

Understanding the Market for U.S. Equity Market Data

Charles M. Jones¹

August 31, 2018

¹ Robert W. Lear Professor of Finance and Economics, Columbia Business School. I am solely responsible for the contents of this paper. I thank Larry Glosten, Frank Hatheway, Terry Hendershott, Stewart Mayhew, Jonathan Sokobin, and Chester Spatt for helpful discussions on these topics. I currently serve on FINRA's Economic Advisory Committee and on Nasdaq's Quality of Markets Committee, and I served as Visiting Economist at the New York Stock Exchange ("NYSE") in 2002–2003. The NYSE provided financial support for this research.

I. Executive Summary

A stock exchange facilitates share trading, in large part by developing computer systems, rules, and processes that allow buyers and sellers to submit orders, trade with each other, and determine a market price for shares listed on those exchanges. In the current market environment, this results in a vast amount of data, which market participants of all types rely on to make investment and trading decisions. Exchanges provide some of this market data to market participants at prices that vary depending on the type of data as well as how the data is used.

This paper provides an analysis of the market for equity market data in the United States. Unlike other data sources, U.S. equity market data is highly regulated by the Securities and Exchange Commission (“SEC”), and recently the SEC has been lobbied by entities arguing that exchanges charge too much. These entities have written comment letters and filed a number of proceedings with the SEC in an effort to reduce the prices of equity market data.

To determine whether these criticisms are valid, this paper provides an economic examination of market data, how it is used, and how it is regulated. This paper also presents data on market data prices and revenues and places them in context. Some of the data is newly public and is analyzed for the first time in this paper. Based on my review and analysis of this data, I show the following:

- Equity market data has value to the consumers of that data because it reflects the price discovery created by exchanges. Data consumers buy this aggregated data not to view their own orders and trades but rather to see the overall state of the orders and trades in a market. Market data products have seen substantial innovation over time, and the ability

to sell exchange proprietary market data products (as well as competition among trading venues) provides exchanges with incentives to continue to innovate.

- Using a variety of metrics, I find that exchange market data revenues are modest and stable over time.
- Exchange equity market data fees are a small cost for the industry overall: the data demonstrates that total exchange market data revenues are orders of magnitude smaller than (i) broker-dealer commissions, (ii) investment bank earnings from equity trading, and (iii) revenues earned by third-party vendors.
- The market is characterized by robust competition: exchanges compete with each other in selling proprietary market data products. They also compete with consolidated data feeds (discussed later in the paper) and with data provided by alternative trading systems (“ATSS”). Barriers to entry are very low, so existing exchanges must also take into account competition from new entrants, who generally try to build market share by offering their proprietary market data products for free for some period of time.
- Although there are regulatory requirements for some market participants to use consolidated data products, there is no requirement for market participants to purchase any proprietary market data product for regulatory purposes.
- There are a variety of data products, and consumers of equity market data choose among them based on their needs. Like most producers, exchanges offer a variety of market data products at different price levels. Advanced proprietary market data products provide greater value to those who subscribe. As in any other market, each potential subscriber takes the features and prices of available products into account in choosing what market data products to buy based on its business model.

- Although the market for U.S. equity market data is highly regulated, the regulatory arrangements have allowed a competitive market for data to operate effectively. This regulatory structure has allowed the development of a large suite of data products with a wide variety of features at differing price levels, and the resulting unparalleled transparency concerning stock trading activity is likely one of the reasons that U.S. equity market quality is the best in the world.

Section II of this paper details some of the ways that market participants use equity market data, and Section III provides a brief introduction to equity market data products.² Section IV discusses the basic economic features of the market for equity market data products. Section V discusses the regulation of equity market data, and Section VI discusses the pricing of equity market data, in particular the evolution of pricing over time and the revenue it actually generates for exchanges (a topic about which there seems to be significant confusion).

II. Introduction: The Many Uses of Equity Market Data

This paper provides an introduction and analysis of the market for stock market data in the United States. Dissemination of market data by U.S. stock exchanges is regulated by the SEC. However, there are several different kinds of equity market data, and this market data is sold and regulated in a variety of ways. A better understanding of the market for market data is

² Because understanding the current regulatory framework is a key part of understanding the overall market for market data, the Appendix provides some historical context and an overview of the National Market System (“NMS”) that underlies the current regulatory framework in the United States.

essential for every market participant, as well as regulators, policy-makers, and academics with an interest in equity markets.

What does a modern stock exchange do? One of its most important roles is to facilitate the trading of shares in publicly listed companies. Today a stock exchange develops and operates sophisticated technology, sets up rules, and puts together a set of trading processes that allow buyers and sellers to learn about the level of trading interest, to submit orders, and to transact with each other. Aggregated together, market participants' buy and sell orders contribute to "price discovery," which is simply the determination of a market price for the shares. Of course, that market price varies from day to day, and even from second to second, because buyers and sellers regularly arrive, depart, or revise the prices and quantities that they are willing to trade.

Bringing together these potential buyers and sellers to engage in price discovery results in a large amount of market data: data on the willingness of traders to buy or sell before transactions take place, and data on transactions that result from the matching of buyers and sellers. This market data is disseminated to market participants through a variety of mechanisms, and it provides information about prices, trading activity, and liquidity in markets.

Equity market data is used in a wide variety of ways. Market participants include institutional money managers, arbitrageurs, hedgers, market makers, operators of other trading venues (such as dark pools), high-frequency traders, individual investors, and others. The market data available to all of these market participants, and the ways in which they respond to the data they receive, form the core of the price discovery process.

Market data obviously informs decisions about whether and what to trade. After a decision to trade has been made, market data enables traders and their brokers to evaluate key

dimensions of current market conditions that inform order submission strategy—the choice of how much and how quickly to trade, whether to place an offer or to hit an existing bid, whether to route an order to one trading venue versus another, and so on. The existence of real-time quote data gives market participants information about the likely prices and quantities available in the market before they make their trading decisions.³

Market data is not just used by traders: after orders have been routed, market data is used by exchanges and other trading venues (such as dark pools) to ensure the executions in those “unlit” venues occur at or within the current National Best Bid and Offer (“NBBO”), which is generally required by SEC rules. Other trading venues use current market quotes from the exchanges as a benchmark for determining execution prices.⁴

Market prices are also used by investors and investment managers to monitor the value of individual positions and portfolios, and by brokers as they monitor customer positions and enforce margin requirements. Real-time market data feeds are used to provide intraday updating of market indices and intraday indicative values for exchange-traded funds, a key component of the index arbitrage process that helps keep index futures and equity prices in line. Real-time

³ In this context, the kinds of market data likely to be useful may vary depending on the nature of the trader and the business it conducts. For most orders from retail customers, the “top-of-book” data available from the consolidated feeds is likely sufficient to provide all information such a trader needs to make a trading decision. For institutional investors using computerized trading algorithms, additional information available in “depth-of-book” feeds may be helpful in some circumstances. For high-frequency traders and others following highly time-sensitive strategies, having a low latency data feed will be important. For others, latency may be less important.

⁴ Midpoint crossing networks, for example, typically allow buyers and sellers to match and transact at the NBBO midpoint, the price at the time of matching that is midway between the national best bid price across all registered exchanges and the analogous national best offer price. Wholesalers make similar use of the NBBO, because they often promise broker-dealers that they will provide price improvement compared to the NBBO on retail order flow that is routed to the wholesalers for execution.

equity market data feeds into option markets, as options market makers generally provide bid and offer prices on options based on the current share price level.

Historical databases of intraday trading and quoting activity are also used by a number of market participants. Historical market data is used to compute execution quality metrics such as effective spreads, price improvement, and speed of execution—metrics that may be used to evaluate market quality at different trading centers or at different times.⁵ Historical data can be used by traders to back-test trading strategies before putting them into operation and by brokers to help optimize their order routing strategies and to evaluate their compliance with best execution obligations. Historical data has been used extensively by the academic community to address a wide range of research topics, and by the SEC, the Financial Industry Regulatory Authority (“FINRA”), and the exchanges to evaluate the impact of rules and changes in market structure. Historical data is also used in the context of regulatory investigations, enforcement actions, and FINRA arbitrations.

Given all of these uses for equity market data, and given the wide range of people and entities with an interest in equity market data, it is unsurprising that a regulatory framework has developed around market data. In addition, regulators focus on market data because researchers have generally found that the availability of information about current bids and offers (called “pre-trade transparency”), and timely reporting of equity market trades (called “post-trade

⁵ For example, trading venues are required to disclose certain market quality metrics under Rule 605.

transparency”), are both important contributors to market quality.⁶ Thus, understanding the current regulatory framework is a key part of understanding the overall market for market data.

III. Equity Market Data Products in the United States

Thousands of publicly traded companies are listed on U.S. equity exchanges that are part of the NMS for trading. There are two main categories of market data products for NMS stocks. Consolidated feeds combine trade and quote data from each trading venue, while each individual exchange offers proprietary market data products that provide additional information about activity at that particular trading venue.⁷

Consolidated feeds provide real-time reporting of all trades in NMS stocks. Consolidated feeds also provide time-stamped “top-of-book” quotes for all NMS stocks, consisting of each exchange’s best (highest) bid price and quantity and its best (lowest) offer price and quantity. This allows market participants to know the NBBO available in the market at any point in time. The consolidated feed is managed by a Securities Information Processor (“SIP”), so consolidated data is sometimes referred to as “SIP data.”

Exchanges have also developed various market data products that they sell directly to subscribers. These generally differ from the SIP feeds. Data products sold by the exchanges include data feeds containing trades and quotes, orders at prices other than the best bid and offer

⁶ Papers that measure the market quality effects of pre-trade transparency in the equity markets include Hendershott and Jones (2005) and Boehmer, Saar, and Yu (2005). The salutary market quality effects of increased post-trade transparency in the U.S. corporate bond markets are documented by Bessembinder, Maxwell, and Venkataraman (2006), Edwards, Harris, and Piwowar (2007), and Goldstein, Hotchkiss, and Sirri (2007).

⁷ Consolidated feeds are administered by the UTP and CTA Plans, which are described in more detail in the Appendix.

(which is typically referred to as “depth-of-book” information), and messages related to price discovery around the opening and closing auctions.

Different market data products offered by the exchanges are designed for different types of market participants with different needs. Some market participants find that the consolidated feeds serve their needs; these participants have little or no need to purchase data directly from exchanges. Institutional brokers and proprietary trading desks may subscribe to some or all exchanges’ depth-of-book data feeds as inputs to their order routing algorithms or to help them work large orders. For example, an executing broker might break up a large order into smaller pieces submitted to multiple venues. Depth-of-book feeds could help that broker decide which venues should get the orders and the prices at which it should submit each order. These feeds would also help the broker readjust the pricing or venue for those orders based on evolving market conditions.⁸

IV. The Economics of the Provision of Equity Market Data by Exchanges

Market data is valuable to subscribers, and this is the basis for a market in equity market data. By developing systems and processes that bring together buyers and sellers, exchanges and other trading venues help create and produce market prices. These market prices, and the resulting market data, are only valuable because the exchanges provide the aggregating and matching services that create them.⁹ To see this concretely, consider for a moment the fact that

⁸ Please see the Appendix for a more detailed discussion of consolidated feeds, proprietary market data products, and the development over time of the underlying regulatory framework.

