Market Efficiency in U.S. Stock Markets: A Study of the Dow 30 and the S&P 30

A Thesis Presented

by

Colin M. Van Oort

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Mathematics

May, 2018

Defense Date: March 30th, 2018
Thesis Examination Committee:

Christopher M. Danforth, Ph.D., Advisor
Bill Gibson, Ph.D., Chairperson
Brian F. Tivnan, Ph.D.
Peter S. Dodds, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of Graduate College

Abstract

The U.S. National Market System (NMS), the largest marketplace in the world for securities and exchange traded funds, suffers from geographic market fragmentation which leads to reduced market efficiency.

Communication lines transmit price updates and other information between geographically isolated exchanges at varying speeds, bounded above by the speed of light. Market participants have access to federally mandated information provided by the Securities Information Processor (SIP) and privately offered information provided by the exchanges, often called direct feeds. These feeds are quantitatively and qualitatively distinct, with the direct feeds tending to provide more information at a faster rate than the SIP feed.

Differences between the SIP and direct feeds can lead to information asymmetries between market participants, which in turn create arbitrage opportunities. Under the market conditions of the NMS in 2016, these arbitrage opportunities occur regularly and many can be captured by market participants with fast connectivity. Several methods exist which allow market participants to reduce their communication latency with trading centers, including the practice of co-location where market participants pay to have their trading infrastructure located in the same building as the matching engines of an exchange.

Such regularly occurring and executable arbitrage opportunities run counter to the Efficient-Market Hypothesis (EMH) in all forms, where even the weak form of the EMH claims that market participants should not be able to systematically profit from market inefficiencies [1, 2].

This thesis investigates the market inefficiencies and related effects introduced by geographic market fragmentation in two baskets of stocks: the Dow Jones Industrial Average (Dow), and the 30 largest stocks by market capitalization in the Standard & Poor's 500 index (S&P 30).

Acknowledgements

I would like to thank Brian Tivnan, Chris Danforth, and Peter Dodds for their invaluable guidance throughout the development process of this thesis, as well as David Dewhurst and John Ring for their assistance in data collection, optimization, and brain storming. Last, but not least, thank you to Tyler Gray for his help with proof-reading and editing.

Table of Contents

Acknowledgements . . . ii
List of Figures . . . vi
List of Tables . . . vii

1 Introduction **1**
1.1 The U.S. National Market System . . . 1
1.2 Regulation NMS . . . 2
1.2.1 Order Protection Rule . . . 2
1.2.2 Access Rule . . . 3
1.2.3 Sub-Penny Rule . . . 3
1.2.4 Market Data Rules . . . 4
1.3 Market Macro-Structure . . . 4
1.4 Data Feeds . . . 5
1.4.1 Thesys Technologies . . . 6
1.5 Market Micro-Structure . . . 6
1.5.1 Continuous Double Auction . . . 7
1.5.2 Frequent Batch Auction . . . 8
1.6 Market Indices . . . 9
1.6.1 The Dow . . . 9
1.6.2 The S&P 500 . . . 10
1.6.3 The S&P 30 . . . 10

2 Methods **11**
2.1 Arbitrage Opportunities . . . 11
2.1.1 Calculation of Arbitrage Opportunities . . . 12
2.2 Realized Opportunity Cost . . . 12
2.2.1 Calculation of ROC . . . 13

3 Results **15**
3.1 Qualities of Arbitrage Opportunities . . . 15
3.1.1 Duration . . . 15
3.1.2 Start Time . . . 18
3.1.3 Magnitude . . . 19
3.2 The Money At Stake . . . 21
3.3 Comparison of the Dow and the S&P 30 . . . 27
3.3.1 Differences in Arbitrage Opportunity Distributions . . . 27
3.3.2 Differences in ROC Statistics . . . 29

4 Conclusion **30**
4.1 Discussion . . . 30
4.2 Comparison With Previous Results . . . 31
4.3 Limitations . . . 32
4.4 Future Work . . . 32
4.5 Conclusion . . . 33

Bibliography **34**

A Glossary **36**
A.1 Market Architecture 36
A.2 Financial Instruments 37
A.3 The Best Bid/Offer 37
A.4 Dislocations 38
A.5 Market Actions 39

B Supplementary Figures **41**

C Supplementary Tables **46**

List of Figures

3.1 Left: Empirical distribution of durations (in seconds) for arbitrage opportunities in stocks that are in the union of the Dow and the S&P 30 plotted with linear axes. Right: Identical distribution plotted with a logged x-axis, arbitrage opportunities with a duration of 0 were excluded. Any observed "0" duration dislocations are the result of timestamps with microsecond resolution and events occurring at sub-microsecond timescales. 16

3.2 Empirical distribution of start times for arbitrage opportunities in stocks that are in the union of the Dow and the S&P 30. Note that on this scale 9:30am corresponds with 34200 and 4:00pm corresponds with 57600, and the histogram is approximately binned by the minute. 18

3.3 Empirical distribution of start times for arbitrage opportunities that are longer than 569 microseconds. Note that on this scale 9:30am corresponds with 34200 and 4:00pm corresponds with 57600, and the histogram is approximately binned by the minute. 19

3.4 Left: Empirical distribution of arbitrage opportunity magnitudes (in USD) for the union of the stocks in the Dow and the S&P 30, plotted with linear axes. Right: Empirical distribution of arbitrage opportunity magnitudes (in USD) for the union of the stocks in the Dow and the S&P 30, plotted with a logged y-axis. 21

3.5 Several truncations of the empirical distribution of arbitrage opportunity magnitudes, allowing for closer inspection of the head of the distribution. From left to right, top to bottom: Magnitudes less than \$1000, \$400, \$50, \$25, \$10, and \$1. 22

B.1 Left: Empirical distribution of durations (in seconds) for arbitrage opportunities in Dow stocks, plotted with linear axes. Right: Identical distribution plotted with a logged x-axis, arbitrage opportunities with a duration of 0 were excluded. Any observed "0" duration dislocations are the result of timestamps with microsecond resolution and events occurring at sub-microsecond timescales. 42

B.2 Left: Empirical distribution of durations (in seconds) for arbitrage opportunities in S&P 30 stocks, plotted with linear axes. Right: Identical distribution plotted with a logged x-axis, arbitrage opportunities with a duration of 0 were excluded. Any observed "0" duration dislocations are the result of timestamps with microsecond resolution and events occurring at sub-microsecond timescales. 42

B.3 Left: Empirical distribution of start times for arbitrage opportunities in Dow stocks, plotted with linear axes. Right: Conditional empirical distribution of start times, where only arbitrage opportunities with a duration longer than 569 microseconds are considered. 43

B.4 Left: Empirical distribution of start times (in seconds) for arbitrage opportunities in S&P 30 stocks, plotted with linear axes. Right: Conditional empirical distribution of start times, where only arbitrage opportunities with a duration longer than 569 microseconds are considered. 43

B.5 Several truncations of the empirical distribution of arbitrage opportunity magnitudes in Dow stocks, allowing for closer inspection of the head of the distribution. Plotted on semi-log axes. From left to right, top to bottom: Magnitudes less than \$1000, \$400, \$50, \$25, \$10, and \$1. 44

B.6 Several truncations of the empirical distribution of arbitrage opportunity magnitudes in S&P 30 stocks, allowing for closer inspection of the head of the distribution. Plotted on semi-log axes. From left to right, top to bottom: Magnitudes less than $1000, $400, $50, $25, $10, and $1. 45

List of Tables

3.1 Summary statistics for the distribution of arbitrage opportunity durations (in seconds) in the stocks of the Dow, the S&P 30, and their union. Summary statistics are computed for all observed arbitrage opportunities (Unfiltered), arbitrage opportunities with a duration longer than 0 microseconds (0 filtered), and arbitrage opportunities longer than 569 microseconds (569 filtered). 17

3.2 Summary statistics for the distribution of magnitudes of arbitrage opportunities (in USD) in the stocks of the Dow, the S&P 30, and their union. Three estimates of the magnitude distribution are provided: an upper estimate, which is calculated using the maximum magnitude of each observed arbitrage opportunity; a lower estimate, which is calculated using the minimum magnitude; and a mean estimate, which is the element-wise mean of the upper and lower estimates. 23

3.3 Summary statistics for the distribution of magnitudes of arbitrage opportunities (in USD) with durations longer than 569 microseconds. 24

3.4 Summary ROC statistics over the 252 trading days in 2016 for the union of stocks in the Dow and the S&P 30. Traded Value, Differing Traded Value, ROC, and ROC Per Share are all dollar values. 25

3.5 Direct comparison between the union of the Dow and the S&P 30, the Dow, and the S&P 30 using a variety of ROC statistics. Traded value, differing traded value, ROC, and ROC per share represent USD amounts, while the other statistics do not have an associated unit. 25

3.6 Estimates of the total arbitrage profits (in USD) available to a perfect arbitrager of the union of the Dow and the S&P 30, the S&P 30, and the Dow in 2016. Estimates are based upon the expected value of the indicated magnitude distribution (in USD), the number of observed actionable (duration $>$ 569 microseconds) arbitrage opportunities in that group in 2016, and the knowledge that the minimum size for a protected quotation is 100 shares. 26

3.7 Statistical tests comparing the distribution of arbitrage opportunity attributes between the Dow and the S&P 30. The Jensen-Shannon Divergence is a distance measure, not a statistical test, so it has no associated p-value. 28

3.8 Summary ROC statistics over the 252 trading days in 2016 for stocks in the Dow. Summary statistics are computed over daily aggregates of ROC statistics. Traded Value, Differing Traded Value, ROC, and ROC Per Share are all dollar values. . . . 29

3.9 Summary ROC statistics over the 252 trading days in 2016 for stocks in the S&P 30. Summary statistics are computed over daily aggregates of ROC statistics. Traded Value, Differing Traded Value, ROC, and ROC Per Share are all dollar values. . . . 29

C.1 Share price (in USD) and market capitalization (in USD) for the stocks under study, computed using the average of four observations taken at the end of each quarter in 2016, as well as rankings of the stocks according to their share price and market capitalization. Since the S&P 30 is a market cap. weighted index, the Market Cap. Rank indicates the relative effect each stock would have on the S&P 30. The Share Price Rank has a similar interpretation for the Dow, since it is a share price weighted index. 47

C.2 Summary ROC statistics over the 252 trading days in 2016 itemized by stock. . . . 56

Chapter 1

Introduction

The U.S. National Market System

Stock markets have existed in the U.S. for nearly its entire lifetime, with the creation of the Philadelphia Stock Exchange (PSX) in 1790 and the first stirrings of the New York Stock Exchange (NYSE) occurring in 1792. However, these early markets were highly fragmented and scarcely regulated, allowing for predatory or manipulative trading and wide deviations in the pricing of assets. As markets became larger, regulatory oversight increased.

The Securities Exchange Act of 1934 formed the basis of federal regulations on stock trading and instituted the Securities Exchange Commission (SEC), the authority responsible for enforcing market regulations and proposing new rules. In addition to its official responsibilities the SEC strives for other goals including the protection of investors, the maintenance of orderly and efficient markets, and the facilitation of capital formation [3]. In 1972 the SEC began its pursuit of a more unified National Market System (NMS), and Congress authorized the SEC to develop the NMS under the Securities Acts Amendments of 1975. These efforts culminated in the a set of regulations titled Regulation National Market System [4], often shortened to Reg. NMS, which aimed to tighten the coupling between disparate exchanges in light of the development of electronic exchange mechanisms which increase trading speeds and tightened bid-ask spreads.

Regulation NMS

Reg. NMS was created with the overarching goal of creating a unified National Stock Market and with two design objectives: to promote competition between markets and between orders, and to serve the interests of long-term investors and listing companies [4]. Reg. NMS is composed of several rules and regulations, the most important of which are summarized below. See [4] for more details.

Order Protection Rule

The Order Protection Rule (Rule 611), also known as the Trade-through Rule, is meant to protect orders from trade-throughs, which occur when a market center matches an order against a local counter-party when a better price is available via a protected quotation displayed by an alternative market center. Note that a "better" price in this context is defined from the perspective of the new order entering the market. Therefore a lower execution price is be considered better for an entering bid (offer to buy), while a higher execution price is be considered better for an entering ask (offer to sell).

A protected quotation is defined in Reg. NMS as a bid or ask quotation that satisfies the following properties: the quotation must be automated, the quotation must be displayed by an automated trading center, and the quotation must offer the lowest ask price or highest bid price among all publicly displayed quotations.

A quotation is considered automated if it may be executed without human intervention (up to the full listed quantity), allows for the correct execution of Immediate-Or-Cancel (IOC) orders against the quotation, immediately provides a response to the sender of an Immediate-Or-Cancel order indicating the execution status of that order, and immediately updates the quotation to reflect any changes to its status.

A trading center is considered automated if it implements systems and procedures that allow it to display automated quotations as defined above, and quotations that do not satisfy the requirements of an automated quotation are identified as manual quotations as quickly as possible.

Trade-throughs are prohibited Under Rule 611, however exceptions are allowed for Intermarket

Sweep Orders (ISO), quotations displayed by markets that fail to meet the reporting requirements for automated quotations, and flickering quotations with multiple prices displayed in a single second.

Access Rule

The Access Rule (Rule 610) concerns itself with setting standards for access to quotations in NMS stocks, and caps the fees that an exchange may charge for accessing its protected quotations at $0.003 per share. Rule 610 allows for the creation and usage of private data feeds, often referred to as direct feeds by market participants since they are offered directly by exchanges rather than through a third party. Rule 610 also prohibits trading centers from displaying quotations which would lock or cross a protected quotation from a different trading center.

A market is said to be locked if the bid-ask spread of that market is zero, in other words there exists a resting bid and a resting ask with identical limit prices. A market is said to be crossed if the bid-ask spread of that market is negative, i.e. there exists a resting bid whose limit price is greater than the limit price of a resting ask, or equivalently a resting ask exists whose limit price is less than the limit price of a resting bid. These effects are the result of coupling geographically fragmented exchanges, since an order that may lock or cross a market would immediately find a counter-party if the two orders were present on the same exchange.

Sub-Penny Rule

The Sub-Penny Rule (Rule 612) prohibits market participants from displaying or accepting quotations for NMS stocks priced in an increment less than $0.01 unless the quotation price is less than $1.00, in which case the minimum increment is $0.0001. Rule 612 is meant to prohibit the practice of "sub-pennying" in which market participants could "step ahead" of a protected quotation by providing a negligible amount of price improvement, allowing the "sub-pennied" order faster execution at effectively no extra cost.

The significance of this rule, with respect to geographic fragmentation and market inefficiencies, is that the minimum increment for the quoted price of a traded instrument sets the minimum profits which are made available by an arbitrage opportunity.

Market Data Rules

Rules 601 and 603 are referred to as Market Data Rules and are meant to promote wide availability of market data, thus providing all market participants with an accurate and reliable source of information on the best prices in NMS stocks. These rules cover the organization of a consolidated data feed for NMS stocks, the reward structure for contributing information to the consolidated data feed, and establishes standards for quote and trade information provided to and provided by the consolidated data feed.

In particular these rules concern the Consolidated Tape Association (CTA) plan which disseminates transaction information for NYSE listed securities, the Consolidated Quotation (CQ) plan which disseminates quote information for NYSE listed securities, and the Nasdaq UTP plan which disseminates consolidated quote and trade data for Nasdaq listed securities. The information provided by the CTA plan and CQ plan forms Consolidated Tape A, and the information provided by the UTP plan forms Consolidated Tape C. There also exists a Consolidated Tape B which reports trade information for stocks listed on regional exchanges. The aggregation of Consolidated Tapes A, B, and C form what is commonly referred to as the SIP feed.

Market Macro-Structure

The NMS, and market systems in general, may be considered a network of market centers connected by a a set of communication channels[1]. Modern electronic market places are composed of computers communicating order information over network infrastructure, which primarily consists of fiber optic cables. Trading centers are individual computers or groups of computers that receive order information from market participants, maintain local order books containing active orders, match orders to execute trades, and more.

In 2016 there were 12 active exchanges in the NMS: the NYSE family of exchanges (NYSE, NYSE ARCA, NYSE MKT), the Nasdaq family of exchanges (Nasdaq, Nasdaq BX, Nasdaq PSX), the BATS family of exchanges (BATS BYZ, BATS BZX, DirectEdge EDGA, DirectEdge EDGX),

[1] See Definition A.1.1 for more details.

