Duke Law Journal

THE NEW STOCK MARKET: SENSE AND NONSENSE

MERRITT B. FOX,† LAWRENCE R. GLOSTEN,††
AND GABRIEL V. RAUTERBERG†††

ABSTRACT

How stocks are traded in the United States has been totally transformed. Gone are the dealers on NASDAQ and the specialists at the NYSE. Instead, a company's stock can now be traded on up to sixty competing venues where a computer matches incoming orders. High-frequency traders (HFTs) post the majority of quotes and are the preponderant source of liquidity in the new market.

Many practices associated with the new stock market are highly controversial, as illustrated by the public furor following the publication of Michael Lewis's book Flash Boys*. Critics say that HFTs use their speed in discovering changes in the market and in altering their orders to take advantage of other traders. Dark pools— off-exchange trading venues that promise to keep the orders sent to them secret and to restrict the parties allowed to trade—are accused of operating in ways that injure many traders. Brokers are said to*

† Michael E. Patterson Professor of Law, NASDAQ Professor for the Law and Economics of Capital Markets, Columbia Law School.

†† S. Sloan Colt Professor of Banking and International Finance, Columbia Business School.

††† Post-Doctoral Research Scholar in the Law and Economics of Capital Markets, Columbia Law School and Columbia Business School; Yale Law School, J.D. 2009. For particularly helpful comments, we are grateful to Daniel Aisen, Kevin Haeberle, Bill Harts, Joseph McCahery, Laura Miller, Joshua Mitts, and participants at the American Law and Economics Association 2015 Annual Meeting, the European University Institute Law Faculty Seminar, the FGV Direito Rio Conference, the City Bar Association, the Massey Prize Symposium, the May 2015 workshop of the Law and Economics of Capital Markets Program, and a 2014 Columbia Law School Faculty Workshop.

mishandle customer orders in an effort to maximize the payments they receive for sending trading venues their customers' orders, rather than delivering best execution.

In this Article, we set out a simple, but powerful, conceptual framework for analyzing the new stock market. The framework is built upon three basic concepts: adverse selection, the principal-agent problem, and a multivenue trading system. We illustrate the utility of this framework by analyzing the new market's eight most controversial practices. The effects of each practice are evaluated in terms of the multiple social goals served by equity-trading markets.

We ultimately conclude that there is no emergency requiring immediate, poorly considered action. Some reforms proposed by critics, however, are clearly desirable. Other proposed reforms involve a trade-off between two or more valuable social goals. In these cases, whether a reform is desirable may be unclear, but a better understanding of the trade-off involved enables a more informed choice and suggests areas in which further empirical research would be useful. Finally, still other proposed reforms are based on misunderstandings of the market or of the social impacts of a practice and should be avoided.

TABLE OF CONTENTS

INTRODUCTION

"The United States stock market, the most iconic market in global capitalism, is rigged."[1] With this provocative statement on *60 Minutes,* Michael Lewis, best-selling author of *Flash Boys: A Wall Street Revolt*,[2] brought to the forefront of public consciousness a growing controversy concerning the way stocks are traded in the United States. Such trading has been totally transformed over the last twenty years. A truly new stock market has developed and not everyone is pleased with the results. This Article addresses these dissatisfactions and, in doing so, develops a framework for analyzing more generally the wide variety of policy issues to which the new stock market has given rise.

The various actors whose interactions make up the new stock market have come in for tremendous scrutiny. Particularly sharp criticism has been aimed at high-frequency traders (HFTs), which are said to use their speed in finding out changes in the market and in altering their own orders to take advantage of other traders in the market.[3] HFTs are believed now to participate in about half of all

---

    1. *60 Minutes: Rigged* (CBS television broadcast Mar. 30, 2014), http://www.cbsnews.com/news/is-the-us-stock-market-rigged [http://perma.cc/7HEF-GXRM].

    2. MICHAEL LEWIS, FLASH BOYS: A WALL STREET REVOLT (2014).

    3. Charles Schwab, founder of the well-known brokerage firm bearing his name, recently suggested, for example, that "[h]igh-frequency traders are gaming the system, reaping billions in the process and undermining investor confidence in the fairness of the markets . . . . It's a

trades.[4] Other features of the new stock market have been the subject of attack as well. "Dark pools" are off-exchange trading venues that promise to keep secret the existence of the orders sent to them and to restrict the kinds of parties allowed to trade.[5] Dark pools are said to often break these promises, to the disadvantage of traders sending orders to these venues.[6] A trader is also hurt if her broker fails to send her order to the trading venue where it will execute at the best price or in the most timely and reliable fashion. Critics suggest that brokers often fail in this way, sending the order instead to the venue that pays the most to the broker through practices such as "payment for order flow" or "maker fees."[7] Polls now indicate that "roughly two-thirds of Americans believe the stock market unfairly benefits some at the expense of others,"[8] a belief that some commentators think explains what has been a sharp drop in the percentage of Americans directly or indirectly owning equities.[9]

---

growing cancer and needs to be addressed." Steven Russolillo, *Schwab on HFT: 'Growing Cancer' That Must Be Addressed*, WALL ST. J.: MONEYBEAT (Apr. 3, 2014, 10:42 AM), http://blogs.wsj.com/moneybeat/2014/04/03/schwab-on-hft-growing-cancer-that-must-be-addressed [http://perma.cc/T5JE-MK7T]. These practices, and criticisms of them, are discussed in more detail in Parts V.A, V.B, and V.C.

4. *See infra* note 22 and accompanying text.

5. *See infra* Part V.E.

6. Sam Mamudi, *UBS Hit With Record Dark Pool Fine for Breaking U.S. Rules*, BLOOMBERG (Jan. 15, 2015, 12:25 PM), http://www.bloomberg.com/news/2015-01-15/sec-fines-ubs-dark-pool-more-than-14-million-for-breaking-rules.html [http://perma.cc/Y53P-G7MC] (imposing on UBS the largest fine ever given a dark pool operator); Sam Mamudi, *Dark Pools Opening Up Amid Increased Scrutiny*, BLOOMBERG BUS. WEEK (May 21, 2014, 12:01 AM), http://www.bloomberg.com/news/articles/2014-05-21/dark-pools-opening-up-amid-increased-scrutiny [http://perma.cc/9XRT-QEG7] (reporting on industry unease with dark pools). These practices, and the criticisms of them, are discussed in more detail in Part V.E.

7. *See, e.g.*, William Alden, *At Senate Hearing, Brokerage Firms Called Out for Conflicts*, N.Y. TIMES: DEALBOOK (June 17, 2014, 9:30 AM), http://dealbook.nytimes.com/2014/06/17/trader-who-called-markets-rigged-tempers-his-critique [http://perma.cc/BLA3-3FC6] (discussing scrutiny of payment for order flow at recent Congressional hearings); Editorial, *The Hidden Cost of Trading Stocks*, N.Y. TIMES (June 22, 2014), http://www.nytimes.com/2014/06/23/opinion /best-execution-and-rebates-for-brokers.html?_r=0 [http://perma.cc/C7TJ-ZGQU] (criticizing the practice of maker-taker fees). These practices, and the criticisms of them, are discussed in more detail in Parts V.F and V.G.

8. *Conflicts of Interest, Investor Loss of Confidence, and High Speed Trading in U.S. Stock Markets: Hearing Before the Permanent Subcomm. on Investigations of the S. Comm. on Homeland Sec. & Governmental Affairs*, 113th Cong. 2 (2014) (statement of Sen. Carl Levin) [hereinafter *High Speed Trading Hearings*].

9. *See* Lydia Saad, *U.S. Stock Ownership Stays at Record Low*, GALLUP (May 8, 2013), http://www.gallup.com/poll/162353/stock-ownership-stays-record-low.aspx [http://perma.cc/8H KC-DBPS] (indicating that in 2013, stock ownership among U.S. adults was at its lowest level since 1998). Lewis attributes this drop, which has occurred in the face of a sharply rising market

Actors in the nation's legal, regulatory, and political arenas have reacted rapidly to the growing furor over the new stock market. The U.S. Department of Justice, the Federal Bureau of Investigation, the Securities and Exchange Commission (SEC), and the Commodity Futures Trading Commission have all confirmed investigations into HFTs.[10] Plaintiffs' class-action lawyers have filed several civil lawsuits based on various controversial market practices.[11] The New York Attorney General has brought a high-profile lawsuit against the major investment bank Barclays, alleging it misrepresented to investors the extent to which its dark pool was free of HFT activity.[12] Several Congressional hearings have been held,[13] after which U.S. Senator Carl Levin wrote to Mary Jo White, the Chair of the SEC,

---

over the last five years, to a sense that the market is unfair. LEWIS, *supra* note 2, at 200–01; *see also* Editorial, *supra* note 7 ("There's no escaping the conclusion that the stock market is not a level playing field where all investors, large and small, have an equal shot at a fair deal.").

10.   *See* Keri Geiger & Patricia Hurtado, *FBI Seeks Help From High-Frequency Traders to Find Abuses*, BLOOMBERG (Apr. 1, 2014, 1:31 PM), http://www.bloomberg.com/news/2014-03-31/fbi-said-to-probe-high-speed-traders-over-abuse-of-information.html [http://perma.cc/WQ67-NMH4]; Sarah N. Lynch & Karen Freifeld, *SEC Chair Discusses Probes into High-Speed Trading*, REUTERS (Apr. 1, 2014, 7:13 PM), http://www.reuters.com/article/2014/04/01/us-congress-sec-highspeed-idUSBREA301RC20140401 [http://perma.cc/4N7G-UCRN]; Douwe Miedema, *U.S. Futures Regulator CFTC Probing Speed Traders*, REUTERS (Apr. 3, 2014, 3:13 PM), http://www.reuters.com/article/2014/04/03/us-hedgefunds-speed-trading-cftc-idUSBREA 321QU20140403 [http://perma.cc/Z4SE-MZL3]; Del Quentin Wilber, Keri Geiger & Patricia Hurtado, *Holder Vows High-Speed Trading Probe to Protect Markets*, BLOOMBERG (Apr. 4, 2014, 11:26 AM), http://www.bloomberg.com/news/2014-04-04/holder-vows-high-speed-trading-probe-to-protect-markets.html [http://perma.cc/A9JC-6HW9] (DOJ).

11.   *See* Amended Complaint at 1 n.1, City of Providence v. BATS Global Mkts., Inc., No. 14-cv-2811-SMF (S.D.N.Y. Sept. 2, 2014) (alleging securities fraud in a class action against every major stock exchange); Complaint at 10, Flynn v. Bank of Am. Corp., No. 14-cv-4321-JMF (S.D.N.Y. June 13, 2014) (alleging manipulation of securities markets by HFTs); Complaint at 4–5, Lanier v. BATS Exch., Inc., No. 14-cv-3745-KBF (S.D.N.Y. May 23, 2014) (alleging breach of contract based on trading venues' differential sale-of-information access to HFTs and other users); Complaint at 3–4, Harel Ins. v. BATS Global Mkts., Inc., No. 14-cv-3608-UA (S.D.N.Y. May 20, 2014) (alleging that trading venues provided HFTs material non-public information in return for kickbacks). CME Group Inc., which owns the world's largest futures market, has been sued by users accusing it of catering to HFTs. *See* Complaint at 1, Braman v. CME Group, Inc., No. 14-cv-2646 (N.D. Ill. Apr. 11, 2014).

12.   Complaint at 2–4, Schneiderman v. Barclays Capital, Inc., No. 451391/2014 (N.Y. Sup. Ct. June 25, 2014).

13.   *See, e.g.*, *High Speed Trading Hearings*, *supra* note 8; *High Frequency Trading's Impact on the Economy*: *Hearings Before the Subcomm. on Sec., Ins., & Inv. of the S. Comm. on Banking, Hous., & Urban Affairs*, 113th Cong. (2014); *Oversight of the SEC's Division of Trading and Markets: Hearing Before the Subcomm. on Capital Mkts. & Gov't Sponsored Enters. of the H. Comm. on Fin. Serv.*, 113th Cong. (2014); *The Role of Regulation in Shaping Equity Market Structure and Electronic Trading*: *Hearing Before the S. Comm. on Banking, Hous. & Urban Affairs*, 113th Cong. (2014) [hereinafter *Role of Regulation Hearings*].

demanding significant changes to market structure and the elimination of "[c]onflicts of interest [that] erode public confidence in the markets."[14]

It is time to step back and take a serious, dispassionate look at how the new stock market functions and the implications of the regulatory choices we face going forward. Legal scholars have done an able job of applying the insights of many economic theories to law. This has not been true, however, of the now well-established field of microstructure economics. Its foundational models of trading behavior in financial markets are rarely cited in legal scholarship and never discussed in depth.[15] The literature of the field itself, although empirically sophisticated, lacks a broad-scope, institutionally nuanced look at the basic dynamics shaping the modern equities market. Thus, we still lack a comprehensive framework for understanding the new stock market. The absence of such a framework acts as a serious obstacle for legislators, regulators, judges, and the public in deciding how to seriously think about regulating our markets. Much is at stake. The performance of the equities market has important effects on the efficiency with which goods and services are produced in our economy and on the real economy's rate of growth. Equities also play a vital role as a place for ordinary individuals to invest their savings. This Article brings the insights of microstructure economics to bear to provide a comprehensive framework for thinking about the new stock market. We demonstrate the usefulness of this framework by applying it to the new market's most controversial practices. Although these practices may seem completely unrelated to each other, they can all be understood through just three basic mechanisms: adverse selection, the principal-agent problem, and a multivenue trading system.[16]

We ultimately conclude that no emergency exists requiring immediate, less-than-fully-considered action. Some reforms proposed

---

14. Letter from Sen. Carl Levin to Hon. Mary Jo White, Chairman, SEC (July 9, 2014), http://www.hsgac.senate.gov/download/levin-letter-to-sec-chairman-mary-jo-white-re-equity-market-structure-july-15_2014 [http://perma.cc/WBK5-G56C].

15. *See, e.g.*, Lawrence R. Glosten & Paul R. Milgrom, *Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders*, 14 J. FIN. ECON. 71 (1985) (providing model of trading behavior under information asymmetries in securities markets); Albert S. Kyle, *Continuous Auctions and Insider Trading*, 53 ECONOMETRICA 1315 (1985) (providing model of trading behavior under information asymmetries in securities markets). The seminal models of Glosten-Milgrom and Kyle have been cited far less than 100 times in any publication covered by Westlaw's "Law Reviews & Journals" database.

16. *See infra* Parts I.D, III.

by critics, however, appear after analysis, to be unambiguously desirable. We conclude, for example, that it would be good to require brokers to pass through maker-taker fees and payment for order flow to their customers. Other proposed reforms involve a trade-off in which an improvement in terms of one worthwhile social goal can only come at a sacrifice of another such goal. In these cases, it may not be obvious whether a reform is, or is not, desirable, but a better understanding of the trade-off involved makes for a more informed choice and may point to areas in which further empirical research would be useful. We find this to be the case with, for example, proposals to briefly delay providing HFTs with information concerning new transactions and quotation changes so that HFTs have no 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 social impact of a given practice. We find this to be the case with, for example, proposals that HFTs must keep their quotes in force for some minimum amount of time and proposals aimed at generally discouraging, or even banning, trading on dark pools.

This Article proceeds as follows. Part I discusses briefly how the stock market has changed, the eight controversial practices we will analyze, and the analytic framework that will guide that analysis. Part II specifies some basic vocabulary; illustrates how, in a multivenue market, the arrival of a market order, the arrival of a limit order, and the cancellation of an already standing limit order each results in a transaction and/or changes the available quotes; and describes how information concerning the quotes and transactions on these venues is collected and disseminated. Part III considers the economics of liquidity supply in the presence of adverse selection, explaining how the most complex and important of our three factors operates.

Part IV sets out the normative criteria for evaluating the social impact of a practice or reform. Part V applies our analytic framework by analyzing each of the new stock market's eight most controversial practices and assessing the ultimate impact of each on the multiple social goals discussed in Part IV. This grand tour of current controversies also serves as an illustration of how the simple analytic framework described above can provide the key to understanding the new stock market more generally. Part VI contains our recommendations, after which we conclude.

I. The New Stock Market: Changes, Controversies, and
Approach

It is important at the outset to see how much the stock market has changed in a relatively short time and to identify the forces that have led to this change. It is useful as well to specify in more detail the most controversial practices associated with the new stock market and to lay out the basics of our approach to these practices and the operations of the new market.

A. How the Stock Market Has Changed

The stock market is an institution that connects potential buyers and sellers of companies' stocks. As recently as the early 1990s, trading in the stock of each publicly traded company of any significance was still largely confined to a *single venue*, either NASDAQ or the New York Stock Exchange (NYSE).[17] At NASDAQ, a dealer was the purchaser of every share sold by a trader and the seller of every share bought by a trader. The dealer did so at quoted prices generated through the calculation and judgment of an *individual human being*. At the NYSE, where there was an actual floor, the specialist for a stock, also a human being, often played a similar dealer role, but in addition 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 floor.[18]

Today, any given stock is potentially traded in each of almost *sixty* competing venues: eleven exchanges and almost fifty dark pools.[19] The NASDAQ dealers and the NYSE specialists are gone. Almost all of these competing trading venues are *electronic limit order books*, in which a trader can post a *limit order*, which is its firm commitment until cancelled, to buy or sell up to a specified number of shares at a quoted price.[20] A computer (the venue's *matching engine*)

---

17. *See generally* George T. Simon & Kathryn M. Trkla, *The Regulation of Specialists and Implications for the Future*, 61 Bus. Law. 217, 225–27 (2005) (explaining the history of NYSE and NASDAQ procedures).

18. *See* Lawrence E. Harris, Trading and Exchanges: Market Microstructure for Practitioners 89–111 (2003) (discussing the structure of the traditional NYSE and NASDAQ exchanges).

19. Laura Tuttle, Alternative Trading Systems: Description of ATS Trading in National Market System Stocks 5–6 (2013), http://www.sec.gov/marketstructure/research/ats_data_paper_october_2013.pdf [http://perma.cc/8J3D-JVBZ].

20. *See infra* Part II.A. For a posted sell limit order, this stated limit price is an "offer." For a posted buy limit order, this stated limit price is a "bid."

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.[21]

Today, HFTs post a majority of the limit orders that are matched in this fashion and result in executed trades.[22] An HFT uses high-speed communications to constantly update its information concerning transactions occurring in each stock that it regularly trades, as well as changes in the buy and sell limit orders posted by others on every major trading venue. This information is automatically fed into a computer that uses algorithms to change the limit prices and quantities associated with the HFT's own limit orders posted on each of the various trading venues.[23] More than three-quarters of all trades in the United States are executed on one or another of these electronic limit order book venues.[24] Most of the remaining trades involve a broker internally matching the buy and sell orders received from retail customers.[25]

B. *Forces for Change and the Role of Regulation*

This transformation to the new stock market is a product of the fantastic increases in the speed of communication and calculation that have arisen from the information-technology revolution. The new stock market's particular structure, though, is due in important part to choices made by Congress and the SEC. The initial impetus for this new market structure goes back to Congress's adoption in 1975 of the National Market System (NMS) amendments to the Securities

---

21.   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. *See infra* Part II.A.

22.   *See* Jonathan A. Brogaard, High Frequency Trading and its Impact on Market Quality 2, 11 (July 16, 2010) (unpublished manuscript), http://www.futuresindustry.org/ptg/downloads/ HFT_Trading.pdf [http://perma.cc/5KCG-JVC4] (finding based on NASDAQ data set that HFTs supply liquidity for 51 percent of all trades and provide the market quotes 50 percent of the time); *see generally* Albert J. Menkveld, *High Frequency Trading and the* New-Market *Makers*, 16 J. FIN. MKTS. 712 (2013) (exploring the role of HFTs as market makers in today's market).

23.   *See* Charles R. Korsmo, *High-Frequency Trading: A Regulatory Strategy*, 48 U. RICH. L. REV. 523, 540 (2014) (defining attributes of HFTs).

24.   LAURA TUTTLE, OTC TRADING: DESCRIPTION OF NON-ATS OTC TRADING IN NATIONAL MARKET SYSTEM STOCKS 11 (2014), https://www.sec.gov/marketstructure/research/ otc_trading_march_2014.pdf [http://perma.cc/VP6V-T6KV].

25.   For a discussion of internalization, see *infra* Parts I.C.8, V.G, and VI.D.

Exchange Act of 1934 (the Exchange Act).[26] Multiple, competing trading venues have the upside of the greater efficiency and higher rate of innovation that are likely to arise from competition. They have the possible downside that orders from potential traders are fragmented among multiple venues, which makes it less likely that willing buyers and sellers can easily find each other and transact. Congress, in its adoption of the NMS amendments, foresaw that improving information technology could significantly reduce this downside by making it easier for traders to see what is going on in each of these venues.[27] The NMS amendments pushed the system to develop in this direction, a push that has been consistently supported by the SEC.[28]

This decision favoring multiple venues is unlikely to be reversed in the foreseeable future. Data concerning the speed of trading, its cost, and the apparent amount of liquidity in the system suggest that the new stock market is a substantial improvement over what came before it.[29] Today's technology, if it instead were operating within a

---

26. Securities Acts Amendments of 1975, Pub. L. No. 94-29, 89 Stat. 97 (codified as amended in scattered sections of 15 U.S.C.) (amending the Securities Exchange Act of 1934, 15 U.S.C. §§ 78a-kk (1976 & Supp. IV 1980)).

27. Alexander P. Okuliar, *Financial Exchange Consolidation and Antitrust: Is There A Need for More Intervention?*, 28 ANTITRUST, no. 2, Spring 2014, at 66, 67 (explaining changes implemented by the SEC to satisfy Congress's vision for Regulation NMS).