⁹ See, for example, J. Harold Mulherin, Jeffrey M. Netter, and James A. Overdahl, “Prices Are Property: The Organization of Financial Exchanges from a Transaction Cost Perspective,” *Journal of Law and Economics* 34 (1991), pp. 591–644.

individual equity market participants are welcome to sell data concerning their own orders and transactions. This does not occur in practice, mainly because there is little value in such disaggregated data. Market data products have value precisely because they aggregate the orders of many market participants and report more than just a small subset of the transactions that result from matching buyers and sellers. To put it another way, market participants are not buying back their own data when they buy market data. What has value, and what they pay for, is to see the entire market: the actionable orders and transactions involving *other* market participants that have been accepted by an exchange.

Market data is a product of an exchange, but it has also been an important driver of exchange innovation. For example, in the early 2000s, the Island Electronic Communications Network (“ECN”) operated a very fast matching engine and distributed a state-of-the-art order-level data feed to market participants. The simplicity, completeness, and speed of the so-called ITCH data feed helped the Island ECN to build market share as it competed with Nasdaq and other established trading venues. When Nasdaq acquired Island’s successor Inet in 2005, Nasdaq adopted much of Inet’s technology, including the ITCH data feed, in part because market participants valued Inet’s matching technology and the associated market data. In fact, the NYSE later developed a similar order-level data feed, in part as a competitive response to the ITCH data feed and to similar feeds being offered by most other exchanges. More generally, the ability to sell market data, as well as the competition among trading venues that has been explicitly encouraged by the SEC, provides incentives for exchanges to innovate in ways that market data consumers value.

Speaking of innovation, it is also important to note that the provision of market data by exchanges is a natural outgrowth of the automation of equity trading. Automated market data

feeds have substituted for manual information flow via humans. Twenty years ago, a large broker-dealer would need dozens of employees scattered around the floor of the NYSE, and those employees would still provide a fraction of the information that is currently provided in a single NYSE data feed.¹⁰ Even the most expensive exchange data feed is cheaper than the average salary and bonus paid to a New York City employee in the securities industry.¹¹

Like most producers, stock exchanges offer a variety of market data products at different price levels. The simplest, most basic products are offered at the lowest prices. For example, consolidated data that is more than 15 minutes old can be easily found on financial websites, because consolidated feed subscribers face no restrictions on the redistribution of these older prices. Financial websites also provide a considerable amount of real-time data at no charge to their users. For example, Google and Yahoo Finance provide real-time last-sale information on all U.S. equities. This real-time information may be sufficient for many investors to make trading decisions.

Comprehensive real-time data comes from the consolidated feed at a cost. The consolidated feed contains a great deal of data that characterizes the essential elements of the national market: the most recent transaction prices from all trading venues, and the best bid and offered prices and quantities at each exchange.

More advanced market data products are offered at higher prices, reflecting their greater value to market participants with specific needs based on how they choose to trade. Exchange order-level data feeds are particularly valuable to active proprietary traders and to users of

¹⁰ See, for example, Ian Domowitz and Benn Steil, “Automation, Trading Costs, and the Structure of the Securities Trading Industry” (working paper, 1997).

¹¹ The Office of the New York State Comptroller reported that the 2016 average salary and bonus for an employee in New York City working in the securities industry was \$375,300. <https://www.osc.state.ny.us/osdc/rpt6-2018.pdf>.

algorithms designed to trade large amounts of stock over periods of time. But exchanges must price proprietary products with care, because overpricing can cause them to lose order flow, and the value of proprietary products is constrained by the existence of the consolidated feeds. For many market participants, exchanges that sell proprietary market data products must compete with the SIPs, because SIP data includes a large subset of each exchange's proprietary data and aggregates together all of the exchanges and other trading venues, thereby reducing the value of any single exchange's proprietary data.

Also, like most producers, stock exchanges face substantial competition from existing rivals and potential new entrants. Currently, there are 13 cash equity exchanges and over 30 ATSs in the United States, with many new entrants in the exchange space over the past 20 years.¹² For example, Cboe, which is now one of the larger U.S. equity exchange operators, manages four exchanges that were previously operated by Bats and Direct Edge. Bats was founded in 2005, and the Direct Edge ECN began in 2007. In addition, a new exchange IEX was just approved in 2016. This competition, and the potential for new entrants, ensures that prices for market data are set in a competitive market.

In fact, one of the important ways that new entrants can compete is by offering free market data. For example, IEX offers real-time depth-of-book and last sale information to subscribers at no cost. Similarly, Bats offered free depth-of-book data for the first several years of its existence, and Arca also offered its depth-of-book data for free for a significant period of time. Exchanges that do sell market data must also consider the effects on their market share of trading. Market participants will decline to purchase market data that is overpriced, and market

¹² FINRA's April 23, 2018 weekly report of ATS trading volume identifies 32 active ATSs.

participants who stop buying an exchange's market data may also decide to route their order flow to other venues. Overpriced market data is not in either side's interest. More generally, institutional investors, broker-dealers, and other professional users of market data are well-situated to advance their own interests in their interactions with exchanges (for example by choosing to divert order flow from exchanges with proprietary data they deem too expensive).

Seen this way, the market for market data is quite similar to the segmented markets for many other products. As an analogy, consider the market for new automobiles. A basic new car (such as the Honda Fit or the Ford Fiesta) can be purchased in the United States for less than \$20,000. Such a vehicle is likely to be fairly small, with a modest number of features, and provides reliable transportation for a small number of passengers. However, such a car provides a great deal of functionality for a relatively low price, and many buyers find that it meets their needs and opt for this choice. For those with an even lower willingness to pay, there are also used cars available at considerably lower prices.

At the higher end of the market, there are automobiles that sell for over \$60,000, such as the BMW 7 series or a Cadillac CTS sedan. These are typically more powerful vehicles with many more features, and these vehicles appeal to buyers with a different set of requirements. Virtually every potential buyer would prefer the higher-end vehicle, but given the price differential, only some buyers—those with the desire for the top-of-the-line performance or features and a willingness to pay the higher price—ultimately choose this particular option.

Still other would-be buyers decide they do not need to purchase a car at all; instead, they might walk or take public transit to satisfy their transportation needs. Analogous investors would make use of free or very low-cost market data alternatives for their investing information needs.

Moreover, there are several different vehicle manufacturers—Ford, GM, Chrysler, Honda, Toyota, BMW, and so on—competing on price, performance, features, styling, and many other dimensions. Most individual buyers purchase just one vehicle at a time. Others, including rental car fleets, acquire many cars from many different manufacturers simultaneously.

Finally, as is the case in equity markets, there are a few entities with business models that require them to purchase a car from each manufacturer. For example, reviewers such as Consumer Reports must purchase vehicles from each well-known manufacturer in order to provide comprehensive reviews, comparisons, and recommendations. It is more expensive for Consumer Reports to purchase a complete range of luxury sedans to review that car class.¹³ However, Consumer Reports does not petition the government to lower luxury sedan prices; in fact, it would probably be embarrassed to do so. It realizes that in a competitive market such as this one, the government does not set car prices but allows the market to operate freely. At the same time, car manufacturers dare not set the price of a luxury sedan too high, because they risk causing buyers to consider an alternative make instead, and even Consumer Reports might decide there is no reason to test the overpriced vehicle that few readers are likely to seriously consider.

V. Market Data and Regulation

For both SIP and proprietary data products, every change in a pricing schedule must be filed publicly with the SEC, and the SEC has the authority to take action to disapprove those fees. Moreover, neither the SIPs nor exchanges can charge fees other than those contained in

¹³ See <https://www.consumerreports.org/cars-how-consumer-reports-tests-cars/>.

their SEC filings. For example, this means that exchanges cannot negotiate different deals with different market data subscribers. In contrast, the prices charged by third-party vendors for market data–related services are unregulated and are not published.

The SEC discussed issues related to equity market data in a 1999 concept release. The SEC stated that Congress “intended to rely on competitive forces to the greatest extent possible to shape the national market system,” but also suggested that Congress believed market forces might not be sufficient to spur the development of a consolidated feed (as opposed to proprietary feeds), and empowered the SEC to ensure “that the essential mechanisms of an integrated secondary trading system are put in place as rapidly as possible.”¹⁴ At least historically, more attention has been paid to SIP data pricing than to proprietary data pricing, likely because brokers need access to the consolidated feed in order to meet certain regulatory obligations. In particular, under Rule 603(c) of Reg NMS, which is sometimes referred to as the Vendor Display Rule, the SEC staff has made clear that broker-dealers must provide a consolidated display of market data when they are providing equity quotation information to customers.¹⁵ Brokers can choose to satisfy the Vendor Display Rule by paying a minuscule \$0.0075 per query to provide a snapshot of the consolidated feed.

A particularly weak argument is that the consolidated feed should be priced based solely on the costs of the SIPs. First, costs should include all of the operating costs the National Market

¹⁴ See SEC Release No. 34-42208.

¹⁵ FINRA Regulatory Notice 15-52, December 2015.

System incurs in facilitating equity transactions, not just the costs specific to the SIPs. In addition, from an economic perspective, a regulated product should be priced to maximize social welfare, which means that prices should reflect the product's overall value, not just all of the costs associated with its production.¹⁶ In the case of market data, the consolidated feed has considerable value to its subscribers along myriad dimensions. Some of this value reflects the public good aspects of the consolidated feed. For example, as described above, some subscribers are midpoint crossing networks that use the resulting prices as the basis for matching buyers and sellers on their networks. Other than relatively small SIP fees (less any rebates from Trade Reporting Facility ("TRF") prints), these networks bear none of the costs of the consolidated price discovery process on which their business models depend. In regulating the pricing of the consolidated feeds, the SEC appropriately and holistically should consider the overall value of the market data being created.

From a regulatory standpoint, proprietary data feeds are fundamentally different from consolidated data feeds. First, there is no regulatory mandate that exchanges sell proprietary data at all. In fact, for many years prior to 2001, depth-of-book data for NYSE-listed stocks was not generally available outside of the specialist's post. Second, depth-of-book data is not necessary or helpful for many types of market participants. For example, according to a 2014 article, only 3.3% of all trades take place outside the NBBO, where depth-of-book information would be

¹⁶ See, for example, Paul L. Joskow and Nancy L. Rose, "The effects of Economic Regulation," in *Handbook of Industrial Organization*, Vol. 2, edited by R. Schmalensee and R. Willig (Elsevier, 1989).

particularly useful.¹⁷ This explains why some market participants do not subscribe to proprietary data feeds at all, and among those who do subscribe, a significant portion subscribe to feeds from some but not all of the exchanges.¹⁸ Although some have argued that depth-of-book data is necessary for a broker to comply with its best execution obligation, the SEC has stated that this is not the case.¹⁹

VI. Market Data Pricing and Revenues

A. Pricing Structure

Exchange market data fees, including fees for consolidated data distributed by the NMS Plans²⁰ and fees for exchange proprietary data, are subject to oversight by the SEC, and all fee changes are submitted as rule changes to the SEC and are published on the SEC's website. The level of transparency regarding exchange equity market data prices is thus extremely high.