The Investors Exchange (IEX), and The Chicago Stock Exchange (CHX). The matching engines for these exchanges are located in data centers in northern New Jersey, more specifically the NYSE Euronext Data Center in Mahwah, the NASDAQ Data Center in Carteret, the Equinix Data Center in Secaucus, and the CenturyLink Data Center in Weehawken. [5] provides an convenient graphical presentation of the geographic configuration of the NMS in addition to information about communication latencies between each location.

Alternative Trading Systems (ATS), which are similar to exchanges but are subject to a different set of regulations [6], may also trade NMS stocks and tend to be less geographically clustered. Forty four ATSs actively traded NMS stocks in 2013 according to [7], thus as many as 56 exchanges and ATSs may have actively traded NMS stocks in 2016 from a variety of geographical locations [8].

This geographic market fragmentation has a tangible effect on the efficiency of the NMS in the form of arbitrage opportunities, which are guaranteed to form since information cannot be transmitted between trading centers faster than the speed of light.

Data Feeds

Electronic market systems generate vast amounts of data, which can be made available to market participants in the blink of an eye, and there exist multiple vehicles for obtaining market data in varying quantity and quality. Data providers include the federally mandated Securities Information Processor (SIP) feed, direct feeds provided by individual exchanges, and third party data providers like Thesys Technologies and Redline Trading Solutions.

The SIP aggregates information about trades and quotes in the NMS. A summarization of these quotes defines the National Best Bid and Offer (NBBO), which is used to provide a notion of "best" price for each security across the NMS. Each exchange in the NMS also provides a "direct feed" which often contains more information than the SIP, such as information about resting orders below the quoted prices, and often provides information more rapidly than the SIP feed.

Thesys Technologies

The data used in this study is derived from an authoritative market dataset provided by Thesys Technologies [9].

Thesys collects data from each of the direct feeds and the SIP feed and combines the information from these sources into a single comprehensive database. Obtaining data through Thesys has two major advantages over using data directly from the SIP and direct feeds.

First, Thesys performs extensive data cleaning and organization making manipulation and interpretation of their data extremely easy. Data inconsistencies, which are somewhat prevalent in certain kinds of market data[2], are reconciled using information from several feeds. Second, Thesys collects all of this data from a single physical observation point in Carteret, New Jersey. Thus the data is synchronized to a single local clock and is available in the exact ordering a market participant at Carteret would have received it. This becomes increasingly important when studying events that occur at a microsecond timescale, where clock synchronization issues may alter the ordering of events and degrade the quality of the data.

Having compiled data from all of the available market feeds, Thesys offers a comprehensive dataset that includes trades, quotes, order flow, and administrative messages. The quality of which is recognized by the SEC, who commissioned Tradeworx (Thesys' parent company at the time) to build their Market Information Data Analytics System (MIDAS) in 2012 [11].

Market Micro-Structure

There are several different mechanisms that can be used to match and execute orders in an electronic exchange, the most common of which is the Continuous Double Auction (CDA). Frequent Batch Auctions (FBA) have recently been proposed as a possible solution to some of the issues introduced by geographic market fragmentation[3].

[2] See inverted timestamps in [10, p. 13] for an example.
[3] See [12, 13] for a detailed analysis of FBAs.

Continuous Double Auction

In a CDA multiple buyers and sellers submit orders to an exchange or auctioneer, which aggregates those orders in a Central Limit Order Book (CLOB) and matches buyers with sellers. The matching of a buyer and seller results in a transaction or trade, which is executed by the exchange on behalf of the market participants when a buy order and a sell order feature compatible prices.

The first two components of the name for this auction type succinctly describe its most prominent features. A CDA, as described above, is said to be a *continuous auction* since market participants may enter orders at an arbitrary time and submitted orders are matched by the exchange as quickly as possible, usually on a best effort basis. Likewise, a CDA is said to be a *double auction* since there are multiple buyers and multiple sellers participating simultaneously, thus forming a double-sided market.

The CLOB, which is maintained by the exchange or auctioneer overseeing the CDA, is a registry of active orders that have yet to find a counter-party (see Definition A.1.3). The orders contained within a CLOB are referred to as resting orders, since they must wait for a counter party before a trade may occur.

Market participants may submit different varieties of orders to an exchange or auctioneer, the most commonly supported order types are the limit order and the market order. A limit order provides the market participant with guarantees about the execution price of their order, but provides no guarantees about when that order will find a counter party. A market order provides the opposite, guaranteeing that a counter party will be found as quickly as possible, but providing no guarantee about the price at which the resulting trade will occur.

A market participant wishing to submit a limit order must provide the exchange or auctioneer with a limit price and a desired quantity. When the exchange or auctioneer receives that order it may be matched with any resting orders on the opposite side of the book whose limit price is at least as good as the limit price of the entering order. If no counter party is immediately available then the entering limit order comes to rest in the order book at its limit price.

A market participant wishing to place a market order need only provide a desired quantity. The exchange or auctioneer may then immediately match the incoming market order with resting orders

on the opposite side of the book until the desired quantity has been matched. It is common to use price-time priority to determine the order in which resting orders are matched, where orders featuring a better limit price (higher for bids, lower for asks) are matched first and orders with an equivalent limit price are matched on a first come, first served basis[4].

Frequent Batch Auction

FBAs appear similar to CDAs from a high level. One may notice that both auction types maintain a CLOB whose purpose and form is identical, though there are several fundamental differences which provide FBAs with theoretical benefits over CDAs.

The first key difference between FBAs and CDAs lies in the matching behavior of FBAs, which match orders in batches after allowing orders to accumulate for some period of time instead of matching orders as quickly as possible. During this accumulation time, orders fill both sides of the book and the accumulation of orders allows for the estimation of supply and demand curves.

The second key difference between the two is seen in how they determine the execution price of an order. After estimating the supply and demand curves as discussed above, the intersection of the supply and demand curves identifies a unique execution price. A FBA may use this information to provide a uniform execution price to all trades that execute in the same batch by allowing bids that cross the identified price to be matched with asks orders that cross the execution price.

This batch matching procedure in combination with the uniform execution price provide the potential for significant price improvement for orders that cross the bid-ask spread. Under certain conditions and assumptions about the availability of market information, FBAs can be shown to promote competition on order price and simultaneously reduce the asymmetries provided by fast information and execution speed. These effects highlight the reasons why FBAs may solve some of the issues introduced by geographic market fragmentation.

[4]See Appendix A.5 for an overview of order types and modifiers, and [14] for more information.

Market Indices

A market index is a theoretical portfolio of stocks whose performance is used to track relative changes in a market. Each index is calculated via a unique formula that usually involves the prices of its constituent stocks. Two well known indices are the Dow Jones Industrial Average (DJIA or Dow), and the Standard & Poor's 500 Index (S&P 500). The Dow and the S&P 500 are meant to track general trends in the NMS and thus are interesting objects of study if one is interested in estimating system-wide properties, such as market efficiency.

The Dow

The Dow is an index of 30 stocks that are selected by the S&P Dow Jones Indices company and is intended to be a total market index that covers all industries except transportation and utilities. The Dow is a price-weighted index and is calculated by summing the prices of a single unit of each of its constituents and then dividing by a proprietary constant which is used to maintain continuity of the index when prices adjust due to stock splits and other events. This calculation may be written as

$$P_{Dow} = \frac{1}{c} \sum_{i=1}^{30} p_i, \tag{1.1}$$

where c is the proprietary constant and p_i is the price of a single unit of stock for the ith constituent [15].

In 2016 the Dow was composed of 3M (MMM), American Express (AXP), Apple (AAPL), Boeing (BA), Caterpillar (CAT), Chevron (CVX), Cisco Systems (CSCO), Coca-Cola (KO), DuPont (DD), Exxon Mobil (XOM), General Electric (GE), Goldman Sachs (GS), IBM (IBM), Intel (INTC), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), McDonald's (MCD), Merck (MRK), Microsoft (MSFT), Nike (NKE), Pfizer (PFE), Procter & Gamble (PG), The Home Depot (HD), Travelers (TRV), United Technologies (UTX), United Health Group (UNH), Verizon (VZ), Visa (V), Walmart (WMT), and Walt Disney (DIS).

The S&P 500

The S&P 500 index is composed of the 500 largest companies by market capitalization listed on the NASDAQ and NYSE stock exchanges and may be calculated via the following equation,

$$P_{S\&P\ 500} = \frac{1}{d}\sum_{i=1}^{500} p_i q_i. \tag{1.2}$$

Where p_i is the price for a single unit of stock for the ith constituent, q_i is the quantity of publicly available shares for the ith constituent, and d is a proprietary divisor which is used to maintain continuity of the index in the presence of company splits, company mergers, and similar events [16].

The S&P 30

The S&P 30 is an index composed of the 30 largest companies in the S&P 500, by market capitalization, constructed for this study with the purpose of direct comparison with the Dow. The S&P 30 varied slightly in 2016, with some of the smaller companies being replaced throughout the year, which makes it unclear which 30 stocks should be considered as the S&P 30 over that period.

In order to make the decision quantitative, the market capitalization of each stock in the S&P 500 is obtained at the end of each quarter in 2016. These four observations per stock are averaged and used to rank the stocks.

The top 30 stocks by this ranking were selected to form the S&P 30 resulting in the following constituents: Alphabet Inc. Class A (GOOGL), Alphabet Inc. Class C (GOOG), Amazon.com (AMZN), Apple (AAPL), AT&T (T), Bank of America (BAC), Berkshire Hathaway (BRK.B), Chevron (CVX), Coca-Cola (KO), Comcast (CMCSA), Exxon Mobil (XOM), Facebook (FB), General Electric (GE), IBM (IBM), Intel (INTC), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), Merck & Co (MRK), Microsoft (MSFT), Oracle (ORCL), PepsiCo (PEP), Pfizer (PFE), Philip Morris International (PM), Proctor & Gamble (PG), The Home Depot (HD), Verizon (VZ), Visa (V), Walmart (WMT), Walt Disney (Dis), and Wells Fargo & Co (WFC). Table C.1 contains the stocks of the Dow and the S&P 30 along with rankings with respect to market capitalization and share price, as well as the measurements that lead to the displayed rankings.

Chapter 2

Methods

From the authoritative dataset provided by Thesys, indicators of market inefficiency such as arbitrage opportunities and realized opportunity costs may be collected. A practical definition of market inefficiency is used in this study, where market conditions that allow market participants to systematically profit from price discrepancies are considered inefficient.

This chapter defines these market inefficiency indicators and describes the procedures used in their collection and processing.

Arbitrage Opportunities

An arbitrage opportunity may be defined as a set of market conditions that allow for profit via the purchase and immediate sale of a good at differing prices. Consider a market system (Definition A.1.1) containing two or more market centers (Definition A.1.2) which are geographically fragmented and connected via two or more data feeds (Definition A.1.6) with differing transmission rates. A dislocation (Definition A.4.2) occurs at any time where a security is quoted at two or more distinct prices among the feeds, and an arbitrage opportunity (A.4.4) is any half open interval of time over which the feeds are dislocated and one of the feeds quotes the best price throughout the entire interval.

Collecting information on arbitrage opportunities provides insights into how tightly coupled the different exchanges and data feeds are, and allows for several estimates of the possible severity of the

market inefficiencies introduced by geographic fragmentation. Additionally, estimates of the latency between exchanges in the market system can be used to filter extremely short arbitrage opportunities in order to provide a characterization of actionable arbitrage opportunities.

Calculation of Arbitrage Opportunities

Each arbitrage opportunity has four properties: a start time, an end time, a maximum value, and a minimum value. Equivalently, a duration may be used in place of an end time when specifying an arbitrage opportunity.

The arbitrage opportunities for a single trading day and a single stock are calculated by first obtaining the complete list of quotes issued by both the SIP feed and the aggregated direct feeds (a combination of the information provided by each exchange) on that date and for that particular stock. Each quote in the SIP feed is matched with the price quoted by the direct feed at the time that the SIP quote was issued, and likewise each quote in the direct feed is matched with the SIP price at the time it was issued. These two sets are then merged and sorted by time of issuance via merge sort, creating a complete view of the dynamics of both feeds throughout the trading day.

Note that quote time-stamps are recorded at the resolution of microseconds (i.e. 1.0×10^{-6} seconds) and that market events like quotations may occur multiple times within the same microsecond. Thus, the stability of merge sort is desirable since it maintains the original ordering of elements with the same time-stamp.

Once this master list of quote information has been assembled, arbitrage opportunities may be identified by iteratively checking that the price quoted by the SIP feed matches the price quoted by the direct feeds. Sequences of one or more consecutive non-matching quotations form arbitrage opportunities, whose attributes may be calculated using the timestamps and prices of the triggering quotes.

Realized Opportunity Cost

An opportunity cost is generally an abstract notion which represents the "cost" associated with a particular choice or set of choices in light of one or more alternatives. Let the realized opportunity

cost (ROC) be defined as the amount of dollars that could have been gained by trading at the price quoted by the aggregated direct feeds instead of the price quoted by the SIP feed at the exact time that a historical trade occurred. This definition provides an estimate for the opportunity cost of using prices quoted by the SIP feed instead of those quoted by the aggregated direct feeds.

Calculation of ROC

The ROC may be calculated by first collecting the information about all trades that occurred in stocks of interest in a desired time period. Each trade is then matched with the quoted prices on both the SIP feed and aggregated direct feeds at the time that it occurred. Trades that did not occur at their respective SIP bid or ask price are removed from consideration. This filtering step is intended to remove trades that were placed using prices quoted by direct feeds or alternative data sources, or trades which were placed with the intent to provide better execution in an alternative metric (i.e., fastest execution, least information leakage, etc.).

The remaining trades in the sample are those that were placed with the goal of best price using the SIP feed to determine what that best price should be. For each of these remaining trades the side of the initiating order is determined based upon whether the execution price matches the quoted bid price or quoted ask price on the SIP feed. For buy-side trades, those where the order which caused the trade was a bid, the ROC for that trade is the difference between the execution price and the best quoted ask price on the aggregated direct feeds. Thus the resulting number is positive if the aggregated direct feeds provided a lower (better) price, and negative if the SIP feed provided a lower (better) price. The calculation for sell-side trades, those where the order which caused the trade was an ask/offer, is found by computing the difference between the best quoted bid price on the aggregated direct feeds and the execution price. This formulation provides an identical interpretation, where a positive value indicates that the price quoted on the aggregated direct feeds is better for the initiator than the price quoted on the SIP feed (since a higher price is better for the initiator of a sell side trade). Trades that occur at a time where the direct feeds and the SIP feed are dislocated will be called differing trades. Differing trades are of interest because trades that occur during dislocations can lead to market inefficiencies.

Finally, the net ROC for a single stock and trading day is calculated by summing the ROC from

each trade.

Chapter 3

Results

Qualities of Arbitrage Opportunities

Before attempting to compare the arbitrage opportunities of two groups of stocks it is helpful to obtain a baseline for what arbitrage opportunities look like and how they behave. In short, this section provides an overview of the distributions of properties of arbitrage opportunities.

Duration

If arbitrage opportunities are to be used as a meaningful indicator of market inefficiency then knowledge of the distribution of the duration of arbitrage opportunities is required along with a notion of the capabilities of market participants. With this information, one can operationalize the definition of market inefficiency provided in Section 2 and begin quantifying the quantity and magnitude of market inefficiency.

Figure 3.1 indicates that the vast majority of dislocations last 1 second or less, but the distribution has a very long tail, indicating that some distributions last for as long as 10000 seconds (almost 3 hours). The left panel of Figure 3.1 fails to display the shape of the distribution since almost all of arbitrage opportunities fall within the first bin, so a logarithmic x-axis is used in the right panel in order to stretch out the contents of the first bin. With the logarithmic x-axis it is revealed that the distribution of durations of arbitrage opportunities is approximately a log-normal distribution, and

Empirical Distribution for the Duration of Arbitrage Opportunities

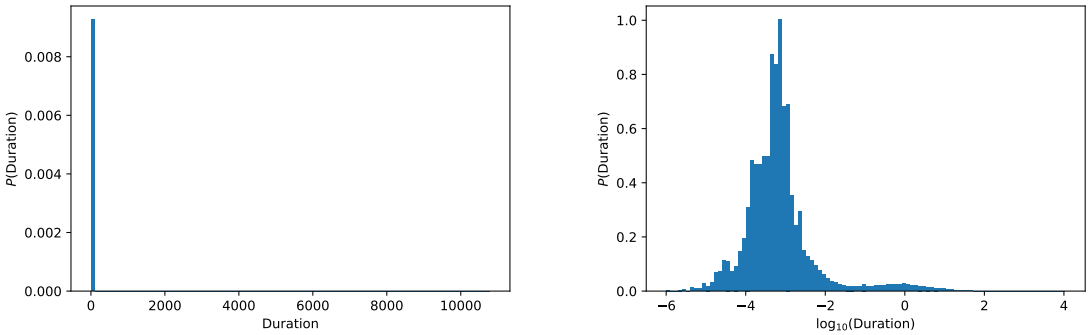

Figure 3.1: Left: Empirical distribution of durations (in seconds) for arbitrage opportunities in stocks that are in the union of the Dow and the S&P 30 plotted with linear axes. Right: Identical distribution plotted with a logged x-axis, arbitrage opportunities with a duration of 0 were excluded. Any observed "0" duration dislocations are the result of timestamps with microsecond resolution and events occurring at sub-microsecond timescales.

that arbitrage opportunities are most likely to have a duration between 100 and 1000 microseconds.

