BYU Law Review

Volume 2018 | Issue 5 Article 4

Winter 4-1-2019

The Stock Exchange as Multi-sided Platform and the Future of the National Market System

Steven McNamara

The Stock Exchange as Multi-sided Platform and the Future of the National Market System

Steven McNamara[*]

Since Regulation National Market System (Regulation NMS) came into force a decade ago, computer technology has transformed the stock markets. While Regulation NMS benefited investors by lowering stated transaction costs, it also created today's complex and fragmented trading system. An increasing amount of trading now occurs off-exchange in dark pools and other "non-lit" venues, and hidden costs proliferate. In addition to the profits taken by high-frequency traders, these include the defensive costs of the technological arms race, the possibility of another "Flash Crash," public suspicions of "rigged" stock markets, reduced allocative efficiency, and rising proprietary data fees paid by stockbrokers and institutional investors. In prioritizing the goal of competition, Regulation NMS failed to take into account the stock exchange's inherent economic nature as a multi-sided platform and the negative effects of setting the existing exchanges into competition with one another. Furthermore, digital technology undermines a number of Regulation NMS's grounding assumptions. Given the nature of modern stock exchange as a digital multi-sided platform, it is time to reconsider the central limit order book (CLOB) proposals made in the 1970s through the early 2000s. An updated proposal for a "virtual CLOB" would allow the current exchanges to remain in existence, thereby avoiding a single monopoly exchange, while eliminating or mitigating many of the most pressing problems of the current system.

* Visiting Associate Professor, The University of Florida, Levin College of Law; Associate Professor of Business Law, Olayan School of Business, The American University of Beirut. The author may be contacted at sm99@aub.edu.lb. He wishes to thank Robert Rhee, Danny Sokol, William Page, Seth Chertok, and conference participants at the 2017 National Business Law Scholars Conference for helpful discussion and comments during the preparation of the Article. Thanks also to the Levin College of Law for a summer research grant that enabled this work.

CONTENTS

INTRODUCTION

Corporate stocks currently trade on over fifty venues in the United States: thirteen stock markets,[1] classified as "exchanges" under the Securities Exchange Act of 1934, and over forty alternative trading systems (ATSs),[2] which include both dark pools and electronic communication networks (ECNs). Significant amounts of stock also trade in the internal matching engines of large broker-dealers. The highly fragmented equity markets are relatively new, since before 2007 the New York Stock Exchange (the NYSE) accounted for 80% of the daily trading volume in NYSE-listed stocks.[3] Now, no one exchange has more than 15% of the trading volume, and no exchange family more than 25%.[4] The fundamental causes of this rapid change of affairs are the implementation of Regulation National Market System (Regulation NMS) and the technological change of the past twenty years. Regulation NMS is a complicated regulatory compromise that knits together the formerly separate stock exchanges into one national market system yet preserves the independence of the individual exchanges. Under Regulation NMS the exchanges are meant to be competing venues for the consummation of trades.

In fact, Regulation NMS has been largely successful in accomplishing its objective of reducing stated transaction costs for investors.[5] While other factors such as technological innovation and the decimalization of prices share credit for this development, the Order Protection Rule of Regulation NMS does force exchanges to

1. *See NYSE, Nasdaq and . . . ? Get to Know the U.S.'s Stock Exchanges, Part 1*, FINRA (Aug. 17, 2016), www.finra.org/investors/nyse-nasdaq-and-get-know-uss-stock-exchanges-part-1.

2. *See Market Structure: Volume and Liquidity Update*, COWEN ATM (Jan. 23, 2017) [hereinafter COWEN ATM], http://www.cowen.com/wp-content/uploads/2017/05/2017-01-23-Cowen-Volume-Liquidity-Update.pdf (listing forty-three dark pools); *see also* GARY SHORTER & RENA S. MILLER, CONG. RESEARCH SERV., R43739, DARK POOLS IN EQUITY TRADING: POLICY CONCERNS AND RECENT DEVELOPMENTS 3 (2014).

3. *See* James J. Angel, Lawrence E. Harris & Chester S. Spatt, *Equity Trading in the 21st Century*, 5 Q.J. FIN. 1, 5 (2015) (stating that the NYSE had 80% market share in its listed stocks in 2003).

4. *See* John Polise, *A Bridge Too Far: A Critical Analysis of the Securities and Exchange Commission's Approach to Equity Market Regulation*, 11 BROOK. J. CORP. FIN. & COM. L. 285, 332 (2017); COWEN ATM, *supra* note 2.

5. *See* Angel et al., *supra* note 3, at 1.

compete with one another through protecting the best-priced quote currently displayed on a registered exchange.

On the other hand, Regulation NMS shares much of the blame for creating a highly fragmented marketplace. Not only is trading dispersed among thirteen different registered exchanges but an increasing amount occurs off-exchange in unlit, "dark" venues. Dark trading has grown in response to the risks of trading in lit venues, where high-frequency traders are able to exploit timing advantages to the detriment of slower traders. The fragmented marketplace drives another cost, the need for traders to purchase in-depth, proprietary market data from the exchanges that goes beyond what is required to be made publicly available by the securities information processor (the SIP). Rising data fees are paid not only by high-frequency traders but by institutional investors and stockbrokers who feel they need to have such data in order to remain competitive and meet their legal duties to clients. Finally, the rise of high-frequency trading (HFT) in the context of fragmented markets contributes to the risk of market crashes, as was experienced in the "Flash Crash" of May 6, 2010. In sum, while stated transaction costs have diminished, hidden costs from fragmentation, rapidly increasing data fees, and the risk of catastrophic market dislocation have generated significant disquiet on the part of stockbrokers, institutional investors, and advocates for the investing public. The historic mission of the Securities and Exchange Commission (the SEC) to provide for investor protection is therefore implicated in the debate over the current market structure under Regulation NMS.

In considering the way forward, it is important to remember that the Order Protection Rule, the lynchpin of Regulation NMS, is itself a compromise. Rule 611 only protects the quotes at the "top-of-the-book," the best-priced quote in the market at the time. It does not provide "depth-of-book" protection to quotes lower down in an exchange's order book, nor does it provide timing protection to previously entered quotes in accordance with the usual price-time priority rules used at the individual exchanges.[6] The various

6. *See* John McPartland, *Recommendations for Equitable Allocations of Trades in High Frequency Trading Environments* 10–11 (Fed. Reserve Bank of Chi., Policy Discussion Paper, Vol. PDP, No. 2013-01, rev. Jul. 2014). *See generally* LARRY HARRIS, TRADING & EXCHANGES: MARKET MICROSTRUCTURE FOR PRACTITIONERS 116–20 (2003) (rule-based order-matching systems).

proposals discussed in the long run-up to the promulgation of Regulation NMS contain alternatives that should now be reconsidered as the digitization of stock trading exacerbates the problems latent within the governing regulatory regime.

In fact, by the time it came into force in 2007, Regulation NMS was already becoming obsolete.[7] The alternative "hard CLOB" proposal floated during the mid-70s, and which briefly reappeared in 2000, may offer a more attractive market structure in the digital age. A virtual central limit order book (CLOB) functioning as a routing system that provides depth-of-book protection could accomplish two goals: First, it would eliminate many of the specific problems that arise from the interaction of various heterogenous groups of traders in the current highly fragmented system. Second, it could form the basis for the governance structure of a key piece of financial markets infrastructure that could mitigate the danger of digital monopoly, or rather oligopoly, among the exchanges. The nature of the stock exchange as a multi-sided platform (MSP) means that as the exchange industry becomes wholly digital, the pressures winnowing the exchanges to a single exchange, or small group of exchanges, intensify. A CLOB as a routing service, but not a single exchange, would allow the separate incumbent exchanges to remain in existence, providing for a large degree of competition. Furthermore, the governance of such a routing service would allow for the interests of non-exchange stakeholders, primarily but not only institutional investors, to be represented. Ultimately, it seems that the future of the stock exchanges will present public policy challenges similar to those posed by other dominant MSPs in the digital era.

This Article has three parts. Part I reviews the recent changes in the stock market in terms of five implicit assumptions of Regulation NMS disrupted by digital technology. Part II then turns to an examination of recent economic thinking on multi-sided platforms in digital business environments. Since the business of running a

7. *See generally* Chris Brummer, *Disruptive Technology and Securities Regulation*, 84 FORDHAM L. REV. 977, 1032–33 (2015); Christopher Groskopf, *The Modern Stock Market Is a Badly Designed Computer System*, QUARTZ (June 15, 2016), https://qz.com/662009/the-sec-tried-to-fix-a-finance-problem-and-created-a-computer-science-problem-instead ("Latency arbitrage is a byproduct of the fact that Reg NMS neglected to take into account the problems associated with building a distributed computer system.").

financial exchange has, like many others, been transformed into a digital one, many of the economic conditions identified as operative in digital MSPs are increasingly applicable to stock markets. Stock markets have some key differences from other types of MSPs, but they do share some important characteristics, in particular the operation of network effects and the use of differential pricing mechanisms across the sides of the platform. Part III draws on the history of the development of the national market system to formulate a proposal for a router functioning as a virtual CLOB. A virtual CLOB would eliminate many of the most important problems with the current system, while also potentially providing a useful governance regime for a critical piece of financial markets infrastructure.

I. THE OBSOLESCENCE OF REGULATION NMS: FIVE ASSUMPTIONS DISRUPTED BY ALGORITHMIC TRADING

Now appears to be a good time to step back from the controversies of recent years regarding high-frequency trading in the stock markets. The Equity Market Structure Advisory Committee, convened in 2015 by the SEC,[8] has failed to issue any recommendations on market structure issues,[9] and the Tabb Group reports that the profits of high-frequency traders have shrunk radically in the past few years, from a high of approximately $7 billion in 2009, to an estimated $1.1 billion in 2016.[10] This radical decline in profitability would seem to indicate that the scale of the

8. *See* Press Release, Sec. & Exch. Comm'n, SEC Announces Members of New Equity Market Structure Advisory Committee (Jan. 13, 2015), https://www.sec.gov/news/press release/2015-5.html.

9. *See* Memorandum from the EMSAC Regulation NMS Subcommittee to the Equity Market Structure Advisory Committee (EMSAC) (April 3, 2017), https://www.sec.gov/spot light/emsac/emaac-regulation-nms-subcommittee-discussion-framework-040317.pdf ("Subject: Framework for Rule 611 & 610 Discussion"); *see also* Rob Daly, *Trade-Through Review Going Slowly*, TRADERS MAG. (Apr. 7, 2017), http://www.tradersmagazine.com/news/buy side/trade-through-review-going-slowly-116104-1.html.

10. *See* Alexander Osipovich, *High-Frequency Traders Fall on Hard Times*, WALL STREET J. (Mar. 21, 2017, 6:30 AM), https://www.wsj.com/articles/high-frequency-traders-fall-on-hard-times-1490092200; *see also* Gregory Meyer, Nicole Bullock & Joe Rennison, *How High-Frequency Trading Hit a Speed Bump*, FIN. TIMES (Jan. 1, 2018), https://www.ft.com/content /d81f96ea-d43c-11e7-a303-9060cb1e5f44.

problems associated with HFT is shrinking.[11] HFT appears to be a maturing industry in which the total amount of profit is declining and less agile entrants are being driven out or acquired by more powerful ones.[12] This decline in profits may lessen the pressure for a regulatory fix. Nevertheless, the changes wrought by algorithmic trading, some of which are highly problematic, are not going away. A useful way into the controversies surrounding Regulation NMS is to examine a number of its grounding assumptions, which make sense in a world of physical trading environments but break down in the relatively disintermediated world of electronic stock trading.

A. Instantaneous Communication Between Markets Is Possible

A bedrock assumption of Regulation NMS is that communication between market centers will occur instantaneously. The drafters of Regulation NMS were in fact aware of timing issues concerning the operation of the stock markets, but the main distinction they had in mind was that of automated versus manual quotations.[13] Manual quotations are those offered by a floor trader, whereas automated quotations are those in a system such as NASDAQ's, which would be acted on without delay as soon as they were entered into the system. Regulation NMS, however, mostly ignores an entirely different category of timing delays: those arising from the time it takes for automated messages to travel from one market center (i.e., a computer server) to another.[14]

The basic informational architecture of Regulation NMS arises out of the interplay between Rule 611's Order Protection Rule and the various "Plans" that are set up to consolidate and disseminate order information and quotations to the market at large. Rule 611

11. An important caveat to this conclusion is that exchanges and other service providers may simply be arrogating an increased share of the profits to themselves through increased charges for the proprietary data, co-location, and technology services required to compete in the HFT arena. *See* Orçun Kaya, *High-Frequency Trading*, DEUTSCHE BANK RES., (May 24, 2016), https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD0000000000454703/Research_Briefing%3A_High-frequency_trading.PDF.

12. *See* Samuel Agini, *High-Frequency Traders Feel the Pressure of One-Way Markets*, FIN. NEWS LONDON (Mar. 17, 2017), https://www.fnlondon.com/articles/high-frequency-traders-feel-the-pressure-of-one-way-markets-20170317.

13. *See*, *e.g.*, Regulation NMS, Exchange Act Release No. 34-51808, 70 Fed. Reg. 37,496, 37,519–20, 37,531 (June 29, 2005).

14. *See infra* note 21.

prohibits "trade throughs" of all "protected quotations" in Regulation NMS stocks.[15] This means that a trading center is required to have policies in place that prevent a trade from being consummated on its facility when a "protected quotation,"[16] an automated quotation at a better price on a registered exchange, is displayed elsewhere in the national market system. In this situation, the order is instead to be routed to that other exchange.

In order for this system to work, one or more SIPs[17] are required to act as central information processors. The SIPs receive and collate the quotes from each exchange and then disseminate the best ones as the national best bid/offer[18] (the NBBO) to the market at large. Rules 602 and 603 govern this activity. Rule 602 requires the exchanges to "make available to vendors the best bid, best offer, and aggregate quotation sizes" for each listed security.[19] Rule 603(b) in turn requires the exchanges to "act jointly pursuant to one or more effective national market system plans" to disseminate "all consolidated information for an individual NMS stock through a single plan processor."[20] These rules call for the basic informational architecture of Regulation NMS, consisting of a central information processor (the SIP) that collates the information it receives from the various market centers and then redistributes it to the market at large as the NBBO.

It gradually became apparent after Regulation NMS came into force that trading conducted within the timing gaps between the market centers, and between the market centers and the SIP, was

15. *See* 17 C.F.R. § 242.611(a)(1) (2018).

16. *See id.* § 242.600(b)(58) (defining "protected quotation").

17. *See* 15 U.S.C. § 78(c)(22)(A) (2012) (defining "securities information processor"). This definition was added to the 1934 Act by the Securities Acts Amendments of 1975. The SIP was seen by the 1975 Amendments as the mechanism that would allow the accomplishment of its third objective, to assure "the availability to brokers, dealers, and investors of information with respect to quotations for and transactions in securities." *Id.* § 78k-1(a)(1)(C)(iii).

18. *See* 17 C.F.R. § 242.600(b)(42) ("National best bid and national best offer means, with respect to quotations for an NMS security, the best bid and best offer for such security that are calculated . . . on a current and continuing basis by a plan processor pursuant to an effective national market system plan.").

19. *Id.* § 242.602(a)(1)(i).

20. *Id.* § 242.603(b).

important.[21] Regulation NMS, however, sidesteps the issue of these timing gaps.[22] During the gestation of Regulation NMS, the main focus of regulators as far as timing was concerned centered on the interplay between manual and automated quotations. Within the decade, however, it became apparent that significant activity occurred within periods that were considerably shorter than the periods of three or ten or thirty seconds that the debate over integrating manual quotations into the automated system considered. These gaps consisted of periods measured in milliseconds (thousandths of a second) and even microseconds (millionths of a second), from one second down to two milliseconds or even less.[23] The inevitable delay between the NBBO as reported by the SIP and what may be the best price at the exchanges prompted financial sociologist Donald MacKenzie to characterize Regulation NMS as assuming a Newtonian world, whereas the contemporary markets are Einsteinian, "in which the time that communication takes, even

21. *See, e.g.*, *An Old SIP in a Modern Market*, TRADERS MAG. (Oct. 1, 2013), http://www.tradersmagazine.com/issues/26_356/An-Old-SIP-In-a-Modern-Market-111619-1.html. Michael Lewis's *Flash Boys* (2014) brought the problem of the staleness of the SIP to public attention, but protagonist Brad Katsuyama (now CEO of IEX) began grappling with the problems of negotiating the latencies between the various exchange servers in 2007. MICHAEL LEWIS, FLASH BOYS 34 (2014). *Wall Street Journal* reporter Scott Patterson also explored these problems in a 2013 book. SCOTT PATTERSON, DARK POOLS: THE RISE OF THE MACHINE TRADERS AND THE RIGGING OF THE U.S. STOCK MARKET (2013).

22. The main discussion of timing differences down to the sub-second level occurs in the context of automated vs. manual quotations. *See* Regulation NMS, 70 Fed. Reg. 37,496 (June 29, 2005). The discussion of the "flickering quotation" exception to the trade-through rule does evidence awareness of the problem of timing gaps but concludes with the statement that "[t]he Commission also notes that opportunities for arbitrage between trading centers displaying different prices for the same NMS stock would exist irrespective of whether the Commission adopted an order protection rule . . . ." *Id.* at 37,523; *see also id.* at 37,528 (acknowledging that intermarket price protection without an opt-out exception may interfere with "extremely short-term trading strategies"); *id.* at n.215 (discussing concerns over "'clock drift' and time lags between different data sources.").

23. *See An Old SIP in a Modern Market*, *supra* note 21 (estimating the latency of the SIP as 800 milliseconds in 2006). In Lewis's *Flash Boys*, message transmission times in 2009 from lower Manhattan to the exchange servers in New Jersey range from four to seven milliseconds. *See* LEWIS, *supra* note 21, at 71. Transmission times from co-located servers at the exchanges are in the 0.2- to 0.35-millisecond range. *See Deutsche Boerse Group's Co-location Service in Cooperation with Equinix*, EUREX EXCHANGE, www.eurexchange.com/exchange-en/technology/co-location-services (last visited Oct. 17, 2018). Processing times at the SIPs are now as low as an astonishing twenty microseconds (0.020 milliseconds). *See* Nicole Bullock & Philip Stafford, *Nasdaq Upgrade Raises Market Data Questions*, FIN. TIMES (Oct. 23, 2016), https://www.ft.com/content/0de0a15c-9933-11e6-8f9b-70e3cabccfae.

at the speed of light, is salient."[24] Regulation NMS implicitly presumes that instantaneous communication between market centers is possible, whereas the development of high-speed computing technology means that gaps of much less than one second are very meaningful.