28. Congress, when the NMS amendments were adopted, expected a proliferation of competing venues. It self-consciously rejected a proposal for an electronic limit order book in which all order flow was directed to a single trading venue, known as a central limit order book (CLOB). *See, e.g.*, S. Rep. No. 94-75, at 12 (1975), *as reprinted in* 1975 U.S.C.C.A.N. 179, 190 (rejecting the 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"); DIV. OF MKT. REGULATION, U.S. SEC. & EXCH. COMM'N, MARKET 2000: AN EXAMINATION OF CURRENT EQUITY MARKET DEVELOPMENTS app. III at 6 (1994) (discussing vigorous industry opposition to the SEC's proposal of a CLOB in the 1970s)*; see also* Milton H. Cohen, *The National Market System—A Modest Proposal*, 46 GEO. WASH. L. REV. 743, 774 (1978) ("But to accord ultimate and total benefit of the auction process to all orders is impossible unless that process is concentrated in one location (which Congress certainly did not set as a goal of the national market system) . . . ."); Lawrence R. Glosten, *Is the Electronic Limit Order Book Inevitable?*, 49 J. FIN. 1127, 1129 (1994) (discussing mechanics of a CLOB); Craig Pirrong, *The Thirty Years War*, 28 REGULATION, no. 2, Summer 2005, at 54, 56 (explaining that a CLOB would effectively function as a public utility and the problems attendant to that status).

29. *See* JAMES J. ANGEL, LAWRENCE E. HARRIS & CHESTER S. SPATT, EQUITY TRADING IN THE 21ST CENTURY: AN UPDATE 11–12 (2013), http://www.q-group.org/wp-content/uploads/2014/01/Equity-Trading-in-the-21st-Century-An-Update-FINAL1.pdf [http://perma.cc/H2Q7-CV3X] (showing 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); *see also* James J. Angel, Lawrence E. Harris & Chester S. Spatt, *Equity Trading in the 21st Century* 7–26

centralized single-venue system, might of course have led to even greater improvements—a possibility that is the subject of continuing debate among academic theorists[30]—but this is entirely a matter of speculation. Moreover, as a matter of political economy, any attempt to reverse the decision for multiple venues would meet stiff resistance from those who have built businesses based on an assumption that the multivenue structure will continue. So, to the extent that the criticisms of the new stock market have merit, the challenge will be to design reforms within the current multivenue system.

C. *The Eight Most Controversial Practices*

Eight practices said to occur within the new stock market have attracted particular controversy. Although they will be analyzed in detail in Parts V and VI of this Article, it is helpful to introduce them at this point, using, in each case, a simple example.

Note at the outset that each of the first three of these practices involves an HFT benefiting itself by taking advantage of having a "co-location" facility at each exchange. Co-location involves the HFT having a computer located right next to an exchange's matching engine. This arrangement allows the HFT to find out about transactions occurring on the exchange, and changes in quoted prices, sooner than other traders. It also allows the HFT to cancel old limit orders posted on the exchange, and submit new ones, very quickly. The HFT's co-location facility at each exchange is connected to its co-location facility at every other exchange by specialized fiber-optic cables, which permit extremely rapid communication among the HFT's co-location facilities at the different exchanges, all of which have their matching engines in northern New Jersey.

(Marshall Research Paper Series, Working Paper No. FBE 09-10, 2010), http://modernmarkets initiative.org/wp-content/uploads/2013/10/Equity-Trading-in-the-21st-Century.pdf [http://perma. cc/5J9P-Y796] (discussing recent technology improvements in trading systems and their impacts).

30. *See generally, e.g.*, Jean-Edouard Colliard & Thierry Foucault, *Trading Fees and Efficiency in Limit Order Markets*, 25 REV. FIN. STUD. 3389 (2012) (discussing the drawbacks of a decentralized trading system when compared with a hypothetical single system); Thierry Foucault & Albert J. Menkveld, *Competition for Order Flow and Smart Order Routing Systems*, 63 J. FIN. 119 (2008) (exploring the effects of market fragmentation on Euronext and the London Stock Exchange); *see also* Regulation NMS, 70 Fed. Reg. 37,496, 37,530 (June 29, 2005) (to be codified at 17 C.F.R. pt. 200) (announcing the adoption of rules governing the dissemination of market data); Pirrong, *supra* note 28, at 54 (discussing Regulation NMS).

1. *HFT electronic front running.* Suppose an institutional investor wishes to buy a very large number of shares of a given stock. The investor breaks the desired quantity into several smaller, but still sizable, marketable orders, each going to a different exchange. Through its co-location facility, an HFT learns of the transaction at the exchange that is reached first by the investor's orders. The HFT's algorithm infers from this information that, quite possibly, similar sizable orders are en route to other exchanges as well. The algorithm instantly sends out signals to make advantageous adjustments in the HFT's limit orders posted on these other exchanges, adjustments that are completed within the tiny interval before the institution's orders reach these other exchanges. Critics of the practice point to the fact that the institutional investor pays more for its shares than if these adjustments had not been made, while the HFT, using its information advantage, is benefited.[31]

2. *HFT slow-market arbitrage.* Suppose that on another day, the same HFT has posted on one exchange buy and sell limit orders that respectively represent the highest bid and lowest offer prices available on any exchange in the market for a particular stock. The HFT's bid and offer are reported as such on the national system for reporting what, across all the exchanges, is the best available bid and offer for the stock. Then, an institutional investor, wishing to sell a substantial quantity of this stock, posts a new limit offer on a second exchange that is lower in price than the HFT's offer. Through the HFT's co-location facilities at the second exchange, the HFT almost instantaneously observes the arrival of this new better offer. There is a short period of time before the national reporting system reflects this new better offer, during which lag the HFT has the possibility of making a certain profit. The HFT leaves standing its offer on the first exchange. During the reporting lag, this offer continues to appear to the market, based on the national reporting system, to be the lowest-price offer available. If a market buy order arrives at this first exchange before the national reporting system reflects the new better offer on the second exchange, the market order will execute against the HFT's offer on the first market at the higher price. The HFT can then repurchase the same quantity of shares on the second exchange at the lower price being offered by the institutional investor, thereby making the HFT a certain profit. Critics point out trading is a zero-

---

31. *See infra* notes 108–09 and accompanying text.

sum game and so the HFT's profit comes at the expense of the other traders in the market.[32]

3. *HFT exploitation of midpoint orders sitting in dark pools.* On yet another day, suppose an institutional trader posts a midpoint limit buy order in a dark pool. This is an order that, until cancelled, will execute against any market sell order that subsequently arrives at the dark pool and will do so at a price equal to what, at the moment of execution, is the midpoint between the best offer and best bid reported to be available on any of the exchanges by the national reporting system. The HFT from our previous examples very rapidly observes, through its co-location facilities, that the quotes have changed on one exchange such that the new best offer on that exchange is lower than the midpoint between what, until that moment, had been the best bid and best offer available on any public exchange. Again, for a short period of time, the national reporting system will not reflect the new better offer that has already been observed by the HFT. The HFT purchases shares at the new better price and then immediately sends a sell order to the dark pool, which executes against the trader's order at the midpoint between the still-official, but now-stale, best offer and best bid reported by the national system. Because the price paid for the shares by the HFT on the exchange is lower than the price at which they are sold to the trader in the dark pool, the HFT makes a profit. Critics point out again that this profit comes at the expense of the other traders in the market.[33]

4. *HFT activities leading to increased volatility and crashes.* In the same period that the new stock market was emerging, with the large role played by HFTs, there was an upsurge in the volatility of share prices. The new market has also had a number of brief crashes and breakdowns in trading, which have been attributed to HFT algorithms receiving market information that leads the HFTs to suddenly exit the market.[34]

5. *Large investment banks in their role as brokers steering orders to their own dark pools.* An institutional investor uses a large investment bank as its broker to handle a buy limit order and the

---

32. *Id.*
33. *See infra* notes 115–16 and accompanying text.
34. *See infra* notes 119–21 and accompanying text.

bank steers the institution's order to a dark pool that the bank operates. The bank's proprietary traders learn through an internal source of the existence of the institution's order. Unless cancelled, this order may sit in the dark pool until such time that the bank's proprietary traders decide it is advantageous for them to send in an order to execute against the institution's limit order (which would mean the execution is disadvantageous for the institutional investor).

6. *Large investment banks in their role as brokers ignoring customer directions to send orders to a specified venue.* Suppose the institution using this investment bank as a broker fears its order will be sent to the bank's dark pool and suffer the fate described above. The institution therefore specifies that its order be sent to an alternative venue. The bank ignores the direction and sends the order to its own dark pool anyway. Even if the trader detects that this has happened, which may be difficult to do, it may not switch brokers because it may feel tied to the large bank due to the free "soft money" research services the bank provides.[35] Market solutions to this large investment bank's violation of its duty to provide best execution for its customer may not work effectively under these circumstances.

7. *Venue "maker-taker" and "taker-maker" fees paid to brokers.* It is common for an exchange to make payments to brokers to prompt the brokers to steer certain kinds of orders in its direction and charge brokers for other kinds of orders they send to the exchange. Under the "maker-taker" model, the exchange pays for certain limit orders it receives that are ultimately executed and charges for each marketable order it receives that executes immediately against the limit orders posted on it. Under the "taker-maker" model, the venue does the opposite. Critics characterize each of these arrangements as a system of bribes. The critics argue they create incentives for brokers to direct

---

35. "Soft money" research consists of ancillary services provided to an institutional investor by a broker free of direct charge (that is, "hard money") in return for that investor directing order flow to that broker. Soft-money arrangements can be desirable from an institutional investor's perspective because it passes the cost of the soft-money research on to the client (in the form of inferior or costlier execution by the broker), rather than as part of the direct cost of the institutional investor's own services to clients. *See* OFFICE OF COMPLIANCE, INSPECTIONS & EXAMINATIONS, U.S. SEC. & EXCH. COMM'N, INSPECTION REPORT ON THE SOFT DOLLAR PRACTICES OF BROKER-DEALERS, INVESTMENT ADVISERS AND MUTUAL FUNDS 6 (1998), http://www.sec.gov/news/studies/softdolr.htm [http://perma.cc/2AL3-5CVN].

customer orders to the venue that pays the highest rebate, rather than the one that delivers best execution for the customer.[36]

8. *Payment for order flow*. For a fee, a brokerage firm may sell to another firm (an "order-execution facility") its full order flow of buy and sell market orders from a certain kind of customer, typically retail investors, who are considered "uninformed." The other firm promises to execute each purchased order at a price that is at least slightly improved over the best offer or bid available in the market at the time the order is placed. Selling order flow in this fashion essentially outsources what a large retail broker might otherwise do internally. This would be matching nearly simultaneous buy and sell orders, buying from the seller at a price slightly over the best bid in the market and selling to the buyer at a price slightly below the best offer in the market, and making the difference between the price paid and the price received as a profit. Critics characterize payment for order flow as another kind of bribe. They argue it creates an incentive for the broker to direct their customer orders to the venue that pays the highest rebate, rather than the one that most improves the prices sellers receive and buyers pay.[37]

D. The Analytic Framework

Most of the criticism of the new stock market simply consists of taking a representative single transaction involving one of these eight practices, showing the transaction benefits one party at the expense of another, and labeling the resulting transfer as "larcenous," "extractive," "predatory," or simply "unfair."[38]

Serious analysis requires digging deeper. There needs to be a consideration of the effects of each of these practices as something that occurs on a repeated basis among competing actors, taking into account the reaction of the various other participants in the market to their knowledge that the practice is transpiring. Additionally, these effects need to be evaluated in terms of their ultimate impact on the multiple social goals—the evaluative criteria discussed in detail in Part IV—that equity-trading markets are expected to serve and that form the justificatory basis for regulation when the markets fall short.

---

36.   *See infra* notes 159–61 and accompanying text.
37.   *See infra* notes 172–74 and accompanying text.
38.   *See, e.g.*, *infra* notes 96, 108 and accompanying text.

Our analytic framework for undertaking this analysis has a surprisingly simple foundation. Though each of the eight most controversial practices seems highly distinct, they can all be understood by reference to just three basic concepts:

Adverse selection. Markets benefit enormously from businesses that compete to post limit orders on venues against which marketable orders can transact, because the availability of these limit orders substantially increases liquidity. These businesses are referred to as "liquidity providers" or "market makers." A professional supplier of liquidity for an issuer's shares—today, typically an HFT—engages in both the frequent purchase and frequent sale of these shares. The liquidity supplier makes money if on average it sells the shares it buys for more than the price it paid. Its biggest problem is *adverse selection*: the possibility that the person who anonymously places an order that executes against the liquidity supplier's quote is doing so because the trader has private information about a stock's value, which is not known to most of the market or to the liquidity supplier. In such a situation, the liquidity supplier will on average lose money on the trade. To survive in a competitive market, the liquidity supplier must set its bid and offer quotes—the limits on the purchase and sell orders it posts on trading venues—aggressively enough to attract business, but not so aggressively that the money it makes from buying from, and selling to, uninformed traders is less than what it loses from engaging in such transactions with informed traders. Liquidity providers, to minimize adverse selection, work to identify which orders come from informed traders. Informed traders, in turn, work to prevent their orders from being so identified.

Principal-agent problems. Most traders are not allowed to send their orders directly to a trading venue due to the need to ensure that contracts involving the exchange of securities for cash are reliably completed. Instead, they must use a broker. The broker needs to be given a certain amount of discretion to be able to handle the order of the trader. *Principal-agent problems* arise because it is impossible to design a contract that cost-effectively assures that the broker (the agent) will act in the best interests of the trader (the principal).[39]

Multiple venues. As discussed above, Congress and the SEC consciously chose to encourage the development of *multiple*

---

39. *See generally* Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305 (1976) (providing the canonical model of a principal-agent problem).

competing venues for the trading of any given issuer's stock, rather than one centralized trading venue. Each of the eight most controversial practices is related, in one way or another, to some aspect of the system that arises from this fundamental policy choice.

The fact that the adverse-selection-driven cat-and-mouse game between liquidity suppliers and informed traders occurs within a world with multiple trading venues, combined with rapid advances in information technology, explains the new stock market's extraordinary complexity and is key to understanding the social consequences of many of its most criticized practices. This complexity, in turn, has created new scope for principal-agent problems between brokers and traders. Although the new stock market feels bewildering, the central claim of this Article is that by understanding how these three factors interact in a competitive environment, a reader can understand most of what is happening.

II. PRIMER ON THE MECHANICS OF THE NEW STOCK MARKET

In order to evaluate the critiques of the new stock market and consider what reforms might be warranted, understanding the mechanics of this market is important. Readers well versed in these mechanics should skip this discussion and move on to Part III.

A. *Vocabulary*

Before tackling the plumbing of an electronic-limit-order-based market, it is worthwhile to specify in more detail some vocabulary that helps to describe both what traders are seeking to accomplish with the orders they send to trading venues and the services that trading venues offer these traders.

1. *Quotes and depth.* Suppose at 1:59:32 PM on July 10, 2014, Maria decides she wants to buy 200 shares of Apple. She contacts her broker and discovers the best quotes for Apple: the national best bid (NBB) is $95.28 and the national best offer (NBO) is $95.29, with a depth, respectively, of 500 and 1,000 shares. In other words, according to the national reporting system, on one or more trading venues, there are one or more buyers willing to pay $95.28 per share for up to 500 shares in aggregate (but no one willing to pay more) and one or more sellers willing to provide up to 1,000 shares in aggregate for $95.29 (but no one willing to charge less).

2. *Market orders, marketable limit orders, and marketable orders.* One possibility is that Maria submits a *market buy order* for 200 shares, that is, an unconditional order to buy at whatever is the best (that is, lowest) price available. Because she places no limit on what she is willing to pay, the order will execute almost immediately. It will do so at $95.29 unless the NBO, as reported by the national reporting system, has changed by the time her order arrives at the trading venue to which it is ultimately sent.[40] If the NBO has changed by that time, the order would execute at the new NBO.

Another possibility is that Maria, knowing the current quotes, but wanting to protect herself in case the NBO moves up too much before her order can execute, places the order for 200 shares but with the caveat that she will not pay more than $95.31. In other words, Maria has submitted a *limit buy order* for 200 shares at $95.31.[41]

Given that the NBO at the time Maria sent the order, $95.29, is below—that is, at least as favorable as—Maria's $95.31 limit, we refer to her order as a *marketable buy limit order*.[42] This is because it will behave just like a market order and execute at whatever is the then-current NBO unless the NBO has changed and has moved to above $95.31 in the brief time it takes her order to arrive at the trading venue to which it is ultimately sent.[43] For this reason, we call both market orders and marketable limit orders *marketable orders*.

3. *Nonmarketable limit orders.* In contrast, a *nonmarketable buy limit order* is a buy limit order with a price limit below the NBO at the time it is sent. It is called nonmarketable because at that moment, no one in the market is willing to sell at a price this low. Similarly, a

---

40. If there are at least 200 shares available at $95.29 at the venue to which Maria's broker sends the order, the order will execute on this venue. If not, NMS Rule 611 requires that the venue have procedures in place to send all, or the unsatisfied part, of the order on to another venue where shares are available at the NBO of $95.29. *See infra* Part II.C. A *market sell order* would work the same way if she instead wished to sell 200 shares, and, unless the NBB changed, would execute almost immediately at $95.28.

41. A *limit sell order* would be a sell order with the caveat that the person placing it would not accept less than a certain price.

42. A sell limit order in which the NBB at the time the order is sent is above—that is, at least as favorable as—the order's limit is referred to as a *marketable sell limit order*.

43. Quotes in fact can move quite quickly. In the ten minutes following when Maria first contacted her broker and noted the best offer of $95.29, the offer was at one point as low as $95.28 and as high as $95.33. If, by the time the order arrives, the NBO had moved above $95.31, Maria's limit order would not execute even though it was considered "marketable" when sent. Note, however, that the order, until its expiration or cancellation, remains a commitment to buy 200 shares at $95.31 or less.

nonmarketable sell limit order is a sell limit order with a price limit above the NBB at the time it is sent, because at that moment, no one in the market is willing to buy at a price this high.

4. *Where bid and offer quotes come from.* The foregoing shows that equity-market trading venues provide a place for market participants to display a variety of different trading interests. In the market we have described, the best offer quote is $95.29 with a depth of 1,000 shares. This is the result of persons who had previously posted still-in-effect nonmarketable sell orders with a limit of $95.29 that aggregate to 1,000 shares (that is, sell orders with a limit price above the NBB, which in our example is $98.28).[44]

5. *Making and taking liquidity.* The persons who have posted these standing nonmarketable limit sell orders have provided Maria with the option to trade immediately at $95.29, an option she can exercise by sending in a marketable order. We say that these persons have provided liquidity or that they are makers of liquidity. Maria, who in our examples takes advantage of this ability to trade immediately by submitting either a market order or a marketable limit order, consumes liquidity. She is a taker of liquidity.

6. *The trade-off between taking and making liquidity.* If Maria is willing to be less aggressive, she can instead attempt to acquire her Apple shares by putting in a nonmarketable limit order to buy 200 shares at $95.28. Then, if the quotes do not change by the time her order reaches the market, she will be adding 200 more shares to the already existing NBB for 500 shares at $95.28. Thus, if she follows this less-aggressive strategy, she can be a maker, not a taker, of liquidity, even though, unlike an HFT, she is not in the business of liquidity supply.

If Maria follows this less-aggressive strategy and a sufficient number of marketable sell orders come in before her offer expires or is cancelled, Maria's limit order will execute and she will have paid a penny less per share. If a sufficient number of such sell orders do not come in, her order will fail to execute. She then runs the risk that she may still want to purchase the shares and that the offer quotes will

---

44. Similarly, the best bid quote is $95.28 with a depth of 500 shares. This is the result of persons who had previously posted still-in-effect nonmarketable buy orders with a limit of $95.28 that aggregate to 500 shares (that is, buy orders with a limit price below the NBO, which in our example is $98.29).

have moved up, in which case she will have to pay more than the $95.29 per share she would have paid had she initially submitted a marketable order. Market orders provide speed and certainty of execution. Limit orders may obtain a better price but are less certain to execute.

B. *The Mechanics of Trading on a Single Venue*

In order to understand the dynamics of a multivenue electronic-limit-order-book market and the standard approach to its depiction, it is convenient to begin the discussion by considering how trades would be depicted if we instead had just a single trading venue.[45] In the Section following, we will add the complications involved with the multivenue system we have today.

1. *Depicting the initial book.* As an example, consider the initial state of an order book for a stock such as XYZ. This book (that is, the collection of standing limit orders) can be depicted as follows:

BIDS		OFFERS	
PRICE	SHARES	PRICE	SHARES
30.48	500	30.50	700
30.46	200	30.51	300
30.45	300	30.52	400
30.44	200	30.57	400

In this case, the best offer is $30.50 with 700 shares available, and the best bid is $30.48 with 500 shares available.[46] This simple

---

45. This is the way a CLOB market would work. As already noted, though the matter is controversial, some commentators believe it would have been better if Congress had required that there be a single venue rather than pushing for the competitive, multivenue system that has in fact developed. *See supra* note 28 and accompanying text.