Table 3.1 provides a more quantitative measure of the center of the distribution, showing the median to be 559 microseconds for the left panel of Figure 3.1 and 571 microseconds for the right panel, where 0 duration arbitrage opportunities were excluded due to the use of the logarithm.

Table 3.1 indicates that the minimum duration for arbitrage opportunities is 0 microseconds, which may seem like an alarming number. This is an artifact of having timestamps with microsecond resolution. If an arbitrage opportunity has a duration that is less than 1 microsecond, then the observed duration is 0 microseconds.

Between Figure 3.1 and Table 3.1 a fairly comprehensive view of the durations of arbitrage opportunities can be formed for the sample under study, but this does not touch upon how long a market participant needs in order to identify and capture an arbitrage opportunity. Recent work has shown that a market participant utilizing data transmitted by fiber optic cables would need a minimum of 570 microseconds to reliably identify and capture arbitrage opportunities at an arbitrary exchange [17]. Table 3.1 indicates that 52.86%, 47.86%, and 49.37% of dislocations are longer than 569 microseconds in the Dow, S&P 30, and their union respectively.

Arbitrage Opportunity Duration Distribution Summary

		Combined	Dow	S&P 30
Unfiltered	count	178,476,945	120,355,462	130,374,761
	mean	0.114859	0.073712	0.114163
	std	5.276667	5.519033	4.822610
	min	0.000000	0.000000	0.000000
	25%	0.000218	0.000216	0.000223
	50%	0.000559	0.000624	0.000541
	75%	0.001156	0.001190	0.001095
	max	10789.825626	10789.825626	10789.825626
0 Filtered	count	175,822,596	118,626,324	128,295,437
	mean	0.116593	0.074786	0.116013
	std	5.316329	5.559103	4.861512
	min	0.000001	0.000001	0.000001
	25%	0.000231	0.000227	0.000237
	50%	0.000571	0.000636	0.000551
	75%	0.001176	0.001202	0.001115
	max	10789.825626	10789.825626	10789.825626
569 Filtered	count	88,119,495	63,617,169	62,392,695
	mean	0.232378	0.139245	0.238271
	std	7.507748	7.590569	6.969153
	min	0.000570	0.000570	0.000570
	25%	0.000763	0.000768	0.000747
	50%	0.001173	0.001127	0.001154
	75%	0.002701	0.002484	0.002584
	max	10789.825626	10789.825626	10789.825626

Table 3.1: Summary statistics for the distribution of arbitrage opportunity durations (in seconds) in the stocks of the Dow, the S&P 30, and their union. Summary statistics are computed for all observed arbitrage opportunities (Unfiltered), arbitrage opportunities with a duration longer than 0 microseconds (0 filtered), and arbitrage opportunities longer than 569 microseconds (569 filtered).

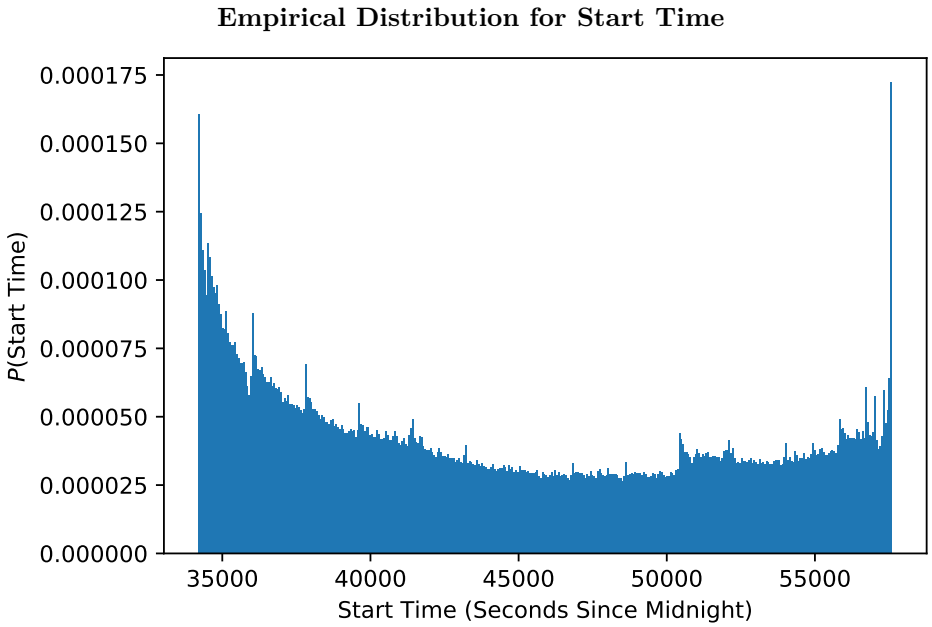

Figure 3.2: Empirical distribution of start times for arbitrage opportunities in stocks that are in the union of the Dow and the S&P 30. Note that on this scale 9:30am corresponds with 34200 and 4:00pm corresponds with 57600, and the histogram is approximately binned by the minute.

Start Time

In addition to the duration of arbitrage opportunities, it is of interest to investigate patterns in the start times of arbitrage opportunities. Figure 3.2, an empirical distribution of arbitrage opportunity start times in the union of the stocks in the Dow and the S&P 30, displays such patterns for 2016.

Important features of this distribution are the rough "U" shape, which indicates that arbitrage opportunities tend to occur at the beginning and end of the trading day; the periodic structure which occurs at approximately 30 minute intervals; and that arbitrage opportunities have the greatest probability of occurring in the last seconds of a trading day.

Following the claim from [17] that market participants need a minimum of 569 microseconds in order to reliably capture an arbitrage opportunity, Figure 3.3 displays the distribution of start times for arbitrage opportunities in the union of the Dow and the S&P 30 which are longer than 569 microseconds. A qualitative comparison of Figures 3.2 and 3.3 reveals only a few small differences between them, indicating that so called "actionable" arbitrage opportunities follow a similar

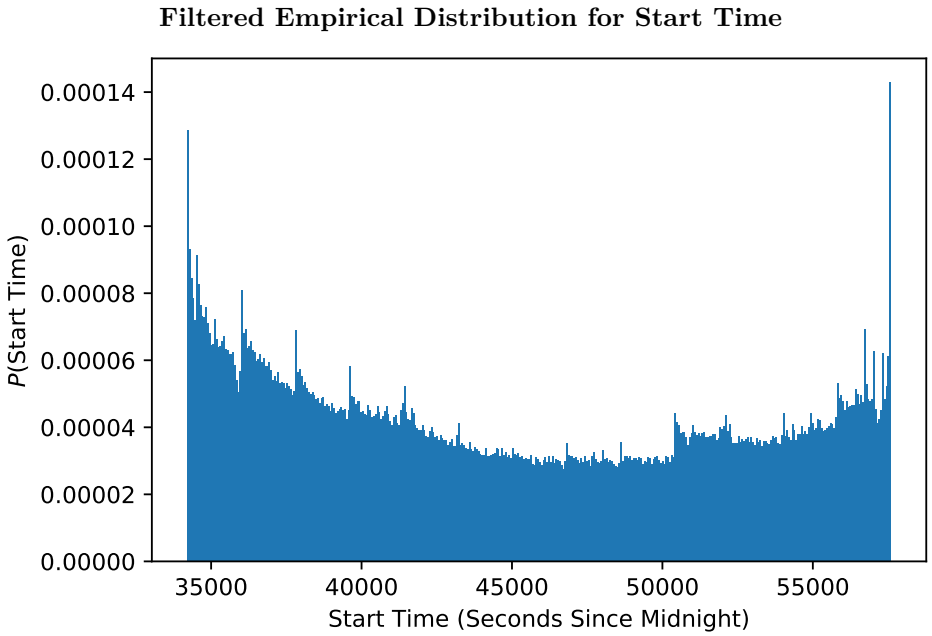

Figure 3.3: Empirical distribution of start times for arbitrage opportunities that are longer than 569 microseconds. Note that on this scale 9:30am corresponds with 34200 and 4:00pm corresponds with 57600, and the histogram is approximately binned by the minute.

distribution of start times. Perhaps the largest difference between the two distributions is that less actionable arbitrage opportunities start in the first or last seconds of the trading day. Additionally, it is worth noting that the periodic structure at approximately 30 minute intervals remains in the conditional distribution and may even be slightly more pronounced.

Magnitude

Under the definition of market inefficiency presented in Section 2, a market participant must be able to systematically profit from a price discrepancy in order for a market inefficiency to occur. In other words, if arbitrage opportunities are not large enough to recover possible losses incurred while executing their capture then they would not be considered a market inefficiency.

Thus it becomes of critical importance to understand the distribution of values that arbitrage opportunities assume. As defined previously, an arbitrage opportunity has two associated values, a minimum value and a maximum value, which are signed numbers obtained by computing the

difference between the prices quoted on the aggregated direct feeds and the SIP feed at a particular time. An argument may be made that regardless of which side of the book is dislocated (bid or ask) and regardless of which feed is favored (direct or SIP) an arbitrager may obtain a profit. If the side of book and feed do not affect the ability to profit from discrepancies between the feeds then all discrepancies may be considered identically as potential market inefficiencies. In the following analysis the minimum and maximum values of an arbitrage opportunity may be reduced to an estimated magnitude.

From the maximum and minimum value of an arbitrage opportunity, various estimates of the magnitude of the arbitrage opportunity may be calculated. In this study an upper estimate, lower estimate, and mean estimate are computed and analyzed jointly. The upper estimate for the magnitude of an arbitrage opportunity is calculated as $\max(|M|, |m|)$ where M is the maximum value of the arbitrage opportunity and m is the minimum value. The lower estimate is calculated in a similar manner, $\min(|M|, |m|)$, while the mean estimate is simply the average of the upper and lower estimates.

With the upper, mean, and lower estimates of arbitrage opportunity magnitude defined the distribution of those estimates may be seen in Figure 3.4. Both panels in Figure 3.4, where the left panel utilizes linear axes and the right panel utilizes semi-log axes, reveal relatively little about the distribution of the magnitudes due to the extremely long tail of the distribution and the vast majority of observations clustered at lower magnitudes.

In order to clearly display the distribution for smaller magnitude arbitrage opportunities a sequence of truncated distributions is shown in Figure 3.5. The lower-left and lower-right panels of Figure 3.5 show that approximately from \$0.05 to \$1.5 the distribution follows a nearly linear scaling in the semi-log axes, which indicates that it may be an exponential distribution with finite size effects at smaller values and sampling noise in the tail. The lower-left panel and middle panels also show that the distribution is extremely noisy and incomplete for larger magnitudes, indicating that there may be too few observations of extreme magnitudes in order to accurately estimate the tail of the distribution.

Table 3.2 provides summary statistics for the distributions shown in Figures 3.4 and 3.5. The quartile statistics for the upper, mean, and lower estimate across all groups of stocks considered

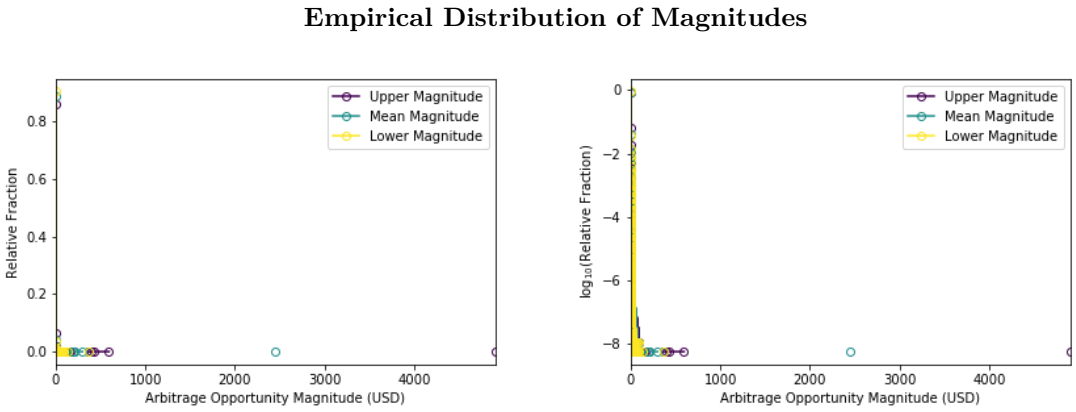

Figure 3.4: Left: Empirical distribution of arbitrage opportunity magnitudes (in USD) for the union of the stocks in the Dow and the S&P 30, plotted with linear axes. Right: Empirical distribution of arbitrage opportunity magnitudes (in USD) for the union of the stocks in the Dow and the S&P 30, plotted with a logged y-axis.

are identically 1 cent, which is to be expected considering that more than 90% of all arbitrage opportunities have a magnitude of 1 cent.

Continuing to investigate actionable arbitrage opportunities, Table 3.3 contains similar summary statistics for the magnitudes of arbitrage opportunities that lasted longer than 569 microseconds. The quartile statistics of the magnitudes of actionable arbitrage opportunities are identical to the unfiltered statistics, and the maximum values of the upper and lower estimates are also unaffected by the filtering, which implies that some of the largest arbitrage opportunities are actionable.

The Money At Stake

Considering that the vast majority of actionable arbitrage opportunities only have a magnitude of $0.01, it may seem that there is no cause for concern. This section estimates the amount of money that is at stake, and shows that arbitrage opportunities compensate for their small magnitude with sheer frequency.

