 1934. The process by which market centers offer rebates for orders from broker-dealers is called *payment for order flow*. This practice remains controversial, since it allows brokers to effectively sell their client's orders to the highest bidder rather than getting them the best execution. At the same time, some have argued that the option to pay for order flow has come with

---

[1] Retail order flow refers to orders for securities that originate from retail investors as opposed to institutional investors such as banks and hedge funds

[2] Internalization is the process by which they execute the order against their own inventory, so they can profit from the spread between your order's payment and the market price

benefits such as greater exchange competition (Angel, Harris, and Spatt, 2011).

According to an industry notice from the National Association of Securities Dealers, now the self-regulatory agency FINRA,

> The traditional non-price factors affecting the cost or efficiency of executions should also continue to be considered; however, broker-dealers must not allow an order routing inducement, such as payment for order flow or the opportunity to trade with that order as principal, to interfere with its duty of best execution. (2001)

The SEC has made similar statements and codified certain requirements for the direction of order flow. Between 2005 and 2007, the SEC updated the Securities Exchange Act of 1934 with Regulation National Market System (Reg NMS); this regulation introduced new rules for order protection, intermarket access, sub-penny pricing, and market data. In particular, according to the SEC's Order Protection Rule of Reg NMS, orders must be executed at the National Best Bid or Offer (NBBO) or better; the best bid is the highest price a buyer will accept and the best offer is the lowest price a seller will accept. This rule establishes price priority, so brokers cannot[3] execute their client's orders at a worse price than the NBBO. However, when multiple exchanges are trading a stock at the NBBO, the broker can choose which venue to route to.

In such a case, brokers may decide where to route based on the rebate they receive for directing order flow. This involves responding to one of two fee structures: either payment for order flow or maker-taker fees. The latter simply involves paying a fee for removing liquidity (sell orders) and receiving a rebate for adding liquidity (buy orders). This fee/payment structure is particularly targeted towards limit orders. On the other hand, payment for order flow consists of purchasing orders independent of whether they are buy or sell; generally, this is targeted towards all orders. For the sake of the retail investor, it would be best if the broker routed their order to maximize execution quality; some factors of execution quality are price improvement, execution speed, execution probability. But, there exists enough leniency within the SEC's definition of a "best execution" under Reg NMS for brokers to route according to their own objectives and offer a post hoc justification.

---

[3] There are exceptions for this rule which are aimed at High Frequency Traders, but this is irrelevant for retail order flow.

## 1.1 Objective

My thesis aims to identify how payments for order flow affect the market share of market/trading centers. Since many firms that process order flow, such as Citadel and Goldamn Sachs, offer other services, I define market share as the percent of orders executed at a specific market center relative to orders executed at all market centers. Understanding the extent to which rebates redirect order flow is necessary for determining their impact on the execution quality of retail orders. If they have an insignificant effect on improving the market share of market centers, then it is unlikely that brokers are prioritizing revenue over execution quality. On the other hand, a significant effect would imply that brokers are trying to capture rebates instead of the best possible execution. This research expands upon previous research in market microstructure that studies the effects of rebates and fees offered firms looking to capture order flow.

## 1.2 Road Map

The remainder of the paper will proceed as follows:

- Section II: Literature Review
- Section III: Data Description
- Section IV: Methodology
- Section V: Results
- Section VI: Conclusion

In the methodology section, I formulate a regression equation to use for my data and decide which variables are relevant. In my results section, I provide my regression outputs and tables for the significance of the functional forms that I test.

# 2 Literature Review

## 2.1 The Theoretical Foundation for Order Rebates

Early research on the effects of payment for order flow is still relatively recent given that the practice emerged in the 1980s. At that time, order flow payments were as high as a penny per share, since major exchanges would trade in 12.5 cent increments. As a result, the early literature, such as Chordia and Subrahmanyam (1995) and Kandel and Marx (1999), argued that payment for order flow was caused by higher tick sizes. However, as tick sizes got smaller, payments did not disappear but simply diminished in size. Now, newer papers like Parlour and Rajan (2001) and Battalio and Holden (2001) present theoretical models for the existence of payments for order flow persisting in continuous price settings.

In particular, Battalio and Holden find that brokers can maximize profits by using their internal data on their clients to identify uninformed order flow; if a broker predicts that their client is uninformed, they internalize the order and profit from the spread. On the other hand, their model assumes that third parties, such as market centers, purchase order flow based on "externally-verifiable" characteristics; this is a realistic assumption given that third parties do not have access to client data before they purchase an order. These factors motivate third parties to pay for order flow, since they will only have access to potentially uninformed order flow if they can compete with internalization. The former paper, Parlour and Rajan (2001), develops another theoretical model consisting of an infinite horizon game of liquidity-provision. This game consists of competing market makers, brokers, and investors; market makers supply liquidity, investors demand or supply liquidity, and brokers act as intermediaries. Their model finds that scenarios where payment for order flow is allowed results in lower consumer and social welfare than the counterfactual. Additionally, its existence results in a welfare transfer from liquidity takers to suppliers, which they define as those submitting market orders versus those sending limit orders. Both models hold the number of market centers fixed, while there also exists literature on the effects of greater competition as well.

4

## 2.2 The Effects of Order Rebates

Angel, Harris, and Spatt (2011) note that Regulation NMS has resulted in significantly greater competition among market centers. However, Dennert (1993) and Duta and Madhavan (1997) find that competition between market centers can actually have mixed effects for retail investors. Firstly, Dennert (1993) models price competition where prices are quoted by competing dealers facing some unknown future demand. Then, introducing informed and uninformed traders to the model, Dennert finds that increasing the number of market makers increases their individual risk exposure. As a result, bid-ask spreads increase and liquidity traders face higher transaction costs. So, there exist cases where liquidity traders, who are generally uninformed, would be better off in monopolistic markets than competitive ones. Next, Duta and Madhavan (1997) present a model where multiple market makers compete with one another resulting in higher spreads; in some ways, this effect is similar to what would occur if the market makers colluded and fixed prices. They suggest that this outcome is due to current institutional arrangements, and investors would be better off if price priority was eliminated, internalization was discouraged, and order flow payments were eliminated.

