Innovations in financial markets and their impacts on market quality

Duc Man Nguyen

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Finance

Finance Discipline Group, UTS Business School
University of Technology Sydney

June 2021

Certificate of original authorship

I, Duc Man Nguyen, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy in Finance in the Business School at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text.

This research is supported by the Australian Government Research Training Program.

Signature

Production Note:
Signature removed prior to publication.

Duc Man Nguyen
Date: 25/06/2021

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my principal supervisor, Professor Talis Putnins, for his tremendous guidance, support, and encouragement. I deeply appreciate Talis's patience during my PhD journey. The skills and knowledge that I have developed under his supervision will be of great value to my future career.

I am especially thankful to Professor Dave Michayluk, Dr. Vinay Patel, and Dr. Christina Nikitopoulos for their support and assistance. I also truly appreciate the help I have received from the administrative staff of the Finance Discipline Group: Andrea Greer, Mala Kapahi, Stephanie Ough, and Duncan Ford.

I gratefully acknowledge the research funding provided by the Capital Markets Cooperative Research Centre and the Ontario Securities Commission where I was a visiting researcher during the inception stage of my doctorate journey. I am also grateful for the support from the Finance Discipline Group and the Graduate Research School at the University of Technology Sydney.

Various chapters in this thesis have benefited from helpful feedbacks by Benjamin Hippert, Jianning Huang, Samarbakhsh Laleh, Juhani Linnainmaa, Matthew Ringgenberg, and Alexandre Rubesam. A special thank you goes to my friends and fellow research students for their support, especially Hao Ming Chen, Prateek Daniels, Anirudh Dhawan, Linh Do, Junquing Kang, Jonathan Karlsen, Marta Khomyn, Mai Luong, Chung Mai, Huong Nguyen, and Tarun Patel.

Finally, I dedicate this thesis to my parents and brother, for their endless love, support, and encouragement.

Working papers and presentations

Chapters 2–4 of this thesis have been concurrently developed as working papers and presented at various academic conferences. The list of working papers and conference presentations is as follows.

1. Nguyen, D.M., Putnins, T., 2021. The decline in US listed stocks: Is the tick size to blame? Unpublished working paper. University of Technology Sydney.
    - 2018 Financial Management Association (FMA) Annual Meeting.
    - 2018 FMA European Conference (Accepted).
2. Nguyen, D.M., Putnins, T., 2021. ETFs and anomalies. Unpublished working paper. University of Technology Sydney.
    - 2018 Australasian Finance and Banking Conference.
    - 2019 SIRCA: Young Researcher Workshop.
    - 2019 FMA Asia/Pacific Conference.
    - 2019 FMA Doctoral Student Consortium.
    - 2019 Asian Finance Association Conference.
3. Nguyen, D.M., Putnins, T., 2021. Passive price pressure. Unpublished working paper. University of Technology Sydney.
    - 2019 Behavioral Finance and Capital Markets (Accepted).

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2SLS	Two-stage least squares
AMEX	American Stock Exchange
CRSP	Center for Research in Security Prices
ETF	Exchange-traded fund
GDP	Gross domestic product
IV	Idiosyncratic volatility
NYSE	New York Stock Exchange
OLS	Ordinary least squares
p.a.	Per annum
REITs	Real estate investment trust
S&P	Standard and Poor's
SEC	Securities and Exchange Commission
SIC	Standard Industrial Classification
US	United States of America
VAR	Vector auto-regression

Abstract

Financial markets are continually evolving. While many developments and innovations can bring substantial benefits to society, some have unintended consequences and are detrimental to the overall market quality. This thesis comprises three studies of recent issues in US financial markets: the declining number of listed stocks, the growth of exchange-traded funds (ETFs), and their impacts on market quality. The findings advance understandings of financial market developments and their effects.

The number of publicly listed companies in the United States (US) has been declining since the 1990s, prompting policymakers to question why US stock markets are becoming less attractive for capital raising. The first study in this thesis investigates the conjecture made by policymakers that microstructure changes in stock markets (a reduction in the minimum price increment, known as the "tick size") have severe impacts on the market liquidity of small companies. However, the empirical results show that there is no evidence of the microstructure changes being harmful. Liquidity has improved for firms in all size groups, small stock valuations have not been adversely affected, and company managers have not sought to restore the previous relative tick sizes through stock splits. Furthermore, evidence from initial public offering (IPO) prices suggests that new issuers are not concerned about reduced tick sizes. The findings have important policy implications in that increasing the tick size, as proposed by US policymakers, is unlikely to stimulate IPOs or reverse the decline in the number of listed companies.