Table 3.4 summarizes the variation in the daily ROC for the stocks in the Dow and the S&P 30 along with information about the number of trades that occurred in those stocks and the total dollar value of the traded stocks. This table provides a slight indication of the staggering amount of money and information that flows through the NMS on a daily basis, where the average stock in

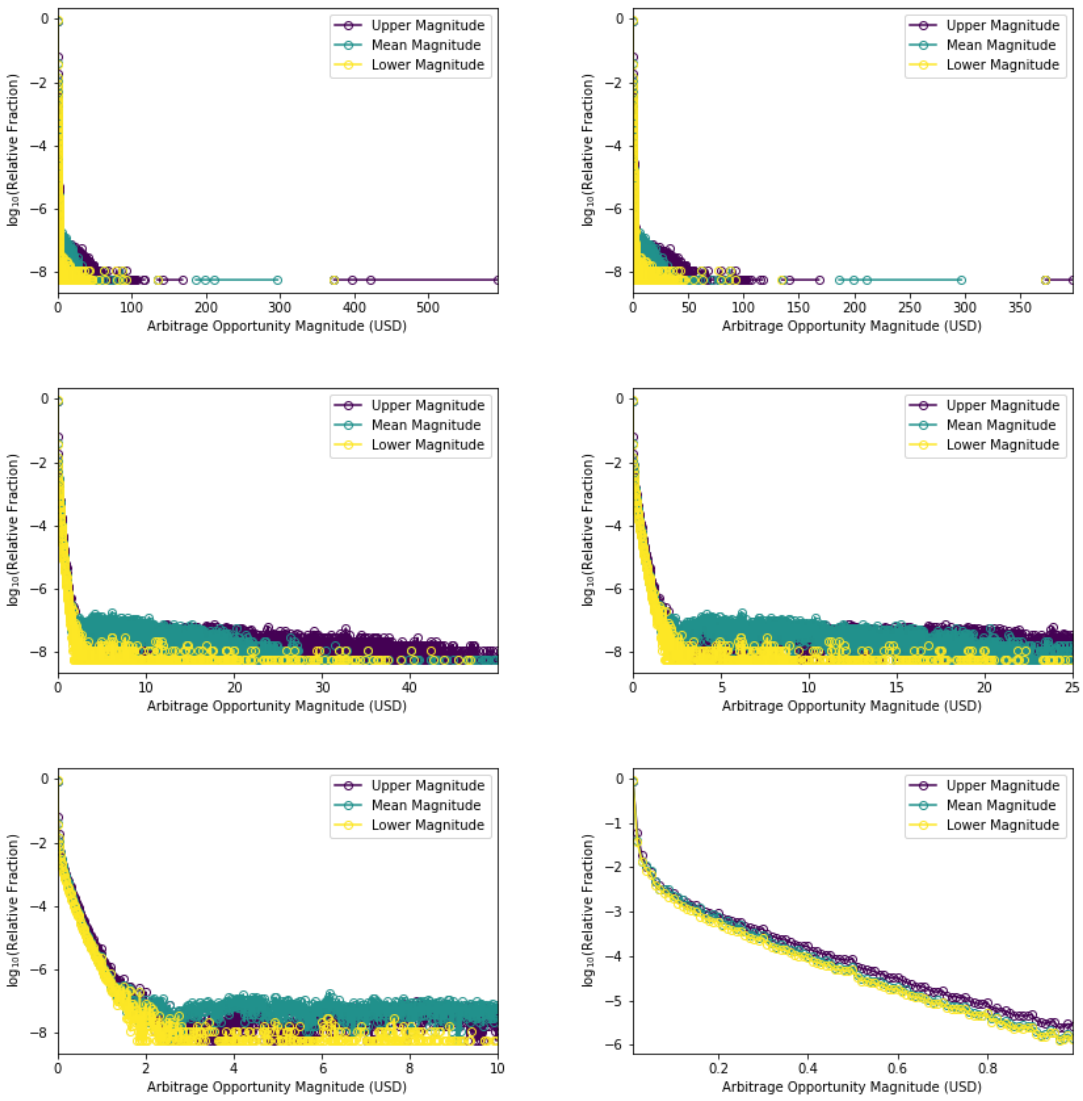

Figure 3.5: Several truncations of the empirical distribution of arbitrage opportunity magnitudes, allowing for closer inspection of the head of the distribution. From left to right, top to bottom: Magnitudes less than $1000, $400, $50, $25, $10, and $1.

Arbitrage Opportunity Magnitude Distribution Summary

		Combined	Dow	S&P 30
	count	178,476,945	120,355,462	130,374,761
Upper Estimate	mean	0.0192	0.0137	0.0207
	std	0.4250	0.5075	0.4705
	min	0.01	0.01	0.01
	25%	0.01	0.01	0.01
	50%	0.01	0.01	0.01
	75%	0.01	0.01	0.01
	max	4,905.69	4,905.69	4,905.69
Mean Estimate	mean	0.0172	0.0124	0.0185
	std	0.2181	0.2581	0.2406
	min	0.01	0.01	0.01
	25%	0.01	0.01	0.01
	50%	0.01	0.01	0.01
	75%	0.01	0.01	0.01
	max	2,452.85	2,452.85	2,452.85
Lower Estimate	mean	0.0152	0.0112	0.0164
	std	0.0548	0.0529	0.0572
	min	0.01	0.01	0.01
	25%	0.01	0.01	0.01
	50%	0.01	0.01	0.01
	75%	0.01	0.01	0.01
	max	372.69	372.69	372.69

Table 3.2: Summary statistics for the distribution of magnitudes of arbitrage opportunities (in USD) in the stocks of the Dow, the S&P 30, and their union. Three estimates of the magnitude distribution are provided: an upper estimate, which is calculated using the maximum magnitude of each observed arbitrage opportunity; a lower estimate, which is calculated using the minimum magnitude; and a mean estimate, which is the element-wise mean of the upper and lower estimates.

Actionable Arbitrage Opportunity Magnitude Distribution Summary

		Combined	Dow	S&P 30
Upper Estimate	count	88,119,495	63,617,169	62,392,695
	mean	0.0208	0.0152	0.0224
	std	0.5984	0.6927	0.6746
	min	0.01	0.01	0.01
	25%	0.01	0.01	0.01
	50%	0.01	0.01	0.01
	75%	0.01	0.01	0.01
	max	4,905.69	4,905.69	4,905.69
Mean Estimate	mean	0.0173	0.0131	0.0186
	std	0.3046	0.3513	0.3426
	min	0.01	0.01	0.01
	25%	0.01	0.01	0.01
	50%	0.01	0.01	0.01
	75%	0.01	0.01	0.01
	max	2,452.85	2,452.85	2,452.85
Lower Estimate	mean	0.0139	0.0109	0.0149
	std	0.0634	0.0660	0.0670
	min	0.01	0.01	0.01
	25%	0.01	0.01	0.01
	50%	0.01	0.01	0.01
	75%	0.01	0.01	0.01
	max	372.69	372.69	372.69

Table 3.3: Summary statistics for the distribution of magnitudes of arbitrage opportunities (in USD) with durations longer than 569 microseconds.

Daily ROC Statistics

	Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
mean	2,173,227.02	22,942,181,028.67	479,578.85	5,484,307,205.40	1,023,810.67	0.0100
std	664,180.20	6,065,524,000.26	197,378.15	1,851,246,298.07	759,800.91	0.0000
min	867,849.00	10,309,350,518.75	134,239.00	1,629,850,369.65	257,330.28	0.0098
25%	1,772,851.00	19,582,689,594.59	362,590.50	4,342,718,696.25	699,310.08	0.0100
50%	1,975,163.50	21,354,839,008.28	429,866.50	5,060,135,078.76	858,644.22	0.0100
75%	2,370,977.50	25,339,612,943.56	535,865.00	6,131,997,746.36	1,082,441.21	0.0100
max	5,015,045.00	50,626,240,421.16	1,443,208.00	13,811,423,929.99	8,652,008.60	0.0101

Table 3.4: Summary ROC statistics over the 252 trading days in 2016 for the union of stocks in the Dow and the S&P 30. Traded Value, Differing Traded Value, ROC, and ROC Per Share are all dollar values.

Aggregated ROC Statistics - 2016

	Combined	Dow	S&P 30
ROC / Traded Value	0.000045	0.000042	0.000044
Diff Trades / Trades	0.2207	0.2230	0.2133
Trades	547,653,210	392,101,579	463,330,929
Traded Value	5,781,429,619,225.01	3,858,963,034,003.48	4,968,476,428,927.45
Differing Trades	120,853,871	87,432,231	98,841,785
Differing Traded Val	1,382,045,415,759.77	900,535,924,962.71	1,150,097,817,041.18
ROC	258,000,289.46	160,213,922.95	220,025,380.71
ROC / Share	0.0100	0.0100	0.0100

Table 3.5: Direct comparison between the union of the Dow and the S&P 30, the Dow, and the S&P 30 using a variety of ROC statistics. Traded value, differing traded value, ROC, and ROC per share represent USD amounts, while the other statistics do not have an associated unit.

the union of the Dow and the S&P 30 on an average trading day experiences more than 2 million traded shares corresponding to nearly $23 billion changing hands. This trading activity is far from perfectly efficient, leading to an average ROC of $1 million and an average ROC per share of $0.01. In other words, the ROC statistic agrees with the analysis of arbitrage opportunity magnitudes above, indicating that the capture of almost any arbitrage opportunity will result in a profit of $0.01 per share.

Over the 252 trading days in 2016 for the union of stocks in the Dow and the S&P 30 more than 547 million trades occurred resulting in the movement of $\tilde{\$}5.8$ trillion worth of stocks. 22.07% of those 547 million trades were potentially inefficient leading to a ROC of approximately $260 million, or 0.000045% of the total traded value. See Table 3.5 for more details.

Since the ROC is estimated using trades that actually occurred, it is possible that it underestimates the amount of market inefficiency since arbitragers may be imperfect, while it is also possible that it overestimates the amount of market inefficiency since it does not account for the time a mar-

Estimates of Total Available Arbitrage Profits

	Estimate	$\mathbb{E}$(Value/Share)	Arb Opps	Total Arb Profits
Combined	Upper Mag.	0.0208	88,119,495	183,288,549.60
	Mean Mag.	0.0173		152,446,726.35
	Lower Mag.	0.0139		122,486,098.05
	Minimum	0.0100		88,119,495.00
S&P 30	Upper Mag.	0.0224	62,392,695	139,759,636.80
	Mean Mag.	0.0186		116,050,412.70
	Lower Mag.	0.0149		92,965,115.55
	Minimum	0.0100		62,392,695.00
Dow	Upper Mag.	0.0152	63,617,169	96,698,096.88
	Mean Mag.	0.0131		83,338,491.39
	Lower Mag.	0.0109		69,342,714.21
	Minimum	0.0100		63,617,169.00

Table 3.6: Estimates of the total arbitrage profits (in USD) available to a perfect arbitrager of the union of the Dow and the S&P 30, the S&P 30, and the Dow in 2016. Estimates are based upon the expected value of the indicated magnitude distribution (in USD), the number of observed actionable (duration > 569 microseconds) arbitrage opportunities in that group in 2016, and the knowledge that the minimum size for a protected quotation is 100 shares.

ket participant would need in order to profit from the inefficiencies it attempts to describe. With this in mind, an estimate of the total amount of money available to a perfect arbitrager may be constructed using the frequency statistics for arbitrage opportunities in combination with an expected arbitrage opportunity magnitude. Using the mean of the upper and lower magnitude estimates for actionable arbitrage opportunities, the total frequency of actionable arbitrage opportunities, and the knowledge that protected quotations have a minimum size of 100 shares, indicates that between \$69.3 million and \$96.7 million were available to a perfect arbitrager of the Dow in 2016. \$93 million to \$139.8 million was available to a perfect arbitrager of the S&P 30, and \$122.5 million to \$183.3 million was available for a perfect arbitrager of the stocks in the union of the Dow and the S&P 30. An even more conservative estimate made using \$0.01 as the expected value of an arbitrage opportunity indicates that nearly \$63.6 million in arbitrage profits were available from the Dow, \$62.4 million from the S&P 30, and \$88.1 million from their union[1].

[1]See Table 3.6 for more information about ROC estimates using properties of arbitrage opportunities.

Comparison of the Dow and the S&P 30

The previous section investigated statistics and distributions derived from a sample-wide view of the 42 stocks under consideration. Once a baseline for the global behavior of these statistics and distributions has been determined, the natural next step is to investigate the behavior of these statistics and distributions in smaller groups of stocks or even individual stocks. The relative comparison of these statistics and distributions between groups of stocks, as well as differences between the sample population and smaller groups of stocks, may provide a deeper understanding of arbitrage opportunities and their causes.

This section compares the Dow and the S&P 30, as defined in Section 1.6. These indices both have 30 constituents, all of which are associated with large U.S. companies, and there is a fair amount of overlap between the two. Of the 42 stocks present in either group: 18 are members of both the Dow and the S&P 30, AAPL, CVX, DIS, GE, HD, IBM, INTC, JNJ, JPM, KO, MRK, MSFT, PFE, PG, V, VZ, WMT, XOM; 12 are members of the Dow and not the S&P 30, AXP, BA, CAT, CSCO, DD, GS, MCD, MMM, NKE, TRV, UNH, UTX; and 12 are members of the S&P 30 and not the Dow, GOOG, GOOGL, AMZN, T, BAC, BRK.B, CMCSA, FB, ORCL, PEP, WFC, PM. The number of shared stocks between these groups and the mechanistic nature of arbitrage opportunities may indicate that the statistics and distributions derived from arbitrage opportunities and ROC should be extremely similar, if not identical.

However, there are significant differences between these statistics and distributions when comparing the Dow and the S&P 30.

Differences in Arbitrage Opportunity Distributions

Following the same order of investigation as Section 3.1, the first comparison made involves the distribution of arbitrage opportunity durations in the Dow and the S&P 30 respectively.

Table 3.1, located in Section 3.1, shows that the mean and median of the duration distributions are different. A two-sample Kolmogorov-Smirnov test (KS test) and a two sample Mann-Whitney U-Test (MWU test) are used to show that the difference between the distributions is significant and

Statistical Comparison of The Dow and The S&P 30

		Statistic	p-Value
Duration	2 Sample KS Test	0.06035	$< 10^{-15}$
	2 Sample MWU Test	8086163730559262	$< 10^{-15}$
	JSD	0.02841	-
Upper Magnitude Estimate	2 Sample KS Test	0.05376	$< 10^{-15}$
	2 Sample MWU Test	7469278737700730	$< 10^{-15}$
	JSD	0.01754	-
Mean Magnitude Estimate	2 Sample KS Test	0.05486	$< 10^{-15}$
	2 Sample MWU Test	7468070356999940	$< 10^{-15}$
	JSD	0.01761	-
Lower Magnitude Estimate	2 Sample KS Test	0.04568	$< 10^{-15}$
	2 Sample MWU Test	7482113302991260	$< 10^{-15}$
	JSD	0.01353	-

Table 3.7: Statistical tests comparing the distribution of arbitrage opportunity attributes between the Dow and the S&P 30. The Jensen-Shannon Divergence is a distance measure, not a statistical test, so it has no associated p-value.

the Jensen-Shannon Divergence (JSD) is used to quantify that difference. The KS test and MWU test are non-parametric tests and may be more appropriate than parametric tests in this case, since the distributions under investigation may not follow common assumptions made by parametric tests. The JSD is an information theoretic measure of the difference between two distributions, and has several convenient properties. Namely, it is symmetric with respect to its inputs (i.e., JSD(X, Y) = JSD(Y, X)) and its value is bounded (it assumes values in the range [0, 1] when computing the difference between a pair of distributions and the base-2 logarithm is used).

When comparing the distributions of arbitrage opportunity durations between the Dow and the S&P 30, the KS test returned a statistic of 0.05953 ($p < 10^{-15}$), while the MWU test returned a statistic of 8086163730559262 ($p < 10^{-15}$), and the JSD between the two distributions measured 0.02841. These results indicate that the two distributions are significantly different, though the magnitude of that difference is relatively small.

Similarly, the mean and median of the distribution of arbitrage opportunity magnitudes differs between the Dow and the S&P 30 regardless of which magnitude estimate is used. Using the same statistical tools as above indicates these distributions are also significantly different, though the differences are again relatively small. See Table 3.7 for more details.

Daily ROC Statistics - The Dow

	Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
mean	1,555,958.65	15,313,345,373.03	346,953.30	3,573,555,257.78	635,769.54	0.0100
std	463,558.93	3,891,299,900.31	146,677.85	1,234,882,079.43	655,911.15	0.0000
min	579,206.00	6,664,671,053.15	89,564.00	1,035,855,029.71	145,205.65	0.0098
25%	1,278,813.25	12,915,031,172.08	262,209.00	2,804,569,367.64	417,485.73	0.0100
50%	1,429,062.00	14,431,597,662.02	309,158.00	3,274,390,601.60	514,856.64	0.0100
75%	1,715,351.25	16,829,521,684.38	387,772.00	3,993,470,514.97	666,268.27	0.0100
max	3,596,006.00	30,999,914,293.66	1,073,029.00	9,428,952,387.10	7,817,684.58	0.0101

Table 3.8: Summary ROC statistics over the 252 trading days in 2016 for stocks in the Dow. Summary statistics are computed over daily aggregates of ROC statistics. Traded Value, Differing Traded Value, ROC, and ROC Per Share are all dollar values.

Daily ROC Statistics - The S&P 30

	Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
mean	1,838,614.80	19,716,176,305.27	392,229.31	4,563,880,226.35	873,116.59	0.0099
std	573,638.98	5,352,699,735.53	165,402.69	1,585,661,474.76	700,878.32	0.0001
min	738,844.00	8,775,273,579.14	112,238.00	1,322,682,479.53	218,378.57	0.0097
25%	1,494,204.00	16,716,749,687.31	292,734.50	3,564,986,358.00	590,618.50	0.0099
50%	1,672,011.50	18,310,212,120.94	348,020.50	4,173,715,014.57	727,405.38	0.0100
75%	2,006,113.25	21,724,903,514.15	445,005.75	5,184,500,948.62	929,452.58	0.0100
max	4,242,910.00	45,174,822,307.48	1,197,001.00	11,532,672,292.09	8,385,367.33	0.0100

Table 3.9: Summary ROC statistics over the 252 trading days in 2016 for stocks in the S&P 30. Summary statistics are computed over daily aggregates of ROC statistics. Traded Value, Differing Traded Value, ROC, and ROC Per Share are all dollar values.