Other papers find similar results regarding the effect of order rebates. Cimon (2016) presents a theoretical model of limit order trading and finds that significant spreads in the fee/rebate for maker-taker venues cause brokers to reroute orders to venues with poor execution quality. This model consists of two periods where brokers decide where to route their clients marketable orders. The paper then provides a theoretical basis for two empirical results: exchanges with higher rebates will endogenously have worse fill rates and, when trading fees are sufficiently different, brokers will prioritize rebates over their clients' orders fill rates. Maglaras, Moallemi, and Zheng (2015) develop a different model where, instead of brokers acting as an intermediary, investors route their own orders and can pay fees or accept rebates for their order flow. The model finds that limit orders sent to high fee venues still face worse execution quality even when investors themselves are making the decision. Additionally, when fill rates between high and low fee venues are significantly close, routing to capture rebates is welfare improving for investors. In

other cases, the rebates offered by high fee venues are not enough to offset their subpar execution quality. In short, both papers imply that maximizing rebate revenue is not equivalent to maximizing execution quality.

Lastly, there exists literature examining the effect of maker-taker and payment for order flow models on execution quality. To begin, we have Battalio, Shkilko, and Van Ness (2016), which examines differences between venues that pay for order flow and those that use the maker-taker model in US options markets. Using trade and quote data for options from OPRA, they find that the cost of liquidity is approximately 40 basis points lower in markets that pay for order flow than those that use maker-taker. Additionally, they find that high priced options face lower spreads in maker-taker exchanges; therefore, they argue that brokers cannot fulfill their best execution obligation if they route all their orders to venues that use the payment for order flow model, which would maximize broker revenue. Analyzing US equity markets instead, Battalio, Corwin, and Jennings (2016), find empirical evidence that brokers who route to maximize rebates cannot be maximizing execution quality. Using a proprietary data set of orders, they develop a model which they then further investigate with the larger TAQ dataset. Battalio, Corwin, and Jennings find that four major retail brokers "appear to route all nonmarketable limit orders to the venue offering the largest rebate." Additionally, based on results from their proprietary data set, they find that limit order execution quality is negatively correlated with an exchange's rebate. Generalizing to the TAQ dataset, they also find that order rebates have greater impacts on low-priced stocks. In any case, these papers develop an empirical foundation for the argument that maximizing rebate revenue does not maximize execution quality.

Based on previous literature, we expect order routing that prioritizes rebate revenue to be subpar in execution quality. As stated above, this has both a theoretical foundation and an empirical foundation. This paper aims to estimate the effect of payment for order flow on the market share of market centers. Given that the latter is measured by order flow itself, this question can also be rephrased as finding the elasticity of the rebate price with respect to the quantity of orders. So, while the literature explains the effects of

revenue-maximizing order routing, this research will estimate the extent of order routing that is actually done to maximize rebates. Thus, the result will further develop the empirical literature and may offer new questions for theory.

# 3 Data Description

For my methodology, I use data from two sources: 605 disclosures and 606 disclosures. After collecting the data, I merge both into a dataset where each observation belongs corresponds to a specific broker, market center, exchange, order type, and time period. Because the data sets are in different time frequencies, I consolidate each to quarterly before merging.

## 3.1 605 Data

Firstly, I obtain the 605 disclosures from each market center's webpage. The SEC requires market centers to disclose information every month regarding the quality of executions for each stock; these disclosures are referred to as 605 filings in accordance with Rule 605 of Reg NMS. The 605 data contains the volume of each stock for each order type that passes through the market center. Additionally, it breaks down this volume into orders that received price improvement, orders that were at-the-quote, and orders that were outside-the-quote. When an order receives price improvement, it means that the buyer or seller received a better price than what they were willing to pay or what the market offered (the NBBO); for instance, if the NBBO for \$AAPL is at 100/101, market centers that execute trades between these prices provide price improvement. This data is further broken down into the number of price improved shares for each stock, the average amount of price improvement, and the average time a price improved order took to execute. Similarly, market centers report the average time and shares of ATQ and OTQ orders as well; ATQ refers to orders at the NBBO while OTQ refers to orders outside the NBBO. This data is necessary in constructing my independent variables: measures of execution quality. Also, since 605 data is reported on a monthly and per stock basis, I convert it to a quarterly and per exchange basis to match the 606 data. The figure below shows the market share of Citadel for stocks at each quantile in market cap; this figure was constructed using the 605 disclosure from Citadel's website and sorted by market cap using securities data from Compustat.

**Market Share of Citadel in 2017**
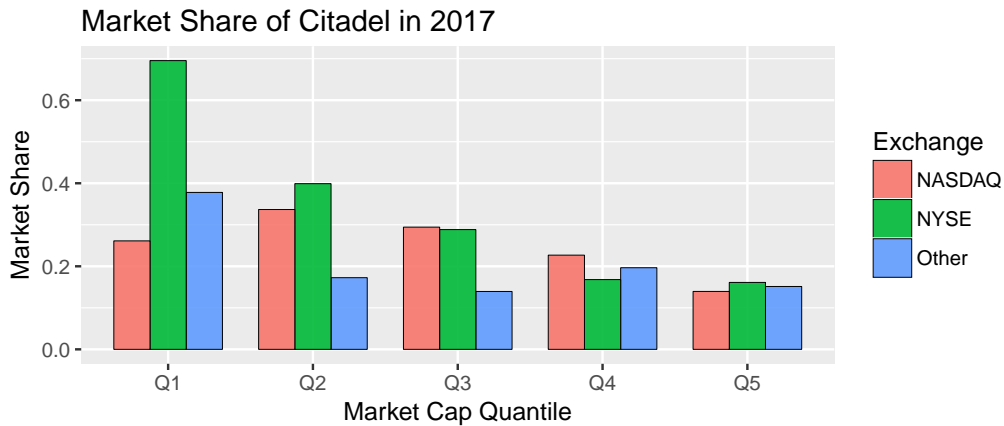


The following figure shows price improvement offered by Citadel from its 605 data. Note that price improvement tends to be higher for market orders, because they are simply orders that investors use to purchase stocks at the market price. It is generally easier to offer price improvement on these, since they must be (legally required to be) executed at either the NBBO or inside of it. On the other hand, limit orders must be executed at a certain price or better; investors specify this price and it may be outside the NBBO, so these orders may not be executed immediately within it. Lastly, orders of type "Other" are not shown, because their price improvement is undefined. These orders greatly vary in their execution requirements, so there is no clear way to measure their price improvement.