¹⁷ Craig W. Holden and Stacey Jacobsen, "Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions," *Journal of Finance* 69, no. 4 (2014), p. 1759.

¹⁸ Initial Decision Release No. 1015, SEC Administrative Proceeding File No. 3-15350, June 1, 2016.

¹⁹ SEC Release 34-59039, pp. 41–42, 75–76. FINRA indicated to its members in November 2015 that "a firm that regularly accesses proprietary data feeds ... for its proprietary trading, would be expected to also be using these data feeds to determine the best market under prevailing market conditions when handling customer orders to meet its best execution obligations." See FINRA Regulatory Notice 15-46, p. 13. However, that FINRA notice does not suggest that firms that do not already subscribe to proprietary feeds for their own internal use would need to start doing so as a result of the notice.

²⁰ Please see the Appendix for a more detailed discussion on the NMS Plans.

1. Consolidated Data Fees

Both the CTA and UTP Plans administer their own fee schedules, and in general the fee schedules do not change frequently.²¹ There are two types of fees: access fees and use fees.

- *Access Fees:* Direct access fees apply for direct connections to the SIP, whereas indirect access fees are charged when data is supplied via a third-party vendor.
- *Use Fees:* Use fees are divided between “display” fees (e.g., “eyeball” usage by a market participant) and “non-display” data (e.g., automated use of the data, such as using the data as an input to an order routing or algorithmic trading system).
 - Display fees are charged per subscriber, with separate rates for professionals and non-professionals.²² Alternatively, users can elect to pay a per-query fee, at a rate of \$0.0075, subject to caps based on the number of queries for non-professionals.
 - Non-display fees are charged based on how the data is used. There are three categories of non-display uses: using data to match buy and sell orders (such as in an electronic trading system or dark pool), using data on behalf of a subscriber’s customers, or using data for a subscriber’s own purposes (such as its own proprietary trading). Each type of usage is charged for separately.

²¹ For example, the CTA Plan from 1987 to 2013 had a fee structure based on 14 pricing tiers. In 2013, CTA updated and simplified the structure to four tiers. See SEC Release No. 34-70010.

²² Users are assumed to be professional unless they meet specific criteria, namely, they are individuals who are not securities professionals and are using the data for personal reasons. Both CTA and UTP Plans charge a monthly rate of \$1 for non-professionals, but for the most part non-professionals do not even pay this modest amount directly, because their brokers usually bear that cost. For professional users of display devices, Tapes B and C charge a flat rate per professional user, while Tape A uses a four-tier system with reduced rates based on the number of professional users.

The fee structures also include various other fees, such as redistribution fees (which are charged to firms that retransmit the data externally) and television ticker display fees (which are tiered based on the number of households that have access). The main types of fees are summarized in Table 1.

2. Exchange Proprietary Data Fees

The exchanges structure their proprietary market data fees in a similar way.²³ For each of the various products offered by the exchanges, the exchanges charge access fees, usage fees, and redistribution fees. Exchanges also apply the same designations for professional and non-professional users, and display and non-display distinctions also apply.

Once they have been put in place, prices for exchange proprietary data products have generally remained stable over time.

For example, NYSE's OpenBook is a proprietary data product that provides frequent snapshots of the entire NYSE order book. It was initially offered in 2002 at a fixed access fee of \$5,000 per month plus a variable fee based on the number of subscribers. The access fee has not changed since inception of the product, and the subscriber fee changed only once, in 2004, from \$50 to \$60 for professional users; the non-professional subscriber fee (\$15/month) has not

²³ Although exchanges' fee schedules are structurally similar, there are nuanced differences between the exchanges. For example, Cboe Global Markets and Nasdaq differentiate between internal and external distribution, whereas NYSE just charges a redistribution fee on top of an access fee. Nasdaq in some cases charges different fees for Nasdaq-, NYSE-, and Amex-listed issues. See, for example, <https://nasdaqtrader.com/Trader.aspx?id=DPUSdata>, <https://markets.cboe.com/us/equities/membership/pricing/> and https://www.nyse.com/publicdocs/nyse/data/NYSE_Market_Data_Pricing.pdf.

Summary of Selected Consolidated Market Data Fees by Tape

	Tape A	Tape B	Tape C
Access Fees			
Direct Access	\$3,000	\$2,000	\$2,500
Indirect Access	\$2,000	\$1,000	\$500
Usage Fees			
Display Only			
Professional	\$19 – \$45/Subscriber	\$23/Subscriber	\$24/Subscriber
Non-Professional*	\$1/Subscriber	\$1/Subscriber	\$1/Subscriber
Per Query*	\$0.0075 per Query	\$0.0075 per Query	\$0.0075 per Query
Non-Display			
For ETS or ATS	\$4,000	\$2,000	\$3,500
Customer Use	\$4,000	\$2,000	\$3,500
Firm Use	\$4,000	\$2,000	\$3,500
Redistribution Fees			
Real Time	\$1,000	\$1,000	\$1,000

Source: CTA Network A Fee Schedule, January 2015; CTA Network B Fee Schedule, January 2015; UTP Plan Network C Fee Schedule, February 2018

Note: All fees are monthly, and are fixed unless indicated otherwise. This table does not represent the complete list of fees charged by the Networks, but does represent the main fee categories. In some cases different fee types are combined for simplicity in comparison, including separate fees for quotes and trades.

*Non-professional and per query fees are typically paid by an end-user's broker, and not by the end-user.

changed at all.²⁴ In 2013, the NYSE started charging a flat fee for all of a subscriber's internal non-display devices instead of requiring subscribers to report the number of non-display devices used.²⁵ Since its inception in 2002, the OpenBook product has been enhanced significantly in terms of speed and volume of data.

ArcaBook is a similar proprietary data product that provides information on the entire NYSE Arca order book. ArcaBook was free for many years (up until 2009), and it now has a fee schedule that is similar to OpenBook's, but with lower fee levels. The ArcaBook access fee is currently \$2,000/month, the professional user display fee is \$60/month, and the non-professional user display fee is \$10/month. There are also redistribution and non-display fees for ArcaBook. Since it became available, the ArcaBook product has been enhanced significantly in terms of speed and volume of data. Nasdaq and Cboe proprietary data product pricing follows a similar pattern.

Exchange market data fee schedules are publicly available, so it is possible to estimate the total costs that would be incurred for proprietary data by various types of market participants. Consider the following hypothetical examples of data costs for different types of firms that subscribe to different packages of data for different uses:²⁶

²⁴ SEC Release No. 34-45138.

²⁵ SEC Release No. 34-69278. In 2009, recognizing that subscribers were incorporating data feeds into their own computer systems, the NYSE changed its unit of count to redefine a subscriber as a unique individual device that receives data, which also introduced the concept of non-display use and required users to report the number of non-display devices. This was introduced as a pilot rule change in 2009 and made permanent in 2010. SEC Release Nos. 34-62038 and 34-59198. In addition, subscribers were able to use managed non-display services as a lower priced option for non-display usage when non-display fees were introduced; managed non-display services were discontinued in 2016.

²⁶ The following examples are calculated based on current market data fee schedules. See https://www.nyse.com/publicdocs/nyse/data/NYSE_Market_Data_Pricing.pdf, <https://www.nasdaqtrader.com/Trader.aspx?id=DPUSdata#tv>, and http://cdn.batstrading.com/resources/membership/US_Market_Data_Product_Price_List.pdf.

- A broker-dealer with no automated use of the data might choose to display on user screens the Nasdaq products Nasdaq TotalView, BX TotalView, and PSX TotalView, for a total cost of \$156 per month per device. If the same broker-dealer took all three NYSE integrated feeds (NYSE, NYSE American, and NYSE Arca), that would add an additional \$140 per month per device. Taking all of Cboe's feeds would add an additional \$100 per device per month. In practice, such a firm might subscribe to data from only a small subset of exchanges, which could lower its cost per device by a considerable amount from the figures above.
- A purely proprietary trading firm with no external customers and fewer than 100 display devices might spend \$59,000 per month for NYSE data, \$59,950 per month for Nasdaq data, and \$32,500 per month for Cboe data.²⁷
- Finally, a global investment bank with a wide range of trading activities might choose to subscribe to all of NYSE Group's proprietary integrated data feeds and the similar feeds for Nasdaq and Cboe. If such a firm were to use these feeds to display limit order books, provide trading algorithms to its institutional investor clients, and support an affiliated dark pool, its total fees would be on the order of \$100,800 per month for NYSE data, \$127,720 for Nasdaq data, and \$37,000 for

²⁷ Assumes firm takes all three NYSE Group integrated feeds, all three Nasdaq TotalView products, and all four Cboe Depth products for 75 display-only devices and for non-display use in one non-display category. For BX and PSX, assumes 250 non-display subscribers at \$55 and \$50 per subscriber.

Cboe data.²⁸ This is an insignificant cost for these types of investment banks, which measure their annual equity trading revenues in billions of dollars.²⁹

Note that I do not have any data on market data charges incurred by individual firms, so these examples are all hypothetical based on exchange fee schedules and assumptions about how market participants choose to use equity market data products. The examples are intended to show the broad range of possible choices and how the costs of market data can be affected by those choices. Market participants choose what business models and trading strategies they pursue and what types of and how much market data to purchase, and those business decisions ultimately determine each market participant's equity market data costs.

B. Consolidated Data Revenues and Allocations

The fees collected by the CTA and UTP Plans for sales of consolidated data, after certain expenses, are distributed back to the Plans' participant exchanges and FINRA. The plan participants can then pass these revenues on to other market participants. For example, some exchanges historically have shared market revenues with specialist firms or other exchange members who routed order flow to the exchanges. FINRA also has a program for rebating market data revenue back to those FINRA members who reported the off-exchange trades. Thus, the ultimate allocation of market data revenue is broader than just the plan participants, and

²⁸ Assumes firm takes all three NYSE Group integrated feeds, all three Nasdaq TotalView products, and all four Cboe Depth products for 120 display-only devices and for non-display use in two non-display categories. For BX and PSX, assumes 250 non-display subscribers at \$55 and \$50 per subscriber.

²⁹ Later in the paper, I estimate total 2015 equity trading revenue of \$47.9 billion for the nine largest investment banks, or an average of \$5.32 billion in equity trading revenues per firm. The \$3.2 million annual data cost from this example is approximately 0.06% of this average revenue figure.

recipients of consolidated market data revenue include broker-dealers who operate dark pools or otherwise execute trades as off-exchange market makers.