In contrast to the declining number of listed stocks, the number of ETFs traded on financial markets has grown remarkably and now accounts for a substantial proportion of stock market capitalization and trading activity. The other two studies in this thesis scrutinize the impacts of this ETF growth on market quality. The second study examines whether ETFs harm informational efficiency by free-riding on the price discovery of active investors. The results show that in contrast to fears raised by opponents of ETFs, the growth of ETFs diminishes the profitability of a broad set of asset pricing anomalies, which is consistent with the increasing informational efficiency of ETFs. Much of this effect occurs because ETFs' stock-lending activities reduce short-selling constraints and thereby allow for a more efficient incorporation of information, reducing mispricing. While anomaly returns have been driven

to almost zero in stocks widely held by ETFs, anomalies persist in stocks with low ETF ownership. This study exploits discontinuities in index inclusion to isolate the causal impact of ETFs.

The third study examines a concern raised by policymakers about the impacts of ETFs, namely price pressure. As ETFs become large, do they cause distortions or "dislocations" of prices when they have to rebalance their portfolios or when they receive large in/outflows? The results show that ETF portfolio rebalancing events do not result in significant distortions of stock prices, in contrast to policymakers' concerns and previously shown index inclusion effects. This study shows that the way ETFs rebalance their portfolios has contributed to the disappearance of the index premium in recent years. However, unexpected investments flow into and out of ETFs, causing temporary price distortions in the underlying stocks held by the ETFs, and these price effects take approximately 40 days to revert.

Overall, this thesis contributes to academic and industry understandings of recent financial market developments.

1. Introduction

Financial markets play an essential role in the economy. If the economy is a car, then financial markets are its engine. Financial markets take fuel in the form of surplus funds held by investors and allocate them to productive projects that help propel the economy. All parts of the economic system work together, allocating resources to their best uses so that individuals and businesses can produce income, goods, and services. Collectively, these improve the overall social welfare.

As with any car, the engine design is of utmost importance. A good model facilitates efficient consumption of resources and provides smooth operation. In contrast, a defective model can lead to a bumpy ride and waste of valuable resources. Some issues are easy to fix, whereas others require considerable time and effort. A small company's failure might not cause much concern, but the collapse of multiple large financial institutions during the 2007 Global Financial Crisis created a serious catastrophe. The latter not only put the economy's "engine" to stop but also necessitated a significant market reform.

The infrastructure of financial markets is, in itself, complex and ever-evolving. All parts are connected to each other and work in harmony most of the time. However, defects arise, and new improvements are made to address these flaws. While the engineers (in this case, regulators) make their best attempts, these improvements might have unintended consequences.

This thesis investigates the recent issues and developments that have occurred in the financial system and analyzes their impacts on market quality and financial market participants. This introductory chapter starts with a brief and nontechnical discussion on the fundamental roles of the financial market. Next, it highlights two major growing concerns that indicate deterioration in the overall market quality: disappearance of listed companies and considerable growth of passive investing in the form of exchange-traded funds (ETFs). While the former issue is assessed using a market microstructure scope, the latter issue is scrutinized using different measures of market quality, including drawing on the empirical asset pricing toolkit.

1.1 What are financial markets and why they matter?

Financial markets have existed for thousands of years and evolved considerably over time. On an average trading day, approximately US$250 billion worth of stocks are exchanged in the US alone. This is equivalent to the annual GDP of Bangladesh in 2020, a country with a population of approximately 160 million. The financial system has four primary roles: matching of buyers and sellers, asset transformation, price discovery, and capital allocation.

1.1.1 Matching buyers and sellers

The most fundamental and essential role of the financial market is to match buyers and sellers and facilitate "gains from trades." Because buyers and sellers willingly engage in transactions, both gain from the trading. The gains to users of financial markets include diversification benefits, risk transfer, saving/borrowing, earning a risk premium, and investment returns.

