2018

High-Frequency Trading and the New Stock Market: Sense And Nonsense

Merritt B. Fox
Columbia University Law School, mfox1@law.columbia.edu

Lawrence R. Glosten
Columbia University Business School, lrg2@gsb.columbia.edu

Gabriel V. Rauterberg
University of Michigan Law School, rauterb@umich.edu

High-Frequency Trading and the New Stock Market: Sense And Nonsense

by Merritt B. Fox, Columbia Law School, Lawrence R. Glosten, Columbia Business School, and Gabriel V. Rauterberg, Michigan Law School*

Stock trading in the U.S. has been totally transformed over the last twenty-five years. The NASDAQ dealers and NYSE specialists are gone; the same stock can now be traded on up to 60 competing venues where computers match incoming orders. But not everyone is pleased with the results.

The new stock market features several controversial participants and practices. High-frequency traders ("HFTs"), which participate in a significant portion of all trades, are criticized as taking advantage of other traders by rapidly adjusting their own orders in response to transactions in a practice known as "electronic front-running." Also under suspicion are "dark pools," which are off-exchange trading venues that promise to keep orders secret and can limit trading to certain kinds of traders.[1] And perhaps most visibly, HFTs have been blamed for events like the infamous "Flash Crash" of May 6, 2010, a period of less than 30 minutes during which the Dow Jones Industrial Average dropped about 1,000 points (representing 9% of its value) and then recovered almost its entire loss. Polls indicate that "roughly two-thirds of Americans believe the stock market unfairly benefits some at the expense of others," a belief that some commentators use to explain a sharp drop in the percentage of Americans directly or indirectly owning equities.[2]

Critics have been vocal. Charlie Munger, vice chairman of Berkshire Hathaway, argued that high-frequency trading is "legalized front-running . . . [that] should never have been able to reach the size that it did."[3] And New York Attorney General Eric Schneiderman has complained that "[w]hen blinding speed is coupled with early access to data, it gives small groups of traders the power to manipulate market movements in their own favor before anyone else knows what's happening."[4] But the most critical and well-publicized attack on the new stock market appeared in Michael Lewis' best-selling book, *Flash Boys: A Wall Street Revolt*. Lewis famously claimed that "[t]he United States stock market, the most iconic market in global capitalism, is rigged."[5]

Regulators reacted rapidly to the furor over the new stock market ignited by Lewis' book. Soon after, the U.S. Department of Justice, the FBI, the Securities and Exchange Commission ("SEC"), and the Commodity Futures Trading Commission all confirmed investigations into HFTs. The New York Attorney General brought a high-profile lawsuit against the major investment bank Barclays, alleging that it misrepresented the extent to which its dark pool was free of HFT activity.[6] And several Congressional hearings were held, after which U.S. Senator Carl Levin wrote to Mary Jo White, the Chair of the SEC, demanding significant changes to market structure and the elimination of "[c]onflicts of interest [that] erode public confidence in the markets."[7]

In this condensed summary of our earlier work, we argue that the issues are more complicated. And because the performance of the U.S. equity market has important effects on not only the investment returns of ordinary individuals, but the overall efficiency and real rate of growth of the U.S. economy, much is at stake in how such issues get resolved and what policy interventions are targeted at them.

We will argue that effective resolution of these controversies must begin with a comprehensive framework for understanding the new stock market. While legal scholars have applied the insights of many economic theories to law, they have largely not done so with the field of market microstructure. This article uses the insights of microstructure economics to provide a framework that relies on two basic mechanisms—adverse selection and the principal-agent problem—to analyze these controversial trading practices as they operate within a multi-venue system. We apply this framework to five of the new market's most controversial practices and evaluate the effects in terms of the ultimate social functions served by the equity markets.

We conclude that some proposed reforms appear unambiguously desirable, such as those requiring brokers to improve their disclosures regarding their execution of

*This is a condensed version of The New Stock Market: Sense and Nonsense, 65 DUKE LAW JOURNAL 191 (2015).

1. Sam Mamudi, *UBS Hit With Record Dark Pool Fine for Breaking U.S. Rules*, Bloomberg, Jan. 15, 2015, http://www.bloomberg.com/news/2015-01-15/sec-fines-ubs-dark-pool-more-than-14-million-for-breaking-rules.html.

2. See Lydia Saad, *U.S. Stock Ownership Stays at Record Low*, Gallup Economy, May 8, 2013, *available at* http://www.gallup.com/poll/162353/stock-ownership-stays-record-low.aspx.

3. Sam Mamudi, *Charlie Munger: HFT is Legalized Front-Running*, Barron's, May 3, 2013, http://blogs.barrons.com/stockstowatchtoday/2013/05/03/charlie-munger-hft-is-legalized-front-running/.

4. Linette Lopez, *New York's Attorney General Has Declared War On Cheating High-Frequency Traders*, Bus. Insider, Sep. 24, 2013, http://www.businessinsider.com/schneiderman-targets-hft-front-running-2013-9.

5. Michael Lewis, *Flash Boys: A Wall Street Revolt* (2014).

customer orders, including those directed to dark pools. But other proposed reforms involve tradeoffs between different social goals, where the most socially desirable outcome is far from clear. In such cases, a better understanding of the tradeoffs involved should make for more informed regulatory choices, while also pointing to where further empirical research would be useful. We find this to be the case with proposals, for example, to briefly delay providing HFTs with information about new transactions and quotation changes, and so reduce HFTs' informational advantages over other traders. Finally, still other proposed reforms are bad ideas that seem to be based on a misunderstanding of how the market really works or of the actual impact of a given practice. We find this to be the case for proposals that would require HFTs to keep their quotes in force for some minimum amount of time, as well as for proposals aimed at generally discouraging, or even banning, trading on dark pools.

How the Stock Market Has Changed

As recently as the early 1990s, publicly traded stocks were still largely confined to trading on a *single venue*, which was either NASDAQ or the New York Stock Exchange. For anyone wanting to buy or sell a stock listed on NASDAQ, a member dealer was the purchaser of every share sold and the seller of every share bought. Dealers provided prices based on their own calculations and judgments as *individual human beings.* At the NYSE, stock "specialists" played a similar dealer role, but also posted quotes sent in by traders willing to buy or sell at stated prices, held auctions, and helped arrange trades by brokers and traders on the NYSE floor.

Today, the NASDAQ dealers and NYSE specialists are gone. Any given stock can be traded on one of almost 60 competing venues, twelve exchanges, and around 30 active dark pools.[8] Most of these competing trading venues, and all of the exchanges, are *electronic limit order books*, where a trader can post as a limit order its firm commitment to buy or sell up to a specified number of shares at a quoted price. A computer (the venue's *matching engine*) matches these posted limit orders with incoming buy and sell market orders, which are orders from traders willing to trade at whatever is the best available price in the market.[9]

Today, HFTs post a significant portion of the limit orders that result in executed trades.[10] An HFT uses high-speed communications and data about activity at venues to constantly update its information about executed trades in each stock that it regularly trades, as well as changes in the buy and sell limit orders posted by others on every trading venue. Computers use this information and proprietary algorithms to change the HFT's own limit orders posted on each of the various trading venues. More than three-quarters of all trades in the United States are executed on one or another of these venues.[11] Most of the remaining trades involve a broker that internally matches the buy and sell orders received from its own retail customers or through over-the-counter market-making.

Forces for Change and the Role of Regulation

The new stock market is partly the product of the information technology revolution, but also partly the result of Congressional and SEC choices. The initial impetus for this new market structure was Congress's adoption in 1975 of the National Market System ("NMS") amendments to the Securities Exchange Act of 1934 (the "Exchange Act").

Multiple, competing trading venues have the advantage of greater efficiency and stronger incentives for innovation. At the same, they have the disadvantage that orders are fragmented among multiple venues, complicating the matching of buyers and sellers. Improving information technology allows traders to manage this complexity by showing what is going on in each of these venues.

Congress' NMS amendments pushed the system to develop in this direction, and this development has been consistently supported by the SEC.[12] And the dramatically increased speed and lower cost of trading that have been documented since then suggest that the new stock market is a substantial improvement over what came before it.[13] Though academic theorists continue to debate whether even greater improvements would have arisen if today's technology were operating within a centralized single venue system, this is entirely a matter of speculation. Moreover, as a matter of political reality, any attempt to centralize the multiple venues would meet stiff resistance from those who have configured their businesses for a multi-venue structure. Thus, we believe

6. Complaint, *Schneiderman v. Barclays*, No. 451391/2014 (N.Y. Sup. 2014).

7. *See* Letter, July 9, 2014, http://levin.senate.gov/download/levin_letter_sec_070914.

8. Laura Tuttle, *Alternative Trading Systems: Description of ATS Trading in National Market System Stocks*, Division of Economic and Risk Analysis 5-6, http://www.sec.gov/marketstructure/research/ats_data_paper-_october_2013.pdf.

9. The computer will also match the limit orders posted on the venue with "marketable limit orders." A buy limit order is "marketable" when it has a limit price greater than or equal to the lowest offer in the market, and a sell limit order is "marketable" when it has a limit price less than or equal to the highest bid.