Prior to 2007, CTA revenues were allocated in proportion to the number of trades reported by each exchange. The SEC established a new revenue allocation formula in 2005 when it adopted Reg NMS. The new formula, which went into effect on April 1, 2007, first allocates revenues across stocks in proportion to the square root of dollar volume, then within each stock allocates 25% of the revenue to plan participants in proportion to the participant's number of trades, 25% in proportion to the participant's share volume, and 50% in proportion to a measure of how often the exchange is offering liquidity in that stock at the NBBO.³⁰

Although the fee schedules described in the previous section have always been public, financial information about the CTA and UTP Plans, including the total amount of fees collected and revenue distributed to participants, has historically not been in the public record, with a few isolated exceptions.³¹

This changed in March 2018, when the CTA and UTP Plans disclosed historical information about the annual revenue distributed to participants going back to 2007, including a decomposition of these distributions for the trade and quote components of the allocation formula. However, it is important to note that this data does not disclose how individual participants share tape revenues with broker-dealers and others. Thus, this data set shows the maximum revenue per participant, not necessarily the amount each participant keeps for itself. In this section, I provide an analysis of this new data set.

³⁰ SEC Release No. 34-51808.

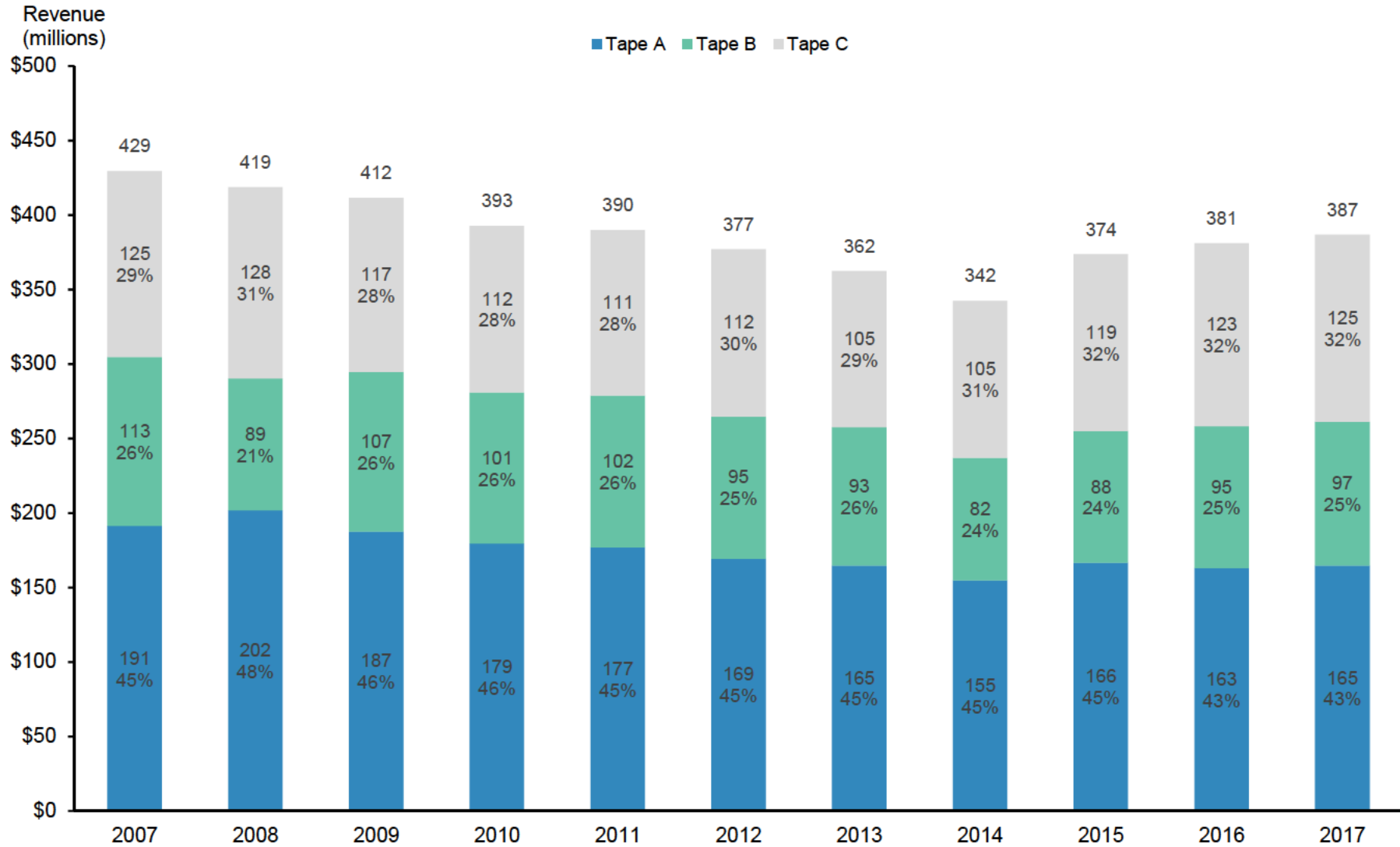
³¹ SEC Release Nos. 34-49325, 34-51808, and 34-61358.

Using the new data released by CTA and UTP, Figure 1 summarizes the aggregate amount of data revenues distributed to plan participants each year from 2007 to 2017, broken down by Tape A, Tape B, and Tape C. As the chart indicates, there has been some fluctuation over the years, but no growth in revenues over time. Total consolidated revenues distributed in 2017 were \$387 million, which is 10% lower than they were in 2007, even without adjusting for inflation. After adjusting for inflation using the CPI-U, consolidated revenues distributed declined by more than 23% over the 10 years ending in 2017. On average over the period from 2007 to 2017, distributed revenues were \$175 million per year for Tape A, \$97 million per year for Tape B, and \$117 million per year for Tape C. On a yearly basis, Tape A constituted between 43% and 48% of total revenues, Tape B constituted between 21% and 26% of total revenues, and Tape C constituted between 28% and 32% of total revenues.

Table 2 shows that consolidated revenues are a small and declining fraction of overall exchange revenues. For example, in 2008 equity SIP revenues were 4% of total NYSE Euronext revenues. By 2017 this percentage had declined to 2% of total parent company revenue. For Nasdaq, consolidated data revenues were 4% of total revenues in 2008, declining to 3% of total revenues in 2017.

Figures 2–4 summarize how the allocation of market data revenues across plan participants has evolved over time for Tapes A, B, and C. Figure 2 shows revenue allocations for Tape A (securities with primary listing on the NYSE). It shows a pattern over time consistent with the well-known increase in fragmentation of volume across trading venues after Reg NMS. Tape A revenues earned by NYSE exchanges have declined since 2007, while Tape A revenues have increased for the Nasdaq exchanges, the Bats/Direct Edge exchanges (acquired by Cboe Global Markets in 2017), and FINRA.

Annual Consolidated (SIP) Equity Market Data Revenue by Tape



Source: UTP Plan Revenue Disclosure Q42017: Trade & Quote Revenue Distributed to Participants; CTA Financial Disclosure on 3/1/18: Tape A Trade & Quote Revenue Distributed to Participants; CTA Financial Disclosure on 3/1/18: Tape B Trade & Quote Revenue Distributed to Participants

Market Data Contributions to Total Exchange Revenue Are Stable Over Time

(NYSE Euronext and ICE, in millions)

		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
NYSE Euronext											
Total Revenues	[A]	\$4,702	\$4,684	\$4,425	\$4,552	\$3,749	\$3,797				
Market Data Revenues	[B]	\$428	\$403	\$373	\$371	\$348	\$353				
US Equity SIP Revenues	[C]	\$168	\$145	\$142	\$131	\$112	\$104				
Other*	[B - C]	\$260	\$258	\$231	\$240	\$236	\$250				
<i>Percentage of Total Revenues</i>											
Market Data Revenues	[B / A]	9%	9%	8%	8%	9%	9%				
US Equity SIP Revenues	[C / A]	4%	3%	3%	3%	3%	3%				
Other*	[(B - C) / A]	6%	6%	5%	5%	6%	7%				
ICE											
Total Revenues	[A]							\$4,352	\$4,682	\$5,958	\$5,834
Market Data Revenues	[B]							\$446	\$470	\$535	\$556
US Equity SIP Revenues	[C]							\$96	\$108	\$108	\$104
Other*	[B - C]							\$350	\$362	\$427	\$452
<i>Percentage of Total Revenues</i>											
Market Data Revenues	[B / A]							10%	10%	9%	10%
US Equity SIP Revenues	[C / A]							2%	2%	2%	2%
Other*	[(B - C) / A]							8%	8%	7%	8%

Source: NYSE Euronext 10-K filing [2008–2012]; NYSE Euronext 10-Q filing [Q3 2013, Q1 through Q3 revenue is extrapolated in order to make the values comparable to the other revenues in the table]; Intercontinental Exchange 10-K filing [2014–2017; CTA Financial Disclosure on 3/1/18: Tape A Trade & Quote Revenue Distributed to Participants, Tape B Trade & Quote Revenue Distributed to Participants; UTP Plan Revenue Disclosure Q42017: Trade & Quote Revenue Distributed to Participants

*The “Other” category includes all revenues associated with market data excluding US Equity SIP data. This includes all proprietary market data for all geographic areas, and includes data from options, futures, indices, and others.

Note: Market Data Revenues represent revenues associated with all asset classes across all geographies. NYSE Euronext Total Revenues and Market Data Revenues from 2008 to 2012 include Euronext revenues after the merger with NYSE on April 4, 2007. For 2013, revenues are calculated by extrapolating Q1 through Q3 data from NYSE Euronext to annual estimates, due to Intercontinental Exchange acquiring NYSE Euronext in November 2013. US Equity SIP Revenues are compiled using recently reported data from CTA and UTP Plans. NYSE Euronext and ICE include tape revenues from New York Stock Exchange, NYSE Amex (starting in 2008), and NYSE Arca.