Differences in ROC Statistics

Table 3.8 and Table 3.9 summarize the distribution of the ROC and several related statistics when aggregated on a daily basis. From these two tables several qualitative observations may be made about the relative efficiency of the Dow and the S&P 30.

The S&P 30 features a greater mean and median for every statistic except for ROC per share when compared with the Dow. This indicates that the stocks of the S&P 30 experience a greater amount of trading volume on average and tend to also experience greater amounts of aggregated market inefficiency, as measured by the ROC statistic. However, the lower ROC per share statistic featured by the S&P 30 indicates that the stocks of the Dow may experience market inefficiencies with slightly larger magnitudes on average. In other words, when prices discrepancies exist between the SIP and direct feeds stocks in the S&P 30 may experience slightly smaller price discrepancies than stocks in the Dow.

Chapter 4

Conclusion

Discussion

From the information presented in this study, several useful heuristics involving arbitrage opportunities may be formed. First, nearly half of all arbitrage opportunities in the sample lasted long enough to be considered actionable under a relatively conservative definition. Second, the vast majority of arbitrage opportunities have a magnitude of $0.01, which is the minimum tick size. However, it is important to note that the tail of the magnitude distribution is long and extreme values occur more frequently than some previous work suggests. Third, arbitrage opportunities are most likely to occur at the start and end of a trading day.

By investigating the relationships between these distributions, a few more heuristics may be formed about the stocks investigated in this sample. Namely, arbitrage opportunities with larger magnitudes tend to occur earlier in the trading day (Kendall's Tau correlation -0.04357, p-value=$< 10^{-15}$), arbitrage opportunities which start later in the trading day tend to last longer (Kendall's Tau correlation 0.05956, p-value=$< 10^{-15}$), and arbitrage opportunities with larger magnitudes tend to last longer (Kendall's Tau correlation 0.03927, p-value=$< 10^{-15}$).

The combination of the ROC statistic and estimates from properties of arbitrage opportunities place the total arbitrage profits available in the 42 stock sample on the order of $258 million for 2016. This number may seem small in relation to the traded volume of those stocks, however it

can be argued that $258 million is extremely large in comparison to the $0 in arbitrage profits that would be available if markets always satisfied the EMH. The existence, quantity, and persistence of arbitrage opportunities as discussed in this study indicates that there are timescales and/or market conditions where the EMH does not hold. The Adaptive Markets Hypothesis is more permissive and allows for the existence of arbitrage opportunities, subject to various evolving market conditions which effect their prevalence [18]. However, even if one assumes that arbitrage opportunities must exist, should their current prevalence be considered acceptable?

Comparison With Previous Results

Bartlett and McCrary claim in [10] that only $11 million were lost to information asymmetries between the SIP and direct feeds when considering the Dow from August 2015 to June 2016, which is approximately an order of magnitude less than the $78 to $170 million estimated in this study. This may be due to differences in the quality of data used, since [10] did not have access to direct feed data and instead attempted to synthesize direct feed-like data using exchange timestamps.

Ding, Hanna, and Hendershott estimate in [19] that $32,000 in arbitrage profits were available in AAPL daily in 2014, which closely matches the daily mean/median of $45,000 shown in Table C.2.

Additionally, Wah found that more than $3 billion in latency arbitrage profits were available in S&P 500 stocks in 2014 [20]. Naively extrapolating the $258 million in arbitrage profits for the 42 stocks considered in this study leads to an estimated profit of $3.1 billion for the S&P 500 in 2016. This extrapolation is the result of scaling up the arbitrage profits by factor proportional to the number of stocks under consideration, i.e. $\frac{500}{42} = 11.9 \approx 12$.

Overall, the estimates of arbitrage opportunity profits calculated in this study agree with previous work in rough scale, strengthening their validity.

Limitations

The working definition of actionable arbitrage opportunities used in this analysis rests upon the estimate found in [17]. However, what is actually actionable may vary depending on the location of a market participant, the existence of faster communication lines, and more. Additionally, the usage of a conditional duration cutoff based upon the location of a market participant and the exchanges required to capture each arbitrage opportunity should provide a more accurate estimate of the truly actionable arbitrage opportunities. Though, adding this level of detail to the calculation of arbitrage opportunities would significantly increase the amount of time and resources required to perform an analysis similar to what is presented in this thesis.

The ROC statistic also has room to grow, since its definition and calculation involve no duration component. More specifically, the ROC statistic would be a more realistic measurement of market inefficiency if it considered whether trades marked as differing trades could reasonably obtain price improvement.

Future Work

This thesis studies arbitrage opportunities in two small groups of stocks over the 252 trading days in 2016, leaving two paths for direct extension. A cross-sectional study that encompasses a greater portion of stocks and other exchange traded instruments would provide a wider view of how market inefficiencies vary based upon the particular instrument under consideration. A longitudinal study that investigates the development of arbitrage opportunities over time may provide insight into evolutionary characteristics of the NMS.

In addition to direct extensions, there are also several related veins of research that may follow from this thesis. One such example is the empirical study of order flow, which could lead to improvements in the understanding of market dynamics and price formation.

Conclusion

Geographic market fragmentation of the NMS has lead to the mechanical creation of market inefficiencies due to information asymmetries between data feeds that necessarily exist due to theoretical limits on information transmission speed. These market inefficiencies are prolific, persistent, and endure long enough to be reliably captured by well equipped market participants, which contradicts portions of the Efficient-Market Hypothesis.

Up to $258 million dollars in arbitrage profits were available between the stocks of the Dow and the S&P 30 in 2016, indicating that the market inefficiencies introduced by geographic market fragmentation may be worse than previously believed. Beyond their immediate impacts, these arbitrage opportunities are potentially more concerning due to the negative impacts they could have on the future of the NMS via a positively reinforcing reduction in market quality as described in [21].

Bibliography

[1] Eugene F. Fama. "Efficient capital markets: A review of theory and empirical work". In: *The journal of Finance* 25.2 (1970), pp. 383–417.

[2] Eugene F. Fama and Kenneth R. French. "Dividend yields and expected stock returns". In: *Journal of financial economics* 22.1 (1988), pp. 3–25.

[3] Securities Exchange Commission. *What We Do*. 2013. URL: https://www.sec.gov/Article/whatwedo.html (visited on 03/29/2018).

[4] U.S. Securities and Exchange Commission. *Regulation NMS*. 2005. URL: https://www.sec.gov/rules/final/34-51808.pdf (visited on 02/28/2018).

[5] Anova Technologies. *Network Map*. 2018. URL: https://www.google.com/maps/d/u/0/viewer?mid=1YCUsndp-vcPxF2a5jSwIYkFHsIw&ll=40.79717387569744%2C-73.95781079073856&z=10l (visited on 03/29/2018).

[6] U.S. Securities and Exchange Commission. *Regulation of Exchanges and Alternative Trading Systems*. 2000. URL: https://www.sec.gov/rules/proposed/34-39884.pdf (visited on 03/01/2018).

[7] Laura Tuttle. *Alternative Trading Systems: Description of ATS Trading in National Market System Stocks*. 2014. URL: https://www.sec.gov/marketstructure/research/alternative-trading-systems-march-2014.pdf (visited on 03/01/2018).

[8] U.S. Securities and Exchange Commission. *Alternative Trading System List*. 2018. URL: https://www.sec.gov/foia/docs/atslist.htm (visited on 03/01/2018).

[9] Thesys Technologies. *Home*. 2018. URL: https://www.thesystech.com/ (visited on 04/05/2018).

[10] Robert P. Bartlett III and Justin McCrary. *How Rigged Are Stock Markets?: Evidence From Microsecond Timestamps*. Tech. rep. National Bureau of Economic Research, 2016.

[11] Federal Business Opportunities. *MIDAS Contract Award Notice*. 2012. URL: https://www.fbo.gov/index?s=opportunity&mode=form&id=80f5ee86812a5a690446b69ccc86d766&tab=core&tabmode=list&= (visited on 02/28/2018).

[12] Eric Budish, Peter Cramton, and John Shim. "The high-frequency trading arms race: Frequent batch auctions as a market design response". In: *The Quarterly Journal of Economics* 130.4 (2015), pp. 1547–1621.

[13] Eric Budish, Peter Cramton, and John Shim. "Implementation details for frequent batch auctions: Slowing down markets to the blink of an eye". In: *American Economic Review* 104.5 (2014), pp. 418–24.

[14] Phil Mackintosh. *Alternative Trading System List*. 2014. URL: https://web.archive.org/web/20160608163426/https://www.kcg.com/uploads/documents/KCG_Demystifying-Order-Types_092414.pdf (visited on 03/02/2018).

[15] S&P Dow Jones Indices LLC. *Dow Jones Industrial Average*. 2018. URL: https://us.spindices.com/indices/equity/dow-jones-industrial-average (visited on 02/27/2018).

[16] S&P Dow Jones Indices LLC. *S&P 500*. 2018. URL: https://us.spindices.com/indices/equity/sp-500 (visited on 02/27/2018).

[17] Brian F. Tivnan et al. "Inefficiencies in the U.S. National Market System: Evidence from the Dow 30". 2018. In progress.

[18] Andrew W. Lo. "The adaptive markets hypothesis: Market efficiency from an evolutionary perspective". In: (2004).

[19] Shengwei Ding, John Hanna, and Terrence Hendershott. "How slow is the NBBO? A comparison with direct exchange feeds". In: *Financial Review* 49.2 (2014), pp. 313–332.

[20] Elaine Wah and Michael P. Wellman. "Latency arbitrage, market fragmentation, and efficiency: a two-market model". In: *Proceedings of the fourteenth ACM conference on Electronic commerce*. ACM. 2013, pp. 855–872.

[21] George A. Akerlof. "The market for "lemons": Quality uncertainty and the market mechanism". In: Elsevier, 1978, pp. 235–251.

Appendix A

Glossary

Market Architecture

Definition A.1.1 (*Market System*)**.** A market system, $m = (V, L)$, may be defined as a network or graph which consists of a set of one or more market centers (V) connected by a set of communication channels (L).

Definition A.1.2 (*Market Center*)**.** A market center is a location, physical or digital, where agents may interact with a market system. A market center, $c = (B, A, I)$, may be defined as a tuple containing a local order book, B, a set of valid actions, A, and a set of traded financial instruments, I.

Definition A.1.3 (*Local Order Book*)**.** The local order book contains information about the unfulfilled orders that have been submitted to a market center, allowing it to accumulate and maintain state. The NMS operates under a Continuous Double Auction (CDA), where the local order book for a single financial instrument can be thought of as two ordered lists of queues where each queue is associated with a price and each list contains queues of bids or offers respectively. Other market mechanisms are also valid, such as Frequent Batch Auctions (FBA).

Definition A.1.4 (*Action Set*)**.** The action set defines the valid actions at a market center. No requirements are imposed on the action set, though a simple real world action set might allow for the submission of limit orders (which guarantee price), market orders (which guarantee execution),

modification of resting orders, and cancellation of resting orders. (i.e. $A = \{$limit order, market order, modify, cancel$\}$)

Definition A.1.5 (*System Activity*)**.** Let the system activity, $\mathbb{A}$, be a chronological list of all actions that are performed in a market system. This includes actions performed by market participants, administrative messages transmitted by regulators, and messages transmitted by the exchange(s).

Definition A.1.6 (*Data Feed*)**.** A data feed, D, is defined to be any subset of of the system activity of a market system (i.e. $D \subseteq \mathbb{A}$). Note that recorded occurrence times of identical events may vary between distinct data feeds due to physical considerations (e.g., speed of light, relativity).

Financial Instruments

Definition A.2.1 (*Security*)**.** A security is a financial instrument that represents partial or total ownership of an object or entity. Securities are fungible; securities belonging to the same "class" have the same value, and therefore are interchangeable. Additionally, the exact value of a security is negotiable. Common varieties of securities include stocks, bonds, and options, all of which may be traded on electronic markets, such as the NMS.

Definition A.2.2 (*Stock*)**.** Stocks, which are sometimes referred to as equities, are a variety of security that represents partial ownership of a publicly traded company. Stocks are a vehicle by which companies can acquire the capital necessary to grow and the secondary market for stocks is the basis of a large portion of the U.S. financial industry.

The Best Bid/Offer

The following definitions assume the existence of a market system, $m = (V, L)$, where each $v \in V$ has an action set, A, where $\{$limit order$\} \subseteq A$, and at least one data feed, D, which contains information about the top of the book at each market venue (i.e., a comprehensive quote feed).

Definition A.3.1 (*Local Best Bid/Offer*)**.** The local best bid for a financial instrument, i, at market venue, v, at a time, t, is given by the tuple (p, q), where p is the maximum price among all

active bids for i in v (as observed via data feed D) and q is the quantity of shares of i currently available at that price at v (i.e. $LBB(D, v, i, t) = (p, q)$). The local best offer is defined similarly, but uses the minimum price among offers on v along with the number of shares at that price (i.e. $LBO(D, v, i, t) = (p', q')$).

Definition A.3.2 (*Best Bid/Offer*)**.** The best bid is similar to the local best bid, but is formed by the maximum price (and quantity available at that price) among resting bids for i among all market venues $v \in V$ (i.e. $BB(D, i, t) = (p'', q'')$).

Similarly, the best offer is formed by the minimum price among resting offers and the number of shares at that price (i.e. $BO(D, i, t) = (p''', q''')$)

Dislocations

The following definitions assume the existence of a market system, $m = (V, L)$, containing two or more market centers, two data feeds, D_1 and D_2, and a financial instrument i that is traded at all $v \in V$.

Note that these definitions are phrased for the best bid, but apply similarly to the best offer.

Definition A.4.1 (*Dislocated Data Feeds*)**.** D_1 and D_2 are said to be dislocated with respect to the best bid of i at a time t if the two feeds do not feature an identical best bid price (i.e. $BB(D_1, i, t).price \neq BB(D_2, i, t).price$).

Definition A.4.2 (*Dislocation*)**.** A dislocation is said to occur between D_1 and D_2 whenever they are *dislocated* with respect to the best bid of i over a half-open interval of time $[a, b)$.

Definition A.4.3. $\Delta BB(i, t) = BB(D_1, i, t).price - BB(D_2, i, t).price$

Definition A.4.4 (*Arbitrage Opportunity*)**.** An arbitrage opportunity with respect to the best bid of i is defined to be any half-open interval of time, $[a, b)$, where D_1 and D_2 are dislocated with respect to the best bid of i and $\text{sgn}(\Delta BB(i, t)) = \text{sgn}(\Delta BB(i, a))\ \forall t \in [a, b)$.

Definition A.4.5 (*Direction*)**.** The direction of an arbitrage opportunity over an interval $[a, b)$ is defined as $\text{sgn}(\Delta BB(g, a))$.

Definition A.4.6 (*Duration*)**.** The duration of a dislocation or arbitrage opportunity over an interval $[a, b)$ is defined as $b - a$.

Definition A.4.7 (*Magnitude*)**.** The magnitude of a dislocation or arbitrage opportunity over an interval $[a, b)$ is defined as $\max_{t \in [a,b)}\{|\Delta BB(g, t)|\}$

Market Actions

The following definitions provide a high-level description of the purpose and details of some common order types, but are not necessarily representative implementations at NMS market centers.

Definition A.5.1 (*Limit Order*)**.** Guarantees market participants an execution price no worse than a provided limit price, but does not provide any guarantees about the timeliness of execution. This may be implemented by placing a received limit order into the price queue associated with the provided limit price on the correct side of the book (bid or ask, as specified by the order), assuming that it did not match with a resting order at a better price.
Fields: Instrument identifier, bid/ask, limit price, desired quantity.

Definition A.5.2 (*Market order*)**.** Guarantees instant execution on a best effort basis, but does not provide any guarantees about the execution price. This may be implemented by matching the market order with the best resting orders on the opposite side of the book until the desired quantity is obtained. A market order may be thought of as a limit order with the limit price set in order to guarantee execution (i.e. 0 for a market ask or infinity for a market bid).
Fields: Instrument identifier, bid/ask, desired quantity

Definition A.5.3 (*Modify*)**.** Allows market participants to update values associated with resting orders and allows for adaptation to changing market conditions. The main usage of this order is to change the number of shares required to fulfill a particular order, since modifying the limit price of order may cause it to lose its place in its current price queue.
Fields: Order identifier, field(s) to modify, new value(s)

Definition A.5.4 (*Cancel*)**.** Allows market participants to remove resting orders from the local book prior to execution.

