

National Best Bid and Offer - NBBO

The best (lowest) available ask price and the best (highest) available bid price to investors when they buy and sell securities. National Best Bid and Offer is the bid and ask price the average person will see. The Securities and Exchange Commission's Regulation NMS requires that brokers must guarantee customers this price.

BREAKING DOWN 'National Best Bid and Offer - NBBO'

The NBBO is updated throughout the day to show the highest and lowest offers for a security among all exchanges and market makers. The lowest ask price and the highest bid price displayed in the NBBO do not have to come from the same exchange. The best bid and ask prices from a single exchange or market maker are simply called "best bid and offer." Traders who want to execute orders larger than those available through the NBBO will want to know the other potential bid and ask prices at which they could execute their orders. They can find these in an exchange or market maker's "depth of book" data. Day traders usually use level 2 market-maker screens to see all the bids and offers for a particular stock.

The Consolidated Quotation System gives the NBBO for securities listed on the New York Stock Exchange, while the Unlisted Trading Privileges Quote Data Feed gives the NBBO for securities listed on the Nasdaq. *A shortcoming of the NBBO system is that the data may not be sufficiently up to date, so investors may not get the prices they were anticipating when their trades are actually executed. This problem is mainly of concern to high-frequency traders, whose trading strategies may fail if their orders aren't executed at a precise desired price. Another NBBO shortcoming is that Regulation NMS is difficult to enforce, because the fast pace of trading and the lack of recorded NBBO prices make it difficult to prove whether an investor received the NBBO price on a trade.*

<http://www.nanex.net/Research/IsNBBOIgnored.html>

Is the National Best Bid or Offer (NBBO) being Ignored?

Executive Summary

The NBBO lies at the heart of Regulation NMS (Reg. NMS) and is the key concept that assures investors are getting the best price when buying or selling stocks. However, due to the recent industry trend that emphasizes speed at all costs, the NBBO, in practical terms, no longer exists. There is no audit trail that can show definitively whether an investor received the best price on their trade. The regulators do not appear to understand the root of the problem because they continue to promote new regulation, when a simple and effective solution exists: enforce Reg NMS. Likewise, getting rid of quotes that only serve to manipulate others is as easy as enforcing Section 9 of the Securities Exchange Act of 1934.

If any new regulations are needed after enforcing existing ones, we think a minimum quote life of 50 ms makes the most sense. Just the discussion of implementing such a rule would expose how deeply flawed the system is today and would be sure to raise a lot of eyebrows. But to be clear, we do not advocate new regulation.

The Death of the NBBO

As the lifetime of a quote approaches zero, the arguments for and against a minimum-quote-life rule become more interesting. Let's suppose, for example, that the day has arrived where a few of the top HFT systems are able to send and cancel quotes in 1 nanosecond (ns). For reference, light travels 30 cm (1 foot) in 1 ns. At this rate, we could see 1 billion quotes per second per stock (qpss). A thousand active stocks trading at this rate would generate 1 trillion quotes per second (qps).

We admit that was extreme, but given the hyperbole from some HFT marketing groups, we couldn't resist. So let's slow things down by a factor of 1,000 and imagine a world where HFT systems can send and cancel a quote in 1 microsecond (us). At this speed we could expect 1 million qpss and a thousand stocks would generate 1 billion qps. Put another way, an active market with a speed limit of 1 microsecond would generate 1 billion quotes per second, requiring everyone receiving CQS to spend about 1,000 times more for telco and equipment.

Still extreme? Let's make things 1,000 times slower again and imagine a world where HFT systems can send and cancel a quote in 1 millisecond (ms). At this speed, we could expect 1,000 qpss and a thousand stocks would generate 1 million qps. Even at this rate, many quotes will expire before leaving exchange networks. Today we frequently see qpss rates of 2,000, with peaks in the 5,000 - 30,000 range. These numbers are growing at alarming rates, and within a year, if left unchecked, recipients of CQS will need to upgrade their telco to 10 gigabit.

It's helpful to keep in mind that section I.C.4 of Reg NMS (page 30) states:

Accordingly, one of the Commission's most important responsibilities is to preserve the integrity and affordability of the consolidated data stream.

And from the same document, page 410:

But in those limited contexts where the interests of long-term investors conflict with short-term trading strategies, the conflict cannot be reconciled by stating that the NMS should benefit all investors. In particular, failing to adopt a price protection rule because short-term trading strategies can be dependent on millisecond response times would be unreasonable in that it would elevate such strategies over the interests of millions of long-term investors – a result that would be directly contrary to the purposes of the Exchange Act.

Perhaps we need to modernize the definition of a quote, or maybe we need a new name for what used to be called a quote. Not long ago, when a trader (or auto-trading software) received a quote marked auto-execute, there was a reasonable expectation of hitting (trading at) that quote — the only real exception being that another trader might beat them to it. Today, under the same circumstances, there is a significant chance that the same auto-execute quote would have already expired in transit and not be honored; the speed-of-light just isn't fast enough.

"No one uses the SIP for the NBBO anymore"

The change in semantics becomes most important when we look at the definition of the NBBO. Per Reg NMS, the NBBO for a stock is defined to mean the best bid/offer sent by a market center to the Security Information Processor known as the SIP (CQS, UQDF). But no one uses the SIP for that anymore, we are often told, which would seem to violate the letter and spirit of Reg NMS. We'd like to note that last year, 2.5 million subscribers spent over \$450 million to receive and process CQS.

So what do they use for the NBBO if not from the SIP?

Each exchange computes the NBBO internally from direct connections to other exchanges. As the speed of trading increases, the likelihood of any two exchanges having the same NBBO decreases. Most of this is because of the pesky speed-of-light limitation.

So how does a trader know whether a trade was routed properly to the exchange with the best price?

He doesn't. It is impossible.

You see, each exchange's view of the other exchange prices only exists in memory on that exchange's machines. It is not recorded. There is no audit trail. Sure, each exchange provides book-level data, but that only includes prices for that one exchange — not the prices that existed on the other exchanges at the time of each order.

"It is impossible to verify that a trade received the best price"

If we are going to allow machines to trade faster and faster, then at the very minimum, there must be audit trail data that includes each exchange's view of top-of-book quotes for every linked exchange trading that stock. In other words, what now only exists in an exchange routing computer's RAM, needs to be captured and made available. The exchange routing simulation video below may help you visualize this. In the video, each box represents one exchange and the price information it has from the other exchanges -- that is the information which needs to be recorded to assure trade through price protection. Essentially, each interlinked exchange would end up recording its view of every exchange's top-of-the-book prices -- exactly what the SIP (the box at the bottom) does now.

Reg NMS has already addressed this issue: page 32:

If the benefits of a fully consolidated data stream are to be preserved, each consolidator would need to purchase the data of each SRO to assure that the consolidator's data stream in fact included the best quotations and most recent trade report in an NMS stock.

"The claim that aggregation causes the SIP to run slower is absurd"

So why not use just use the SIP? Auditing, maintaining, and improving one system is much better than 14 (one for each exchange). Well, because as some, including the SEC, have said: The SIP is not as fast because it has to aggregate information from the other exchanges.

Wait a minute.

How can an exchange ensure every order receives the best prices if it doesn't aggregate information from all the other exchanges? Take a close look at the simulation video and note that every exchange must essentially replicate the SIP's aggregation function. There is no way around this. Which means the claim that aggregation causes the SIP to run slower is absurd. Quote rate overload is the primary cause of latency in the SIP and direct feeds.

Besides, there are numerous places where Reg NMS language is expecting that order routing would reference the SIP such as this (emphasis ours) on page 314:

For instance, modifications to order routing and execution systems will need to be made to route and execute orders in compliance with the requirements of the Rule to prevent trade-throughs of protected quotations (which include, for instance, the ability to recognize quotations identified in the consolidated quotation system as manual quotations on a quotation-by-quotation basis).

And on page 423, the use of the adjective displayed referring to prices in the SIP:

Intermarket sweep orders must, by definition, be routed to execute against the full displayed size of protected quotations...

There is more.

A computer that aggregates information from other computers uses an aggregation policy that details how it selects between messages coming in at the same time from multiple computers (things like time-slice period, scheduling quantum, size of input queues, overflow behavior, etc.). Let's just say it's very complicated and has many dependencies: it's easy to make mistakes (or hide bias). The aggregation policy is important because it affects the 3rd criterion in NBBO selection: price, size, time. The SIP's aggregation policy is something that is fairly easy to verify. How easy is it to obtain and verify the 14 aggregation policies used by the 14 other exchanges? How do you know if exchange X's aggregation policy treats other exchanges equally?

It gets worse.

We have discovered a few anomalies when analyzing aggregation characteristics of CQS, which is the SIP for stocks listed on NYSE, AMEX, and NYSE Arca (home of many ETFs). One disturbing anomaly is that the SIP appears to be ignoring the timestamp in the quotes it receives from an exchange, using instead the actual time the SIP receives the quote, even if the timestamp in the quote is over 5 minutes late. That is, we found an example of the SIP treating 5 minute old data as real-time, affecting the NBBO in thousands of stocks. We will be publishing the results of our analysis shortly. The main point, is that this type of detection is impossible to carry out for individual exchange aggregation behavior; and if it was, it would take 14 times more work.

"We found an example of the SIP treating 5 minute old data as real-time, affecting the NBBO in thousands of stocks"

Even more disturbing, the SIP applies the timestamp to quotes and trades after these messages have been throttled and queued, so it is impossible to detect or measure latencies occurring within the system. We have published numerous studies on this problem because it was one of the prime factors in the Flash Crash on May 6, 2010. Recipients couldn't detect significant delays until the system suddenly became overloaded. If timestamps reflected the actual time a quote was generated, we would have known how dangerously close to saturation the SIPs had become weeks earlier. Sophisticated trading firms with direct exchange feeds had to be aware of this.

We understand how difficult it can be to grasp these problems. Understanding complex networked systems where the speed-of-light is the dominant source of latency is hard. When the fate of your nation's financial infrastructure is at stake, we think regulators should have a solid understanding of these issues.

Exchange Routing

The video below shows how orders flow between linked exchanges trading one stock. Orders are represented by triangles, and begin from the edge of the display and feed into one exchange. That exchange then updates and checks its internal table to determine whether to route the order to another exchange with a better price, execute it locally, or simply update the top of the book for that exchange. Only orders that affect the top of the exchange book are included to reduce clutter. The order is then transmitted to every other exchange making a market in the stock (via premium direct exchange feeds) as well as CQS which is shown at the bottom. When other exchanges receive the order, they update their internal tables in order to keep track of the other exchanges' top-of-the-book quotes.

The price data within each box therefore represents how that exchange views the prices on other exchanges trading that stock. Unlike this simulation which assumes a perfect world: due to the variations in distance and connection quality, system load, and other real-world imperfections, exchanges won't always have identical NBBO information. As update rates increase, the percentage of time that all exchanges have identical NBBO information will rapidly drop to zero.

A Proposal for Solving the "Payment for Order Flow" Problem

Allen Ferrell*

An issue that has increasingly occupied the attention of the Securities and Exchange Commission is "payment for order flow." This is the practice whereby securities markets compete for orders placed by brokers by providing side payments to brokers in return for brokers promising to send them investors' orders. Does this create inefficient nonprice competition between securities markets? This Article argues that it does, that all the proposed solutions (including the SEC's disclosure requirements) miss the mark, and that the problem is really a result of the SEC's regulation of the prices at which investors' orders must be filled. As this paper will show, permitting brokers to credit investors' orders with the National Best Bid or Offer (NBBO) price regardless of any price improvement realized on these orders would ensure an efficient allocation of investors' orders across securities markets.