10. *See* Jonathan A. Brogaard, *High Frequency Trading and its Impact on Market Quality*, Working Paper (2010), which finds using the NASDAQ data set that HFTs supply liquidity for 51% of all trades and provide the market quotes 50% of the time. *See generally* Albert J. Menkveld, *High-Frequency Trading and the New-Market Makers*, 16 J. Fin. Markets 712 (2013).

11. Laura Tuttle, *OTC Trading: Description of Non-ATS OTC Trading in National Market System Stocks*, Division of Economic and Risk Analysis, https://www.sec.gov/marketstructure/research/otc-_trading_march_2014.pdf.

12. Congress, when the NMS amendments were adopted, expected that there would be a proliferation of competing venues. It self-consciously rejected a proposal for an electronic limit order book where all order flow was directed to a single trading venue, known as a central limit order book ("CLOB"). *See, e.g.*, S. Rep. No. 75, 94th Cong., 1st Sess., 12 (1975), 1975 U.S.C.C.A.N. 179, 190 ("Senate Report") (rejecting role for "the SEC . . . as an 'economic czar' for the development of a national market system" and noting that "a fundamental premise of the bill is that . . . a national market system . . . will depend upon the vigor of competition within the securities industry").

13. *See* James J. Angel, Lawrence E. Harris & Chester S. Spatt, *Equity Trading in the 21st Century: An Update* 11-12 (2013), which reports significant increases in the speed of execution, decreases in the bid-ask spread, decreases in commissions, and increases in the number of quotes per minute.

the policy favoring multiple venues is unlikely to be reversed in the future, and any reforms enacted will take place within the current multiple venue system.

The NMS amendments also included broad provisions for consolidating information in the U.S. stock market. The SEC requires trading venues to have systems (called SIPs); there is one for NASDAQ listed stock and one for securities listed either on the NYSE or elsewhere, which provide the best bid and best offer quotes for each stock traded. The SIPs aggregate this information into consolidated books with the best offer and best bid for a stock at each of the venues where it trades, along with the corresponding sizes, and they make this quote information available to the public on fair and reasonable terms. At any given time, the best bid and best offer on this consolidated book represents the official National Best Bid (NBB) and National Best Offer (NBO), which together make up the NBBO.

But because the national reporting system's updates lag slightly behind any change in the best bid or offer available, HFTs can use their co-location, private data feeds, and superior information technology infrastructure to carry out their practices of electronic front running, slow market arbitrage, and dark pool mid-point order exploitation. During the lags, they can cancel standing limit orders and post new ones—which means that the quotes in the consolidated book may no longer in fact be available.

Five of the Most Controversial New Practices

In the rest of this article, we focus mainly on five new stock market practices that have attracted particular controversy. The first three practices are made possible by the HFTs' "co-location" of computers right next to each exchange's matching engine, which allows the HFTs to learn about trades and adjust their limit bids and offers sooner than some other traders. (It must be stressed, though, that many traders other than HFTs also co-locate, acquire high-speed communications, and use private data feeds.) HFTs can cancel old limit orders and submit new ones very quickly. The HFT's co-location facility at each exchange is connected to all its other co-location facilities through specialized fiber optic cables that have their matching engines in northern New Jersey.

1. *HFT electronic front running.* Let's examine the practice of electronic front running through a simple example involving just one HFT, called Lightning, and two exchanges, BATS Y and the NYSE. Lightning has co-location facilities at the locations of the BATS Y and NYSE matching engines. These co-location facilities are connected with each other by high-speed communications equipment.[14] An actively managed institutional investor, Smartmoney, decides that Amgen's future cash flows are going to be greater than its current price suggests. The NBO is $48.00, with 10,000 shares being offered at this price on BATS Y and 35,000 shares at this price on NYSE. Smartmoney decides to buy a substantial block of Amgen stock and sends a 10,000 share market buy order to BATS Y and a 35,000 share market buy order to NYSE.[15] The 35,000 shares offered at $48.00 on NYSE are all from sell limit orders posted by Lightning.

The order sent to BATS Y arrives at its destination first and executes. Lightning's co-location facility there learns of the transaction very quickly. An algorithm infers from this information that an informed trader might be looking to buy a large number of Amgen shares and thus may have sent buy orders to other exchanges as well. Because of Lightning's ultra-high speed connection, it has the ability to send a message from its BATS Y co-location facility to its co-location facility at NYSE, which in turn has the ability to cancel Lightning's 35,000 share $48.00 limit sell order posted on NYSE. All this can happen so fast that the cancellation would occur before the arrival there of Smartmoney's market buy order. If Lightning does cancel in this fashion, it has engaged in "electronic front running."

Critics of the practice assert that this allows HFTs to benefit at the expense of institutional investors.

2. *HFT slow market arbitrage.* Suppose that the HFT Lightning has a limit sell order for 1,000 shares of IBM at $161.15 posted on NYSE. This quote represents the NBO at the moment. Institutional investor Mr. Lowprice then posts a new 1,000 share sell limit order for IBM on EDGE for $161.13. The national reporting system is a bit slow, and so a short period of time elapses before it reports Lowprice's new, better offer. Lightning's co-location facility at EDGE very quickly learns of the new $161.13 offer, however, and an algorithm sends an ultra-fast message to Lightning's co-location facility at NYSE informing it of the new offer. During the reporting gap, though, Lightning keeps posted its $161.15 offer. Next, Ms. Stumble sends a marketable buy order to NYSE for 1,000 IBM shares. Lightning's $161.15 offer remains the official NBO, and so Stumble's order could legally transact against it. Lightning's co-location facility at NYSE then sends an ultra-fast message to the one at EDGE instructing it to submit a 1,000 share marketable buy order there. This buy order transacts against Lowprice's $161.13 offer. Thus, within the short period before the new $161.13 offer is publicly reported, Lightning has been able to sell 1,000 IBM shares at $161.15 and purchase them at $161.13, for what appears to be a riskless $20 profit.

14. *See supra* Subsections I.A and I.C for a discussion of exchange matching engines and HFT co-location facilities.

15. This example fleshes out the story by Michael Lewis of how electronic front running could occur with Amgen stock in such a situation. Lewis, Flash Boys, at 33-34. Lewis asserts that the HFT could profit at the expense of others by cancelling its quotes on another exchange, but he does not discuss exactly why it would be profitable for the HFT to do so. Nor does he analyze how the quotes initially available might be different if the practice of electronic front running were eliminated. The discussion that follows fills in these holes.

3. *HFT exploitation of mid-point orders sitting in dark pools.* On yet another day, suppose that an institutional trader posts a mid-point limit buy order in a dark pool that, until cancelled, would execute against any market sell order that subsequently arrived at the dark pool and at a price equal to the mid-point between the best offer and best bid reported by the national reporting system on any of the exchanges. Through its speedy co-location facilities, the HFT Lightning would observe that the new best offer on that exchange is lower than the mid-point between what, until that moment, had been the best bid and best offer available on any public exchange. Because of the national reporting system's brief lag, Lightning could buy shares at the new better price and then immediately send a sell order to the dark pool, which executes against the trader's order at the mid-point between the still official, but now stale, best offer and best bid reported by the national system. Since the price paid for the shares by Lightning on the exchange is lower than the price at which they are sold to the trader in the dark pool, Lightning makes another guaranteed profit at the expense of the other traders in the market.

4. *Concerted selling by HFTs during market downturns, leading to increased volatility and crashes.* There was an upsurge in the volatility of share prices and a few brief crashes and breakdowns in trading as the new stock market was emerging. Such volatility and crashes have been attributed to the sudden exit of HFTs from the market after receiving new market information.

5. *Large investment banks* in their role as brokers steering orders to their own dark pools. A large investment bank steers an institutional customer's buy limit order to its own dark pool in a way that is unobservable by other traders. The bank's proprietary traders learn through an internal source of the institution's buy order, giving them the option to fill the institution's limit order even when that would be disadvantageous for the customer.

Undertaking a Serious Analysis

Most of the criticism of the new stock market simply shows that in retrospect a given transaction benefited one party at the expense of another, finds an advantage that favors the former, and labels the resulting transfer as "larcenous," "predatory," or simply "unfair." Serious analysis requires digging deeper, especially since these practices occur repeatedly between competing actors who generally understand what is happening and take into account the reaction of the other market participants. We offer an informal equilibrium analysis of these practices.

While each of the controversial practices seems fairly distinct, they can all be understood by reference to two basic dynamics at work in today's market structure:

Adverse selection. Limit orders substantially increase liquidity. Those who provide the orders are referred to as "liquidity suppliers" or "market makers." A professional supplier of liquidity—in this case, let's assume it's an HFT—buys and sells shares frequently, and it makes money if on average it sells the shares it buys for more than the price it paid. As discussed below, a major problem HFTs face is *adverse selection*: the possibility that another trader has private information about a stock's value that is not known to most of the market or to the liquidity supplier. When dealing with such traders, liquidity suppliers will on average lose money because the better informed will sell to the HFT only when it is willing to pay too much and buy from the HFT only when it is willing to sell for too little. To survive, liquidity suppliers must set bids and offers aggressively enough to attract business, but not so aggressively that they lose more money trading with informed traders than they make from uninformed traders. To minimize their losses from adverse selection, liquidity providers try to identify orders that are coming from informed traders. At the same time, informed traders try to prevent their orders from being so identified in order to buy or sell shares at the best possible prices.