In early financial markets, traders had to be physically present to trade. They verbally negotiated the terms of trades, such as price, quantity, and delivery. Nowadays, millions of trades between traders worldwide are executed electronically within milliseconds with the help of computers and algorithms. Traders can find counterparties with minimal effort through security exchanges. Currencies are used to measure the values of financial products, which have become standardized with clear specifications.

Financial markets also allow resource pooling. Companies have access to a large pool of capital collectively supplied by small individual investors, managed investment funds, and financial institutions. This leads to the second important role of financial markets.

1.1.2 Transformation of size, maturity, and risk

Financial markets help transform the size of funds. They enable large borrowers to borrow considerable money from a collection of small savers and lenders. Using stocks and bonds, companies can raise capital through security exchanges and financial institutions. However, without such products, companies have to reach out to investors by themselves. This process can be very costly and inefficient.

Financial markets also transform the maturity of funds. For example, small savers and lenders might not want to lock in their savings for a long time because of unexpected liquidity needs, such as buying a house, car, or traveling. In contrast, large corporations prefer to pay off their loans after a long time, as investments can take a while to pay off. Without financial markets, this mismatch in maturity can impede corporation investments and dampen economic growth.

Another critical characteristic of capital that financial markets help to transform is risk. A high level of uncertainty increases the cost of capital, discouraging companies from borrowing and investors from putting money into the market. Consequently, valuable projects are forgone and economic growth is dampened. Parties with opposite risks can reduce their risks by trading with each other using financial securities, such as forwards and futures.

Trading in a wide range of assets can facilitate risk diversification for investors' portfolios. The seminal paper by Markowitz (1952) shows that an investor can reduce the total risk of their portfolio by investing in assets that are less than perfectly correlated. Loss from one asset can be offset by a gain from another. Consequently, the overall performance of the portfolio will be less volatile as the idiosyncratic risk of individual stocks is reduced. Today, financial markets provide access to a broad range of securities at very low costs. The recent growth of passive investing, including ETFs, offers investors investments that have built-in diversification benefits. Investors can gain exposure to the whole market or industry instead of investing in individual underlying stocks, which is very costly and takes significant effort. In contrast, diversification reduces risk, which lowers the cost of capital, increases the capital supply for companies, and facilitates economic growth.

1.1.3 Price discovery

The third important role of financial markets is price discovery, which is the "efficient and timely incorporation of the information implicit in investors trading into market prices" (Lehmann, 2002). This process enables prices to converge toward their true "fundamental" values. Informative prices allow investors and corporate managers to make informed decisions that are crucial for asset allocation.

Financial markets facilitate the price discovery mechanism through the buyer–seller interaction. Just like any democratic system, traders express their opinion or valuation of the asset by voting with their money through the course of trading. The price with the most votes is perceived as the true fundamental price. If an investor believes that an asset is underpriced (i.e., the current price is lower than its "true" price), they will vote by buying the asset. Alternatively, if a seller believes that the asset is overvalued, they will agree to sell. This process continues until the asset price reflects the entire information (i.e., a price is "discovered").

Financial markets are more than just platforms for investors to trade. They include media vendors and research institutions that help uncover news and dissect information about assets, companies, and commodities. This makes information available to the market so that market participants can "vote" on the value of that information.

1.1.4 Capital allocation

Finally, one of the most critical roles of financial markets is the efficient allocation of scarce resources. To achieve this, financial markets need to be informationally efficient and liquid. They need to perform the above roles well. Buyers and sellers need to be matched easily so that each can accomplish their objectives, including liquidity needs and risk diversification. Companies should be able to access a large pool of capital to fund their activities, which creates jobs and increases economic output. Prices need to be informative so that investors can differentiate between good and bad investments and capital can flow to the most productive and promising ventures. Market pressure can encourage corporate managers to take up value-adding projects instead of bad ones (Luo, 2005). All these can maximize the return on capital and prevent limited capital from being wasted. The overall result is "improved total welfare."

Wurgler (2000) studies 65 countries and shows a positive relation between the quality of the financial market and the efficiency of capital allocation. Countries with developed financial sectors allocate more capital to expanding industries and less capital to declining industries. Therefore, it is critical to thoroughly investigate the recent issues and developments that have occurred in financial markets.