B. Limit Orders Are Good, and Regulation Should Operate to Protect Them

A second assumption is that limit orders are good and that regulation should operate to protect them.[25] While order types have mushroomed in the past decade, limit orders and market orders are the two basic order types used by equities traders, as well as the building blocks for more complex types. A market order is a simple command to a broker to go into the market and purchase (or sell) a given number of shares at the prevailing best price. A limit order, on the other hand, is an order offering to buy (a bid) or sell (an offer) a given amount at a given price.[26] Assume that the NBBO for General Motors (GM) stock is currently $42.50 (bid) and $42.55 (offer). A trader wishing to sell GM, but who thought the current best price to buy was too low, could submit a new limit order to sell 1000 shares at $42.53. This limit order would now establish a new best offer of $42.53, and the order would remain open until either a buyer responded to this limit order and "hit" the quote, or the trader cancelled it.

Limit orders play several important roles. Most importantly, they set the prevailing prices. By announcing a price at which traders are willing to buy or sell, limit orders are the basis for market orders, since traders submitting a market order usually do so on the basis of the current NBBO. And the combined limit orders in the various order books of the exchanges constitute the "depth"

24. Donald MacKenzie, A Sociology of Algorithms: High-Frequency Trading and the Shaping of Markets 41–42 (June 2014) (unpublished manuscript), http://www.sps.ed.ac.uk/__data/assets/pdf_file/0004/156298/Algorithms25.pdf.

25. *See* Regulation NMS, 70 Fed. Reg. at 37,505 (discussing commenters who "stress[] that limit orders are the cornerstone of efficient, liquid markets and should be afforded as much protection as possible"). For a critical appraisal, see *Hearing on Equity Market Structure: A Review of SEC Regulation NMS Before the Subcomm. on Capital Mkts. & Gov't Sponsored Enters. of the H. Comm. on Fin. Servs.*, 113th Cong. (2014) [hereinafter Lofchie] (written testimony of Steven Lofchie).

26. *See* Michael Morelli, *Regulating Secondary Markets in the High Frequency Age: A Principled and Coordinated Approach*, 6 MICH. BUS. & ENTREPRENEURIAL L. REV. 79, 82 n.81 (2016).

of liquidity in a particular stock. This pool of liquidity is one of the central reasons for forming a stock market, as greater liquidity in a stock greatly increases its value to traders, who often wish to exit an investment quickly and at the highest price available. A deep pool of liquidity makes it possible to do so. Not only does it signal a willing buyer at a pre-established price but depth of liquidity ensures that an individual small purchase or sale will have a small-to-minimal effect on the prices of the remaining shares. Finally, a limit order itself is a type of option: it provides other traders the opportunity to make a transaction at a set price for as long as the quote remains on the exchange's order book.[27] Today, it is the limit order's function as an option that threatens its economic viability.

Posting a limit order always entailed the risk that another trader would see a bid as too high or an offer as too low. This judgment could be based on fundamental analysis, inside information, short-term predictions of market momentum, or other factors. But whatever the reason, if such a judgment were correct, in the zero-sum game of trading the person posting the limit order would suffer a loss if her quote were hit. The risk for a trader submitting a limit order (or conversely, its value to the market at large) is a function of three factors: the discrepancy of the price in the order versus the "true" value of the stock, the number of shares at stake, and the amount of time the order remains open.[28] The rise of HFT appears to have increased the average magnitude of the first and third factors. High-speed computing technology greatly increases the ability of high-frequency traders to find and act on information, including price momentum information. This heightens the adverse selection risk associated with limit orders.[29] It also appears that in the new world of HFT, the effective measure of time is greatly sped up as compared to the old manual markets. Now, a far greater volume of orders is entered per second, with most quickly cancelled, than in the past.[30] Since the "volume clock" of the current markets is running so much faster than it used to, submitting limit

27. *See* Regulation NMS, Exchange Act Release No. 51808, 70 Fed. Reg. 37,496, 37,526–27 (June 29, 2005).

28. *See* Lofchie, *supra* note 25 at 6.

29. *See* Merritt B. Fox, Lawrence R. Glosten & Gabriel V. Rauterberg, *The New Stock Market: Sense and Nonsense*, 65 DUKE L.J. 191, 238–42 (2015) (discussing "slow-market arbitrage").

30. *See* David Easley, Marcos M. López de Prado & Maureen O'Hara, *The Volume Clock: Insights into the High-Frequency Paradigm*, 39 J. PORTFOLIO MGMT. 19 (2012).

orders in an old-fashioned way, without the wherewithal to monitor all markets and information sources and quickly pull the order if negative information is spotted, becomes significantly riskier than before. Because of these changes in how the stock markets function, the risk inherent in limit orders has increased.

The natural response of many non-HFT traders has been to migrate away from the lit markets. This is largely because dark pools do not publicly post quotes for all to see and react to. By not posting the quote, any information entailed in the limit order is not revealed to the market at large, although there do appear to be ways to access it.[31] Furthermore, a number of dark pools have marketed themselves as inhospitable to HFT, or as actively preventing HFT shops from trading in their pools.[32] Even though such claims have not always been truthful, dark pools do function as "protective coves," as opposed to the "naked bazaars" of the lit markets.[33] The growth of dark trading will be explored in greater detail in section II.B, but at this point it is important to realize that it is driven by the costs, including risk, of trading the old-fashioned way with limit orders on an exchange. The grounding assumption of Regulation NMS, that resting limit orders are good and that regulation should serve to protect them, may remain true in a normative sense, but increasing adverse selection risk undermines the economic basis many traders previously had for using them.

C. The Securities Information Processor (the SIP) Presents Useful Information

A corollary to the first assumption of instantaneous communication between markets is that the SIP presents useful information. The SIP is the market institution responsible for collecting,

31. *See* Rob Curran, *Watch Out for Sharks in Dark Pools*, WALL STREET J. (Aug. 19, 2008, 11:59 PM), https://www.wsj.com/articles/SB121911298392752051. *See generally* Gregory Scopino, *The (Questionable) Legality of High-Speed "Pinging" and "Front Running" in the Futures Markets*, 47 CONN. L. REV. 607 (2015).

32. *See* Scott Patterson & Bradley Hope, *Barclays Dark Pool Drew Early Alarms*, WALL STREET J. (July 20, 2014, 7:41 PM), https://www.wsj.com/articles/barclays-dark-pool-drew-early-alarms-on-fast-trading-1405898592; Press Release, Sec. & Exch. Comm'n, Barclays, Credit Suisse Charged with Dark Pool Violations (Jan. 31, 2016), https://www.sec.gov/news/pressrelease/2016-16.html.

33. *See* Lofchie, *supra* note 25, at 5.

collating, and distributing the quotes that make up the NBBO.[34] Regulation NMS conceives of the SIP as the central informational node in the market system, which should operate to place all traders and investors on an even playing field, as well as to prevent exploitation of traders through the exercise of market power over the valuable market data by the exchanges. A crucial assumption of Regulation NMS therefore is that the information presented by the SIP is useful to market participants. As with the second assumption above, the failure of this assumption is a matter of degree—it is not entirely false, nor is it entirely true.

The SIP is the descendant of the "consolidated tape" which provided transaction data to the market at large in the pre-computerized era.[35] While it was defined in the 1975 Amendments,[36] it has been woven into the fabric of Regulation NMS. Rules 602 and 603 concern the provision and dissemination of quotations in Regulation NMS stocks. Rule 602 requires that an exchange collect quotes from broker-dealers and make those quotes available to "vendors."[37] Rule 603 mandates that each exchange act pursuant

34. There are presently three SIPs for cash equities: Tape A (covering NYSE-listed stocks), Tape B (covering stocks, structured products, and ETFs listed on neither the NYSE or NASDAQ), and Tape C (covering NASDAQ-listed stocks). The Tape A and Tape B SIPs are governed by the Consolidated Tape Association, a consortium of exchanges, and managed by the NYSE. The Tape C SIP is governed and administered by NASDAQ. *See* Larry Tabb, *Latency Arbitrage and the Problem with the SIP*, TABBFORUM (July 19, 2016), https://tabbforum.com/opinions/latency-arbitrage-and-the-problem-with-the-sip. The OPRA SIP covers exchange-traded securities options. *See OPRA Overview*, OPRA, https://opradata.com/overview/opra_over.jsp (last visited Dec. 30, 2018).

35. *See* Jerry W. Markham, *High-Speed Trading on Stock and Commodity Markets – From Courier Pigeons to Computers*, 52 SAN DIEGO L. REV. 555, 601 n.284 (2015); Dale A. Oesterle, *Regulation NMS: Has the SEC Exceeded Its Congressional Mandate to Facilitate a "National Market System" in Securities Trading?*, 1 N.Y.U. J.L. & BUS. 613, 630–35 (2005).

36. Section 3(6) of the Securities Acts Amendments of 1975 added section 3(a)(22) to the Exchange Act:

> The term "securities information processor" means any person engaged in the business of (i) collecting, processing, or preparing for distribution or publication, or assisting, participating in, or coordinating the distribution or publication of, information with respect to transactions in or quotations for any security (other than an exempted security) or (ii) distributing or publishing (whether by means of a ticker tape, a communications network, a terminal display device, or otherwise) on a current and continuing basis information with respect to such transactions or quotations.

15 U.S.C. § 78c(a)(22)(A).

37. The term "vendor" here refers to a securities information processor. *See* Regulation NMS, Exchange Act Release No. 34-51808, 70 Fed. Reg. 37,496, 37,574 n.688 (June 29, 2005).

to "one or more . . . national market system plans to disseminate consolidated information, including a national best bid and best offer, on quotations for and transactions in national market stocks."[38] A SIP is established under a "national market system plan."[39] Rule 603(a)(1) provides that "[a]ny exclusive processor, or any broker or dealer with respect to information for which it is the exclusive source, that distributes information . . . to a securities information processor shall do so on terms that are fair and reasonable."[40] Rule 603(a)(2) requires that "[a]ny national securities exchange, . . . broker, or dealer that distributes information with respect to quotations for or transactions in an NMS stock to a securities information processor . . . shall do so on terms that are not unreasonably discriminatory."[41] The SIP is therefore the entity under Regulation NMS that is charged with collating and distributing this information to the market at large. Also important is the distinction between an exclusive source of information providing it on terms that are "fair and reasonable," and all other providers doing so on terms that are "not unreasonably discriminatory."[42]

The development of high-speed trading led to great reductions in the latency (the time it takes for messages to be transmitted) of messages to and from the servers of the exchanges, which has had the corresponding effect of making the NBBO as communicated by the SIP increasingly "stale."[43] In addition to the simple effect of decreases in message transmission times increasing the relative significance of processing times at the SIP, co-location services and the proprietary, enriched data feeds offered by the exchanges have contributed greatly to its obsolescence. Co-location services allow HFT firms to rent space in the facilities that host an exchange's servers, so messages will reach an exchange server in the least

38. 17 C.F.R. § 242.603(b) (2018).
39. *Id.* § 242.600(b)(43).
40. *Id.* § 242.603(a)(1).
41. *Id.* § 242.603(a)(2).
42. *Id.* § 242.603(a); *see also* Fox et al., *supra* note 29, at 270–71.
43. *See* Concept Release on Equity Market Structure, Exchange Act Release No. 34-61358, 75 Fed. Reg. 3593, 3601 (Jan. 21, 2010) (estimating latency of the SIP at approximately five milliseconds); Shengwei Ding, John Hanna & Terrence Hendershott, *How Slow Is the NBBO? A Comparison with Direct Exchange Feeds*, 49 FIN. REV. 313, 320 (2014) (estimating latency of the SIP as approximately 1.5 milliseconds in 2012).

possible amount of time.[44] Co-location is the most obvious example of traders positioning themselves in a way that makes the SIP stale, but proprietary data feeds and the SEC's interpretation of Rule 603(a)(2) compound its effects. Proprietary data feeds are delivered directly to co-located servers. They provide a greater depth of information than the SIP, whose NBBO just presents the best bid and offer nationally. Proprietary data feeds show an exchange's complete "book" in each stock, i.e., all the limit orders going deeper into the book away from the best prices.[45] Order cancellation information is also available, which allows high-frequency traders the ability to draw a picture of the direction of the market in a particular stock. The deeper and more detailed information offered by the proprietary data feeds is essential for HFT, which requires a more complete picture of the market to engage in many short-term trading strategies.[46]

A key regulatory factor for proprietary data feeds is Rule 603(a)(2), which states that an exchange "that distributes information with respect to quotations for or transactions in an NMS stock to a securities information processor, broker, dealer, or other persons shall do so on terms that are not unreasonably discriminatory."[47] The SEC interprets this to allow an exchange to send information from its server to co-located servers, typically of high-frequency traders, simultaneously with its transmission to the SIP.[48] This may have seemed sensible in 2005, but with the development of high-speed transmission technology, and the

44. *See* Brummer, *supra* note 7, at 1030.

45. *See* SAL ARNUK & JOSEPH SALUZZI, BROKEN MARKETS: HOW HIGH FREQUENCY TRADING AND PREDATORY PRACTICES ON WALL STREET ARE DESTROYING INVESTOR CONFIDENCE AND YOUR PORTFOLIO 111–17 (2011). Regulation NMS rejected a depth-of-book requirement for the SIP, stating that the basic information of prices, sizes, and market identifiers for the NBBO would be sufficient for retail investors, though the SEC also foresaw other traders purchasing deeper information from market centers. *See* Regulation NMS, Exchange Act Release No. 34-51808, 70 Fed. Reg. 37,496, 37,569 (June 29, 2005).

46. *See* IRENE ALDRIDGE, HIGH-FREQUENCY TRADING: A PRACTICAL GUIDE TO ALGORITHMIC STRATEGIES AND TRADING SYSTEMS 53–74 (2d ed. 2013) (discussing the use of both Level I and the deeper Level II categories of data by HFT'ers).

47. 17 C.F.R. § 242.603(a)(2).

48. Regulation NMS, Exchange Act Release No. 34-51808, 70 Fed. Reg. 37,496, 37,567 (June 29, 2005) ("Stated another way, adopted 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."); *see also* Fox et al., *supra* note 29, at 269–71.

corresponding lack of technological development at the SIP,[49] it meant that by the time messages were received at the SIP, processed there, and then released to the market at large, the NBBO was significantly stale.[50] The provision of enriched data feeds to co-located servers greatly reduces the value of the information offered by the SIP, creating a two-tiered marketplace. High-speed traders pay extra to be at the front of the line, while those relying on the NNBO as communicated through the SIP are in back.

There are indications that the problems associated with the SIP may be lessening, however. In 2016, NASDAQ upgraded the technology of the SIP it operates, reducing latency times from more than one millisecond to just over half of a millisecond today.[51] Reductions in this latency time are important, although in principle there will always be some delay in the NBBO as reported by the SIP in our current system. Improvements at the SIP could also be a contributing factor in the declining profits of HFT firms and are important for the effective operation of the national market system. But there will always be an irreducible delay in the propagation of the SIP, so to some extent the assumption that the SIP presents valuable information will always be somewhat wrong. The fact that many traders are willing to pay substantial amounts for co-location and proprietary data feeds casts doubt on this assumption.

D. Sufficient Incentives Exist for Market Makers to Supply Liquidity

A fourth assumption of Regulation NMS is that sufficient incentives exist for market makers to provide liquidity. In the old NYSE "specialist system," where specialists were granted a monopoly position as market maker in the stocks they covered, there was sufficient economic incentive to act as a market maker in times of market stress. The demise of the specialist system and its

49. *See An Old SIP in a Modern Market*, *supra* note 21.

50. *See supra* note 43 and accompanying text.

51. Although processing times at the upgraded SIP are in the 50-microsecond range, including transmission times to other market centers would place the latency of the SIP in the 500-microsecond range. *See* Ivy Schmerken, *Nasdaq OMX Won Over SIP Committee with Latency Reductions & Tech Upgrades*, INFORMATIONWEEK WALLSTREET & TECH. (Nov. 7, 2014, 12:35 PM) http://www.wallstreetandtech.com/infrastructure/nasdaq-omx-won-over-sip-committee-with-latency-reductions-and-tech-upgrades/d/d-id/1317316.html; Tabb, *supra* note 34.

replacement by "Designated Market Makers" (or DMMs) (as well as "Supplemental Liquidity Providers" or SLPs) illustrates the difficulty of using the old regulatory categories in the new world of digital markets. A related concern is the health of the trading "ecosystem" in this new world, specifically the diversity of players in the system and the decline in the numbers of various players, particularly broker-dealers.

In the NYSE's specialist system, specialists were granted the right to see the order book in a particular stock.[52] While they were not supposed to trade on it, it is hard to believe that this information did not constitute a valuable advantage in their market-making activities.[53] In exchange for advance knowledge of the order book in the stocks it covered, the specialist assumed the obligation to make a market in times of market stress, to be willing to buy those stocks when investors would flee, and to sell when they all wanted to buy. Specialists were therefore charged with standing on the opposite side of the market from the general direction of the herd. Economically, the positional advantages of being a specialist appear to have more than compensated them for this role, and the specialist system endured until the NYSE was thoroughly dominated by electronic traders.[54] The specialist system at the NYSE was phased out in 2008 and replaced with the new category of "Designated Market Maker."[55]

Since then, the market has struggled to find the appropriate balance of positional advantage and affirmative obligations for DMMs. The signal event in this quest was the "Flash Crash" of

52. *See generally* HARRIS, *supra* note 6, at 494–513; JOEL SELIGMAN, THE TRANSFORMATION OF WALL STREET: A HISTORY OF THE SECURITIES AND EXCHANGE COMMISSION AND MODERN CORPORATE FINANCE 335–44 (3d ed. 2003).

53. *See* CHARLES R. GEISST, WALL STREET: A HISTORY 223 (1997); *see also* Stanislav Dolgopolov, *Regulating Merchants of Liquidity: Market Making from Crowded Floors to High-Frequency Trading*, 18 U. PA. J. BUS. L. 651, 667–68 (2016) (describing both the privileges and obligations of market makers in the pre-Regulation NMS era); Hans R. Stoll, *Electronic Trading in Stock Markets*, 20 J. ECON. PERSP. 153, 158–59 (2006) (detailing enforcement actions against specialists abusing their privileges).

54. *See* Dolgopolov, *supra* note 53, at 662–67 (analyzing the economic forces affecting the provision of liquidity in a specialist system).

55. *See* Order Approving a Proposed Rule Change to Create a New NYSE Market Model, Exchange Act Release No. 34-58845, 73 Fed. Reg. 64,379 *passim* (Oct. 24, 2008); *see also* Evangelos Benos & Anne Wetherilt, *The Role of Designated Market Makers in the New Trading Landscape*, 2012(Q4) BANK OF ENG. Q. BULL. 343 (2012).