I. Introduction

Among the most intractable and pressing issues in securities regulation today concerns the widespread practice by securities markets of paying brokers to route investors' orders to them for execution.^[1] In a typical "payment for order flow" arrangement, a securities market will pay a broker anywhere from one to three cents for each buy or sell order the broker sends to it.^[2] These "kickbacks" have angered numerous market entities, large and small: from the New York Stock Exchange (NYSE),^[3] which has watched as competing markets have captured a substantial and increasing portion of its equity trading, to individual investors who have filed lawsuits alleging that they paid too much for stock purchases or received too little from stock sales due to the corrupting influence of these side payments on their brokers.^[4] Concerns over payment for order flow have been intensifying, in part thanks to studies documenting that small investors' orders—which constitute the bulk of the orders that are routed as a result of these arrangements—are often filled by dealers at prices inferior to those available on the NYSE and other exchanges.^[5]

Regulators and legislators have not ignored this practice—far from it. For example, an entire section of the Securities and Exchange Commission's (SEC) study, *Market 2000: An Examination of Current Equity Market Developments*—the SEC's first comprehensive analysis of securities markets and their regulation since the early 1970s—is devoted solely to questions relating to order-flow payments and whether investors are receiving "best execution" of their orders.^[6] Order-flow payments have also been the subject of Congressional hearings,^[7] an industry roundtable discussion hosted by the SEC,^[8] as well as SEC releases and regulations.^[9]

Defenders of order-flow payments have not stood idly by as others have attacked the practice. Defenders have tapped into popular anti-regulatory logic to argue that these side payments are the product of healthy free-market competition between securities markets for

investors' orders.^[10] They believe that investors, and not regulators, should be the ultimate judge of whether a broker who routes orders to a particular market for execution is offering an attractive service. As in any competitive industry, they argue, brokers will cater to the preferences of their customers or risk losing business to a competitor. The logic of this position sounds right. Who would argue that a car manufacturer that received from its muffler supplier an annual rebate on its purchases in the form of a cash payment is thereby placed in a conflicted position? Presumably any rebate would be passed along to the car manufacturer's customers in the form of lower prices. Applied to the payment for order flow context, the same reasoning suggests brokers will use payments from securities markets to lower their commission rates. On this view, the real danger lies in the NYSE and other exchanges exploiting the current controversy to impose new, burdensome regulations on their increasingly successful competitors.^[11]

Against the backdrop of these differing points of view, this Article will argue that the conflict of interest brokers face as a result of these side payments is a serious problem to which there exists a fairly simple solution that has yet to be considered. In order to understand the nature of the problem, as well as the solution this Article will propose, it might be helpful to consider a market analogous, for relevant purposes, to the brokerage market. Suppose a condominium owner wishes to sell her condominium and knows that she can get at least its "listed price." Perhaps with some effort, a buyer who is willing to pay more than the listed price could be found. There are any number of reasons why such a buyer might exist: the buyer might have a pressing need for the particular type of condominium being offered, or the listed price (for whatever reasons) might understate the condominium's true market value. Of course, such a buyer might not exist. But in order to capitalize on the possibility of finding a buyer willing to pay more than the listed price, the condominium owner could hire an agent, with experience in these matters, to conduct a search for these buyers. The dilemma facing the owner is that it might be extremely difficult, if not impossible, to know whether the agent has done a thorough job. Maybe the agent was extremely diligent but there happened to be no buyers willing to pay more than the listed price. On the other hand, perhaps the agent did find a buyer willing to pay a small premium over the listed price, but there existed a large number of buyers who were willing to pay a much larger premium and who could have been found with a minimum amount of additional effort on the agent's part. Without re-doing the job the agent was hired to do in the first place, the owner may never know.

An investor wishing to sell a security faces a dilemma not unlike that of the condominium owner.^[12] A securities broker, like the real estate agent in the analogy, acts as an investor's agent—finding the securities market offering the highest price.^[13] There are often many markets to which a broker could potentially send an order. Much like buyers of condominiums, different markets might, and often do, offer different prices. The analog to a condominium's listed price in the securities context is the National Best Bid or Offer price (NBBO), a price that is publicly disseminated over the Consolidated Quotation System (CQS) or the NASDAQ Quotation Dissemination Service (NQDS).^[14] The NBBO is based on the various bids (and offers) at which securities markets publicly indicate they are willing to purchase (and sell) securities.^[15] *Under current law, a broker must ensure that an investor receives at least the NBBO. For a variety of reasons, however, the best bid disseminated by the CQS or NQDS will often not be the best price an investor can actually*

receive.^[16] In other words, a securities market often will place a lower bid price on the CQS than the actual price investors would receive if their orders were sent there. An investor placing an order with a broker, much like the condominium owner, really has no way of knowing whether the broker has actually found the best price as opposed to just the NBBO.

Indeed, a securities investor confronts, in some ways, a more serious problem than the condominium owner. The condominium owner usually does not have to encounter a marketplace in which her agent, along with others, are systematically offered kickbacks by buyers who wish to avoid having particular condominiums sold to other buyers who may be willing to pay higher prices. A securities investor should worry. In payment for order flow arrangements, a broker will usually promise, in return for the cash payments, to send all small "nonprofessional" orders (i.e., orders placed by relatively unsophisticated investors) to a particular dealer who will then automatically fill these orders at the NBBO, regardless of the opportunities available on other markets. The securities investor's problem is further aggravated by the fact that many small investors are probably not even aware that a broker has a range of choices of where to send an order. It would be as if the condominium owner believed that the buyer found by her agent is the only buyer that exists.

The failure of brokers to act as loyal agents of investors has three adverse effects on efficiency. First, auction markets, such as the NYSE, are in large part institutionally incapable of offering side payments to brokers and are, therefore, systematically disadvantaged by such nonprice competition. In order to receive payments, brokers will avoid sending orders to auction markets even if they offer investors, as they often do, the best prices. Second, even if all markets were capable of offering side payments, there would nevertheless be a distortion where any given order was routed. The distortion would only cease if securities markets, at that moment, could quickly adjust the size of their side payments to reflect the full amount they were capable of rebating. While there could conceivably exist an equilibrium where this happens, there is no reason to believe that this happy state of affairs will be the norm, especially in an environment as fluid and fast-paced as that of the securities markets. Third, there is a distortion in whether an investor will place an order in the first place. Investors will not have the necessary information to make informed adjustments in their holdings, whether this consists of buying or selling stock. In particular, there will be socially excessive trading if traders do not realize the full true costs of trading, such as the opportunity cost of foregone auction market prices.^[17]

Putting aside the distortions that order-flow payments create, there are distributional concerns implicated as well. Overwhelmingly, it is small investors, not sophisticated market participants, who have suffered from inferior prices received as a result of order-flow payments. While it is true that a significant percentage of individual investors' funds are invested through pension plans and mutual funds, a substantial portion of investment still takes place in the form of individual investors purchasing and selling stocks through brokers.^[18] This is likely to remain true for the foreseeable future.^[19]

The efficiency and distributional consequences that result from payment for order flow jeopardizes the two fundamental goals of securities regulation: ensuring that the securities industry is efficient, and that it treats all investors, small as well as large, fairly.

Unfortunately, the current regulatory regime intended to align investors' and brokers' interests is flawed. So too are the reform proposals that have commonly been advanced.^[20] There is, however, a straightforward solution, yet to be considered, which would completely remove a broker's temptation to place its interests over those of its customers in routing orders.

The solution to the problem is this: If a broker could choose to provide a small investor with whatever the NBBO was at the time of order execution, irrespective of the price that was actually received, the conflict of interest created by order-flow payments would be resolved. To see how this proposal would work, consider two alternative ways the agent of the condominium owner could be compensated. The agent might charge a flat percentage of whatever the sales price of the condominium happens to be, say five percent, or she could charge a certain percentage of the listed price, say two percent, and keep for herself any amount the buyer was willing to pay above the listed price. The first payment schedule could very well lead to a lackluster search by the condominium owner's agent. After all, the agent would only enjoy five percent of the gains that result from exerting additional effort to find a buyer willing to pay more than the listed price. Under the latter payment schedule, however, the agent would have the proper incentive to find the buyer offering the highest price. Failure to do so would only come at the agent's expense. This would hold true regardless of the extent of the owner's knowledge of possible buyers or whether a buyer could offer the agent a kickback to sell the condominium to him rather than someone else. In economic terms, the misalignment of incentives has been removed since the agent internalizes all the effects of its choices at the margin. This is the standard way agency problems are resolved.

Allowing brokers to credit small investors with just the NBBO at the time of execution, call it the "NBBO pricing option," is equivalent to the "two percent of the listed price/everything above that to the agent" payment schedule mentioned above. Such a scheme, which would run afoul of current law, would ultimately redound to the advantage of investors. The brokerage industry is highly competitive and there is little doubt that investors, small and large alike, are aware of the size of the commission rates brokers charge.^[21] Any benefits brokers received as a result of finding superior prices would be passed along to investors in the form of lower commission rates.

Under the proposal outlined in this Article, a securities broker could choose which payment schedule—the current one or the NBBO pricing option—it wished to use for small orders. Small investors who are confident that they can monitor brokerage selection of securities markets could select a broker who did not use the NBBO pricing option. Investors who are not as confident in their ability to make an intelligent choice, or not even aware that such a choice exists, could select a broker based just on its commission rate, as many currently do, and still be assured that their broker has their best interests in mind. Brokers who are able to offer the lowest commission rates will tend to be the ones utilizing the NBBO pricing option. The SEC should help investors who do not want to base their decision solely on commission rates by gathering and releasing statistics that indicate how effective brokers are in finding the best possible prices.^[22] This step, incidentally, should be taken regardless of whether the proposal in its entirety is adopted or not.

Besides removing the conflict of interest created by order-flow payments, and the associated adverse efficiency and distributional consequences, the proposal has other benefits. Other areas of potential divergence between brokers' and investors' interests will also be resolved. In particular, the question of how much expense a broker should ideally incur in attempting to locate and route an order to the securities market capable of providing the best price will also be effectively resolved. As a result, some of the complicated current regulatory structure that has been erected to deal with these problems could be discarded with the proposal's implementation. Finally, an important side effect of the proposal would be to reduce intermarket fragmentation of securities trading.[\[23\]](#)

The preceding discussion of the problem of broker opportunism obviously leaves important points in need of further explanation and elaboration. Part II of the Article provides the details necessary for understanding the institutional framework in which investors, brokers, and securities markets operate. Parts III and IV will then, respectively, examine the conflict of interest that can arise when a broker acts as an investor's agent in deciding where an order should be sent for execution and the adverse consequences of this conflict. The current regulatory attempts to address this problem, and their shortcomings, will be dealt with in Part V. Part VI will provide a further explanation of the proposal and address possible criticisms, while Part VII will take a critical look at the most popular competing reform proposals.

II. Institutional Framework

A. The Role of Brokers

Although by no means are all orders channeled to securities markets through brokers,[\[24\]](#) most trades, especially smaller ones, begin with an investor placing an order with a broker who then ensures that the order is executed.[\[25\]](#) In return for a commission, a broker acts as an investor's agent in selecting which securities market should receive the order for execution. For many stocks, especially ones that are among the more actively traded NYSE-listed securities, there will be a number of markets from which to choose. A broker, for instance, might very well have the following choice of markets:

- (1) the NYSE;
- (2) one of the five regional exchanges;[\[26\]](#)
- (3) the "upstairs" market;[\[27\]](#)
- (4) an over-the-counter dealer who trades in the particular stock;[\[28\]](#)
- (5) a proprietary trading system such as Instinet;[\[29\]](#)
- (6) a foreign exchange such as the London Exchange;[\[30\]](#)
- (7) acting as a dealer by filling the order itself (internalization); and

(8) internally crossing investors' orders (agency crosses).