1.2 Recent developments in financial markets and concerns about their effectiveness

1.2.1 Systematic decline in the number of listed stocks

One of the major roles of financial markets is to provide companies access to capital. Traditionally, the following two primary sources of funds are available for firms: debt and equity. Debt includes loans from banks and debt securities issued to the market, such as bonds, debentures, and promissory notes. Equity comprises the contribution from founders and stocks issued to the market and traded on stock exchanges.

There are many benefits to being a public company. First, the equity market offers a large pool of capital and investors, ranging from large fund managers to individuals. Second, going public helps a company's valuation to be established. The company's securities can be publicly traded, and their liquidity can be considerably enhanced. In addition, stocks and options can be used in employee and executive compensation packages to better align their interests with those of shareholders. Third, corporate managers can learn about important investments from the market by looking at the stock price reactions. Finally, market pressure encourages managers to take on value-adding projects.

Considering the important role of public listing on stock markets, an alarming trend in US capital markets is the decline in the number of listed companies during the past two decades. Figure 1.1 below shows the number of stocks listed on US stock markets over time. In 1996, there were approximately 8,000 listed companies; this number has been steadily decreasing and reached approximately 4,000 today. The declining trend persists even after adjusting for the population growth. This phenomenon has also occurred in other developed countries, such as the UK and Canada. However, the declining trend started first, and has been the steepest, in the US. Thus, investigating the trend in the US is expected to offer great insights.

Figure 1.1: Decline in the number of listed stocks in the US

This figure shows trends in the number of listed stocks in the US from 1975 to 2016. The sample includes all firms listed on AMEX, NASDAQ, and NYSE. The number of firms per capita is displayed on the secondary axis in units of millions.

This tendency is concerning because the total number of US companies (both private and public) is rising, but fewer companies are choosing the equity market as the source of capital, which is the ultimate purpose of the market. This trend is pervasive across all industries. In one of his first speeches in 2017, the Securities and Exchange Commission (SEC) Chairman Jay Clayton stated, "… the effects of [this declining trend] on the economy and society are, in two words, not good."[1] This concern raises an important question: Is the key market for capital raising failing?

1.2.2 Reduction in tick size as a potential driver of the declining number of listed stocks

One possible explanation for the declining number of listed stocks that US policymakers, including the US Congress, have proposed is that small stocks are neglected in the secondary market because of decreases in the minimum price increment (known as the "tick size") and the subsequent repercussions of this structural change in the market. From 1997 to 2001, the tick size in US stock markets was reduced from 1/8th of a dollar to 1/100th

[1] https://www.sifma.org/wp-content/uploads/2019/09/2019-Capital-Markets-Fact-Book-SIFMA.pdf

(one cent) for most stocks. Figure 1.1 above shows that the peak of the number of listed stocks coincides with the introduction of tick size reduction (as shown by the two verticle lines). Thus, US policymakers have argued that the reduction in tick sizes may have diminished the profitability of market making in small stocks, leading to a decline in liquidity provision for these stocks and a subsequent decline in information obtained from analysts. Policymakers conjecture that the deterioration of market quality for small stocks, as well as their low liquidity, means that small companies cannot enjoy the benefits of being public. This discourages these companies from listing or forces them to wait until they are big enough to do so. According to the proponents of this argument, changes to the secondary market structure (tick size) are to be blamed for this.

While, for some market observers, it may seem far-fetched that something as small as a change in the price increments in markets can explain the decline in listed stocks, the chain of reasoning proposed by policymakers illustrates the subtle and complex ways in which financial market innovations can have far-reaching consequences. These conjectures were taken seriously enough that, as a result, the US SEC launched a recent market design experiment, the 2016 Tick Size Pilot, which sought to investigate the liquidity implications of changing tick sizes.

Is there any evidence to support or refute the conjecture that the tick size changes contributed to the declining number of listed stocks? Chapter 2 of this thesis examines this issue. Particularly, it analyzes the liquidity of small stocks vs. big stocks by using several measures to determine whether the market quality for small stocks has deteriorated. It also examines whether tick size changes have influenced the decisions of corporate managers. Finally, it provides insights into whether secondary market liquidity can promote public listings. Various trends that we measure are consistent with the notion that the tick size changes contributed to the declining listed stocks. The evidence contradicts the key conjectures proposed by policymakers.