May 6, 2010, when the stock market lost over a trillion dollars in approximately thirty-six minutes.[56] Subsequent smaller crashes have also occurred.[57] These sudden crashes reveal the susceptibility of the stock markets to large and sudden swings in prices. The failure demonstrated here is that the predominant liquidity providers in the current market, high-frequency traders, are skittish, cancelling their quotes and exiting as soon as they detect a whiff of market instability. From their point of view, such behavior is necessary in order to protect against standing on the wrong side of the market and suffering significant losses.[58] From the perspective of the market as a whole, the provision of liquidity by electronic traders lacks the incentives of the old specialist system, which tended to mitigate market swings. This has led to complaints about "phantom liquidity," liquidity that appears on the order books of the exchanges but is liable to vanish at the first signs of stress.[59]

In the current system, the DMMs and SLPs should function as a backstop of liquidity provision when high-frequency traders withdraw from the market. In the Flash Crash and many smaller "mini flash crashes," they do not appear to have taken up this role. The current system at the NYSE, as well as many other exchanges, relies on DMMs to support trading by offering them certain benefits, but not an advance look at trades coming into the stock

56. *See* COMMODITY FUTURES TRADING COMM'N & SEC. & EXCH. COMM'N, FINDINGS REGARDING THE MARKET EVENTS OF MAY 6, 2010 (2010) [hereinafter CFTC & SEC FINDINGS].

57. *See* Maureen Farrell, *Mini Flash Crashes: A Dozen a Day*, CNN MONEY (Mar. 20, 2013, 1:56 AM), https://money.cnn.com/2013/03/20/investing/mini-flash-crash/.

58. *See, e.g.*, CFTC & SEC FINDINGS, *supra* note 56, at 48 ("In general . . . it appears that the 17 HFT firms traded with the price trend on May 6 and, on both an absolute and net basis, removed significant buy liquidity from the public quoting markets during the downturn.").

59. *See, e.g.*, *The Cost of Phantom Liquidity*, THEMIS TRADING BLOG (Aug. 14, 2012), http://blog.themistrading.com/2012/08/the-cost-of-phantom-liquidity/; *see also* Graham Partington, Richard Philip & Amy Kwan, *Is High Frequency Trading Beneficial to Market Quality?* (Ctr. Int'l Fin. & Regulation, Working Paper No. 083, 2015) (arguing that HFTs, more often than not, supply liquidity on the "thick" side of the book where it is needed least). *But see* Jesse Blocher, Rick Cooper, Jonathan Seddon & Ben Van Vliet, *Phantom Liquidity and High Frequency Quoting*, 11 J. TRADING 6 (2016) (analyzing NASDAQ data to argue that "phantom liquidity" is merely HFT firms continually adjusting their quotes); Avanidhar Subrahmanyam & Hui Zheng, *Limit Order Placement by High-Frequency Traders*, 16 BORSA ISTANBUL REV. 185, 205 (2016) (finding that HFT increases liquidity provision in times of high volatility).

exchange.[60] Instead, DMMs can trade at parity with those orders but are provided special quoting and fee rebates, particularly for less active securities.[61] These incentives do not always appear to have been enough to prompt them to step into the breach.[62] The root cause of the problem appears to be the very unforgiving and tightly coupled environment created by algorithmic trading and low latency times. Any disadvantageous position will be quickly and ruthlessly converted into a gain by a trader on the opposite side of the market. In the old system, with broker-dealers on the floor of an exchange, a specialist could stand by a position and, presumably, other traders would trade against it. The limited number of traders on the floor and the timing gaps between outside information reaching floor traders, however, seem to have provided a natural limit to how much a liquidity-providing stance could be exploited. This in turn lent a certain measure of "stickiness" to the market. In a digital environment, these natural limits are greatly lessened, if not eliminated entirely. Digital markets challenge the assumption that market makers have sufficient incentive to supply liquidity in all market conditions.

E. The Stock Exchanges Are Properly Positioned to Be Self-Regulatory Organizations

A fifth assumption under considerable strain is that the exchanges are properly positioned to be self-regulatory organizations (SROs). While the factors undermining this role are complex, and not entirely attributable to technological change, the complaint that

60. In the current electronic market, an advance look at the limit order book wouldn't even make sense. Nowadays, traders purchasing proprietary data feeds have knowledge of such orders as soon as they come into the exchange and can act on these orders or modify their own orders based on this information. In the old specialist system, the specialists were the only participants with such knowledge. *See* Terrence Hendershott & Pamela C. Moulton, *Automation, Speed, and Stock Market Quality: The NYSE's Hybrid*, 14 J. FIN. MARKETS 568, 570–71 (2011).

61. *See* Benos & Wetherilt, *supra* note 55; *Designated Market Makers*, NYSE (2018), https://www.nyse.com/publicdocs/nyse/listing/fact_sheet_dmm.pdf.

62. *See* JOINT CFTC-SEC ADVISORY COMM. ON EMERGING REGULATORY ISSUES, RECOMMENDATIONS REGARDING REGULATORY RESPONSES TO THE MARKET EVENTS OF MAY 6, 2010, at 10 (2011), http://www.sec.gov/spotlight/sec-cftcjointcommittee/021811-report.pdf. *But see* Adam D. Clark-Joseph, Mao Ye & Chao Zi, *Designated Market Makers Still Matter: Evidence from Two Natural Experiments*, 126 J. FIN. ECON. 652 (2017).

the exchanges privilege their best customers at the expense of others goes to the heart of the problems with the current system. The SRO system envisioned in the Exchange Act, as well as the 1975 Amendments, and even Regulation NMS, springs from the world of open outcry trading with trading rooms, specialists, and floor brokers, as well as off-exchange traders.

The 1934 Act classifies any "national securities exchange" as an SRO.[63] An SRO, in turn, is tasked with enforcing compliance with the 1934 Act and its regulations, as well as the exchange's own rules among its members.[64] Section 6(b)(5) outlines the objectives an exchange's rules must advance:

> The rules of the exchange are designed to prevent fraudulent and manipulative acts and practices, to promote just and equitable principles of trade, to foster cooperation and coordination with persons engaged in regulating, clearing, settling, processing information with respect to, and facilitating transactions in securities, to remove impediments to and perfect the mechanism of a free and open market and a national market system, and, in general, to protect investors and the public interest; and are not designed to permit unfair discrimination between customers, issuers, brokers, or dealers, or to regulate by virtue of any authority conferred by this title matters not related to the purposes of this title or the administration of the exchange.[65]

These more general goals of the SRO system made sense in terms of the trading environment of the time. Exchanges were cooperative institutions originally formed by the stockbrokers themselves, and up until the 1990s, and 2006 in the case of the NYSE, remained member-owned institutions. Most importantly, a stock exchange possessed significant market power, if not outright monopoly, either in its geographical region or with respect to the particular securities traded on it.[66] This market power gave it great authority over its members, as the threat of enforcement and

63. Securities Exchange Act of 1934 § 3(a)(26), 15 U.S.C. § 78c(a)(26) (2012).
64. Securities Exchange Act of 1934 § 6(b)(1), 15 U.S.C. § 78c(a)(3)(A).
65. Securities Exchange Act of 1934 § 6(b)(5), 15 U.S.C. § 78f(b)(5).
66. *See* Onnig H. Dombalagian, *Demythologizing the Stock Exchange: Reconciling Self-Regulation and the National Market System*, 39 U. RICH. L. REV. 1069, 1091 (2005).

ultimately expulsion could deprive a member of lucrative trading privileges.

The shift from cooperative ownership to the demutualized, corporate stock exchange fundamentally upset this balance. While there were numerous instances of exchanges failing to properly police their members in the pre-corporate era[67] (with the resultant harms falling on the investing public), the interests of the owners in a cooperative exchange are aligned with those of its primary customers, the broker-dealers, since these are the same parties.[68] A cooperative exchange possessing market power, whose option of expulsion amounted to a significant threat that the exchange could exercise with little harm to itself, would have an interest in regulating its members for the larger success of the entire cooperative.[69]

In the case of a corporate exchange, however, the interests of the owners and the customers are not aligned. Here, the stockholders will only favor actions that contribute directly to profitability, while the brokers will be mainly interested in paying as low a fee as possible.[70] The expense of regulation will come to seem a cost that contributes to neither party's immediate goals, and any longer-term benefit to a strong regulatory function may seem too remote to motivate significant devotion of time and resources. Compounding the effects of demutualization is Regulation NMS, which sets the exchanges in direct competition with one another for order flow. Now, the exchanges are incentivized to cater to their customers, particularly their best ones, the high-frequency traders, not to enforce irksome discipline. In this environment, regulation amounts to a public good that all financial market participants benefit from, but no one exchange wants to fund.[71] Individual exchanges are therefore tempted to free-ride on the regulatory efforts of others, partaking in the general benefits of strong

67. *See, e.g.*, SELIGMAN, *supra* note 52, at 335–44.

68. *See* Andreas M. Fleckner, *Stock Exchanges at the Crossroads*, 74 FORDHAM L. REV. 2541, 2591 (2006).

69. *See* David S. Evans, *Governing Bad Behavior by Users of Multi-sided Platforms*, 27 BERKELEY TECH. L.J. 1201, 1231–35 (2012) [hereinafter Evans, *Governing Bad Behavior*].

70. *See* Yesha Yadav, *The Failure of Liability in Modern Markets*, 102 VA. L. REV. 1031, 1091–92 (2016).

71. *See* Jonathan R. Macey & Maureen O'Hara, *From Markets to Venues: Securities Regulation in an Evolving World*, 58 STAN. L. REV. 563, 576 (2005).

regulatory enforcement while shirking the responsibility of paying for it. In this new environment, the stiff competition for order flow operates in conjunction with the demutualized, corporate status of the exchanges to place regulatory efforts on the back burner.

The desperate competition for order flow in fact explains why the exchanges have been accused of privileging high-frequency traders. In the analysis of David Evans, a multi-sided platform, which the stock exchange is, requires the ability to exclude parties that violate the platform's code of conduct.[72] A properly functioning platform has an incentive to police bad behavior, because doing so will increase the total value of the platform to its users as well as the profits to its owners. In the present environment, the lack of market power of any one exchange means that high-volume customers are so valuable that exchanges are tempted to cater to them at the expense of their other, less technologically sophisticated ones. There is evidence to support this admittedly strong assertion. In 2015, the BATS exchange paid a $14 million fine to settle an SEC enforcement action that the Direct Edge exchange had selectively disclosed how its "Hide Not Slide" order type functions.[73] The allegation was that BATS had informed high-frequency traders of the details regarding the functionality of the Hide Not Slide order but kept them from other market participants. Trader Haim Bodek further alleged that many of the complicated "special order types" are in fact developed by the exchanges working directly with HFT firms, and their complexity and arguably unfair functionalities, such as queue-jumping features, enable HFT firms to profit at the expense of other traders.[74] On a more general level, the development of co-location services and enriched data feeds, both used by HFT firms to gain speed and informational advantages, creates a two-tiered marketplace in which one set of traders profits at another's expense. Indeed, a number of commentators have

72. *See* Evans, *Governing Bad Behavior*, *supra* note 69, at 1220.

73. *See* Bradley Hope, *BATS to Pay $14 Million to Settle Direct Edge Order-Type Case*, WALL STREET J. (Jan. 12, 2015, 6:12 PM), https://www.wsj.com/articles/direct-edge-exchanges-to-pay-14-million-penalty-over-order-type-descriptions-1421082603.

74. *See* Haim Bodek, *Locked Markets, Priority and Why HFTs Have an Advantage: Part 2: Hide & Light*, TABBFORUM (Oct. 16, 2012), https://tabbforum.com/opinions/locked-markets-priority-and-why-hfts-have-an-advantage-part-2-hide-and-light; *see also* PATTERSON, *supra* note 21, at 204–05.

characterized the profits made in this manner as a tax or rent on long-term investors, calling into question the ultimate social utility of high-frequency trading that relies solely on positional advantage.[75]

From the perspective of the exchanges, however, it is necessary to cater to their most important customers. The corresponding failure of the exchanges as SROs is that they then permit and even profit from "unfair discrimination between customers," in the language of Exchange Act section 6(b)(5).[76] While this failure is not the exclusive result of technological change, the development of HFT in an environment of corporate exchanges set into direct competition with one another casts doubt on the assumption that the exchanges can properly function as SROs in the manner the Exchange Act intends.

The failure of these five assumptions shows the increasing tension between Regulation NMS and the digital stock markets it must now govern. With its origins in the Securities Acts Amendments of 1975, Regulation NMS was formulated for an earlier age. While it is neither an unalloyed success nor a complete failure, its emphasis on setting the exchanges into competition with one another faces considerable challenges for the future. The nature of the stock exchange as a platform business, and how platforms appear to operate in the digital environment, encapsulates these challenges.

II. THE STOCK EXCHANGE AS DIGITAL MULTI-SIDED PLATFORM

What type of business is exchange trading, and how might this affect the project of revamping the regulatory structure of the national market system? The stock exchange is fundamentally a multi-sided platform, a business through which disparate groups come together to transact. The literature on MSPs illuminates a

75. *See* Fox et al., *supra* note 29, at 241–44; Albert S. Kyle & Jeongmin Lee, *Toward a Fully Continuous Exchange*, 33 OXFORD REV. ECON. POL'Y 650, 651 (2017) ("The discreteness of today's market design creates rents for traders who have invested in costly superior technology."); John C. Coffee, Jr., *High Frequency Trading Reform: The Short Term and the Longer Term*, COLUM. L. SCH. BLUE SKY BLOG (July 21, 2014), http://clsbluesky.law.columbia.edu/2014/07/21/high-frequency-trading-reform-the-short-term-and-the-longer-term/; *see also* Yesha Yadav, *Insider Trading and Market Structure*, 63 UCLA L. REV. 968, 971 (2016) ("[T]he design of modern automated markets departs from past practice by expressly giving meaningful informational advantages to select, high-speed algorithmic traders.").

76. Securities Exchange Act of 1934 § 6(b)(5), 15 U.S.C. § 78f(b)(5) (2012).

number of the key features of stock markets both in their traditional and digital forms. The stock exchange as MSP is also crucial to the debate surrounding the pricing of exchange services through "maker-taker" rebates, as well as governance in the exchange space in the coming decades. Part II surveys the MSP literature and then looks at the current fragmented markets with a view to how the exchanges reflect the usual economic characteristics of MSPs. The stock exchange is a unique type of MSP, bearing some but not all of its typical characteristics.

Part II argues for the following propositions: (1) The stock exchange itself is an MSP, and the debate over pricing for its services should acknowledge this. Proposals to ban "maker-taker" rebates therefore are in considerable tension with the underlying economics of the digital stock exchange. (2) Exchange trading, both in its pre-digital and current forms, is affected by strong centripetal forces that tend to concentrate trading on one venue or a small number of venues. Indeed, some economists have viewed financial exchanges as natural monopolies. (3) Regulation NMS disrupts the natural operation of the liquidity network effect. While customers have benefited from the increased competition among the exchanges, which has driven down costs, a corollary of increased competition is the reduced incentive for exchanges to function effectively as SROs, as discussed above. (4) As competition has reduced the ability of exchanges to derive revenue from the provision of trading services, they have increasingly become data-provision businesses. Disputes surrounding data pricing have therefore come to the fore in recent years, particularly with the ongoing *NetCoalition* litigation.[77] While the conventional economic view is that each stock exchange functions as an individual MSP, to the extent that they may be acting in concert regarding data pricing, the exchanges begin to look like a collection of nodes that constitute a single super-MSP. Of course, it is just this question of monopoly, or rather, oligopoly, that is at issue in *NetCoalition*. The question of whether exchange trading really is a natural monopoly, as many economists suggest, and what an appropriate regulatory response might be, introduces Part III.

77. *See* NetCoalition v. SEC, 615 F.3d 525, 527 (D.C. Cir. 2010).

A. The Economics of the Stock Exchange

1. Why trade on an exchange?

It is common to see businesses of one type cluster together. In New York City, think of the theater district, the diamond district, the Fulton fish market, or Wall Street. In the case of a marketplace or souk, having a number of vendors selling similar goods in one place is beneficial for both buyers and sellers. Buyers can survey the different product offerings side by side and negotiate amongst the various sellers to find the one offering the best deal. Conversely, sellers "want to be where the action is," benefitting from the increased traffic generated by many buyers. Buyers and sellers thus interact in a way that creates a feedback loop, where more sellers attract more buyers, and vice versa.[78] Search costs are lowered for all by congregating trade in one location.

Trading in stocks also tends to converge on a single venue. Financial economists have built various models that provide a formal explanation why this is so.[79] The model of Marco Pagano divides traders into two classes, large traders interested in liquidity and speculators.[80] Liquidity traders such as institutional investors will need to sell an asset at a certain time. To obtain the best price, they seek the market that can absorb their order with the least amount of adverse effect on the price they will be offered, i.e., the most liquid market. While speculators are primarily interested in volatility, which offers the possibility of buying low and selling high, they are also attracted to the market with the greatest volume of trade because it offers the most potential matches. A feedback loop thus arises between liquidity traders and speculators.[81] The

78. *See generally* Glenn Ellison & Sara Fisher Ellison, *Lessons About Markets from the Internet*, 19 J. ECON. PERSP. 139, 142–43 (2005) (discussing agglomeration of trade in a variety of businesses including stock exchanges).

79. *See, e.g.*, Anat R. Admati & Paul Pfleiderer, *A Theory of Intraday Patterns: Volume and Price Variability*, 1 REV. FIN. STUD. 3 (1988); Marco Pagano, *Trading Volume and Asset Liquidity*, 104 Q.J. ECON. 255 (1989); Craig Pirrong, *A Theory of Financial Exchange Organization*, 43 J.L. & ECON. 437 (2000) [hereinafter Pirrong, *A Theory*]; Craig Pirrong, *The Organization of Financial Exchange Markets: Theory and Evidence*, 2 J. FIN. MKTS. 329 (1999) [hereinafter Pirrong, *Organization of Financial Exchange Markets*]; Lester G. Telser, *Why There Are Organized Futures Markets*, 24 J.L. & ECON. 1 (1981).

80. *See* Pagano, *supra* note 79, at 255–57.

81. *Id.* at 256.

model of Pagano shows that, under stylized conditions, including that trade is costless, trade will converge on a single market.[82] Complementing Pagano's model is that of Anat Admati and Paul Pfleiderer, which explains why trade in specific stocks tends to congregate at various times of the day, such as the periods immediately after the market opens and just before its close.[83] In their model, liquidity traders want to trade when they believe the market will be thickest, thus attracting informed traders as well.[84] These models therefore provide a theoretical explanation for the congregation of trade on specific markets and at specific times.

They also imply that, in the simplest situation, trading in financial instruments is a natural monopoly.[85] This corresponds to the commonly observed phenomenon that trading in a particular stock or financial instrument tends to cluster in one location or venue. While regulation at either the stock exchange level, such as former NYSE Rule 390, or at the level of the market as a whole, such as the Order Protection Rule of Regulation NMS, can encourage or discourage such clustering, the models imply that in the abstract world of financial theory, exchange trading is a natural monopoly. Perhaps more intuitively, a number of important economists have viewed a single exchange as in principle the most efficient arrangement for securities trading.[86] The models account for fragmentation of trading as well. According to Pagano, where transaction costs differ in markets, and large traders fear adverse price changes, they

82. *Id.* at 262.