**Percent of Orders Price Improved by Citadel**
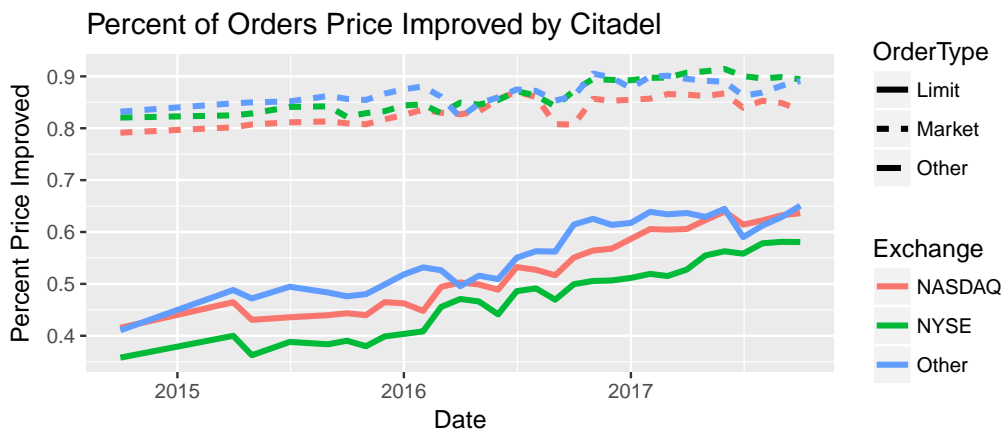


The next figure shows expected price improvement, which is the product of percent of orders price improved and average price improvement; so, this is the amount of money that an investor would save from having an additional share routed to a market center. This figure also shows the availability of the 605 data. Not every market center has

publicly available 605 data from before last year; by Rule 605 of Regulation NMS, they are only obligated to post the last three quarters of data. Since the amount the data across the time dimension varies with each market center, my regression will be unbalanced panel.



Expected Price Improvement over Time
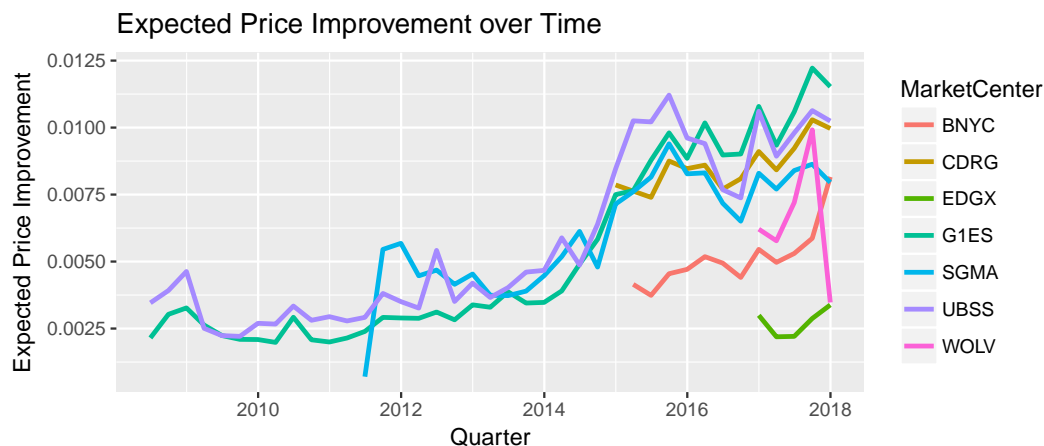
## 3.2   606 Data

Similarly, the SEC requires broker-dealers to disclose where they route their orders and the average payment that they receive; these disclosures are referred to as 606 filings. The 606 data shows the percent of orders, by order type, that was routed to each market center. For example, the following is a table from TD Ameritrade's 606 disclosure for NASDAQ stocks during the third quarter of 2017.

| Routing Venue | Total Orders (%) | Market Orders (%) | Limit Orders (%) | Other Orders (%) |
|---|---|---|---|---|
| TD Ameritrade Clearing | 77% | 85% | 74% | 72% |
| Citadel Execution Services | 13% | 9% | 14% | 17% |
| Virtu Americas, LLC | 6% | 4% | 6% | 8% |
| Citi Global Markets | 3% | 0% | 5% | 0% |
| Two Sigma Securities, LLC | 1% | 2% | 1% | 3% |

Also, note that TD Ameritrade redirects a majority of its orders to its own clearing corporation, which itself produces a 606 disclosure documenting its order flow; this is a fairly common practice among the larger brokers. In order to account for order flow from a broker to its subsidiary clearing corp, I take the order flow payments made to the clearing corporation (found in their seperate 606 filing), weight them by the share of orders sent to the clearing corp, and add them back to the order flow payments to

the broker. This ensures that order flow sent to subsidiaries is accounted for in the final dataset. For instance, if TD Ameritrade Clearing sends 50% of its orders to Citadel, then the total percent of orders routed to Citadel is 51.5% (77% * 50% + 13%).

Along with this information, the disclosure states the average payment received by TD Ameritrade from each routing venue; this is the payment for order flow. However, this variable is a deterministic function of the order routing already present in the data. That is, market centers pay different amounts for different types of orders. The average payment reported by brokers is the total payment received divided by the number of shares routed to a certain market center. So, if a market center offers higher rebates for market orders than limit orders and a broker routes more market orders their way in some future time period, then the average payments reported would increase even though the actual rebates offered may not have changed. This makes it impossible to draw any information from the average payments statistic without an endogeneity problem; the data for the actual rebate values offered by market centers is unavailable, so there is no way of extracting any information from average payments. Instead, I simply keep track of which brokers offer order rebates with a dummy variable; brokers must disclose if they are receiving payment for order flow, so this data is available. I use this later on for a subgroup analysis. In total, the 606 data provides order routing data, or "market share" within brokers, for each market center specified by exchange and order type on a monthly basis.