There is no easy answer to the question of where a broker should send an order. Which market will offer the best price for an order will often depend on an order's particular characteristics, such as whether the order is so large that the trade is unable to clear at the current market price (liquidity costs) or the extent to which potential contra-parties view the trade as motivated by private information about the stock's likely future value (perceived informational content of an order). For instance, the "upstairs" market is typically used for executing large orders in order to minimize the cost of liquidity.[\[31\]](#) Another example is the use of proprietary trading systems that allow traders to mask the size of their trade until a contra-party is found—a beneficial feature if large trades are viewed by market participants as more likely to be informationally motivated.[\[32\]](#)

Important as it is, price is not the sole factor a broker needs to consider. There are some investors who place primary importance on establishing their desired positions quickly. There are any number of reasons why an investor might be willing to pay a premium, in the form of an inferior price, for the provision of immediate liquidity. An investor might want to take advantage of a temporary divergence in prices for the same asset in different markets[\[33\]](#) or may be attempting to put in place an intermarket hedge where a delay of even a few seconds in establishing one of the requisite positions could be very costly.[\[34\]](#) Also, an investor may simply want to avoid being exposed to the risk that the security's price will change while the best possible contra-party offer is being sought.[\[35\]](#)

Brokers handling orders placed by this type of investor can turn to various markets to satisfy the need for immediate liquidity. Brokers can send orders to dealers who can immediately fill them by trading against their own account, although such trades will often have to settle for a lower price as compensation to the dealer for providing this service. Automated systems have been developed which enhance the ability of investors to execute quickly trades at the best available dealer price. NASDAQ's Small Order Execution System (SOES), for example, allows brokers to have trades, usually in increments of one hundred shares, automatically routed to the dealer offering the best posted quotation for execution, unless the trade can be crossed with a superior limit order stored in the SOES Limit Order File.[\[36\]](#)

A broker more concerned with getting the best price for an investor, on the other hand, can send an order to one of the exchanges, which are generally considered slower than a dealer, but often offer better prices.[\[37\]](#) Brokers can also rely on floor traders to handle an order with instructions that vary depending on an investor's need for immediacy. Floor traders are sometimes given almost complete discretion in deciding when, and even if, an order is executed, depending on the trader's sense of the market (so-called "not held" orders). Brokers can also attempt to get the best possible price, although at the expense of immediacy, by utilizing various new trading structures. The Arizona Stock Exchange, for example, batches orders over time for simultaneous execution at a single price.[\[38\]](#)

Price and the need for immediacy do not exhaust the list of factors that a broker should consider in deciding how to handle and route orders. Perhaps the most important additional factor is the cost savings that accrue from having uniform procedures for handling large

numbers of small orders. It is often simply not worth tailoring placement and execution services to meet each small investor's individual preferences. Indeed, it might not be economical for a broker to always get the best possible price for a particular order due to the cost of routing each order to the market which offers the best price at a particular point in time.[\[39\]](#) The impressive number of small orders that are routed for execution by automated handling systems operated by the exchanges—systems that are typically designed to receive smaller orders electronically from brokers and route them to the exchange floor for execution, is a powerful testament to the cost effectiveness of these systems.[\[40\]](#)

Perhaps the best example of this is the NYSE's SuperDOT system. This system accepts market orders up to 30,099 shares and limit orders up to 99,999 shares from member firms and routes them directly to the exchange floor where the specialist, the exchange-designated dealer in a security, attempts to get the best price for the order.[\[41\]](#) The SuperDOT system handles approximately 85% of all NYSE trades accounting for 33% of NYSE share volume.[\[42\]](#) Larger orders needing more specialized treatment are typically not set through the SuperDOT system, but are instead handled by floor brokers who can provide tailored execution services.[\[43\]](#) The other exchanges have similar small-order handling systems available for their members and, unlike SuperDot, execution is completely automated.[\[44\]](#) Small orders for stocks traded on the NASDAQ are typically handled by similar systems—either SOES or small-order handling systems, usually good for orders of up to 2000 shares, operated by over-the-counter dealers.[\[45\]](#)

Besides avoiding the costs of providing more individualized treatment, these automated systems result in other savings as well. They enable brokers to avoid the errors in handling that would inevitably occur if the process were not automated. Manual handling of large numbers of trades can result in a significant number of errors as the "back-room" crisis in the late 1960s amply demonstrated.[\[46\]](#) Small-order handling systems also minimize exchange service fees. Specialists participating in SuperDOT, MAX, P-Coast, and PACE generally do not charge commissions for handling small orders that are routed to them through these systems.[\[47\]](#)

In short, there are a number of factors that will go into what constitutes the appropriate market for a particular order.[\[48\]](#) The most important considerations typically are:

- (1) price;
- (2) the need for immediacy;
- (3) costs of providing individualized treatment;
- (4) reliability of the routing and execution system; and
- (5) exchange service fees.

The moral of all this is simple: Whether a broker has chosen the appropriate market for a particular order can be very difficult to ascertain, especially by an outside observer, given all the various, and sometimes competing, considerations involved.

B. The Securities Markets

1. *Differences in Market Structure: Auctions and Dealers*

Among the most important decisions a broker must make is whether to send an order to an auction or dealer market for execution. In an auction market, investors trade with each other by submitting an order to a broker who then, acting as the investor's agent, brings the order to a centralized location—historically a physical floor—to interact with other investors' orders. Investors' buy and sell orders are matched on the exchange floor according to a complex set of trading rules.^[49] On the NYSE, for example, market orders are filled, depending on who offers the best price, by either the specialist, floor traders (the "crowd") who represent other investors' orders or trade on their own accounts, or the specialist's electronic book which stores investors' limit orders until they can be executed. In dealer markets, on the other hand, investors trade not with each other but with a dealer who maintains its own portfolio of stocks. Dealers maintain a "spread," the difference between the price the dealer is willing to sell the stock (the offer) and the price the dealer is willing to buy (the bid), as compensation for their intermediation.^[50]

While important, the distinction between auction and dealer markets often gets blurred in practice. The NYSE is not a pure auction market since it relies on specialists, as do the other exchanges, who have an affirmative obligation to trade against their own account when necessary to balance out temporary mismatches in supply and demand for a particular stock.^[51] On the other hand, the SEC has recently required NASDAQ dealers, as well as specialists, to display the price and full size of customer limit orders in their quotations when these orders represent an improvement over the dealer's quotation.^[52] Investors' orders routed to NASDAQ, as in an auction market, can now directly interact without dealer intermediation.

2. *Competition Between Auction and Dealer Markets*

Competition between auction and dealer markets for investors' orders is fierce. While auction markets have historically dominated the bulk of equity trading in the United States, times have changed. Although the NYSE has many competitors, easily its most potent competition for order flow comes from dealers—whether it is NYSE member firms who internalize investors' orders by trading against them on their own account,^[53] specialists on the regional exchanges,^[54] or third-market dealers. Regardless of whether a company decides to list its stock on an auction market, dealers can still compete for orders in that stock.^[55]

The differences in the market structure of auction and dealer markets have important consequences for how these markets compete for order flow. To a significant extent, auction markets are institutionally incapable of providing side payments to brokers for routing orders to them. Nor can brokers, by and large, internalize order flow by vertically

integrating with exchange specialists. This is due to the simple fact that auction markets are characterized by the matching up of investors' orders without dealer intermediation. There is simply no dealer who participates in almost every trade and should, therefore share with a broker the profits arising from order execution.[\[56\]](#) While there are exchange dealers, such as floor traders and specialists, they participate in only a minority of all exchange trades.[\[57\]](#) As James Shapiro, an economist for the NYSE, explained:

Under current conditions, it is hard to see how an auction market, like the NYSE can compete effectively by paying for order flow, for the simple reason that neither the exchange nor the specialists on the floor earn the entire spread, as off-exchange dealers do. The spread on the floor is dissipated among customers. . . . [I]t would be impossible for NYSE specialists to pay as much for order flow as 'pure' dealers do.[\[58\]](#)

Dealer markets, in sharp contrast, make extensive use of side payments to attract order flow. As of 1993, according to SEC estimates, between fifteen and twenty percent of all orders in NYSE-listed stocks were routed pursuant to cash for order flow arrangements.[\[59\]](#) Even this figure does not tell the full story. Cash payments are not the only kind of side payments dealers can offer brokers in exchange for order flow. Dealers also offer brokers soft-dollar payments, such as sharing investment research or providing clearance services, in exchange for order flow. Sometimes the consideration for routed order flow is order flow itself. A broker-dealer will send order flow to another broker-dealer for execution in return for that broker-dealer returning the favor. Profit-sharing agreements can also operate as an inducement for routing order share. Some regional specialists have entered into "joint venture" agreements whereby the specialist agrees to share a portion of its profits with a broker in return for the broker routing orders to it.[\[60\]](#)

As important as these transfers, monetary or otherwise, are to dealer markets in attracting order flow, focusing exclusively on them would be misleading. Dealers that are vertically integrated with brokers can have orders routed to them, without having to make any payments, by merely having their brokerage divisions send orders to them for execution. Side payments are unnecessary since such transfers would function only to transfer value from one division of the firm to another. The broker-dealer will internally capture any profits that would have been transferred from the broker if they were not integrated. In other words, internally routing order flow can be the economic equivalent of a payment for order flow arrangement.

Vertical integration can be achieved in a variety of ways. Partly as a result of NYSE Rule 390's limitations (soon to be lifted) on NYSE members' ability to provide over-the-counter dealer services, many member firms in the last twenty-five years have acquired regional specialists to whom they send orders. With the exception of the CSE, the BSE, and the PSE, all the exchanges assign only one specialist to a security. A broker who sends order flow to an exchange where it owns the only specialist assigned to the stock can be assured that, to the extent there is specialist participation in the exchange's trades, orders it sends to the exchange in that stock will be internalized. Moreover, both the CSE and the BSE, which do not have monopoly specialists, have implemented SEC approved "preferencing" programs whereby a broker can substantially improve the probability that an affiliated specialist on one of those exchanges will fill any orders that are routed there.[\[61\]](#) A dealer "preferenced" by a broker will typically receive the broker's orders for execution instead of

competing with exchange dealers so long as it offers a price that is as good as its competitors.^[62] While "preferencing" is often a way for a broker-dealer to internalize order flow, exchange specialists can also pay a broker for preferencing it,^[63] much as over-the-counter dealers make payments for order flow. Brokerage firms not only have the option of internalizing order flow through the use of affiliated regional specialist units, but also, due to exchange rule changes on the NYSE and AMEX in 1986, can acquire specialist units on the national exchanges.^[64] Since 1986 a number of national exchange specialist units have been acquired by large brokerage firms.^[65] Of the NYSE's thirty-eight specialist units, ten are affiliated with broker-dealers making markets in 1,061 of the 3,308 securities traded on the NYSE. Six out of AMEX's thirteen specialist units are affiliated with broker-dealers accounting for 403 out of the 896 securities traded on the AMEX. These broker-dealers often have systems that automatically direct order flow to an affiliated specialist.^[66] Finally, brokerage firms can internalize order flow by sending orders overseas, on the so-called "fax market," for execution by a foreign dealer subsidiary.

III. Brokers' Conflict of Interest

In order to place in sharper focus the nature of the relationship between brokers and investors and where it can go awry, brokers will be thought of as offering a product—placement and execution services—to potential customers—investors wishing to buy or sell stock.^[67] The price of this product is the broker's commission rate, while its quality can be thought of as how well the broker does in selecting a securities market. A broker offers a low-quality product if it could obtain better prices for investors' orders but, for whatever reason, chooses not to do so.

With this conceptualization, the question of brokers' conflict of interest reduces to this: Under what circumstances will brokers only offer a low-quality product to customers who want, and are willing to pay for, higher quality? If this does occur, an agency problem exists—a broker's interests would then be in conflict with those of its customers. The literature on agency problems demonstrates that conflicts can arise, even in a competitive market, when three conditions are present: (1) some customers, although aware of price, do not monitor the quality of a product; (2) sellers can roughly distinguish between monitoring and nonmonitoring customers; and (3) providing quality is costly.^[68] In such a market, high-quality products will tend to be under-supplied as the provision of quality will not receive a sufficient premium to warrant the cost. The larger the number of nonmonitoring customers, the greater the ability to distinguish between monitoring and nonmonitoring customers; or, the more costly quality is, the higher the likelihood that nonmonitoring customers will not be provided the level of quality they might want.

All three conditions are satisfied in the market for placement and execution services to an extent sufficient to create a substantial divergence between brokers' and investors' interests. To show this, each of the three conditions will be examined in turn.

A. Investors' Ability to Monitor Brokers

The market for placement and execution services is highly competitive. There is no question that brokers compete fiercely with each other for business.^[69] Securities markets,

in turn, compete strenuously with each other for the order-flow brokers' control. The competition for order flow between auction and dealer markets is one illustration of this. Auction and dealer markets compete among one another as well.[\[70\]](#) As global competition for equity trading continues to intensify, competition will only increase.

This competition has been reflected in the substantial reduction in overall commission revenues since the abolition in 1975 of fixed commissions. In the last twenty years, NYSE members' securities commissions, for instance, have declined from approximately fifty percent of total revenues to only fifteen percent.[\[71\]](#) The downward pressure on commission rates is not surprising since they tend to be well-advertised and easy to compare. Rates typically consist of flat fees for trades up to a certain size and a cent or two for each share traded above that level.[\[72\]](#)

Evaluating the quality of competing brokerage products, on the other hand, is often far more difficult. While it is true that investors can easily judge brokers by certain criteria, such as ease of access to account information, determining whether an order has received the best possible price is far more involved.[\[73\]](#) It may very well be prohibitively expensive for many small investors to acquire the necessary expertise and information to make a meaningful judgment about whether a broker has sent their orders to the appropriate market.[\[74\]](#) Indeed, many small investors are probably unaware that they are uninformed.[\[75\]](#)

It is true that there are firms, albeit few in number and small in size, to whom investors can turn for brokerage evaluation.[\[76\]](#) Unfortunately, these companies use substantially different methodologies in evaluating and comparing brokerage services; whose approach is the best is a matter of considerable dispute.[\[77\]](#) In any event, these firms tend to cater to institutional investors and their special trading needs.