83. *See* Admati & Pfleiderer, *supra* note 79.

84. *Id.* at 5.

85. *See id.* at 33 ("In equilibrium, discretionary . . . trading is typically concentrated. . . . [, and i]nformed traders trade more actively in periods when liquidity trading is concentrated."); Pagano, *supra* note 79, at 262 ("Thus, if trade is costless (or equally costly), all traders tend to concentrate on a single market."); Pirrong, *Organization of Financial Exchange Markets*, *supra* note 79, at 330 ("[T]rading of particular financial instruments exhibits strong natural monopoly characteristics . . . .").

86. *See, e.g.*, Fischer Black, *Toward a Fully Automated Stock Exchange*, FIN. ANALYSTS J., July–Aug. 1971, at 28, 29 ("It appears that the market for a single stock is most efficient if all orders for the stock come in to a single point, so that all potential buyers can be exposed to all sell orders, and all potential sellers can be exposed to all buy orders."); George J. Stigler, *Public Regulation of the Securities Markets*, 37 J. BUS. 117, 129 (1964) ("The greater the number of transactions in a security concentrated in one exchange, the smaller the discontinuities in trading and the smaller the necessary inventories of securities. As a result the price of a security will almost invariably be 'made' in one exchange."); *see also* Lawrence R. Glosten, *Is the Electronic Open Limit Order Book Inevitable?*, 49 J. FIN. 1127 (1994).

will trade on the market (or search off-exchange) with higher transaction costs.[87] Pagano provides the example of off-exchange trading in the Italian stock market; another example would be the "upstairs" market at the NYSE, which was the precursor to today's dark pools. Fragmentation therefore occurs when the combination of transaction costs and different sizes of pools of liquidity make it advantageous for some traders to turn to a second market or to individualized search.

A final theoretical observation to note is Craig Pirrong's argument that natural monopoly can coexist with supra-competitive profits in the world of member-owned financial exchanges.[88] By restricting their numbers, the members of a cooperative exchange will increase each individual seat holder's return. The cost of competing with the exchange however may prevent a competitor from arising: "Under plausible conditions, exchanges have enough members to make it unprofitable for competing exchanges to form, but fewer members than is socially optimal."[89] Pirrong surveys the prices for seats on various financial exchanges, finding evidence that the exchanges had done exactly that.[90] Of course, the major impetus for Regulation NMS was to reduce the monopoly profits of the NYSE members. While the solution has led to problems of its own, which are the subject of this Article, it is important to remember the original problem it was intended to deal with: monopoly rents accruing to members in the form of high transaction costs.

87. *See* Pagano, *supra* note 79, at 256.

88. *See* Pirrong, *Organization of Financial Exchange Markets*, *supra* note 79, at 339.

89. *Id.* at 330.

90. *Id.* at 349–53. Pirrong uses the measure "Tobin's *q*," defined as "the ratio of the market price of financial claims outstanding against a firm to the replacement value of the firm's assets," to investigate whether the price of seats on an exchange capitalizes economic rents. In a perfectly competitive industry, with no barriers to entry, Tobin's *q* should be 1.00. Looking at the prices for seats on the Chicago Board of Trade; the Chicago Mercantile Exchange; the New York Cotton Exchange; the Coffee, Sugar, and Cocoa Exchange; the New York Stock Exchange; and the Chicago Board Options Exchange, Pirrong finds that the *q* ratio in each case exceeded 1.00 and was typically above 2.00. The *q* ratio for the NYSE ranged from a high of 4.05 in 1987 to a low of 1.25 in 1991. These Tobin's *q* ratios thus provide evidence that the profits earned by exchange members constitute economic rents generated by restricting membership.

2. Multi-sided platforms

Over the past fifteen years economists have begun to study platform businesses or "multi-sided platforms."[91] MSPs are businesses that mediate between two or more groups of customers. A crucial swath of businesses in the economy perform this function in many different ways—commonly cited examples include recent digital platforms such as Amazon, Uber, and the like, but pre-digital examples include credit cards, shopping malls, financial exchanges, and even newspapers. While there are many types of MSPs, any business that allows for the direct interaction of two or more disparate groups of consumers has features of a multi-sided platform.[92] Financial exchanges have long been recognized as an important type of MSP, although there has been little specific study of them as a particular type of MSP.[93]

MSPs are businesses that exhibit strong "indirect network effects." This means that users on one side of the platform are attracted to, and benefit from, a large number of users on the other side.[94] In the case of a shopping mall, individuals shopping for new clothes will be attracted to a mall with many clothing retailers. Likewise, investors looking to buy stock on a stock exchange will

91. While many papers, including those of Rochet & Tirole, refer to "two-sided markets," the term "multi-sided platforms" seems more appropriate because many of these markets have a multiplicity of customers. Stock exchanges, for example, cater not only to buyers and sellers of securities but to issuers; there are also purchasers of market data, who may or may not be buyers and sellers. The platform itself is also important. For terminological discussion, see Ellison & Ellison, *supra* note 78, at 146; David S. Evans, *Antitrust Issues Raised by the Emerging Global Internet Economy*, 102 NW. U. L. REV. 1987, 1994 n.31 (2008) [hereinafter Evans, *Antitrust Issues*]; Andrei Hagiu & Julian Wright, *Multi-sided Platforms*, 43 INT'L J. INDUS. ORG. 162 (2015); Jean-Charles Rochet & Jean Tirole, *Two-Sided Markets: A Progress Report*, 37 RAND J. ECON. 645, 645 n.1 (2006) [hereinafter Rochet & Tirole, *Two-Sided Markets*].

92. *See* Andrei Hagiu, *Multi-sided Platforms: From Microfoundations to Design and Expansion Strategies* (Harvard Bus. Sch., Working Paper No. 07-094, 2006); Jean-Charles Rochet & Jean Tirole, *Platform Competition in Two-Sided Markets*, 1 J. EUR. ECON. ASS'N 990 (2003) [hereinafter Rochet & Tirole, *Platform Competition*].

93. *See* Jose Miguel M. Abito, A Two-Sided Market Analysis of Stock Exchanges (2006) (unpublished Honours B.Soc.Sci. thesis, National University of Singapore); *see also* Michael Castelle, *The Platform as Exchange: Financial Metaphors for the Regulation of Marketplace Platforms*, OXFORD INTERNET INST. IPP2016, at 10 (2016); Evans, *Antitrust Issues*, *supra* note 91, at 1996; Evans, *Governing Bad Behavior*, *supra* note 69, at 1231–35; Jonathan D. Levin, *The Economics of Internet Markets* (Nat'l Bureau of Econ. Research, Working Paper No. 16852, 2011).

94. *See* Hagiu, *supra* note 92, at 3; *see generally* Michael L. Katz & Carl Shapiro, *Systems Competition and Network Effects*, 8 J. ECON. PERSP. 93, 97–98 (1994).

be attracted to an exchange with a large number of sellers posting offers. The hallmark of an MSP is the operation of indirect network effects whereby a group on one "side" of the platform is affected by the number of users on the other side. (Direct network effects on the other hand are where customers are affected by the number of users on the same side—for example, a person signs up for telephone service because many others have signed up as well, and it will be easy to communicate with them.[95])

The economics literature specifies the essence of an MSP more precisely, however. In the pioneering articles of Jean-Charles Rochet and Jean Tirole, the defining characteristic of an MSP is that the total amount of activity conducted on the platform is affected not just by the total price charged for the service but by the breakdown of fees allocated to each side. According to Rochet and Tirole, "a two-sided market [is] one in which the volume of transactions between end-users depends on the structure and not only on the overall level of the fees charged by the platform."[96] This specification is crucial for the economics of MSPs and illuminates many commonly observed phenomena surrounding platforms. For example, why do most large American cities have a weekly newspaper distributed free of charge to commuters? Why do bars often host a "ladies' night" where women are charged half price for drinks? In the analysis offered by the MSP literature, these pricing strategies make sense in that they increase the volume of participants on one side of the market to a level that will cause the overall value created by the platform to increase.[97] As Rochet and Tirole note, these differences in pricing often lead to disputes about whether a group is being charged too much—as we see in the disputes over fees and rebates at the stock exchanges.[98] In the

95. *See* Evans, *Antitrust Issues*, *supra* note 91, at 1993 n.27.

96. Rochet & Tirole, *Two-Sided Markets*, *supra* note 91, at 646; *see also* Rochet & Tirole, *Platform Competition*, *supra* note 92, at 1018 ("[T]he volume of transactions on and the profit of a platform depend not only on the total price charged to the parties to the transaction, but also on its decomposition.").

97. *See* David S. Evans, *The Antitrust Economics of Free*, 7 COMPETITION POL'Y INT'L 71, 75 (2011).

98. *See* David S. Evans, *Some Empirical Aspects of Multi-sided Platform Industries*, 2 REV. NETWORK ECON. 191, 193–97 (2003) ("Optimal prices are not proportional to marginal costs . . . .") [hereinafter Evans, *Some Empirical Aspects*]; Rochet & Tirole, *Platform Competition*, *supra* note 92, at 1018–19; *see also infra* Section II.B.3.a ("Maker-taker pricing").

examples above, MSPs function as matchmakers that operate to reduce search costs for disparate groups of users.[99] In the language of economics, the positive externalities of the MSP are "internalized" by the network in the form of lower search costs.[100] The platform has a valuable role to play by offering much lower search costs to participants than the available alternatives.

A second important feature of MSPs is that they often lower search costs by certifying or monitoring participants on one or multiple sides of the platform.[101] An example would be seller certifications on Amazon.com, where sellers of books or other products are rated as to their reliability, or Airbnb.com, where hosts are rated as to their service levels. On the stock exchange, these certifications have long been important—a broker holding a seat on the NYSE, or an NASD (now FINRA) member trading on NASDAQ, possessed an important qualification.[102] Furthermore, this certification was connected to an important set of rules that policed conduct on the exchange, providing a guarantee (not always realized, of course) to customers of legal and ethical conduct by stockbrokers. While it is questionable whether the exchanges are presently incentivized to police their members' conduct, this traditionally has been an important aspect of the stock exchange's role as an MSP.[103]

A final point to note is the distinction between "multi-homing" and "single-homing."[104] Multi-homing is where one side to the exchange can easily use competing networks, such as a consumer who carries both Mastercard and American Express credit cards. Single-homing, on the other hand, is where it is only feasible for a

99. The matchmaker role played by MSPs reflects the fact that they solve bargaining problems between parties in situations in which the Coase theorem does not apply. *See* Evans, *Governing Bad Behavior*, *supra* note 69, at 1203; Rochet & Tirole, *Two-Sided Markets*, *supra* note 91, at 649.

100. *See* Evans, *Some Empirical Aspects*, *supra* note 98, at 192–93.

101. *See* Hagiu, *supra* note 92, at 5–6, 15; Evans, *Governing Bad Behavior*, *supra* note 69, at 1218–26.

102. *See* Lofchie, *supra* note 25.

103. *See* Yesha Yadav, *Oversight Failure in Securities Markets*, CORNELL L. REV. (forthcoming 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2754786; Lofchie, *supra* note 25.

104. *See* Mark Armstrong, *Competition in Two-Sided Markets*, 37 RAND J. ECON. 668, 669–70 (2006); Rochet & Tirole, *Two-Sided Markets*, *supra* note 91, at 659–60.

group of users to use one network: issuers listing their stock on a particular exchange, or shoppers in a regional city with only one shopping mall, would be single-homing. The importance of single-homing is that in cases where all users must single-home, the network looks like a utility-type situation where monopoly is often the most efficient structure.[105] In the case of a stock exchange, however, buyers and sellers don't have to single-home, and they typically multi-home by monitoring prices on many venues and consummating a transaction on the one displaying the most favorable price. As discussed further below, the stock exchanges have a complex institutional landscape with some but not all features of classic MSPs. While issuers (typically but not always) single-home, the other groups of exchange customers typically multi-home.

3. The effects of digitization

As the examples of Amazon, Microsoft, Airbnb, Expedia, Uber, and others indicate, MSPs are rapidly increasing in importance as information technology and the internet develop.[106] Many of the most important businesses in the world right now are MSPs, and most established industries are in various, often near-complete, stages of transformation due to the digitization of commercial life. Facebook, Instagram, and Twitter have essentially created the entirely new industry of social media, while Amazon has transformed the preexisting retail sales industry. At the opposite end of the spectrum from social media—largely a new creation—are the stock exchanges. While the transformation of financial markets is thoroughgoing, it is the result of the application of computer technology to a highly detailed, preexisting business and regulatory framework, not the creation of an entirely new business.[107] As Part I above explores, digitization creates significant

105. *See* Armstrong, *supra* note 104, at 670 ("[W]here all users must single-home. . . . the efficient outcome is for all agents to use the same platform.").

106. *See* Hagiu, *supra* note 93, at 2; Günter Knieps & Johannes M. Bauer, *The Industrial Organization of the Internet*, *in* HANDBOOK ON THE ECONOMICS OF THE INTERNET 23, 31–32 (Johannes M. Bauer & Michael Latzer eds., 2016).

107. *See* Castelle, *supra* note 93 (arguing that regulators grappling with the "emergent organizational and regulatory complexities" of MSPs should draw on the example of securities regulation, which has gradually encompassed electronic stock exchanges).

tension within this framework, and key assumptions no longer hold good. Nevertheless, the application of advanced computer technology does not have the same character as in other industries which have been newly created out of the whole cloth of the internet, such as social media, or just radically transformed by it, such as retail. While the "death of distance" and increasing returns to scale enabled by the internet have radically transformed many traditional businesses,[108] their effect on the financial exchanges appears less profound. The preexisting business and regulatory structure already enabled a high degree of connection with investors not present at an exchange through brokers holding seats there, and an infrastructure that could handle a substantial volume, and increase in volume, of transactions.

The transformation of the exchanges by advanced information technology may be more subtle than in other industries, then. Nevertheless, its effects are significant. First, with the digitization of financial markets has come a decreasing need for various categories of financial intermediaries. Information technology fuels disintermediation here because there is less of a role for specialized parties in electronic stock markets.[109] The "death of the specialist" on the NYSE is the most obvious example of this: no longer do we need specialists in each listed stock to act as matchmakers on the trading floor.[110] The death of the specialist really traces back to NASDAQ's quote-driven market, which operated with broker-dealers communicating with one another telephonically, unlike the open outcry auction market at the NYSE.[111] In addition to dispensing with the specialists, the current market is disintermediating further, with a steep decline in the number of broker-dealers in

108. *See* Justus Haucap & Torben Stühmeier, *Competition and Antitrust in Internet Markets*, *in* HANDBOOK ON THE ECONOMICS OF THE INTERNET 183, 184–88 (Johannes M. Bauer & Michael Latzer eds., 2016); Knieps & Bauer, *supra* note 106, at 24.

109. *See generally* Brummer, *supra* note 7, at 1024–31; ONNIG H. DOMBALAGIAN, CHASING THE TAPE: INFORMATION LAW AND POLICY IN CAPITAL MARKETS 182 (2015); *see also* Fischer Black, *Toward a Fully Automated Stock Exchange, Part II*, 27 FIN. ANALYSTS J., Nov.–Dec. 1971, at 24, 86–87.

110. *See* Jerry W. Markham & Daniel J. Harty, *For Whom the Bell Tolls: The Demise of Exchange Trading Floors and the Growth of ECNs*, 33 J. CORP. L. 865, 897–900 (2008).

111. *See* Stoll, *supra* note 53, at 159–60.

recent years.[112] With more and more retail investing going through online brokerages such as E-Trade and Charles Schwab, there is less business for the local, smaller brokerage houses. As Pirrong noted in 2000, computer-driven markets drive disintermediation, and the development of the markets since then illustrates this.[113]

A second effect of digitization is the rapid development and application of special order types. Throughout most of the history of the exchanges, orders came in two basic forms: a limit order, which was a standing order to buy or sell a certain number of shares at a fixed price, and a market order, which was an immediate order to buy or sell a certain number of shares at the current market price.[114] As the exchanges digitized, becoming essentially complicated, rule-based matching engines housed in computer servers,[115] it became much easier to develop and apply complicated and specialized order types. Now orders such as Midpoint Immediate or Cancel, Adding Liquidity Only Limit, Day ISO (Intermarket Sweep Order), Non-Displayed Limit, and various types of auction orders, activated at the open or close of trading, are offered by the exchanges.[116] Many of the special order types developed in recent years are designed to take advantage of, or guard against, various local conditions in the electronic markets.[117] The development of complex order types is an example of the innovation and customization available to operators of online, algorithmic businesses.[118] It also reflects the fact that digitization allows for vastly

112. *See* Hester Pierce, *Dwindling Numbers in the Financial Industry*, BROOKINGS INSTITUTION (May 15, 2017), https://www.brookings.edu/research/dwindling-numbers-in-the-financial-industry/ (detailing a decline of 5892 broker-dealers registered with the SEC in 2007 to 3989 in 2017); *see also* Richard Johnson, *Brokers Adapt to Shrinking Equity Commissions*, GREENWICH ASSOCIATES (July 27, 2017), https://www.greenwich.com/equities/brokers-adapt-shrinking-equity-commissions.

113. *See* Pirrong, *A Theory*, *supra* note 79, at 460.

114. *See supra* note 26 and accompanying text.

115. *See* Paul G. Mahoney & Gabriel Rauterberg, *The Regulation of Trading Markets: A Survey and Evaluation* 4 (Univ. of Va. Sch. of Law, Law and Economics Research Paper Series No. 2017-07, 2017), https://ssrn.com/abstract=2955112 ("It is tempting to think of a stock market as a facility, physical or virtual, but it is better described as a set of rules and procedures pursuant to which investors buy and sell securities.").

116. *See*, *e.g.*, NYSE Rule 13 (2018) (listing and defining various order types).

117. *See* ALDRIDGE, *supra* note 46, at 44–45; Morelli, *supra* note 26.

118. *See* Levin, *supra* note 93, at 11.

increased information flows.[119] The panoply of sophisticated order types both requires a great amount of data to function, and in turn produces a great amount of data available for use, at a price, by financial markets participants at large.