Fields: Order identifier

Definition A.5.5 (*Immediate Or Cancel*)**.** Often shortened to IOC, this is a modifier which may be applied to any order rather than a stand alone order type. The modifier indicates that the associated order should be executed immediately upon receipt or canceled if immediate execution is not possible.

Appendix B

Supplementary Figures

Distribution of Arbitrage Opportunity Durations in Dow Stocks

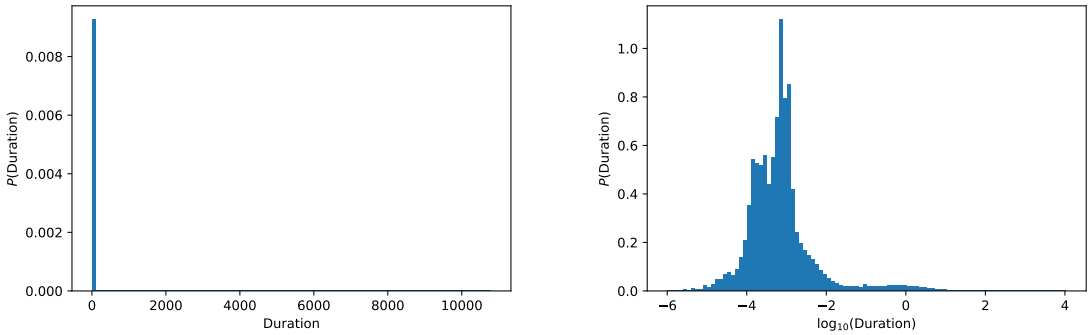

Figure B.1: Left: Empirical distribution of durations (in seconds) for arbitrage opportunities in Dow stocks, plotted with linear axes. Right: Identical distribution plotted with a logged x-axis, arbitrage opportunities with a duration of 0 were excluded. Any observed "0" duration dislocations are the result of timestamps with microsecond resolution and events occurring at sub-microsecond timescales.

Distribution of Arbitrage Opportunity Durations in S&P 30 Stocks

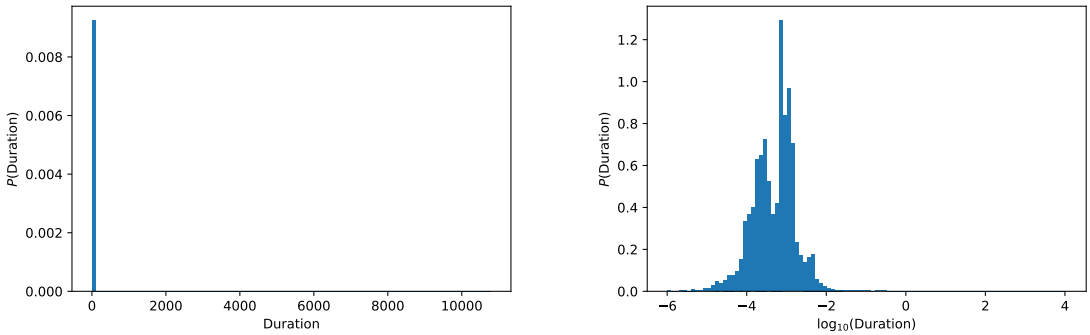

Figure B.2: Left: Empirical distribution of durations (in seconds) for arbitrage opportunities in S&P 30 stocks, plotted with linear axes. Right: Identical distribution plotted with a logged x-axis, arbitrage opportunities with a duration of 0 were excluded. Any observed "0" duration dislocations are the result of timestamps with microsecond resolution and events occurring at sub-microsecond timescales.

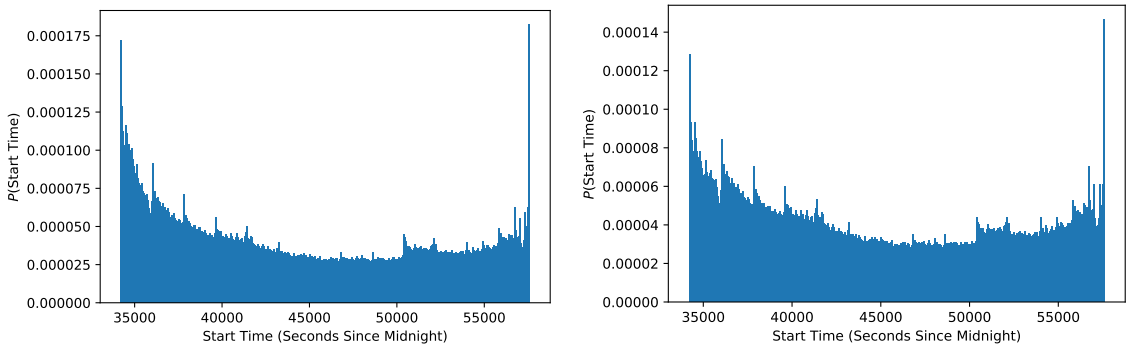

Figure B.3: Left: Empirical distribution of start times for arbitrage opportunities in Dow stocks, plotted with linear axes. Right: Conditional empirical distribution of start times, where only arbitrage opportunities with a duration longer than 569 microseconds are considered.

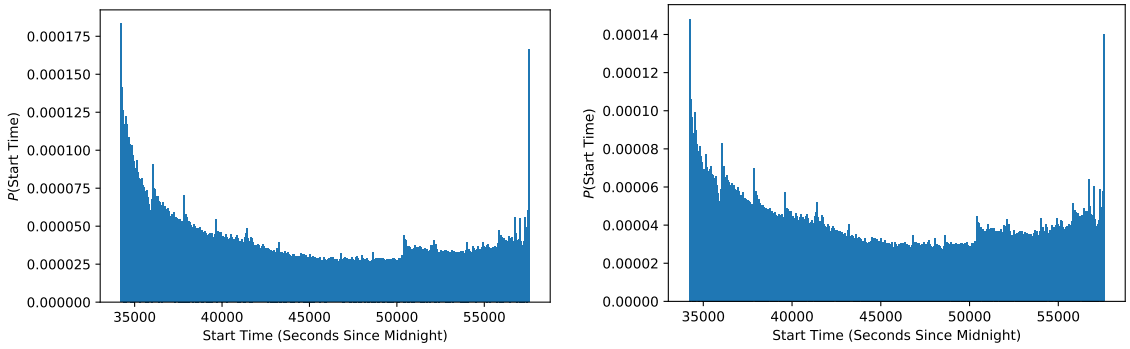

Figure B.4: Left: Empirical distribution of start times (in seconds) for arbitrage opportunities in S&P 30 stocks, plotted with linear axes. Right: Conditional empirical distribution of start times, where only arbitrage opportunities with a duration longer than 569 microseconds are considered.

Truncated Distributions of Arbitrage Opportunity Magnitudes in Dow Stocks

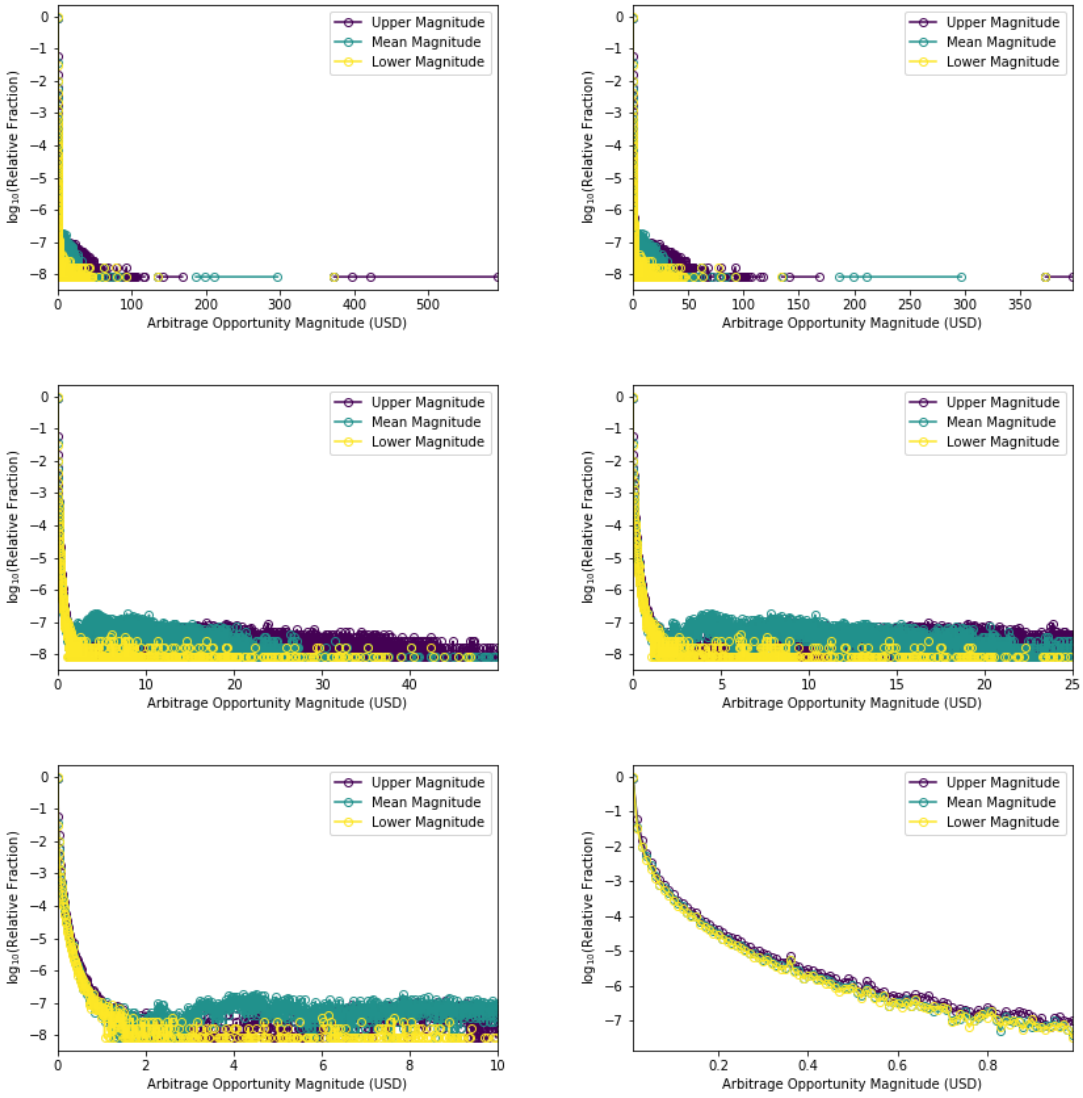

Figure B.5: Several truncations of the empirical distribution of arbitrage opportunity magnitudes in Dow stocks, allowing for closer inspection of the head of the distribution. Plotted on semi-log axes. From left to right, top to bottom: Magnitudes less than $1000, $400, $50, $25, $10, and $1.

Truncated Distributions of Arbitrage Opportunity Magnitudes in S&P 30 Stocks

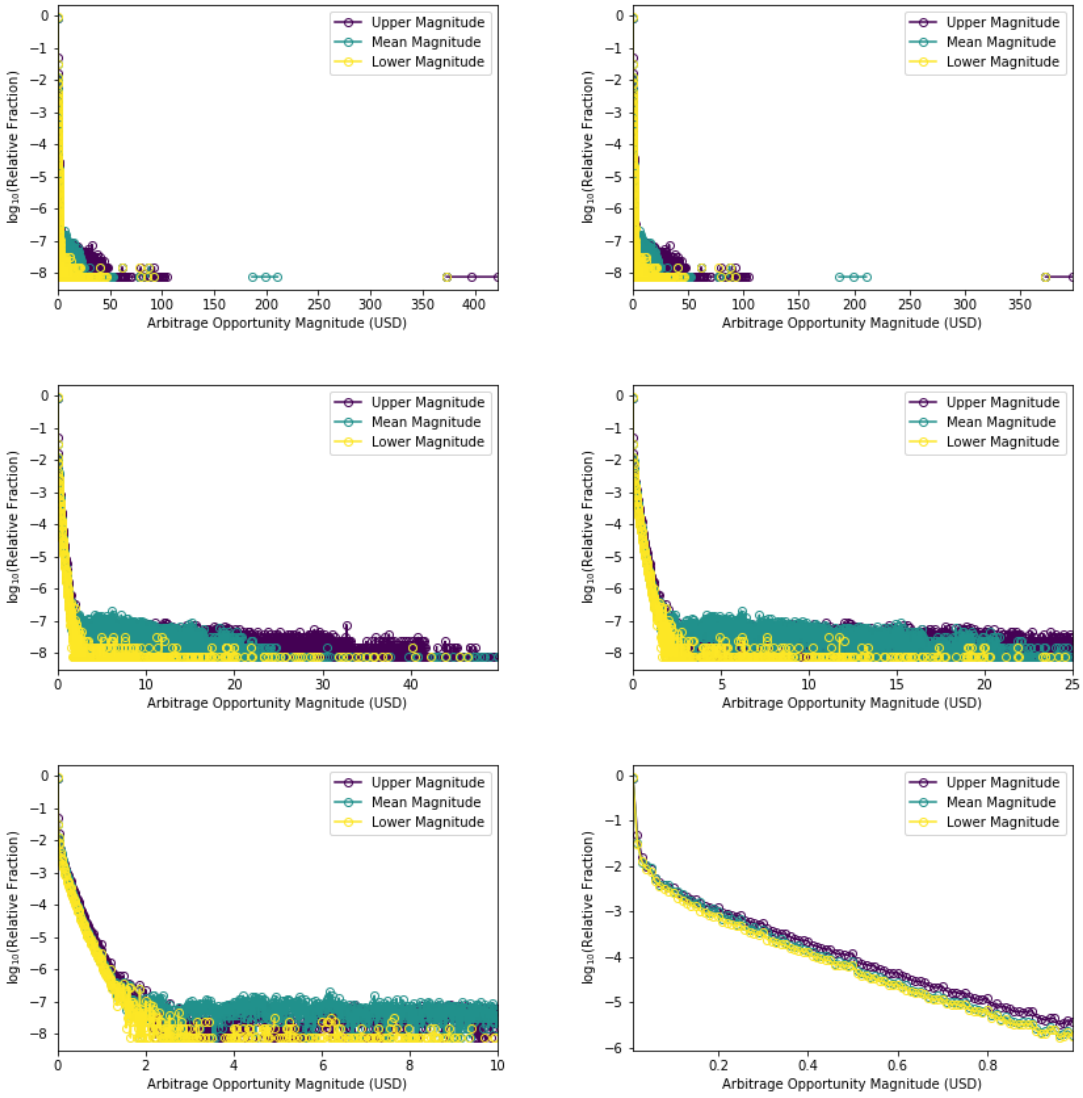

Figure B.6: Several truncations of the empirical distribution of arbitrage opportunity magnitudes in S&P 30 stocks, allowing for closer inspection of the head of the distribution. Plotted on semi-log axes. From left to right, top to bottom: Magnitudes less than $1000, $400, $50, $25, $10, and $1.