46. Note that the lowest offer of $30.50 is above the highest bid of $30.48. If that were not the case, then the seller and buyer should transact given that the seller would be willing to accept less than the buyer is willing to pay. Thus, under normal circumstances, the lowest offer resting on a venue should exceed the highest bid.

description of the "top of the book" is all the information the typical retail investor receives. Notice, however, that this description does not fully describe the market. For example, this simple top-of-the-book description does not reveal that investors have bid 500 additional and offered 700 additional shares within three cents of the best market quotes. In other words, information about the depth of the book beyond the best bid and offer is revealed only by a full order book. Notice also that even with the fuller depiction set out above, one cannot tell whether the book consists of seven offers of 100 shares each at $30.50 or one offer of 700 shares.

2. *Depicting a marketable limit buy order.*   Suppose that Anna decides she wishes to buy 400 shares of XYZ, but is not willing to pay more than $30.60 per share. Accordingly, she instructs her broker to submit a limit order to buy 400 shares with a limit price of $30.60. Because $30.60 is above the best offer of $30.50 and more than 400 shares are available at $30.50, her order is marketable and would transact immediately at a price of $30.50.

Regulation NMS requires that a report of the *executed transaction*—a sale of 400 shares at $30.50—be sent almost immediately to a publicly disseminated last-trade data stream that forms the national reporting system for transactions.[47] The venue is also allowed to simultaneously send the same last-trade report to persons, including co-locating HFTs, that contract with it to receive a direct feed.[48] Regulation NMS also requires that a report of the changes in the *quotes*—the reduction in the number of shares offered at $30.50 from 700 to 300—be sent to a publicly disseminated quote stream that forms the national reporting system for quotes[49] (and again, the venue may simultaneously send this information to contracting persons such as co-locating HFTs with a direct feed).[50] The new order book, reflecting this change, would appear as follows:

---

47.   *See* 17 C.F.R. § 242.601(b)(2) (2015).

48.   *Id. See infra* Part VI.A.3 (providing a fuller description and discussion of proposed reforms to this practice).

49.   *See* 17 C.F.R. § 242.602 (2015).

50.   *See, e.g.*, Concept Release on Equity Market Structure, Exchange Act Release No. 61,358, 75 Fed. Reg. 3594, 3601 (Jan. 14, 2010) (to be codified at 17 C.F.R. pt. 242).

BIDS		OFFERS	
PRICE	SHARES	PRICE	SHARES
30.48	500	30.50	**300**
30.46	200	30.51	300
30.45	300	30.52	100
30.44	200	30.57	400

3. *Depicting a nonmarketable limit buy order.* Suppose that another investor, Dave, prompted by this new state of the book, decides he wishes to buy 200 shares of XYZ, but is not willing to pay more than $30.49 per share. Accordingly, he instructs his broker to submit a limit order to buy 200 shares with a limit price of $30.49. Because $30.49 is below the best offer of $30.50, his order is nonmarketable and no transaction will occur. Instead, his limit order will be posted on the limit order book. Because he is expressing his willingness to buy at $30.49, it becomes a bid for 200 shares at this price. Because $30.49 is higher than the previous best bid, Dave's limit order becomes the new best bid, thereby reducing the spread between the best bid and the best offer by a penny (though also reducing depth at the best bid to 200 shares). The new book reflecting the posting of this new order would appear as follows:

BIDS		OFFERS	
PRICE	SHARES	PRICE	SHARES
30.49	**200**	30.50	300
30.48	500	30.51	300
30.46	200	30.52	100
30.45	300	30.57	400
30.44	200		

Information about the new state of the book would be disseminated as in the last example.

4. *Symmetry for sell orders*. We have presented the dynamics associated with buy orders: both marketable buy orders, which reduce what is available on the offer side of the book, and nonmarketable buy orders, which add to the bid side of the book. The situation is symmetric for sell orders—both marketable sell orders, which reduce what is available on the bid side of the book, and nonmarketable sell orders, which add to the offer side of the book.

C. The Mechanics of Trading on Multiple Exchanges

With multiple exchanges, the order book dynamics are similar, but the routing of the order can be much more complicated. Part of the complication comes from Regulation NMS Rule 611, which requires that a marketable sell order—regardless of the trading venue to which it is originally sent—execute at a price equal to the best bid available on any exchange in the country, and a marketable buy order at the best offer.[51] To see how this works, consider the following consolidated-limit-order book, which aggregates the quotes from all the exchanges in the country. The aggregate number of shares bid or offered at any given exchange is identified by a single letter corresponding to that exchange.

---

51.   Federal regulation requires trading venues to establish procedures reasonably designed to prevent the purchase or sale of a stock at a price inferior to the lowest offer or highest bid, respectively, which is disseminated on the national reporting system for quotations. *See* 17 C.F.R. § 242.611(a)(1) (2015) (establishing the rule); *id.* § 242.600(b) (defining relevant terms). Regulation NMS is the most important body of federal regulation governing trading in the stock market.

BIDS		OFFERS	
PRICE	SHARES	PRICE	SHARES
30.48	Q 300 P 200	30.50	Q 500 P 100 Z 100
30.46	Q 200	30.51	Q 500 D 100
30.45	Z 300	30.52	Q 400
30.44	Z 200	30.57	Q 400

Now consider Maria wanting to purchase 1,000 shares. One way to accomplish this is for her broker to send the following market buy orders: 800 shares to Q, 100 shares to P, 100 shares to Z. Assuming the quotes are still good by the time her order arrives at these respective venues, she would pay an average price of $30.503. If speed were important and there was reason to think that the order would reach Q first, it might appear to be better to send the whole 1,000-share order to Q and pay the slightly higher average price of $30.505. Because of Rule 611,[52] however, sending the whole order to Q would not have this result. Instead, Q is required to have a system that would forward 100-share orders to each of P and Z, at which shares were also available at the NBO of $30.50. These orders would execute on these venues at this price. On Q, 500 shares would execute at $30.50 and the remaining 300 at $30.51. Again, the average price would be $30.503.

The preceding discussion, however, assumes that everyone involved is instantly aware of all newly executed transactions and all changes in quotes. It also assumes that orders can be sent and cancelled instantaneously. Things become more interesting when we drop these unrealistic assumptions. Consider first how Maria's broker became aware of the quotations and how Q knows about offers on P and Z at $30.50. Traditionally, each exchange independently provided

---

52. *Id.* § 242.611(a)(1).

quotation and transaction information. As discussed, however, the 1975 NMS amendments to the Exchange Act included broad provisions for consolidating information in the U.S. stock market, which reflected the congressional vision of an electronically linked market made up of competing venues trading in the stock of the same issuers.[53] Full realization of this vision, including rules leading to the construction of the national reporting system for quotes and transactions, took thirty years, culminating in the SEC's adoption of Regulation NMS.[54]

According to the rules, a trading venue must participate in reporting plans with the SEC, which must approve these plans.[55] The plan must provide that there is a system by which the best bid and best offer quotes posted on the venue for each issuer's stock traded there is furnished to an exclusive processor. The price and size of all transactions in each of these stocks executed on the venue must also be furnished. The exclusive processor consolidates all of this information with the information the processor acquires concerning each of the same stocks from the other venues where they trade. From all this, the exclusive processor constructs a consolidated book depicting the best offer and best bid for a stock, at each of the venues at which it trades and the corresponding sizes. The exclusive processor must make this quote information, as well as price and size information concerning the latest executed transactions in the stock, available to the public on terms that are fair and reasonable.[56] At any point in time, the best bid and best offer on this consolidated book represents the official NBB and NBO.

It takes a short period of time for the national reporting system to reflect any change in the quotations on a venue or any transaction executed on that venue. Thus, the national reporting system lags slightly behind any change in what is the best bid or offer available at any venue. As we noted in our initial descriptions of electronic front running, slow-market arbitrage, and dark-pool-midpoint-order exploitation, HFTs can—through co-location, private-data feeds, and

---

53. Section 11A(c)(1)(B) provides, *inter alia*, that the SEC should "assure the prompt, accurate, reliable, and fair . . . distribution" of transaction information. Securities Acts Amendments of 1975, Pub. L. No. 94-29 (June 4, 1975), 89 Stat. 97, 115 (1975).

54. Regulation NMS, Exchange Act Release No. 51,808, 70 Fed. Reg. 37,496 (June 25, 2005).

55. 17 C.F.R. §§ 242.600(b)(21), 242.601, 242.603, 242.608 (2015) (establishing plan requirement and rules).

56. *See id.* §§ 242.601, 242.602, 242.608.

superior information-technology infrastructure—become aware of changes in offers or bids, or of newly executed transactions, before the information becomes available to the public from the exclusive processor.[57] During this brief reporting lag, they can act on this information by cancelling standing limit orders and posting new ones. Thus, by the time Maria sees the quotes depicted in the sample consolidated book above, they may no longer in fact be available. The same is true for venue Q at the time her order arrives there.

III. THE ECONOMICS OF LIQUIDITY PROVISION

A professional supplier of liquidity for an issuer's shares—today, typically an HFT posting buy and sell limit orders—engages in both the frequent purchase and frequent sale of these shares. In doing so, it stands ready to buy and sell shares up to stated amounts at stated prices. The liquidity supplier makes money if on average it sells the shares it buys for more than the price paid. It might appear that doing so is easy, even in markets with a one cent spread: buy at the bid and sell at the offer and make a half cent per share on every transaction. Do this for a billion shares and pretty soon you are talking about real money. In fact, however, it is not so easy.

The persons with whom a liquidity supplier trades generally do not reveal their identities. The possibility always exists that the person (the "trader") who places a marketable order that executes against the liquidity supplier's quote is doing so because the trader has private information not known to most of the market or to the liquidity supplier.[58] An informed trader of this kind will buy from the liquidity supplier when her private information suggests that the stock's value is above the liquidity provider's offer. And she will sell to the liquidity supplier when her private information suggests the value is below the liquidity provider's bid. In such transactions, the liquidity supplier sells at prices below the value of the stock and buys at prices above the value of the stock, not a formula for success. Despite this, the liquidity supplier, if skillful, can still make money on a net basis, because the remaining traders with whom it transacts do not possess private information and the liquidity supplier can profit on these transactions.

---

57. *See supra* Part I.C.

58. *See* HARRIS, *supra* note 18, at 158 (discussing analyses indicating that in most markets adverse selection accounts for the majority of the bid-ask spread).

A. *Kinds of Private Information and Their Sources*

There are three primary kinds of private information, which we will label *inside information*, *announcement information*, and *fundamental value information.*

1. *Inside information.* Inside information has its ultimate origins from within some institutional source. Frequently, this institutional source is the issuing company of the stock itself. This is information the institution seeks to prevent from becoming public or from being the basis of trading by others. Trading on such information is, under many circumstances, illegal under Section 10(b) of the Exchange Act and Rule 10b-5.[59] The existence of cases of successful prosecutions under these provisions shows that such information is the basis of at least some of the informed trading that occurs in the market.[60]

2. *Announcement information.* Announcement information is information that has only just been publicly revealed, for example a government statistic about the economy or a company's earnings announcement. A trader who acts on this information extremely quickly, before other traders and the liquidity suppliers themselves can react, is also an informed trader.

3. *Fundamental value information.* Fundamental value information is a superior estimate of an issuer's future cash flows based on a person gathering bits of publicly available information and analyzing that information in a sophisticated way. The traders whose trades are informed due to this kind of information include hedge funds, actively managed mutual and pension funds, nonprofit institutions, and very wealthy individuals with actively managed portfolios. Liquidity suppliers are vulnerable to trades based on these superior estimates because liquidity suppliers tend to specialize in the business of supplying liquidity. Thus, they generally do not engage in their own fundamental analysis.

B. *Adverse Selection*

Whatever the source of an informed trader's private information, the liquidity provider will be subject to adverse selection and lose

---

59. 15 U.S.C. § 78j(b) (2012); 17 C.F.R. § 240.10b-5 (2015).

60. *See, e.g.*, United States v. O'Hagan, 521 U.S. 642, 649–66 (1997) (reinstating an insider-trading conviction).

money when it buys at the bid from informed sellers or sells at the offer to informed buyers. As long as there are enough uninformed traders willing to suffer the inevitable expected trading losses of always buying at the offer and selling at the bid, however, the liquidity provider can break even.[61] The spread simply needs to be large enough between the bid and offer that the losses accrued by transacting with informed traders are offset by the profits accrued from transacting with uninformed investors.[62]

Two useful ways exist of thinking about the calculations that liquidity providers need to perform to survive in a competitive market. The first, sometimes referred to as the "accounting perspective,"[63] is based on the proposition that for a liquidity supplier to survive in business, what it loses from transacting with the informed traders must be offset by what it gains from transacting with uninformed traders. The second, sometimes referred to as the "information perspective,"[64] relates to how a liquidity supplier rationally should update its estimate of a stock's value depending on whether the next order to transact against its quotes is a buy or a sell. Each of these two perspectives leads to the same bid-ask spread.[65]

1. *The accounting perspective.* At a point in time, let P be the market's assessment of the value of a share of stock given current information. If A and B are respectively the offer and bid, and the market consists entirely of traders with no private information, then the liquidity provider's expected profits are (A – P) from buyers and (P – B) from sellers. A liquidity provider receives A from an uninformed buyer and gives up a share of stock worth, given current information, P. Similarly, the liquidity provider pays out B to an uninformed seller and receives a share of stock worth P.

Now suppose, however, that some traders in the market may be informed. An informed trader's private information leads to a

---

61. The term "break even" is used here to include an ordinary market return on capital that would be considered "profit" from an accounting perspective.

62. It is possible for the market to break down so there is no trade. The smaller the portion of trading attributable to uninformed traders, the bigger the spread needs to be to compensate for the losses from the informed traders. But the bigger the spread, the fewer the uninformed investors willing to tolerate the associated trading losses.

63. *See* HARRIS, *supra* note 18, at 320–21 (discussing the accounting and information perspectives).

64. *Id.*

65. *See generally* Glosten & Milgrom, *supra* note 15.

different, on average more accurate, appraisal of the stock's value than the market assessment of P. Informed traders will buy if their appraisal of the stock value, V, exceeds A, the offer. They will sell if their estimate of value, V, is below B, the bid. The liquidity supplier knows that if a buyer is informed, its view of V will, on average, be more accurate than the view of others in the market. Therefore the liquidity supplier rationally expects that if it unknowingly sells at A to a person who is informed, on average V is greater than A. It will similarly expect that if it unknowingly buys at B from a person who is informed, on average V is less than B. Hence, expected profits from transactions with informed buyers are negative, as are profits from transactions with informed sellers.

The final input to the calculation is the likelihood of informed and uninformed traders. On average, buy and sell orders from uninformed traders will be equal in number. If trading takes place on the basis of *positive* private information, the informed traders will submit buy orders but not sell orders.

A concrete example is useful here. Suppose that over the next short interval of time the market knows there will be an announcement. If it contains good news, the apparent value of the stock will rise from $60.00 to $61.00. If it contains bad news, the apparent value will instead fall to $59.00. To those without private information, it is equally likely that the news will be good as that it will be bad. Informed traders, however, know what the announcement will be. One percent of the order flow is expected to come from informed traders, who will buy if what they know is good and sell if what they know is bad. The uninformed are equally likely to buy or sell.

Then, the probability that a *buy* is from an informed trader is (.5 x .01) (the likelihood the information is positive multiplied by the percentage of trades that will be informed), whereas the probability that it is from an uninformed trader is (.99 x .5) (the percentage of traders that are uninformed multiplied by the even chance they will be buyers rather than sellers). Thus, for the offer quote not to be a losing proposition, it must be at least as big as the A that solves:

$$(.5 \text{ x } .01)(A - \$61.00) + (.99 \text{ x } .5)(A - \$60) = 0$$

Solving this equation, A, the offer, must be $60.01. By the same reasoning, for the bid quote not to be a losing proposition, it should

be $59.99, that is, a spread of two cents in a competitive market of liquidity suppliers.[66]

2. *The information perspective.*   The second way to view the quoting problem is the following. A liquidity supplier knows it is possible that the next marketable order that arrives will be from an informed trader. The liquidity supplier knows that if the next marketable order to arrive is a buy, it is possibly motivated by positive private information and there is no chance it is motivated by negative private information. Similarly, if the next order to arrive is a sell, it is possibly motivated by negative private information and there is no chance it is motivated by positive private information. Thus, whichever is the kind of order next to arrive, its arrival will alter the liquidity supplier's estimate of the stock's value: up if it is a buy order and down if it is a sell order. The offer and the bid are set in advance of knowing which it will be, with the offer being contingent on the arriving order being a buy and the bid on it being a sell. Thus, when a liquidity supplier is deciding on its offer price, it knows an informed trader will only accept that offer if the information was positive, and that acceptance would cause the liquidity supplier to revise its estimate upward. So, for a transaction with a buy order to be regret free, the liquidity supplier must set the offer to reflect this upward revision in advance. The same logic applies for setting the bid: to be regret free it must reflect the downward revision that would accompany the arrival of a sell order.[67]

3. *The pattern of transaction prices in the presence of informed trading.*   This second approach highlights an important characterization of rational liquidity provision in a market with private information. Liquidity suppliers will be constantly updating valuations in response to transactions. With a sufficient number of trades, the market price will come to reflect private information. The behavior of rational liquidity providers thus reflects a kind of

---

66. For expositional simplicity, we are ignoring here other determinants of the bid-ask spread, including inventory costs and the marginal costs for the personnel and facilities necessary to be in the liquidity supply business. *See generally* Ananth Madhavan, *Market Microstructure: A Survey*, 2 J. FIN. MKTS. 205, 212–23 (2008) (discussing empirical studies of various determinants and microstructure models generally). None of the other determinants undermine the conclusions in this Article that flow from this analysis.

67. Using the example in the text and applying Bayes' Rule leads to exactly the same calculation of the bid and ask as was generated by the accounting perspective. *See* Glosten & Milgrom, *supra* note 15, at 93–94.

"invisible hand": simply as a result of their efforts to avoid losses to informed traders, liquidity providers are repeatedly revising their quotes so they come to fully reflect informed traders' information, making stock prices genuinely informative. In our example, suppose the news known by the informed traders is good. Over a period of time, both marketable buy and marketable sell orders will arrive at trading venues, but there will be more buys than sells. As a result, although there will be ups and downs in the offers and bids as the estimate of value moves up and down with the arrival of each buy and sell order, the ups will predominate and the midpoint between the bid and offer will trend upward toward $61. Similarly, if the news known by the informed traders is bad, the midpoint will trend downward toward $59. Empirical evidence strongly supports the results of these adverse-selection models: analyses of intraday changes in quotes and in the prices of executed transactions consistently show that they respond to the pattern of buy and sell orders at the time.[68]

IV. The Evaluative Criteria

Now that the reader has a basic understanding of how the new stock market works, we can introduce the criteria with which we will evaluate today's most controversial market practices and which can be used more generally to analyze stock market public policy issues. The parties engaged in each of these controversial practices do so in a competitive market on a repeated basis and the other actors in the system generally take this fact into account in determining their own actions. The question then is how the existence of the practice affects the system as a whole in terms of its ultimate impact on the multiple social goals that equity-trading markets are expected to serve and that form the justificatory basis for regulation when the markets fall short.

---

68.   *See generally* Kalok Chan, Y. Peter Chung & Herb Johnson, *The Intraday Behavior of Bid-Ask Spreads for NYSE Stocks and CBOE Options*, 30 J. Fin. & Quantitative Analysis 329, 334 (1995) (suggesting that adverse selection is an important determinant of the intraday behavior of bid-ask spreads); Lawrence R. Glosten & Lawrence E. Harris, *Estimating the Components of the Bid-Ask Spread*, 21 J. Fin. Econ. 123 (1988) (testing 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).

A. Goals

A number of social goals animate discussion of secondary equity markets[69] and their regulation: (i) promoting the efficient allocation of capital so it goes to the most promising new investment projects in our economy; (ii) promoting the efficient operation of the economy's existing productive capacity; (iii) promoting the efficient allocation of resources between current and future periods so as to best satisfy the needs of firms seeking funds for real investments (trading the promise of future dollars to obtain current dollars) and the needs of savers seeking to forgo current consumption in order to enjoy future consumption (trading current dollars to obtain the promise of future dollars); (iv) promoting the efficient allocation among investors of the risks associated with holding securities so that the volatility in the cash flows generated by productive enterprises is borne by risk-averse investors in a way that generates the least disutility; (v) fostering an overall sense of fairness; (vi) economizing on the real resources society devotes to the operation of the trading markets and to the enforcement and compliance costs associated with their regulation; and (vii) fostering innovation that over time can improve the capacity of the system to serve these preceding goals.

Any particular practice in the market may, of course, have a positive impact in terms of some of these seven goals and a negative impact in terms of others. It is nevertheless desirable to structure the market for the secondary trading of equities so no unnecessary trade-offs occur—that is, so it satisfies each goal to the fullest extent possible without compromising one or more others—and to identify the nature of the remaining unavoidable trade-offs so intelligent choices can be made.

B. Market Characteristics that Impact on These Goals

The stock market's operations relate to these social goals in complex ways that result from its interacting characteristics. The two most important characteristics are share-price accuracy and liquidity. The impact of any given practice on the goals above is most easily evaluated through a two-step process, first assessing the effect of the practice on each of these two market characteristics and then

---

69. In the primary market, stocks are purchased from the company issuing those stocks, whereas in the secondary market, traders buy and sell stocks from each other. Stock exchanges are fundamentally secondary markets.

identifying the effect of the characteristic on the goals. For each of the controversial practices, we also identify the wealth transfers it predictably generates. We do this for two reasons. First, much of the criticism of the modern stock market as "unfair" seems to pivot on its perceived wealth-transfer effects. Second, because understanding how a practice affects the wealth of various market participants is essential to understanding how it affects their behavior, and consequently, liquidity and price accuracy.