Finally, the development of the internet has allowed new entrants to compete against established players.[120] This is true of the stock exchanges, although after the rush to compete against the NYSE in the wake of Regulation NMS, it appears that the opposite has been true, and the markets have settled into a stable allocation of market share among a few major players.[121] In 2005, the NYSE was still a monopoly exchange, handling approximately 80% of the trading in NYSE-listed shares. A number of new entrants quickly arose, including BATS and Direct Edge, and the NASDAQ encroached on the NYSE as well. After a few years of growth, a wave of consolidation occurred, with BATS acquiring the Direct Edge exchanges and the dominant exchanges within the NASDAQ and NYSE families each holding on to 15–20% of the market.[122] While there is intense competition among the exchanges, at present it appears difficult for new entrants to break into the industry. The Investors Exchange (IEX), headed by *Flash Boys* protagonist Brad Katsuyama, is an important case in point.[123] Despite marketing itself as a more ethical alternative to the major stock exchanges, it has not been able to capture more than 3% of the market.[124] IEX appears to be a cautionary tale for would-be entrants. The reasons for its struggles are complex but may be evidence that, like other online businesses, the exchange business shows a heightened tendency to winnow down to one, or a small number, of players in a "winner-takes-all," or more accurately "winner-takes-most," tournament.[125]

119. *See* Nicholas Economides, *The Impact of the Internet on Financial Markets*, 1 J. FIN. TRANSFORMATION 8 (2001).

120. *See* Fleckner, *supra* note 68, at 2566–67; Knieps & Bauer, *supra* note 106, at 24.

121. *See* discussion *infra* Section II.B.2.

122. *See supra* note 4 and accompanying text.

123. *See infra* Section III.B.6.

124. *See* Nicole Bullock, *IEX Chief Sticks to Principles in Battle for Presence*, FIN. TIMES (June 1, 2017), https://www.ft.com/content/4c805dd6-449f-11e7-8519-9f94ee97d996.

125. *See* Economides, *supra* note 119, at 9; Evans, *Antitrust Issues*, *supra* note 91, at 2003.

B. The Exchanges Under Regulation NMS — A Network of MSPs or a Super-MSP?

In the decade since Regulation NMS came into force, the most important, and controversial, effects of digital technology on the stock market have come from HFT. Both the activities of high-frequency traders themselves and the response of the exchanges and others, such as dark pool operators and institutional traders, have played a role. The story of this transformation bears strong evidence that the exchanges are MSPs. Most obviously this is seen in the development of "maker-taker" pricing schemes at the exchanges, but the use of digital technology to implement special order types, as well as launch new exchanges, also fits the pattern of MSPs in the digital business environment. Meanwhile, these economic and technological developments have played out under the new regulatory regime of Regulation NMS. Two provisions of Regulation NMS are most important for the stock exchanges as MSPs: By "socializing order flow" the Order Protection Rule (Rule 611) weakens the liquidity network effect, and the Access Rule (Rule 610) provides the parameters within which the exchanges can offer rebates to broker-dealers for providing order flow, i.e., liquidity. The operation of these rules in the new environment has resulted in intense competition among the exchanges. Traders have paid lower fees, as intended, but in response the exchanges have sought alternative revenue sources in the form of higher fees for their proprietary data. The ensuing disputes over the exercise of market power on the part of the exchanges leads back to the question of monopoly—or now, oligopoly—power that was the original impetus for the 1975 Securities Acts Amendments. Profits the exchanges can no longer make from the provision of trading services are now supplanted by profits from the sale of information generated by those trading services.

Section II.B surveys the current state of the markets, with an eye on the nature of the stock market as MSP. Market fragmentation has increased greatly in recent years, particularly in the migration of trading to dark pools and other non-exchange venues; fragmentation among the lit markets is less important, however, as four exchanges, grouped into three exchange families, hold the bulk of the market share. Nevertheless, the response of the exchanges to the current environment bears cause for concern, as

certain techniques appear to tilt the playing field to the benefit of some and disadvantage of others. The current state of the markets is complex. Regulation NMS cuts against the nature of the stock market as an MSP, with its strong pressures toward centralization, to the benefit of customers through increased competition. On the other hand, it creates an environment that allows a host of other costs and inefficiencies to take root, jeopardizing the SEC's mission of investor protection.

1. Dark fragmentation

While stock trading is now fragmented among thirteen exchanges, a significant share of trading occurs outside of the exchanges entirely. The most basic type of fragmentation then is between the "lit" markets, or the registered exchanges, and the "dark" markets, or "dark fragmentation."[126] At present, about 40% of stock (by dollar volume) is traded off-exchange in a variety of venues.[127] Stock also trades off-exchange in the internalization engines of broker-dealers such as the large banks, and in the private, "upstairs" market.[128] Dark pools are probably the best known of these venues. They are known as "dark" venues because, unlike at the exchanges, the quotes (buy and sell orders) submitted to them are not publicly displayed. When a trade is made, however, that information is sent to the SIP and so contributes to the price discovery process after the fact.[129] Dark pools are governed by Regulation Alternative Trading Systems (Regulation ATS) under the 1934 Act and are required to register as broker-dealers.[130]

The volume of stock traded off-exchange has risen considerably in recent years. At the time Regulation NMS came into force, it was estimated that only 4% of shares traded in alternative venues.[131] By

126. *See* Ryan J. Davies & Erik R. Sirri, The Economics and Regulation of Secondary Trading Markets 15 (Jul. 20, 2017) (working paper) (on file with Columbia University, The New Special Study of the Securities Market), www.law.columbia.edu/sites/default/files/microsites/capital-markets/davies_sirri_20170720_final_draft.pdf.

127. *See Equities LiquidityMatrix*, TABBFORUM (Sept. 14, 2018), https://tabbforum.com/liquidity-matrix?ticket=ST-15372207594461-IJd8RwED8FV9IZop0j3uSVBWoc7a9YiYhpUHUL4c.

128. *See* COWEN ATM, *supra* note 2.

129. *See* SHORTER & MILLER, *supra* note 2, at 2.

130. *Id.* at 5.

131. *See* Scott Patterson, *'Dark Pools' Face Scrutiny*, WALL STREET J. (June 5, 2013, 9:55 PM), https://www.wsj.com/articles/SB10001424127887324069104578527361102049152.

2016, that amount had climbed to approximately 36%.[132] Of this 36%, approximately 13% traded in dark pools, with the other 23% trading in other off-exchange venues.[133] Fueling the rise of alternative trading venues is the search for lower transaction costs, as well as a desire to avoid interacting with high-frequency traders. Because orders submitted to dark pools are not publicly posted or otherwise distributed, it is in principle more difficult for high-frequency traders to trade against these orders or otherwise exploit them. While trading in dark pools clearly entails risks,[134] investors are attracted to them because trading on the exchanges also entails significant risk of adverse selection, as well as higher fixed costs.[135] As Stephen Lofchie suggests, instead of thinking of dark pools as something sinister or bad, it might be more appropriate to think of them as "protective coves."[136]

The rise of dark pools and off-exchange trading has been controversial and certainly carries with it considerable negative externalities. Most importantly, traders submitting quotes are not contributing to the price discovery process surrounding a stock but are instead free-riding on prices generated on the lit markets. Insofar as the stock exchange is a mechanism for allocating capital efficiently, significant amounts of dark trading detract from this important public function.[137] On the other hand, the rise of dark pools has coincided with the growth of HFT and must be seen as an effort on the part of investors to avoid the risks and costs of the lit markets. Turning back to the economic theory discussed above, the model of Pagano predicts that fragmentation will occur when different groups of investors have different cost and risk profiles.[138]

132. *See* COWEN ATM, *supra* note 2.

133. *Id.*

134. *See, e.g.*, Press Release, Sec. & Exch. Comm'n, *supra* note 32; *see also* Christopher Mercurio, *Dark Pool Regulation*, 33 REV. BANKING & FIN. L. 69 (2013).

135. *See* SHORTER & MILLER, *supra* note 2, at 4–5; Lofchie, *supra* note 25, at 4–6; Yadav, *supra* note 103, at 33–35; *see also* Frank Hatheway, Amy Kwan & Hui Zheng, *An Empirical Analysis of Market Segmentation on U.S. Equities Markets* 52 J. FIN. & QUANTITATIVE ANALYSIS 2399, 2401 (2017) ("[A]dverse selection risk on dark venues is 60% to 80% less than that on lit markets . . . .").

136. Lofchie, *supra* note 25, at 5.

137. *See* DOMBALAGIAN, *supra* note 109, at 62; Hatheway et al., *supra* note 135, at 2424–25; *see also* Yesha Yadav, *How Algorithmic Trading Undermines Efficiency in Capital Markets*, 68 VAND. L. REV. 1607 (2015).

138. *See* Pagano, *supra* note 79, at 268–69.

The rise of dark pools must be understood in light of the risks posed to slower, longer-term investors by trading on the exchanges. While in principle it may be best for trading to occur on lit markets, it is difficult to blame investors for wanting to trade in the dark markets in search of lower transaction costs and less risk of adverse selection. Nevertheless, dark fragmentation has significant drawbacks—it can result in negative effects on market quality, and if financial markets information is a public good, insofar as only post-trade information is reported, dark trading fails to contribute to its creation and free-rides on prices created by the lit markets.[139] And these more general considerations are buttressed by the instances in which dark pool operators failed to safeguard their venues for all their customers.[140]

2. Visible fragmentation

The 64% of trading volume that occurs on the "lit" markets is itself fragmented among the thirteen different stock exchanges. Of these markets, the data provided by Cowen ATM indicates that no single exchange has more than 14% of the total.[141] It must also be noted that eleven of the exchanges are grouped under three holding companies that offer different venues, each catering to different groups of investors through different pricing mechanisms. The NYSE group operates the NYSE, NYSE Arca, and NYSE American[142] exchanges, as well as the NYSE National, which traces its existence back to the National Stock Exchange and before that the Cincinnati Stock Exchange.[143] NASDAQ operates the NASDAQ

139. *See* DOMBALAGIAN, *supra* note 109, at 27–28; *see also* J. Harold Mulherin, Jeffry M. Netter & James A. Overdahl, *Prices Are Property: The Organization of Financial Exchanges from a Transaction Cost Perspective*, 34 J. L. & ECON. 591 (1991).

140. *See, e.g.*, Press Release, Sec. & Exch. Comm'n, *supra* note 32. *See generally* Brian P. Baxter, *The Securities Black Market: Dark Pool Trading and the Need for a More Expansive Regulation ATS-N*, 70 VAND. L. REV. 311 (2017).

141. *See* COWEN ATM, *supra* note 2.

142. Formerly the NYSE MKT, the NYSE American is intended to compete with IEX and includes a 350-microsecond "speed bump" and a reliance on midpoint orders. Such a move is ironic, considering the NYSE's vociferous opposition to IEX's application to become a registered exchange. *See* John D'Antona, *NYSE American Marketplace Launches*, TRADERS MAG. (July 25, 2017), http://www.tradersmagazine.com/news/ecns_and_exchanges/nyse-american-marketplace-launches-116480-1.html.

143. *See NYSE Regulation,* NYSE, https://www.nyse.com/regulation (last visited Dec. 30, 2018).

(with 14.0% of volume, the single largest exchange),[144] NASDAQ BX, and NASDAQ PSX exchanges.[145] And Cboe Global Markets operates the BZX (Bats-Z), BYX (Bats-Y), EDGA, and EDGX exchanges.[146] These three groups therefore comprise the dominant operators of exchange services in the American market. Finally, there is the Investors Exchange, or IEX, and the Chicago Stock Exchange, CHX.[147] Given that IEX and CHX together currently account for less than 4.0% of market volume, the market for equity trading services exhibits a high degree of consolidation when considered from the point of view of ownership of the various exchanges.

Upon the promulgation of Regulation NMS, but before it came into force in 2007, the bulk of trading still occurred on the NYSE: the NYSE held on to nearly 80% of the trading in NYSE-listed stocks.[148] It is important to note that NYSE Rule 390, which had prevented NYSE members from trading NYSE-listed stocks anywhere other than on the NYSE, was repealed in 2000.[149] Despite that rule change, the NYSE held on to the vast bulk of American stock trading.

By 2016, the market was split into the following venues:[150]

144. *See* COWEN ATM, *supra* note 2.

145. *See Nasdaq Equities Market Data*, NASDAQ, https://business.nasdaq.com/intel/GIS/Market-Data-Equities.html (last visited Dec. 30, 2018).

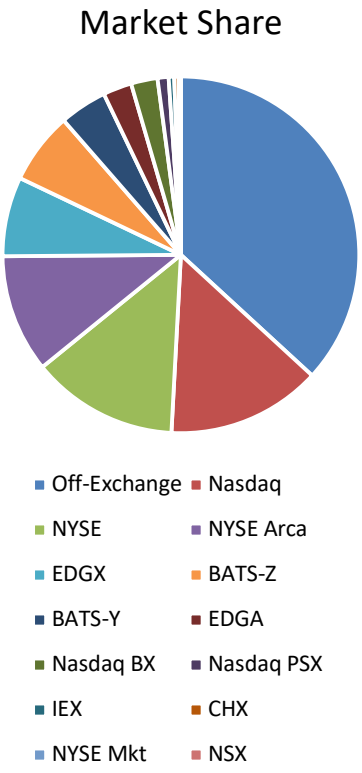
146. *See Cboe U.S. Equities*, CBOE, https://markets.cboe.com/us/equities/overview (last visited Dec. 30, 2018).

147. *See Last Trading Day*, IEX TRADING GROUP (Oct. 18, 2018), https://iextrading.com/stats/; John McCrank, *Chicago Stock Exchange Says Seeking New Potential Buyers*, REUTERS (Mar. 6, 2018, 11:17 AM), https://www.reuters.com/article/us-chicagostocexchange-m-a/chicago-stock-exchange-says-seeking-new-potential-buyers-idUSKCN1GI2I0.

148. *See* Concept Release on Equity Market Structure 6, Exchange Act Release No. 34-61358, 75 Fed. Reg. 3593, 3595 (Jan. 21, 2010).

149. *See* Oesterle, *supra* note 35, at 614.

150. COWEN ATM, *supra* note 2; data from BATS Market Data.



Exchange	Market Share
Off-Exchange	36.8%
Nasdaq	14.0%
NYSE	13.3%
NYSE Arca	10.7%
EDGX	7.2%
BATS-Z	6.5%
BATS-Y	4.3%
EDGA	2.6%
Nasdaq BX	2.4%
Nasdaq PSX	1.0%
IEX	0.5%
CHX	0.4%
NYSE Mkt	0.2%
NSX	0.0%

As of 2016, the fragmentation among venues is fairly even, with the NASDAQ having the single-greatest volume of 14.0%. When the exchanges are grouped together by ownership, no single group possesses more than a quarter of the total market: the four exchanges owned by the NYSE group have 24.2%, the NASDAQ group has 17.4%, and the Cboe group has 20.6%. Although there is significant fragmentation among the thirteen venues, and far more alternative venues, the degree of concentration at the group level in the lit markets is suggestive of oligopoly.

3. The response of the exchanges

The review presented above indicates that on the level of the trading system as a whole, the pressures toward fragmentation are counteracting the inherent network effect of exchange trading. Regulation NMS has only been in force for a decade, and it must be

remembered that this is a very short period of time in the entire sweep of the economic history of the United States—the NYSE, for example, traces back to the Buttonwood Agreement of 1792.[151] In addition, the past twenty years have been ones of rapid and ongoing technological change, so it is too early to predict which forces will ultimately prevail. With the advent of Regulation NMS, the exchanges are locked in a fierce battle for market share and, ultimately, survival. In order to meet their primary objective of attracting liquidity, they offer not only rebates for liquidity "makers" but other enticements as well. And to make up for lost revenue from trading, they have raised fees for their proprietary data feeds. The intense debate surrounding both the various means of enticing traders—particularly the high-volume, high-frequency ones—to their exchanges, as well as increasing market data fees, suggests that the forces that have disciplined MSPs in other industries have not been operative on the exchanges. The reasons why this may be so highlight the nature of exchange trading as a natural monopoly and the effects of knitting together a collection of competing MSPs into a single, linked trading system.

From the point of view of the exchanges, in the past decade they have been cast into a fight for their very survival.[152] The charts above indicate the rapid decline in market share of the NYSE since 2005, which had long held a de facto monopoly position in trading of NYSE-listed securities.[153] Not only were the major investment banks setting up dark pools as alternative venues for institutional traders who previously relied on the "upstairs market" at the NYSE for block trades, broker-dealers began to execute trades internally. The exchanges responded to the competition in a variety of ways.

151. *See* Stuart Banner, *The Origin of the New York Stock Exchange, 1791–1860*, 27 J. LEGAL STUD. 113, 115 (1998).

152. *See* Brummer, *supra* note 7, at 1007; Fleckner, *supra* note 68, at 2574; *see also* Amir N. Licht, *Stock Exchange Mobility, Unilateral Recognition, and the Privatization of Securities Regulation*, 41 VA. J. INT'L L. 583 (2001) (exploring competitive pressures on exchanges in the pre-Regulation NMS environment).

153. The NYSE maintained its monopoly position in the trading of NYSE-listed securities in part through the operation of NYSE Rule 390, which prohibited "off-board" trading by NYSE members of NYSE securities. *See* Mark Borrelli, *Market Making in the Electronic Age*, 32 LOY. U. CHI. L.J. 815, 838–40 (2001).

a. Maker-taker pricing. In the competition for order flow, they followed the early ECNs and adopted "maker-taker" pricing.[154] While recently the subject of significant controversy,[155] maker-taker pricing was pioneered in the 1990s by Island, an early ECN.[156] It involves the exchange issuing a rebate to a trader posting a quote, the maker of liquidity, that is then hit by another trader, the taker, who buys or sells at the posted price.[157] Maker-taker fees are calculated off of the fee permitted under the Access Rule of Regulation NMS, Rule 610,[158] which allows exchanges to charge a fee to traders of up to 0.30¢ per share, or 30¢ per 100 shares. If an exchange charges a 0.30¢ fee, a typical rebate paid to the maker of liquidity would be 0.25¢. This means that the taker of liquidity will pay 0.30¢ for the trade, the maker of liquidity will receive a rebate 0.25¢, and the exchange will pocket the difference of 0.05¢.

Maker-taker pricing has a number of effects on securities markets.[159] Most importantly, it can incentivize brokers to route customer orders to the exchanges paying the best rebates, arguably causing them to violate their duty of best execution.[160] It also distorts prices, thereby reducing transparency in the markets, as it creates a discrepancy between the posted price of a stock and the

154. *See generally* Larry Harris, Maker-Taker Pricing Effects on Market Quotations (Nov. 14, 2013) (working paper) (on file with USC Marshall School of Business).

155. *See, e.g.*, *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 Comm. on Homeland Sec. & Governmental Affairs*, 113th Cong. 7–9 (2014) (statement of Robert Battalio, Professor of Finance, University of Notre Dame); Curt Bradbury & Kenneth E. Bentsen Jr., *How to Improve Market Structure*, N.Y. TIMES (July 14, 2014, 7:03 AM), https://dealbook.nytimes.com/2014/07/14/how-to-improve-market-structure/; Jonathan Macey & David Swensen, Opinion, *Wall Street Profits by Putting Investors in the Slow Lane*, N.Y. TIMES (July 18, 2017), https://www.nytimes.com/2017/07/18/opinion/wall-street-brokers-rebates-kickbacks.html; *see also* Stanislav Dolgopolov, *The Maker-Taker Pricing Model and Its Impact on the Securities Market Structure: A Can of Worms for Securities Fraud?*, 8 VA. L. & BUS. REV. 231 (2014).