Besides using third-party evaluations, another standard way for uninformed customers to compensate for their inadequate monitoring is to free-ride off of those who do monitor. In other words, simply buy the same product, at the same price, that monitoring customers do. This is, however, unlikely to be a successful strategy for small investors in the market for placement and execution services. What might constitute the optimal handling of orders placed by sophisticated traders, who are likely to monitor brokerage quality, will often not be the same for the orders of smaller, less informed investors. The extent to which the handling of an order so often depends on factors correlated with an investor's likely sophistication is striking—trade size, an order's perceived informational content, the need to quickly establish a position, and the opportunity cost of using automated routing and handling systems all play important roles.[\[78\]](#) Moreover, a small investor might very well have to incur substantial costs in order to determine whether she is in fact receiving the same services as monitoring investors, who often require fairly individualized treatment.

The lack of monitoring by small investors is more than mere speculation. Small orders have been routinely routed to securities markets offering inferior prices. One study analyzed the prices that all orders in 500 NYSE-listed stocks received in 1988 and 1989.[\[79\]](#) The study found that in 1988 traders who placed orders for less than 400 shares and had their orders routed to a non-NYSE securities market received, on average, 1.07 cents less per share than

similarly sized orders sent to the NYSE. For small orders routed to third-market dealers, the disparity between NYSE prices and prices received was even larger, at 1.51 cents per share.^[80] The data from 1989 is similar: for orders less than 400 shares non-NYSE securities markets offered, on average, prices lower by 1.22 cents per share, while third market dealers' prices were lower by an average of 1.58 cents per share.^[81] Disturbingly, it is smaller orders that have been increasingly sent by brokers to non-NYSE securities markets, often in exchange for cash payments.^[82]

In contrast, the relative performance of the NYSE and non-NYSE securities markets for orders larger than 400 shares was very different. The overall disparity in 1988 for orders in the 500-900 share range was -0.22 cents per share—non-NYSE securities markets actually offered slightly better prices—while in 1989 the difference was .45 cents.^[83] Non-NYSE securities markets performed even better when executing trades in the 1,000–1,900 share class.^[84]

This failure to monitor brokerage quality is consistent with growing empirical evidence that noninstitutional investors often do not have even a basic understanding of how financial markets work. Consider just a few of many possible pieces of evidence. Individual investors tend to increase their purchases of mutual funds due to recent good performance, even though this is a very poor indicator of future returns.^[85] Similarly, commodity funds are sold to individual investors on the basis of exceptional return performance before issuance, but often fail to deliver similar results after they become public.^[86] These examples do not necessarily indicate that these investors are irrational but rather might highlight investors' responses to the cost of acquiring information.^[87] Whatever the reason, uninformed behavior is far from uncommon.

B. Brokers' Ability to Identify Investors Unable to Monitor

A lack of monitoring by uninformed investors is not a sufficient condition for the existence of an agency problem. A number of investors will undoubtedly find monitoring worthwhile, at least partially, especially if they trade often or in large amounts.^[88] Brokers will have to be able to identify, at least roughly, into which group an investor falls. To the extent that they cannot but still offer a product of inferior quality, they run the risk of losing the business of investors who do monitor. The greater the ability to identify nonmonitoring investors, the greater the incentive to reduce the quality provided to these investors as a way of cutting unnecessary costs.

As it turns out, brokers are often readily able to identify into which category an investor likely falls. Besides observing the characteristics of an order, such as trade size, brokers and more sophisticated investors often have long-standing personal relationships and frequent interactions.^[89] Involved discussions between brokers and institutional investors over commission rates, as well as other matters, occur often. Order flow arrangements between dealers and brokers typically include an agreement that brokers will not send the dealer "professional" orders—an agreement obviously premised on the ability of brokers to segregate orders based on investors' sophistication.

C. Cost of Quality

The final question is whether providing quality is costly. If it is, then brokers will tend to provide only low quality products, i.e., routing investors' orders to securities markets which are not necessarily offering the best price, regardless of nonmonitoring investors' wishes.^[90]

There are various costs that a broker must incur in attempting to route a particular order to the market offering the best possible price. First, there are the costs of searching for the market. Checking the prices offered by the various markets conducting trading in any given security, and routing individual orders based on this information, may be a time-consuming task if done for every order.^[91] Automatically routing small orders to a particular securities market, irrespective of the price being offered, can eliminate these costs. Next, routing orders to some markets can incur various handling fees.^[92] Proprietary trading systems, for instance, often charge a fee for using their systems.^[93]

Most importantly, the securities market offering the best price will not necessarily be the market offering the largest side payment. Indeed, the ability to offer a large side payment can be the direct result of profits arising from executing orders at an inferior price. When this occurs, brokers who route orders to the securities market offering the best price will incur the opportunity cost of these foregone side payments.^[94] This represents a real cost to the broker regardless of whether a side payment is in the form of cash, a nonmonetary benefit, or internally captured dealer profits.

Since the provision of quality is often costly, brokers will tend to offer only low-quality services to small investors. In turn, this implies that securities markets, when they can, will attempt to attract the order flow brokers' control not by competing on providing the best price but on providing the largest side payments. This has three distinct adverse effects on efficiency as well as worrisome distributional consequences.

IV. Adverse Consequences of the Conflict of Interest

A. Efficiency Consequences

1. *Penalization of Certain Types of Securities Markets*

As the discussion in Part II emphasized, auction markets are, to a significant extent, institutionally incapable of offering side payments to brokers.^[95] Nor can they, for the same reasons, enable broker-dealers to accomplish the economic equivalent through internalization of order flow. Dealer markets, on the other hand, suffer from no such handicap. Through order-flow arrangements, dealers are able to clear small orders at the lowest bid or highest offer legally permissible—the NBBO—and return a portion of the profits in the form of side payments to the brokers who routed the orders there in the first place.^[96] Given this difference in market structure, small orders are more likely to be routed by brokers in search of side payments to dealer markets, irrespective of the prices available on the NYSE.^[97] The inability of auction markets to translate their often superior prices into side payments to brokers creates a distortion in where orders are routed. Any given order will not necessarily be sent to the securities market that values that order the most.

The statistics paint a troubling picture for the NYSE. Its dealer competitors have been increasingly successful in attracting order flow in NYSE-listed stocks from the NYSE, especially in small orders. As a percentage of all the transactions in NYSE-listed stocks, the third market's share increased from 2.39% in 1976 to 10.77% in 1995. For the regional exchanges, their percentage of NYSE-listed stock transactions has risen from 11.53% in 1976 to 19.01% in 1995. The third market's and regional exchanges' increases in NYSE-listed trading as measured by share volume is far less impressive—implying that many of the orders increasingly being sent to them are small in size.^[98] In contrast, the percentage of transactions in NYSE-listed stocks occurring on the NYSE dropped from 86% in 1976 to 70.22% in 1995.^[99] As of 1996, the NYSE had only 47% of market share for orders less than 1,000 shares.^[100] Some observers have even predicted that if current trends continue, within a decade less than half of all trades in NYSE-listed stocks will occur on the NYSE.^[101]

The popularity of the NYSE's dealer competitors is troubling since several studies have documented that the NYSE often offers prices significantly better than the NBBO. One study examined transactions in 500 NYSE-listed stocks over a two-year period.^[102] Prices received in non-NYSE transactions, as compared to trades conducted on the NYSE, were inferior, on average, by 0.69 cents per share in 1988 and 0.98 cents in 1989. Third-market dealers, for all order sizes, executed trades at prices inferior to those obtained on the NYSE.^[103] Another study analyzed every trade in NYSE-listed stocks occurring in 1989.^[104] It found that NYSE executions enjoyed an average price improvement over non-NYSE trades of 0.79 cents per share.^[105] A third study examined prices received by 40,865 market orders electronically routed to securities markets by three large retail brokerage firms on April 25, 1991 and August 20, 1991.^[106] It found that similarly sized orders in the same stock received an average of 3.5 cents per share more on the NYSE as opposed to the regional exchanges.^[107]

The general ability of auction markets often to offer better prices than dealers is not surprising. One of the advantages of an auction market is its ability to directly cross market orders thereby avoiding the spread that dealers charge.^[108] All three studies referred to above attributed the NYSE's superior performance, in large part, to the percentage of trades executed within the NBBO spread. Roughly 62% of all NYSE trades in 1988 and 1989 that occurred when the NBBO spread was one-quarter, for instance, were executed within the spread, as compared to approximately 38% for third-market trades.^[109] Studies of spreads on the London Exchange, a dealer market, for stocks listed on continental auction markets have reached similar conclusions, finding that auction markets often provide superior prices.^[110] It should be noted, however, that the NYSE's lead in executing trades within the spread is probably not now as significant as when the studies were conducted, given the increased use of price improvement programs by third-market dealers.^[111]

In contrast to auction markets, dealers are normally the contra-party to orders sent to them for execution and, accordingly, will maintain a spread to cover the costs of their intermediation services. A dealer will, for example, have to unwind stock positions acquired as the result of trading against its account, incurring the risk that the price of the stock will decline during the interim. These inventory risks are especially acute when orders are placed by investors with private information or when the stock's price is volatile.

There are also costs incurred in maintaining a continuous presence in a market, such as committing capital and monitoring trades.[\[112\]](#) And, of course, a dealer will need to make a competitive return on its expenditures, otherwise it will not be worthwhile to stay in business.

If small investors desire the best possible price, and not the immediacy that dealer intermediation provides, they are being forced to pay for a service they do not want as a result of the inability of auction markets, as a general matter, to offer side payments. Auction markets, even if they provide the best execution for small investors' orders, are penalized for their inability to compete with their dealer competitors. The increasing portion of small trades in exchange-listed securities diverted to dealers, and their execution at inferior prices, underscores the seriousness of the problem.

2. Distortion in Where a Given Order Is Sent, Even Among Securities Markets Capable of Offering Side Payments

Building on the condominium analogy—where an owner relies on an agent to find the buyer willing to pay the largest amount over the condominium's listed price—suppose that all potential buyers can offer bribes to the agent. Further, suppose that the agent and the potential buyers will have numerous interactions, over a long period of time, concerning the terms of sale of other condominiums whose owners the agent also represents. The question is: Should the condominium owner be concerned if her agent has promised a particular buyer that all the condominiums the agent handles are always sold to him at the listed price? Now add the fact that the agent receives from the buyer a cash "gratuity" for the promise as well as an annual vacation at the buyer's beachhouse.

Most owners would be concerned. Somehow the explanation that she will enjoy a lower commission rate as a result of these bribes would fail to provide much comfort. Yet the same argument has often been made in the securities market context. Investors, some say, are the ultimate beneficiaries of any side payments brokers receive given the competitive nature of the securities industry. Securities markets will compete away any profits they make in executing small investors' orders by offering larger and larger side payments to brokers, while brokers will compete with each other by offering lower and lower commission rates.[\[113\]](#)

Despite the competitive nature of the securities industry, order-flow arrangements will result in inefficiencies. Inefficiencies can occur even in situations where all securities markets can offer side payments, if the size of the side payments cannot be adjusted as quickly or as precisely as securities quotations. Under these conditions, the size of a securities market's side payment would be a less accurate indicator of the optimal market than the prices securities markets would quote if they did not have to compete to offer the largest side payments. As a result, any given order is less likely to be routed to the appropriate market than it would be in the absence of side-payment competition.

In terms of the condominium analogy, the distortion created by the arrangement between the real estate agent and the buyer is that, for any given condominium, another buyer might value the condominium more than the buyer providing the gratuity and vacation. If part of

the arrangement is that all condominiums are automatically sold to the buyer providing perks to the broker, regardless of the offers of others for any given condominium, there is a distortion in the selection of the buyer. Condominiums will not necessarily be sold to the buyers who value them the most. The ability of higher-valuing buyers to offer larger bribes does little good if the agent is already committed to selling the condominium to someone else. Moving one step back to analyze the situation at the point before the agent has entered into one of these arrangements, there is still a distortion if it is easier to locate the buyer willing to pay the highest price in a situation where buyers compete on price rather than on providing the largest gratuity or the most desirable vacation. It seems likely that it would be more difficult and time-consuming for an agent to compare the relative value of various types of bribes, and for buyers to precisely adjust the value of their bribes to reflect how much they are willing to pay for any particular condominium, than to compare and adjust bids.