1. *Share-price accuracy*. Price accuracy relates to the accuracy with which the market price of an issuer's shares predicts the issuer's future cash flows.[70] Because the price of any new share offering by a publicly traded issuer will be determined largely by the price of its already outstanding shares in the stock market, more accurate stock market prices will lead to capital being more likely to go to the issuers with the most promising new real investment projects. Share price also influences the availability of new project funding from other outside sources and the willingness of managers to use internal funds for investment, and so greater price accuracy assists the efficient allocation of capital in these other ways as well.[71]

More generally, accurate share prices help reveal managers who are performing poorly both in terms of their deployment of internal funds for new investment projects (again assisting the efficient allocation of capital), and in terms of their management of the issuer's current assets (assisting the efficient operation of the economy's existing productive capacity).[72]

Over time, more accurate share prices today also likely lead to a greater sense of fairness on the part of investors because they will experience fewer negative surprises.[73]

---

70. *See* HARRIS, *supra* note 18, at 206–14 (discussing the social benefits of accurate stock prices).

71. Merritt B. Fox, *Civil Liability and Mandatory Disclosure*, 109 COLUM. L. REV. 237, 260–64 (2009).

72. *Id.* at 258–60.

73. 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, efficient, but relatively inaccurate, prices would result in as many positive surprises as negative ones. To many investors, however, the negative surprise is likely to be more salient. So, when a negative surprise materializes, it generates a sense of grievance even though, ex ante, a positive surprise was equally likely.

2. *Liquidity*. A second characteristic is how liquid the market is. Liquidity is a multidimensional concept that relates to the size of a trade, the price at which it is accomplished, and the time it takes to accomplish the trade.[74] 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. The more liquid the market is, however, the less severe are these trade-offs.[75] For a small retail purchase or sale of stock, the spread between the NBO and the NBB is a good measure of liquidity because the trader can effect the buy or sell transaction immediately at those respective prices and, in essence, will be paying half the spread to do so. For larger orders, how much is available at prices not too inferior to the NBO or NBB (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. To start, the prospect of greater liquidity promotes more efficient allocation of resources over time. Consider this first in terms of enterprises seeking new capital to devote to real investment projects. In essence, they are purchasers of current dollars in return for the promise of future dollars. The more liquid an issuer's shares are, the more valuable their shares are to hold for any given level of expected future cash flow.[76] Thus, when an issuer offers shares in the primary market, the more liquid investors anticipate the shares will be in the future, the higher the price, all else equal, at which the issuer can sell its shares. Hence, the lower the issuer's cost of capital.[77]

In welfare economics terms, just like a tax, illiquidity results in a wedge between the value of what the savers (the purchasers of future

---

74. *See* HARRIS, *supra* note 18, at 394–410 (analyzing liquidity).

75. *Id.*

76. For a purchaser of the shares in the primary market—the sellers of current dollars in return for the promise of receiving future dollars—more liquidity means it is less costly to sell her shares in the future to provide for future consumption because the bid will be less below the midpoint between the bid and the offer. In addition, more liquidity means that buyers in the market at the time of this sale would value the shares more highly so that this midpoint will be higher. This is because it is less expensive to buy the shares in the sense that the offer will be less above the midpoint and again it will be less expensive for these buyers to sell at yet some further point in the future because the bid then will be less below the midpoint.

77. 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 & Haim Mendelson, *Asset Pricing and the Bid-Ask Spread*, 17 J. FIN. ECON. 223, 230 (1986); Yakov Amihud & Haim Mendelson, *Liquidity and Asset Prices: Financial Management Implications*, 17 FIN. MGMT. 5, 6 (1988).

dollars) expect to receive in the future and what the entrepreneurs or issuers (the suppliers of future dollars in the form of future dividend streams) expect to give up in the future.[78] As a result, illiquidity results in the less efficient allocation of resources over time. This wedge prevents certain transactions from occurring that would have occurred if the shares were expected to be more liquid. The fact that, absent this wedge, the issuer and savers would have willingly entered into these transactions means the transactions prevented by illiquidity are ones that would have made both parties better off on an expected basis. These lost transactions are projects with expected returns that are lower than the marginal project that gets funded in a world with a degree of illiquidity, but that nevertheless are high enough to make some people feel that, absent liquidity concerns, sacrificing their current dollars for the projects' promises of future ones would be worthwhile.[79]

   b. Greater share-price accuracy.   More liquidity also lowers the transaction costs associated with speculative trading based on acquiring a variety of bits of publicly available information and analyzing them to make more accurate predictions of an issuer's cash flows, that is, creating fundamental value information.[80] Thus, it stimulates such activity and in the process increases share-price accuracy, with the attendant benefits in terms of more efficient capital allocation and utilization of existing productive capacity discussed just above.

   c. More efficient allocation of risk.   Greater liquidity also promotes the more efficient allocation of risk.[81] Constant change in the world means that what constitutes an individual's optimal portfolio, in terms of diversification and of the individual's relative degree of risk aversion, is always shifting. By making both the purchase and sale of securities less expensive, greater liquidity allows the individual investor to cost-effectively adjust her portfolio over time to keep it closer at each moment to what is optimal for her.

---

   78.   *See* THIERRY FOUCAULT, MARCO PAGANO & AILSA RöELL, MARKET LIQUIDITY: THEORY, EVIDENCE, AND POLICY 307 (2013) (discussing how illiquidity acts as a wedge between transaction prices and the fundamental value of assets).

   79.   HARRIS, *supra* note 18, at 214–15.

   80.   *See id.* at 206–14 (discussing social value of liquidity).

   81.   *Id.*

V. ANALYZING THE EIGHT MOST CONTROVERSIAL NEW STOCK MARKET PRACTICES

The usefulness of our analytical framework for assessing the new stock market can be demonstrated by its application to the new stock market's eight most controversial practices.

A. *Electronic Front Running*

So-called "electronic front running" involves a situation in which an HFT, before others in the market, learns of a transaction that has occurred at one exchange and alters its quotes on other exchanges given the possibility that similar orders may still be in transit heading toward other exchanges. The HFT races ahead of these orders still on their way to the other exchanges and, before they arrive at their destinations, changes its quotes on these other exchanges.[82]

Electronic front running has been harshly criticized. For example, Charlie Munger, vice chairman of Berkshire Hathaway, has objected that high-frequency trading is "legalized front-running[,] . . . and it should never have been able to reach the size that it did."[83] Similarly, 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."[84] *Flash Boys*, published after these comments, makes electronic front running its principal focus.[85] Even more

---

82. *See supra* Part I.C.1. For fascinating empirical evidence suggesting that electronic front running occurs, see Vincent van Kervel, Liquidity: What You See Is What You Get? 2–6 (May 2012) (unpublished Ph.D. thesis, Tilburg University), http://www.rsm.nl/fileadmin/home/Department_of_Finance__VG5_/LQ5/VanKervel.pdf [http://perma.cc/RBX2-P9QP] (modeling and gathering empirical evidence that transactions on venues are followed by limit order cancellations on competing venues).

83. Sam Mamudi, *Charlie Munger: HFT is Legalized Front-Running*, BARRON'S: STOCKS TO WATCH (May 3, 2013, 1:25 PM), http://blogs.barrons.com/stockstowatchtoday/2013/05/03/charlie-munger-hft-is-legalized-front-running [http://perma.cc/NS8R-V7QN].

84. Linette Lopez, *New York's Attorney General Has Declared War on Cheating High-Frequency Traders*, BUS. INSIDER (Sept. 24, 2013, 2:41 PM), http://www.businessinsider.com/schneiderman-targets-hft-front-running-2013-9 [http://perma.cc/EX85-TCH3].

85. *See, e.g.*, LEWIS, *supra* note 2, at 108, 126, 172. Since the publication of the book in the spring of 2014, a host of other commentators have chimed in with criticism equating electronic front running with traditional illegal front running. *See, e.g.*, Ellen Brown, *Computerized Front Running: Another Goldman-Dominated Fraud*, HUFFINGTON POST: THE BLOG (June 26, 2010, 5:12 AM), http://www.huffingtonpost.com/ellen-brown/computerized-front-runnin_b_548148.html [http://perma.cc/X379-8R3P] (stating that electronic front running contributes to "the manipulation of markets for economic and political ends"); Gene Marcial, *High-Frequency*

recently, a prominent class-action suit was filed against all of the nation's exchanges that, in support of its claim of fraud, includes allegations that the exchanges cooperated with HFTs in facilitating electronic front running.[86]

Substantively, all these criticisms focus on the fact that when an HFT engages in an act of electronic front running, 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, relates to a situation involving a customer giving her broker an order to handle. Then the broker, which has a legal duty to its customer not to use knowledge of its customer's order to its own advantage, breaches this duty by engaging in a trade on its own behalf that executes ahead of the customer's order.[87] In contrast, when an HFT engages in the practice labeled as "electronic front running," it has no preexisting relationship with the trader placing the order that the HFT detects and so no relationship between the two could give rise to a duty on the part of the HFT akin to what a broker owes its customer. The matter of whether rules should prevent HFTs from engaging in this practice is of course an appropriate issue for policy analysis, as is

---

Trading Mainly Hurts the Traders and Short-Term Investors, FORBES (Apr. 2, 2014, 4:06 PM), http://www.forbes.com/sites/genemarcial/2014/04/02/high-frequency-trading-mainly-hurts-the-traders-and-short-term-investors [http://perma.cc/4DTH-SN7W] (stating that "front-running on Wall Street, which is what high-frequency trading is all about and what it really intends to be, is old news," and arguing that only the speed with which HFTs front run other investors is new).

86. Amended Complaint at 26, City of Providence v. BATS Global Mkts., Inc., No. 14-cv-2811-SMF (S.D.N.Y. Sept. 2, 2014).

87. Traditional front running is prohibited under the common law, federal law, and industry self-regulatory standards. *See* Opper v. Hancock Sec. Corp., 250 F. Supp. 668, 676 (S.D.N.Y. 1966), *aff'd*, 367 F.2d 157 (2d Cir. 1966) (holding front running to be illegal under principles of agency and federal law). The SEC prosecutes front running as a violation of Section 10(b) of the Exchange Act and Rule 10b-(5). *See, e.g.*, Complaint at 13, SEC v. Bergin, 2015 WL 4275509 (N.D. Tex. May 23, 2013) (No. 3:13-cv-1940), 2013 WL 2400793 (charging trader for front running clients under Section 10(b)); *see also* Concept Release on Equity Market Structure, Exchange Act Release No. 61,358, 75 Fed. Reg. 3594, 3609 (proposed Jan. 14, 2010) (to be codified at 17 C.F.R. pt. 242) (discussing the illegality of front running). FINRA Rule 5270(a) prohibits trading in a security if an individual has "material, non-public market information concerning [a customer's] imminent block transaction in that security," until that block transaction has become public. *FINRA Rules*, FIN. INDUS. REGULATORY AUTH., at Rule 5270 [hereinafter FINRA Rules], http://finra.complinet.com/en/display/display_viewall.html?rbid=2403&element_id=607&record_id=609&filtered_tag= [http://perma.cc/78SE-RWUZ] (last amended Sept. 3, 2013).

being undertaken here. It is important to note, however, that the practice involves no breach of duty or mutually-agreed-upon terms between contracting parties, nor does it involve any obvious breach by HFTs of the federal anti-fraud laws.

1. *An example.* We will examine the practice of electronic front running through use of an example. For simplicity of exposition, just one HFT, Lightning, and two exchanges, BATS Y and NASDAQ, are involved. Lightning has co-location facilities at the respective locations of the BATS Y and NASDAQ matching engines. A high-speed fiber-optic cable connects these co-location facilities with each other.[88]

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 $148.00, with 10,000 shares being offered at this price on BATS Y and 35,000 shares at this price on NASDAQ. 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.[89] The 35,000 shares offered at $148.00 on NASDAQ 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 NASDAQ, which in turn has the ability to cancel Lightning's 35,000 share $148.00 limit sell order posted on NASDAQ. All this can happen so fast that the cancellation would occur before the arrival there of Smartmoney's market buy order. If

---

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

89. This example fleshes out the story by Michael Lewis of how electronic front running could occur with Amgen stock in such a situation. LEWIS, *supra* note 2, 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.

Lightning does cancel in this fashion, it has engaged in electronic front running.

Why might Lightning wish to cancel its sell limit order on NASDAQ? One possibility is that given its inference that a large market buy order is likely soon to arrive at NASDAQ, 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 at $148.02. 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 the NASDAQ limit order book has room so that the $148.02 offer price is still more attractive to potential buyers than any other offers with respect to Amgen already posted on NASDAQ. Suppose, for example, that prior to Lightning's cancellation, the next best offer on NASDAQ was 15,000 shares at $148.01 and the best offer after that was 20,000 shares at $148.02. 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 $148.02 offer by Lightning.

This cautionary note, though, hides a more critical point: Lightning may wish to cancel its $148.00 sell limit order even if in fact the book contains no room 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 money it makes by buying from, and selling to, uninformed traders is less than what it loses by engaging in such transactions with informed traders. At the time it posted its sell limit order, Lightning calculated $148.00 as 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 great enough that $148.00 is no longer the optimal price at which to offer to sell shares. In that case, Lightning will be better off cancelling its $148.00 limit offer on NASDAQ.

2. *Wealth-transfer considerations*. To see the distributive effects of electronic front running, we need to consider how the world would differ if the practice were eliminated. As a first cut for this discussion of the practice's wealth effects, and for the discussion below of its efficiency effects, we will make the assumption, later relaxed, that only three kinds of market participants exist: HFTs, informed traders who trade on the basis of fundamental value information, and uninformed traders.[90]

a. Electronic front running narrows spreads. As the analysis of the example makes clear, the practice of electronic front running by HFTs makes orders by large purchasers and sellers somewhat less anonymous in the sense that the practice 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. The primary cost of being a liquidity supplier—the losses incurred from dealing with informed traders—would therefore go up. Accordingly, HFTs would increase their initially posted spreads to compensate.

Going back to our example, if Lightning were not able to electronically front run, it might have initially posted its limit sell order for 35,000 shares at $148.01 instead of $148.00. For the same reasons, it would also have a lower bid. So if, with electronic front running being allowed, its bid would have been $147.96, without the practice its bid might instead have been $147.95.

b. Electronic front running helps uninformed investors and hurts informed investors. If electronic front running were eliminated, uninformed traders and informed traders will each suffer from the resulting larger spreads—the higher offers and lower bids—because for both it will be more expensive 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, the starting point again is the fact that the elimination of electronic front running would make it more difficult

---

90. *See supra* Part III.A for a discussion of the different types of private information. We will revisit this discussion later with a more nuanced analysis that focuses on the fortunes of each of the three kinds of informed traders: fundamental value information traders, announcement traders, and inside-information traders. *See infra* Part V.A.4.

for HFTs to detect indications of possible informed orders, and so more informed trades would execute against their quotes. Trading is a zero-sum game. Thus, if HFTs did not increase their spreads in response to the end of the practice, the gains enjoyed by informed investors would just equal the increased losses suffered by HFTs. In fact, however, if electronic front running is eliminated, then HFTs will increase their spreads. They will do so by an amount just sufficient to cover what these losses would otherwise be.[91] This is because, as we learned in Part III, the economic pressures on HFTs operating in a competitive market require them to set their spreads at a level such that they just break even.

The increased spreads will be borne by all traders, informed and uninformed alike, because the HFTs cannot condition their exchange-posted limit orders on the identity of the person who sends the market orders against which their limit orders execute.[92] This means 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, without electronic front running, HFTs would find it harder to detect indications of possible trading on private information and as a result would increase their spreads. Informed traders would get all of the gains from being better able to hide the informed nature of their trades. But they pay, through the increased spreads, only part of the added costs incurred by HFTs as a result of entering into more losing transactions. The rest of these added costs are borne by uninformed investors, who receive no such benefit. So, electronic

---

91. For reasons of expository simplicity, this statement assumes that the increase in spreads would not decrease the volume of trading. Therefore, it is assumed that the increase in the absolute number of HFT trades with informed traders that would occur from the elimination of electronic front running would be the same with the increase in spreads as without. 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. This simplification does not alter the basic logic of the analysis in the text, however.

92. Regulation NMS Rule 610(a) precludes exchanges from restricting access to trading on their facilities. *See* 15 U.S.C. § 78f(b)(2) (2012) (providing that "the rules of [a registered] exchange [must] provide that any registered broker or dealer . . . may become a member of such exchange"); 17 C.F.R. § 242.610(a) (2015) (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).

front running benefits uninformed investors and harms informed ones.

  c. The ultimate incidence of electronic front running. Electronic front running has been regularly attacked as harming "ordinary investors."[93] Our analysis, however, suggests that this attack is unmerited. To start, consider retail investors, the paradigmatic ordinary investors. Retail investors generally lack any significant private information and hence are properly assumed to be uninformed. Uninformed investors, as we have just seen, 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: their purchases and sales are not prompted by any kind of private information. Rather they purchase 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, helps, not hurts, these funds and derivatively their ordinary investors.

  What, though, about people who invest in managed mutual or pension funds? They too are presumably mostly ordinary persons.[94] These entities do fundamental value research and thus have the potential of being informed investors. The analysis above suggests that electronic front running hurts informed investors. Though these funds can be expected to enjoy gains from the elimination of electronic front running, these gains might well not be passed on to the ordinary people who invest in them. The investment industry and those who work in it each appear to operate in fairly competitive markets. To the extent that these markets are in fact competitive, much of whatever above-market returns are generated by these institutions' informed trading will be captured in the form of higher

---

  93. *See, e.g.*, LEWIS, *supra* note 2, at 104; *see also* Amended Complaint at 93–95, City of Providence v. BATS Global Mkts., Inc., No. 14-cv-2811-SMF (S.D.N.Y. Sept. 2, 2014).

  94. *See* INV. CO. INST., 2013 INVESTMENT COMPANY FACT BOOK 90 (53d ed. 2013), http://www.ici.org/pdf/2013_factbook.pdf [http://perma.cc/KWT6-JAN4] (stating that mutual funds are primarily owned by individual investors). Indeed, it appears to be these particular ordinary investors that Michael Lewis has in mind when arguing that electronic front running takes money from ordinary folks on Main Street and gives it to HFTs. *See* LEWIS, *supra* note 2, at 81, 102, 108, 172.

fees or salaries for the professionals who make the actual investment decisions.[95] This suggests that any gains in these entities' trading returns that might result from the elimination of electronic front running are likely to go primarily to increase the fees and salaries of the professionals who make the actual investment decisions, not to the ordinary persons on whose behalf they trade. So even these ordinary investors likely are not hurt by electronic front running.

The beneficiaries of electronic front running, according to the critics of the practice, are the exchanges and the HFTs themselves.[96] Here, the critics 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.[97] If electronic front running were eliminated tomorrow, HFT co-location facilities would be worth less to the HFTs and they might consequently not be willing to pay as much in fees. This might reduce the rents collected by the exchanges. Any such reduction in rents would hurt the exchanges, at least in the short run.[98] The exchange business, however, has become much more competitive than in the past, making the exchanges' longer-run ability to collect rents questionable.[99] In the 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, to the extent that the exchange business has in fact become competitive, any revenues lost from co-location fees would eventually need to be made up through higher charges to investors who trade on the exchange.

---

95. Jonathan B. Berk & Richard C. Green, *Mutual Fund Flows and Performance in Rational Markets*, 112 J. POL. ECON. 1269, 1271 (2004).

96. *See, e.g.*, LEWIS, *supra* note 2, at 126, 176; *see also* Amended Complaint at 6, 26, 93–97, *City of Providence*, No. 14-cv-2811-SMF (S.D.N.Y. Sept. 2, 2014).

97. *See, e.g.*, N.Y. STOCK EXCHANGE, NYSE PRICE LIST 2015, at 18–20, https://www.nyse.com/publicdocs/nyse/markets/nyse/NYSE_Price_List.pdf [http://perma.cc/GGB9-VYSU] (defining co-location fees).

98. The impact of eliminating any of these practices, however, is uncertain because HFTs desire co-location for a number of reasons. *See* Charles M. Jones, *What Do We Know About High-Frequency Trading* 10, 26 (Columbia Bus. Sch. Research Paper, Paper No. 13-11, 2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2236201 [http://perma.cc/EC79-K4UR] (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). If sophisticated investors extensively use co-location and smart routers, it could also reduce, or possibly eliminate, the incidence of electronic front-running.

99. Nu Ri Jung, *The Present and Future of the Financial Services Industry: Convergence, Consolidation, Conglomeration, and Collaboration*, 29 QUINNIPIAC L. REV. 729, 740 (2011) (describing competitive pressures driving changes in stock exchanges).

Similarly, the lower volume of HFT business that would result from the elimination of electronic front running would reduce the profits of firms now in the HFT business and thus lower the value of their existing assets. But in the longer-run future, investments in the industry can be expected to earn a competitive return, with or without the practice. Persons with abilities and skills that are uniquely valuable to the business of HFTs would, however, suffer both a short- and longer-term diminution in their wealth positions from its elimination.

3. *Efficiency considerations.* 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 enhancing and one efficiency diminishing.

a. *Improved share price accuracy.* We have just seen that informed traders would be net gainers from the elimination of electronic front running. Their cost of trading would go up from the increase in spreads, but this would be more than compensated for by the more advantageous trades they can make with HFTs because of the reduced ability of HFTs to detect indications of possible informed trading. In the simplified world we are analyzing in this first cut at the problem, the only informed traders are persons who trade on fundamental value information. These are the speculative investors who make money by searching out bits of publicly available information, analyzing what has been gathered in a sophisticated way, and coming up with a superior estimate of a share issuer's future cash flows than is implied by the current market price of its shares. Hedge funds, actively managed mutual funds, pension funds, and endowments of nonprofits are examples of such informed traders.