# 4    Methodology

## 4.1    Model Formulation

Suppose that brokers route orders based on a the market center's execution quality, and its payment for order flow; realistically, these are the only possible variables a profit-maximizing firm would take into account besides for some sort of personal relationship with a market center. That sort of effect will be fixed over time, so it can be modeled with a constant term. So, in total, we have fixed effects represented by $\alpha$, execution quality represented by $X$, and payment for order flow denoted by $Z$; these variables determine $Y$ which denotes market share.

$$Y_{i,j,k,l,t} = \alpha_{i,j,k,l} + \lambda_1 \cdot \beta \cdot X_{j,k,l,t} + \lambda_2 \cdot \gamma \cdot Z_{j,l,t} \tag{1}$$

The subscript $i$ denotes the broker, $j$ denotes the market center, $k$ denotes the exchange, and $l$ denotes the order type. Additionally, $\lambda_1$ represents the weight put on a linear function of execution quality and $\lambda_2$ represents the weight put on a linear function of payment for order flow; I assume $\lambda_1 = (1 - \lambda_2)$, so brokers either weight towards execution quality or rebates. Theoretically, if a broker has strong monopoly power, it would have $\lambda_1$ closer to 0; on the other hand, if all of its executions were monitored for quality it may have $\lambda_2 = 0$. In short, brokers weight routing orders to provide the best service for its investor clients (execution quality) and/or the best product for market center clients (order flow).

## 4.2    Assumptions

There are several issues with running this regression as it is. Firstly, $Z$, payment for order flow, is unknown because there is no data available for it. Secondly, the lambda coefficients are, in actuality, dependent on the broker; brokers have different objectives so they likely vary. Thirdly, the unobservable fixed effects may be correlated with $X$. So, I regress whatever model I use by considering both random effects and fixed effects,

and choose the more effective model using a Hausman test. Thus, for now, suppose that equation (1) can be estimated consistently. Then, suppose we simplify the model to:

$$Y_{i,j,k,l,t} = \lambda_1 \cdot \beta \cdot X_{j,k,l,t} + \lambda_2 \cdot \gamma \cdot Z_{j,l,t} \tag{2}$$

This still cannot be estimated with the presently available information. So, instead, consider the equation without the inclusion of the variable $Z$ for payment for order flow.

$$Y_{i,j,k,l,t} = \beta \cdot X_{j,k,l,t} \tag{3}$$

Estimating this equation for a broker that does not receive payment for order flow will produce consistent estimates of $\beta$, since they put no weight on their non-existent payment for order flow; that is, this equation is equivalent to (2) for "unpaid" brokers because they cannot route to maximize rebates if they don't accept rebates to begin with. Hence, the $\beta$ estimate for equation (2) will be equivalent to that in (3) for unpaid brokers, since there's no omitted variable bias.

However, using this equation for a broker that does receive payment for order flow may produce inconsistent estimates since the functional form is incorrect. This can occur when we have a non-zero coefficient on $Z$ and a correlation between $Z$ and $X$. I believe the latter is highly unlikely. This is because $Z_{j,l,t}$ states the level of rebate offered for order flow at some point in time $t$. For it to have a causal relationship with $X_{j,k,l,t}$, we would need market centers either reacting to execution quality changes by changing rebates and/or responding to rebate revenue with changes in execution quality. The latter is impossible, because market centers cannot simply choose how well trades are executed on their markets. Better trade-matching algorithms may lead to better executions, but market centers cannot "undo" investments into improving their execution quality nor are these straightforward investments to make. So, if rebate revenue increased, a market center cannot cut costs by reducing execution quality. Additionally, improvements in execution quality tend to be a result of technological process (eg: faster computers); its non-trivial for a market center to improve execution quality by any premeditated amount.

On the other hand, it is also unlikely that market centers can respond to execution quality changes by changing their rebates. This is because market centers cannot change their rebates very often; there's a significant amount of paperwork involved and it can take several quarters for changes in the rebate structure to actually go through. So, market centers cannot react well to changes in execution quality by changing the prices they offer for rebate flow. Moreover, it is very difficult to predict when execution quality will fall, and, even if predicted, it may be a result of systemic problems in the market rather than being particular issue with the market center itself. In short, $cov(Z, X)$ is likely 0. Moreover, brokers are not supposed to, legally speaking, take rebates into account when routing orders; that is, $\lambda_2$ should be 0 for all brokers. I test for the possibility that $\lambda_2 \neq 0$ but assume that $cov(Z, X) = 0$.

So, I proceed with using equation (2), since it will not have omitted variable bias. We can then get an estimate of $\beta$ for unpaid brokers, $\beta_0$, and another for brokers that accept rebates, $\beta_1$. Also, note that any estimates of $\beta$ will actually be of $\lambda_1 \cdot \beta$. Then, running a Wald test on $\beta_0 \neq \beta_1$ would tell us if they differ in weighting execution quality. Recall that $\lambda_1 = 1$ for unpaid brokers, but it still may be less than 1 for brokers that accept rebates. So, the estimated coefficients of $\beta_1$ should be weaker than those of $\beta_0$. For instance, brokers should try to maximize price improvement for their clients; so we may see a positive coefficient for price improvment in $\beta_0$, but, since paid brokers may have $\lambda_0 < 1$, we may see a slightly less positive coefficient for $\beta_1$. On the other hand, brokers should minimize the time orders take to execute, so we should see the same effect but with negative coefficients for that variable. Essentially, the Wald-test between $\beta_0$ and $\beta_1$ would amount to checking if $\lambda_1$ for unpaid and rebate-accepting brokers are the same (assuming the true $\beta$ for both subgroups is the same). I also check how specific coefficients compare between $\beta_0$ and $\beta_1$. This avoids the assumption that both groups have the same $\beta$. Instead, this would give us the specific effects of how accepting rebates affects a broker's order routing incentives.

In short, I have to run several tests for any set of $X$ variables that I use to proxy for execution quality.