Order-flow arrangements often have the same troubling characteristics. It is not uncommon for brokers to enter into various types of long-term arrangements with particular dealers, such as "joint venture" agreements,[\[114\]](#) making it more difficult for an outside dealer to successfully compete for the order flow of the broker. More importantly, it is likely to be far easier for a securities market to indicate quickly and accurately the value it places on a particular security at any point in time by adjusting its quotations rather than altering, for instance, the amount of free investment research services it provides to a particular broker over the course of a year. Indeed, nonprice competition is likely to be of far more concern in the securities context than in the condominium example as valuations placed on securities change at a much quicker pace than condominiums. Moreover, while it is straightforward for a broker to compare quotations, it is likely to be far more difficult to assess the relative value of various types of side payments.

This distortion suggests that the provision of side payments by securities markets should be of concern not only in the context of exchange-listed securities traded in the third market, which has been the focal point of the debate over order-flow payments, but also for securities traded exclusively over-the-counter, like most NASDAQ securities. Worryingly, payments for order flow in NASDAQ securities are commonplace and tend to be larger than payments offered for order flow in exchange-listed securities.[\[115\]](#) Interestingly, proprietary trading systems, where no order-flow payments typically occur, have been especially successful in attracting order flow in NASDAQ securities from sophisticated, institutional investors.[\[116\]](#) Any proposed solution to the distortions caused by order-flow payments should address this component of the problem.

3. Distortion in Whether Orders Will Be Placed

Even if all securities markets could instantaneously and costlessly translate their ability to offer price improvement into side payments, there is still a third distortion caused by these side payments. The more side payments a broker receives, the lower its commission rate is likely to be. The full true cost of trading, as a result, will not be fully reflected in the broker's commission rate.

This is problematic if small investors believe that brokers buy or sell securities on their behalf at market value, while commission rates represent the cost of making that adjustment in their holdings. The fierce competition between brokers on the basis of commission rates suggests that this is so. This belief is mistaken because it ignores the fact that commission rates will not reflect a portion of trading costs, such as foregone auction market price improvement opportunities. If trading costs appear lower than they actually are, socially excessive trading is the likely result. Those who argue that order-flow payments are unobjectionable because they reduce commission rates ignore the distortion created by misleading price signals.[\[117\]](#)

B. Distributional Consequences

Nonmonitoring small investors, along with auction markets, bear the brunt of these inefficiencies. The data contained in one study, summarized above in Part III, suggests that investors who place orders in NYSE-listed stocks for less than 400 shares are the ones who tend to have their trades misrouted.[\[118\]](#) It is interesting to note that the average individual investor's order is for 300 shares.[\[119\]](#) Order-flow payments occur not only for orders in exchange-listed securities but for over-the-counter securities as well. Trading in over-the-counter securities is driven more by individual investors and less by institutional ones as compared to exchange-listed securities. Unfortunately, estimating the cumulative impact of these inefficiencies on small investors is highly speculative. Not only would one have to know the foregone price improvement opportunities, but also the losses that occur due to inefficient nonprice competition[\[120\]](#) and excessive trading.[\[121\]](#) Whatever the true losses, there is ample cause for concern given the popularity of order-flow payments.

V. The Current Regulatory Regime and Its Shortcomings

There are different types of regulatory responses to markets characterized by significant numbers of customers who do not monitor the quality of a product. One approach attempts to directly attack the underlying problem by reducing the cost to customers of acquiring and acting on information concerning quality. Subsidizing the production of information or mandating its public dissemination are two common ways of accomplishing this approach. This can increase efficiency because private production and use of information can result in sub-optimal amounts of information being available to most investors. Informed investors confer a positive externality on uninformed investors who are able to free ride off their monitoring activities.[\[122\]](#) Another popular regulatory response involves mandating minimum standards of quality, at least where the parties have not reached an agreement to the contrary.

The current regulatory regime employs both approaches in attempting to ensure that orders are handled with consideration for investors' interests in mind despite the lack of monitoring by small investors. There are requirements that brokers provide information detailing their receipt of order-flow payments as well as minimum standards of brokerage quality. Three main components constitute the current regulatory regime: the Intermarket Trading System's (ITS) trade-through prohibitions (basically requiring that securities markets offer prices at least as favorable as the NBBO for exchange-listed securities);

brokerage "best execution" obligations; and disclosure requirements. These requirements, and their shortcomings, will be examined in turn.

A. The ITS Trade-Through Prohibitions

In order to understand the ITS trade-through prohibitions, and why they are inadequate to the task, their role in the SEC's "national market system" must first be understood.

1. *The National Market System*

The SEC, in its letter transmitting the *Institutional Investor Report* to Congress in 1971,^[123] and Congress, in the 1975 amendments to the Securities Exchange Act of 1934,^[124] called for the establishment of a "national market system" (NMS) for securities. While never a particularly clear concept, a national market system would basically "hold together in some fashion what might otherwise be an unduly fragmented market in multiply-traded common stocks."^[125] The national market for securities would, it was hoped, facilitate:

- (i) economically efficient executions;
- (ii) fair competition among brokers and dealers, among exchange markets, and between exchange markets and markets other than exchange markets;
- (iii) public availability of quotation and transaction information;
- (iv) an opportunity to obtain best execution; and
- (v) an opportunity to obtain execution without dealer intervention to the extent consistent with economically efficient executions and the opportunity to obtain best execution.^[126]

All these goals are jeopardized by the provision of order-flow payments and order-flow internalization. Facilitating price competition between securities markets (iii) and the opportunity to obtain best execution (iv) are, perhaps, the aims most obviously implicated by a broker's incentive to maximize its side payments at the expense of finding the best possible price for an investor's order. This incentive, as well as the resulting distortion on an investor's decision about whether to place an order in the first place, also damages the goal of ensuring economically efficient executions (i). Finally, the systematic penalization of auction markets created by this form of nonprice competition can hardly be said to encourage fair competition between securities markets (ii) or an opportunity to obtain execution without dealer intervention (v). The question is: What regulatory safeguards does the NMS structure erected by the SEC provide against these dangers?

The core components of the national market system, as it has evolved in the last twenty-five years, are three electronic communications systems connecting various securities markets in different ways: the Consolidated Tape, the CQS, and the ITS. The Consolidated Tape disseminates securities transaction information within ninety seconds, whether the trades

occur on an exchange or over-the-counter, for practically all exchange-listed securities.[\[127\]](#) Supplementing the Consolidated Tape are several reporting systems operated by NASDAQ which also report, in real time, last sales information for almost all the over-the-counter securities, i.e., securities not listed on any exchange.[\[128\]](#)

The CQS, on the other hand, publicly disseminates the best bid and offer (NBBO), derived from the quotations exchanges and over-the-counter dealers provide, for essentially all exchange-listed securities.[\[129\]](#) Exchanges are required to furnish the quotations and sizes—the number of shares that a quotation is good for—at which member brokers and dealers indicate that they are willing to trade at.[\[130\]](#) The NASD, just as the exchanges, reports the highest bid and lowest offer, along with their associated quotation sizes, offered by over-the-counter dealers for these securities.[\[131\]](#) A broker or dealer who indicates it is willing to trade a security at a given price is normally obligated to execute orders at that price up to the reported quotation size.[\[132\]](#) Supplementing the CQS, the NQDS publicly disseminates quotations, including the highest bid and lowest offer, for NASDAQ securities, i.e., NASDAQ/National Market System securities and NASDAQ Small-Cap securities.[\[133\]](#) And just as with their CQS quotations, over-the-counter dealers must stand by these quotations.[\[134\]](#)

The ITS is a market-to-market routing system which enables orders to be transferred for execution from one securities market, including all the exchanges, to another based on who is offering the best price as reported on the CQS. The ITS links over-the-counter dealers with the exchanges through the ITS interface with the NASD's Computer Assisted Execution System (CAES). The CAES, in turn, is an automated execution system that enables over-the-counter dealers to transfer orders in exchange-listed securities between each other for execution. All third market dealers are required to be members of the CAES and, hence, are also linked to the exchanges.[\[135\]](#) The ITS/CAES interface, though, is limited to "Rule 19c.3" securities which means that an order cannot be transferred for execution between an over-the-counter dealer and an exchange through the ITS if it is for a stock listed on the NYSE before April 26, 1979.[\[136\]](#) However, third-market dealers who wish to have the same access as exchange members can themselves become a member of an exchange. For instance, Madoff Investment Securities, by far the largest third-market dealer, is a member of the CSE. Among the exchanges, orders can be transferred for execution between each other through the ITS for the vast majority of exchange-listed securities, whether they are 19c.3 securities or not.[\[137\]](#)

Under ITS rules, a securities market offering a price inferior to the NBBO for a security traded on the ITS must either match that price for any orders it receives or send a "commitment to trade" to the market posting the best price. If the commitment is accepted, then the order is executed on that market.[\[138\]](#) Each ITS market, in compliance with ITS requirements, has adopted "trade-through" rules which prohibit securities markets from executing an order at a price worse than the price quoted on the CQS by an ITS participant market—over-the-counter dealers with respect to 19c.3 stocks and the exchanges—without first attempting to have it executed by the market offering the superior price.[\[139\]](#) Nevertheless, trade-throughs still occur, although infrequently.[\[140\]](#)

There are some structural limitations to the ITS which hamper its effectiveness. Most importantly, the NBBO reported on the CQS is not necessarily the best ITS price for purposes of the ITS trade-through prohibitions. The fundamental question, however, putting aside any limitations which could be remedied by an extension of the ITS network, is whether a guarantee of the NBBO is sufficient protection for investors' orders.

2. *Why a Guarantee of the NBBO Is Insufficient*

Many commentators initially thought that the combination of the CQS and the ITS, with its routing system and trade-through prohibitions, would result in securities markets having both the ability and the obligation to always get the NBBO for investors' orders in listed securities. Security markets, in turn, would compete for the order flow that brokers control by revealing on the CQS the highest bid and lowest offer for a security that they could possibly provide.

This has not happened. The failure of the national market system to facilitate quote competition is reflected in the SEC's finding that the implementation of the ITS system had no "statistical effect on the quality [as measured by spreads and volatility] of the primary market" prices.^[141] Spreads between ITS and non-ITS stocks do not differ in any systematic fashion. Exchanges competing for NYSE-listed securities execute anywhere from 34.7% (CSE) to an incredible 92.1% (PHLX) of a security's dollar trading volume when neither its posted bid nor its offer is the NBBO.^[142]

There are a number of reasons why security markets have been generally unwilling, or even unable, to compete for order flow through posting competitive quotations on public forums such as the CQS. Indeed, some of these reasons also explain why over-the-counter dealers have refused to post the best prices they are capable of providing for public dissemination by the NQDS for NASDAQ securities.

a. *Crossing Market Orders*

One of the main reasons why the NBBO may not reflect the actual price available on the NYSE, as well as other auction markets, is the ability of auction markets to directly cross market orders at prices within the NBBO spread.^[143] In a study of all transactions on the NYSE in September of 1993, 28% of all trades in NYSE stocks occurred between the NYSE's posted bid and offer. When the NBBO was greater than the minimum tick of 1/8, 66% of all NYSE trades occurred within the spread.^[144] Whether a market order will be crossed is very difficult to predict as this will depend on the other orders that happen to be routed to the exchange floor at a particular point in time.

b. *Monitoring Costs*

Another reason is due to the risk a dealer or a specialist runs of being "picked off" if it posts a price on the CQS or NQDS which it is then obligated to stand by. If market conditions change, and the trader is unable to change its quotation quickly enough, the trader might very well have to buy (or sell) stocks for more (or less) than they are worth. The ITS routing system enables other market traders to take advantage of outdated, or even ill-

advised, quotations as they have the choice whether to match the NBBO or transfer the order to the market offering the best price. If the current NBBO is off-base for some reason, the trader who posted the quotation will attract unwanted order flow from those taking advantage of the error. On the other hand, if the NBBO accurately reflects market conditions then competing markets will likely choose to match the NBBO for a particular order and retain the business for itself. In short, there is a "winner's curse" associated with competing for order flow through posting quotations on the CQS and NQDS.