Because these informed traders buy when their superior estimate of share value suggests that a stock is underpriced and sell when it indicates a stock is overpriced, their activities make share prices more accurate. The elimination of electronic front running would make it more profitable for these traders to engage in their activity and so they will do more of it. As a result, prices will be more accurate. As we have seen, more accurate prices benefit the economy by helping to allocate the economy's scarce capital to the most promising potential real investment projects and by improving the utilization of the economy's existing productive capacity through optimizing the signals provided to management about investment decisions and the signals

given to boards and shareholders about the quality of management decisions.[100]

   b. *Reduced resources going to HFT activities*.   The second positive effect from eliminating electronic front running relates to the productive resources that are currently being devoted to undertaking the practice, including the skills and abilities of highly sophisticated technical personnel, advanced computers, and fiber-optic networks. Though HFTs are notoriously secretive, HFT Virtu Financial, Inc. (Virtu) did make certain public disclosures in the run up to its now-postponed IPO.[101]   In 2013 alone, Virtu reported spending approximately $65 million on communication and data processing and $78 million on employee compensation and payroll taxes.[102] Because Virtu has only 151 employees, this means they pay an average salary of about $517,000.[103] Virtu is just one of several large HFTs and there are many smaller ones as well.[104] With infrastructure and human capital no longer needed to support electronic front running, they would be freed up to increase other productive activities in the economy.

   c. *Allocation of resources over time and allocation of risk*.   The elimination of electronic front running, by widening spreads, would make the market for equities less liquid. This has an unambiguously negative effect on the efficient allocation of resources over time. As we have seen, the prospect that an issuer's shares will have less

---

   100.   *See supra* Part IV.B.1. Implicit in this analysis is that the improvements in the real economy from more accurate prices in terms of better capital allocation and better utilization of the economy's existing productive capacity are greater than the value of the additional real resources that are brought to the task of gathering and analyzing bits of publicly available information. Ample empirical evidence suggests that accurate price signals do in fact have efficiency-enhancing effects on managerial decisions. *See* FOUCAULT ET AL., *supra* note 78, at 361–68 (collecting relevant empirical studies). Theory also suggests that accurate financial information will often be underproduced due to its status as a public good. *See, e.g.*, Alex Edmans, Itay Goldstein & Wei Jiang, *The Real Effects of Financial Markets: The Impact of Prices on Takeovers*, 67 J. FIN. 933, 938 n.6 (2012).

   101.   *See* Virtu Financial, Inc., Registration Statement Under the Securities Act of 1933 (Form S-1) (Mar. 10, 2014), https://www.sec.gov/Archives/edgar/data/1592386/000104746914 002070/a2218589zs-1.htm [http://perma.cc/F38S-CC2D].

   102.   *Id.* at 73.

   103.   *Id.* at 73, 105.

   104.   John McCrank, *Exclusive: SEC Targets 10 Firms in High Frequency Trading Probe— SEC Document*, REUTERS (July 17, 2014, 5:19 PM), http://www.reuters.com/article/2014/07/17/ us-sec-investigation-highfrequencytradin-idUSKBN0FM2TW20140717   [http://perma.cc/67AZ-HX5D] (discussing some of the largest high-frequency trading firms).

liquidity in the secondary trading market increases the issuer's cost of capital. Just like a tax, illiquidity results in a wedge between the value of what the savers (the purchasers of future dollars) expect to receive in the future and what the entrepreneurs or issuers (the suppliers of future dollars in the form of future dividend streams) expect to give up in the future. This blocks transactions that both parties would otherwise have found advantageous if the market for the stock was expected to be more liquid, and hence diminishes economic welfare.[105]

Less liquidity would similarly have an unambiguously negative effect on the efficient allocation of risk. The greater transaction costs deter each investor from adjusting as finely her portfolio when circumstances alter what would be optimal in terms of diversification and suitability to her risk preferences.

4. *Taking other kinds of informed traders into account.* The preceding discussion assumes that the only informed traders are ones trading on the basis of fundamental value information. In fact, we know of two other types of private information that can give a trader a significant advantage: announcement information and inside information. Taking account of these additional kinds of private information does not change the conclusions above that electronic front running has positive effects on uninformed investors as well as on the efficiency with which risk is allocated and resources are allocated over time, nor does it alter the fact that electronic front running has negative social effects in terms of the real resources it consumes. It also does not change the conclusion that it has a negative impact on informed traders as a group. But, depending on one's assessment of the parameters involved, taking account of these additional kinds of private information may well change the conclusion above concerning the impact of electronic front running on fundamental-value-information traders and hence of the practice's impact on price accuracy.

The issue is as follows. Suppose that electronic front running is much more helpful at enabling liquidity suppliers to respond to trading based on announcement information and inside information than to trading based on fundamental value information. Suppose as well that trading on the basis of these other two kinds of information constitutes a large portion of all informed trading. Then traders informed by these other two types of information will enjoy most of

---

105.   *See supra* Part IV.B.2.

the increased trading gains from the elimination of the practice. If electronic front running were eliminated, HFTs would need to increase spreads sufficiently to cover their corresponding increased trading losses, most of which would be due to inside-information and announcement-information traders. With the elimination of electronic front running, fundamental-value-information traders will thus have to pay as much extra per trade from the increased spread as traders on the other two kinds of private information, but will only get 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.

A key factor in determining the likelihood of this possibility is the susceptibility of fundamental-value-information trading to detection by electronic front running relative to that of trading on the basis of the other two kinds of private information. Announcement-information traders are particularly susceptible because they need to do all of their trading in a very short period of time. They therefore need to engage in larger transactions, which are easier for HFTs to detect and to which it is easier for HFTs to react. Fundamental-value traders, in contrast, may often have days to complete their planned purchases or sales and can break the total amount they wish to transact into small packets that look more like the trades of uninformed traders. But we would need to know much more to resolve the question definitively. Existing empirical research is not very enlightening concerning several other important factors: the proportion of informed trades based on each of the three kinds of private information, the average value of the information associated with each, and the exact sensitivity of the trading patterns associated with each of the three kinds of informed traders to detection by electronic front running.

If further empirical research ultimately suggests that electronic front running actually helps, not hurts, fundamental-value-information trading, it would suggest that the practice, contrary to our earlier analysis, actually helps share-price accuracy by making the business of fundamental-value-information trading more rewarding. In contrast, announcement-information trading is not important in terms of the social benefits that are derived from share-price accuracy because the information will be reflected in price very quickly even

without the trading. Inside-information trading is likely not socially useful either.[106]

B. *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.[107]

Slow-market arbitrage was a target of criticism in *Flash Boys*,[108] which in turn reflected growing discontent among commentators in the years preceding the book's publication.[109] The practice has also formed a basis for litigation. For example, in the City of Providence's class action against all the exchanges for their cooperation with HFTs, the complaint alleges that HFTs engaged in slow-market arbitrage

---

106. Persons trading on the basis of confidential nonpublic information neither worked to develop the information, nor paid someone else to work to develop it. Whether these trades are legal or not depends on the circumstances, but legality aside, the gain the trader enjoys at the expense of other investors would be hard to justify as representing a socially useful incentive. Such information usually becomes public relatively quickly and thus would have been reflected in price soon anyway. Yet, as we have seen, the existence of the practice of trading on its basis decreases liquidity, which discourages the activities of those who trade on the basis of information that does take work to develop. So, on a net basis, trading on the basis of nonpublic confidential information that took no work to develop is, if anything, likely to be socially harmful. *See* Michael J. Fishman & Kathleen M. Hagerty, *Insider Trading and the Efficiency of Stock Pric*es, 23 RAND J. ECON. 106, 110 (1992); Zohar Goshen & Gideon Parchomovsky, *On Insider Trading, Markets, and "Negative" Property Rights in Information*, 87 VA. L. REV. 1229, 1238–43 (2001).

107. *See, e.g.*, LEWIS, *supra* note 2, at 172 (depicting an example of putative slow-market arbitrage).

108. *Id.*

109. *See, e.g.*, Tyler Durden, *"Do It Yourself" Latency Arbitrage: How HFTs Can Manipulate the NBBO at Whim Courtesy of NYSE Empty Quote Gluts*, ZEROHEDGE (Aug. 23, 2010, 9:29 AM), http://www.zerohedge.com/article/do-it-yourself-latency-arbitrage-how-hfts-can-manipulate-nbbo-whim-courtesy-nyse-quote-stuff [http://perma.cc/G858-EM6F]; *Latency On Demand?*, NANEX (Aug. 23, 2010), http://www.nanex.net/FlashCrash/FlashCrashAnalysis_LOD.html [http://perma.cc/AD2Y-3Q9G] (discussing discrepancies between NYSE quotes in the public quotation system and its private feeds and the potentially manipulative gaming of those feeds by HFTs).

and claims the practice "generate[d] billions of dollars more a year in illicit profits than front-running."[110]

1. *An example.* To understand the practice in more detail, let us return to our HFT Lightning. Suppose that 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. 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 transacts 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 $20.00 profit.

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.

---

110. Amended Complaint at 6, City of Providence v. BATS Global Mkts., Inc., No. 14-cv-2811-SMF (S.D.N.Y. Sept. 2, 2014).

2. *Wealth-transfer effects.* 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? 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 might 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—that is, 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.

The results from this example can be generalized. The persons who are hurt by HFTs engaging in the practice of slow-market arbitrage on an ongoing basis are regular traders, both informed and uninformed ones.[111] In contrast to electronic front running, in which the practice 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.

3. *Efficiency considerations.*   In most situations, arbitrage activities, at least if they do not consume any real resources, have positive economic-welfare effects. The actions of arbitrageurs equilibrate prices in two markets, each of which has its own group of potential participants, and as a result, presumptively welfare-enhancing transactions are entered into that otherwise would not have occurred.[112] However, as the example shows, 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 instead is the only gainer from the so-called arbitrage activity. Regular traders, both informed and uninformed, are in fact losers because their cost of trading goes up. So the normal presumption in favor of activities carrying the label "arbitrage" does not apply here.

In fact, even if slow-market arbitrage consumed no real resources, it would have an unambiguously negative impact on welfare. Consider first the effect of the increased effective cost of trading for informed traders. Slow-market arbitrage, by raising the effective cost of trading for informed traders, makes it less rewarding

---

111.   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.

112.   The arbitrageurs do this by buying in the low-price market, thereby putting upward pressure on the price there, and selling the same item in the high-price market, thereby putting downward pressure on the price there. As a result, there is a group of potential sellers in the initially low-price market who would not find it worthwhile to sell at the initial lower price, but who do find it worthwhile to sell at the higher equilibrating price. And there is a group of potential buyers in the initially high-price market that are in the exact mirror image situation. The transactions entered into by these two groups of people, which would not occur but for arbitrage, are presumptively welfare enhancing because they are entered into voluntarily by both parties to each of the transactions.

to seek out bits of publicly available information and to analyze their implications in a sophisticated way. This reduces share-price accuracy, which, as we have seen, would in turn have negative effects on the allocation of capital for new real investment projects and the efficient utilization of existing productive capacity. As for the increased effective cost of trading on uninformed traders, it has the now-familiar negative effects on the efficient allocation of resources over time and on the efficient allocation of risk.[113]

Slow-market arbitrage in fact does consume real resources, which is another efficiency consideration. If it were the only HFT practice dependent on co-location facilities and ultra-fast connections, it would use substantial amounts of real resources that could otherwise be usefully employed increasing the production of other goods and services. If HFTs were to continue the practice of electronic front running, however, the marginal cost in real resources of engaging in slow-market arbitrage as well is probably fairly low.

C. *HFT Exploitation of Midpoint Orders*

A trader will often submit to a dark pool a "midpoint" limit buy or sell order, the terms of which are that it will execute against the next marketable order with the opposite interest to arrive at the pool and will do so at a price equal to the midpoint between the best publicly reported bid and offer at the time of execution.[114] Midpoint orders appear to have the advantage of allowing a buyer to buy at well below the best offer and sell well above the best bid. It has been noted for a number of years, however, that traders who post such orders are vulnerable to the activities of HFTs,[115] a point that was also picked up in *Flash Boys*.[116] Midpoint-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

---

113. *See supra* IV.B.2.

114. Robert P. Bartlett, III & Justin McCrary, Dark Trading at the Midpoint: Pricing Rules, Order Flow and Price Discovery 5 (Feb. 12, 2015) (unpublished manuscript), http://www.stern. nyu.edu/sites/default/files/assets/documents/2%20Bartlett%20and%20McCrary%20Shall%20 We%20Haggle.pdf [http://perma.cc/87G8-DFVB].

115. *See, e.g.*, Alex Paley, *Navigating the Dark Pool Landscape*, *in* ACCESSING FRAGMENTED LIQUIDITY 46 (2010), https://autobahn.db.com/microSite/docs/Navigating_Dark_ Pool_Landscape.pdf [http://perma.cc/8LC3-NVTA].

116. LEWIS, *supra* note 2, at 113–18.

pool that contains midpoint limit orders with the opposite interest that transact at a price equal to the midpoint between the now-stale best publicly reported bid and offer.[117]

1. *An example.* Let us bring back again our HFT, Lightning. Suppose that the NBB and NBO for IBM are $161.11 and $161.15, 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 midpoint 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 Lightning has a chance to make 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 at least one midpoint buy limit order posted at Opaque for IBM, it will execute against Lightning's order at $161.13, halfway between the now-stale, but still official, NBB of $161.11 and NBO of $161.15. If there are no such midpoint buy orders posted at Opaque, nothing is lost.

Assume that at least one such midpoint buy order exists aggregating to at least 1,000 shares so that 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 1,000 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.

---

117. *See supra* Part I.C.3.

2. *Wealth-transfer and efficiency considerations.* The regular practice of HFT exploitation of dark pool midpoint orders provides rents to HFTs—they are able to make profitable trades they otherwise would not be able to do. This activity yields no prosocial incentive effect because it is unrelated to the positive social function we have attributed to HFTs: providing liquidity in a world with both uninformed and informed traders. Because trading is a zero-sum game, if the HFTs gain, certain regular traders must lose. Because of the practice, the expected cost of trading using midpoint orders in a dark pool goes up. This would hurt those who are deterred from using dark pool midpoint orders because of this higher cost of trading, as well as those who still do use them but have to incur these higher costs.

The efficiency effects of the practice closely resemble the efficiency effects of the abuses by investment banks and dark-pool operators that are the subject of later discussion.[118] Suffice it to say here, dark pools are a place for uninformed traders to lower their cost of trading by finding other uninformed traders with which to trade. Midpoint exploitation undermines their ability to do this at least to some extent. The practice increases the effective cost of trading for those uninformed traders who use dark pools. This hurts not only those who use dark pools despite the higher effective cost of trading but also those who would have used them but for this higher cost. For the same reasons as discussed with respect to the earlier practices, this will reduce the efficiency of both the allocation of resources over time and the allocation of risk in the economy. At the same time, as discussed more fully below, to the extent that the practice steers more uninformed traders to trade in 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.

D. *High-Frequency Trading and Volatility*

Though much of the controversy about high-frequency traders has focused on their trading strategies, a different, but also important, strain of criticism has alleged a causal connection between HFT activity and greater volatility in equity markets.[119]

---

118. *See infra* Part V.E.

119. *See, e.g.*, Bob Dannhauser, *Debating Michael Lewis' 'Flash Boys': High-Frequency Trading Not All Bad*, SEEKING ALPHA (Apr. 7, 2014, 11:42 AM), http://seekingalpha.com/

1. *General increase in volatility*. One criticism is that HFTs have made the markets more volatile on an ongoing basis month in and month out. Michael Lewis, in *Flash Boys*, for example, asserts that the intraday price volatility of the stock market was 40 percent greater between 2010 and 2013 than it was between 2004 and 2006, and associates this change with the enactment of Regulation NMS and the rise of HFTs.[120]

Lewis's use of this comparison to draw an inference about HFTs' influence on volatility seems deeply mistaken. One big problem is that the years 2004–2006 are a poor comparison sample because they had uncharacteristically low volatility, below any other two-year period from 1998 to 2012.[121] Another big problem is that the years 2010–2013 are also uncharacteristic, as they followed the most severe financial crisis since the Great Depression and would be expected to show high volatility due to the increased uncertainty associated with the fundamental values of securities.[122] A better comparison sample would be 2012 to the present, which shows market volatility that is generally lower than the 1990s and early 2000s, despite the greatly increased role of HFTs in the latter period. As one prominent article has noted, current "[i]ntra-day volatility is below the levels of the pre-electronic 1990s."[123] As best one can tell so far, no serious evidence shows a causal link between the rise of HFTs and ongoing increased volatility. HFTs rose to prominence during a period of greater volatility, which was due to extraneous causes rather than the HFTs themselves. In addition, no general theoretical reasons give rise to an expectation of greater volatility due to HFT activity. Instead, the

article/2129253-debating-michael-lewis-flash-boys-high-frequency-trading-not-all-bad   [http://perma.cc/R4K3-46YB] (describing flash crashes as "disconcerting and damaging to investor trust"); Kit R. Roane, *How NYSE Plans to Use 'Flash Crash' to Reclaim Its Glory*, CNN MONEY (May 12, 2010, 9:38 AM), http://money.cnn.com/2010/05/12/markets/NYSE_flash_crash.fortune/index.htm [http://perma.cc/Z5GG-UT5X] (reporting attempts by NYSE executives to connect HFTs with the flash crash).

120.   LEWIS, *supra* note 2, at 112.

121.   FABIO PANETTA, ET AL., BANK FOR INT'L SETTLEMENTS, THE RECENT BEHAVIOUR OF FINANCIAL MARKET VOLATILITY 1 (2006), http://www.bis.org/publ/bppdf/bispap29.pdf [http://perma.cc/SX42-V3JF].

122.   *See* ANGEL ET AL., *supra* note 29, 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, 1 (2001) (finding significant spikes in volatility during periods of major economic crisis).

123.   ANGEL ET AL., *supra* note 29, at 2.

majority of academic evidence on the subject suggests that HFTs reduce volatility.[124]

2. *The Flash Crash.* More interesting is a second claim: that HFTs exacerbate volatility in a very extreme manner when there has been some kind of disruption in the market, such as the infamous May 6, 2010 "Flash Crash." The Flash Crash occurred within a window of less than thirty minutes during which the Dow Jones Industrial Average (DJIA)—a benchmark of general market performance—dropped about 1,000 points, losing 9 percent of its value, and then recovered almost its entire loss.[125] In this incident, the DJIA suffered the greatest one-hour decline in its history,[126] and several individual stocks displayed astonishing volatility. Accenture, for instance, fell from trading at $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.[127] Apple, as another example, at one moment traded for almost $100,000 per share.[128]

The Flash Crash was widely taken to "highlight[] the risks of electronic trading" as NYSE's then-head of operations suggested.[129] In

---

124.   *See, e.g.*, Joel Hasbrouck & Gideon Saar, *Low-Latency Trading*, 16 J. FIN. MKTS. 646, 648 (2013) (finding that HFT activity reduces volatility); Jonathan Brogaard, Thibaut Moyaert & Ryan Riordan, High Frequency Trading and Market Stability 1–2 (May 2014) (unpublished manuscript), http://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETI NGS/2014-Rome/papers/EFMA2014_0242_fullpaper.pdf [http://perma.cc/DUG7-UQFT]. *But see* Sandrine Jacob Leal, Mauro Napoletano, Andrea Roventini & Giorgio Fagiolo, *Rock Around the Clock: An Agent-Based Model of Low- and High-Frequency Trading* 3 (Université Nice Sophia Antipolis GREDEG Working Papers Series, Working Paper No. 2014-21, Jan. 2014), http://www.gredeg.cnrs.fr/working-papers/GREDEG-WP-2014-21.pdf [http://perma.cc/ UQF5-WYUW].

125.   STAFFS OF THE CFTC & SEC, FINDINGS REGARDING THE MARKET EVENTS OF MAY 6, 2010, at 1 (2010) [hereinafter FLASH CRASH REPORT], http://www.sec.gov/news/studies/2010/ marketevents-report.pdf [http://perma.cc/YC3Y-XTP8]; Alexandra Twin, *Glitches Send Dow on Wild Ride*, CNN MONEY (May 6, 2010, 7:36 PM), http://money.cnn.com/2010/05/06/markets/ markets_newyork [http://perma.cc/KE3Y-Z272].

126.   Tom Lauricella & Peter A. McKay, *Dow Takes a Harrowing 1,010.14-Point Trip*, WALL ST. J. (May 7, 2010, 12:01 AM), http://www.wsj.com/articles/SB100014240527487043707070 4575227754131412596 [http://perma.cc/VEN2-XA2H].

127.   *Id.*

128.   Tom Lauricella & Scott Patterson, *Legacy of the 'Flash Crash': Enduring Worries of Repeat*, WALL ST. J. (Aug. 6, 2010, 12:01 AM), http://www.wsj.com/articles/SB1000142405274 8704545004575353443450790402 [http://perma.cc/2XXJ-MSLY]. Many of the most outlandish transactions executed during the Flash Crash were later cancelled or "broken" by regulators. *See* Brian Korm & Bryan Y.M. Tham, *Why We Could Easily Have Another Flash Crash*, FORBES (Aug. 9, 2013, 7:58 AM), http://www.forbes.com/sites/deborahljacobs/2013/08/09/why-we-could-easily-have-another-flash-crash [http://perma.cc/WXGA-XG8W].