156. *See* PATTERSON, *supra* note 21, at 157–60.

157. *See* Harris, Maker-Taker Pricing, *supra* note 154, at 2.

158. 17 C.F.R. § 242.610 (2018).

159. *See* Harris, Maker-Taker Pricing, *supra* note 154; Haim Bodek & Stanislav Dolgopolov, *Deconstructing Maker-Taker, Part 1: A Gordian Knot for Market Structure?*, TABBFORUM (June 30, 2014), https://tabbforum.com/opinions/deconstructing-maker-taker-part-1-a-gordian-knot-for-market-structure.

160. *See* Robert Battalio, Shane A. Corwin & Robert Jennings, *Can Brokers Have It All? On the Relation Between Make-Take Fees and Limit Order Execution Quality*, 71 J. FIN. 2193, 2197 (2016).

real price received by the seller.[161] And it is deeply implicated in the development of HFT generally and the special order types it uses.[162] Nevertheless, it is most obviously a way to attract liquidity by offering a discount to traders bringing liquidity to an exchange or other venue. It therefore reflects Rochet and Tirole's insight that the volume of transactions carried out on an MSP is sensitive to the allocation of trading costs between the parties interacting there.[163] Similar to a website that offers news content for free to readers while charging advertisers a fee, maker-taker pricing is an acknowledgement of both the value of liquidity suppliers and that liquidity takers will pay to access that resource. This is not to deny that rebates to makers can have deleterious second-order effects, as Jonathan Macey and David Swensen, Haim Bodek and Stanislav Dolgopolov, and others point out.[164] Rebates can tempt brokers to violate their duty of best execution, as well as distort the efficiency of the stock market by causing the price in a quote to deviate from the true price paid to the seller. Nevertheless, maker-taker pricing fundamentally appears to reflect the fact that the greatest value of transactions on an exchange will occur when parties on one side of the platform, the suppliers of liquidity, are charged less than those on the other side for transacting on the exchange.[165]

161. *See* Harris, Maker-Taker Pricing, *supra* note 154, at 3 (explaining that maker-taker creates a transparency problem by causing net spreads to diverge from quoted spreads). *But see* Yong Chao, Chen Yao & Mao Ye, Tick Size Constraints, Two-Sided Markets, and Competition Between Stock Exchanges 35 (2015) (working paper), https://www.sec.gov/comments/265-29/26529-37.pdf (explaining that by reducing the effective tick (real pricing), maker-taker reduces frictions).

162. *See* Maureen O'Hara, *High Frequency Market Microstructure*, 116 J. FIN. ECON. 257, 262 (2015).

163. *See supra* notes 96–100 and accompanying text. Despite the intuitive nature of the connection between maker-taker pricing and Rochet & Tirole's insights on MSP pricing, it is not widely discussed in the market microstructure literature. *But see* Evans, *Antitrust Issues*, *supra* note 91, at 293–94; Levin, *supra* note 93, at 6; Chester Spatt, *The New Realities of Market Structure and Liquidity: Where Have We Been? Where Are We Going?* 9–10 (May 3, 2016) (presented at the Federal Reserve Bank of Atlanta 2016 Financial Markets Conference, "Getting a Grip on Liquidity: Markets, Institutions, and Central Banks").

164. *See supra* notes 155, 159.

165. An interesting variation is the "inverted exchange," where liquidity takers (not makers) are paid a rebate. Inverted exchanges are attractive to brokers handling trades for institutional or retail investors using market orders. Currently, the NASDAQ BX, BYX, EDGA, and NYSE National exchanges offer taker-maker pricing. *See* Alexander Osipovich, *'Inverted' Model Said to Be Considered for NYSE's Newest Exchange*, WALL STREET J. (Feb. 28,

b. Special order types, co-location, and enriched data feeds. In addition to maker-taker pricing, the exchanges have developed a variety of other features to lure traders to their venues, including special order types, flash orders, and co-location. All of these are typically used by high-frequency traders in conjunction with proprietary data feeds offering a greater depth of information than the SIP provides. In the new world of trading, high-frequency traders have become the exchanges' best customers, responsible for over half of trading volume in 2016.[166] In order to attract them, exchanges developed special order types that facilitated their trading techniques. These orders went far beyond the basic market orders and limit orders and were designed, allegedly in conjunction with the HFT shops, to allow HFT firms to take advantage of the particular conditions and regulatory restraints governing exchange trading.[167] A well-known example is the Hide Not Slide order developed by the Direct Edge exchanges, EDGA and EDGX. Rule 610(d) of Regulation NMS prohibits an exchange posting an order that "locks" or "crosses" a quote on another exchange.[168] This means that if one exchange currently displays the national best bid (a buy order) in a stock of $12.60, another exchange cannot post an offer (a sell order) at the same price. Such an order would "lock" the market, representing a fundamentally irrational state of affairs: a trader posting that order should immediately hit the previously posted bid. Similarly, a sell order of $12.59 would "cross" the market and represents an even more irrational state of affairs. In order to prevent such situations, Rule 610(d) prohibits "locks and crosses." (Note however that maker-taker pricing incentivizes posting a locked or crossed order due to the desire to avoid paying an access fee and to receive a rebate.[169]) The Hide Not Slide order

2017, 5:33 AM), https://www.wsj.com/articles/inverted-model-said-to-be-considered-for-nyses-newest-exchange-1488277981.

166. *See* RENA S. MILLER & GARY SHORTER, CONG. RESEARCH SERV., R44443, HIGH FREQUENCY TRADING: OVERVIEW OF RECENT DEVELOPMENTS 2 (2016) (stating in 2016 that HFT accounted for "roughly 55% of trading volume in U.S. equity markets").

167. *See generally* Stanislav Dolgopolov, *High-Frequency Trading, Order Types, and the Evolution of the Securities Market Structure: One Whistleblower's Consequences for Securities Regulation*, 2014 J.L. TECH. & POL'Y 145, 147–54 (2014).

168. 17 C.F.R. § 242.610(d) (2018) (prohibiting "[l]ocking or crossing quotations").

169. *See* Ivy Schmerken, *Nasdaq's Battle over Locked Crossed Markets*, INFORMATIONWEEK WALLSTREET & TECH. (Apr. 15, 2003, 12:22 PM), http://www.wallstreetandtech.com/exchanges

was formulated to put its user at the head of the line once the market moved.[170] Because it allowed its user to be first in line once the market unlocked, or uncrossed, it was seen as facilitating queue-jumping on the part of HFT. In addition to complaining about its fundamental mechanics, critics such as Haim Bodek complained that the exchanges hid the details of their operation from slower, institutional traders.[171] Indeed, Direct Edge, by then acquired by BATS, was fined by the SEC for improper disclosure of its Hide Not Slide order in 2015, substantiating Bodek's claims.[172]

Other means used by the exchanges to attract order flow from high-frequency traders include offering "flash orders" to traders,[173] which allow high-frequency traders to trade on incoming quotes before they are transmitted to the SIP, and co-location services and enriched data feeds.[174] While much more could be said about all of these developments, they have given rise to considerable controversy. From the point of view of the exchanges, they are merely offering a service open to all who are willing to pay for it, and it has long been considered acceptable for businesses to offer in-depth or enhanced information to customers willing to pay for it. On the other hand, the cumulative effect of these services offered to high-frequency traders is to create a two-tiered marketplace.[175] In the fight for market share, the exchanges appear to have colluded with

/nasdaqandrsquos-battle-over-locked-crossed-markets/d/d-id/1255842.html (exploring the interaction of liquidity rebates and locked and crossed markets).

170. *See* Scott Patterson & Jenny Strasburg, *For Superfast Stock Traders, a Way to Jump Ahead in Line*, WALL STREET J. (Sept. 19, 2012, 5:59 PM), https://www.wsj.com/articles/SB10000872396390443989204577599243693561670.

171. *See, e.g.*, Haim Bodek, *HFT Checkmate – The Alpha in Order Types*, TABBFORUM (Dec. 31, 2013), https://tabbforum.com/opinions/hft-checkmate-the-alpha-in-an-order-type.

172. *See* Hope, *supra* note 73; *cf.* Matt Levine, *'Hide Not Slide' Orders Were Slippery and Hidden*, BLOOMBERG: OPINION (Jan. 12, 2015, 5:35 PM), https://www.bloomberg.com/opinion/articles/2015-01-13/hide-not-slide-orders-were-slippery-and-hidden.

173. *See* Lawrence E. Harris & Ethan Namvar, *The Economics of Flash Orders and Trading*, 14 J. INV. MGMT. 74 (2016).

174. *See* Jacob Adrian, *Informational Inequality: How High Frequency Traders Use Premier Access to Information to Prey on Institutional Investors*, 14 DUKE L. & TECH. REV. 256, 266–68 (2016).

175. *See* Steven McNamara, *The Law and Ethics of High-Frequency Trading*, 17 MINN. J. L. SCI. & TECH. 71, 113–14 (2016).

their best customers to allow them to profit at the expense of their slower, less important ones.[176]

It is interesting to compare the practices of high-frequency traders and the stock exchanges with the MSPs discussed by David Evans in "Governing Bad Behavior By Users of Multi-sided Platforms."[177] Evans shows that, in general, MSPs will act to maximize the value of their platform (and by extension, the profits of the platform itself) by developing governance mechanisms that reduce harmful behavior on the part of participants. An example of this would be the governance of trading practices at the London Stock Exchange in the 1700s through the 1900s.[178] If unchecked, market manipulation and insider trading would lead to distrust of the exchange and a reluctance of investors to trade there, so the exchange itself prohibited such conduct. The main mechanism available to a platform to police members who engage in manipulative or abusive activity is to exclude them—the "Bouncer's Right."[179]

Query however whether the exchanges are currently incentivized to exclude traders who create negative externalities for others, particularly if those traders are responsible for a large fraction of the total trading volume. The exchanges, in an intense competition for order flow, lack the real ability to exclude traders, particularly their most active ones. Furthermore, they are tempted to work with those more active and very sophisticated traders in profiting from the less agile order flow of institutional traders, in effect taking a cut of the profits gleaned from HFT activity. This is most obvious with the increasing fees charged for co-location, but many practices of the exchanges arguably amount to taking a cut of the profits HFT makes off of slower traders.

It is also interesting to reflect on the role of Regulation NMS here. The operation of the Order Protection Rule, which mandates

176. *See, e.g.*, PATTERSON, *supra* note 21, at 204–05 ("High-speed firms worked hand in hand with the trading networks to create exotic order types that would behave in very specific ways."); LEWIS, *supra* note 21, at 163 ("By giving HFT what it wanted (speed, in relation to the rest of the market; complexity only HFT understood; and payment to brokers for their customers' orders, so that HFT had something to trade against), the new stock exchanges had stolen market share from the old stock exchanges.").

177. Evans, *Governing Bad Behavior*, *supra* note 69.

178. *See id.* at 1232–35.

179. *See id.* at 1221 (citing Lior Jacob Strahilevitz, *Information Asymmetries and the Rights to Exclude*, 104 MICH. L. REV. 1835 (2006)).

routing an order to the exchange currently posting the NBBO, appears to interfere with the natural operation of the liquidity network effect, which would otherwise lead to the concentration of trading on one venue or a small number of venues.[180] With less competitive pressure on the exchanges, the ability to exclude would likely be much greater, as the examples Evans provides suggest.[181] The London Stock Exchange was to some extent a monopoly MSP in its market,[182] and to some degree Facebook (for social media) and Google (online search) are as well. One of the drawbacks of Regulation NMS then is that it appears to cut against the incentives the stock exchanges have traditionally had to police the conduct of those trading on them. By knitting together the various exchanges in a single, competitive system, Regulation NMS appears to have incentivized bad behavior on the part of the exchanges themselves, albeit in a competition in which their very existence is at stake.

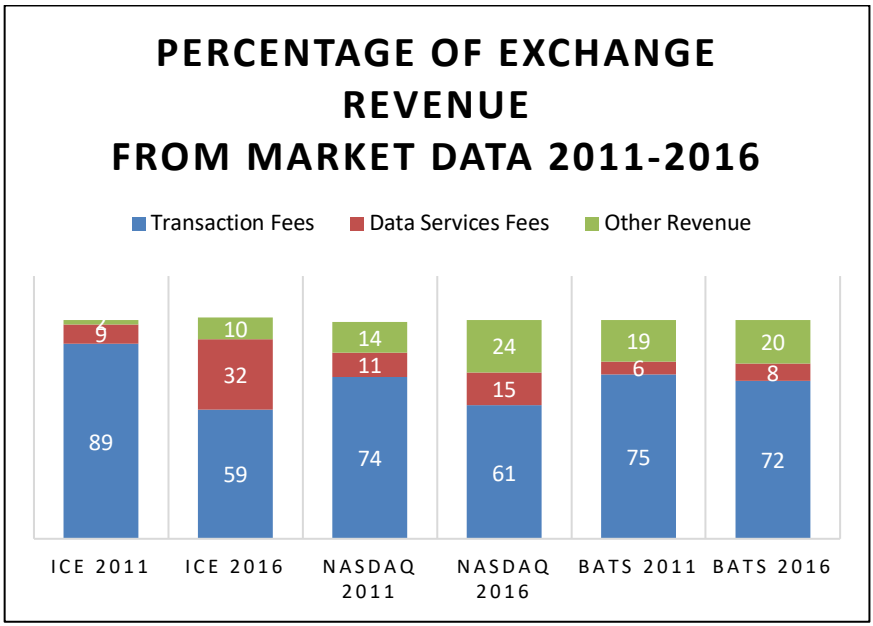
c. Rising data fees and the NetCoalition *litigation*. The phenomena reviewed so far involve exchanges competing with one another for order flow, that is, market share in their core business of facilitating transactions. While transaction revenue has been under pressure for decades, revenue from the sale of market data has filled the gap, accounting for an increasing share of exchange revenue in recent years[183]:

180. *See generally* Daniel M. Gallagher, Comm'r, Sec. & Exch. Comm'n, Remarks to the Georgetown University Center for Financial Markets and Policy Conference on Financial Markets Quality (Sept. 16, 2014); Spatt, *supra* note 163.

181. *See* Evans, *Governing Bad Behavior*, *supra* note 69, at 1226–40.

182. *See* R.C. Michie, *The London Stock Exchange and the British Securities Market, 1850–1914*, 38 ECON. HIST. REV. 61, 75 (1985).

183. *See SEC Review Threatens Growth of Exchanges' Lucrative Market Data Fees*, S&P GLOBAL MKT. INTELLIGENCE (Nov. 7, 2016, 9:45 AM), https://www.spglobal.com/marketintelligence/en/news-insights/research/sec-review-threatens-growth-of-exchanges-lucrative-market-data-fees (data source: SNL Financial).



The steady increase in revenue from information services, amounting to approximately 62% growth over the past five years,[184] has prompted criticism that the exchanges are effectively acting as a monopoly, or more accurately, oligopoly, supplier of trading data.[185] In earlier decades, of course, the NYSE was a de facto monopoly, and its members profited from both high commissions and wide spreads on stock trades.[186] In the current environment, traders and other parties who use the proprietary data sold by the exchanges again charge that the exchanges are acting in a monopolistic fashion, this time by exploiting their market power to

184. *See* Robin Wigglesworth, Nicole Bullock & Gregory Meyer, *Costly Data Battle Heats Up Between Traders and Equity Exchanges*, FIN. TIMES (July 5, 2016), https://www.ft.com/content/785092ec-33d8-11e6-ad39-3fee5ffe5b5b.

185. *See, e.g.*, HEALTHY MKTS. ASS'N, U.S. EQUITY MARKET DATA: HOW CONFLICTS OF INTEREST OVERWHELM AN OUTDATED REGULATORY MODEL AND MARKET PARTICIPANTS (Nov. 2017); U.S. DEP'T OF THE TREASURY, A FINANCIAL SYSTEM THAT CREATES ECONOMIC OPPORTUNITIES: CAPITAL MARKETS 63 (Oct. 2017); Dave Michaels, *Traders Want to Know What Exchanges Earn from Market Data*, WALL STREET J. (Dec. 6, 2017, 6:20 PM), https://www.wsj.com/articles/traders-want-to-know-what-exchanges-earn-from-market-data-1512592201; Larry Tabb, *Fight over Market Data Fees Is Going to Get Ugly*, TABBFORUM (Dec. 28, 2016), https://tabbforum.com/opinions/fight-over-market-data-fees-is-going-to-get-ugly.

186. *See* McNamara, *supra* note 175, at 78–81.

charge unjustified rates to traders who need this data to compete or even to comply with a broker's duty of best execution. The debate over market data fees is intense and ongoing, with an important 2016 SEC decision in the long-running *NetCoalition* (now *SIFMA*) litigation currently on appeal.[187] It also raises the question of whether the exchanges, which are knit together into a single, virtual marketplace through the national market system, in fact constitute a sort of super-MSP. While the exchanges compete vigorously for market share, they do appear to be acting in concert in gradually raising their market data fees.[188]

The market data litigation began in response to the SEC's 2006 order approving the imposition of fees by NYSE Arca for its "ArcaBook" data, which provides traders with complete information concerning the depth of liquidity in a stock.[189] The petitioner, NetCoalition, was a group of twenty internet companies and SIFMA, which represented over 600 trading firms.[190] NetCoalition argued that the SEC erred in approving the fees, because they violated Regulation NMS Rule 603(a)'s requirement that prices for "non-core" data be both "fair and reasonable" and "not unfairly discriminatory."[191] The Court of Appeals for the D.C. Circuit

187. *See* NetCoalition v. SEC, 615 F.3d 525, 527 (D.C. Cir. 2010).

188. While not at issue in the *Netcoalition* litigation, the gradual rise in prices first by one exchange and then others is suggestive of "tacit collusion" or "conscious parallelism." While "[c]ircumstantial evidence can establish an antitrust conspiracy," purely tacit collusion is not actionable under current U.S. antitrust law. *In re* Text Messaging Antitrust Litig., 782 F.3d 867, 871 (7th Cir. 2015) (quoting *In re* Text Messaging Antitrust Litig., 630 F.3d, 622, 627–29 (7th Cir. 2010)); *see also* Edward J. Green, Robert C. Marshall & Leslie M. Marx, *Tacit Collusion in Oligopoly*, *in* 2 THE OXFORD HANDBOOK OF INTERNATIONAL ANTITRUST ECONOMICS 464 (Roger D. Blair & D. Daniel Sokol eds., 2015).

189. Order Setting Aside Action by Delegated Authority and Approving Proposed Rule Change Relating to NYSE Arca Data, Exchange Act Release No. 34-59039, 73 Fed. Reg. 74,770–01 (Dec. 9, 2008).

190. *See NetCoalition*, 615 F.3d at 527.