Appendix C

Supplementary Tables

Stock Information and Rankings

Symbol	Share Price	Market Cap.	Market Cap. Rank	Share Price Rank
AAPL	108.3650	588,674,250,000	1	15
AMZN	724.1100	342,605,250,000	7	3
AXP	65.0700	60,897,056,768	39	29
BA	136.0575	85,909,215,232	37	9
BAC	16.1350	164,734,500,000	21	42
BRK.B	148.5300	366,148,750,000	6	7
CAT	83.4650	48,739,393,536	41	23
CMCSA	65.4150	158,159,250,000	27	28
CSCO	29.7750	149,579,500,000	31	41
CVX	105.2125	198,394,250,000	18	17
DD	67.1225	58,571,190,272	40	27
DIS	98.5525	158,958,750,000	26	19
FB	117.9250	338,086,000,000	8	13
GE	31.1225	282,407,750,000	10	40
GOOG	746.5400	519,468,000,000	3	2
GOOGL	765.7350	519,468,000,000	2	1
GS	176.5700	75,577,461,760	38	4
HD	130.9700	162,088,250,000	24	10
IBM	157.0175	150,224,000,000	30	6
INTC	34.7925	164,545,500,000	22	38
JNJ	115.7100	317,181,750,000	9	14
JPM	68.5600	248,467,000,000	12	26
KO	43.8750	189,596,750,000	19	35
MCD	120.7750	104,620,209,360	34	11
MMM	174.1375	105,269,000,000	33	5
MRK	57.9500	160,298,750,000	25	30
MSFT	56.5350	442,765,000,000	4	31
NKE	55.0375	92,671,447,504	35	32
ORCL	39.8925	164,379,750,000	23	37
PEP	105.4550	152,207,500,000	28	16
PFE	32.8000	199,846,250,000	17	39
PG	85.2025	228,141,500,000	14	22
PM	97.1350	150,685,250,000	29	20
T	41.3800	254,488,500,000	11	36
TRV	118.1800	34,257,640,448	42	12
UNH	142.5350	135,615,000,000	32	8
UTX	103.4675	86,217,687,040	36	18
V	77.8425	184,334,000,000	20	24
VZ	53.8200	219,440,750,000	16	33
WFC	48.7700	246,265,250,000	13	34
WMT	70.6875	219,609,750,000	15	25
XOM	88.7175	368,008,250,000	5	21

Table C.1: Share price (in USD) and market capitalization (in USD) for the stocks under study, computed using the average of four observations taken at the end of each quarter in 2016, as well as rankings of the stocks according to their share price and market capitalization. Since the S&P 30 is a market cap. weighted index, the Market Cap. Rank indicates the relative effect each stock would have on the S&P 30. The Share Price Rank has a similar interpretation for the Dow, since it is a share price weighted index.

Daily ROC Statistics - 2016 - Itemized By Stock

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
AAPL	mean	174,820.85	2,542,188,952.00	34,316.58	483,265,898.89	45,852.81	0.0090
	std	68,897.09	1,040,923,482.82	20,556.77	280,321,422.73	27,275.35	0.0015
	min	54,824.00	983,856,430.54	2,112.00	35,009,317.38	2,773.35	0.0033
	25%	129,830.00	1,872,512,861.35	23,498.75	340,459,129.85	32,088.96	0.0094
	50%	156,198.50	2,272,037,106.11	32,741.00	452,246,993.35	43,204.10	0.0097
	75%	199,793.25	2,870,019,105.28	42,674.75	599,146,631.96	57,647.05	0.0098
	max	517,270.00	8,280,915,338.59	103,885.00	1,596,912,962.05	138,331.08	0.0100
AMZN	mean	15,247.79	792,584,183.03	5,323.36	295,465,467.21	36,209.44	0.0617
	std	7,332.26	335,570,554.73	2,779.11	131,110,532.55	27,276.77	0.0256
	min	5,896.00	318,618,478.15	936.00	54,345,232.16	2,925.31	0.0058
	25%	10,367.50	552,563,427.87	3,569.25	207,214,150.84	19,174.63	0.0501
	50%	13,622.50	709,879,550.69	4,842.50	286,035,318.63	28,451.90	0.0622
	75%	18,410.50	950,115,228.13	6,699.75	370,801,518.10	45,197.83	0.0762
	max	54,111.00	2,526,198,199.88	16,193.00	794,727,936.33	150,997.03	0.1264
AXP	mean	32,348.46	250,614,304.97	9,086.69	71,464,081.61	11,622.14	0.0100
	std	16,110.77	143,031,721.64	4,434.64	36,283,858.01	7,156.73	0.0002
	min	11,095.00	90,438,986.65	2,219.00	19,241,382.52	2,666.91	0.0089
	25%	22,756.50	168,209,590.34	5,999.50	49,149,197.52	7,672.38	0.0099
	50%	26,835.00	207,178,850.49	7,476.00	57,481,058.84	8,987.37	0.0100
	75%	37,067.75	277,456,051.09	10,905.75	86,485,488.61	13,792.36	0.0101
	max	159,135.00	1,468,245,304.80	31,507.00	302,294,385.78	75,473.73	0.0114
BA	mean	20,749.26	288,851,358.58	7,071.25	100,312,506.47	10,955.15	0.0126
	std	10,435.29	154,859,396.19	3,027.87	41,626,820.43	6,235.13	0.0014
	min	4,220.00	60,869,511.81	1,209.00	22,712,059.97	2,404.84	0.0095
	25%	14,825.75	202,629,761.69	4,865.00	69,859,447.55	6,607.66	0.0115
	50%	18,904.00	260,864,798.97	6,613.50	95,165,851.81	9,608.32	0.0124
	75%	24,641.25	339,733,518.51	8,877.75	123,131,081.82	13,061.48	0.0134
	max	101,159.00	1,496,951,020.26	19,630.00	303,000,376.46	47,010.92	0.0176
BAC	mean	121,342.03	1,226,417,338.14	16,790.38	172,754,559.88	133,884.80	0.0097
	std	63,081.55	671,525,306.26	9,444.79	103,812,115.35	209,464.23	0.0027
	min	47,092.00	446,519,241.11	3,829.00	44,580,585.21	26,745.03	0.0017
	25%	82,477.75	788,211,503.98	10,822.25	104,229,472.11	65,699.05	0.0094
	50%	100,742.50	1,022,475,309.71	14,338.00	141,426,121.62	94,437.26	0.0100

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	75%	140,287.75	1,421,516,325.93	21,049.50	207,433,016.43	144,912.96	0.0100
	max	427,260.00	4,871,049,389.04	55,669.00	796,038,773.20	3,041,654.25	0.0328
BRK.B	mean	20,319.82	279,507,272.57	6,127.18	95,881,014.18	9,043.00	0.0115
	std	6,968.24	96,585,788.86	2,236.23	34,040,025.15	7,767.70	0.0008
	min	9,369.00	128,393,899.66	2,833.00	47,226,477.55	3,639.32	0.0097
	25%	15,692.00	211,781,989.51	4,642.50	72,927,387.38	5,853.34	0.0110
	50%	19,071.00	265,611,819.32	5,569.50	87,606,321.13	7,193.05	0.0114
	75%	22,880.50	317,882,557.02	7,111.75	110,321,518.18	9,692.30	0.0118
	max	60,437.00	759,560,842.28	17,403.00	271,880,756.87	94,453.60	0.0149
CAT	mean	30,586.73	269,579,023.84	9,239.74	81,143,988.26	11,986.17	0.0104
	std	11,384.23	107,296,519.47	3,721.88	30,953,831.34	8,988.83	0.0005
	min	7,660.00	72,342,016.91	2,283.00	24,025,499.19	2,847.50	0.0075
	25%	22,670.50	204,956,948.21	6,684.75	58,633,048.01	7,680.36	0.0101
	50%	28,267.00	245,802,664.58	8,451.00	76,013,433.31	10,301.82	0.0103
	75%	36,304.25	323,347,949.42	10,730.00	95,123,308.69	13,455.58	0.0106
	max	77,886.00	964,799,514.35	22,381.00	222,261,612.89	100,244.92	0.0129
CMCSA	mean	60,041.05	451,728,101.18	13,017.85	91,436,280.33	14,855.92	0.0095
	std	24,892.69	183,551,477.13	7,898.99	51,910,212.49	9,307.13	0.0011
	min	22,158.00	165,676,133.42	852.00	6,599,640.92	994.62	0.0026
	25%	44,361.50	333,282,222.44	9,101.00	66,019,849.44	10,217.30	0.0097
	50%	53,545.00	404,921,832.91	12,379.00	89,831,688.91	13,897.14	0.0099
	75%	67,946.50	511,880,757.71	16,574.75	117,696,447.05	18,747.66	0.0100
	max	182,217.00	1,323,876,760.25	44,721.00	299,063,397.63	52,791.06	0.0103
CSCO	mean	77,364.30	493,693,519.98	11,555.12	74,134,548.33	26,409.30	0.0099
	std	33,235.82	207,062,395.07	8,173.36	50,695,319.75	19,401.61	0.0009
	min	31,865.00	182,535,557.30	660.00	4,502,758.25	1,461.77	0.0059
	25%	58,015.00	367,489,467.23	6,881.00	46,381,850.87	15,394.81	0.0099
	50%	68,328.50	444,190,912.86	10,643.00	70,264,638.23	23,922.43	0.0100
	75%	86,368.50	548,980,902.84	14,364.50	92,558,544.11	32,439.85	0.0100
	max	307,808.00	1,702,786,754.09	58,922.00	316,907,129.91	130,317.79	0.0213
CVX	mean	44,441.79	462,460,384.39	12,439.81	134,648,014.78	17,036.07	0.0101
	std	17,816.08	164,739,606.44	6,693.09	56,874,558.81	16,228.50	0.0005
	min	13,879.00	144,582,207.22	2,377.00	28,830,654.89	2,456.15	0.0054
	25%	32,594.50	346,722,417.91	8,344.75	97,784,127.50	9,772.81	0.0099

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	50%	39,655.50	430,819,298.93	10,794.50	122,659,276.18	12,993.21	0.0101
	75%	53,123.50	538,846,282.98	14,257.25	158,567,573.75	18,798.64	0.0103
	max	148,515.00	1,263,782,534.87	50,186.00	423,871,063.95	190,901.32	0.0122
DD	mean	18,036.06	132,521,012.09	4,913.74	37,476,052.09	6,342.15	0.0101
	std	8,759.67	63,295,360.33	2,764.13	19,045,796.05	4,403.69	0.0005
	min	5,262.00	40,582,912.43	773.00	6,491,584.69	832.86	0.0033
	25%	12,017.00	89,557,470.32	3,123.50	24,611,595.27	3,912.60	0.0100
	50%	15,462.00	114,690,819.55	4,104.50	32,687,710.05	5,066.52	0.0101
	75%	20,793.50	155,332,165.60	5,499.00	44,016,133.46	7,243.21	0.0103
	max	52,298.00	418,605,566.86	15,217.00	113,435,890.12	42,392.91	0.0113
DIS	mean	41,156.78	495,392,306.10	10,535.97	129,495,234.75	39,331.64	0.0100
	std	15,686.14	208,901,188.83	4,550.24	53,409,308.00	323,220.91	0.0007
	min	17,030.00	203,854,389.74	2,633.00	36,156,641.31	3,221.94	0.0056
	25%	31,892.25	374,803,933.38	7,657.50	97,517,442.34	10,284.53	0.0098
	50%	36,745.50	430,039,189.49	9,220.50	114,691,273.09	13,055.17	0.0100
	75%	45,623.25	558,118,900.42	11,989.00	144,599,501.14	17,818.25	0.0101
	max	124,145.00	1,659,028,038.95	32,212.00	369,007,239.36	5,138,897.26	0.0180
FB	mean	106,411.67	1,734,512,266.21	24,430.35	382,746,277.09	32,710.88	0.0089
	std	52,106.73	851,612,136.44	14,883.35	218,536,974.53	20,257.94	0.0017
	min	35,721.00	569,582,684.01	2,091.00	30,130,052.95	2,077.65	0.0026
	25%	74,555.00	1,215,942,426.39	16,876.00	273,664,681.48	21,883.93	0.0093
	50%	89,947.50	1,512,724,410.09	22,846.00	368,933,944.10	30,956.82	0.0096
	75%	122,280.50	1,918,383,918.51	30,009.25	474,905,736.65	40,920.19	0.0098
	max	410,069.00	6,838,289,214.05	77,754.00	1,288,151,142.55	104,657.25	0.0101
GE	mean	83,963.26	741,830,493.33	12,828.05	119,789,470.34	44,606.60	0.0095
	std	35,661.52	309,010,418.40	8,012.21	69,234,760.11	71,027.27	0.0015
	min	27,905.00	290,466,991.56	2,653.00	33,035,264.68	7,844.38	0.0044
	25%	59,365.25	514,268,974.25	7,603.00	74,660,777.81	23,089.88	0.0094
	50%	74,767.50	670,261,948.27	10,589.00	96,944,717.80	29,655.10	0.0099
	75%	96,517.00	876,527,780.45	14,826.00	141,143,821.44	45,447.71	0.0100
	max	236,395.00	1,961,985,442.37	49,675.00	427,596,291.36	1,020,533.87	0.0261
GOOG	mean	7,531.45	366,335,942.32	3,011.42	158,709,931.57	20,788.61	0.0732
	std	3,658.82	182,235,732.35	1,689.38	88,251,186.14	16,264.74	0.0303
	min	2,249.00	124,655,348.07	309.00	14,933,581.17	1,081.41	0.0105

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	25%	5,132.25	248,778,207.21	1,962.25	106,767,269.08	11,265.70	0.0555
	50%	6,801.00	329,863,400.24	2,809.50	149,899,630.13	17,390.24	0.0739
	75%	8,621.25	423,295,739.44	3,763.75	193,904,103.00	26,178.89	0.0904
	max	27,533.00	1,380,483,276.83	10,540.00	534,738,339.83	93,508.71	0.1669
GOOGL	mean	7,702.79	386,290,407.80	3,139.58	167,728,031.42	21,666.72	0.0763
	std	3,691.26	193,618,235.38	1,787.47	95,018,077.41	16,290.66	0.0305
	min	2,577.00	166,653,629.71	389.00	22,136,548.67	1,488.76	0.0113
	25%	5,385.25	272,598,797.58	2,021.25	109,053,238.60	12,006.00	0.0607
	50%	6,607.50	329,884,754.37	2,887.00	155,883,718.95	17,816.25	0.0770
	75%	8,442.75	428,950,041.69	3,770.50	200,237,315.54	27,440.66	0.0947
	max	27,052.00	1,363,711,173.41	11,422.00	615,242,804.55	88,509.46	0.1689
GS	mean	16,072.52	266,630,735.82	6,039.60	100,455,871.70	12,632.51	0.0190
	std	6,759.46	124,491,943.62	2,299.09	38,813,844.47	7,817.49	0.0042
	min	5,914.00	106,821,197.38	1,908.00	43,864,040.44	4,126.93	0.0068
	25%	11,672.50	178,117,450.68	4,400.50	72,797,279.72	8,094.88	0.0161
	50%	14,285.00	224,601,809.66	5,593.50	89,806,560.53	10,478.40	0.0181
	75%	18,995.50	329,800,789.60	7,207.25	123,468,835.06	14,081.00	0.0211
	max	50,816.00	857,877,495.97	14,393.00	247,177,637.53	72,612.29	0.0482
HD	mean	27,728.62	366,840,862.69	8,920.89	123,442,984.04	12,744.17	0.0109
	std	8,963.39	127,799,636.38	3,551.59	45,816,593.86	13,953.86	0.0008
	min	13,006.00	165,434,810.90	2,515.00	36,439,575.01	2,864.63	0.0087
	25%	21,668.50	276,025,739.12	6,473.00	92,772,210.99	7,812.18	0.0103
	50%	25,747.00	339,935,686.44	8,234.00	115,952,305.04	9,863.20	0.0107
	75%	31,341.50	416,697,720.23	10,762.00	146,870,527.07	13,201.84	0.0112
	max	64,114.00	1,031,531,952.92	22,597.00	291,592,154.20	186,403.78	0.0141
IBM	mean	19,503.60	283,053,487.10	6,540.58	97,629,157.53	10,322.91	0.0124
	std	7,762.60	121,204,978.80	3,031.09	41,935,403.95	11,852.61	0.0013
	min	6,168.00	83,951,134.35	1,493.00	24,252,638.00	2,042.64	0.0100
	25%	14,595.00	209,597,644.74	4,586.00	71,732,705.50	5,972.78	0.0116
	50%	17,729.00	252,167,134.97	5,852.50	89,826,517.09	7,844.79	0.0122
	75%	22,431.25	328,204,236.66	7,532.75	111,719,286.16	10,385.29	0.0131
	max	59,625.00	972,131,459.03	21,810.00	299,050,973.50	111,628.46	0.0164
INTC	mean	88,012.92	539,061,461.61	13,623.27	80,485,200.80	24,652.76	0.0095
	std	32,133.18	218,280,102.40	8,604.73	50,349,950.02	16,048.53	0.0013