Principal-agent problems. Most traders, even institutional ones, need brokers, and brokers often exercise substantial discretion when handling customer orders. *Principal-agent problems* arise from conflicts of interest between the broker (the agent) and the investor (the principal) because the investor usually cannot perfectly observe the broker's effort and skill.

Multiple venues. Finally, it's important to recognize that although managing adverse selection and principal-agent problems were challenging under the old single-venue market system, such challenges have been greatly increased by the existence of many competing venues for stock trading, which have made possible all of the controversial practices listed above.

In sum, the adverse-selection-driven cat and mouse game between liquidity suppliers and informed traders occurs within multiple trading venues and in the context of rapid information technology advances that have created extraordinary complexity as well as new scope for principal-agency problems between brokers and traders. By understanding how these three factors interact in a competitive environment, you can understand most of what is happening.

The Economics of Liquidity Provision

It might appear that a professional liquidity supplier such as an HFT could make money easily, even in markets with a one cent spread. For example, simply buying at the bid and selling at the offer to make a half cent per share on every transaction for a billion shares should yield a tidy—and apparently riskless—profit of $50 million. In fact, it is not so easy.

Liquidity suppliers generally do not know whom they are trading with. There is always the possibility that the trader who places a marketable order that executes against

the liquidity supplier's quote is doing so because the trader has nonpublic information regarding the value of the stock that is not known to the liquidity supplier.[16] Despite this informational disadvantage, the liquidity supplier can still make money on a net basis if it makes enough profit from the remaining traders who do not possess private information.

We identify three primary kinds of private information, which we label "inside information," "announcement information," and "fundamental value information."

1. Inside information. Inside information originates from some institutional source, such as the company issuing the stock itself. The institution usually seeks to prevent this information from becoming public, however, and trading on such information is, under many circumstances, illegal under Section 10(b) and Rule 10b-5 of the Exchange Act. Successful prosecutions under these provisions show that such information is behind at least some of the trading that occurs.

2. Announcement information. This is information that is disclosed publicly, such as a government statistic about the economy or a company earnings announcement. Traders who act on such information very quickly, before other traders and liquidity suppliers themselves can react and adjust their quotes, can earn trading profits. We refer to these as announcement informed traders.

3. Fundamental value information. Some investors use fundamental value information to produce more accurate estimates of an issuer's future cash flows based on sophisticated analysis of publicly available information. Traders with this kind of information include hedge funds, actively managed mutual and pension funds, non-profit institutions, and very wealthy individuals with actively managed portfolios. Liquidity suppliers can also be vulnerable to fundamental valuation traders because they specialize in supplying liquidity but do not do fundamental analysis themselves.

Adverse Selection

Liquidity providers, as already noted, lose money when trading with informed sellers or buyers. But the liquidity provider can still break even if the bid and offer spread is large enough that losses from informed traders are offset by the profits from trading with uninformed investors.

There are two ways to think about the calculations that liquidity providers need to survive in a competitive market. The first, sometimes referred to as the "accounting perspective," subtracts a liquidity supplier's losses from transacting with the informed traders from its gains from transacting with uninformed traders to determine the spread. The second, sometimes referred to as the "information perspective," focuses on how liquidity suppliers update their estimates of a stock's value in anticipation of whether the next order to transact against its quotes is a buy or a sell. Because a liquidity supplier knows that the next marketable order may come from an informed trader, that order will alter the liquidity supplier's estimate of the stock's value—and the adjustment will be up if it is a buy order and down if it is a sell order.[17]

Moreover, as liquidity suppliers constantly update their valuations in response to the inflow of buy and sell orders, the market comes to reflect private information. If the news possessed by the informed traders is on balance favorable, there will be more buys than sells, and the bid and offer quotes will trend upward. But if the news known by the informed traders is bad, the mid-point between the bid and ask will trend downward.[18]

The Evaluative Framework

HFTs and investment banks trade in a competitive market on a repeated basis, and the other actors in the system generally take this fact into account in their own actions. The question for policymakers and regulators, then, is how their practices affect the multiple social goals that equity trading markets are expected to serve.

The most important social goals of secondary equity markets are generally thought of as including: (1) promoting the efficient allocation of capital so that it goes to the most promising new investment projects in our economy; (2) promoting the efficient operation of the economy's existing productive capacity; (3) promoting the efficient allocation of resources between current and future periods; (4) allocating capital and risk among risk-averse investors in ways consistent with their capabilities and resources; (5) fostering an overall sense of fairness among market participants; and accomplishing all these objectives while (6) economizing on the real resources used in trading markets, including enforcement and compliance costs, and (7) encouraging valuable innovation in the system.

Two central characteristics of a stock market affect its ability to deliver on these social goals and serve as proxies for their success in so doing—namely, share price accuracy and liquidity.

1. Price accuracy. An accurate share price does a reasonably good job of predicting the issuer's future cash flows. Because the price of any new share offering will be determined largely by the price of its already outstanding shares, more accurate stock market prices will encourage capital

16. *See* Lawrence E. Harris, *Trading and Exchanges* 158 (2002), which summarizes a body of theory and evidence suggesting that in most markets adverse selection accounts for the majority of the bid-ask spread.

17. See Lawrence R. Glosten and Paul R. Milgrom, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," 14 J. Fin. Econ. 1, (1985).

18. *See* Lawrence R. Glosten and Lawrence E. Harris, "Estimating the Components of the Bid-Ask Spread," 21 J. Fin. Econ. 123 (1988), which estimates a model in which the bid-ask spread is divided into an adverse selection component and a transitory component due to inventory costs, clearing costs, and other factors. .

to flow to the issuers with the most promising investment projects. Share prices also influence the availability of new project funding from other outside sources and the willingness of managers to use internal funds for investment, creating another link between share price accuracy and the efficient allocation of capital. And, finally, more accurate share prices also tend to create a greater sense of fairness among investors to the extent that they experience fewer very large negative surprises.[19]

2. Liquidity. Liquidity is a multi-dimensional concept that relates to the size of a trade, the price at which it is accomplished, and the time it takes to complete the trade. Generally, the larger the size of the purchase or sale and the faster one wishes to accomplish it, the less desirable will be the price. But the more liquid the market, the less severe are these tradeoffs.

For a small retail purchase or sale of stock, the spread between the national best offer (NBO) and the national best bid (NBB) is a good measure of liquidity because the trader can buy or sell immediately at those prices and, in essence, will be paying half the spread to do so. For larger orders, the quantity of stock that is available at prices that are not too far above the NBO or too far below the NBB (both indicators of the "depth of the book") will become relevant as well.

Liquidity also has an impact on a number of social goals:

a. More efficient allocation of resources over time. The more liquid an issuer's shares, all other things equal, the more valuable they are. In this sense, greater liquidity can be seen as reducing the issuer's cost of capital, thereby encouraging it to take on more investment.[20]

b. Greater share price accuracy. To the extent more liquidity also lowers the transaction costs associated with trading based on fundamental, value-based investment strategies—that is, acquiring and analyzing publicly available information to make more accurate predictions of an issuer's cash flows and earnings—an increase in liquidity can also lead to more accurate share prices.

c. More efficient allocation of risk. Constant change means that the optimal portfolio, in terms of diversification and of each investor's relative degree of risk aversion, is always shifting. Greater liquidity increases individual investors' ability to make cost-effective adjustments of their portfolios over time.

Analyzing the Five Most Controversial New Stock Market Practices

Electronic Front Running

So-called "electronic front running" involves an HFT learning of a transaction that has occurred at one exchange and adjusting its quotes at other trading venues. The most obvious reason for doing this is that the HFT has inferred that orders similar to the one that executed may still be in transit heading towards other exchanges; and the HFT, for reasons that will be discussed, may want to avoid transacting with those orders.

All of the criticisms of HFT electronic front running focus on the fact that the HFT can be expected to be better off and some other trader involved worse off. It should be noted at the outset, however, that the HFT practice labeled as "electronic front running" is distinctly different from the kind of behavior that has traditionally been termed "front running." Traditional front running, which is clearly illegal, refers to the practice of a broker—who bears a legal duty to its clients not to use their orders to its own advantage—trading ahead of said orders to realize a gain. In contrast, when an HFT engages in "electronic front running," it has no preexisting relationship with the trader akin to what a broker is obligated vis-a-vis its customer. And the practice thus involves no breach of duty or mutually agreed upon terms between contracting parties, nor does it involve any breach by HFTs of the federal anti-fraud laws. (A better term for the practice might be "inter-venue order cancellation," but we will stick with the popular term here.)

Our analysis of electronic front-running is somewhat involved, so a few summary points are in order. Basically, permitting electronic front-running enables liquidity providers to more easily adjust their quotes at trading venues in response to information about quotes and transactions that they receive at any given venue. Essentially, such adjustments make it more difficult for traders who want to *transact rapidly in large size* before liquidity providers can do exactly that—adjust their quotes. Investors who wish to purchase or sell only a small volume of stock will be indifferent to order cancellation since they can simply transact at the top of the book at one venue. Also largely unaffected by the adjustments of liquidity providers are those investors who wish to transact in significant volume, but have a considerable period of time to do so, and so can simply send in small orders over an extended period of time.