- Test whether inclusion of fixed effects is necessary

    - $H_0 : cov(\alpha, X) = 0$

    - Hausman Test on Equation (2) with Random Effects versus Equation (2) with Fixed Effects

    - Can do this test for unpaid brokers without assuming $cov(X, Z) = 0$

- Tests for differences in order routing objectives and $\lambda_2 = 0$

    - Assuming $cov(X, Z) = 0$, $H_0 : \beta_0 = \beta_1$

    - Will test individual coefficients as well

    - Wald Test between coefficients of estimates for paid and unpaid brokers

## 4.3  Variable Selection

With respect to execution quality, I have access to data on price improvement and spreads through the 605 disclosures. Firstly, considering price improvement, I have date on the number of price-improved shares, the average amount of price improvement, and the average time that price-improved shares took to execute. The latter two I consider in my regressions, and I use the former to construct the percent of shares that were price-improved. Additionally, I include an interaction term between percent price-improved and average price improvement; this variable is the expected amount of price improvement. Additionally, since the data shows the average time that all types of executions took, I construct a variable for the average time that a share takes to execute at a market center. These variables are denoted in my regression results as $PrImp\_Pct$, $PrImp\_AvgAmt$, $PrImp\_ExpAmt$, $PrImp\_AvgT$, and $All\_AvgT$. I do not include both $PrImp\_AvgT$ and $All\_AvgT$ in the same regression, since they are highly collinear. Additionally, I consider average effective spreads, the average difference between the NBBO spread (ask - bid) and the trade price, in my regression; tighter spreads generally indicate better executions. In total, I run several regressions using some combination of terms for the level of price improvement, a term for average time taken, and a term for spreads; I use the variables with the best explanatory power to construct the $X$ variable.

15

Additionally, as mentioned earlier, I run a pooled regression of all brokers and then another two for rebate-accepting brokers and unpaid brokers. This makes the weaker assumption that the coefficient on execution quality, $\beta$, is the same within the two subgroups rather than the same for all brokers. Doing so allows me to check the difference between the two coefficients to determine the how order routing changes in the presence of rebates. Furthermore, I run regressions for each broker as well to see if there are rebate-accepting brokers who still route as well as the average unpaid broker or vice versa; this means finding $\beta$ for each broker separately.

Lastly, I set up random effects regression as well. Specifically, I set up my random effects as:

$$\alpha = \mu + U_j + W_{ij} + R_{ijk} + S_{ijkl}$$

Here, $\alpha$ is the mean market share of a market center for all exchange and of all order types. Next, I include market center specific random effects, $U_j$, and then include random effects for each additional dimension. For $U, W, R, S$, I disclude one observation from each to prevent multicollinearity with $\mu$. Including random effects allows me to identify which market centers receive a significantly higher percentage of orders independent of execution quality. I test whether to use this model or fixed effects by running a Hausman test.

# 5 Results

## 5.1 Estimation Results

For the following tables, the value in the parenthesis is the standard (robust) error. The data set for all the tables is the same; orders of type "Other" were dropped since their execution quality statistics for price improvement are undefined. In the random effects table, there are binary variables present for each random effect that I've omitted from the output. Lastly, for the fixed effects table, the data set is demeaned within each "market center, broker, exchange, order type" subset.

### 5.1.1 All Brokers

Table 1: All Brokers (Random Effects)

| | Market Share | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | 0.089*** | | 0.093*** | |
| | (0.025) | | (0.022) | |
| PrImp_AvgAmt | 3.359*** | | 3.162*** | |
| | (0.800) | | (0.791) | |
| PrImp_ExpAmt | | 9.531*** | | 9.478*** |
| | | (1.519) | | (1.562) |
| PrImp_AvgT | −0.009 | −0.014* | | |
| | (0.008) | (0.007) | | |
| All_AvgT | | | −0.0003 | −0.0004 |
| | | | (0.0003) | (0.0003) |
| AvgEffecSpread | −1.473** | −0.757 | −1.365* | −0.743 |
| | (0.542) | (0.478) | (0.539) | (0.480) |
| Constant | 0.005 | 0.036** | −0.0005 | 0.032* |
| | (0.023) | (0.014) | (0.020) | (0.013) |
| Observations | 3,273 | 3,273 | 3,273 | 3,273 |
| $R^2$ | 0.690 | 0.689 | 0.690 | 0.689 |
| Adjusted $R^2$ | 0.650 | 0.650 | 0.650 | 0.649 |
| Residual Std. Error | 0.129 | 0.129 | 0.129 | 0.129 |
| F Statistic | 17.409*** | 17.398*** | 17.411*** | 17.384*** |

*Note:*        *p<0.05, **p<0.01, ***p<0.001

## Table 2: All Brokers (Fixed Effects)

| | Market Share | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | −0.027 | | −0.017 | |
| | (0.047) | | (0.046) | |
| PrImp_AvgAmt | 2.769*** | | 2.615*** | |
| | (0.609) | | (0.615) | |
| PrImp_ExpAmt | | 10.186*** | | 9.992*** |
| | | (2.570) | | (2.561) |
| PrImp_AvgT | −0.011 | −0.013 | | |
| | (0.009) | (0.009) | | |
| All_AvgT | | | −0.0001 | −0.0003 |
| | | | (0.0004) | (0.0003) |
| AvgEffecSpread | −2.561*** | −1.437 | −2.573*** | −1.477* |
| | (0.745) | (0.735) | (0.729) | (0.732) |
| Constant | 0.000 | 0.000 | 0.000 | 0.000 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Observations | 3,273 | 3,273 | 3,273 | 3,273 |
| $R^2$ | 0.019 | 0.020 | 0.020 | 0.017 |
| Adjusted $R^2$ | 0.018 | 0.019 | 0.018 | 0.016 |
| Residual Std. Error | 0.217 | 0.217 | 0.217 | 0.217 |
| F Statistic | 16.213*** | 21.762*** | 16.405*** | 19.019*** |