Market Data Contributions to Total Exchange Revenue Are Stable Over Time

(Nasdaq and Bats, in millions)

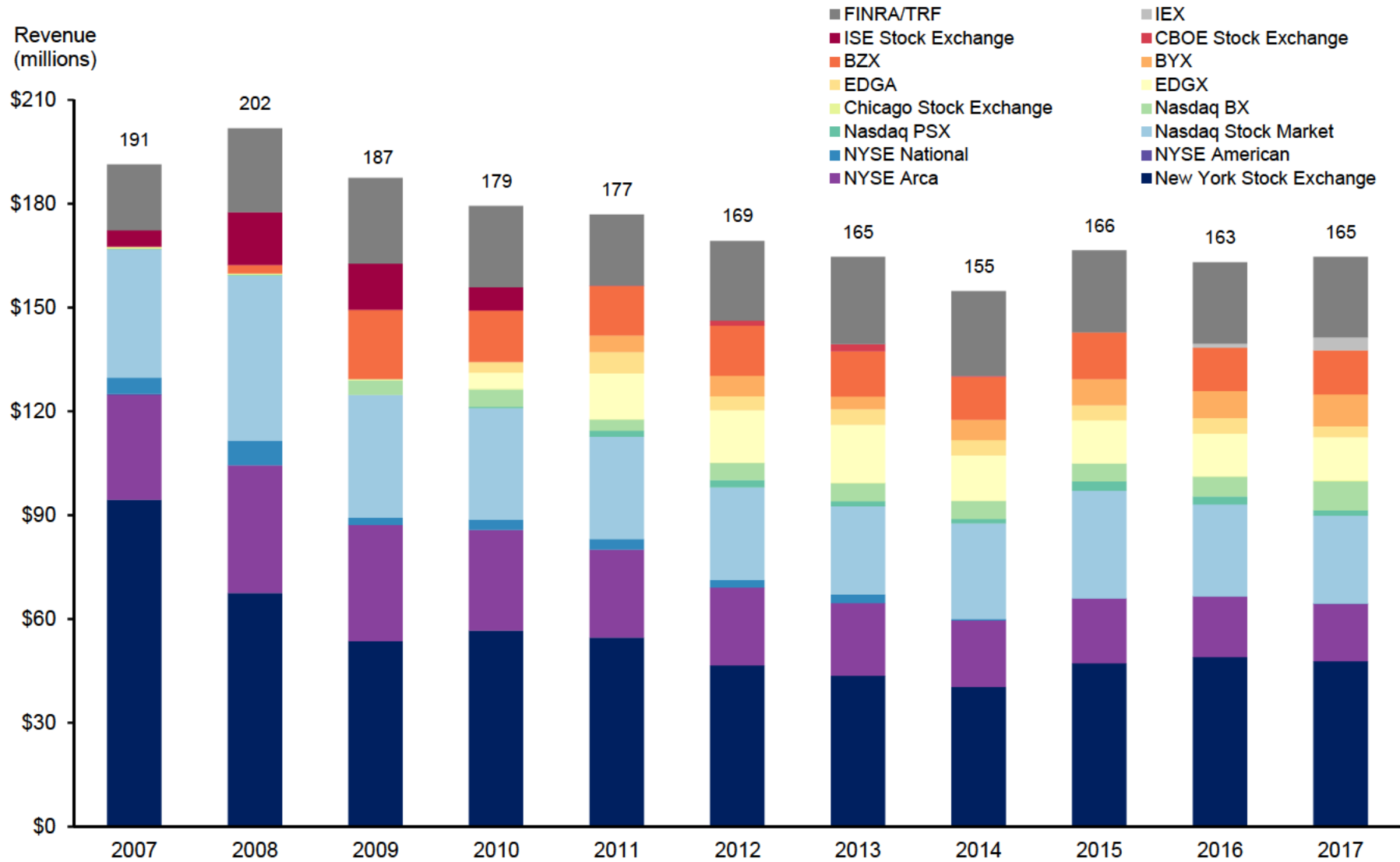
		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Nasdaq											
Total Revenues	[A]	\$3,650	\$3,410	\$3,191	\$3,438	\$3,120	\$3,211	\$3,500	\$3,403	\$3,705	\$3,965
Market Data Revenues	[B]	\$330	\$325	\$313	\$333	\$337	\$362	\$384	\$399	\$427	\$454
US Equity SIP Revenues	[C]	\$135	\$114	\$105	\$100	\$100	\$92	\$93	\$102	\$102	\$107
Other*	[B - C]	\$195	\$211	\$208	\$233	\$237	\$270	\$291	\$297	\$325	\$347
<i>Percentage of Total Revenues</i>											
Market Data Revenues	[B / A]	9%	10%	10%	10%	11%	11%	11%	12%	12%	11%
US Equity SIP Revenues	[C / A]	4%	3%	3%	3%	3%	3%	3%	3%	3%	3%
Other*	[(B - C) / A]	5%	6%	7%	7%	8%	8%	8%	9%	9%	9%
Bats											
Total Revenues	[A]								\$1,779	\$1,869	\$2,229
Market Data Revenues	[B]								\$131	\$146	\$165
US Equity SIP Revenues	[C]								\$100	\$103	\$100
Other*	[B - C]								\$31	\$43	\$65
<i>Percentage of Total Revenues</i>											
Market Data Revenues	[B / A]								7%	8%	7%
US Equity SIP Revenues	[C / A]								6%	5%	4%
Other*	[(B - C) / A]								2%	2%	3%

Source: Nasdaq 10-K filing [2007–2017]; BATS Global Markets 10-Q Filing [Q3 2016]; BATS Global Markets Press Release [Q4 2016]; CBOE 10-K filing [2017]; CTA Financial Disclosure on 3/1/18: Tape A Trade & Quote Revenue Distributed to Participants, Tape B Trade & Quote Revenue Distributed to Participants; UTP Plan Revenue Disclosure Q42017: Trade & Quote Revenue Distributed to Participants

*The “Other” category includes all revenues associated with market data excluding US Equity SIP data. This includes all proprietary market data for all geographic areas, and includes data from options, futures, indices, and others.

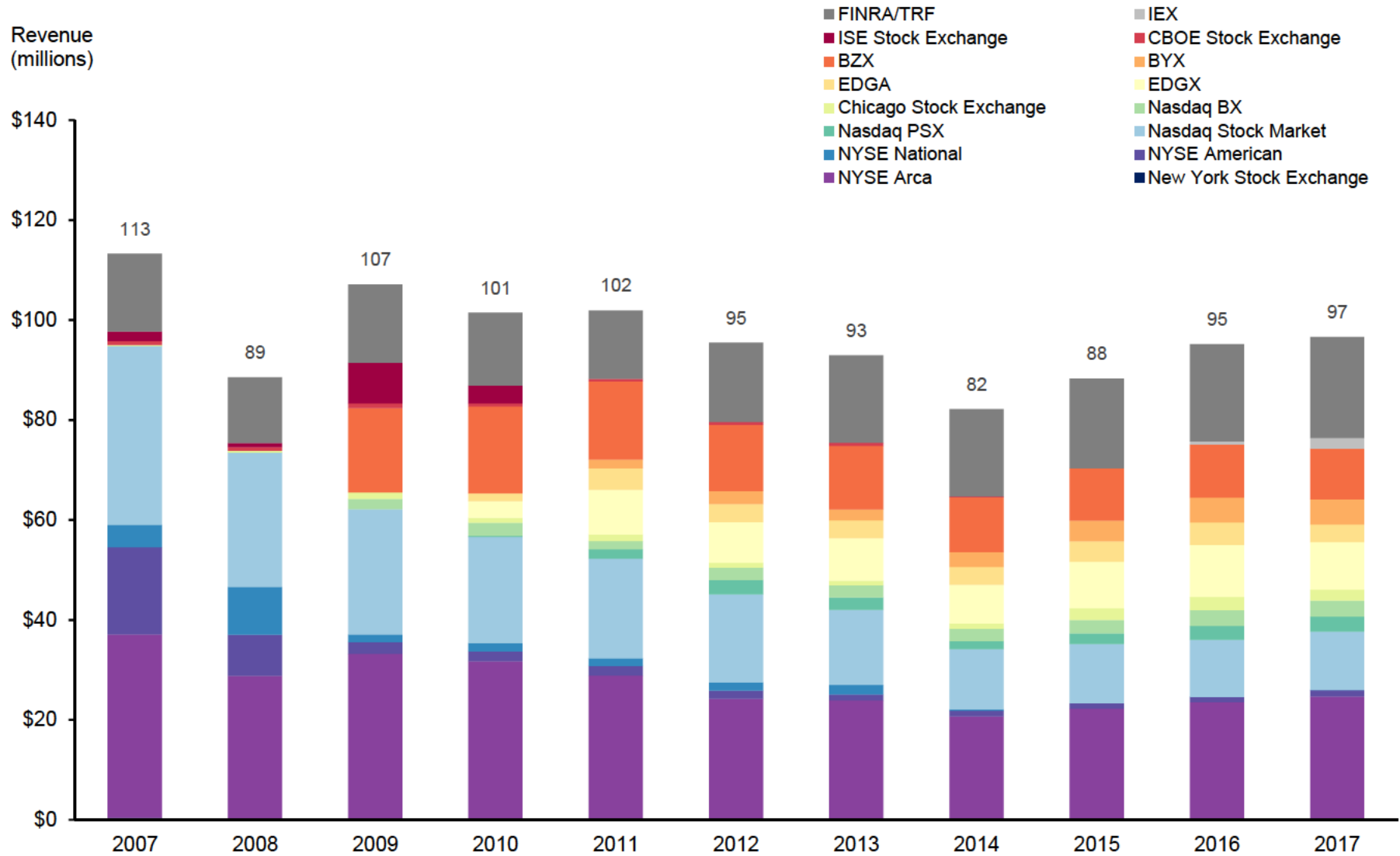
Note: Market Data Revenues represent revenues associated with all asset classes across all geographies. Bats Total Revenues and Market Data Revenues for 2015 and 2016 are calculated by combining nine months of financial reporting ending September 30 from the Bats Global Markets, Inc. 10-Q filed November 8, 2016 and three months of financial reporting ending December 31 from the Bats Global Markets, Inc. February 9, 2017 Press Release. Bats Total Revenues for 2017 are populated from the Cboe Global Markets, Inc. 10-K filed February 22, 2018. US Equity SIP Revenues are compiled using recently reported data from CTA and UTP Plans. Nasdaq includes tape revenues from Nasdaq, Nasdaq BX (starting in 2009), and Nasdaq PSX (starting in 2008). Bats includes tape revenue from BZX, BYX, EDGA, and EDGX.