Thus, the distributional and efficiency consequences of electronic front-running turn on precisely *who* is interested in transacting rapidly in large size and with whom they so trans-

19. In an efficient market, the market price, whether it is relatively accurate or inaccurate, is an unbiased predictor of an issuer's future cash flows. If it is inaccurate, it is just more likely to be far off, one way or the other, from how things ultimately turn out. Thus an efficient, but relatively inaccurate, price would result in as many positive surprises as negative ones. To many investors, the negative surprise is likely to be more memorable. So when a negative surprise materializes, it generates a sense of grievance even though, ex ante, a positive surprise was equally likely.

20. The cost of capital is lower because the prospect of a smaller bid/ask spread results in the same issuer's expected future cash flow being discounted to present value at a lower discount rate. *See* Yakov Amihud and Haim Mendelson, "Asset Pricing and the Bid-Ask Spread," 17 *Journal of Financial Economics* (1986).

act. Because trade data is anonymous, our analysis must rely on a stylized characterization of market participants based on the available empirical data and the implications of trading needs.

A. Wealth transfer considerations. To see the distributive effects of electronic front running, we will begin by assuming that there are only three kinds of market participants: HFTs, informed traders who trade on the basis of fundamental value information, and uninformed traders.

Why might Lightning wish to cancel its sell limit order on NYSE? One possibility is that given its inference that a large market buy order is likely soon to arrive at NYSE, Lightning wishes to submit, in place of its cancelled order, a new sell limit order for the same number of shares at a higher price—say, $48.02. If Lightning does so and Smartmoney's buy order executes against this new higher quote, the HFT will be better off, and Smartmoney worse off, by $.02 per share.

Note, though, that the HFT will be able to improve its position in this way only if there is room in the NYSE limit order book so that the $48.02 offer price is still more attractive to potential buyers than any other offers with respect to what Amgen already posted on NYSE. Suppose, for example, that prior to Lightning's cancellation, the next best offer on the NYSE was 15,000 shares at $48.01 and the best offer after that was 20,000 shares at $48.02. The price and time priority rules would mean that Smartmoney's buy order would execute against these other two standing offers, not against any new $42.02 offer by Lightning.

This cautionary note, though, hides a more critical point: Lightning may wish to cancel its $48.00 sell limit order even if in fact there is no room in the book to improve its position by selling to Smartmoney at a higher price. Recall that to survive in a competitive market, a market maker like Lightning must set its quotes aggressively enough to attract business, but not so aggressively that the profit it makes when buying from, and selling to, uninformed traders is less than what it loses by engaging in such transactions with informed traders. $48.00 was what Lightning calculated at the time it posted its sell limit order to be the optimal price for an offer of 35,000 shares, based on what it knew then about the likelihood of the existence of positive private information. Now, however, Lightning knows something more: a large buy order has transacted on BATS Y. This will cause Lightning to revise upward its assessment of the likelihood that private information suggests that the value of a security is higher than the market previously thought. The upward revision is very possibly large enough that $48.00 is no longer the optimal price at which to offer to sell shares. In that case, Lightning will be better off cancelling its $48.00 limit offer on NYSE.

As this example suggests, the fundamental distributional effect of permitting electronic front-running is thus to enable liquidity providers to reduce their losses to informed traders who are attempting to trade rapidly in large size.

Further, the ability of liquidity providers to reduce losses to the informed has two significant consequences:

1. Electronic front running narrows spreads. The availability of electronic front running by HFTs allows HFTs to better detect the possibility that informed market orders are headed for their limit orders. If HFTs did not have the ability to learn these things and alter their standing limit orders accordingly, they would know that a larger percentage of the trades that will execute against their limit orders will come from informed traders. And the primary cost of being a liquidity supplier—the losses incurred from dealing with informed traders—would therefore go up. Accordingly, HFTs would widen their initially posted bid/ask spreads to compensate.

2. Electronic front running helps uninformed investors and hurts informed investors. If electronic front running were eliminated, uninformed traders and informed traders would both suffer from the resulting larger spreads—the higher offers and lower bids—because it will be more expensive for both to trade. For uninformed traders, that is the end of the story. Informed traders, however, would get a more-than-compensating benefit. To see why, note that because eliminating electronic front running would make it more difficult for liquidity providers to detect informed traders, HFTs would increase their spreads sufficiently to cover the expected trading losses against informed traders, but not so much as to undermine their competitive position.[21] And because the increased spreads will be borne by all traders, informed and uninformed alike, the higher spreads paid by the uninformed traders will effectively "subsidize" the informed traders who would otherwise have incurred even larger spreads.[22] And this means that informed traders come out ahead; the gains they would have enjoyed without the increase in spreads are not fully dissipated by the extra they must pay because the spreads in fact are increased. The rest of what HFTs need to break even comes from uninformed traders, who must pay the increased spread too.

In sum, electronic front running benefits uninformed investors and harms informed ones who seek to trade rapidly in large size.

B. The ultimate incidence of electronic front running. Electronic front running has been regularly attacked as harming "ordinary investors."[23] Our analysis, however,

21. This statement assumes that the increase in spreads would not decrease the volume of trading but, in fact, an increase in spreads makes trading more costly, suggesting that the volume would be lower with the increase in spreads than without it.

22. Reg. NMS precludes exchanges from restricting access to trading on their facilities. *See* Regulation National Market System Rule 610(a), 17 C.F.R. § 242.610(a) (2005) (prohibiting "national securities exchange[s] [from] . . . prevent[ing] or inhibit[ing] any person from obtaining efficient access" to trading against the buy and sell quotes posted on exchanges); Securities and Exchange Act Section 6, 15 U.S.C. § 78f (1934) (providing that "the rules of [a registered] exchange [must] provide that any registered broker or dealer . . . may become a member of such exchange").

23. *See, e.g.*, Lewis, Flash Boys, at 104;

suggests that this is mistaken. Retail investors generally lack any significant private information and are assumed to be uninformed. Small uninformed investors are helped, not hurt, by electronic front running.

Most of the persons whose money is invested in index-based mutual funds and pension funds would also presumably count as ordinary investors. These entities too, by definition, are uninformed traders. The purchases and sales of such funds are not prompted by any kind of private information; they simply buy all the stocks in the index when they receive a net inflow of investor funds and sell all stocks in the index when the volume of investor redemptions is sufficient to result in a net outflow of funds. Again, electronic front running, by narrowing spreads and reducing the cost of trading, generally helps, not hurts, these funds and their ordinary investors. However, insofar as index funds sometimes find themselves needing to trade rapidly in large size, they too will suffer from the availability of electronic front-running.

Critics have pointed out that the beneficiaries of electronic front running are the exchanges and the HFTs themselves[24]—and here they are closer to the mark. An exchange charges HFTs fees for permitting co-location: namely, the right to place the HFT's server very near the exchange's matching engine. If electronic front running were eliminated tomorrow, HFT co-location facilities would be worth less to the HFTs and this may reduce the rents collected by the exchanges. Any such reduction in rents certainly would hurt the exchanges, at least in the short run.[25] In the much longer run, the revenues of firms in a competitive industry can be expected to just equal their costs, including an ordinary market return on capital. Thus, any revenues lost from co-location fees would need to be made up through higher charges to investors who trade on the exchange.

C. Efficiency considerations. Recall that the fundamental effect of electronic front-running is to make it harder to trade rapidly in large size without liquidity providers adjusting their quotes. Assessing the efficiency consequences of this means understanding how the relevant participants are affected.

Elimination of electronic front running would have three effects in terms of the efficient operation of the economy, two of which would appear to be efficiency-increasing and one efficiency-reducing.

- Improved share price accuracy. Elimination of electronic front running would make it more profitable for informed traders to generate new private information and so they will do more of it, thereby making prices more accurate.
- Reduced resources going to HFT activities. Eliminating electronic front running would reduce the productive resources currently devoted to it, including highly sophisticated technical personnel, advanced computers, and fiber optic networks.[26]
- Allocation of resources over time and allocation of risk. By widening spreads, elimination of electronic front running would make the equities market less liquid. This has an unambiguously negative effect, both on share prices and capital allocation, and on the efficient allocation of risk throughout the economy.

D. Taking other kinds of informed traders into account. As mentioned, in addition to fundamental value information, two other types of private information can give a trader a significant advantage: announcement information and inside information. These additional kinds of private information do not change the conclusions above that electronic front running has positive effects on uninformed investors or that electronic front running consumes real resources. But, taking account of these additional kinds of private information may well change the conclusion above about the impact of electronic front running on fundamental value information traders and hence the impact on price accuracy.

One might conclude that eliminating electronic front running would help traders with announcement information and inside information more than traders with fundamental value information. If HFTs need to increase spreads sufficiently to cover their increased trading losses, fundamental value information traders would have to pay as much extra per trade as traders on the other two kinds of private information, but would get only a small portion of the additional trading gains. It is thus quite possible that fundamental value information traders will gain less than they pay in increased spread and thus will be hurt by the elimination of the practice.

This is because fundamental value traders are less susceptible to detection by electronic front runners than the other two kinds of private information traders. Announcement information traders need to do all of their trading quickly and therefore need to do larger transactions, which are easier for HFTs to detect and react to. Fundamental value traders, by contrast, often spread their planned purchases or sales over several days or weeks, and so break the total amount they wish to transact into small packets that look more like the trades of uninformed traders. Admittedly, we would need to know much more to make this characterization definitively, but the longer the time period before other market participants get wind of the information possessed by an informed trader, the less that trader's incentive to trade in substantial

24. *See, e.g.*, Lewis, Flash Boys, at 126, 176.