This is more than a theoretical possibility. The NASD conducted a study of its SOES system which automatically routes small orders for execution to the NASDAQ dealer offering the best price. The study revealed that twenty-nine firms inundate the system with heavy trading at points when these firms thought the best posted price was too generous. Not surprisingly, NASDAQ dealers, in order to protect themselves, widened their spreads.^[145] The CSE, which automatically executes orders at the NBBO, has experienced similar problems.^[146]

Underlying these problems is the basic inability of a reporting system, which relies on simple combinations of quotations and sizes, to capture all the trading interest that exists at any given point in time.^[147] Traders who are willing to trade under a specific set of circumstances but not others will often be unable satisfactorily to qualify their quotations on the CQS or NQDS. Even if a reporting system could qualify quotations in this potentially infinite number of ways, it would be prohibitively expensive for a trader always to specify its trading interest in such detail. The very existence of brokers, who must exercise their judgment as to how an investor's order should be handled, underscores the costs to a trader of providing exact specifications of how its trading preferences vary under different conditions. Indeed, the impressive size of the "upstairs" market can be attributed in large part to the high cost of expressing trading interest on an exchange floor, despite exchanges being a far more flexible forum for expressing trading interest than the CQS.^[148]

c. Informationally Motivated Order Flow

Ill-advised or outdated quotations are not the only way traders can be picked off. There is the chance that the order flow attracted by a superior price on the CQS or NQDS is from investors with private information about a security's likely future value. Other traders have an incentive not to match the NBBO for orders they think are information-based, sending them instead to the trader posting the NBBO for execution. Market participants, whether they are floor brokers, specialists, or over-the-counter dealers have various ways of sorting out whether orders are informationally motivated or not and adjusting their prices accordingly.^[149] Making these kinds of judgments about order flow being transferred over the ITS or SOES is much more difficult. In any event, it often will be too late by this point because the securities market will have to stand by their quotations.

d. Raising Competitors' Costs

Some have also argued, somewhat more controversially, that the NYSE does not always post its best prices on the CQS as a way of raising its competitors' costs. In order for the

NYSE's competitors to attract order flow in NYSE-listed stocks from sophisticated traders, the argument goes, these traders will have to be assured that they are receiving prices at least as good as those available in the primary market. If the NYSE is able to conceal its true price, competitors will have to incur the cost of ascertaining the NYSE's actual prices. Regional specialists have reported that they are forced to incur the cost of submitting their own orders to the NYSE to uncover the NYSE's real prices, which are not available on the CQS.^[150] A study of NYSE specialists' practices revealed that they fail to display on the CQS approximately one-half of all limit orders that are superior to the best CQS quotation.^[151]

B. Duty of "Best Execution"

The SEC has explained that a broker has "a duty to seek to obtain best execution for customer orders, which is understood to mean that a broker-dealer must obtain the most favorable terms available under the circumstances for a customer's transaction."^[152] Of course, everything hinges on what exactly a broker can and cannot do "under the circumstances." The only guidance the SEC has provided, beyond this statement, is a non-exhaustive list of factors with uncertain weight.^[153] Other formulations of the duty of "best execution," whether given by the courts,^[154] exchanges,^[155] the NASD,^[156] the common law of agency,^[157] or commentators^[158] are almost as amorphous. This is not surprising, given the complicated nature of choosing the appropriate market.^[159] This imprecision in the definition of the duty is reflected by the fact that violation of the duty has few legal consequences outside a rather circumscribed set of circumstances, despite it being a "cornerstone of market integrity."^[160] Outside these well-defined circumstances, the SEC, the NASD, and the courts, at least until recently, have been loathe to second-guess how a broker weighed all the relevant, and often competing, factors that go into a judgment as to what constitutes "best execution" for an order.

The well-recognized situations where a broker's handling of an order violates its duty of "best execution" fall into five general categories: (1) failing to obtain even the NBBO for an order (an occurrence which is already partially protected against through the ITS trade-through prohibitions); (2) failure to consummate transactions in a reasonable period of time;^[161] (3) engaging in unauthorized transactions;^[162] (4) unnecessarily using a second party, usually a second broker, in executing an order;^[163] (5) failing to disclose potential conflicts of interest, such as whether the broker was also acting as principal in a trade.^[164] Beyond these specific circumstances, the SEC has limited itself to periodically issuing largely hortatory statements explaining that a broker "must at least make periodic assessments of the quality of competing markets to assure that it is taking all reasonable steps under the circumstances to seek out best execution of customers' orders."^[165] Attempts to further regulate brokerage handling of orders by invoking state law, mainly state common-law agency principles, have, to date, failed.^[166] In other words, where an investor's order is routed is basically up to the broker.

The Third Circuit's recent unanimous en banc opinion in *Newton v. Merrill Lynch*,^[167] could be read as a substantial extension of the duty of "best execution" beyond these traditional categories. In this case, a class-action suit was brought by an investor contending that various broker-dealers, who over a two-year period automatically executed investors'

orders at the NBBO, had violated their obligation to provide "best execution" since superior prices were available on other markets. While admitting that the broker-dealers had handled the orders in a way that had "never been held to be fraudulent by any court or regulator," the court concluded that the district court's grant of summary judgment in favor of the broker-dealers was inappropriate, given that a "reasonable trier of fact could conclude that the defendants misrepresented that they would execute the plaintiffs' orders so as to maximize the plaintiffs' economic benefit."[\[168\]](#) The court's reasoning casts a seemingly long shadow over any contestable decision made by a broker in handling an investor's order. Since anything that a trier of fact might consider a broker's failure to "maximize [a] plaintiff[s] economic benefit" would potentially violate the duty of best execution.

The impact of *Newton*, however, is unlikely to be as sweeping as some of its reasoning might suggest for several reasons. First, it is unclear whether future courts will read *Newton* so broadly. Conceivably, courts will limit *Newton* to situations where a broker-dealer consistently automatically executes orders at the NBBO over a substantial period of time, even when better prices were often available with a minimum of effort. There would be powerful reasons for such a cautious approach. Allowing best-execution claims to proceed in more complicated situations could prove unworkable. One such situation would be if a broker routinely executes orders executed at the NBBO after having attempted to achieve a better price in at least some limited manner. The ability of juries to make consistent and reasonably accurate judgments about such matters as whether checking another securities market for quotations "would have added substantial expense and delay to the execution of [investors'] orders"[\[169\]](#) is dubious. Indeed, this is one of the reasons why the Ruder Committee, a committee created by the NASD to look into payment for order flow issues, suggested that the duty of best execution should presumptively be considered satisfied when orders are cleared at the NBBO.[\[170\]](#)

Second, central to the court's conclusion that there was a material issue of fact as to whether there was a violation of best execution was its belief that "a trier of fact could infer that the defendants' acceptance of the orders was reasonably understood as a *representation* that they would not be executed at the NBBO price when better prices were reasonably available elsewhere."[\[171\]](#) Of course, if the broker had made a different representation before acceptance of an order, there could be no misrepresentation. To avoid liability, brokers would just have to inform investors beforehand of contestable practices. In fact, for the most contestable practice of brokers, acceptance of order-flow payments, there are mandated disclosure requirements, some of which are triggered before any transactions occur.[\[172\]](#)

C. Disclosure of Payments for Order Flow

In response to the concern that investors are often not receiving the best possible price for orders diverted by brokers to dealers, a fear not laid to rest by the ITS trade-through prohibitions and the duty of best execution, the SEC adopted rules requiring the disclosure of various items of information by brokers who receive order-flow payments for all securities whether they be exchange-listed or over-the-counter.[\[173\]](#)

1. *The Disclosure Requirements*

The payment for order flow disclosure rules require brokers to provide information on three occasions: opening of a new account, annually, and trade confirmation. When an investor opens a new account, the broker must disclose in writing whether it receives order-flow payments, a detailed description of the nature of the compensation received, and the policies for determining where to route investors' orders that are the subject of payment for order flow. The disclosure also includes a description of the extent to which orders can be executed at prices superior to the NBBO. This information must be updated and distributed to all the broker's customers on an annual basis.[\[174\]](#) Finally, in each trade confirmation, a broker who accepts payments for order flow must include a statement to that effect.[\[175\]](#) Payments for order flow are defined as "any monetary payment, service, property, or other benefit that results in remuneration, compensation, or consideration to a broker or dealer from . . . [any broker or dealer, or exchange] in return for the routing of customer orders."[\[176\]](#)

In addition to these disclosures, the SEC has recently adopted regulations requiring a broker to report, on a quarterly basis, the identity of the ten trading venues to which the broker has routed the largest number of orders as well as a description of the broker's relationship with the trading venues, including payment for order flow and profit-sharing arrangements.[\[177\]](#) Moreover, investors can receive, upon request, detailed information concerning the routing of their own orders. It is hoped that this information will enable investors to make more informed choices in selecting a broker.

2. The Effectiveness of the Disclosure Requirements

The disclosure requirements placed on brokers are largely off-target. Even detailed disclosure of payments for order flow and a broker's treatment of purchased order flow would provide an incomplete picture. An investor should not necessarily be concerned with brokerage receipt of side payments per se, but rather the extent to which the broker is forgoing price-improvement opportunities as a result. The disclosed information does not shed much light on the answer to this question.

It is important to realize that price-improvement opportunities can vary depending on such factors as the specific security being traded—more heavily traded securities may have a different probability of being crossed within the spread on the NYSE floor than those less traded—and the time of day trading is occurring. Indeed, an investor would need to know not only the foregone price-improvement opportunities, given the circumstances, but also the cost of capitalizing on these opportunities, such as the size of the search costs and exchange-service fees, in order to make a complete analysis.[\[178\]](#)

The SEC's approach has other serious shortcomings, which further undermine the disclosure requirements' effectiveness. First, the disclosure requirements do not adequately deal with internalization of order flow and, therefore, address only a portion of the problem. This is a problem generic to a regulatory approach relying on disclosure of order-flow payments. It would be extremely difficult for a regulator to measure the implicit order-flow payment given in exchange for internalized order flow. Second, order-flow payments are defined by regulation as a monetary inducement exceeding a securities market's service fee. An inducement is an inducement whether it takes the form of a service-fee reduction or

a cash rebate. Finally, it seems likely that at least a significant portion of small investors would fail to read, to understand, and to act upon the disclosed information when selecting a broker.[\[179\]](#)

VI. The Proposal and Possible Objections

A. The Proposal and Its Benefits

To summarize the discussion above, the conflict of interest that presently compromises brokers' selection of a securities market arises out of small investors' inability to monitor these decisions by simply comparing commission rates.[\[180\]](#) If small investors could somehow monitor brokers' choices of securities markets, even indirectly, the divergence of interests between brokers and investors would cease. Should those interests be harmonized, brokers who route orders to the appropriate securities market would be rewarded with small investors' business. And securities markets, in turn, would compete more by posting competitive quotations than on the basis of who can offer the largest side payments.

Reconciling brokers' and investors' interests is simple to achieve: allow (but not require) brokers to credit small investors' orders with whatever the NBBO happens to be at the time of execution regardless of the actual price received. Call this the NBBO pricing option.[\[181\]](#) "Small investors' orders" should be defined to include the orders of those investors least likely to be able to monitor anything other than commission rates. One could attempt roughly to identify this group by the size of the orders they place with their brokers, for example, orders whose sizes are 400 shares or less. Empirical evidence indicates that it is for this class of investors that the agency problem is most significant. Investors whose orders exceed this amount, by and large, do not appear to have them routinely misrouted.[\[182\]](#)

A broker that adopts the NBBO pricing option will have a powerful incentive to route small orders to the securities market offering the best possible price, perhaps a better price than the NBBO, since a failure to do so would come only at its expense. Of course, if the market offering the best possible price for a security also charged high exchange-service fees, or if the benefits of the best price were otherwise economically outweighed, the broker might decide that sending the order to another securities market would be preferable. The same outcome might also occur if a broker determined that finding the market offering the best price was too expensive. In any event, brokers employing the NBBO pricing option would have the proper incentive to make the most efficient decision, since they will bear the full consequences, both good and bad, that the selection of a particular securities market would entail.