129.   Lauricella & McKay, *supra* note 126 (quoting Louis Pastina, executive vice president and head of operations at NYSE Euronext's New York Stock Exchange).

the years since, blame has been persistently attributed to HFTs and commentators have suggested that HFTs generally increase the severity of market crashes.[130] The report eventually issued by federal regulators, however, explained the Flash Crash not as the result of HFT predation, but as the result of a liquidity crisis caused by a large sell order that triggered a flight of liquidity from the market. This flight involved HFTs, but only in the sense that many HFTs are market makers that left the market in response to the large sell order.[131] This temporary disappearance of the HFTs removed substantial liquidity.[132]

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 answer to this question returns us to the overarching theme of this Article—that comprehension and intelligent regulation of the modern stock market is impossible without a thorough appreciation for the role of adverse selection in shaping the provision of liquidity.[133] Volatility of the kind involved in the Flash Crash is directly connected to adverse selection. The reason is that a large, aggressive sell (or buy) order suggests to liquidity providers that the order submitter may have important private information. HFTs know that if this apparent private information in fact turns out to exist, then they will lose money from trading with that order and so they will widen their spreads.[134] If the threat of being adversely selected by the order becomes extreme enough, many or all liquidity providers will temporarily exit from the market altogether and prices will, as a result, fluctuate widely in the absence of quotes reflecting any plausible estimate of a security's

---

130. *See, e.g.*, 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 [http://perma.cc/44XN-Q68A] (suggesting that "computer-driven high frequency trading is partially responsible for accelerating stock gyrations"); Andrew Smith, *Fast Money: The Battle Against the High Frequency Traders*, THE GUARDIAN (June 7, 2014, 4:00 AM), http://www.theguardian.com/business/2014/jun/07/inside-murky-world-high-frequency-trading [http://perma.cc/4NY8-37HA].

131. Andrei A. Kirilenko, Albert S. Kyle, Mehrdad Samadi & Tugkan Tuzun, The Flash Crash: The Impact of High Frequency Trading on an Electronic Market 25–26 (Sept. 24, 2014) (unpublished manuscript), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1686004 [http://perma.cc/N8CQ-X82Z].

132. FLASH CRASH REPORT, *supra* note 125, at 6.

133. This Article focuses on HFTs as liquidity providers, and ample evidence shows that they play this role. *See, e.g.*, Menkveld, *supra* note 22, at 712.

134. *See supra* Part III.B.

fundamental value.[135] This, in essence, is what happened on a large scale during the Flash Crash.

The behavior of HFTs during the Flash Crash was not predatory; it was simply unheroic. Perceiving the large sell order to have a higher probability of being motivated by private information, given its size and aggressiveness,[136] HFTs removed their quotes to minimize their trading losses, and liquidated the long positions they had accumulated, exacerbating pressures on price declines.[137] Because HFTs provide a large share of liquidity, in their absence, the only quotes left lay far from the true price of a security; that is, the present value of its future cash flows.[138]

3. *Wealth-transfer considerations.* Assessing the wealth transfers resulting from gyrations, such as in the Flash Crash, is equivalent to asking who wins and loses when HFTs stop providing liquidity. Despite suggestions by critics of predatory behavior, HFTs cannot make money if they do not trade. Among traders, the losers are persons 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 the persons who posted previously way-out-of-the-money limit orders against which these market orders transacted.

4. *Efficiency considerations.* Events such as the Flash Crash seem bound to occur from time to time with an HFT-dominated system for providing liquidity. The old NYSE specialist system, in which the specialist was supposed to "lean against the wind" to provide liquidity may have been less prone to such problems. So perhaps was the dealer system more generally, in which human beings made the trading decisions. These occasional brief moments of total collapse of liquidity do not really seem very important in terms of our touchstones for efficiency, however. Very brief sharp deviations of share prices from fundamental values do not seriously undermine the

---

135. FLASH CRASH REPORT, *supra* note 125, at 2–3.

136. *See* David Easley, Marcos M. 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).

137. FLASH CRASH REPORT, *supra* note 125, at 29; *see* Jones, *supra* note 98, at 26.

138. FLASH CRASH REPORT, *supra* note 125, at 45–57; *see* FOUCAULT ET AL., *supra* note 78, at 2 (discussing the "fundamental value" of a security as a market consensus about its appropriate price).

role of share prices in aiding the efficiency with which capital is allocated to new real investment projects and with which existing productive capacity is utilized. Accuracy most of the time is what matters. And investors can protect themselves from extreme results by using orders with limits that would appear to make them marketable. They can stay briefly out of the market without seriously undermining the efficient allocation of resources over time or the efficient allocation of risk. The modern stock market's overall performance in terms of liquidity provision and operational costs is far better than the market of the past, which matters more for the ultimate social goals promoted by a well-functioning equity market.

E.   *Dark Pools and the Fate of Customer Orders*

The next two controversial practices involve the fate of customer orders that end up in dark pools.[139] The first alleged practice involves large investment banks, which are both important providers of brokerage services and operators of most of the largest dark pools.[140] They are accused of routing their brokerage customers' orders to the banks' own dark pools even when the orders will receive inferior execution there.[141] Related to this first practice is the claim that it is common for a dark-pool operator to misrepresent the nature of other parties' trading in its pool in order to induce brokerage customers to agree to have their orders sent to this pool.

---

139.   *See, e.g.*, Nathaniel Popper, *As Market Heats Up, Trading Slips into Shadows*, N.Y. TIMES, Mar. 31, 2013, at B1; *Shining a Light on Dark Pools*, THE ECONOMIST: SCHUMPETER (Aug. 18, 2011, 4:08 PM), http://www.economist.com/blogs/schumpeter/2011/08/exchange-share-trading [http://perma.cc/22EM-RGS7]; David Zeiler, *How High-Frequency Traders Use Dark Pools to Cheat Investors*, MONEY MORNING (Apr. 22, 2014), http://moneymorning.com/2014/04/22/how-high-frequency-traders-use-dark-pools-to-cheat-investors [http://perma.cc/V6LL-6VSX] (discussing how HFTs may be abusing dark pool access).

140.   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, Michael 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* RHODRI PREECE, CFA INST., DARK POOLS, INTERNALIZATION, AND EQUITY MARKET QUALITY 14–15 (2012). All of the ten largest brokers on NYSE are also global investment banks. *See NYSE Broker Volume*, NYSE MARKET DATA, http://www.nyxdata.com [http://perma.cc/YC2V-XUAG].

141.   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, *supra* note 2, at 123. Inferior execution could also occur on a dark pool if the counterparties trading there were especially informed or were given information about the existence of the customer limit orders posted there.

The second alleged practice involves large investment banks ignoring their brokerage customers' instructions to direct their orders to specific venues, and instead routing the orders to their own dark pools, where, again they receive an inferior execution.[142] The market may not solve the problem, critics of the practice continue, because customers have difficulty detecting the practice and, even when they do, the customers are reluctant to switch brokers because of their dependence on soft-money services the banks provide their most loyal customers.[143]

Dark-pool operators have been no strangers to legal actions involving allegations that they have engaged in these kinds of practices. The SEC has brought a number of successful proceedings against dark-pool operators.[144] 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 United States) misrepresented to users the involvement of HFTs in LX; the informational advantages given to HFTs; and that Barclays, as a broker, impartially sought to route orders for best execution, when it actually disproportionately routed client orders to its own pool.[145]

We do not know whether any of these practices is in fact widespread. As discussed below, they are clearly illegal and their wealth-transfer and efficiency effects appear completely negative. We will suggest in Part VI some policy reforms that would make enforcement more effective if, despite the illegality of these practices, evidence emerges that they are in fact widespread.

1. *Understanding the function of dark pools.* Recall that 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.[146] Moreover, it has the ability to restrict who can post

---

142. *See, e.g.*, LEWIS, *supra* note 2, at 182–87, 264–65.

143. *Id.* at 102–03, 214–15.

144. *See, e.g.*, Liquidnet Inc., Exchange Act Release No. 72,339, 2014 WL 2547522 (June 6, 2014); Pipeline Trading Systems, Exchange Act Release No. 65,609, 2011 WL 5039038 (Oct. 24, 2011).

145. Complaint at 2–4, Schneiderman v. Barclays Capital, Inc., No. 451391/2014 (N.Y. Sup. Ct. 2014).

146. *See* 17 C.F.R. § 242.602(b)(1) (2015) (defining scope of reporting requirements); *id.* § 242.600(b)(65) (defining broker-dealer); *id.* § 242.600(b)(73)(ii)(A) (defining subject security).

limit orders and submit marketable orders.[147] Dark pools, despite their nefarious-sounding moniker, can provide useful, legitimate services to their customers. They arose because of the more liberal regulatory environment established by the NMS Amendments to the Exchange Act and the information-technology revolution. The key force driving their rise—as with so many other institutions and practices within the new stock market—was concern to mitigate adverse selection. A dark pool's most valuable characteristic, from this perspective, is to provide a venue where uninformed buyers and sellers, seeking to trade substantial amounts of stock, can minimize the movement of prices against them and transact at prices potentially much better than the NBO and NBB.[148] These advantages arise from the fact that quotes on a dark pool are not publicly displayed and because dark-pool operators have the ability to exclude traders.

In terms of serving these functions, the ideal dark pool would be one in which both the parties posting limit midpoint orders and parties sending in marketable orders are completely uninformed. The midpoint is a substantially better price for the buyer than the NBO, and it is the same for the seller relative to the NBB. The system begins to break down to the extent that the parties posting limit orders are in fact informed. This is because their counterparties—the parties submitting marketable orders—would be disadvantaged by being in a dark pool because they would not be able to see from the size of the posted limit orders that an informed party might be on the other side. It similarly begins to break down to the extent that the parties submitting the marketable orders are informed. This is because they will only transact against the limit orders in the dark pool when their information suggests that the midpoint is a price that makes the transaction advantageous to them, which means it is a price that makes the transaction disadvantageous to the person posting the limit offer. Thus, the dark-pool operator provides a service to the extent that it can effectively monitor both the parties posting the

---

147. *See* Regulation of Exchanges and Alternative Trading System Rule 301(b)(5), 17 C.F.R. § 242.301(b)(5) (2015); Concept Release on Equity Market Structure, Exchange Act Release No. 61,358, 75 Fed. Reg. 3594, 3614 (proposed Jan. 14, 2010) (to be codified at 17 C.F.R. pt. 242) ("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.").

148. *See, e.g.*, PREECE, *supra* note 140, at 12–13.

midpoint limit orders and the parties sending in marketable orders to assure that each side is relatively unlikely to be informed.[149]

2. *Wealth-transfer and efficiency considerations.* To the extent that a dark pool does not function in accordance with the ideal described above, an order sent there may execute at less-desirable terms than if it were sent to another venue. An investment bank that operates a dark pool has intimate knowledge of the extent to which it in fact falls short of this ideal. If a brokerage unit of an investment bank sends a trader's order to the bank's own dark pool when the broker knows, or should know, that the order would receive superior execution elsewhere, the bank gains from the extra volume of trade in its dark pool and in other possible ways,[150] and the customer loses from the inferior terms of execution. The same result is likely to arise if the trader, having tried to determine where its order is most likely to get best execution, instructs its broker that the order be sent to a venue other than the bank's dark pool, but the instruction is ignored. The same result is also likely if the bank operating the dark pool misrepresents to customers the nature of the parties allowed to trade on the bank's dark pool, in order to create the impression that there exists less danger of informed counterparties there than is in fact the case. Such a misrepresentation is likely to attract orders that could execute on better terms elsewhere. All of these results generalize if these failures are common practices: they make investment banks richer and traders poorer.

At the simplest level, the negative efficiency consequences arising from these broker or dark-pool-operator failures are the same as the efficiency justifications more generally for legal remedies against those who breach contracts or engage in misrepresentation. A broker has a duty of best execution in the way it routes a customer's order.[151] This requires the broker to exercise "reasonable diligence to

---

149. The operator provides a similar service to the extent that it keeps out HFTs that engage in midpoint order exploitation.

150. 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. Even more nefarious inducements are suggested by the criticisms, such as exploitation of orders by a broker's HFT affiliate that has improperly been given details about orders.

151. This duty exists both as a matter of state common law of agency and under the rules of the Financial Industry Regulatory Authority. *See, e.g.*, *In re* Merrill Lynch Sec. Litig., 911 F.

ascertain the best market" for a transaction to ensure an order receives a price "as favorable as possible under prevailing market conditions."[152] 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: that the parties voluntarily entered into the transaction no longer leads to the presumption that it can be expected to advance the interests of both and that it is thus efficiency enhancing. The same analysis applies to a broker disregarding customer instructions as to where to route an order, which is also clearly illegal.[153] When a misrepresentation induces a party to enter a contract—in this case agreeing to have an order sent to a dark pool that has different counterparties than represented—the same problem is again created: the transaction no longer carries the presumption of being efficiency enhancing. Finally, if an investment bank that is both a broker and a dark-pool operator provides information concerning a customer order to its trading affiliates (or anyone else), it would be violating its agency duties of confidentiality, provisions of Regulation ATS, and probably its own marketing material.[154] Again, one can view these duties as default provisions in a contract, the breach of which robs the transaction of the presumption that it is efficiency enhancing.

Supp 754, 760 (D.N.J. 1995), *rev'd on other grounds sub nom.* Newton v. Merrill, Lynch, Pierce, Fenner & Smith, Inc., 135 F.3d 266 (3d Cir. 1998) (describing fiduciary duties of brokers); *id.* at 769 ("A broker-dealer's duty to seek to obtain the best execution of customer orders derives from the common law agency [duty] of loyalty . . . ." (quoting DIV. OF MKT. REGULATION, *supra* note 28, at V-1)); *see also* DIV. OF MKT. REGULATION, *supra* note 28, at 21 ("A broker-dealer has a duty to seek to obtain the best execution for its customer orders.").

152.   FINRA Rules, *supra* note 87, at Rule 5310(a)(1). Reasonable diligence requires a broker to consider "the character of the market for the security (e.g., price, volatility, relative liquidity, and pressure on available communications)"; "the size and type of transaction"; "the number of markets checked"; the "accessibility of the quotation"; and "the terms and conditions of the order which result in the transaction, as communicated to the member and persons associated with the member." *Id.* at Rule 5310(a)(1)(A)–(E).

153.   Ignoring a principal's instructions would violate the agent's duties noted above as well as an agent's specific duty of obedience to the principal. *See, e.g.*, Gagnon v. Coombs, 654 N.E.2d 54, 61 (Mass. App. Ct. 1995) (discussing agent's duty of obedience). Though many retail brokers do not provide customers any choice as to execution venue, some do, and brokers for institutional customers would be expected to provide options in this respect. *See, e.g.*, *SmartRouting*, INTERACTIVE BROKERS, https://www.interactivebrokers.com/en/index.php?f=16 85 [http://perma.cc/K5H7-DAPM] (providing comparison of security prices offered by Interactive Brokers versus industry average).

154.   *See* RESTATEMENT (THIRD) OF AGENCY § 8.05(2) (2006) (outlining agent's duty of confidentiality). Rule 301(b)(10) of Regulation ATS requires an alternative trading system (ATS) (essentially, a nonexchange stock-market trading venue, such as a dark pool) to protect users' confidential information. 17 C.F.R. § 242.301(b)(10) (2014).

If we look at the efficiency question from the more nuanced mode of inquiry that we have been generally using in this Article, the story becomes a bit more complicated. If these practices are in fact widespread, then they increase the effective costs of trading for those uninformed traders whose orders go to dark pools. They also discourage some traders from using dark pools who would use them, and who would enjoy lower costs of trading as a result, if these practices did not occur. For the same reasons as discussed with respect to the earlier HFT practices, these effects will reduce the efficiency of both the allocation of resources over time and the allocation of risk in the economy.

On the other hand, these practices often probably improve share-price accuracy and hence enhance the real economic-efficiency effects that flow from that. This is true to the extent that informed investors get advantages trading in dark pools because these practices lower their cost of trading. It is also true to the extent that uninformed investors are scared off from the dark pools because of these practices. When more uninformed investors instead trade on exchanges, the spreads are lower for informed investors. The uninformed investors trading on exchanges subsidize informed investors and this subsidy encourages those in the business of generating and trading on fundamental value information. Even if the positive efficiency effects from the practices dominate the negative ones, however, which we have no reason to believe, we would not recommend abandoning the traditional rules of upright commercial practice that are breached by these practices. Other methods exist for achieving the same thing, for example by prohibiting or limiting dark pools.[155]

F. *Maker-Taker and Taker-Maker Fees*

Recall that "maker-taker fees" refer to a situation in which a trading venue pays a fee (a rebate to the broker) for each nonmarketable limit order it receives that is ultimately executed, and charges a fee to the broker for each marketable order it receives that executes against a limit order posted on the venue. With "taker-

---

155. *See* Kevin S. Haeberle, *Stock-Market Law and the Accuracy of Public Companies' Stock Prices*, COLUM. BUS. L. REV. (forthcoming 2015) (manuscript at 25–27, 44–45), http:// papers.ssrn.com/sol3/papers.cfm?abstract_id=2410422 [http://perma.cc/FF72-HEAC] (arguing that off-exchange trading venues increase the probability of adverse selection on exchanges).

maker fees" the venue does the opposite.[156] In a typical maker-taker structure, a venue charges in the range of \$.0025–\$.0030 per share for marketable orders and pays rebates for executed nonmarketable orders in the range of \$.0020–\$.0025 per share.[157] In contrast, the taker-maker scheme will typically charge executed nonmarketable limit orders \$.0007–\$.0010 per share and rebate \$.0002–\$.0005. The actual fees and rebates vary somewhat across exchanges and vary through time,[158] but under either scheme the venue typically nets in the neighborhood of \$.0005 per share.

The maker-taker and taker-maker fee structures have been subject to vigorous criticism. In a letter to the SEC, Senator Charles Schumer argued that they "create[d] a conflict of interest, as brokers may be incentivized to execute trades on a particular venue even if that venue is not offering the best price," creating "room for brokers to arguably put their own interests ahead of their clients by maximizing the rebates they receive from exchanges."[159] Senator Carl Levin, asking a TD Ameritrade executive about the fact that TD Ameritrade had directed virtually all nonmarketable orders to a single trading venue, which happened to offer the highest rebate for such orders, suggested, "[y]our subjective judgment as to which market provided best execution for tens of millions of customer orders a year allowed you to route all of the orders to the market that paid you the most . . . . I find that to be a frankly pretty incredible coincidence."[160] Since those hearings, a class-action lawsuit has been

---

156.	*See supra* Part II.C.7.

157.	*See generally* Lawrence E. Harris, Maker-Taker Pricing Effects on Market Quotations 2 (Nov. 14, 2013) (unpublished manuscript), http://www.securitiesmosaic.com/gateway/sec/speech/hujibusiness_Maker-taker.pdf [http://perma.cc/DP97-SCJP] (discussing typical maker-taker pricing scheme).

158.	The actual fees and rebates imposed must be publicly available and must be the same across all stocks traded on the exchange. The fees and rebates for any exchange are available on the exchange's website. *See generally* Securities Exchange Act of 1934 § 19(b)(3)(A)(ii)(4), 15 U.S.C. § 78s(b)(3)(A)(ii) (2012); 17 C.F.R. § 240.19b-4(f)(2) (2013).

159.	Letter from Sen. Charles Schumer to Hon. Mary Schapiro, Chairman, SEC (May 10, 2012). *Flash Boys* added its own criticism, declaring that "[t]he maker-taker system of fees and kickbacks used by all of the exchanges was simply a method for paying the big Wall Street banks to screw the investors whose interests they were meant to guard." LEWIS, *supra* note 2, at 168–69.

160.	*See* William Alden, *supra* note 7. A recent New York Times editorial suggested that maker-taker fees are "corrupting" brokers, who "under the guise of making subjective judgments about best execution, . . . were routinely sending orders to venues that paid the highest rebates," and concluded by calling for greater regulation or elimination of maker-taker fees. *See* Editorial, *supra* note 7 (using TD Ameritrade data from the fourth quarter of 2012).

launched against TD Ameritrade, alleging that it violated its duty of best execution in this way.[161]

1. *Modeling the fee structures.* To see how these fee structures work, we will start with a model that makes several simplifying assumptions. These assumptions are that the maker rebate and taker fee are the same and equal r, that there is a single, consolidated trading venue, that all traders (not just HFTs, the traders are in the business of liquidity supply) submit their orders directly to the trading venue rather than doing so through a broker, and that regular traders (that is, all traders except HFTs) are rational actors who are well informed about the terms of trade available in the market. Under these assumptions, maker-taker fees and rebates turn out to be entirely benign: they have no effect on how liquidity is supplied or on anyone's wealth position. We will then go on to consider the consequences of relaxing these assumptions.

a. Liquidity suppliers. Let us look at things first from the liquidity suppliers' point of view. These are HFTs, the entities in the competitive business of supplying liquidity through posting nonmarketable limit orders, as well as regular traders who post such orders in order to transact at better prices than they would get from marketable orders. Consider first the HFTs. Let S equal half of the spread HFTs need to cover their costs associated with making a market given the possibility of informed trading. Let P be the consensus value of a security at the time a quote is made. Absent a liquidity maker rebate, the limit orders posted by the HFTs at the trading venue will put the offer at P + S and the bid P – S.[162] When there is rebate of R (R < S), the offer will be P + S – R. This is because an HFT will receive R every time a limit order it posted transacts, and so receiving P + S – R is, from an economic point of view, the equivalent to receiving P + S if there were no rebate (that is, the amount the HFT would need to receive to break even without a rebate). Under the same logic, the bid will be P – S + R, which is the economic equivalent for the HFT of paying the breakeven price of P – S per share in a world without rebates.[163] Rational regular traders

---

161. *See, e.g.*, Complaint at 20, Klein v. TD Ameritrade Holding Corp., No. 14-cv-5738 (D.N.J. Sept. 15, 2014), 2014 WL 5018542.