25. However, the impact of eliminating any of these practices is uncertain because HFTs desire co-location for a number of reasons. *See* Charles M. Jones, *What Do We Know about High-Frequency Trading*, Columbia Business School Research Paper No. 33-36, at 10, 26 (2013) (discussing that HFTs seek co-location to minimize their latency in learning of quote changes and in altering their quotes and analyzing empirical evidence that the introduction of co-location improves liquidity).

26. While high-frequency traders are notoriously secretive, HFT Virtu Financial, Inc. ("Virtu") did make certain public disclosures in the run up to its now postponed IPO. In 2013 alone, Virtu reported spending approximately $65 million on communication and data processing and $78 million on employee compensation and payroll taxes. Since Virtu has only 151 employees, this means they pay an average salary of about $517,000. Virtu is just one of several large HFTs and there are many smaller ones as well. *See* Form S-1 of Virtu Financial, Inc., https://www.sec.gov/Archives/edgar/data/1592386/0001047469140-02070/a2218589zs-1.htm#dm16701_business.

volume quickly. An announcement trader must trade quickly, since the signal that makes them informed has just become public. Not so with the fundamental value informed trader.

In fact, further research may well suggest that electronic front running actually helps, not hurts, fundamental value information trading. And to the extent this is so, we would have to modify our earlier conclusion that electronic front running would reduce share price accuracy.

Slow Market Arbitrage

Slow market arbitrage can occur when an HFT has posted a quote representing the NBO or NBB on one exchange, and subsequently someone else posts an even better quote on a second exchange, which the HFT learns of before it is reported by the national system. If, in the short time before the national report updates, a marketable order arrives at the first exchange, the order will transact against the HFT's now stale quote. The HFT, using its speed, can then make a riskless profit by turning around and transacting against the better quote on the second exchange. (For an illustrative example of slow-market arbitrage, see Section 1 of the APPENDIX.)

A. Wealth transfer effects. In contrast to electronic front running, which decreases the effective cost of trading for uninformed traders but increases it for informed traders, slow market arbitrage increases the effective cost of trading for all regular traders, informed and uninformed.

B. Efficiency considerations. Although arbitrage usually has positive economic welfare effects, slow market arbitrage has little in common with ordinary arbitrage. Slow market arbitrage adds a third party, the liquidity supplier, whose only social purpose is to facilitate trades between regular traders, but who are the only gainers from the so-called arbitrage. Regular traders, both informed and uninformed, are losers because their cost of trading goes up. So the normal presumption in favor of activities carrying the label "arbitrage" does not apply here.

Even if slow market arbitrage consumed no real resources, it would have an unambiguously negative impact on welfare. By raising the effective cost of trading for informed traders, slow market arbitrage makes it less rewarding for fundamental investors to seek out publicly available information and analyze their implications in a sophisticated way. This reduces share price accuracy, with all the negative effects already described.

HFT Exploitation of Mid-Point Orders

A trader will often submit to a dark pool a "mid-point" limit buy or sell order, the terms of which require that it be executed against the next marketable order with the opposite interest to arrive at the pool and at a price equal to the mid-point between the best publicly reported bid and offer at the time of execution. Mid-point orders appear to have the advantage of enabling uninformed investors to buy at well below the best offer and to sell well above the best bid.

It has been noted for a number of years, however, that the traders who post such orders are vulnerable to the activities of HFTs.[27] Mid-point order exploitation again involves an HFT detecting an improvement in the best available bid or offer on one of the exchanges before the new quote is publicly reported. The HFT puts in an order to transact against the new improved quote, and then sends an order reversing the transaction to a dark pool that contains mid-point limit orders with the opposite interest that transact at a price equal to the mid-point between the now stale best publicly reported bid and offer. (For an illustrative example of such mid point-order exploitation, see Section 2 of the APPENDIX.)

A. Wealth transfer and efficiency considerations. HFT exploitation of dark pool mid-point orders clearly provides rents to HFTs. There is no social benefit from this activity since it is unrelated to the main positive function that we have attributed to HFTs—namely, providing liquidity in a world with both uninformed and informed traders. Since trading is a zero-sum game, if the HFTs gain, certain regular traders must lose.

The economic function of dark pools is to provide a place for uninformed traders to lower their costs by trading with other uninformed traders. By undermining the ability of such traders to do this, mid-point exploitation by HFTs hurts not only those who use dark pools but also those who would have used them but for this higher cost. This will reduce the efficiency of both the allocation of resources over time and the allocation of risk in the economy.

Nevertheless, to the extent that the practice steers more uninformed traders to the exchanges, it leads to a narrowing of spreads on the exchanges, thereby reducing the cost of fundamental value information trading and thus improving share price accuracy.

High-Frequency Trading and Volatility

When making his case that HFT activity causes greater volatility in equity markets, Michael Lewis asserts that the intra-day price volatility of the stock market was 40% greater between 2010 and 2013 than it was between 2004 and 2006, and associates this change with the enactment of Reg. NMS and the rise of HFTs.[28] But there is a major problem with this comparison: the years 2004-2006 were ones of uncharacteristically low volatility, below that of any other two-year period from 1998 to 2012.[29] And the years 2010-2013 are also unrepresentative in the sense that they came in the wake of the most

27. *See, e.g.*, Alex Paley, *Navigating the Dark Pool Landscape*, Deutsche Bank 46 (2010) https://autobahn.db.com/microSite/docs/Navigating_Dark_Pool_Landscape.pdf. This point that was also noted in *Flash Boys,* p. 113-118.

28. Lewis, Flash Boys at 112.

29. Bank for International Settlements Papers No. 29, *The Recent Behaviour of Financial Market Volatility* 1 (2006).

severe financial crisis since the Great Depression and thus significantly increased uncertainty about the fundamental values of securities.[30] A more useful and revealing comparison would have shown that market volatility during the period 2012 to the present, even with the expanded HFT activity, was considerably lower than the volatility experienced during the comparably long (and more representative) period of the 1990s and early 2000s.

In sum, there is little serious evidence of a causal link between HFTs and ongoing increased volatility: HFTs, as just noted, rose to prominence during a period of greater volatility that was attributable to economic causes that had little to do with the HFTs themselves. And there is also no theoretical reason for expecting HFT activity to increase general, ongoing volatility. Indeed, the majority of academic evidence on the subject suggests that the activity of HFTs reduces such volatility.[31]

A. The Flash Crash. More interesting and plausible is the claim that HFTs exacerbate volatility during market disruptions, such as the infamous May 6, 2010 "Flash Crash." The Flash Crash occurred during a period of less than 30 minutes in which the Dow Jones Industrial Average dropped about 1,000 points (representing 9% of its value) and then recovered almost its entire loss. This was the DJIA's greatest one-hour decline in history and several individual stocks displayed astonishing volatility. [32] Accenture, for instance, fell from $39.98 at 2:46 p.m. to one cent at 2:49 p.m., only to return to $39.51 by 2:50 p.m.

The Flash Crash was widely taken to "highlight the risks of electronic trading," as suggested in a report by NYSE's then head of operations.[33] And in the years since, other commentators have also blamed HFTs for the severity of market crashes.[34] However, the report eventually issued by federal regulators explained the Flash Crash not as the result of HFT predation, but as the result of a liquidity crisis caused by a series of large sell orders that triggered a flight of liquidity from the market. This flight involved HFTs, but only in the sense that many HFTs are market makers who left the market in response to the large sell orders. This temporary disappearance of the HFTs removed substantial liquidity.[35]

The crucial question is: Why would a large market sell order trigger a flight by HFTs, when the business of HFTs is to provide liquidity to persons submitting marketable orders? The short answer is that, as we have seen, adverse selection shapes the provision of liquidity.[36] The Flash Crash is directly connected to adverse selection. A large, aggressive sell (or buy) order suggests to liquidity providers that the order submitters may have important private information. If that is correct, then HFTs will lose money from trading that order and so they will widen their spreads. If the adverse selection threat becomes extreme enough, many or all liquidity providers will temporarily exit from the market altogether and prices will fluctuate widely.[37] This happened on a large scale during the Flash Crash.

In sum, the behavior of HFTs during the Flash Crash was not predatory; it was simply self-preserving and unheroic.[38] Moreover, the history of human market makers' responses to crises is largely consistent with this episode.

B. Wealth transfer considerations. The wealth transfers resulting from gyrations such as the Flash Crash are the same as those that occur at other times when HFTs stop providing liquidity. The losers are the traders who put in market sell orders for stocks that temporarily went way down and market buy orders for stocks that temporarily went way up. The winners were those who posted previously way-out-of-the-money limit orders against which these market orders transacted.

C. Efficiency considerations. Events such as the Flash Crash receive a lot of public attention, but such occasional brief moments of total collapse of liquidity are not ultimately very important in terms of the performance and efficiency of the overall economy—though, if large and frequent enough, they could have important effects on investors' confidence in the market. But barring that possibility, such sharp but very brief deviations of share prices from fundamental values do not seriously undermine capital allocation; it is accuracy most

30. *See* Angel et al., *supra* note 15, at 11-12; *see also* John Y. Campbell, Martin Lettau, Burton G. Malkiel & Yexiao Xu, *Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk*, 56 J. Fin. 1 (2001) (finding significant spikes in volatility during periods of major economic crisis).