1.2.3 Growth of passive investing and the era of ETFs

The fund management industry forms a fundamental part of financial markets. It collects money from a wide range of investors and invests in virtually all securities in the market. Investment funds can enhance the efficient allocation of resources by supplying capital to good companies while withdrawing money from bad ones. These funds have a significant impact on market quality and can be categorized into two groups: active and passive funds.

Active funds are actively managed by skillful managers who conduct intensive research to extract valuable information. Their active trading in mispriced securities facilitates market efficiency by making the prices more informative. The primary goal of active funds is to produce positive returns or beat a benchmark, such as the market. In contrast, the main goal of passive funds is to closely track a benchmark regardless of its performance. They often follow a straightforward set of rules, such as a buy-and-hold strategy. Because passive funds are not actively traded, their fee is lower than that of their active counterparts. Thus, passive funds offer investors considerable diversification due to the number of stocks they hold.

Over the past two decades, a fundamental shift in the fund management landscape has occurred worldwide: strong growth in passive investment, fueled by efficient and low-cost investment vehicles, such as ETFs, with a corresponding decline in active management. Competition for capital and performance has led investors to seek low-cost products, such as index funds and ETFs. Overall, passive funds have been empirically proven to outperform active funds (Sharpe, 1991; Malkiel, 2003). This performance difference, coupled with the low fee, has caused considerable capital to flow out of active funds and into ETFs. Bloomberg reported in 2019 that the funds managed by passive managers had exceeded those managed by active managers.[2] Figure 1.2 below illustrates the cumulative flow from active and passive funds over the last decade. As seen in the figure, active funds have experienced a substantial outflow, whereas ETFs and index mutual funds have captured most of these flows.

[2] https://www.bloomberg.com/news/articles/2019-09-11/passive-u-s-equity-funds-eclipse-active-in-epic-industry-shift

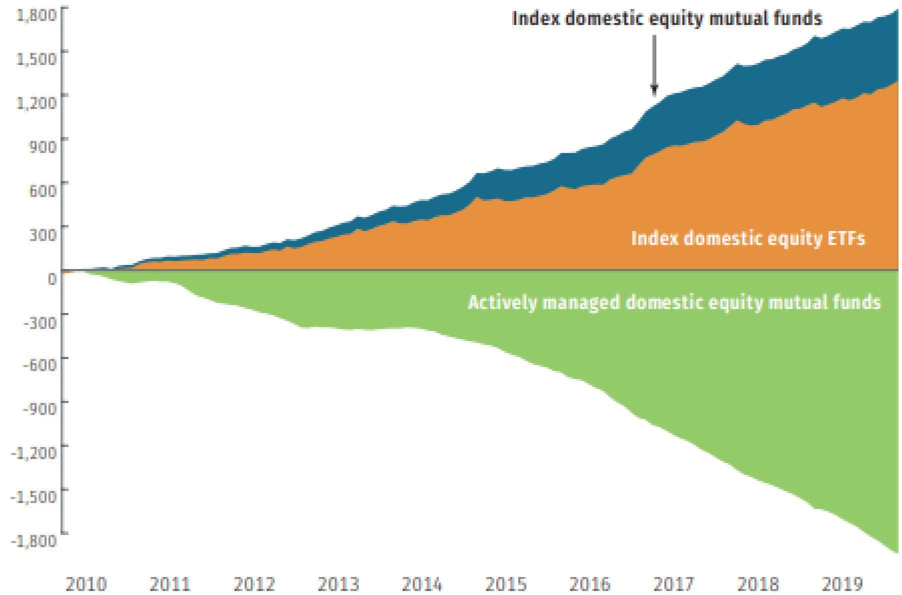


Figure 1.2: Cumulative flows from active to passive funds in the US

This figure presents the cumulative outflows from equity mutual funds and inflows into index equity mutual funds and ETFs in the US from 2010 to 2019. The vertical axis indicates fund flows in USD (billion). Data are obtained from the Investment Company Institute.[3]

The popularity and rapid growth of passive investing have alarmed market observers and regulators, who have raised concerns about numerous potential detrimental effects on market efficiency. Despite the many benefits that passive investing offers for investors, market commentators often claim that index funds and ETFs are free-riding on the price discovery provided by active funds, which is harmful for market quality. Some have argued that passive investment funds are "parasites" and even "worse than Marxism."[4] Even John Bogle, who is often referred to as the father of index funds, has warned that the growth of passive investing might not "serve the national interest."[5] Academic literature on this topic is growing and is not yet conclusive. Chapters 3 and 4 of this thesis contribute to this growing debate by investigating different channels through which ETFs influence market efficiency.