162. *See supra* Part III.B.

163. For expository simplicity, the analysis assumes the tick size (the minimum difference allowed by the market between one price and the prices above and below) is infinitesimal,

who submit nonmarketable limit orders will set the limit that they specify so, relative to the limit they would have set in the absence of a rebate, they adjust a buy order limit down by R and a sell order up by R.

b. Liquidity takers. Now consider liquidity takers, that is, traders who submit marketable orders. On one hand, because of the adjustments to the available offers and bids described above, posted by liquidity suppliers in reaction to receiving rebates, the offer price which traders submitting marketable orders pay for shares, and the bid price at which they sell them, are each improved by R relative to what they would have paid or received in the no-rebate world, respectively. On the other hand, this narrowing of the spread is exactly offset by the fee R they must pay the venue for each share bought or sold.

c. The benign effect of the fees and rebates given the simplifying assumptions. This analysis shows that the combination of maker rebates and taker fees leads to precisely the same terms of trade from an economic point of view as would prevail in their absence. So there is no reason to think, at least under the assumptions we have employed so far, that anyone would behave differently than in a world without rebates. This same analysis holds for taker rebates and maker fees, because, if they are equal, they simply correspond to a negative R in the expressions above.

2. *Relaxing the assumptions of the model.* Consider what happens when we relax various assumptions of the model.

a. Relaxing the assumptions about the trading venues. Relaxing the simplifying assumptions we made about trading venues does not change the conclusions. We started with the assumption that the maker rebate and the taker fee are equal. In reality, they are not because the venue needs to be paid a fee for the service of providing a place to trade. This fee is the difference between the two.[164] The size of the venue's fee is irrelevant to the current discussion, and, with it

---

whereas, pursuant to NMS Rule 612, for most stocks it is in fact a penny, but including this complication in the analysis leaves the conclusion largely unchanged.

164. The rebate paid to liquidity makers is generally smaller than the fee charged to the liquidity taker in the maker-taker structure, and the opposite in the taker-maker structure.

taken out, the maker rebate and taker fee are by definition always going to be equal.

We also assumed a single, consolidated trading venue. The real world, with competing venues, should work the same way, however, at least as we continue to hold on to our assumptions about the nature of the traders in the market, because both liquidity suppliers and liquidity takers, when deciding whether to send an order to any given venue, can make the same calculations as in the model above that the venue's fee and rebate system will affect bids and offers in a way that just offsets these fees and rebates.

b. Relaxing the assumptions about regular traders.   The assumptions about traders may play a more critical role. To start, we assumed that regular traders send their orders directly to the market without broker intermediation. That means that in maker-taker venues, traders who submit nonmarketable limit orders would receive the rebate directly and those who submit marketable orders would pay the fee directly (and the reverse for taker-maker venues). In fact, few regular traders are allowed to send orders directly to trading venues; they must use a broker. So, for a regular trader, the broker is the one who receives the rebate and pays the fee. And typically, nothing in the trader's contract with its broker provides that the rebates and fees be passed through directly to the trader,[165] nor is this required by regulation.

If we continue to assume competitive exchanges, a competitive market for brokerage services, and regular traders who are rational and sufficiently informed to be able to effectively monitor the quality of service their brokers are delivering, the effects of the maker-taker fee structure (and the taker-maker one) remain benign. As set out below, the maker rebates offered by a given venue will lead to commensurately lower brokerage fees for nonmarketable limit orders sent to this venue and the taker fees will lead to commensurately higher brokerage fees for marketable orders there. So the rebates and fees are simply passed on indirectly through their effects on brokerage fees.

---

165. *High Speed Trading Hearings*, *supra* note 8, at 8 (statement of Robert Battalio, Professor of Finance, University of Notre Dame). Interactive Brokers offers such a contract, but it also offers a fixed-price-for-execution contract. *See Commissions: Stocks, ETFs and Warrants—Overview*, INTERACTIVE BROKERS, https://www.interactivebrokers.com/en/?f= commission [http://perma.cc/3KKK-WPUR].

Consider first a trader who wishes her broker to submit on her behalf a nonmarketable buy limit order. This buy order will be competing with bids posted by HFTs. Recall that each HFT will increase its bid price commensurately to the size of the rebate, if any, at the venue where it is posting its offer. The trader's limit order, to have the same likelihood of timely execution as it would in a venue without a rebate, would need to have a commensurately higher limit price as well. If the trader's limit order ultimately does transact, the cost of buying a share to the trader is her limit price plus B_{nm}, where B_{nm} is the broker's commission for nonmarketable orders posted on that venue. If C_{nm} is the cost to the broker for the act of servicing a nonmarketable order, but the broker receives R as a rebate, then the broker's net cost is $C_{nm} - R$. In a competitive brokerage market with informed consumers of the service, the forces of competition would assure that $B_{nm} = C_{nm} - R$, whatever R is at the venue. When R is higher, the trader will need to submit a commensurately higher limit price to get the same likelihood of timely execution, and so would need a commensurately lower brokerage fee to come out even.

Now consider a trader who wishes its broker to submit on his behalf a marketable buy limit order. The price he will need to pay will be $P + S + B_{m}$, where B_{m} is the broker's commission for nonmarketable orders posted on this venue. If C_{m} is the cost to the broker for the act of servicing a marketable order, and the broker must pay a taker fee of R as a rebate, the broker's net cost is $C_{m} + R$. In a competitive brokerage market with informed consumers of the service, the forces of competition would assure that $B_{m} = C_{n} + R$, whatever R is. When R is lower, S will be commensurately higher, and so the brokerage fee will need to be lower for the trader to end up paying the same net price for a share.[166]

This story, though, has two flaws, which raise, but do not prove, the possibility that rebates affect how players in the market behave. First, at least for most retail trades, brokerage fees are fixed on a per-execution basis and do not depend on the venue in which a particular order transacts. Second, it is not so easy for a trader to monitor the performance of his broker. With brokerage commissions invariant to where an order is sent, a broker has an incentive to send a nonmarketable limit order to the venue with the highest maker rebate and marketable orders to the venue with the lowest taker fee.

---

166.   Mirror images of these stories apply to nonmarketable and marketable sell orders.

This problem is potentially more acute with a nonmarketable limit order because determining whether or not it received best execution is more difficult to figure out than with marketable orders. The quality of a limit-order execution includes whether it executes or not, how long it takes to get executed, and how likely the market is to move in a disadvantageous direction following an execution. A recent paper sheds some light on this question.[167] To take the extremes at the time of their data, Edge X had a rebate rate of $.0030 per share and provided a fill rate (partial or full) of 54 percent, an average fill speed of 111 seconds, and a good-fill ratio (measured as the proportion of time the midpoint of the market quotes five minutes after a transaction was above the transaction price at the bid or below the transaction price at the ask) of 49 percent. In contrast, Boston, with a rebate rate of -$.0014 per share had a fill rate of 74.5 percent, a fill speed of thirty-three seconds and a good-fill ratio of 55 percent. Four retail brokerage houses sent roughly half of their limit orders to Edge X and only Interactive Brokers spread its limit orders around.[168] Evidently, nonmarketable limit orders posted on high-rebate venues are not getting very good execution. Yet that is where many retail brokerages are sending nonmarketable limit orders when the customer does not specify a venue for execution.

3. *Wealth-transfer considerations.* Even if traders are unable to monitor the quality of broker execution completely, especially with regard to nonmarketable limit orders, the level of competition among trading venues and among brokers probably assures that neither type of enterprise is making excess returns as a result of the maker-taker and taker-maker rebate and fee structures that abound. In terms of traders, there appears to be excess liquidity on the maker-taker venues and this benefits large traders who place marketable orders there. Those who are harmed are the ordinary traders whose nonmarketable limit orders get inferior execution. Putting a cost on inferior execution is difficult.

4. *Efficiency considerations.* Extra liquidity of the kind that appears to be generated by venue rebates and fees may be valuable to

---

167. Robert Battalio, Shane Corwin & Robert Jennings, Can Brokers Have It All? On the Relation Between Make-Take Fees & Limit Order Execution Quality 10 (Dec. 13, 2013) (unpublished manuscript), http://www3.nd.edu/~scorwin/documents/BattalioCorwinJennings_20 131213_SSRN.pdf [http://perma.cc/83NY-JZ3A].

168. *Id.* at 11.

informed traders and hence may add to share-price accuracy. The welfare effects of poor execution quality for traders placing nonmarketable limit orders are harder to trace through without a better sense than we have of who they are and what they are trying to accomplish. They too may be informed traders, or they may be uninformed traders who think they are informed. In any event, as with dark-pool misconduct, it is illegal for brokers not to provide best execution, which is an implicit term in their contracts with customers. Even if their failure to do so results in greater price accuracy, allowing this breach of an understanding between broker and customer does not seem the right way to increase share-price accuracy.

G. *Purchase of Order Flow*

Many brokers sell their order flow to "internalizers" such as Citadel and KCG Americas, which are trading venues that match incoming buy orders with nearly simultaneous incoming sell orders, buying from the sellers and selling to the buyers.[169] The payments the brokers receive from the internalizers are referred to as "payment for order flow."[170] Internalizers typically only pay for marketable orders. They generally agree to provide the broker's customers with nominal price improvement, with shares purchased from sell-order customers at perhaps $.0001 over the NBB and shares sold to buy-order customers at that amount below the NBO. Battalio and his coauthors show that nine out of the ten retail brokers studied send virtually all marketable orders to order-flow purchasers.[171]

Payment for order flow has been heavily criticized as potentially "creat[ing] . . . conflicts of interest" between brokers, who seek to maximize revenue from selling order flow, and customers, who might receive better execution elsewhere.[172] Indeed, in a letter to SEC Chair White, Senator Carl Levin stated that "[c]onflicts of interest erode public confidence in the markets" and that "payments [for order flow] create another incentive for brokers to maximize their own profits at

---

169. HARRIS, *supra* note 18, at 520–22 (defining and discussing internalization).

170. Allen Ferrell, *A Proposal for Solving the "Payment for Order Flow" Problem*, 74 S. CAL. L. REV. 1027, 1028 (2001).

171. Battalio et al., *supra* note 167, at 10.

172. *High Speed Trading Hearings*, *supra* note 8, at 33 (statement of Joseph P. Ratterman, CEO, BATS Global Markets); *id.* at 2 (opening statement of Sen. Carl Levin ("'[P]ayment for order flow,' can add up to untold millions, and almost every retail broker keeps these payments rather than passing them on to clients")).

the expense of best execution of customer orders," and should be eliminated.[173]

1. *Wealth effects.* If the internalizer and broker markets are each competitive, then the analysis of the wealth effects of payment for order flow is very simple. There are none. The internalizers pay the brokers what they are able to make above their costs (plus a market return on capital) from executing the buy and sell orders on the promised terms of slight price improvement. What the brokers receive reduces the costs of providing brokerage services and in a competitive brokerage market these savings are passed on through lower commissions.

If, on the other hand, the internalization market is not competitive, then internalizers are the monopsonist equivalent of oligopolists and may have the market power to pay less for order flow than what they make above their costs (plus a market return on capital) from executing orders at a slight improvement over the NBO and NBB. What brokers then pass on to customers in the form of lower brokerage fees would be insufficient, leaving traders worse off. Alternatively, the internalization market may be competitive, but the broker market may not be. In that case, the brokers may sell the order flow for an appropriate price but only pass a portion of those savings on to customers. If so, then just like with maker-taker fees, payment for order flow could create a genuine agency problem by aligning brokers' incentives with receiving payments for order flow, whereas better execution may be obtained on an exchange. In this noncompetitive version of the market, customers receive inferior execution because their orders are not routed to exchanges and only receive a portion of the savings obtained.

2. *Efficiency effects.* Liquidity suppliers on the exchange do not get the opportunity to interact with internalized order flow. This alters the adverse-selection environment on the exchange.[174] By removing uninformed order flow from exchanges, the probability that liquidity providers face informed traders increases, thereby increasing the spread that fundamental-value traders need to pay. This, as we have seen, reduces their incentives to engage in this activity and thus reduces share-price accuracy. On the other hand, it is possible that

---

173.   Letter from Sen. Carl Levin to Hon. Mary Jo White, Chairman, SEC *supra* note 14.

174.   *See* Haeberle, *supra* note 155, at 44–45.

execution at terms slightly better than NBO and NBB, when combined with lower brokerage fees resulting from the indirect passing on of at least part of the payment for order flow, reduces the effective cost of trading for uninformed traders. If so, the practice of brokers selling order flow to internalizers improves the efficiency with which resources are allocated over time and the efficiency with which risk is allocated. Still, for these efficiency gains to be fully realized, internalizers must pay competitive rebates for order flow, and these must be fully passed on to traders one way or another.

VI. RECOMMENDATIONS

Potential regulatory responses, prompted by the preceding survey of controversial practices in the new stock market, fall into four groups. The first three controversial practices that we reviewed—electronic front running, slow-market arbitrage, and exploitation of dark pool midpoint orders—all involve HFTs acquiring information concerning newly executed transactions and changing their quotes ahead of others in the market. The next practice, the relation of HFTs to market volatility involves the entire market's reliance on this form of liquidity supply. The next two practices relate to dark-pool operations and their investment-bank operators as brokers. And the last two involve payments to brokers relating to the customer orders they are handling.

A. HFT Speed in Obtaining Market Information

Because electronic front running, slow-market arbitrage, and exploitation of dark-pool-midpoint orders all involve HFTs acquiring market information ahead of others, once we have assessed the social desirability or undesirability of each of these practices, any possible reform will need to take account of its effect on all three. Below we will make such an assessment of each of these practices and then discuss possible reforms.

1. *Would it be desirable to eliminate electronic front running?* Persons transacting in stocks have always played a cat-and-mouse game in which each tries to figure out what the others are doing. One function of regulation is to step in and prohibit particular informational advantages when such intervention can lead to improved social outcomes. The question is whether the informational advantages HFTs obtain from electronic front running call for such an

intervention. The first step to answering this question requires a policy analysis that compares a world with and without electronic front running. Based on what we know at the moment, the matter is too close to call, with considerations pointing in both directions. The unfairness case against electronic front running is weak. It is the efficiency case for and against the practice that is too close to call.

    a. Actual unfairness. The rhetoric of the critics of electronic front running focuses primarily on what they see as the unfairness of the resulting wealth transfers.[175] Deeper analysis, however, shows that a compelling case for the elimination of the practice cannot be made on this basis of perceived unfairness. As we have seen above, the practice actually appears to benefit ordinary people to the extent that they invest directly in the market as retail customers. The same is true to the extent that such ordinary people have channeled savings to mutual funds or pension funds that invest all or a portion of their funds in equities on an index basis and execute their purchase and sale transactions on exchanges. This is because retail investors are largely uninformed and index investing is by definition uninformed. The elimination of electronic front running would reduce liquidity, which, as we have seen, would make uninformed trading more expensive and the uninformed traders would enjoy no countervailing gain from the inability of market makers to detect transactions at other exchanges.[176] To the extent that ordinary people invest through actively managed mutual and pension funds, which, by definition, seek to be informed traders, any net benefit that would arise from the practice's elimination is, as discussed above, likely to be substantially

---

    175. *See supra* notes 83–84.

    176*.* It should be noted that a significant portion of retail marketable orders and index-based institutional orders execute off exchanges in venues where the trades can be identified as largely uninformed. *See* Order Granting Approval to Proposed Rule Change, as Modified by Amendment No. 1, To Establish the Retail Price Improvement Program on a Pilot Basis until 12 Months from the Date of Implementation, Exchange Act Release No. 68,937, at 17, 78 Fed. Reg. 12,397, 12,404 (Feb. 22, 2013); PREECE, *supra* note 140, at 3 ("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, when, in the absence of the practice, the spreads would be wider to reflect the greater risk that the HFTs are dealing with informed traders. In reality, however, the spreads are barely smaller in these off-exchange executions (that is, there is only a small amount of "price improvement"). As analyzed *supra* Part V.G, 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.

captured by the persons running the funds in the form of the higher fees and salaries they would be able to command.

  b. Efficiency. A stronger case can be made for the elimination of electronic front running on efficiency grounds. As at least our first-cut efficiency analysis of electronic front running shows, elimination of the practice would involve an unavoidable efficiency trade-off. There would be efficiency gains in the form of better capital allocation and utilization of the economy's existing productive capacity, which would arise from increases in price accuracy.[177] But there would be efficiency losses in the form of less-efficient allocation of resources over time and allocation of risk, which would arise from decreases in liquidity.[178] A priori, however, there is no obvious reason for choosing the point in this trade-off associated with the elimination of the practice over the point associated with the continuation of the practice.

  Also, our more nuanced analysis, which takes account of the presence of announcement-information traders and inside-information traders, suggests it is quite possible that eliminating electronic front running would in fact reduce, rather than improve, price accuracy, so no trade-off occurs and the practice is unambiguously efficiency enhancing.[179]

  c. Resources consumed. What is clear is that prohibiting electronic front running would save the very substantial human and material resources currently supporting the practice that would be freed up to be used elsewhere in the economy in some more clearly productive way. These savings suggest that if a reasonably cost effective way could be found to eliminate the practice and that no market forces are likely to take us to the same result in the foreseeable future, such regulation would be desirable, unless we

---

  177. *See supra* Part V.A.2.

  178. *See supra* Parts V.A.2 and V.A.3.

  179. *See supra* Part V.A.4. It is worth noting, however, that the example of electronic front running we address deals with a market maker who defensively removes a quote when a transaction at another exchange is observed. It is possible to imagine an example in which the market maker removed its initial quote, and then posts an inferior quote against which an incoming order nonetheless transacts, because of a lack of quote competition at the top of the book. If this higher quote simply reflects the market maker seizing an opportunity to charge a higher price, then it is just a wealth transfer with no social benefits. In fact, the higher cost of trading can lead to a variety of negative effects. Conversely, if the higher price represented an upward adjustment in the quote for risk-management reasons, then our analysis might mirror that in the text above.

affirmatively believe the practice is efficiency enhancing on a net basis in terms of the factors discussed just above.

   d. Appearance of unfairness.   Although our analysis above suggests that electronic front running does not actually result in unfairness, HFT practices of this sort are clearly viewed by a substantial portion of the public as being unfair.[180] Much of this perception is of course due to the very criticism of HFT practices that this Article seeks to subject to more serious analysis. Normally, the better response to misunderstanding is education, not a change in what people are allowed to do when the activity does not in fact pose a problem. Still, this perception of unfairness may be very hard to eradicate. If one were persuaded that an efficiency analysis leans toward the conclusion that the practice is on balance socially undesirable, the existence of an unfounded but persistent sense of unfairness can add to the desirability of its prohibition. Such a perception 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, absent having this sense that something unfair was going on, would find equities to be an investment vehicle that suits some of their needs.

   2. *Adding slow-market arbitrage and dark-pool-midpoint-order exploitation to the analysis.*   There are presumably large synergies in terms of the use of real productive resources between electronic front running, slow-market arbitrage, and dark-pool-midpoint-order exploitation. So resource use is not an independent consideration with regard to these other two speed-based HFT practices. Still, these two practices each seem unquestionably undesirable.

   No fairness argument can be made that slow-market arbitrage helps any ordinary investors: it hurts all regular traders, uninformed and informed alike, by increasing their effective cost of trading.[181] And because of this, its economic welfare effects are unambiguously negative as well.[182] The increased effective cost of trading for informed traders means that it is less rewarding to seek out bits of publicly available information and to analyze their implications in a

---

   180.   *See supra* notes 83–84.

   181.   *See supra* Part V.A.2.

   182.   *See supra* Part V.A.3.

sophisticated way. As a consequence, share-price accuracy, with its beneficial effects on the real economy, is reduced. The increased effective cost of trading on uninformed traders has negative effects on the efficient allocation of resources over time and on the efficient allocation of risk.

For slow-market arbitrage to occur, however, the same HFT must have posted a quote that is at the top of the book of an exchange; discover that the best quote available for that stock nationally has changed; an order must transact against the HFT's now-stale quote; and the HFT must be able to transact against the new best quote before anyone else can. This seems like it would be an unusual circumstance, although it is worth further empirical study.

Dark-pool-midpoint-order exploitation hurts uninformed investors and in so doing again has negative effects on the efficient allocation of resources over time and on the efficient allocation of risk.[183] It arguably helps share-price efficiency by deterring uninformed traders from using dark pools so they use exchanges instead, thereby causing spreads to lower and reducing the effective cost of trading for fundamental-value-informed traders. But if it is in fact good social policy to push uninformed traders into the market to subsidize such informed trading, other methods could be utilized to accomplish this goal more directly, such as prohibiting or limiting the use of dark pools.

The policies and procedures of the trading venue IEX illustrate the possibility that private-market solutions can midpoint-order exploitation without the need for regulatory intervention. IEX imposes a 350-microsecond delay before transactions can be placed through its matching engine during which it utilizes private-data feeds from all exchanges to check whether the midpoint of the best available quotes has changed and thus ensure that a midpoint order is not executing at a stale NBO or NBB.[184] The inducement for dark pools to implement such reforms is that traders would find the reformed venues more attractive places to send their orders.

---

183.   *See supra* Part V.C.2.