Brokers who utilize the NBBO pricing option and route orders to securities markets offering the best prices would be able, as a result of the option, to offer lower commission rates than they otherwise could. Brokers who are not as effective in selecting the appropriate securities market, perhaps by accepting side payments for routing orders smaller than the size of the price improvement available on another securities market, would have to charge higher commission rates to compensate for this misallocation. Given the highly competitive nature of the brokerage industry, and the emphasis on low

commission rates, this is likely to prove costly.[\[183\]](#) The fact that discount brokers have been using their sizable order-flow payments to help reduce some of their commission rates is telling.[\[184\]](#)

With the resolution of brokers' conflict of interest, due to the availability of the NBBO pricing option, securities markets will no longer be rewarded by brokers for the mere fact that they offered side payments in lieu of better prices. Indeed, to the extent that prices can adjust more quickly and at less cost than side payments—the source of one of the inefficiencies created by the current incentive structure—markets will have a new reason to compete on the basis of price. Posting quotations is likely to be, as a general matter, the fastest and cheapest method for a dealer to signal the exact times at which it is able to offer a price better than those offered by its auction-market competitors.[\[185\]](#) Increased competition based on who can offer the best prices would also have the beneficial effect of removing the current penalty that auction markets labor under as a result of their inability to offer unproductive side payments. To the extent that order-flow payments, or any other broker-securities market arrangement, serve legitimate purposes—such as helping dealers realize economies of scale, rather than exploiting a broker's conflict of interest—they would continue to be offered.[\[186\]](#)

If brokers using the NBBO pricing option would properly handle investors' orders so as to offer the lowest possible commission rates to small investors, this solution would appear only to aggravate the third inefficiency resulting from order-flow payments, namely that investors will not understand the full costs of trading by merely looking at brokers' commission rates. Would commission rates then become even less reflective of trading costs? The answer to this question will depend on the effect the proposal would have on the NBBO. If securities markets compete more on price and less on side payments, the NBBO would become more reflective of the true market value of securities than it currently is. Commission rates would tend to rise as diverting orders to dealers and clearing them at the NBBO became less and less profitable. At the extreme, if the NBBO were perfectly reflective of market value, commission rates would be a very reliable indicator of the full cost of trading.[\[187\]](#) Ironically, brokers' efforts to minimize the commissions they charge, under the NBBO pricing option, would provide additional incentives for securities markets to compete on the basis of price rather than side payments.

It seems likely that there will be some number of investors who place small orders (whether the cut-off is 400 shares or some other amount) who will have both the ability and the desire to monitor their brokers directly, instead of relying on a broker who uses the NBBO pricing option. Perhaps such investors have specific execution needs or are sophisticated traders who sometimes trade in small increments. For whatever reason, if an investor decided that it was worthwhile to incur monitoring costs despite the availability of merely using brokers utilizing the NBBO pricing option, then she should be able to do so.

Brokers who wish to attract the business of these types of investors could elect not to handle orders pursuant to the NBBO pricing option. In order to effectuate the NBBO-pricing-option proposal, with its element of investor choice, brokers should have to commit to use, or not to use, the option of clearing small orders at the NBBO for a substantial period of time. That way, it will be clear to investors who monitor directly, and not simply

through commission rates, how their orders will be handled. The SEC can play a useful role in helping monitoring investors wisely select a broker who does not use the NBBO pricing option. The SEC has recently started down this path by requiring securities markets, although not brokers, to release statistics measuring their price-improvement performance.[\[188\]](#)

Although it is worthwhile, disclosure alone will not resolve the agency problem. Most brokers have traditionally charged a flat rate for investors' orders regardless of whether they are limit or market orders, despite the fact that limit orders are more expensive to handle.[\[189\]](#) This is still common, even though new order-execution requirements make handling limit orders even more expensive relative to market orders. Many investors apparently prefer considering a single commission rate even at the cost of paying higher prices than they would otherwise have to pay if there were separate commission rates for market and limit orders. Even the most user-friendly disclosure statements are more difficult to understand than comparing commission schedules.

Small investors who cannot or do not care to monitor brokers' choices of securities markets can safely select brokers based on their commission rates, much like they are probably doing already. Brokers who utilize the NBBO pricing option, and therefore who have no incentive to mishandle orders, will tend to have lower commission rates, to the extent that the NBBO is not fully reflective of a security's true market value, than brokers who do not. Any advantages enjoyed by brokers as a result of using the NBBO pricing option will likely accrue to the ultimate benefit of their investors.[\[190\]](#)

The NBBO pricing option, and its disclosure regime, would be an inexpensive solution to the agency problem. The requirement that a broker who uses the NBBO pricing option actually provide the NBBO at the time of order execution is easily verifiable and not very burdensome. Presumably, most brokers already keep price-improvement statistics for their own self-evaluations.[\[191\]](#) The ITS trade-through prohibitions, as well as the application of the duty of best execution to brokerage routing decisions, would no longer be necessary. Rather, investor protection would be provided through competition. Securities markets would compete with each other in providing the most attractive execution for the orders investors place with brokers.[\[192\]](#)

B. The Problem of Intermarket Fragmentation

Conflict-of-interest issues aside, some are still likely to be concerned over the fragmentation of order flow across several different securities markets.[\[193\]](#) One view is that all order flow in a security should be concentrated in one trading venue.[\[194\]](#) The proposal would only partially address these concerns over order-flow fragmentation. It would reduce the problem (assuming it is a problem), given that its likely effect is to allow the NYSE to capture more of the order flow in its listed stocks given the greater price-improvement opportunities that are usually present there. Fragmentation of order flow in NYSE-listed stocks has been especially significant in the market for small orders. Nevertheless, significant inter-market fragmentation of order flow would undoubtedly still exist after the proposal's implementation.

In considering this concern, it should be pointed out that competition between markets for order flow was anticipated and approved by Congress when it was amending the securities laws in 1975.[\[195\]](#) Securities markets competing for order flow inevitably means intermarket order-flow fragmentation. Congress' policy has merit. Concentrating order flow in one market center would tend to reduce competitive pressures to innovate. Intermarket competition for order flow in NYSE-listed securities has had a beneficial effect in forcing the NYSE to adopt new technologies and trading practices. Competing markets have also served investors by catering to the needs of particular types of traders, needs that would not necessarily have been satisfied if there were only one market. Moreover, there is a real limit to the extent to which the SEC, putting aside the question of Congressional intent and regulatory authority, can force order flow into one locale. Investors will still have the option of moving their trades overseas or purchasing stock derivatives through which they can readily replicate stock positions.

More to the point, the debate over the appropriate amount of order flow consolidation is really a much more general debate than the one about the agency problem created by certain forms of nonprice dealer competition for order flow. One can be worried about the agency problem these forms of dealer competition create without having to commit oneself to a particular position on the more general issue. Only specific claims about how nonprice dealer competition for order flow aggravates problems associated with intermarket fragmentation need to be considered in evaluating a proposed solution to the agency problem.[\[196\]](#)

There has been one such concern that has often been expressed in the debate over diversion of order flow to dealers. If payment for order flow and internalization are mechanisms for ensuring that dealers trade only against investors without any private information about the likely future value of securities (uninformed investors), hence capturing the most profitable portion of exchange-listed order flow ("cream skimming"), a proposal which did not address this segregation would provide a basis for criticism. Dealer cream skimming is problematic if auction markets, like the NYSE, cannot also segregate informed and uninformed order flow much as they cannot provide order-flow payments or completely internalize order flow. The result of an auction market's inability to segregate would be its having to charge a larger spread than a dealer in order to account for the increased probability of trading against an informed investor.[\[197\]](#) The loss of uninformed order flow would force the NYSE, among others, to increase further their spreads, making cream skimming even more attractive. Diversion of uninformed order flow would beget further segregation in a reinforcing cycle resulting in the continued deterioration in the quality of auction-market spreads.

While this scenario is theoretically possible, it is unclear whether it is actually occurring for several reasons. First, it is difficult to determine whether small orders are being diverted to dealers due to cream skimming or because they are placed by small investors who do not monitor execution quality. This is because nonmonitoring investors also tend to be uninformed investors. The empirical evidence is mixed as to whether the NYSE's competitors are competing primarily by attracting a "safer" uninformed order flow. Studies have shown that price discovery generally takes place on the NYSE, which suggests that informed order flow is primarily sent there for execution.[\[198\]](#) More specifically, one study

has found that the informational content of orders cleared on the CSE is significantly lower than orders on the NYSE.^[199] On the other hand, as Professor Coffee has pointed out, some empirical research indicates that informed trades are concentrated in medium-sized orders. This is interesting because orders of this size are an area where the regional exchanges have been particularly successful in attracting order flow away from the NYSE through price competition.^[200] Another study found that when Madoff Investment Securities, by far the largest third-market dealer,^[201] begins to offer payments for order flow in a particular stock, the stock's spread actually decreases.^[202] This is the opposite of what one would expect according to the cream-skimming hypothesis.

The contention that dealers are primarily competing by cream skimming is undercut by the fact that third-market trading is concentrated in the 400 most actively traded NYSE- and AMEX-listed securities. It is important to realize that an increase in the probability of informed trading does not have the same importance to a dealer in actively traded stocks as it does for a dealer in less frequently traded stocks. In a heavily traded stock, a dealer is able to offset any position, whether acquired from informed or uninformed traders, by quickly trading against the constant incoming stream of buy and sell orders. For less frequently traded stocks, a dealer will have to charge a higher spread given the increased length of time it must hold a position acquired from a potentially informed trader. This increases the likelihood of the market incorporating the previously private information during the period the dealer holds the stock resulting in dealer losses.^[203] Empirical studies have documented that losses from informed traders are significantly more important in determining dealer spreads for inactively traded stocks as compared to stocks with higher trading volume.^[204] Accordingly, cream-skimming, ensuring that one trades only against uninformed investors, is likely to be a much more profitable strategy in the least frequently traded stocks.

Even assuming that dealers are currently competing by cream skimming, it is questionable whether exchange specialists and floor brokers cannot similarly offer better prices for orders they believe are uninformed once they are sent to the floor for execution. Some commentators have suggested that specialists also have the ability to discriminate in their pricing, much as over-the-counter dealers, between informed and uninformed orders.^[205] If this is true, then under the regime established by the proposal, a broker will not have an incentive to send her uninformed order flow to dealers as opposed to an auction market, such as the NYSE, which can similarly offer prices reflecting an order's likely informational content.

The bottom line is that it is presently very difficult to tell whether (a) dealers are currently competing through cream skimming and rebating to brokers part of the profits from this strategy, or (b) auction-market participants, such as specialists and floor brokers, cannot also adjust their prices once an order is sent to an exchange based on its likely informational content. If either of these is not the case, then the proposal would probably adequately deal with a dealer's incentive to trade against only uninformed traders. The present proposal at least partially addresses cream skimming concerns by providing a greater incentive than currently exists to direct small orders, which are often placed by nonmonitoring investors to auction markets.

C. Other Possible Objections to the Proposal

There are at least three likely criticisms of the proposal as a solution to brokers' conflict of interest in selecting a securities market: (1) the proposal presupposes that price is the only consideration; (2) brokers will be forced to shoulder unwanted price risk; and (3) the proposal will actually render the NBBO less reflective of true market value than it already is.

1. *Price Is Not Everything*

Price is not necessarily everything.^[206] In particular, while dealer markets may not always offer the same price improvement opportunities as auction markets, dealers provide a service that auctions do not—immediate execution of orders. This enables investors quickly to establish a position at the prevailing market price before it changes. If small investors want immediacy, even at the cost of price-improvement opportunities, then the proposed regime will not necessarily be superior. Orders that are currently being filled by dealers might be sent to the NYSE as a result of this proposal, even where investors would prefer to lock-in the prevailing market price.

Several points can be made in response to this concern. The vast majority of small investors probably care most about getting the best possible price.^[207] The reasons that often lead investors to be willing to pay for immediacy—arbitrage opportunities, intermarket hedging, and short-lived private information concerning a security's future value^[208]—are typically important only for sophisticated investors who can already ensure that they receive the brokerage services that they need. Moreover, the relative speed of execution of small orders routed to dealer and auction markets is often not significant. For example, the SEC has found that the NYSE's execution for small orders is only eleven to thirteen seconds slower than small order execution on NASDAQ.^[209] Some dealers' execution of orders are also not immediate. Madoff Investment Securities, for example, exposes market orders to other markets for thirty seconds, in an attempt to obtain a better price, before executing the order itself.^[210] Small investors for whom the delay associated with auction market executions is too long always have the option of submitting marketable limit orders. Marketable limit orders are orders whose limit price is the current prevailing market price. By placing such orders, these investors can ensure that they receive the current market price while still enabling the average small market order investor to receive the best possible price for her order.