[3] https://www.ici.org/system/files/attachments/pdf/2020_factbook.pdf
[4] https://www.ft.com/content/ab7e1c90-f879-11e2-92f0-00144feabdc0
https://www.cnbc.com/2016/08/24/passive-investing-is-worse-than-marxism-bernstein-strategist-claims.html
[5] https://www.wsj.com/amp/articles/bogle-sounds-a-warning-on-index-funds-1543504551

1.2.4 ETFs and return anomalies

As mentioned in the previous subsection, there is an ongoing debate on the effect of ETFs on market efficiency. One proxy for market inefficiency is the profitability of return anomalies. A return anomaly refers to a phenomenon in which investors can outperform the market by following simple investment rules. The momentum anomaly is a classic example, whereby stocks that have performed well over the past six months or one year tend to continue to do so the next month, whereas stocks that have performed poorly tend to continue to underperform. The ability to earn excess risk-adjusted returns from such simple trading strategies is a symptom of market inefficiency. Therefore, the magnitude of anomaly returns is a way to gauge the informational efficiency of financial markets.

In contrast to claims made in the media and by supporters of active funds that ETFs exacerbate mispricing in stocks, Chapter 3 shows that ETFs can improve market efficiency by lowering anomaly returns. Evidence suggests that at least part of this unintended consequence of ETFs arises through their tendency to lend out the shares that they hold as a way of earning a lending fee that can be used to lower their net management fees. This stock-lending activity relaxes short-sell constraints in stocks, allowing investors to borrow and short-sell stocks at a lower cost. This helps negative information to be impounded more efficiently into prices and reduce mispricing.

1.2.5 ETFs and passive price pressure

As the amount of funds controlled by ETFs has grown substantially and passive investing has approached 50% of all assets under management, further concerns have been raised about the potential for large-scale distortions of asset prices when these enormous funds rebalance their holdings, such as around index reconstitution events. The significance of these events has grown with an increasing amount of money flowing into the ETFs tracking these indices. This, coupled with the growth in trading by high-frequency traders and algorithms, can increase the severity of distortions in prices and cause market instability. For example, on June 28, 2019, $170 billion worth of stocks were traded within the last few seconds of the day. Funds tracking the Russell 1000, 2000, and 3000 indices rushed to buy and sell stocks in those seconds so that they could minimize their tracking errors. These large orders in a tiny time window might have put the market under stress. The Federal Reserve raises concerns that ETFs can amplify the index inclusion effect, which in turn negatively affects the performance of ETFs and that of the investors holding them.[6]