191. 17 C.F.R. § 242.603(a) (2018). "Core data" is required to be reported to the SIP by the exchanges, and consists of

> (1) last sale reports, which include the price at which the latest sale of the security occurred, the size of the sale and the exchange where it took place;
> (2) the current highest bid and lowest offer for the security, along with the number of shares available at those prices, at each exchange; and
> (3) the 'national best bid and offer,' or NBBO, which are the highest bid and lowest offer currently available in the country and the exchange(s) where those prices are available.

NetCoalition, 615 F.3d at 529.

rejected NetCoalition's argument that the SEC should use a "cost-based approach" and not a "market-based approach" but agreed with the Petitioner that the SEC had not presented evidence sufficient to justify the fee imposition. The court also stated that it did not mean that "a cost analysis is irrelevant. On the contrary, in a competitive market, the price of a product is supposed to approach its marginal cost," and "[s]upracompetitive pricing may be evidence of 'monopoly,' or 'market,' power."[192] The court found that even though a market-based approach is sufficient, the SEC failed to "require NYSE Arca to substantiate its market data costs."[193] General statements that order flow competition was "fierce"[194] failed to justify "the SEC's conclusion that order flow competition constrains market data prices."[195] The court also found that the SEC had lacked sufficient evidence to conclude that purchasers would substitute another product for the NYSE Arca depth-of-book offering "instead of paying a supracompetitive price."[196] The order was thus vacated and remanded to the SEC for further proceedings.

After the D.C. Circuit determined in 2013 that it lacked jurisdiction over the case due to the new section 19(b)(3)(C) of the Exchange Act, the case was returned to the SEC, where it was assigned to Chief Administrative Law Judge Brenda Murray.[197] As directed by the D.C. Circuit, Judge Murray focused on evidence that competition constrains the pricing of non-core market data by the exchanges. Expert witness testimony established that of 350,000 professional subscribers to NASDAQ's data services, 30,000 of

"Non-core data" is all other data generated by the exchanges, and most importantly includes quotes deeper in the order book, allowing a trader to see all shares available at an exchange. *Id.* at 529–30; *see also* Regulation NMS, 70 Fed. Reg. 37,496, 37,567 (June 29, 2005).

192. *NetCoalition*, 615 F.3d at 537 (citing Tejas Power Corp. v. FERC, 908 F.2d 998, 1004 (D.C. Cir. 1990)).

193. *Id.* at 538.

194. *Id.* at 539.

195. *Id.* at 541.

196. *Id.* at 544.

197. Exchange Act section 19(b)(C)(3) was amended by section 916 of The Dodd-Frank Act of 2010, entitled "Streamlining of filing procedures for self-regulatory organizations." It allows changes to rules setting fees to take effect immediately upon filing with the SEC, and removes the ability of the SEC to "abrogate" a rule change within sixty days of filing with the SEC. Such a change effectively shifts the burden of any decision as to the validity of an SRO rule change from the SRO to the SEC. *See* Rules of Practice, Exchange Act Release No. 34-63723, 76 Fed. Reg. 4066 (Jan. 24, 2011).

them purchased the most comprehensive "Total View" service.[198] Furthermore, of these 30,000 subscriptions, 5000 were for servers operated by an estimated 100 HFT firms. These firms were responsible for approximately 90% of trading on NASDAQ.[199] Because the total amount of trading is concentrated in a relatively small number of firms, these firms allegedly held "the upper hand" in negotiations with NASDAQ. While the exchanges' expert witnesses conceded that HFT firms would require data feeds from all exchanges, Judge Murray found that for the vast majority of exchange customers, one exchange's non-core data was a substitute for another's.[200] This conclusion is supported by the fact that trading in most stocks is dispersed across a number of exchanges, not concentrated on one.[201] Judge Murray also noted that there was substantial evidence that the exchanges considered competitive pressure and pushback from clients when considering rate increases.[202] There was also evidence that one trader diverted order flow from NASDAQ in response to a rate increase in 2012.[203]

A number of points are important to note. First, SIFMA expert witness David Evans relied on the concept of an MSP to argue that the exchanges cross-subsidize their highly competitive transaction businesses by charging high prices for market data, where demand is more inelastic.[204] NYSE expert witness Terrence Hendershott conceded that the exchanges are "multi-product" firms but resisted their characterization as MSPs. This is likely because the MSP literature indicates that it is natural for one side to be charged a higher price, whereas the thrust of the Defendants' argument was that competition for order flow, in fact, constrains market data pricing.[205] Second, Judge Murray lumps together high-frequency traders with all other traders when considering the entire class of data customers, yet focuses on them in isolation in order to

198. *See* Sec. Indus. & Fin. Mkts. Ass'n, Release No. 1015, 114 SEC Docket 1388, at 12 (ALJ June 1, 2016) (initial decision).
199. *Id.*
200. *See id.* at 33.
201. *See id.*
202. *See id.* at 34.
203. *See id.* at 37–38.
204. *See id.* at 24–25.
205. *See id.* at 17.

demonstrate the power they have over the exchanges. Certainly for the HFT firms depth-of-book data from every exchange is not optional, so for their most important traders, one exchange's offering is decidedly *not* a substitute for another's.[206] While Judge Murray concedes this point, because HFT firms are by far the exchanges' most important customers, it is in considerable tension with the insistence on the substitutability of depth-of-book from the various exchanges.[207] Third, Judge Murray rejects the notion that a statistical or econometric review of pricing data is necessary here: "Statistical evidence is not required to resolve every dispute."[208] Although the exchanges were only able to point to one instance of a trader diverting order flow from NASDAQ in response to a price increase, evidence that the exchanges devoted sufficient resources to marketing their data offerings and to considering customer reaction to fee increases suffices to show that "depth-of-book prices are constrained by order flow competition."[209] Nor are estimated profit margins of 70–85% on data products determinative, despite the statement of the D.C. Circuit that "in a competitive market, the price of a product is supposed to approach its marginal cost."[210] Even conceding accounting questions as to true marginal cost, such margins seem high. Finally, Judge Murray seems out of touch with HFT practice when she determines that because "nearly 97% of all trades occur at or within the NBBO . . . most customers do not require any sort of depth-of-book data."[211] While a trade may occur

206. *See, e.g.*, U.S. DEP'T OF THE TREASURY, *supra* note 185, at 63 ("[T]he market for proprietary data feeds is not fully competitive. For use in making routing and trading decisions for active or institutional size order flow, data from one exchange's feed cannot substitute for data from another exchange's feed.").

207. *See* Sec. Indus. & Fin. Mkts. Ass'n, Release No. 1015, 114 SEC Docket 1388, at 36 (ALJ June 1, 2016) (initial decision). Judge Murray acknowledges that high-frequency traders may require depth-of-book data from all exchanges, but she emphasizes that such traders "reflect only a small percentage of all market participants." *Id.*; *cf.* NetCoalition v. SEC, 615 F.3d 525, 540 (D.C. Cir. 2010) (pointing out the tension between the SEC's belief that the exchanges face significant competitive pressures in selling their market data and its insistence that "depth-of-book data is simply not very important to most traders, even professionals.").

208. Sec. Indus. & Fin. Mkts. Ass'n Release No. 1015, 114 SEC Docket 1388, at 42 (ALJ June 1, 2016) (initial decision).

209. *Id.*

210. *NetCoalition*, 615 F.3d at 537.

211. Sec. Indus. & Fin. Mkts. Ass'n, Release No. 1015, 114 SEC Docket 1388, at 36 (ALJ June 1, 2016) (initial decision). This determination relies on expert witness Terrence

within the parameters of the NBBO, that does not mean that depth-of-book data was not instrumental in determining whether or not it was a good bet.[212] Indeed, if 90% of trades on NASDAQ are conducted by HFT shops, which use depth-of-book data intensively, one suspects that more than a mere 3% of trades rely on such data. The decision may result in another appeal to the federal courts.[213] Given the amount of revenue and the intrinsic importance of the issues at stake, the outcome will be closely watched.

This review of the economics of exchange trading illustrates the dual nature of the current national market system.[214] In the main, the exchanges function as a linked network of competing MSPs. Regulation NMS breaks the natural operation of the network effect of an individual exchange as an MSP, contributing to the existence and survival of multiple exchanges. The resulting competition is fierce, leading to conduct on the part of the exchanges that has on occasion amounted to offering up their less-savvy customers to their savvier and best ones, the HFT shops. On the other hand, insofar as the exchanges have market power in their pricing of data,

---

Hendershott's testimony that most traders do not need depth-of-book data: "Hendershott again cited the fact that 96.7% of trades occur at the NBBO prices that are provided by core data." *Id.* at 18.

212. Expert witness Hendershott "conceded that depth-of-book data may be useful to certain market participants, such as high-frequency traders or traders who rely on algorithmic computer models." *Id.* For discussion of this usefulness, see Jonathan Brogaard, Terrence Hendershott & Ryan Riordan, *High-Frequency Trading and Price Discovery*, 27 REV. FIN. STUD. 2267, 2300 (2014) ("The results show that HFTs' order flow is correlated with information embedded in the limit order book.").

213. Judge Murray affirmed her 2016 decision on Dec. 21, 2017. *See* Sec. Indus. & Fin. Mkts. Ass'n, Release No. 5414, File No. 3-15350 (ALJ Dec. 21, 2017), https://www.sec.gov/alj/aljorders/2017/ap-5414.pdf. As this article was going to press, the SEC reversed Judge Murray's decision. The SEC Commissioners found that NASDAQ and NYSE Arca had not met their burden of proof in demonstrating that the proposed fee increases were fair and reasonable. *See* Sec. Indus. & Fin. Mkts. Ass'n, Exchange Act Release No. 84423, 2018 WL 5023228, at *3 (Oct. 16, 2018). This decision represents a new and important stage in this very long-running legal battle. While it does not declare the fee increases unjustified, it does represent a significant reversal for the exchanges and may portend new regulatory scrutiny of their data fee businesses. *See SEC Issues Landmark Order Rejecting Nasdaq and NYSE Arca Market Data Fee Increases*, SIDLEY UPDATES (OCT. 29, 2018), https://www.sidley.com/en/insights/newsupdates/2018/10/sec-issues-landmark-order-rejecting-nasdaq-and-nyse-arca-market-data-fee-increases.

214. O'Hara and Ye capture this dual nature with their observation that "while U.S. equity markets are spatially fragmented, they are, in fact, virtually consolidated into a single market with many points of entry." Maureen O'Hara & Mao Ye, *Is Market Fragmentation Harming Market Quality?*, 100 J. FIN. ECON. 459, 460–61 (2011).

and to the extent they exercise it in tacit agreement with one another, the system as a whole has the loose character of an oligopoly and thus a sort of super-MSP.

III. THE STOCK EXCHANGE AS VIRTUAL CENTRAL LIMIT ORDER BOOK

Given the nature of the stock exchange as a digital MSP, what might the stock exchange of the future look like? Perhaps more speculatively, what should it look like? In recent years, a multitude of reforms have been proposed, including batched auctions,[215] banning maker-taker payments,[216] curbing high rates of order cancellation,[217] subjecting high-frequency traders to broker-dealer regulation,[218] and even banning high-frequency trading altogether.[219] Some of these reforms might be beneficial, while others would seem to be mere band-aids. The notion of banning HFT altogether is nonsensical—HFT is not one thing, but many, and putting the information technology genie back in the bottle is as implausible in the financial markets as it would be anywhere else.

Since the fragmentation of the current system gives rise to the bulk of the problems of the current market, and the stock exchange as digital MSP carries with it a strong tendency toward centralization, consideration of a market structure offering a controlled return to centralization is useful at this point. In fact, the not-too-distant history of regulatory reform provides a blueprint for how such a centralization might be achieved: through the implementation of a central limit order book or "CLOB." Part III sketches a virtual CLOB and discusses a number of specific problems it would either solve or mitigate.

215. *See* Eric Budish, Peter Cramton & John Shim, *The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response*, 130 Q.J. ECON. 1547 (2015).

216. *See* Macey & Swensen, *supra* note 155.

217. *See* GARY SHORTER & RENA S. MILLER, CONG. RESEARCH SERV., R43608, HIGH-FREQUENCY TRADING: BACKGROUND, CONCERNS, AND REGULATORY DEVELOPMENTS (2014).

218. *See* Mary Jo White, Chair, SEC, Address at the Sandler O'Neill & Partners, L.P. Global Exchange and Brokerage Conference: Enhancing Our Equity Market Structure (June 5, 2014), https://www.sec.gov/news/speech/2014-spch060514mjw.

219. *See* Michael Shields, *ECB's Nowotny Calls for High-Frequency Trade Ban*, REUTERS BUS. NEWS (Sept. 13, 2012, 6:52 AM), https://www.reuters.com/article/us-ecb-nowotny-supervision/ecbs-nowotny-calls-for-high-frequency-trade-ban-idUSBRE88C0O020120913.

A. Regulatory Design

Implementation of a CLOB could take two basic forms. In the first and simplest, the exchange could exist as a single institution which would handle all stock trading. An obvious objection is that the exchange would then have a monopoly position and so be able to charge monopoly rents for its services.[220] A monopoly exchange would also have little reason to innovate, and as a "single point of failure" would likely face heightened cybersecurity risks.[221] Because of the inherent problems of a monopoly exchange, it is tempting to look at more nuanced proposals that amount to the institution of a "virtual CLOB."

Instead of moving to a single exchange, a central routing service for orders could be set in place that would function as a virtual CLOB while allowing the current exchanges to remain in existence. An expansion of the current Consolidated Quotation System could accomplish this. Instead of merely collecting the best quotes from each exchange, and then selecting the best of these to present as the NBBO, a routing service would collect the entire book of quotes in each stock from the various exchanges. They would then be compiled into a national consolidated order book. This system-wide order book would be made available to traders, and all incoming orders would be transacted on the basis of the consolidated order book. The prices and times as recorded in the consolidated order book would be the basis for trades; when a preexisting quote was hit, it would be sent back to the hosting exchange for processing. A central routing service would therefore function as the single determinative node in the system, eliminating problems which arise from fragmentation. Such a proposal would also by definition do away with the need for the Order Protection Rule. Since there would be a single, central file of all orders, and all incoming orders would trade against this one book, it would by definition provide "depth-of-book" protection for incoming orders

220. *See* Oesterle, *supra* note 35, at 636–37 (explaining that the mid-twentieth century dominance of the NYSE allowed it to generate monopoly rents).

221. *See* Kristin N. Johnson, *Cyber Risks: Emerging Risk Management Concerns for Financial Institutions*, 50 GA. L. REV. 131 (2015); Tom C.W. Lin, *Compliance, Technology, and Modern Finance*, 11 BROOK. J. CORP. FIN. & COM. L. 159 (2016).

as well as, of course, top-of-the-book protection.[222] This would mean that protection against trade-throughs would extend to all quotes, not just the very best quote in the market at the time, as in the present system.

In addition to depth-of-book protection, a routing service as virtual CLOB would also enforce strict time priority in the market, giving it the character of a "hard CLOB."[223] This would be unlike the current rule, which protects just price priority. Any quote previously entered, therefore, from whatever exchange, at the same price, would stand ahead in line of later entered quotes at the same price. In this way a central routing service would implement the standard rule of price-time priority used at the individual exchanges on the level of the national market system as a whole, something the current system doesn't do.[224]

The basic principles of a virtual CLOB were in fact presented at various points by the SEC from the early 1970s until the promulgation of Regulation NMS.[225] In 1976 the SEC set forth the principles of a "Composite Central Limit Order Repository" for public comment, and in 1978 it stated its belief that a "central limit order file" presented the best means to achieve national limit order protection under the 1975 Amendments. The 1978 Release stated that:

> The objectives of a Central File are relatively simple: to make available a mechanism in which public limit orders can be entered and queued for execution in accordance with the auction trading principles of price and time priority and by means of which such orders can be assured of receiving an execution prior to the execution of any other order by a broker or dealer in any market at the same or an inferior price (determining that price by reference to the price required to be reported in the consolidated

222. *See* Stoll, *supra* note 53, at 171–72 (explaining how the Order Protection Rule only provides top-of-the-book protection).

223. *See* Oesterle, *supra* note 35, at 644; *see also* Robert L.D. Colby & Erik R. Sirri, *Consolidation and Competition in the U.S. Equity Markets*, 5 CAP. MKTS. L.J. 169, 177–78 (2010).

224. *See supra* note 6 and accompanying text.

225. *See* Osterle, *supra* note 35; Junius W. Peake, *Entropy and the National Market System*, 1 BROOK. J. CORP. FIN. & COM. L. 301 (2006).

> system pursuant to Rule 17a–15, in the event of a completed transaction).[226]

This proposal was not meant to eliminate the various exchanges, but it would create a central limit order "book" or "file" that would have the effect of extending both depth-of-book and time priority to the national system at large.

The proposals in the late 1970s met with vociferous opposition from the NYSE and other exchanges, which inferred that their implementation would likely spell the end of trading on the exchange floor.[227] As a compromise, the Inter-Market Trading System or ITS was developed, which allowed for, but did not mandate, the routing of orders to exchanges displaying better quotes.[228] The SEC floated the idea of a hard CLOB once again in 2000 in its request for comment on issues relating to market fragmentation, but has not broached the idea since.[229] Most importantly, Regulation NMS pulls back from implementation of a CLOB. Rule 611 and the other rules offer a very limited version of a CLOB, with only top-of-the-book quotations given protection.[230]

As with the earlier CLOB proposals, an updated proposal should not abolish the individual exchanges but should instead develop a routing service on the model of the existing Plans operating the SIPs. With a virtual CLOB as routing service, now the entire book of each exchange's limit offers in a security would be collated with every other exchange's book to present a consolidated limit offer book or file. This book would be available to market participants, just as each exchange currently offers depth-of-book information as part of its proprietary information offerings. And just as with the CTA and CQS Plans, revenue from the sale of this information would be allocated to the various exchanges. Such a

226. Development of National Market System, Exchange Act Release No. 34-14416, 43 Fed. Reg. 4354, 4359 (Feb. 1, 1978) (footnotes omitted).

227. *See* Oesterle, *supra* note 35, at 639; Peake, *supra* note 225, at 308.

228. *See* Oesterle, *supra* note 35, at 639.

229. *See* Fragmentation Concept Release, Exchange Act Release No. 34-42450, 65 Fed. Reg. 10,577 (Feb. 28, 2000).

230. *See* Stoll, *supra* note 53; *see also* Craig Pirrong, *SLOB vs. CLOB*, STREETWISE PROFESSOR (Aug. 10, 2013, 1:00 PM), https://streetwiseprofessor.com/slob-vs-clob/ (referring to the current system as a "Simulacrum Limit Order Book"); Roger D. Blanc, *Intermarket Competition and Monopoly Power in the U.S. Stock Markets*, 1 BROOK. J. CORP. FIN. & COM. L. 273, 283 (2006) (calling the Order Protection Rule "Trade-Through Rule Lite").

plan would present the same information as the proprietary data offerings of the exchanges, although the use of this information would presumably change given the great reduction in fragmentation such a proposal entails. Since the central order book would be the single book or file used to govern the allocation of all orders in the system, it would function as the central nervous system of the entire national market system, with important implications for the provision of co-location services and proprietary data services, as discussed below.