Annual Consolidated (SIP) Equity Market Data Revenue: Tape A



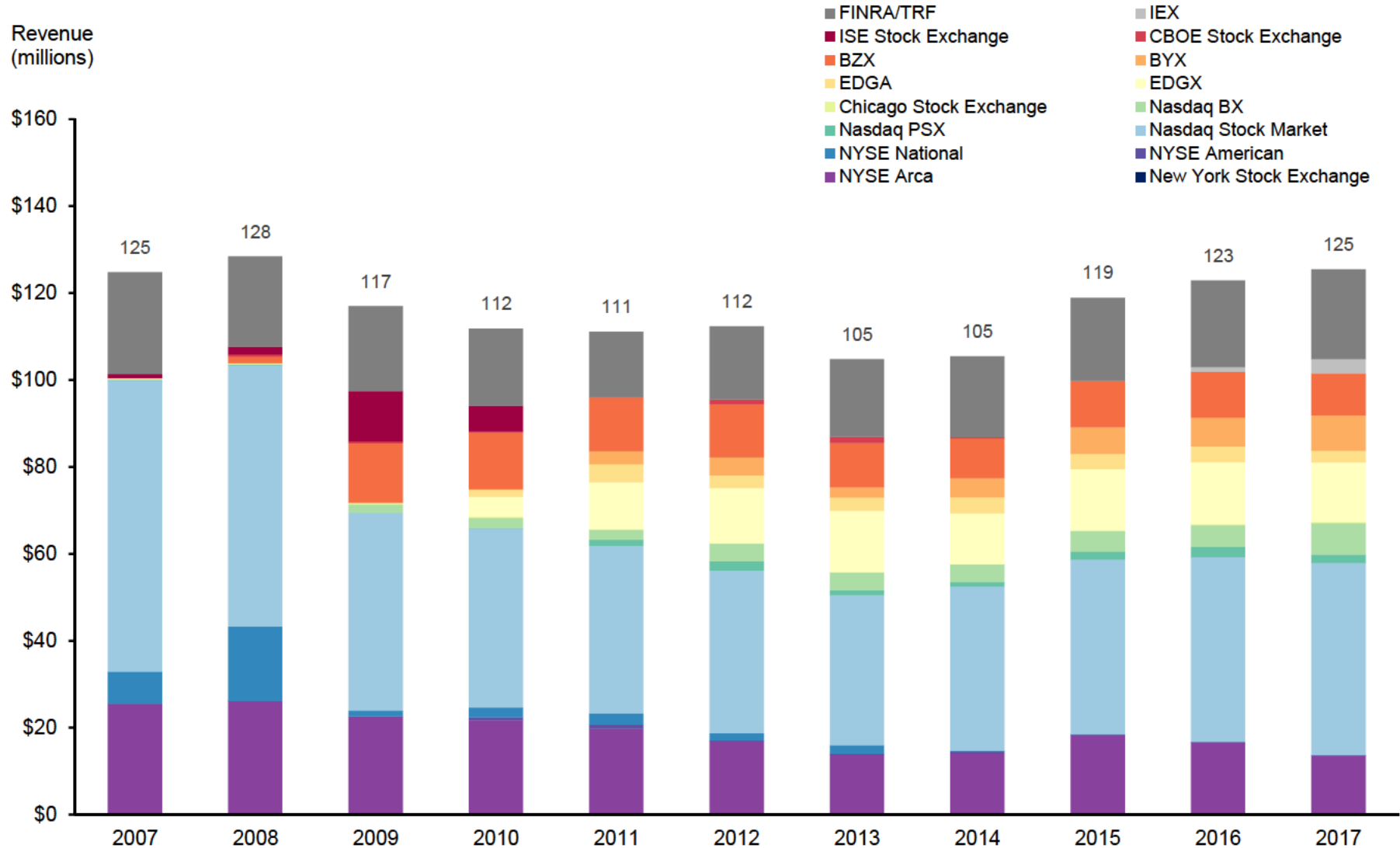
Source: CTA Financial Disclosure on 3/1/18: Tape A Trade & Quote Revenue Distributed to Participants

Annual Consolidated (SIP) Equity Market Data Revenue: Tape B



Source: CTA Financial Disclosure on 3/1/18: Tape B Trade & Quote Revenue Distributed to Participants

Annual Consolidated (SIP) Equity Market Data Revenue: Tape C



Source: UTP Plan Revenue Disclosure Q42017: Trade & Quote Revenue Distributed to Participants

Figure 3 shows the evolution of revenue allocation for Tape B (securities with primary listing on exchanges other than NYSE or Nasdaq). In 2007, a significant portion of Tape B revenues was earned by NYSE Arca and the American Stock Exchange. The combined Tape B revenue for current NYSE exchanges has decreased since then, as have Tape B revenues for the three current Nasdaq exchanges. Bats/Direct Edge exchanges and FINRA have gained market share over this interval and have seen an increase in Tape B revenues since 2007.

Figure 4 provides a similar chart for Tape C, which consists of securities with primary listing on Nasdaq. In 2007, Tape C revenue was mostly shared by Nasdaq, NYSE Arca, and FINRA. Since that time, the revenue earned by the Nasdaq exchanges has decreased, offset by increases in revenues by the Bats/Direct Edge exchanges and FINRA.

The data released by CTA and UTP also provides a breakdown between revenue distributed for the quote and trade components of the allocation formula. Based on the allocation formula that became effective in 2007, 50% of distributed revenues is allocated based on trading activity (number of trades and number of shares) and 50% based on quoting activity. Thus, across the entire industry, the amount of revenues distributed from the quote component equals the amount distributed from the trade component.

However, FINRA is not often used as a channel for displaying quotes, so its revenues are derived almost exclusively from trades.³² Consequently, FINRA's share of the overall market revenues, reflected in the charts above, does not reflect its market share of trade reports. For example, in 2017, FINRA captured 16.6% of all consolidated market data revenue and 33.2% of

³² When market participants use FINRA's Alternative Display Facility ("ADF") for quoting, FINRA does earn quote credit. In recent years, however, FINRA has not received any allocation of quote revenue. FINRA's share of quote revenue across all networks was approximately 2.14% in 2014, 0.14% in 2015, and zero in 2016 and 2017.

the trade revenue. Because the exchanges are competing with FINRA for trade revenue but not for quote revenues, the exchanges derive more than half their SIP revenues from the quote component. This is an important distinction, because it reflects the value of a key aspect of the price discovery process.

In summary, Figures 2–4 show that while total consolidated revenues distributed have stayed roughly constant since 2007, primary listing exchanges NYSE, Nasdaq, and AMEX/NYSE American have experienced reduced allocations, in large part due to new trading venue entrants. New ECNs would successfully capture market share in trading and then become registered exchanges through mergers or through exchange registration. This happened with Archipelago in the early 2000s, which gained access to tape revenue by affiliating with and then acquiring the Pacific Stock Exchange (although in this case the NYSE Group ultimately recaptured that portion of the market data allocation when it acquired Arca Ex in 2006). This happened again with the development of the Bats ECN, which became a registered exchange in 2008 and launched a second exchange in 2010, and two Direct Edge ECNs, which became registered exchanges in 2010.

Off-exchange trading also provides a significant source of competition for consolidated market data revenues. FINRA's competing TRFs—the FINRA/Nasdaq TRF and the FINRA/NYSE TRF—pass through the majority of their market data revenue from the CTA and UTP Plans to broker-dealer market centers that report trades to the TRFs. ATSS and broker-dealers trading as principal (including internalizers and wholesale purchasers of retail order flow) report their trades in this way. Thus, it is not just exchanges that receive revenue from consolidated feeds, but also dark pools, ATSS, and internalizers (who collectively receive tens of

millions of dollars annually in market data revenue rebates).³³ In fact, the total dollar amount of market data distributed to FINRA members who report off-exchange trades to a TRF has increased over time as the off-exchange share of trading has increased.

Market data rebates to broker-dealers reporting off-exchange trades serve two important disciplining roles. First, they effectively reduce the net amounts that off-exchange market centers pay for market data. Second, these rebates create an additional form of competition. In competing vigorously for order flow, exchanges can and do recognize that they must offer a trading product that is attractively priced relative to an alternative that may include market data rebates.

C. Exchange Market Data Revenues

Exchanges receive equity market data revenue from the sale of proprietary data and from the sale of SIP data. Although the exchanges do not provide itemized details of their exchange market data revenues, total market data revenues (which include market data revenues from securities exchanges and other sources as well) are reported in the financial disclosures of exchanges' parent companies. These disclosures indicate that total market data revenue is a small portion of overall reported revenue, and has remained roughly constant over time as a percentage of those total revenues.

Table 2 provides data on total market data revenues (across all asset classes and all geographies) of exchange groups as a percentage of total revenues over time for the three major ownership groups, as reflected in their financial disclosures. Market data revenue reported by

³³ In 2017, the TRFs collected over \$64 million in revenues from Tapes A, B, and C. According to FINRA Rule 7610B, 85% or more of revenues are shared with FINRA members whose market share is at least 0.1%. See http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=7355.

ICE (and its predecessor NYSE Euronext), which includes market data for equity and non-equity products, both inside and outside the United States, has remained between 8% and 10% of total revenue from 2008 to 2017. Likewise, Nasdaq's market data revenue has remained between 9% and 12% over the same time period. For Bats, market data revenue accounted for 7% to 8% of revenues from 2015 to 2017.

From these financial disclosures, it is also possible to place a strong upper bound on the revenues from the sale of equity securities exchange proprietary data. Table 2 also shows that proprietary data accounts for at most \$65 million of 2017 revenue at Bats, which is 3% of its overall revenue that year. For Bats, it is clear that proprietary data is a significantly smaller source of revenue compared to consolidated data. For NYSE and Nasdaq, equity securities exchange proprietary data revenues have been discussed in recent earnings calls. During the 3Q17 ICE earnings call, for example, ICE management stated that "the sales of NYSE real-time equity data products [i.e., proprietary market data products] are expected to be less than \$90 million in annual revenue to us and their growth has been relatively stagnant. These products account for approximately 2% of ICE's annual revenue." For the same quarter, Nasdaq provided a slide in its earnings presentation noting that U.S. equity proprietary depth products generated \$101 million in trailing 12-month revenue, compared to \$120 million for its share of consolidated data fees. Thus, it appears that for all three major U.S. stock exchange groups, proprietary equity market data actually provides less revenue to these firms than consolidated data.

D. Third-Party Vendors

Industry research reports, such as those by Burton-Taylor and Atradia,³⁴ provide detailed information about the costs of market data and related services to investors. For example, Burton-Taylor reports that in 2016, the total revenue earned by third-party vendors for market data–related services was over \$12 billion.³⁵ To put this number into perspective, this is over 10 times as much revenue as all the major exchanges combined earned for both proprietary and consolidated data during the same period. Exchange market data revenues across all asset classes and geographies (which is much more than market data revenue from just U.S. equity markets) total about \$1.1 billion in 2016—a small fraction of the over \$12 billion paid by market participants for real-time and trading data–related services during that time period.³⁶

E. Market Data Revenues as a Friction in Investment Performance

Retail and other equity investors might look at these market data revenues and conclude that they are significant costs that could contribute to higher brokerage commissions, greater mutual fund fees, and other drags on an investor’s overall investment performance. However,

³⁴ Burton-Taylor provides an annual report called “Financial Market Data/Analysis: Global Share & Segment Sizing.” Atradia published a research study in August 2010 called “The Cost of Access to Real Time Pre & Post Trade Order Book Data in Europe.”

³⁵ “Financial Market Data/Analysis: Global Share & Segment Sizing,” Burton-Taylor, 2017, p. 139. Note that this figure does not include exchange market data fees, and only includes fees paid to vendors themselves, denoted as “Real-Time & Trading Data.”

³⁶ Note that these three exchange groups operate a variety of financial markets, including options markets, futures markets, and others. Their financial statements do not separately break out U.S. equity market data fees, so the 2016 total U.S. *equity* market data fees are below \$1.108 billion, and probably substantially so. Similarly, the \$12.465 billion revenue number for third-party vendors applies to all financial markets, not just U.S. equity markets.

the data does not bear this out: the aggregate cost of equity market data is very tiny compared to the amounts invested in the stock market.