*Note:* *p<0.05, **p<0.01, ***p<0.001

### 5.1.2 Rebate-Accepting Brokers

Table 3: Rebate-Accepting Brokers (Random Effects)

| | Market Share | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | 0.063** | | 0.053** | |
| | (0.020) | | (0.019) | |
| PrImp_AvgAmt | 1.253* | | 1.425** | |
| | (0.536) | | (0.536) | |
| PrImp_ExpAmt | | 2.921* | | 3.024* |
| | | (1.249) | | (1.284) |
| PrImp_AvgT | 0.016* | 0.010 | | |
| | (0.007) | (0.006) | | |
| All_AvgT | | | 0.0004 | 0.0003 |
| | | | (0.0002) | (0.0002) |
| AvgEffecSpread | −0.468 | −0.285 | −0.591 | −0.323 |
| | (0.337) | (0.303) | (0.335) | (0.304) |
| Constant | 0.038* | 0.077*** | 0.049** | 0.078*** |
| | (0.018) | (0.010) | (0.016) | (0.010) |
| Observations | 706 | 706 | 706 | 706 |
| $R^2$ | 0.792 | 0.788 | 0.791 | 0.788 |
| Adjusted $R^2$ | 0.758 | 0.754 | 0.757 | 0.753 |
| Residual Std. Error | 0.049 | 0.050 | 0.049 | 0.050 |
| F Statistic | 23.337*** | 23.002*** | 23.233*** | 22.974*** |

*Note:* $^*$p<0.05, $^{**}$p<0.01, $^{***}$p<0.001

## Table 4: Rebate-Accepting Brokers (Fixed Effects)

| | Market Share | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | −0.027 | | −0.017 | |
| | (0.047) | | (0.046) | |
| PrImp_AvgAmt | 2.769*** | | 2.615*** | |
| | (0.609) | | (0.615) | |
| PrImp_ExpAmt | | 10.186*** | | 9.992*** |
| | | (2.570) | | (2.561) |
| PrImp_AvgT | −0.011 | −0.013 | | |
| | (0.009) | (0.009) | | |
| All_AvgT | | | −0.0001 | −0.0003 |
| | | | (0.0004) | (0.0003) |
| AvgEffecSpread | −2.561*** | −1.437 | −2.573*** | −1.477* |
| | (0.745) | (0.735) | (0.729) | (0.732) |
| Constant | 0.000 | 0.000 | 0.000 | 0.000 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Observations | 706 | 706 | 706 | 706 |
| $R^2$ | 0.020 | 0.016 | 0.021 | 0.016 |
| Adjusted $R^2$ | 0.015 | 0.012 | 0.016 | 0.012 |
| Residual Std. Error | 0.099 | 0.099 | 0.099 | 0.099 |
| F Statistic | 3.629** | 3.754* | 3.807** | 3.879** |

*Note:*                      *p<0.05, **p<0.01, ***p<0.001

### 5.1.3 Unpaid Brokers

Table 5: Unpaid Brokers (Random Effects)

| | Market Share | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | 0.085** | | 0.090** | |
| | (0.031) | | (0.028) | |
| PrImp_AvgAmt | 4.360*** | | 3.972*** | |
| | (1.078) | | (1.060) | |
| PrImp_ExpAmt | | 11.428*** | | 11.007*** |
| | | (1.912) | | (1.965) |
| PrImp_AvgT | −0.014 | −0.017 | | |
| | (0.010) | (0.009) | | |
| All_AvgT | | | −0.001 | −0.001* |
| | | | (0.0003) | (0.0003) |
| AvgEffecSpread | −1.964** | −0.836 | −1.748* | −0.807 |
| | (0.761) | (0.669) | (0.758) | (0.670) |
| Constant | −0.00005 | 0.018 | −0.007 | 0.015 |
| | (0.030) | (0.018) | (0.026) | (0.018) |
| Observations | 2,567 | 2,567 | 2,567 | 2,567 |
| $R^2$ | 0.663 | 0.663 | 0.664 | 0.663 |
| Adjusted $R^2$ | 0.623 | 0.623 | 0.623 | 0.623 |
| Residual Std. Error | 0.143 | 0.143 | 0.143 | 0.143 |
| F Statistic | 16.356*** | 16.391*** | 16.374*** | 16.395*** |

*Note:* *p<0.05, **p<0.01, ***p<0.001

## Table 6: Unpaid Brokers (Fixed Effects)

| | Market Share | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | −0.027 | | −0.017 | |
| | (0.047) | | (0.046) | |
| PrImp_AvgAmt | 2.769*** | | 2.615*** | |
| | (0.609) | | (0.615) | |
| PrImp_ExpAmt | | 10.186*** | | 9.992*** |
| | | (2.570) | | (2.561) |
| PrImp_AvgT | −0.011 | −0.013 | | |
| | (0.009) | (0.009) | | |
| All_AvgT | | | −0.0001 | −0.0003 |
| | | | (0.0004) | (0.0003) |
| AvgEffecSpread | −2.561*** | −1.437 | −2.573*** | −1.477* |
| | (0.745) | (0.735) | (0.729) | (0.732) |
| Constant | 0.000 | 0.000 | 0.000 | 0.000 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Observations | 2,567 | 2,567 | 2,567 | 2,567 |
| $R^2$ | 0.019 | 0.021 | 0.017 | 0.017 |
| Adjusted $R^2$ | 0.018 | 0.019 | 0.016 | 0.015 |
| Residual Std. Error | 0.231 | 0.230 | 0.231 | 0.231 |
| F Statistic | 12.636*** | 17.954*** | 11.353*** | 14.403*** |

*Note:*        *p<0.05, **p<0.01, ***p<0.001

## 5.2 Specification Tests

### 5.2.1 Random Effects versus Fixed Effects

The table below shows the results of each Hausman test for each subgroup and regression of the data. The columns correspond to the functional forms used in the regressions.

Table 7: Hausman Test Results
(Random Effects versus Fixed Effects)

|  | Wald Stat | | | |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| All Brokers | 29.46*** | 0.76 | 3.64 | 3.09 |
| Rebate-Accepting | 7.45 | 1.03 | 13.85** | 7.81* |
| Unpaid | 36.45*** | 0.66 | 4.16 | 3.68 |
| *Note:* | | | *p<0.05, **p<0.01, ***p<0.001 | |

For the following sections, I analyze the coefficients using whatever functional form is supported by the Hausman test; when $p < 0.05$, I use the fixed effects model for a regression.