While this proposal moves much closer to implementing a single market for securities, it should be remembered that it does not mandate the institution of a single exchange. Leaving the current exchanges in existence would preserve competition in the market for exchange services, helping to avoid the obvious problems of a single monopoly exchange. The governance of the routing system would present a primary challenge. Just as with the CTA and the CQS, any entity providing information to the market at large has elements of a public utility,[231] and complaints have been raised about the governance of the current plans.[232] The SEC should ensure that a wide variety of interests, including those of non-high-frequency traders such as institutional and retail investors, are represented there.

B. Problems Solved by a Virtual Central Limit Order Book

The implementation of a CLOB would immediately cause many of the most important problems since the implementation of Regulation NMS to vanish. It would also significantly mitigate others, although it would also create new governance challenges.

231. *See* Oesterle, *supra* note 35, at 625 (explaining Congress's understanding in 1975 that an exclusive cross-market information processor is "a public utility" that "should be regulated accordingly").

232. *See* Bradley Hope, *Nasdaq Cancels Deal to Operate Data Feed*, WALL STREET J. (Jan. 14, 2014, 6:32 PM), https://www.wsj.com/articles/nasdaq-to-stop-operating-securities-information-processor-1389727668; Ivy Schmerken, *SIFMA Criticizes SIP Selection Process for Lack of Transparency*, INFORMATIONWEEK WALLSTREET & TECH. (Oct. 17, 2014, 9:38 AM), http://www.wallstreetandtech.com/data-management/sifma-criticizes-sip-selection-process-for-lack-of-transparency/d/d-id/1316720.html.

And by reducing adverse selection risk and giving depth-of-book protection, dark fragmentation may also decrease as well.[233]

1. The staleness of the SIP disappears

Along with the implementation of a single routing service would come a single source of trading information. Since the information presented by the virtual CLOB would be the single source of determinative information on stock quotes, there would be no need for a separate SIP. Currently, there can exist a true NBBO that is not yet reflected in the NBBO presented by the SIP.[234] With a central routing service offering a national order book, the problem of stale quotes and the NBBO would vanish.

2. Arbitrage due to fragmentation lessens dramatically

Because the time lag between the servers of the various exchanges is meaningful, HFT firms can engage in "latency arbitrage," exploiting the timing gaps between the exchanges and the SIP.[235] While the profits generated by this activity appear to be in decline, it is a symbolically important phenomenon that is captured by the phrase "front-running." Illegal front-running occurs when a broker trades ahead of its client's order, before a price change occurs that would be caused by that order. Front-running is against both the rules of the exchanges and the common law applicable to broker-dealers and is a violation of a broker's fiduciary duty to its client.[236]

Because an HFT firm almost always trades on a proprietary basis, not on behalf of a client, a high-frequency trader engages in "front-running" only in the sense that it discerns market activity and is able to trade ahead of it. When a trader is not front-running a client order, such activity is not illegal. Nevertheless, it is symbolically important.[237] The narrative of *Flash Boys* centers

233. *See* Colby & Sirri, *supra* note 223, at 178.

234. *See* Ding et al., *supra* note 43.

235. *See* Fox et al., *supra* note 29, at 226–42.

236. Front-running on the part of broker-dealers is illegal under both the common law and SRO regulations. *See* Opper v. Hancock Sec. Corp., 250 F. Supp. 668 (S.D.N.Y. 1966), *aff'd*, 367 F.2d 157 (2d Cir. 1966); *FINRA Rule 5270*, FINRA, http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=10860 (last visited Jan. 23, 2019). For a discussion of the problems HFT poses for insider trading law, see Yadav, *supra* note 75.

237. *See* McNamara, *supra* note 175.

around such front-running: when protagonist Katsuyama cannot hit listed quotes because the HFT firms posting them pull them ahead of his orders, he claims the markets are "rigged."[238] Even though electronic front-running on the part of high-frequency traders is legal, it strikes many long-term investors as fundamentally unfair.

By definition, a virtual CLOB would eliminate arbitrage activity generated by the timing gaps between the exchanges. While it would not eliminate all arbitrage—think of arbitrage between market centers in Europe or Asia and those in America, between commodities and options markets in Chicago and stock exchange servers in suburban New Jersey, or even between dark pools and a virtual CLOB—arbitrage solely between different American stock markets would be eliminated.

3. True "depth-of-book" protection implemented

A limit order book will have all available limit orders in a stock listed on either side of the NBBO, thereby displaying, in a single place, all the available shares in a given stock and their various prices. Generally speaking, as prices fall away from the NBBO there will be more shares available because, the greater amount a purchaser will be willing to pay, the more shares it would be offered, and vice versa. The total number of shares available for purchase comprises the "depth-of-book" that constitutes the liquidity in a stock.[239] In a CLOB, because all shares are listed on a single book, a buyer or seller has true depth-of-book protection for their orders.

In the years preceding the implementation of Regulation NMS, certain traders and economists advocated for a true depth-of-book system, whereby the Order Protection Rule would require an exchange to send orders to any other exchange posting a better quote, no matter how deep in the order book those quotes were.[240]

238. *See* LEWIS, *supra* note 21, at 40.

239. *See* Spatt, *supra* note 163, at 4–5.

240. *See, e.g.*, *The Emerging Structure of U.S. Securities Markets and the Appropriate Role for Regulation: Hearing Before the S. Comm. on Banking, Housing, & Urban Affairs*, 106th Cong. 56–61 (2000) (statement of Henry M. Paulson, Chairman and CEO, Goldman Sachs & Co.); Junius W. Peake, Comment Letter on Proposed Rule Change and Amendment No. 1 Thereto

The incumbent exchanges, on the other hand, fought this proposal vigorously.[241]

Economically, a true depth-of-book system would represent the best state of affairs for traders, as it would allow them to access all available shares at the best prices. A single limit order book by definition would implement such a system, since all available orders to buy and sell would be collated into a single list. Not only would arbitrage between the various exchanges be eliminated but all orders up and down the book would be present to traders, who would be assured they would receive all available shares at the best possible prices.

4. Complex order types reduced (and prohibition on locks and crosses eliminated)

As the exchanges attempted to cater to their customers in the past decade, the variety of order types mushroomed.[242] Many of these new order types were designed to allow high-frequency traders, the exchanges' best customers, to position themselves advantageously as prices quickly moved around. The complex order types therefore came under fire from other traders. At best, they appeared to add significant complexity to the market microstructure and, at worst, to allow high-frequency traders a way to cut in line ahead of others.

Traditionally there were two basic order types: the limit order and the market order. As the national market system was implemented, a number of other order types became common. To name just a few, these include intermarket sweep orders ("ISOs"), pegged midpoint orders, flash orders, and orders designed to position traders at the top of the market when the market "unlocks" or "uncrosses" such as the notorious Hide Not Slide.[243] In order to prevent the irrational and confusing condition of locked or crossed markets, Rule 610(d) of Regulation NMS prohibits an exchange

Relating to Enhancements to the Exchange's Existing Automatic Execution Facility (Sept. 22, 2004), https://www.sec.gov/rules/sro/nyse/nyse200405/jwpeake092204.pdf; Black, *supra* note 86.

241. *See* Oesterle, *supra* note 35, at 639.

242. *See supra* notes 116–119 and accompanying text.

243. *See supra* notes 167–169 and accompanying text.

from posting a bid (buy) order at a price higher than the current best offer (sell order) or an offer at a price lower than the current best bid.[244]

In the current race for speed, however, traders often profit from being first in line when a market unlocks or uncrosses, as they will then be able to capture a "maker" rebate.[245] With the Hide Not Slide order, a trade could submit an order that would be hidden if it would lock or cross the market, yet would be placed first in line when the market changed. While this order allowed an HFT shop to eke out profits in such conditions, it appeared to critics that the exchanges were allowing them to "cut in line," so to speak.[246] While there would likely remain other complex order types, by doing away with multiple markets, a CLOB would undercut the need for order types that are predicated on a fragmented marketplace. Since locks and crosses would not be possible in a single market, the need for complex order types to negotiate them will disappear, as well as other orders designed to negotiate between different markets, such as intermarket sweep orders.

5. Broker-dealer best execution problems diminish

The duty of best execution on the part of stockbrokers serves as a backstop to the Order Protection Rule. This duty mandates that a broker send its client's order to the exchange where it will receive the "best execution."[247] Note that best execution is not defined solely in terms of price; speed and the desire to move markets as little as possible when placing an order are also legitimate considerations for brokers making routing decisions.

The inducements for order flow from brokers offered by the exchanges, ATSs, and large broker-dealers running internal matching engines illustrates the value of these orders. The inducements are primarily maker-taker rebates, but flash orders and other

244. 17 C.F.R. § 242.610(d) (2018) (prohibiting "[l]ocking or crossing quotations").

245. *See* Schmerken, *supra* note 169.

246. *See* Patterson & Strasburg, *supra* note 170.

247. *See* Fragmentation Concept Release, Exchange Act Release No. 34-42450, 65 Fed. Reg. 10,577, 10,584 (Feb. 28, 2000); *see also* Polise, *supra* note 4, at n.207; *FINRA Rule 5310*, FINRA, http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=10455 (last visited Jan. 23, 2019).

enticements to get high-frequency traders "to play in one exchange's sandbox" also play a role.[248] While maker-taker rebates have received criticism for their corrupting influence on stock-brokers, under an economic analysis they would seem to be a natural consequence of the value of liquidity.[249] Maker-taker rebates should therefore not be seen as an overriding problem with today's trading system. Also important are payments to brokers for order flow by large broker-dealers, such as Charles Schwab or Citadel, who then settle trades in their internal matching engines.[250] In addition to contributing to the movement of orders away from the lit exchanges to dark venues, academic studies indicate that such orders frequently do not receive the best price.[251] In such instances, broker-dealers will likely not have met their duty of best execution. On the other hand, it would seem that disclosure of maker-taker rebates on the part of brokers to their clients should be sufficient to deal with any disquiet due to maker-taker rebates, as these rebates naturally arise given the stock exchange's character as an MSP.

Nevertheless, were a virtual CLOB instituted, problems surrounding the broker's duty of best execution should diminish considerably. For a start, the routing decision would be reduced to the primary decision of whether to trade in the lit markets, which would use a single limit order book, or in a dark pool or other non-exchange venue. Furthermore, by offering strict price and time protection to all lit orders, a virtual CLOB should be expected to diminish the incentive to trade off-exchange.[252] The simplification of the markets through a virtual CLOB would greatly reduce the danger of brokers sending orders to venues where they were at risk of receiving less than best execution.

6. Complexity due to speed bumps disappears

A virtual CLOB would eliminate the problem of the increasing complexity of the market system through the implementation of

248. ARNUK & SALUZZI, *supra* note 45, at 84.
249. *See supra* note 163 and accompanying text.
250. *See* Fleckner, *supra* note 68, at 2570.
251. *See, e.g.*, Battalio et al., *supra* note 160.
252. *See* Colby & Sirri, *supra* note 223, at 178.

"speed bumps" by IEX and other exchanges. In response to the outcry over HFT front-running, IEX designed a system specifically designed to protect pegged midpoint orders from exploitation by high-frequency traders.[253] Their system subjects all incoming orders to a 350-microsecond "speed bump." When an order enters the IEX system, it traverses a thirty-eight-mile length of coiled optical fiber cable that delays the order for 350 microseconds (millionths of a second). Updates of pricing information from the other exchanges and the SIP are not however subject to this delay.[254] That means that incoming orders receive the benefit of any price changes occurring in the national market system at large. Most importantly, this delay prevents high-frequency traders from reacting to incoming orders before those orders can be filled, because any HFT order would itself be 350 microseconds behind the current, true NBBO. This system is meant to allow traders to operate without the adverse selection risk that trading on the normal lit exchanges entails.

Before the SEC approved IEX's application to become a registered exchange in June 2016,[255] the established exchanges and a few other players, such as Citadel, waged an intense battle to stop it.[256] A key objection was that the speed bump introduced a new and significant element of complexity into the trading environment.[257] By intentionally delaying incoming orders, the argument went, IEX's new system makes an already complicated system more complicated yet, and reduces the reliability of quotes in general. On legal grounds, the objectors in fact had a plausible argument that because of the built-in delay in IEX's system, it

253. *See* Andrew Upward, *IEX: An Intriguing Wrinkle in the National Market System*, WEEDEN & CO. (Mar. 17, 2016), www.weedenco.com/market-intelligence/market-structure.

254. *See* Investors' Exchange LLC, Comment Letter Regarding Investor's Exchange LLC Form 1 Application (Release No. 34-75925; File No. 10-222) (Nov. 13, 2015), https://www.sec.gov/comments/10-222/10222-20.pdf.

255. *See* Press Release, Sec. & Exch. Comm'n, SEC Approves IEX Proposal to Launch National Exchange, Issues Interpretation on Automated Securities Prices (June 17, 2016), https://www.sec.gov/news/pressrelease/2016-123.html.

256. *See* Annie Massa, *IEX Outduels Citadel, NYSE as 'Flash Boys' Exchange Approved*, BLOOMBERG (June 17, 2016), https://www.bloomberglaw.com/product/blaw/document/O8XZWL6JTSKT?bc=W1siU2VhcmNoIFJlc3VsdHMiLCIvcHJvZHVjdC9ibGF3L3NlYXJjaC9yZ.

257. *See* Dave Michaels, *Michaels's Take: Post-IEX Win, Speed-Bump Fight Goes On*, WALL STREET J. (July 1, 2016, 7:31 AM), https://www.wsj.com/articles/michaelss-take-post-iex-win-speed-bump-fight-goes-on-1467372715.

should not qualify as an "automated trading center," with its quotes entitled to "protected" status under Rule 611. The objectors focused on language from the Regulation NMS Adopting Release, which states that "[t]he term 'immediate' precludes any coding of automated systems or other type of intentional device that would delay the action taken with respect to a quotation."[258] The SEC, however, ruled in favor of IEX, determining that the 350 micro-second delay was merely "de minimis."[259] Setting aside the obvious questions—If the speed bump were de minimis, why would IEX build an entire business on it? And why such vehement opposition from the established players?—the SEC stated that the principle of providing for competition in the exchange space justified the introduction of this new and complex system. As denouement, the exchanges revealed that their objections were not in fact based on any important principle. They have since announced plans to offer exchanges with speed bumps themselves.[260]

Whatever the final import of this battle, by doing away with the opportunity for arbitrage between the various exchanges, a CLOB would eliminate the need for speed bumps as a defensive mechanism offered by an exchange.

7. Avoidance of piecemeal reforms

A final and not insignificant benefit of the reform proposed here would be the avoidance of smaller piecemeal reforms. A number of such reforms have been proposed: An access fee pilot program that would experiment with smaller fees under Rule 610, thereby lowering the permissible maker-taker rebates;[261] batched auctions;[262]

258. Regulation NMS, Exchange Act Release No. 34-51808, 70 Fed. Reg. 37,496, 37,534 (June 29, 2005).

259. Inv'rs Exch., LLC for Registration as a Nat'l Sec. Exch., Exchange Act Release No. 34-78101, 114 SEC Docket 2064, at 77 (June 17, 2016), https://www.sec.gov/rules/other/2016/34-78101.pdf.

260. *See* John D'Antona Jr., *NYSE American Marketplace Launches*, TRADERS MAG. (July 25, 2017), http://www.tradersmagazine.com/news/ecns_and_exchanges/nyse-american-marketplace-launches-116480-1.html.

261. *See* SEC EQUITY MKT. STRUCTURE ADVISORY COMM., RECOMMENDATION FOR AN ACCESS FEE PILOT (July 8, 2016), https://www.sec.gov/spotlight/emsac/recommendation-access-fee-pilot.pdf; *Missing: The Access Fee Pilot Proposal*, THEMIS TRADING BLOG (Nov. 29, 2017), http://blog.themistrading.com/2017/11/missing-the-access-fee-pilot-proposal/.

262. *See* Budish et al., *supra* note 215.

revisiting the SEC's interpretation of Rule 603(a)(2), which enables co-location and proprietary data feeds by allowing the exchanges to transmit data to HFT servers at the same time it is sent to the SIP, thereby greatly aggravating the problem of the staleness of the SIP;[263] banning maker-taker rebates;[264] and subjecting high-frequency traders to broker-dealer regulation,[265] among others.

All these reforms would involve considerable regulatory study, proposals, and public comment, and would each entail second-order effects as well as significant new complexity in the trading system and its governing law. Most importantly, they are all intended to mitigate problems that fundamentally arise from the current system's fragmented character. The institution of a virtual CLOB would eliminate or mitigate many of the most important problems arising from this fragmentation and do so in a way that greatly simplifies the national market system.

CONCLUSION

This Article argues for a renewed look at the concept of a CLOB in light of the dysfunctions of the digital stock markets under Regulation NMS. Such a market structure would mitigate or eliminate many of the most important recent problems of the exchanges. It also comports with the economic nature of the stock exchange as an MSP. Given the investment of the exchanges and other important players in the current fragmented system, the institution of a CLOB is currently implausible as a matter of political economy. Nevertheless, this option should be kept in mind in future debate concerning the organization of the stock markets. Regulation NMS has only been in force for a decade, but the transformation of financial markets by digital technology has already put the national market system under considerable strain. Although the competition Regulation NMS incentivizes has reduced transaction costs considerably, its fragmented system has given rise to many other less obvious costs. Most importantly, the rising share of data revenues generated by the exchanges functions as a sort of tax on traders, and by extension investors, who are forced to pay for

263. *See Direct vs SIP Data Feed*, NANEX (Apr. 4, 2014), http://www.nanex.net/aqck2/4599.html; *see also* Fox et al., *supra* note 29, at 270–71.

264. *See* Macey & Swensen, *supra* note 155.

265. *See* White, *supra* note 218.

enhanced data feeds if they are to remain competitive and meet their duty of best execution.

It must be remembered that the need for this market information, like many of the more particular problems of fragmentation reviewed above, is itself a creation of Regulation NMS. While it is unlikely a CLOB will be instituted anytime soon, the economic and political events of the past decade caution against investing too much in the status quo. In a number of other areas of the economy, citizens, consumers, legislators, and regulators are beginning to grapple with the implications of dominant MSPs in areas as diverse as internet search, retail, and social media and news. While the particular regulatory landscape inhabited by the stock exchanges differs greatly from these other examples, dominant MSPs pose a fundamental challenge to the system of American capitalism, which has relied on a large number of enterprises to generate broadly distributed benefits from competition. Recognition of the powerful tendencies toward centralization in the world of digital capitalism and a willingness to implement governance structures with a wide variety of inputs may be necessary in the years ahead to preserve a workable system of capitalism in the digital environment of the future.