To see this, consider the \$1.1 billion of revenue reported by the three major exchange groups under the market data category (which includes all asset classes and geographies) in 2016 relative to the overall size of the U.S. equity market, which was \$30.15 trillion at the end of 2016.³⁷ Recall that this figure applies to all financial assets and jurisdictions where NYSE, Nasdaq, and Cboe operate, so it overstates U.S. equity market data revenue (likely by a substantial amount), whereas the size figure for the U.S. equity market is in fact limited to U.S. equities. Even so, this market data revenue figure represents less than 0.004% of the market capitalization of U.S. stocks, and the true ratio is probably substantially lower than 0.004% considering that the market data revenue figure (the numerator) includes data revenues from non-U.S. and non-equities markets. Equity market data has considerable value, as noted above, but even if its cost were considered as a simple drag on investment performance, the cost of equity market data would subtract far less than one basis point from overall investor performance each year.

This figure is also minimal compared to other standard sources of “drag” in investment performance: the overall amount charged in commissions, fees charged by investment managers, and so on. For example, I collected data on commissions charged by the retail brokerage sector. Together, the six firms in the Bloomberg Intelligence U.S. Retail Brokerage Competitive Peers Index reported \$10.0 billion in commission and related revenue in 2016. These firms alone take about 10 times as much in commissions from the subset of investors who use them as all market

³⁷ Center for Research in Security Prices (CRSP) database total value of listed equity securities as of December 30, 2016.

data generates from *all* market participants. By making this comparison, I do not mean to suggest that these brokerage firms have inappropriate commission levels. In fact, these broker-dealers also seem to engage in robust competition for customers, constantly improving their technological infrastructures and service delivery, while providing value in the form of equity transactions in return for small fees. My only point in drawing these comparisons is that market data costs are quite modest in comparison to other costs incurred by equity market participants.

Exchange market data costs are also small relative to overall broker-dealer equity trading revenues. For example, in the first nine months of 2015 the nine largest investment banks earned a total of \$35.9 billion from their equities trading operations.³⁸ This amounts to an annualized total of \$47.9 billion, assuming that the banks generated revenues at the same rate.³⁹ In contrast, in 2015 the total market data revenue earned by NYSE, Nasdaq, and Cboe (for all asset classes and geographies) was \$1.1 billion. Thus, total exchange market data revenues were less than 2.3% of equities trading revenues for just these nine investment banks. Since the numerator includes non-equity market data revenue, and the denominator includes only nine firms, this 2.3% percentage overstates (and probably substantially so) the fraction of equity trading revenues spent on equity market data by broker-dealers in aggregate. In short, exchange equity market data is a very small cost for the securities industry overall.

³⁸ Christina Rexrode, “The New Kid on the Stock-Trading Block: Citigroup,” *Wall Street Journal*, January 10, 2016.

³⁹ This value is calculated by dividing the reported values of equities trading revenue found in the *Wall Street Journal* article by 0.75 to estimate annual revenue. Although each of the banks does not report a separate value for equities trading revenue in their financial statements, the extrapolated number appears to be in the right ballpark based on relevant reported categories.

VII. Conclusion

The data on equity market data revenues is clear. Revenues from the consolidated feed are modest, totaling \$387 million in 2017. These revenues are lower than they were 10 years ago, while the consolidated feed has gotten considerably faster. Scaled by the over 1.5 trillion U.S. shares that changed hands in 2017, consolidated feed revenues amount to at most two hundredths of a cent per traded share. Exchanges are selling their own proprietary market data, but their overall market data revenues are relatively small, and they have remained approximately constant as a percentage of overall exchange revenues. Finally, market data revenues are small compared to some of the other costs that market participants face. Third-party vendors have overall real-time and trading data revenues that are over 10 times exchange market data revenues. Broker-dealer commission revenue is similarly much larger than exchange market data revenue. When aggregated together, annual exchange market data revenues are at most 0.4 basis points of the U.S. equity market capitalization, so they are truly a rounding error when it comes to calculating overall investment performance.

The economics of equity market data are also clear. Market data is clearly valuable to a wide variety of market participants for a wide variety of reasons, and basic economic principles dictate that the producers of that market data should be compensated for that value, which the existing regulatory system accomplishes. Although most broker-dealers are required to subscribe to it, consolidated market data also has public good aspects, and like other public goods, consolidated market data might be underpriced without regulatory oversight. The SEC is capable of taking into account all of these considerations.

For proprietary exchange data feeds, the main question is whether there is a competitive market for proprietary market data. More than 40 active exchanges and alternative trading

systems compete vigorously in both the market for order flow and in the market for market data. The two are closely linked: an exchange needs to consider the negative impact on its order flow if it raises the price of its market data. Furthermore, new entrants have been frequent over the past 10 years or so, and these venues often give market data away for free, serving as a check on pricing by more established exchanges. These are all the standard hallmarks of a competitive market.

Appendix — More Details on Market Data

There is a long history of stock exchanges and vendors selling market data. After stock ticker technology was introduced in 1867, ticker companies sold access to equity market data. For example, New York Quotation Co. and Gold and Stock Telegraph both disseminated quotation data from the NYSE. New York Quotation Co. became owned and controlled by the NYSE in 1890, and was given the exclusive right to provide equity market data to NYSE members.⁴⁰

The modern era of equity market data began with the overhaul of securities market regulation in the early 1970s. This was a time of intense legislative and regulatory action, including a focus on the fragmentation of trading across primary exchanges, regional exchanges, and third-market (off-exchange) trading. A series of studies, reports, and hearings involving the SEC, the exchanges, advisory committees, and congressional committees culminated in a new regulatory framework built around the core principles of the legislatively mandated National Market System.⁴¹

An important component of this new regulatory framework was the development of a system for channeling trade and quote data from each trading venue into consolidated feeds. This was accomplished by creating joint industry plans (“NMS Plans”),⁴² including the

⁴⁰ For a description of the mechanics of how the tickers worked in the early twentieth century, see Sereno Pratt “The Work of Wall Street,” (1912), pp. 182–184.

⁴¹ See Section 11(a) of the Securities Exchange Act of 1934. For a detailed summary of the regulatory activity at that time, see Robert Colby et al., “The National Market System: A Selective Outline of Significant Events,” 1985.

⁴² An NMS Plan is a consortium of self-regulatory organizations (including registered securities exchanges and FINRA) that come together as “participants” under the plan’s governing documents as a mechanism for coordinating compliance with a particular regulatory mandate. The plans themselves are advised by various committees of market participants and are governed by committees made up of the plan members.

Consolidated Tape Association Plan, the Consolidated Quotation Plan, and the UTP Plan, described below.

Interestingly, consolidated equity market data is not mandated in many other jurisdictions around the world. For example, many European stocks are traded on multiple stock exchanges in the European Union, but each stock exchange there distributes its data as it sees fit and is not required to channel trade and quote data into consolidated feeds. In these jurisdictions, third-party vendors are typically the consolidators, aggregating individual exchange feeds together for use by market participants.

A. Consolidated Data

1. National Market System Plans

Under the U.S. regulatory framework developed in the early 1970s, certain trade and quote data must be disseminated through consolidated data feeds administered by NMS Plans regulated by the SEC.⁴³ Market participants, media outlets, and others subscribe to the consolidated data feeds to obtain data on current market quotes and trade reports. The NMS Plans collect fees from sales of consolidated data and distribute the revenues, net of certain expenses, back to the plan participants. Since 2007, revenues have been allocated among participants based on a formula established by the SEC in connection with the adoption of Reg

⁴³ This requirement is laid out in Rule 603(b) (17 CFR 242.603(b)).

NMS.⁴⁴ In some cases, NMS Plan participants then pass a portion of these revenues on to other market participants through rebate programs.⁴⁵

The two organizations responsible for overseeing the dissemination and sales of consolidated data for U.S. equity markets are the Consolidated Tape Association (“CTA”) and the UTP Plan.⁴⁶ They oversee the process under which trade and quote information is collected from the NMS Plan participants, consolidated, and disseminated to subscribers.

The CTA oversees the operations of the Consolidated Tape System (“CTS”), launched in 1974, and the Consolidated Quote System (“CQS”), launched in 1978.⁴⁷ The members or “participants” of the CTA Plan and CQ Plan include every registered stock exchange and FINRA.⁴⁸ (See Appendix Table A).

Trade and quote data for securities with a primary listing on the NYSE are distributed through CTA’s Network A (also known as Tape A), and trade and quote data for securities with primary listing on another non-Nasdaq exchange are distributed through CTA’s Network B (Tape B).⁴⁹ Historically, Network B consisted of securities listed on the American Stock Exchange (now known as NYSE American). After Archipelago Exchange became part of the

⁴⁴ For a description of the current formula, see SEC Release No. 34-51808.

⁴⁵ For a description of the history of market data rebate programs, see Cecilia Caglio and Stewart Mayhew, “Equity Trading and the Allocation of Market Data Revenue,” *Journal of Banking & Finance* 62 (2016), pp. 97–111.

⁴⁶ More information about these plans, including governing documents, is available on their websites, www.ctaplan.com and www.utpplan.com. For information about the NMS Plan tasked with overseeing collection and distribution of data in the options market, see www.opradata.com.

⁴⁷ For governing documents, see the Consolidated Tape Association Plan and the Consolidated Quotation Plan (“CQ Plan”).

⁴⁸ As of March 2018, there are 16 participants: New York Stock Exchange, NYSE Arca, NYSE American, NYSE National, Nasdaq Stock Market, Nasdaq BX, NASDAQ PSX, ISE Stock Exchange, CBOE Stock Exchange, BZX Equities, BYX Equities, EDGA Equities, EDGX Equities, the Chicago Stock Exchange, the Investors’ Exchange, and FINRA.

⁴⁹ Note that it is the primary listing venue, not the trade or quote venue, that determines the reporting network. Thus, trades and quotes on securities with a primary listing on the NYSE are distributed through Network A, even if the trade or quote occurred on another exchange.

NYSE Group in 2006, NYSE Arca became a popular listing venue for exchange-traded funds and structured products. More recently, Cboe's BZX exchange has adopted a similar listing strategy, and as of 2018, Network B includes securities with primary listings on NYSE Arca, NYSE American, and BZX. When the CTA was developed in the early 1970s, FINRA's predecessor, the National Association of Securities Dealers, operated a nascent system called NASDAQ for dealers to post quotes for stocks not listed on any exchange. The data distributed by the CTA did not include these stocks. The UTP Plan was developed to oversee the dissemination and sales of market data for stocks listed on Nasdaq, through a data channel known as Network C (Tape C).⁵⁰ Today, trade and quote data for securities with a primary listing on the Nasdaq exchange are distributed through Network C.

The 1975 regulatory framework also created the concept of a Securities Information Processor, or SIP, an entity registered with the SEC that is responsible for handling the mechanics of disseminating consolidated market data.⁵¹ Accordingly, consolidated data is sometimes referred to as "SIP data." The SIP for the CTA is the Securities Industry Automation Corporation ("SIAC"), now a subsidiary of NYSE Group, and the SIP for the UTP Plan is Nasdaq.

2. Trade and Quote Data

The CTA and UTP Plans govern the collection and initial distribution of consolidated market data. Subscribers (including third-party vendors) to the consolidated data feeds have

⁵⁰ Note that the CTA and UTP feeds do not provide trade or quote data for securities that are quoted on the OTC Markets (formerly known as the "Pink Sheets") or FINRA's OTC Bulletin Board. Market data feeds are available for such stocks from OTC Markets, but these are not considered NMS Securities, and OTC market data distribution is not governed by an NMS Plan. See www.otcmarkets.com.