### 5.2.2 Variable Exclusions

With respect to all brokers, most of the variables were significant. In the first two regressions, PrImp_AvgT, the average execution speed of price improved shares, was not significant. Similarly, All_AvgT, the average execution speed of any share, was insignificant as well. Average effective spread was highly significant in the first regression, but insignificant in the later three. All measures of price improvement were significant except in the first regression.

For rebate-accepting brokers, price improvement measures were significant in the first two regressions but not in the later two. On the other hand, spreads and execution speed were significant in the later two regressions but not the first two. However, for paid brokers, price improvement measures were highly significant in all regressions. Spreads

were significant in the first and third regression but not the others. Execution times were barely significant in all four.

Based on these results, it is difficult to rule out a particular regression model. The significance of execution speeds were relatively weak in all the models. However, these coefficients are important to keep in the model, since brokers report their execution speeds to draw in clients to use their service. Whether or not they are significant, they should be in the model, because there is likely still a causal link between order routing and execution speed. So, for the following sections, I continue my analysis with all four.

### 5.2.3 Subgroup Analysis

The two main subgroups are unpaid and paid brokers. I expect unpaid brokers to have higher coefficients on variables that increase execution quality and lower coefficients on those that decrease it. That is, the coefficients on percent price improved and average price improvement should be positive, and the coefficients on spreads and execution time should be negative.

Table 8: Wald Test between Coefficients
(Paid versus Unpaid Brokers)

| | Wald Stat | | | |
| --- | --- | --- | --- | --- |
| | (1) | (2) | (3) | (4) |
| PrImp_Pct | 1.86 | | 7.16** | |
| PrImp_AvgAmt | 4.22* | | 5.32* | |
| PrImp_ExpAmt | | 9.90** | | 17.63*** |
| PrImp_AvgT | 5.36* | 4.50* | | |
| All_AvgT | | | 17.58*** | 18.24*** |
| AvgEffecSpread | 2.90 | 0.34 | 0.02 | 0.70 |
| *Note:* | | *p<0.05, **p<0.01, ***p<0.001 | | |

For each subgroup, I compare coefficients using a consistent estimator. For example, when looking at function form (1), I use coefficients from the random effects model for

rebate-accepting brokers and those from the fixed effects model for unpaid brokers. In the table below, I've listed the Wald Stats ($(b_1 - b_0)^2/var(b_1 - b_0)$) for each pair of coefficients; I assume that $cov(b_0, b_1) = 0$ for all. The p-value for the tests corresponds to the probability of the null $H_0 : b_0 = b_1$.

In the above table, all cases of a significant difference between coefficients were those where unpaid brokers were prioritizing execution quality more than unpaid brokers. For instance, the coefficient on PrImp_AvgT, average execution time for a price improved share, was lower for unpaid brokers than paid brokers; this meant unpaid brokers put more weight on execution speed; the same applies to All_AvgT as well. Similarly, unpaid brokers put significantly more priority on price improvement in all the regressions except (1) for the PrImp_Pct coefficient. Lastly, for effective spreads, there was no significant difference between how well paid and unpaid brokers prioritized sending orders to market centers with lower spreads.

## 5.3  Coefficient Results

I expect the coefficients on PrImp_Pct, PrImp_AvgAmt, and PrImp_ExpAmt to all be positive, since more price improvement corresponds to better execution quality. On the other hand, the coefficients for the average time variables and spreads should be negative, since lower values of these correspond to better execution quality.

For the pooled regression of all brokers, the signs of the coefficients were all as expected except in the first regression. In regression (1), the sign on the PrImp_Pct was slightly negative; however, it was insignificant so I don't draw any conclusions from this. The coefficient on PrImp_ExpAmt in regressions (2) and (4) was approximately 9.5, which means that increasing the expected amount of price improvement by $0.01 will increase the market share of a market center by 1%. This seems reasonable, since an extra $0.01 is a meaningful amount of price improvement. Additionally, the coefficient on this variable in both regressions was extremely significant with a t-stat over 6. The coefficients on the average time variables were all negative but not very significant. This may be because different market centers have different objectives. For example, Citigroup's market center

primarily accepts large limit orders, so they would have high execution times but that would be a result of the orders that they're accepting. For future regressions, it may be important to control for the order sizes that a firm accepts when using execution time in a regression. Lastly, the coefficients on spreads were negative as expected; the coefficients regression (1) and (3) were significant but not the others. This may be a result of including expected price improvement in (2) and (4), since the two seem correlated. This correlation seems to be apparent in the below graph where the two variables were demeaned within each unique broker, market center, exchange, and order type combination.



Expected Price Improvement versus Effective Spreads (demeaned)
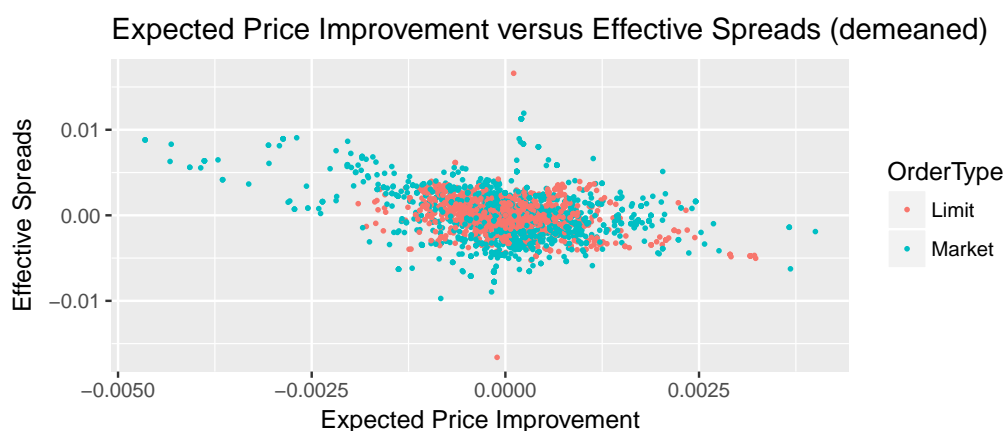
With respect to the two subgroups, all coefficients tended to be slightly less significant. For both brokers, all the signs on the coefficients were the same as for the pooled regressions. Between the two subgroup, unpaid brokers had coefficients of a higher magnitude than paid brokers for all but one coefficient (PrImp_Pct in regression (1)). This was hypothesized, because unpaid brokers do not face an agency problem and will route orders to maximize their clients' order execution quality. The t-stats for unpaid brokers were relatively higher than those for paid brokers. This may be due to payment for order flow adding more noise in order routing decisions for brokers accepting rebates. For example, the coefficient on PrImp_ExpAmt was highly significant ($p < 0.00001$) for unpaid brokers but not as significant ($p < 0.05$) for paid brokers. The significance of the differences between the two subgroups' coefficients was summarized in Table 8.