31. *See, e.g.*, Joel Hasbrouck & Gideon Saar, *Low-Latency Trading*, 16 J. Fin. Markets 646 (2013), which finds that HFT activity reduces volatility. See also Jonathan Brogaard, Thibaut Moyaert & Ryan Riordan, *High-Frequency Trading and Market Stability*, Working Paper (May 2014).

32. *See* Tom Lauricella & Peter A. McKay, *Dow Takes a Harrowing 1,010.14-Point Trip*, Wall S.J., May 7, 2010 and Tom Lauricella & Scott Patterson, Legacy of the 'Flash Crash': Enduring Worries of Repeat, WALL S.J., Aug. 6, 2010, http://www.wsj.com/articles/SB10001424052748704545004575353-443450790402. Many of the most outlandish transactions executed during the Flash Crash were later cancelled or "broken" by regulators. See Deborah L. Jacobs, *Why We Could Easily Have Another Flash Crash*, FORBES, Aug. 9, 2013, http://www.forbes.com/-sites/deborahljacobs/2013/08/09/why-we-could-easily-have-another-flash-crash/.

33. *Id*.

34. *See, e.g.*, Andrew Smith, *Fast Money: The Battle Against the High Frequency Traders*, Guardian, June 7, 2014, http://www.theguardian.com/business/2014/jun/07/inside-murky-world-high-frequency-trading; Michael Ono, *High Frequency Trading May Magnify Market Woes*, ABC News, Aug. 11, 2011, http://abcnews.go.com/Business/high-frequency-trading-accelerating-market-woes/story?id=14280847 (suggesting that "computer-driven high frequency trading is partially responsible for accelerating stock gyrations").

35. Flash Crash Report at 6.

36. This article focuses on HFTs as liquidity providers, and there is ample evidence they play this role. *See, e.g.*, Albert J. Menkveld, *High-Frequency Trading and the New-Market Makers*, 16 J. Fin. Markets 712 (2013).

37. Flash Crash Report at 2-3.

38. David Easley, Marcos López de Prado & Maureen O'Hara, *The Microstructure of the 'Flash Crash': Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading*, 37 J. Portfolio Mgmt. 118, 120-26 (2011) (suggesting that order flow was especially informed and hence toxic for market makers in the period preceding the Flash Crash). Perceiving the large sell order to have a higher probability of being motivated by private information, given its size and aggressiveness, HFTs removed their quotes to minimize their trading losses, and liquidated the long positions they had accumulated, exacerbating pressures on price declines; 31 of Flash Crash Report at 29. Because HFTs provide a large share of liquidity, in their absence, the only quotes left lay far from the true price of a security. See Flash Crash Report at 45-57.

of the time that matters. The modern stock market's overall performance in terms of liquidity provision and operational costs is far better than the market of the past.

Dark Pools and the Fate of Customer Orders

Large investment banks, which are both important brokers and operators of dark pools,[39] have been accused of directing their brokerage orders to their own dark pools even when the orders will receive inferior execution there.[40] Dark pool operators are also alleged to misrepresent the nature of other parties' trading in their pools in order to induce brokerage customers to use the pools. Customers have difficulty detecting such practices; and even when they do, they are allegedly reluctant to switch brokers because they depend on "soft money" services from the banks.[41]

We do not know whether any of these practices is widespread,[42] though it's worth noting that the SEC has brought a number of successful proceedings against dark pool operators.[43] These practices are clearly illegal, and their wealth transfer and efficiency effects appear completely negative. If evidence emerges that they are in fact widespread, we would suggest policy reforms designed to make enforcement of the current laws more effective.

A. Understanding the function of dark pools. A dark pool, like an exchange, is typically an electronic limit order book; but unlike an exchange, it does not publicly reveal the limit orders that are posted on it. Dark pool operators restrict who can post limit orders and submit marketable orders.[44] Despite their nefarious-sounding moniker, dark pools can provide useful, legitimate services to their customers. Such pools were initially created with the aim of limiting adverse selection costs by providing a venue where uninformed buyers and sellers could trade substantial amounts of stock at prices potentially much better than the NBO and NBB.[45] The mid-point is a substantially better price for the buyer than the NBO, and it is the same for the seller relative to the NBB.

The ideal dark pool would be one where the parties posting limit mid-point orders and sending in marketable orders are completely uninformed. The system begins to break down when dark pool traders are informed. Since informed traders will transact against limit orders in the dark pool only when the mid-point price looks advantageous to them, such trades are likely to be disadvantageous to the person posting the limit offer. Thus, the dark pool operator provides a service if it can effectively monitor the parties posting the mid-point limit orders and the parties sending in marketable orders to ensure that both sides are relatively unlikely to be informed.[46]

B. Wealth transfer and efficiency considerations. An order sent to a less than ideal dark pool may execute at less desirable terms than at another venue. If an investment bank sends a trader's order to its own dark pool knowing that the order would receive superior execution elsewhere, the bank gains and the customer loses.[47] The same result is likely if the bank ignores customer instructions or if it misrepresents the nature of the parties allowed to trade on the bank's dark pool to create the impression that there is less danger of informed counterparties than there really is. All of these effects make investment banks richer and traders poorer. What's more, brokers have a legal duty of best execution in routing their customers' orders, one that should be enforced as vigorously, but also as cost effectively, as possible.[48] Moreover, such practices are inefficient for the simple reason that fraud, misrepresentation, and failure to carry out customer orders as directed all end up undermining the voluntary nature of transactions, and thus the underlying premise that trade is mutually beneficial and so welfare enhancing.

Recommendations

Potential regulatory responses to these five practices can be seen as falling into three groups: (1) proposals designed to limit the negative effects of front-running and other practices associated with HFTs' speed and informational advantages;

39. An underlying premise of these criticisms is that the largest investment banks are also among the most prominent brokers and dark pool operators. For instance, Lewis often discusses dark pools as being operated by Wall Street banks, which is accurate—six of the ten largest dark pools are run by major investment banks, *see* Rhodi Preece, *Dark Pools, Internalization, and Equity Market Quality*, CFA Institute Codes, Standards, and Position Papers 14-15 (2012). All of the ten largest brokers on NYSE are also global investment banks. *See* NYSE Market Data, *NYSE Broker Volume*, http://www.nyxdata.com/ (last checked Jan. 16, 2015).

40. Michael Lewis, for example, claims that dark pool operators sell access to their trading venues to HFTs—without disclosing this practice to other users—and that these HFTs then exploit other traders. Lewis, Flash Boys at 123. Inferior execution could also occur on a dark pool if the counterparties trading there are especially informed or were given information about the existence of the customer limit orders posted there.

41. *Id*. at 102-03, 214-15.

42. Most recently, New York Attorney General Eric Schneiderman filed a civil suit against Barclays alleging that Barclays' dark pool, Barclays LX (then the second largest in the U.S.) misrepresented to users the involvement of HFTs in LX, the informational advantages given to HFTs, and that Barclays, as a broker, claimed that it treated all venues the same based on quality, while it actually disproportionately routed client orders to its own pool. See Complaint, *Schneiderman v. Barclays*, No. 451391/2014 (N.Y. Sup. 2014), at ¶¶ 1-2.

43. *See, e.g.*, *In the Matter of Pipeline Trading Systems LLC*, SEC Release No. 33-9271; *In the Matter of Liquidnet Inc.*, SEC Release No. 33-9596.

44. *See* Regulation of Exchanges and Alternative Trading System Rule 301(b)(5), 17 C.F.R. § 242.301(b)(5) (1997); Concept Release on Equity Market Structure, Securities Exchange Act Release No. 34-61358, 17 C.F.R. § 242, at 72 ("As [trading systems] that are exempt from exchange registration, [off-exchange platforms] are not required to provide fair access [to all traders] unless they reach a 5% trading volume threshold in a stock, which none currently do[es]" and that "[a]s a result, access to . . . [these platforms] . . . is determined primarily by private negotiation.").

45. *See, e.g.*, Rhodi Preece, *Dark Pools, Internalization, and Equity Market Quality*, CFA Institute Codes, Standards, and Position Papers 12-13 (2012).

46. The operator provides a similar service to the extent that it keeps out HFTs that engage in mid-point order exploitation.

47. A broker can make money off transactions occurring on its dark pool for several additional reasons. If it is executing marketable orders on its dark pool, then a broker will receive its commission without having to subtract the "taker" fee charged marketable orders on most exchanges. If the broker is internalizing orders on its own dark pool and transacting against them as principal, then it can make half the spread on each trade. And then there are the more nefarious inducements suggested by the criticisms, such as exploitation of orders by a broker's HFT affiliate that has improperly been given details about orders.

48. This requires the broker to exercise "reasonable diligence to ascertain the best market" for a transaction to ensure an order receives a price "as favorable as possible under prevailing market conditions." In essence, the duty of best execution is a default term in the contract between the broker and its customer. Its violation leads to the same efficiency concerns that any other breach would: the fact that the parties voluntarily entered into the transaction no longer leads to the presumption that it can be expected to

(2) proposals intended to limit the effects of HFTs on stock market volatility; and (3) proposals intended to limit abuses by dark pools.

Proposals to Regulate HFT Speed in Obtaining Market Information

Such proposals are designed to limit any negative effects of three of the controversial practices we have focused on: electronic front running, slow market arbitrage, and exploitation of dark pool mid-point orders.