⁵¹ See Section 11(a) of the Securities Exchange Act of 1934.

contractual limitations on their ability to redistribute consolidated market data for a period of 15 minutes. Data more than 15 minutes old is considered “historical data” and subscribers, including third-party vendors, can use that data as they wish, including redistributing or reselling the data without any payments to the CTA and UTP Plans. Databases of historical consolidated trade and quote data are widely used by market participants, academics, and regulators for research purposes and forensic analysis, in the form of the NYSE’s TAQ database and analogous products sold by other vendors such as Thomson Reuters Tick History.

Trade data in the consolidated feed includes the ticker symbol, time stamp, execution price, number of shares executed, information about the reporting venue, and various condition codes indicating special circumstances. Trades must be reported regardless of whether they are executed on an exchange, executed on an ATS (i.e., a dark pool or ECN), or executed by an internalizer (a broker that fills a client’s order using its own inventory) or wholesale market maker. Prior to October 31, 2013, trades for fewer than 100 shares (known as odd-lot trades) were not reportable.⁵²

Trades executed on an exchange are reported with an exchange identifier. Trades executed off-exchange are reported to FINRA, typically through a Trade Reporting Facility, or TRF. These trades are identified on the consolidated feed as having been reported through a TRF, but the execution venue is not identified.⁵³ Thus, it is possible to identify which trades were executed off-exchange, but not whether the off-exchange trades were internalized, routed to a wholesaler, or executed on an ATS or dark pool.

⁵² SEC Release Nos. 34-70793 and 34-70794.

⁵³ SEC Release No. 34-61358.

Quote data included in the consolidated feed includes time-stamped “top-of-book” quotes from each exchange, including exchange best (lowest) offer price, number of round lots available at the best offer, best (highest) bid price, number of round lots available at the best bid, and various condition codes indicating special circumstances. The consolidated feed also contains quotes displayed by off-exchange market makers or ATSS on FINRA’s ADF and information about market conditions such as limit up/limit down events and trading halts.

B. Exchange Proprietary Market Data Products

Exchanges have also developed various market data products that they sell directly to subscribers. These data products generally differ from the SIP feeds. Data products sold by the exchanges include data feeds containing trades and quotes, depth-of-book information, and messages related to price discovery around the opening and closing auctions. Other data products sold by the exchanges include historical trade, quote, and order book data at all price levels, daily data summarizing trading activity by security, and reference data including information about securities, corporate actions, and indices.

Different market data products offered by the exchanges are designed for different types of market participants with different needs:

- Some market participants find that the consolidated feeds are sufficient; these participants have little or no need to purchase data directly from exchanges.
- Institutional brokers and proprietary trading desks may subscribe to some or all exchanges’ depth-of-book data feeds as inputs to their order routing algorithms or to help them work large orders. For example, an executing broker might break up a large order into smaller pieces submitted to multiple venues. Depth-of-book feeds could help that broker decide which venues to send the orders to and the prices at

which it should submit each order. These feeds would also help the broker readjust the pricing or venue for those orders based on evolving market conditions. For this purpose, “level data,” which summarizes the total amount of liquidity displayed at each price, may be sufficient.

- Other market participants, such as high-frequency trading firms, may be implementing market making operations or other trading strategies that rely on having low-latency access to order book information, or more granular information about the orders in an exchange’s book. For these market participants, the exchanges offer proprietary feeds with order-level data.
- Finally, some market participants may be interested in back-testing trading strategies or order submission strategies, for which highly granular historical data products can be useful.

1. Exchange Trade and Quote Feeds

Prior to 2005, SEC rules prohibited exchanges from distributing trade reports through channels other than the consolidated feed.⁵⁴ The reforms adopted as part of Reg NMS in 2005 permitted exchanges to distribute trade reports through direct feeds, and more generally provided a regulatory framework for all sales of data through direct feeds.⁵⁵ Shortly after Reg NMS was adopted, there was an increase in the use of proprietary data feeds by market participants to get access to trades and top-of-book quote information faster than they could get it through SIPs. As

⁵⁴ See SEC Rules 11Aa3-1(c)(2) and 11Aa3-1(c)(3), which were rescinded with the passage of Reg NMS in 2005 (see SEC Final Rule Release No. 34-51808).

⁵⁵ See Rule 603 of Reg NMS. For a discussion of this change, see SEC Release No. 34-49325.

described below, SIP latencies have decreased substantially in recent years due to technological improvements.

When it proposed and adopted Rule 603, the SEC stated that the rule meant that exchanges are prohibited from distributing data through direct channels “on a more timely basis” than they make the same data available to the SIPs. The SEC also clarified explicitly that this does not mean that an exchange must delay dissemination of its direct feeds in an attempt to synchronize the arrival of the feeds to end users. Rather, the SEC interprets Rule 603 as prohibiting an exchange from “transmitting data to a vendor or user any sooner than it transmits the data to a Network processor.”⁵⁶ There is no rule governing the timing of when any data purchaser receives data.

In the last decade, there have been dramatic improvements in the latency for both quotes and trades. In February 2018, for example, the average latency for quotes reported through the SIPs was 0.09 milliseconds for Tape A and Tape B securities and 0.017 milliseconds for Tape C securities. These quote latencies represent a significant reduction since the first quarter of 2010, when the average latency was 4.04 milliseconds for Tape A and Tape B securities and 5.42 milliseconds for Tape C securities. There have been similar improvements in trade-reporting times. The average latency for trades reported through the SIPs fell from 6.46 milliseconds in the first quarter of 2010 to 0.15 milliseconds in February 2018 for Tape A and Tape B securities, and from 6.06 milliseconds to 0.017 milliseconds for Tape C securities over the same period.⁵⁷

⁵⁶ SEC Release No. 34-51808, pp. 269–271.

⁵⁷ “Key Operating Metrics of Tape A&B U.S. Equities Securities Information Processor (CTA SIP),” *Consolidated Tape Association*, Q4 2017; “UTP Q1 2018 - February TAPE C QUOTE METRICS,” *Unlisted Trading Privileges*, February 2018; U.S. Equities Securities Information Processor (UTP SIP) Key Quarterly Operating Metrics of Tape C,” *Unlisted Trading Privileges*, Q4 2015.

2. Depth of Book Data

The market data products sold directly by exchanges include real-time limit order book information. Although SIP data contains quotes displaying the number of shares available at each exchange's best bid and offer (top-of-book quotes), the direct data feeds available from exchanges include "depth-of-book" information about displayed liquidity at other price levels below the exchange's best bid and above the exchange's best offer.

Some depth-of-book data products include only aggregate information about the number of shares available at each price point, whereas others provide more granular information on individual orders. Some depth-of-book products provide an updated view of the limit order book at fixed time intervals, whereas others are updated in event time.

Historically, limit order book information for NYSE-listed stocks was available only at the specialist's post on the floor of the exchange. The introduction of NYSE's OpenBook in 2002 was the first time that market participants off the trading floor could see the number of shares available in the NYSE's order book at price levels outside the NYSE's best bid and offer quotes.

When it was originally launched, OpenBook was distributed only through third-party vendors, included the aggregate number of shares available at each bid and offer price provided, and was updated every 10 seconds.⁵⁸ Over time, the OpenBook product has improved markedly both in terms of speed and granularity. Today, NYSE offers OpenBook Aggregated, a feed similar to the original OpenBook product but updated every second, and OpenBook Ultra, which is updated with every limit order event in real time.

⁵⁸ SEC Release No. 34-45138.

For Nasdaq stocks, market data feeds summarizing the top-of-book liquidity (Level 1) and quotes from individual dealers at all prices (Level 2) have long been available to market participants. Level 2 quotes first became broadly available to public market participants with the development of the Nasdaq Quotation Dissemination Service in 1983.⁵⁹ Currently, Nasdaq's main depth-of-book product is TotalView, which shows full depth at each price level for any security that can be traded at Nasdaq. TotalView also shows odd-lot orders, as well as order imbalance information for opening and closing auctions each day, for IPOs, and for the reopening of trading after trading halts.

Cboe has similar real-time product offerings which include top-of-book and depth-of-book data for the BZX, BYX, EDGA, and EDGX exchanges. Customers can purchase trade and quote data, last sale data, or a composite product that offers both, along with aggregated depth-of-book data. Cboe also offers historical market data for its quote, trade, and depth products.

C. Third-Party Vendors

Market data is widely available from third-party vendors. These vendors provide integrated access to a wide variety of services to assist their clients in their trading activities. The vendors' services include access to the real-time market data that SIPs and exchanges provide, as well as reference and valuation data, analytics, news, independent research, and trading platforms. Investment professionals rely on the technology from third-party vendors to not only access market data, but to interact with it and to trade. There is a large market for these services, and they generate substantial revenues. As detailed elsewhere in the paper, the revenues generated by third-party vendors from selling their services are an order of magnitude

⁵⁹ SEC Release No. 34-79863.

larger than the revenues generated by SIPs and exchanges through sales of consolidated and proprietary data.

CTA and UTP Plan Participants

Participant	Predecessors	Reporting Code
New York Stock Exchange		N
NYSE Arca	Pacific Exchange/Archipelago Exchange (–2006)	P
NYSE American	American Stock Exchange (–2008); NYSE Alternext US/NYSE Amex/NYSE MKT (2008–2017)	A
NYSE National	Cincinnati Stock Exchange (–2003); National Stock Exchange (2003–2011)	C
Nasdaq Stock Market		T/Q
Nasdaq BX	Boston Stock Exchange (–2008)	B
Nasdaq PSX	Philadelphia Stock Exchange (–2008)	X
BZX	BATS Z/Bats BZX (2005–2017)	Z
BYX	BATS Y/Bats BYX (2005–2017)	Y
EDGA	EDGA/Bats EDGA (1998–2017)	J
EDGX	EDGX/Bats EDGX (1998–2017)	K
Chicago Stock Exchange	Midwest Stock Exchange (–1993)	M
The Investors Exchange		V
FINRA	NASD (–2007)	D

Source: CQ Plan - Composite as of May 3, 2018; CTA Plan - Composite as of May 3, 2018; UTP Plan Effective as of January 9, 2018; SEC Self-Regulatory Organization Rulemaking Website: <https://www.sec.gov/rules/sro.shtml>

Note: Participants currently receiving Plan revenues are included. The list of predecessors is not exhaustive. The Chicago Board Options Exchange (Cboe) and the International Securities Exchange (ISE) are also listed as participants in the CTA and UTP Plans. Cboe and ISE are active options exchanges. At one time they operated stock exchanges, known as the Cboe Stock Exchange and the ISE Stock Exchange, respectively, but these exchanges are no longer operational: Cboe Stock Exchange has not generated any market data revenues since 2014, and ISE Stock Exchange has not generated any market data revenues since 2010.