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	min	25,392.00	174,808,926.57	668.00	3,512,129.76	906.38	0.0031
	25%	66,319.50	409,452,090.75	8,564.50	53,076,046.07	15,370.02	0.0098
	50%	81,767.00	493,100,646.52	13,526.00	79,046,604.94	23,962.08	0.0099
	75%	100,219.25	601,791,580.25	17,608.50	104,796,930.94	32,243.13	0.0100
	max	233,578.00	1,765,833,707.79	48,079.00	318,483,188.44	91,380.43	0.0101
JNJ	mean	41,248.16	516,784,968.61	10,117.01	132,739,127.27	15,971.14	0.0099
	std	13,010.19	163,195,302.28	4,751.53	54,033,725.20	24,562.10	0.0003
	min	15,606.00	194,794,413.45	2,156.00	34,113,674.01	3,046.94	0.0082
	25%	32,847.50	413,846,348.61	7,231.25	98,042,130.74	8,347.87	0.0097
	50%	38,411.50	483,292,741.16	8,718.00	117,582,458.80	10,975.76	0.0099
	75%	45,961.50	586,813,347.06	11,288.00	153,593,921.79	16,623.58	0.0100
	max	94,603.00	1,244,615,527.23	32,165.00	338,562,051.69	362,771.34	0.0106
JPM	mean	88,003.57	801,423,694.85	21,356.75	193,852,644.59	29,550.37	0.0098
	std	39,466.22	360,958,601.04	11,483.72	91,730,373.77	14,749.77	0.0003
	min	30,040.00	331,806,293.97	4,953.00	58,788,624.46	7,089.25	0.0069
	25%	61,325.75	565,821,050.04	13,638.00	130,610,914.09	19,065.50	0.0097
	50%	77,139.00	711,684,130.50	17,913.50	171,373,698.77	25,663.01	0.0098
	75%	101,690.75	948,789,239.22	25,153.00	232,200,018.53	34,390.20	0.0099
	max	256,973.00	3,004,137,079.38	70,052.00	646,651,792.53	92,386.71	0.0101
KO	mean	52,120.74	406,264,869.51	10,086.25	81,371,474.40	18,263.90	0.0099
	std	19,287.46	161,269,975.11	4,577.72	33,628,799.23	8,429.36	0.0004
	min	19,958.00	185,384,176.07	3,156.00	30,732,830.23	7,111.74	0.0066
	25%	39,138.50	301,353,437.57	7,209.25	59,076,462.53	13,153.95	0.0099
	50%	47,536.50	368,857,020.57	8,995.00	74,612,460.50	16,482.05	0.0100
	75%	58,796.00	463,324,316.41	11,326.50	92,283,870.91	20,688.83	0.0100
	max	151,901.00	1,308,364,552.46	30,895.00	222,649,014.88	88,890.33	0.0100
MCD	mean	28,809.30	380,847,318.26	7,442.77	103,499,997.57	10,822.55	0.0103
	std	9,250.20	146,529,362.71	2,681.91	39,288,571.37	10,847.35	0.0005
	min	9,911.00	117,553,924.00	2,479.00	32,522,381.94	2,926.51	0.0091
	25%	22,526.25	277,305,454.74	5,422.50	75,412,153.02	6,484.23	0.0100
	50%	26,999.50	355,968,666.10	7,088.50	98,050,825.62	8,795.32	0.0102
	75%	33,173.25	455,898,847.39	8,601.50	121,862,623.83	11,289.40	0.0105
	max	72,028.00	1,044,773,633.09	20,018.00	265,940,261.14	114,279.57	0.0126
MMM	mean	11,365.37	167,307,657.17	3,636.52	57,734,183.69	12,063.44	0.0134

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	std	3,901.98	56,357,516.03	1,769.07	23,315,655.68	102,963.80	0.0018
	min	3,704.00	42,376,029.54	852.00	12,737,335.17	1,268.98	0.0100
	25%	8,870.50	128,614,072.44	2,564.00	42,198,399.30	3,302.92	0.0122
	50%	10,484.00	156,620,977.62	3,148.00	53,116,097.50	4,412.75	0.0131
	75%	13,011.00	192,588,435.88	4,113.00	67,265,920.46	6,182.75	0.0142
	max	27,168.00	374,180,512.59	11,339.00	141,420,561.60	1,638,916.42	0.0211
MRK	mean	52,065.45	404,241,094.10	12,269.51	97,773,420.29	17,435.31	0.0099
	std	21,247.82	198,964,179.96	6,450.80	49,864,763.12	10,578.05	0.0005
	min	18,727.00	139,953,296.94	4,541.00	37,156,365.39	5,935.82	0.0051
	25%	39,157.50	302,275,577.81	7,789.00	64,970,853.90	10,991.95	0.0098
	50%	46,619.50	360,181,487.45	10,518.00	84,618,791.63	15,107.30	0.0099
	75%	58,293.50	460,791,595.64	13,950.25	113,701,288.72	19,811.20	0.0100
	max	232,717.00	2,584,131,245.57	46,595.00	456,348,016.09	112,089.37	0.0156
MSFT	mean	141,856.07	1,190,901,402.50	24,761.04	203,129,267.74	36,706.48	0.0095
	std	63,588.22	533,057,860.34	17,480.90	139,593,661.04	25,629.62	0.0010
	min	37,036.00	459,917,664.02	1,070.00	8,596,209.02	1,253.33	0.0040
	25%	102,602.75	837,915,412.53	14,050.50	122,690,779.06	22,117.84	0.0096
	50%	124,327.50	1,053,837,918.53	22,263.00	180,976,495.62	32,474.83	0.0098
	75%	156,482.00	1,373,674,878.44	32,070.75	262,037,990.81	48,649.14	0.0099
	max	456,106.00	4,125,126,448.00	98,307.00	950,946,403.87	138,913.71	0.0102
NKE	mean	46,386.10	377,535,172.78	10,935.36	89,164,054.37	18,227.11	0.0098
	std	15,357.30	145,806,631.19	3,796.12	32,750,534.77	20,652.07	0.0003
	min	13,818.00	84,721,641.01	2,885.00	18,676,669.13	3,523.46	0.0072
	25%	37,737.50	295,226,871.17	8,613.25	69,144,244.49	12,031.69	0.0097
	50%	42,544.00	344,601,219.52	9,822.00	80,592,736.50	14,390.93	0.0099
	75%	51,532.50	424,862,753.77	12,534.00	102,100,476.22	18,212.90	0.0099
	max	121,962.00	1,195,681,284.35	28,410.00	232,923,873.16	280,266.40	0.0103
ORCL	mean	57,682.71	396,195,204.98	12,588.18	89,349,468.90	23,090.45	0.0099
	std	24,998.32	175,437,821.39	6,925.16	45,328,759.95	13,613.79	0.0002
	min	18,500.00	115,616,594.93	3,976.00	25,275,961.69	6,296.00	0.0073
	25%	41,096.25	279,206,895.10	8,045.50	60,516,930.73	14,818.30	0.0099
	50%	49,330.50	357,457,037.05	10,544.50	76,162,800.69	18,686.12	0.0100
	75%	67,876.75	454,604,488.64	14,292.75	103,662,921.13	26,306.23	0.0100
	max	199,694.00	1,519,364,917.16	49,799.00	320,608,854.11	114,033.10	0.0101

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
PEP	mean	25,878.99	288,145,967.63	6,362.81	77,538,268.01	14,135.13	0.0099
	std	7,302.44	87,491,020.05	2,685.75	30,490,556.29	95,874.35	0.0003
	min	9,697.00	79,276,595.25	1,345.00	15,580,968.88	1,432.80	0.0087
	25%	20,767.00	225,251,680.54	4,648.50	57,988,011.12	5,579.44	0.0098
	50%	24,447.00	272,443,376.61	5,773.50	72,276,964.98	7,056.84	0.0100
	75%	29,553.00	331,719,958.42	7,293.75	90,162,371.68	9,004.46	0.0100
	max	51,761.00	639,054,176.76	19,175.00	213,566,531.33	1,527,179.69	0.0109
PFE	mean	91,040.68	692,324,391.87	13,862.73	110,715,986.10	31,625.70	0.0097
	std	49,256.08	473,362,104.74	6,672.49	60,406,629.99	16,222.08	0.0006
	min	32,599.00	212,898,806.65	4,422.00	28,855,501.39	8,447.34	0.0059
	25%	59,097.50	426,001,630.46	9,726.75	74,658,424.30	21,093.39	0.0098
	50%	80,628.00	611,356,656.02	13,270.50	103,824,254.59	29,745.09	0.0099
	75%	109,044.50	783,454,707.50	16,379.00	129,458,335.77	36,810.46	0.0100
	max	474,221.00	5,427,524,575.47	56,238.00	602,885,333.66	145,936.99	0.0100
PG	mean	50,438.27	570,844,223.65	11,760.85	134,139,256.91	17,786.87	0.0097
	std	26,464.80	419,733,122.26	5,828.25	70,501,433.31	13,441.76	0.0005
	min	19,980.00	185,431,171.67	3,696.00	40,926,831.50	4,789.61	0.0052
	25%	34,682.50	362,299,048.00	7,530.75	87,831,864.68	10,158.57	0.0097
	50%	43,215.50	456,219,304.50	10,168.50	113,664,173.39	14,134.84	0.0098
	75%	57,796.00	612,257,033.06	14,163.00	160,102,759.35	20,335.25	0.0099
	max	181,697.00	3,330,428,860.98	38,467.00	460,594,145.16	111,040.42	0.0101
PM	mean	25,200.89	258,662,053.72	6,473.35	70,383,748.59	7,599.05	0.0100
	std	7,898.99	89,749,492.42	2,363.46	24,104,316.15	2,884.22	0.0003
	min	11,615.00	102,582,018.39	2,283.00	24,491,174.21	2,688.89	0.0089
	25%	20,070.75	198,951,589.31	4,900.00	52,936,626.30	5,660.13	0.0099
	50%	23,642.50	242,915,800.43	5,932.50	65,503,961.62	7,015.20	0.0100
	75%	29,210.50	303,743,337.55	7,560.50	82,154,739.29	8,735.59	0.0102
	max	70,626.00	730,181,888.71	17,838.00	179,429,529.04	20,452.43	0.0118
T	mean	74,002.79	628,431,010.70	14,536.90	128,793,809.07	34,263.92	0.0097
	std	29,190.92	263,898,792.99	6,237.26	53,639,641.10	19,413.10	0.0008
	min	35,829.00	273,472,237.02	4,703.00	43,628,154.84	10,233.73	0.0043
	25%	54,496.25	445,793,490.15	10,359.00	90,668,166.03	22,220.20	0.0098
	50%	67,941.50	579,350,257.07	13,088.00	116,808,479.97	29,783.94	0.0099
	75%	84,811.50	719,130,698.87	17,054.50	148,454,938.65	38,760.70	0.0100

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	max	235,785.00	2,319,764,912.52	45,820.00	402,701,802.86	164,375.07	0.0121
TRV	mean	10,544.19	106,389,400.10	3,568.88	39,506,286.77	4,441.92	0.0115
	std	3,416.58	36,241,051.32	1,447.62	15,393,415.99	2,794.22	0.0011
	min	3,018.00	27,592,851.46	771.00	7,628,101.68	964.98	0.0095
	25%	8,487.25	82,492,360.44	2,705.00	29,702,903.75	2,990.61	0.0107
	50%	9,965.50	101,071,670.21	3,334.50	37,475,398.81	3,837.38	0.0112
	75%	12,010.25	122,831,584.80	4,172.25	46,650,233.97	4,933.57	0.0119
	max	27,468.00	294,476,802.95	11,339.00	107,591,813.81	28,594.17	0.0165
UNH	mean	17,446.67	228,660,097.56	5,642.65	77,377,042.02	7,680.73	0.0119
	std	5,246.70	81,435,935.28	2,011.09	27,032,487.15	4,216.99	0.0013
	min	6,412.00	89,234,548.68	1,849.00	26,225,274.13	2,378.89	0.0098
	25%	14,129.00	173,512,633.11	4,413.50	59,089,553.72	5,357.03	0.0110
	50%	16,636.00	214,637,619.41	5,371.50	75,046,042.50	6,539.35	0.0116
	75%	19,932.50	260,900,529.94	6,717.75	92,546,473.56	8,912.00	0.0125
	max	41,842.00	725,532,688.10	15,652.00	218,550,591.96	30,826.07	0.0178
UTX	mean	24,903.26	263,375,122.23	8,217.23	88,158,366.16	17,510.92	0.0108
	std	12,739.37	141,586,417.29	4,913.50	47,017,041.73	109,897.32	0.0007
	min	5,358.00	49,595,310.95	1,315.00	13,549,323.03	1,579.70	0.0081
	25%	16,977.00	182,655,118.69	4,942.75	59,862,917.81	6,260.83	0.0103
	50%	21,463.00	229,710,235.45	7,034.50	77,766,133.91	8,629.70	0.0107
	75%	27,806.50	295,642,614.43	9,447.25	101,084,369.56	11,780.77	0.0111
	max	86,284.00	1,144,629,181.20	29,297.00	275,444,139.17	1,749,683.12	0.0137
V	mean	48,950.33	460,497,961.48	13,097.62	122,925,302.52	16,818.68	0.0099
	std	17,793.66	170,963,876.45	6,162.95	50,951,708.77	9,603.52	0.0004
	min	23,142.00	162,781,451.21	3,273.00	30,311,313.92	3,873.41	0.0048
	25%	36,797.00	351,926,962.18	9,092.75	88,092,316.21	11,157.63	0.0098
	50%	44,660.00	411,627,578.41	11,552.00	112,507,821.59	14,624.21	0.0099
	75%	56,347.00	531,962,904.82	14,735.00	139,117,677.04	18,457.99	0.0100
	max	128,775.00	1,261,830,529.49	42,661.00	355,125,487.75	85,584.79	0.0106
VZ	mean	62,098.01	494,149,523.80	13,525.58	109,544,287.76	51,450.08	0.0097
	std	23,339.73	185,520,474.01	5,963.55	44,308,281.09	427,124.11	0.0007
	min	29,671.00	204,408,079.42	5,039.00	41,836,595.67	6,539.31	0.0050
	25%	46,137.50	362,469,762.67	9,445.50	77,277,337.22	14,070.98	0.0098
	50%	55,823.50	449,943,857.51	11,922.00	100,842,957.95	19,546.38	0.0099

Continued on next page

Symbol		Trades	Traded Val	Diff Trades	Diff Traded Val	ROC	ROC/Share
	75%	71,345.00	574,383,307.99	15,340.50	128,753,322.82	27,179.16	0.0100
	max	147,919.00	1,264,130,771.03	36,340.00	266,067,716.19	6,798,041.07	0.0109
WFC	mean	95,906.39	820,025,907.36	20,824.21	179,965,091.36	39,793.20	0.0098
	std	45,245.25	431,043,219.13	11,031.12	96,747,201.74	31,259.59	0.0012
	min	37,409.00	318,637,186.95	5,456.00	54,749,684.29	9,945.62	0.0067
	25%	67,742.00	560,383,265.96	13,249.50	117,577,200.44	22,116.46	0.0097
	50%	83,345.00	685,095,920.22	18,570.50	151,911,063.59	30,478.22	0.0099
	75%	109,693.50	942,672,549.65	24,114.00	212,347,399.32	44,038.04	0.0100
	max	403,924.00	3,956,788,572.60	83,572.00	795,479,504.39	320,409.40	0.0273
WMT	mean	49,823.30	448,218,124.11	11,786.63	109,524,107.26	19,815.12	0.0099
	std	20,042.63	187,614,765.33	5,642.63	49,138,231.17	26,412.77	0.0003
	min	20,706.00	211,540,076.99	3,709.00	34,219,678.86	4,605.33	0.0062
	25%	36,156.25	325,522,820.91	7,935.00	74,826,489.33	10,770.16	0.0098
	50%	44,622.50	399,048,171.16	10,657.00	99,148,394.58	14,630.73	0.0099
	75%	57,546.50	520,989,325.68	13,105.25	125,786,171.91	19,936.97	0.0100
	max	156,021.00	1,562,166,750.41	36,698.00	361,429,655.92	246,675.56	0.0105
XOM	mean	64,074.02	670,862,447.91	17,774.64	188,657,442.78	35,104.83	0.0097
	std	28,483.97	265,569,760.30	10,924.37	93,450,720.08	127,162.28	0.0006
	min	21,646.00	201,555,090.63	4,205.00	46,296,094.35	4,953.61	0.0051
	25%	46,888.00	496,895,704.13	11,816.25	129,373,837.81	15,072.46	0.0097
	50%	55,080.50	593,690,988.09	14,020.00	162,976,539.44	18,862.46	0.0098
	75%	74,045.75	786,147,285.48	19,397.50	211,792,668.60	31,372.43	0.0099
	max	209,816.00	1,761,362,028.61	75,421.00	613,405,517.24	2,003,841.58	0.0104

Table C.2: Summary ROC statistics over the 252 trading days in 2016 itemized by stock.