1. Would it be desirable to eliminate electronic front running? The unfairness case against electronic front running is weak. And it is unclear whether the informational advantages that HFTs obtain from electronic front running call for regulatory intervention on efficiency grounds. Based on what we know at the moment, the matter may be too close to call.

a. Actual unfairness. Electronic front running actually appears to benefit ordinary retail investors, including those who own mutual fund investments or pension funds that invest in indices and trade on exchanges. Retail investors are largely uninformed, and index investing is by definition uninformed.[49] The elimination of electronic front running would likely reduce liquidity for such investors, making uninformed trading more expensive without any gains for the uninformed traders from the increased anonymity.

b. Efficiency. Elimination of electronic front running could arguably produce efficiency gains from better capital allocation arising from increases in price accuracy. But such gains, as just noted, would come at the expense of reduced liquidity, leading to less efficient capital and risk allocation. And on balance, it is not clear that elimination would increase efficiency. Our more nuanced analysis, which considers the roles in price discovery played by announcement information traders, suggests that eliminating electronic front running would reduce, not improve, price accuracy. In terms of its effects on various kinds of informed traders, electronic front-running makes it more difficult for announcement traders to be profitable, but does not affect the profitability of fundamental value trading. Because announcement trading harms liquidity but is of little benefit from a price accuracy perspective—since real economy decision making obviously occurs on a much longer time scale than mere milliseconds—electronic front-running probably improves price accuracy.

c. Appearance of unfairness. While our analysis suggests that electronic front running does not actually result in unfairness, a substantial portion of the public still views HFT practices as unfair. Under normal circumstances, the best response to misunderstanding is education, not prohibition of an activity that does not in fact pose a problem. Still, an unfounded but persistent sense of unfairness is demoralizing: it simply makes people feel bad to think that a major social institution is corrupt. It also discourages direct and indirect ownership of equities by persons who, without this sense that something unfair was going on, would find equities to be a suitable investment vehicle. More empirical study of market confidence could make a valuable important contribution to more effective securities policymaking. If the perception of information asymmetries prevents a substantial amount of retail participation in equities, regulation designed to maintain or increase such confidence may indeed be worthwhile.

2. What happens to the case for eliminating electronic front running when slow market arbitrage and exploitation of dark pool mid-point orders are added to the analysis? These two practices both seem unquestionably undesirable. Slow market arbitrage hurts all regular traders, uninformed and informed alike, by increasing their effective cost of trading. Its economic welfare effects are unambiguously negative as well. The exploitation of dark pool mid-point orders by HFTs hurts uninformed investors and misallocates resources and risk. And even if it may be good social policy to push uninformed traders out of dark pools and onto exchanges, there are more direct ways of doing it than allowing HFTs to profit in this particular fashion.

3. What does this imply about current proposals to regulate HFTs? When evaluating measures to prevent electronic front running and other speed-based practices, we lean toward reforms that would reduce HFTs' informational advantages, provided it can be done at relatively low cost and would reduce or eliminate slow market arbitrage and mid-point order exploitation while not interfering with electronic front-running.

Consider two regulatory proposals that aim to curb high frequency quoting activity.[50] The first provides financial disincentives for high-volume quoting, such as NYSE Euronext's recent surcharge on each order above a 100:1 order-to-trade ratio. If mandated by regulation, such fees would widen spreads and reduce depth by making it harder for market makers to control adverse selection and inventory risks through their quoting strategies.

A second proposal would impose a minimum time-in-force for quotes, prohibiting them from being canceled, within, for example, 100 milliseconds of submission. But

advance the interests of both and that it is thus efficiency enhancing. This duty exists both as a matter of state common law of agency and under the rules of the Financial Industry Regulatory Authority ("FINRA").

49. It should be noted that a significant portion of retail marketable orders and index-based institutional orders execute off exchanges and in venues where the trades can be identified as largely uninformed. *See* SEC Release No. 34-68937; File No. SR-NASDAQ-2012-129, February 15, PG. 17 (2013; Rhodi Preece, *Dark Pools, Internalization, and Equity Market Quality*, CFA Institute Codes, Standards, and Position Papers 3 (2012) ("Internalization is also thought to account for almost 100% of all retail marketable order flow."). In a fully competitive market, the spreads associated with these trades should not include a significant adverse selection component. Thus, they should be unaffected by whether or not electronic front running occurs on the exchanges, where, in the absence of the practice, the spreads would be wider to reflect the greater risk that the HFTs are subjected to in dealing with informed traders. In reality, however, the spreads are barely smaller in these off exchange executions (i.e., there is only a small amount of "price improvement"). As analyzed below, why this is the case will affect the conclusion of whether wider spreads on the exchange in fact are passed on to the retail customer.

50. *See* Jones, *supra* note 31, at 42-51.

the costs of such a regulation in terms of liquidity could be substantial. It sets a floor on the length of the option offered by liquidity providers to liquidity takers, which increases their chance of being "picked off" by informed traders and so would tend to widen spreads as liquidity providers increase the cost of liquidity in response.

Another much-discussed proposal calls for replacing the current market trading structure that features continuous two-sided (i.e. buy and sell) auctions for each security with frequent batch auctions, say, every 100 milliseconds.[51] Batch auctions would consist of uniform-price, sealed-bid auctions conducted at discrete time intervals. But if frequent batch auctions have the potential to eliminate the value of minute speed advantages, their effectiveness in so doing would depend on how they are implemented. To eliminate such advantages, every exchange would have to hold its auction simultaneously. If auctions were sufficiently frequent and held at different times at each exchange, then intra-exchange exploitation of tiny speed differences could persist, including electronic front-running. We consider this an intriguing proposal, but it would be difficult to implement on a system-wide basis.

We think there is an approach to ending HFT information speed advantages that is simpler both in terms of implementation and of achieving the needed legal changes. None of these three practices would be possible if private data feeds did not make top-of-the-book quote and transaction data effectively available to some market participants before others. Thus, one potential regulatory response to the problem posed by HFT activity is to require that private dissemination of quote and trade information be delayed until the exclusive processor under the Reg. NMS scheme, referred to as the "SIP," has publicly disseminated information from all exchanges.

Rule 603(a) of Reg. NMS already prohibits exchanges from "unreasonably discriminatory" distribution of market data.[52] The SEC has interpreted this to mean that privately "distributed data could not be made available on a more timely basis [to private clients] than core data is made available to a Network processor [the SIP]… Rule 603(a) prohibits an SRO or broker-dealer from transmitting data to a vendor or user any sooner than it transmits the data to a Network processor."[53] This interpretation of the "unreasonably discriminatory" distribution language appears to permit core data information to reach HFTs more rapidly than the public recipients of the SIP as long as the signal to the HFT and the signal from the SIP went out at the same time. And the SEC, in its choice of enforcement actions, has confirmed this interpretation as well.[54] Nonetheless, the language of 603(a) is plausibly open to requiring that best quote and transaction data arrive at the same time for all traders. Such a regulation, if effectively enforced, would have the effect of limiting, though not completely eliminating, the informational advantage of HFTs. And by so doing, some of the liquidity benefits of electronic front running for uninformed traders would be preserved, while significantly reducing the ability of HFTs to conduct their slow market arbitrage and dark pool activities.

HFTs and Volatility

Overall, there is little evidence that HFT activities increase market volatility on an ongoing basis. The connection between HFTs and episodic volatility is not attributable to predatory behavior by HFTs, but rather to their rational withdrawal from the market at certain moments of stress.

There are nonetheless a number of existing proposals that address the alleged link between HFT activity and volatility. These proposals fall into two groups: one seeks to limit trading volatility generally and would incidentally affect HFTs; the second seeks to target a specific link between HFTs and volatility.

The first group includes SEC-governed single-stock circuit breakers, which impose a five-minute trading halt if the price of a specific stock moves by more than 10% within five minutes. This gives liquidity providers breathing room to consider whether order imbalances actually reflect information. Similarly, the SEC has also approved a "limit up-limit down" plan that suspends trading in a stock if transactions move more than a certain amount, often 5%, away from the security's average price over the last five minutes. These are both moderate proposals that should help limit future crashes.

The second set of proposals assumes that market makers should have stronger liquidity-providing obligations than they currently do. In the wake of the Flash Crash, exchanges have already imposed a range of affirmative obligations on institutionally identified market makers at their venues. For instance, the NYSE has "designated market makers" who have specific obligations to help maintain an orderly and continuous trading market in particular stocks. Some commentators want HFTs to have legal responsibilities resembling those of the pre-2005 NYSE specialists.

We understand the desire for liquid markets even during periods of extreme volatility. But any system that requires liquidity providers to take heavy losses during periods of extreme adverse selection must compensate them for doing

51. For example, Budish et al.'s proposal, which was endorsed by New York's Attorney General Eric Schneiderman. *See* Eric B. Budish, Peter Cramton, & John J. Shim, *The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response*, Fama-Miller Working Paper; Chicago Booth Research Paper No. 14-03 (December 23, 2013), http://ssrn.com/abstract=2388265.

52. *See* 17 C.F.R. § 242.603(a)(2). Section 11A(c)(1) of the Exchange Act authorizes the Commission to regulate market data.