# 6 Conclusion

Brokers must choose to route orders either to maximize execution quality or maximize some other objective; legally, by the SEC's Regulation NMS, there should be no factor besides execution quality in their order routing decisions. However, although this is in violation of the law, brokers may route to to maximize rebates and then claim that they maximized execution quality after the fact. Moreover, it is incredibly difficult to prove that order flow was not routed to maximize execution quality after it is sent. The particular state of the market at the time an order is sent affects the location where it would receive the best execution. When brokers receive rebates, one might suspect that their order routing suffers even though, in theory, it is possible for a broker to accept rebates while still prioritizing execution quality.

However, based on the results, there does seem to be a significant difference between the order routing objectives of rebate-accepting brokers and unpaid brokers. Specifically, rebate-accepting brokers are relatively less inclined to route orders towards firms that offer more price improvement and faster execution speeds. These two components are likely the most important factors of execution quality, and paid brokers such as TD Ameritrade even advertise these factors to clients. Since the coefficients of these factors were significantly lower in magnitude compared to unpaid brokers, payment for order flow is likely causing an agency problem with rebate-accepting brokers. Rather than providing the best service for their clients, they seem to be maximizing their profit which is to be expected.

In response, brokers may make the argument that payment for order flow helps reduce commissions. This a valid point, since commissions tend to be significantly larger than the amount of price improvement that an investor may receive on their order. However, at the same time, it is likely that the reduction is significantly less than the loss of execution quality. Most retail investors are unaware of how stocks are actually executed and what types of price improvement they should expect. As a result, we can expect the elasticity of retail demand for a market center to be fairly low with respect to execution quality. On the other hand, commissions are straightforward and well known. So, the elasticity

of consumer demand with respect to commissions is likely much higher. Thus, brokers are unlikely to pass on all of their revenue from rebates back to retail investors. In short, the results from these regressions can still be used to argue that retail investors receive a net loss in welfare from routing to rebate-accepting brokers.

Understanding the extent of this issue is important as well. In the results section, I found a highly significant difference in the coefficient for expected price improvement between the two subgroups; the difference between the two from regression (2) was 8.51 with a standard error of 2.28. Now, suppose there was a paid and unpaid broker, each of which routed 100 million shares in the previous quarter (less than a third of what Citadel receives). The results suggest that a market center would have to offer the paid broker $0.0851 more in price improvement to receive another 1% of its orders; this would amount to an extra $3.4 million annually. Repeating this with the results of regression (4) outputs an extra $7 million annually. In reality, market centers cannot offer varied levels of execution quality to specific brokers; orders sent to their exchange are given what they deem to be the best execution irrespective of their source. So, the paid broker in this case would be failing to provide $7 million in price improvement for its clients. Furthermore, paid brokers, as a whole process, billions of shares every year; this means that the total welfare losses for investors from subpar execution quality is much higher. So, overall, these results find that payment for order flow significantly and negatively effects the order routing decisions of rebate-accepting brokers.

# References

Angel, J.J., Harris, L.E., & Spatt, C. S. (2011). Equity Trading in the 21st Century. Quarterly Journal Of Finance, 1(1), 1-53.

Battalio, R., S. Corwin, S., & Jennings, R. (2016). Can Brokers Have It All? On the Relation between Make-Take Fees and Limit Order Execution Quality. Journal Of Finance, 71(5), 2193-2238. doi:10.1111/jofi.12422

Battalio, R., Shkilko, A., & Van Ness, R. (2016). To Pay or Be Paid? The Impact of Taker Fees and Order Flow Inducements on Trading Costs in U.S. Options Markets. Journal Of Financial & Quantitative Analysis, 51(5), 1637. doi:10.1017/S0022109016000582

Battalio, R., & Holden, C. W. (2001). A simple model of payment for order flow, internalization, and total trading cost. Journal Of Financial Markets, 433-71. doi:10.1016/S1386-4181(00)00015-X

Chordia, T., & Subrahmanyam, A. (1995). Market Making, the Tick Size, and Payment-for-Order Flow: Theory and Evidence. The Journal Of Business, (4), 543.

Cimon, D. (2016). Broker routing decisions in limit order markets. Bank of Canada. *http://www.bankofcanada.ca/wp-content/uploads/2016/11/swp2016-50.pdf*

Dutta, P.K., & Madhavan, A. (1997). Competition and Collusion in Dealer Markets. The Journal Of Finance, (1), 245. doi:10.2307/2329563

Kandel, E., & Marx, L.M. (1999). Payments for Order Flow on Nasdaq. The Journal Of Finance, (1), 35.

Financial Industry Regulatory Authority, 2001. NASD Notice to Members 01-22. http://www.finra.org/industry/notices/01-22

Dennert, J. (1993). Price Competition between Market Makers. The Review Of Economic Studies, (3), 735.

O'Hara, M. & Ye, M. (2011) Is market fragmentation harming market quality? Journal of Financial Economics 100, 459-474

Maglaras, C., Moallemi, C., & Zheng, H. (2015). Optimal Execution in a Limit Order Book and an Associated Microstructure Market Impact Model. Columbia Business School Research Paper No. 15-60.

Parlour, C.A., & Rajan, U. (2003). Payment for order flow. Journal Of Financial Economics, 68379-411. doi:10.1016/S0304-405X(03)00071-0

Securities and Exchange Commision, 2005. SEC adopts regulation NMS and provisions regarding Investment Advisers Act of 1940. *https://www.sec.gov/news/press/2005-48.htm*