2. *Inefficient Transfer of Price Risk*

Another possible objection is that the proposal forces brokers, instead of the security holder, to bear the risk that an order does not receive the expected price improvement. Brokers will have to set their commission rates in light of expected price improvement opportunities which could turn out to be inaccurate. If a broker is not the most efficient bearer of this risk, the proposal could result in sub-optimal shifting of risk.

This is unlikely to be a significant problem. It is improbable that brokers will have to bear much risk as a result of having to set its commission rates before knowing which orders

actually receive price improvement over the NBBO. A broker will only have to estimate the average price improvement for a very large number of orders. While it might be unknown whether a particular order will enjoy price improvement, this does not mean that there is substantial uncertainty as to what the average price improvement will be for hundreds of thousands of orders.[\[211\]](#)

Moreover, even if it turns out that brokers are being forced to bear unwanted price risk, there are market mechanisms which will enable them to shift the risk to another party. Indeed, the most obvious example is the typical payment for order flow agreement where a broker is guaranteed the NBBO plus a fixed payment per share.[\[212\]](#)

3. Rendering the NBBO Less Reflective of True Market Value

Given the importance of the NBBO under this proposal, some might be concerned that market participants will have an even stronger incentive not to express their actual trading interest on the CQS or the NQDS. Specifically, one might argue that a broker-dealer who is considering whether to improve upon the NBBO, will now have to consider the costs this will impose on its brokerage business as the size of the expected price improvement, which before was retained by the broker, will be reduced by an improved NBBO. If the broker-dealer, as a result of its desire to make the expected price improvement as large as possible, decides not to improve on the NBBO, the NBBO will become even less accurate than it already is. Investors will end up with less information about the true value of the securities they hold or wish to buy—a worrisome development.

Perhaps the easiest response to this possibility is to point out that a similar incentive already exists under the present regime. Broker-dealers who automatically execute order flow at the NBBO have an incentive not to improve upon the CQS or NQDS because they will also have to provide the improved NBBO to their customers. This incentive also currently exists for dealers who receive order flow from brokers pursuant to order-flow agreements which typically only require that orders are filled at the NBBO.[\[213\]](#)

Moreover, to the extent that nonvertically integrated dealers currently set the NBBO, the NBBO would remain as (in)accurate as it is now. Dealers who are not affiliated with any broker will be unconcerned with the effect improved NBBO's, and hence smaller expected price-improvement opportunities, will have on brokerage profits associated with handling small investors' market orders. In addition, whenever a dealer or specialist is handling a limit order, which improves upon its NBBO quotation or size, they have no choice, but to reflect the limit order in their posted NBBO price.[\[214\]](#)

VII. Competing Reform Proposals: A Critical Appraisal

A. Banning Payment for Order Flow

Banning order-flow payments, the most draconian of the three popular reform proposals, derives a great deal of support from common law principles of agency which generally require not only disclosure but also a principal's consent before an agent may enjoy any profits arising from the principal-agency relationship.[\[215\]](#) Given the difficulty of receiving

ex ante consent from each investor, this common law requirement could result in the effective prohibition of the practice.^[216] For similar reasons, some have argued that payments for order flow are bribes or kickbacks, which are illegal under most state commercial bribery statutes,^[217] and even RICO.^[218]

Although it might be easy to prohibit monetary payments, it would be much more difficult to stop nonmonetary payments; and it would be impossible to deal with the likely consequence of such a prohibition—increased internalization where no payments need be made. This is already happening under the much more lenient disclosure regime now in place. In response to the SEC's disclosure requirements, Charles Schwab is purchasing dealer subsidiaries apparently so it can state to its customers that it does not accept payments for order flow.^[219] Incentives to integrate vertically would be greatly augmented by a flat prohibition on side payments.

Regardless of whether the proposal is practical or not, it ignores the fact that order-flow payments can serve legitimate purposes. It encourages brokers to send order flow in bulk, enabling dealers to realize economies of scale in making a market for a security.^[220] Tellingly, order flow agreements almost always specify that the broker, in order to receive any compensation, must route to the dealer a large minimum number of orders, usually at least 100,000 orders per month.^[221] This practice, for obvious reasons, can also reduce brokerage search costs. Finally, side payments have also served as a way for the NYSE's competitors to get the attention of brokers who have traditionally sent their order flow to the NYSE.^[222] Under the proposal advocated in this Article, dealers could still offer side payments to brokers in order to realize these potential benefits.

B. Decimalization

Perhaps the most popular regulatory recommendation is "decimalization" which would enable markets to quote on the CQS spreads in increments of 1/100 of a dollar instead of the historical tick of 1/8.^[223] Supporters of this proposal analogize the minimum-tick requirement to a fixed-commission regime, i.e., requiring quotations to be stated in rigid increments is a form of price-fixing. A minimum-tick regime prevents more efficient markets from attracting order flow by offering superior prices within the tick. As was the case with the NYSE's fixed-commission regime, various ways of circumventing this artificial constraint have arisen. Payments for order flow is one such avenue. Suppose the spread, as determined by market forces, would be \$1/16 but a dealer is forced to quote in increments of \$1/8. A dealer could offer a spread of \$1/8, the minimum spread it is able to offer, as well as a payment to a broker for order flow of \$1/16. The willingness to make such payments is a direct result of the inability to merely post the spread the dealer would offer in the absence of the minimum-tick requirement. Abolish the tick rules, some believe, and the incentive to make payments for order flow will disappear.^[224]

Whether decimalization is an unmitigated benefit, as some of its more enthusiastic advocates suggest, is not so clear. A reduction in the minimum tick can, for a variety of reasons, reduce the liquidity and transparency of a securities market. Empirical studies of markets that have reduced their minimum ticks have documented that there is reason for concern.^[225] It is fair to say that whether decimalization of the U.S. securities markets will

have a net beneficial impact is a complicated and contested question. However, whether decimalization will have detrimental side effects is really besides the point. Decimalization, for better or worse, is being phased in.[\[226\]](#)

The relevant question is whether decimalization will eliminate the misalignment of investor and broker interests, obviating the need for further regulatory action. The answer to this question is probably no. While the minimum-tick rules undoubtedly contributed to the ineffectiveness of the CQS system in reflecting current trading interest, as the previous discussion has shown, it is certainly not the only factor.[\[227\]](#) Moreover, since dealer markets do not enforce time priority, as auction markets usually do, the size of the minimum tick is much less likely to change dealer practices, especially in NASDAQ securities which tend to be traded only over-the-counter.[\[228\]](#) Interestingly, reduction of minimum ticks in auction markets has not typically resulted in less internalization.[\[229\]](#)

C. Midwest Stock Exchange's Proposed Rebate Scheme

The Midwest Stock Exchange (now Chicago) in a petition for rulemaking[\[230\]](#) proposed that brokers be required to rebate to investors' brokerage accounts any payments for order flow they receive. The SEC wisely rejected the proposal.[\[231\]](#)

Crediting investors' accounts with the payments received by a broker would be unworkable. Uniform procedures for estimating the value of nonmonetary payments, such as the sharing of investment research or a dealer ensuring that a broker can participate as an underwriter in an upcoming public offering, would have to somehow be established. Moreover, the rebate scheme would run into difficulty because payments for order flow are often deferred, creating the possibility that some investors would no longer have an account with the broker when payments for order flow were actually received. Even assuming these difficulties were surmountable, a broker easily could circumvent a mandatory rebate scheme through internalization. This raises, in turn, the general point that any rebate scheme would have great difficulty in treating even-handedly the different ways dealers can share profits with brokers, whether it is through payments or internalization. To attempt to construct a rebate regime that could do this would be, at best, very burdensome.

Putting aside its administrative difficulties, a rebate scheme would still be a poor solution to the agency problem. The Midwest Stock Exchange's proposal rests on the assumption that payments are not already being "rebated" to investors through lower commissions, which is improbable. Erecting a complicated mandatory rebate scheme to displace market forces already at work would likely only create costs without any offsetting benefit. More fundamentally, the proposal suffers from the same basic shortcoming that the current SEC disclosure requirements do. An investor needs to know not only the size of the rebate a broker is offering but the value of the foregone price improvement. If an investor does not have this information, a rebate scheme would do little to encourage brokers to send order flow to auction markets that are incapable of providing the payments that would then be rebated to investors' accounts.

VIII. Conclusion

Justice Stone famously asserted after the Crash of 1929 that "most of [the financial industry's] mistakes and its major faults will be ascribed to the failure to observe the fiduciary principle, the precept as old as holy writ, that 'a man cannot serve two masters.'" [232] Even if this is an overstatement, it still holds more than a kernel of truth. While investor sophistication has undoubtedly increased over the last sixty-five years, the complexity of the securities markets and the choices facing investors' agents has also grown. Agency problems, and their proper regulatory treatment, will continue to constitute an important set of issues in securities regulation. The "payment for order flow" problem provides a timely reminder of this fact.

<https://web.eecs.umich.edu/srg/?p=1546>

Arguments about "front running"

Posted on | April 13, 2014 | *Comments Off*

One of the brilliant rhetorical devices deployed by Michael Lewis in his public interviews about *Flash Boys* is referring to some HFT practices as "legal front running". By inserting the "legal", he takes off the table any accusations of lawbreaking. Nevertheless, defenders of HFT seem to cling to legalistic definitions of "front running" in making their arguments that HFT algos do not front-run. As an example, let me take on Rishi K. Narang, founding principal of T2AM, who denounced the study authored by Elaine Wah and myself in a recent commentary published by CNBC:

Some allege that HFTs front-run other players' orders because HFTs have access to direct feeds from various exchanges, while "regular" investors generally rely on the SIP-provided national best bid-offer (NBBO), which aggregates all the various exchanges' order books, but which also arrives with considerable latency (due not to conspiracy, but to outdated technology). As a recent University of Michigan report claims, "[b]y anticipating future NBBO, an HFT algorithm can capitalize on cross-market disparities before they are reflected in the public price quote, in effect jumping ahead of incoming orders[...]". This is blatantly false.

Narang's argument is that HFTs cannot front-run because they are not specifically trading ahead of one of their own customers in order to then trade with that customer.

When an order is placed, it takes some time to be reflected in the NBBO. But that order is already in the market before the HFT can see it, even on the direct feed, by definition. HFTs never know what a customer's order is before it's in the market. HFTs have no customers.

The distinction between whether the order is from an HFT's own customer or not would be relevant to a legal allegation of front-running, but that is not in question here (and I certainly would not be qualified to speak on such legal points). The other part of the argument is that the order is already in the market, so there is no front-running. However, a lot hinges here on what we mean by "already in the market".

The HFT tactic we investigated in our paper is *latency arbitrage* (what Michael

Lewis calls "slow market arbitrage"), in which the HFT profits on a price disparity between fragmented markets, arising due to latency in the NBBO. Suppose an order to sell exists in exchange A at \$100.00, but the outdated NBBO ASK quote is \$100.02. If an order to buy at \$100.01 comes into exchange B (and stays there based on the outdated NBBO), the HFT can arbitrage between the two markets, buying in A and selling in B, pocketing \$0.01.

So is Narang arguing that this cannot happen, or that it does not amount to "in effect jumping ahead of incoming orders"? There is in fact much evidence that it can and does happen, so we are probably arguing about semantics of "in effect jumping ahead". Ultimately what we call it does not matter very much, but I don't see how one can deny that the HFT is exploiting its information advantage to take the sell order on exchange A before the incoming order at B sees it. The statement that the order at B is "already in the market" is true in a sense, but somewhat misleading. "The market", as we know, is not a centralized entity, but rather fragmented over a plethora of exchanges and non-exchange trading venues.

Back to Narang:

What's actually happening behind the scenes may be frustratingly complicated, but it's not immoral, unethical, harmful or illegal. Nor does it cost the non-HFT anything.

As noted, we are not arguing about legality. Nor did our paper suggest anything about morality or ethics. We did focus on harm, however, and found that indeed latency arbitrage is harmful to market efficiency. Not only does it cost the non-HFT (it's quite obvious that some transfer of surplus must be at play in latency arbitrage), we found that the non-HFT costs exceed the HFT profits. In other words, it's a negative-sum game. That is not even counting the costs to the HFTs themselves for achieving their speed advantage.