53. *See* Regulation NMS, 70 Fed. Reg. 37,496, 37,567 & 37,569 (June 29, 2005) (adopting release), http://www.sec.gov/rules/final/34-51808fr.pdf.

54. In a 2012 proceeding, the SEC found that NYSE had been sending market data, including best bids and offers, to private subscribers *before* it sent that data to the SIP, and fined NYSE $5 million. *See In the Matter of New York Stock Exchange LLC, and NYSE Euronext*, Securities Exchange Act Release No. 67857, 2012 WL 4044880 (2012), http://www.sec.gov/litigation/admin/2012/34-67857.pdf.

so at other times. Determining the value of that compensation is extremely difficult, and these designated liquidity providers will be the prime targets of informed traders during crises.[55] Thus, we are skeptical about such proposals, especially because the wealth transfers and efficiency consequences of episodic volatility are not as substantial as many critics seem to believe.

Dark Pools

Our analysis suggests that the regulatory focus here should be on ensuring disclosure of whether customer orders are being routed to the venues offering best execution and whether order routing directions are being ignored. FINRA has recently requested comment on several new proposed rules promoting greater disclosure. Dark pools should provide FINRA with more extensive order book information for the Order Audit Trail System ("OATS") that helps FINRA carry out its surveillance activities. Brokers could be required to disclose what percentage of orders routed to their venue were executed there, at what price, and what instructions, if any, were associated with those orders.[56]

Certainly, more could be done to strengthen the stock market's mandatory disclosure regime. Brokers are not currently required to disclose to customers on their transaction confirmation slips the venue in which an order was executed,[57] even though such records must already be retained and would provide customers with the ability to check whether their requests were being followed. To be effective, these disclosure proposals would have to enable the SEC or private litigants to reveal inaccurate broker disclosures, and customers would need to examine and act on those disclosures. If we have reason to worry that they will not, the SEC should conduct periodic audits to verify the accuracy of these confirmations.

Conclusion

This article provides a comprehensive framework for understanding a number of controversial players and practices in the new stock market, including high frequency traders, electronic front running, and dark pool operators. We argue that the issues raised by such practices can fundamentally be understood through just two basic mechanisms—adverse selection and the principal-agent problem—as they play out in the context of a multi-venue trading system.

We briefly assess the likely effectiveness of a variety of potential reforms to current market structure. We agree, for example, that brokers should be required to disclose more information about their effectiveness in carrying out the orders of their customers, particularly those directed to dark pools. We disagree with proposals that HFTs be required to keep their quotes in force for some minimum amount of time, and with proposals aimed at generally discouraging, or even banning, trading on dark pools. These are bad ideas that seem to be based on a misunderstanding of how the market really works or of the actual social impact of a given practice. In other cases that involve complicated trade-offs, it may not be obvious whether a reform is desirable, but our framework allows for a better understanding of the tradeoff involved, and thus a more informed choice—and it may have the added benefit of pointing to where further empirical research would be useful. We find this to be the case with proposals to briefly delay providing HFTs with information about new transactions and quotation changes, so that HFTs have no advantages over other traders.

MERRITT FOX is Michael E. Patterson Professor of Law and NASDAQ Professor for the Law and Economics of Capital Markets at Columbia Law School, as well as Co-Director of the School's Center for Law and Economics and Co-Director of the Columbia Law School/Business School Program in the Law and Economics of Capital Markets.

LAWRENCE R. GLOSTEN is the S. Sloan Colt Professor of Banking and International Finance at Columbia Business School and an adjunct professor at Columbia Law School.

GABRIEL RAUTERBERG is an Assistant Professor of Law at the University of Michigan Law School. His research interests are in financial trading markets, contracts, and corporate law.

55. *See* Angel et al., *supra* note 15, at 33.

56. Economist James Angel, among others, has called for greater disclosure by brokers, suggesting that "brokerage firms themselves disclose execution quality directly to their customers." Testimony of James J. Angel, *The Role of Regulation in Shaping Equity Market Structure and Electronic Trading: Hearing Before the S. Comm. on Banking, Housing, and Urban Affairs* at 7 (2014).

57. Brokers do have limited disclosure requirements under Reg. NMS. Rule 605 requires trading venues to provide monthly reports with various measures of execution quality, and Rule 606 requires broker-dealers that route customer orders to provide quarterly reports that identify at an aggregate level the venues where client orders are executed. *See* 17 C.F.R. § 242.605-606.

APPENDIX

Section 1

Example of slow-market arbitrage

An example.

It is worth noting that the first step in this story—Lowprice's posting of the $161.13 offer on EDGE—does not guarantee that Lightning can make this profit. No marketable buy order may arrive at NYSE during the reporting gap. Also, even if one does, by the time Lightning is able to submit its marketable buy order at EDGE, some other person may already have submitted a buy marketable order to EDGE that picks off the $161.13 offer. This becomes particularly likely if, as is the case in the real world, there are a number of HFTs besides Lightning with co-location facilities at EDGE and at the other exchanges. Depending on the nature of their own respective offers posted on various exchanges, one or more of these other HFTs may be competing with Lightning to pick off the one $161.13 offer.

Who is helped and who is hurt in the example above, and what are the larger distributive consequences with slow market arbitrage as an ongoing practice?[58] In the example, the first thing to note is that Ms. Stumble, the person who, during the reporting gap, submits the marketable order that transacts against Lightning's stale $161.15 offer, is not harmed by Lightning's slow market arbitrage activities. Stumble would have suffered the same fate if Lightning had not engaged in slow market arbitrage because that course of action would have also left the $161.15 offer posted on NYSE, and so Stumble's buy order would still have transacted against it.

Still, someone must be worse off: Lightning is better off than if it had not engaged in the slow market arbitrage, and trading is a zero-sum game. To see who this worse off person may be, consider first why Lightning is better off. Lightning is in the business of buying and selling shares, not holding on to long or short positions for any significant period of time. So it needs to reverse quickly each transaction it enters. Here, it sold shares when Stumble's order transacted against Lightning's $161.15 offer on NYSE. To reverse this transaction, Lightning needed to buy shares. By engaging in slow market arbitrage, it did so by seizing the best offer in the market—Lowprice's $161.13 offer on EDGE—before others in the market even knew the offer was available. If Lightning had not detected this new offer ahead of others and seized it, Lightning's reversal of the situation would occur through posting a bid that a marketable order transacts against. We know from Part III that the sale of the shares at $161.15 and their repurchase at this newly posted bid would each, on an expected basis, be a break-even transaction. By successfully engaging in slow market arbitrage, Lightning instead made a certain $.02 profit per share sold and purchased.

To figure out who is hurt from Lightning engaging in slow market arbitrage—i.e., detecting the $161.13 offer and seizing it—consider who would have been better off if Lightning had posted a new buy limit order instead of seizing Lowprice's $161.13 offer. The person or persons helped would come from one of two groups of potential liquidity takers. One group is potential sellers who submit marketable sell orders: the posted bid that Lightning would need filled would improve the terms for the marginal seller. The other group is potential buyers who submit marketable buy orders: the opportunity by members of this group to seize Lowprice's $161.13 offer, which was better than anything else available in the market at the time, would improve terms for the marginal buyer.

Section 2

Example of HFT Exploitation of Mid-Point Orders

1. An example. Let us bring back again our HFT, Lightning. Suppose that the NBO and NBB for IBM are $161.15 and $161.11, respectively, and each are for 1,000 shares and are posted on NYSE by HFTs other than Lightning. Then the $161.15 offer is cancelled and a new 1,000 share offer is submitted at $161.12. Lightning, through its co-location facilities at NYSE, learns of these changes in advance of their being publicly reported. During the reporting gap, the official NBO remains $161.15.

Lightning knows that mid-point orders for IBM are often posted on Opaque, a well-known dark pool, and Lightning programs its algorithms accordingly. Because Opaque does not disclose what is in its limit order book, Lightning cannot know, however, whether at this moment any such orders are posted on Opaque, and, if there are, whether they are buy orders or sell orders. Still there is the potential for making money.

Using an ultra-fast connection between the co-location facility at NYSE and Opaque, a sell limit order for 1,000 shares at $161.13 is sent to Opaque with the condition attached that it cancel if it does not transact immediately (a so-called "IOC" order). This way, if there was one or more mid-point buy limit orders posted at Opaque for IBM, they will execute against Lightning's order at $161.13, half way between the now stale, but still official, NBB of $161.11 and NBO of $161.15. If there are no such mid-point buy orders posted at Opaque, nothing is lost.

Assume that there are one or more such mid-point buy orders aggregating to at least 1,000 shares and so Lightning's sell order of 1,000 shares transacts at $161.13. Lightning's co-location facility at NYSE is informed of this fact through Lightning's ultra-fast connection with Opaque. A marketable buy order for 1000 shares is sent almost instantaneously to NYSE, which transacts against the new $161.12 offer. Thus, within the short period before the new $161.12 offer on NYSE is publicly reported, Lightning has been able to execute against this offer, purchase 1,000 IBM shares at $161.12, and sell them at $161.13, for what appears to be a $10.00 profit.

58. In the example, if Lightning did not engage in slow market arbitrage, it is possible that it would be another HFT engaging in slow market arbitrage, not an ordinary trader, who would transact against the $161.13 offer. The ultimate question we are asking, however, is what would happen if no HFT engaged in the practice.