184.   *IEX True Price™ Matching Engine*, IEX TRADING, http://www.iextrading.com/trading/trueprice [http://perma.cc/DTN2-NYEF]. A representative of Goldman Sachs has also said that there is significant client pressure for Goldman to ensure its dark pool has access to private data feeds from exchanges. *See 2015 Workshops: The New Stock Market: Sense and Nonsense*, COLUM. L. SCH. (May 14, 2015), http://web.law.columbia.edu/capital-markets/previous-workshops/2015 [http://perma.cc/67WP-PL4Q].

3. *Measures to prevent electronic front running and other speed-based practices.*   When we combine our ambivalence concerning electronic front running with our clear negative evaluation of slow-market arbitrage and midpoint-order exploitation, and add in the substantial resources that HFTs consume undertaking these three practices, we lean toward favoring reforms that would eliminate HFT advantages in obtaining information ahead of others in the market, if such a reform were relatively low cost. The recent controversy surrounding HFTs—and electronic front running in particular—has resulted in a spate of proposals for addressing their activity, which offer a range of potential benefits and costs.

Two proposals address HFT activity in general by taking aim at high-frequency quoting activity.[185] 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.[186] These fees may simply be a sensible response to HFTs externalizing the cost of bandwidth use. Things get more complicated, however, if the fees are higher than what is necessary to tax this externality and are aimed at reducing practices such as electronic front running. HFTs revise quotes for many reasons other than information they learn from electronic front running and, assuming HFTs are often revising quotes above the 100:1 rate, then fees on quotes in numbers exceeding this limit would create disincentives for all such revisions. Thus, they can be expected to widen spreads and reduce depth because they make it harder for market makers to control adverse-selection and inventory risks through their quoting strategies.[187] That is not a concern if these fees are imposed by an exchange in active competition with other exchanges: customers can decide whether they like the trade-offs implied by the ultimate results. But it would be a concern if mandated by regulation as a way to stop electronic front running.

A second proposal regarding HFT conduct would impose a minimum time in force for quotes, prohibiting them from being canceled, within, for example, 100 milliseconds of submission.[188] Again, HFTs cancel orders faster than 100 milliseconds for reasons other than electronic front running. The costs of such a plan in terms

---

185.   *See* Jones, *supra* note 98, at 42–51.
186.   *Id.* at 45.
187.   *Id.* at 46.
188.   *Id.* at 47.

of liquidity provision could be substantial. It sets a floor on the length of the option offered by liquidity providers to liquidity takers, increasing their chance of being adversely selected and so widening spreads.[189]

Another much-discussed proposal involves altering the current market-trading structure.[190] Stock exchanges currently conduct continuous two-sided (that is buy and sell) auctions for each security. A recent proposal, endorsed by New York Attorney General Eric Schneiderman,[191] is to replace this with frequent batch auctions, say, every 100 milliseconds. Batch auctions would consist of uniform-price, sealed-bid auctions conducted at discrete time intervals. Echoing our own concerns about electronic front running, the proposal argues that the current structure permits frequent technical arbitrage opportunities based on speed, creating a socially wasteful arms race to exploit these opportunities.[192] Frequent batch auctions would eliminate the value of minute speed advantages. Though the proposal may have significant merit, much would depend on implementation. To eliminate electronic front running, every exchange would need to have its auction (nearly) simultaneously. If auctions were sufficiently frequent and at different times at each exchange, then intraexchange exploitation of tiny speed differences might persist, including electronic front running.

We think an approach to ending HFT information speed advantages exists that is simpler both in terms of implementation and in terms of achieving the needed legal changes. None of these three practices would be possible if private data feeds did not make market 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 Regulation NMS scheme, referred to as the "SIP," has publicly disseminated information from all exchanges.

---

189.   *Id.*

190.   *See* Eric B. Budish, Peter Cramton & John J. Shim, *The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response* 6 (Chicago Booth Research, Paper No. 14-03, Feb. 17, 2015), http://ssrn.com/abstract=2388265 [http://perma.cc/937M-8KLE].

191.   *See* Linette Lopez, *New York Attorney General Endorses A Radical Change To The Way The World Trades Stocks*, BUSINESS INSIDER (Mar. 18, 2014, 4:42 PM), http://www.businessinsider.com/schneiderman-endorses-batch-auctions-2014-3 [http://perma.cc/2KC3-8Q7J].

192.   Budish et al., *supra* note 190, at 12–20.

Rule 603(a)(2) of Regulation NMS prohibits exchanges from "unreasonably discriminatory" distribution of market data.[193] In its adopting release for Regulation NMS, the SEC outlined its interpretation of that provision, which is 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."[194] Core data is composed of last-trade reports and each exchange's current highest bids and lowest offers for each security, from which its NBB and NBO is ascertained.[195]

This interpretation of the "unreasonably discriminatory" distribution language of Rule 603(a)(2) appears to say that it is permissible for core-data information to reach an HFT more rapidly than the public recipients of the SIP as long as the signal sending the data to the HFT did not precede the signal sent to the SIP. The exchanges and the HFTs, in agreeing to their co-location arrangements, have assumed this interpretation to be correct. The SEC, in its choice of enforcement actions, has confirmed this interpretation as well. No actions have been brought against co-location arrangements in which the signal sent to HFTs did not precede the signal sent to the SIP. Indeed, the SEC, in a 2010 Concept Release, acknowledged the existence of exchanges' widely known practice of submitting data simultaneously to the SIP and private feeds and that, as a result, private feeds will reach subscribers far faster than the SIP distributes its data.[196] In a 2012 proceeding, however, the SEC found that the 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.[197]

---

193. *See* 17 C.F.R. § 242.603(a)(2) (2015). Section 11A(c)(1) of the Exchange Act authorizes the Commission to regulate market data. 15 U.S.C. § 78k-1(b) (2012).

194. *See* Regulation NMS, 70 Fed. Reg. 37,496, 37,567 & 37,569 (June 29, 2005) (adopting the release).

195. *See* 17 C.F.R. §§ 242.601, 242.602 (2015) (requiring exchanges to report last sales—price and size of the most recent trades—and current best bids and offers); NetCoalition v. SEC, 615 F.3d 525, 529 (D.C. Cir. 2010) (discussing the core data regime).

196. *See, e.g.*, Concept Release on Equity Market Structure, Exchange Act Release No. 61,358, 75 Fed. Reg. 3594, 3601 (Jan. 21, 2010) (discussing how the consolidation processing time of the SIP "means that [private] data feeds can reach end-users faster than the consolidated data feeds.").

197. *See* New York Stock Exchange LLC, Exchange Act Release No. 67,857, 2012 WL 4044880, *1–3, *12 (Sept. 14, 2012).

The language of Rule 603(a)(2) could plausibly be interpreted in a contrary fashion: sending the signal simultaneously to an HFT and to the SIP arguably is an "unreasonably discriminatory" distribution of core data to the end users given that it is predictable that some will consistently receive it faster than others. This interpretation of Rule 603(a)(2)'s language already has its advocates. The market research firm Nanex has repeatedly insisted that the exchanges are in standing violation of Regulation NMS for this reason.[198] Interestingly, this focus on the time at which information reaches end users rather than the time of a public announcement is the approach the courts and the SEC have traditionally taken with respect to when, for purposes of the regulation of insider trading, information is no longer nonpublic.[199] Thus the SEC's ability to alter its interpretation of Rule 603(a)(2) may be the path of least legislative or regulatory resistance to prohibiting electronic front running. One may feel, however, that too much has already been invested in reliance on the SEC's apparent original interpretation for a prohibition to be imposed without the normal procedures of an administrative-agency rule change. If so, then that process can be followed and the rule amended.

B. *HFTs and Volatility*

Overall, we concluded that no evidence exists of a relationship between HFT activities and general increases in market volatility.[200] We conclude as well that the connection between HFTs and episodic volatility is not due to predatory behavior on their parts, but rather their rational withdrawal from the market at certain moments of stress.[201]

Nonetheless, a number of existing proposals seek to address the alleged link between HFT activity and volatility. These proposals fall into two groups: one group seeks to ameliorate trading volatility generally and would incidentally affect HFTs; the second group seeks to target a specific link between HFTs and volatility.

---

198.   *See Direct vs. SIP Data Feed*, NANEX (Apr. 4, 2014), http://www.nanex.net/aqck2/4599.html [http://perma.cc/SY69-GKTS]; *HFT Front Running, All The Time*, NANEX (Sept. 30, 2013), http://www.nanex.net/aqck2/4442.html [http://perma.cc/7XJE-KCU2].

199.   *See, e.g.*, SEC v. Tex. Gulf Sulphur, 401 F.2d 833, 854 (2d Cir. 1968); Investors Mgmt. Co., Exchange Act Release No. 9207, 1971 WL 120502, at *8 (July 29, 1971).

200.   *See supra* Part V.D.1.

201.   *See supra* Part V.D.2.

Consider the first group. Soon after the Flash Crash, the SEC phased in single-stock circuit breakers, which impose a five-minute trading halt if the price of a specific stock moves by more than 10 percent within five minutes.[202] This trading pause is designed to give liquidity providers breathing room to consider whether order imbalances actually reflect information or not.[203] Similarly, the SEC has also approved a "limit up-limit down" plan that pauses trading in a stock if transactions move more than a certain amount, often 5 percent, away from the security's average price over the last five minutes.[204] These are moderate proposals, which should have salutary effects in moderating future crashes.

The second set of proposals tackle the important question of whether market makers, however they be defined, should have stronger affirmative 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.[205] Some commentators want to go further down this road and to impose on HFTs legal responsibilities resembling those imposed on the institutionally designated market makers of yesteryear, such as the specialists of the pre-2005 NYSE.[206]

There is an obvious attraction to proposals that might moderate the flight of liquidity provision from the market during periods of extreme volatility. The historical evidence, however, suggests that

---

202. Order Granting Accelerated Approval to Proposed Rule Change to Amend FINRA Rule 6121, Exchange Act Release No. 62,251, 75 Fed. Reg. 34,183, 34,184 (June 10, 2010); Order Granting Accelerated Approval to Proposed Rule Changes Relating to Trading Pauses Due to Extraordinary Market Volatility, Exchange Act Release No. 62,252, 75 Fed. Reg. 34,186, 34,187 (June 10, 2010).

203. Order Granting Accelerated Approval to Proposed Rule Change to Amend FINRA Rule 6121, 75 Fed. Reg. at 34,186; Order Granting Accelerated Approval to Proposed Rule Changes Relating to Trading Pauses Due to Extraordinary Market Volatility, 75 Fed. Reg. 34,189.

204. Order Approving, on a Pilot Basis, the National Market System Plan to Address Extraordinary Market Volatility, Exchange Act Release No. 67,091, 77 Fed. Reg. 33,498, 33,499, 33,501 n.36 (May 31, 2012).

205. *See* NYSE, DESIGNATED MARKET MAKERS (2014), https://www.nyse.com/publicdocs/nyse/markets/nyse/designated_market_makers.pdf [http://perma.cc/5346-RB9K].

206. KHALDOUN KHASHANAH, IONUT FLORESCU & STEVE YANG, IRRC INST., ON THE IMPACT AND FUTURE OF HFT 14–17 (2014), http://irrcinstitute.org/pdf/HFT-Academic-White-Paper.pdf [http://perma.cc/25DN-JPEB].

strong paper obligations have proved insufficient in the past to motivate market makers to continue supplying liquidity during periods of extreme volatility.[207] Commentators have also noted problems with strengthening such affirmative obligations, emphasizing that any system that requires liquidity providers to take heavy losses during periods of extreme adverse selection must compensate them for doing so at other times. At least two other problems result: first, determining the value of that compensation is extremely difficult, and second, during times of crisis these designated liquidity providers will be the prime targets of informed traders.[208] Thus, we are skeptical about such proposals, especially given our conclusion that the consequences of episodic volatility in terms of wealth transfer and efficiency are not substantial.[209]

C. Dark Pools

We examined two practices associated with an investment bank playing the dual roles of a broker for customers' orders and as an operator of a dark pool. The first potential problem concerns brokers directing customer orders to their own dark pools even when the customer receives inferior execution there. The second concerns brokers ignoring client instructions to direct an order to a specific venue and instead routing it to the broker's own dark pool.

A series of regulatory proposals that seek to rein in the growth of dark-pool volume generally could affect these two practices, but with the added and potentially undesirable effects of affecting access to dark-pool trading in general. These proposals include: eliminating retail investors' access to dark pools and reserving them solely for institutional investors;[210] requiring that dark pools offer price

---

207. *See* Jones, *supra* note 98, at 13–11, 38 ("While some observers suggested greater obligations for market-makers, experience in other rapid downdrafts, including the stock market crash of October 1987, when Nasdaq market-makers and others refused to answer their phones or provide market-making activity, indicates that market-makers will almost always choose to withdraw from the market in the face of such extreme volatility."); *see also* Ian Domowitz, *Take Heed The Lessons From The 1962 Flash Crash*, INFORMATIONWEEK: WALLSTREET & TECH. (June 21, 2010, 5:15 PM), http://www.wallstreetandtech.com/exchanges/take-heed-the-lessons-from-the-1962-flash-crash/a/d-id/1263651? [http://perma.cc/P23Q-56JT] (discussing liquidity problems during a prior flash crash).

208. *See* Angel et al., *supra* note 29, at 33.

209. *See supra* Parts V.D.3, V.D.4.

210. *See* Philip Stafford, *HK Plans Retail Investor Dark Pool Ban*, FIN. TIMES (Feb. 27, 2014, 11:53 PM), http://www.ft.com/intl/cms/s/0/32ab7298-9fba-11e3-94f3-00144feab7de.html#axzz3I1I B1vhn [http://perma.cc/D4AX-CS3W].

improvement before an order can be routed to one;[211] and caps for the volume that can transact on dark venues.[212] All of these proposals seem insufficiently targeted, however, if the problems with dark pools are solely the potential conflicts of interest noted above, rather than more fundamental features.

More targeted and less intrusive would be a reform proposal focused on disclosure, designed to assist customers in determining whether their orders are being routed to venues offering best execution and whether order-routing directions are being ignored. The Financial Industry Regulatory Authority (FINRA) has recently taken steps in the direction of greater disclosure, requesting comment on several new proposed rules.[213] The most relevant proposal would require dark pools to provide FINRA with more extensive order-book information than they currently provide to the Order Audit Trail System (OATS).[214] OATS is an order-tracking system designed to assist FINRA's surveillance activities.[215] Also, 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. Economist James Angel, for example, has called for greater disclosure by brokers, suggesting that "brokerage firms themselves . . . disclose execution quality directly to their customers."[216]

---

211.   *See* Nicole Bullock, *Financial Regulators Probe Dark Pools*, FIN. TIMES (Sep. 16, 2014, 12:07 AM), http://www.ft.com/intl/cms/s/0/50428000-220d-11e4-9d4a-00144feabdc0.html#axzz3F U6kj2LE [http://perma.cc/EM92-86BS].

212.   *See* Philip Stafford & Alex Baker, *Europe Deal Will Cap 'Dark Pools' Trading*, FIN. TIMES (Nov. 21, 2013, 5:26 PM), http://www.ft.com/intl/cms/s/0/ac3ce0f8-52ce-11e3-8586-00144 feabdc0.html#axzz3CAa8BKBz [http://perma.cc/K4LB-DXXS].

213.   *Update: FINRA Board of Governors Meeting*, FINRA (Sept. 19, 2014), https://www. finra.org/Industry/Regulation/Guidance/CommunicationstoFirms/P600807   [http://perma.cc/JN 99-HVY8]; *see* Sarah N. Lynch, *Wall Street Regulator Unveils Proposals for Dark Pools, Bond Markets*, REUTERS (Sept. 19, 2014, 3:21 PM), http://www.reuters.com/article/2014/09/19/us-financial-regulations-finrarules-idUSKBN0HE25G20140919 [http://perma.cc/WLF9-2Q8E].

214.   *Update: FINRA Board of Governors Meeting*, *supra* note 213 ("The Board authorized the publication of a Regulatory Notice requesting comment on a proposal to require alternative trading systems (ATSs) to provide FINRA with order book information that is not currently reported by the ATS to the Order Audit Trail System (OATS), with such information to be reported to FINRA using existing OATS interfaces.").

215.   *See generally* ORDER AUDIT TRAIL SYSTEM (OATS) OATS BASICS, FINRA, http:// www.finra.org/sites/default/files/AppSupportDoc/p016184.pptx [http://perma.cc/TB65-TYPP].

216.   *Role of Regulation Hearings*, *supra* note 13, at 55 (statement of James J. Angel, Associate Professor, Georgetown University McDonough School of Business). Several other commentators have also called for greater disclosure by dark pools. *See, e.g.*, *id.* at 33 (statement of David Lauer, President and Managing Partner, KOR Group LLC) (urging that the SEC

Certainly, more could be done to strengthen the stock market's mandatory disclosure regime. Currently, brokers are not required to disclose to customers on their transaction confirmation slips the venue in which an order was executed.[217] The cost of requiring disclosure of execution venue should not be taxing, as records of where execution occurred must already be retained. Such disclosures would provide customers with the ability to check whether their requests were being followed.

Proposals based on disclosure, however, share at least two vulnerabilities. First, if brokers are submitting inaccurate disclosures, then the SEC or private litigants must feasibly be able to reveal such conduct. Second, disclosure to customers will only be effective if customers are in fact examining and acting on those disclosures, which may not be the case. In an effort to mitigate this problem, the SEC could conduct periodic audits to verify whether routing was being accurately completed.

D. *Payments to Brokers in Connection With Customer Orders*

The last two controversial practices we considered were maker-taker/taker-maker fees and payment for order flow. Each raises principal-agent problems between traders and their brokers.

With regard to maker-taker and taker-maker fees, we saw that with different venues providing different rebates and charging different fees—something we observe in the real world—if brokerage commissions do not vary depending on the venue to which an order is sent, which again in the real world they do not, the broker has an incentive to send nonmarketable limit orders to the venue with the highest rebate and marketable orders to the venue with the lowest fee.[218] In each case, the venue to which the order is sent will probably not provide best execution, a proposition for which there is, as we have seen, some empirical evidence. If customers were perfect monitors of their brokers, this incentive might not matter, but they

---

update Rules 605 and 606 to increase ATS reporting); *id.* at 23 (statement of Tom Wittman, Executive Vice President and Global Head of Equities, NASDAQ OMX) (recommending that brokerage firms should disclose execution quality directly to customers).

217.   Brokers do have limited disclosure requirements under Regulation 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. §§ 242.605–242.606 (2015).

218.   *See supra* Part V.F.2.

are not perfect monitors, especially with respect to limit orders. This leads us to the conclusion that rebates should be passed directly through to customers and fees charged to them, each independent of whatever commission the broker chooses to charge. Such a reform would not guarantee best execution, but it would eliminate an incentive for poor execution.

With regard to payment for order flow, if the market for internalization services and for brokerage services are both sufficiently competitive, internalization with payment for order flow promises retail market order traders as low an effective cost of trading, when brokerage commissions are counted as part of the calculation, as they are going to be able to get.

If one or both of the markets is not fully competitive, however, the practice leads to these traders having higher effective costs of trading than could be achieved by some other arrangement. Remember that we are looking at a situation in which brokers have the ability to segregate out retail order flow, which is uninformed, and match buyers and sellers without the adverse-selection concerns that generate much of the spread on exchanges. So the cost of execution should be very low. Indeed, it is even possible that these traders would be better off if all their trades would be sent to the exchanges, because now the execution of retail market orders execute at prices only slightly better than the NBB or NBO on exchanges that are deprived of internalized order flow. If they had this additional uninformed order flow, the NBB and NBO would be lower.

The key question is whether brokers pass on to customers the substantial payments they receive for order flow in the form of lower commissions given that internalizers only offer nominal price improvement. We do not know the answer to this question, and it is certainly possible that it happens. Still, passing through the payments would solve the problem, if it were a problem. Such a reform should also not be very costly. In essence, this appears to be a situation in which the cure is sufficiently cheap that "if it might be broke, fix it."

CONCLUSION

Over the last two decades, the stock market has been completely transformed. Driven by regulatory change and the information-technology revolution, the structure of the market, the behavior of its participants, and the character of liquidity provision have all undergone dramatic and highly controversial changes. This Article

provides a comprehensive framework for understanding the new stock market and the social impact of the activities occurring within it. We demonstrate the utility of this framework by applying it to eight of the new market's most controversial practices. These practices—which include activities of high-frequency traders such as electronic front running, the behavior of dark-pool operators, and payments by trading venues to brokers in return for sending their customer orders—may seem completely unrelated to each other, but we show that they can all be understood through just three basic mechanisms: adverse selection, the principal-agent problem, and a multivenue trading system.

We come to a number of conclusions concerning the proposals for reform. We agree, for example, with recommendations that brokers should be required to pass through maker-taker fees and payment for order flow to their customers. We disagree, for example, with proposals that HFTs must keep their quotes in force for some minimum amount of time and 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. Yet other proposed reforms involve a trade-off in which an improvement in terms of one worthwhile social goal can only come at a sacrifice of another such goal. In these cases, it may not be obvious whether a reform is, or is not, desirable, but our framework allows for a better understanding of the trade-off involved, makes for a more informed choice, and may point to where further empirical research would be useful. We find this to be the case with, for example, proposals to briefly delay providing HFTs with information concerning new transactions and quotation changes, so HFTs have no advantages over other traders.

In many ways, this Article is just a beginning. As the new stock market continues to evolve, issues arising out of it will not be in short supply. For example questions about whether the minimum stock-price differential ("tick size") should be larger than the current one cent and accusations that HFTs are sending in and cancelling quotes at a high rate of speed simply to clog the system are each issues that have been recently getting increasing attention. Surely many other issues will soon be coming at us from over the horizon. Thus, the framework developed here should have continued utility for some time to come.