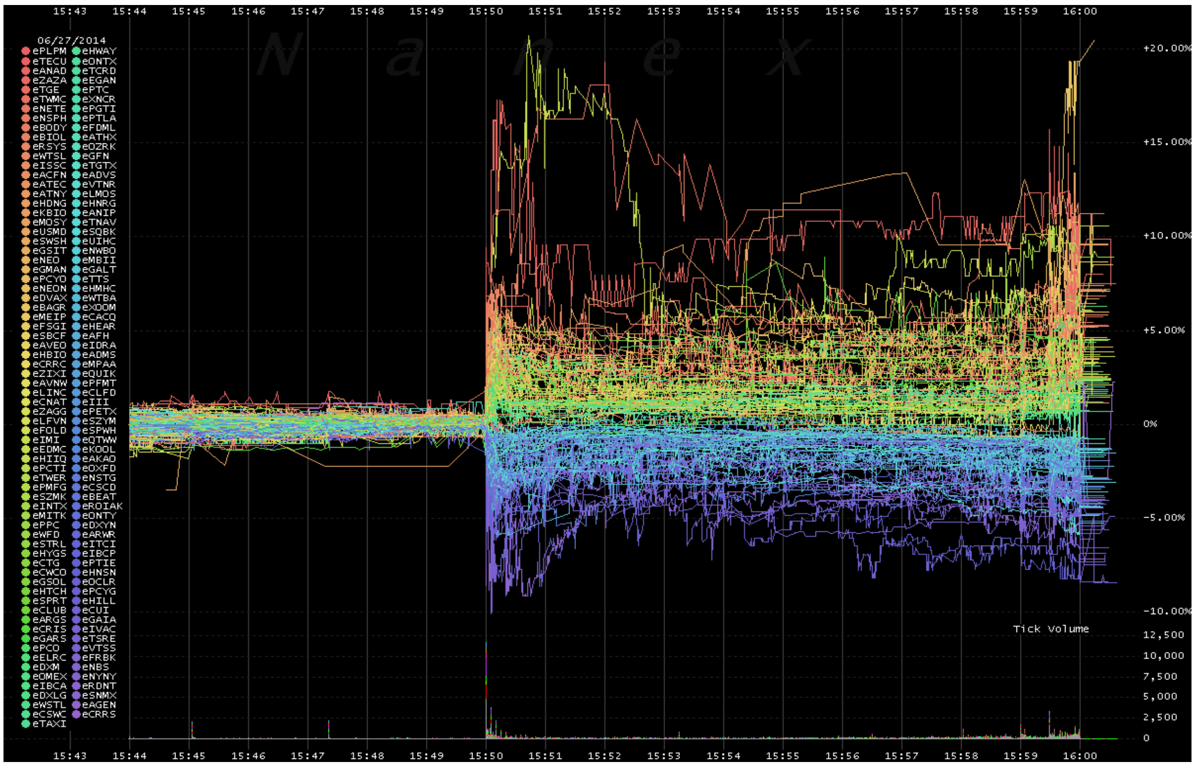


Figure 1.3: Trading around the Russell reconstitution

This figure presents the price movements of stocks related to the Russell 1000 and 2000 indices on June 27, 2014. Each line indicates one of the 143 stocks that moved by 2% or more in a few seconds at 15:50:00. The list of stocks is included on the left of the figure. The vertical axis on the right shows the percent change in stock prices from 15:50:00. Data are obtained from Nanex. [7]

Figure 1.3 below illustrates a chaotic trading scene that occurred around a reconstitution event of the Russell indices on June 27, 2014. Each line indicates one of the 143 stocks related to the reconstitution. Most of the yellow lines indicate stocks that got deleted, whereas most blue lines indicate stocks that got added. The frantic buying and selling started at 15:50, when NASDAQ released the closing imbalance numbers, just 10 min before the market closed. Many of these stocks experienced a price jump or drop of more than 5% within just a few minutes, illustrating the basis of the regulatory concerns about the potential dislocation of prices during intense periods of passive fund rebalancing.

[6] https://www.federalreserve.gov/econres/feds/files/2018060pap.pdf
https://www.morningstar.com/articles/941510/what-happens-if-everyone-indexes

[7] https://www.nanex.net/aqck2/4658.html

Chapter 4 examines how large changes in ETF holdings of the underlying stocks impact stock prices and whether there is evidence regarding temporary price distortions. Do changes in ETF holdings of stocks create distortions in stock prices? Do the sources of these changes in holding matter (fund inflows/outflows vs. index rebalancing)? Do these changes in ETF holdings contain any fundamental information? The answers to these questions will add insights into the ongoing debate on the impacts of ETFs on market quality.

1.3 Thesis outline

This thesis is organized into five chapters. The introductory chapter provides a nontechnical background on the importance of financial markets, recent developments, and their potential concerns. The next three chapters investigate three of the recent and substantial issues in financial markets, addressing the following questions.

a. Chapter 2: Have changes in the tick size contributed to the decline in the number of listed companies in the US?
b. Chapter 3: How do ETFs impact informational efficiency, specifically the "anomaly" returns in asset pricing models?
c. Chapter 4: Do ETFs create price distortions ("price pressure") around index rebalancing events?

The last chapter summarizes the findings of this thesis and provides suggestions for future research.

2. The decline in US listed stocks: Is the tick size to blame?

2.1 Introduction

Policymakers have been questioning whether US equity markets have become ineffective at enabling companies to raise capital, and if so, why. One concerning trend is that the number of listed companies in the US has dropped significantly from 7,885 in 1996 to around 3,742 in 2015. Doidge, Karolyi, and Stulz (2017) show that this decline is in part due to a lower number of IPOs and in part due to a higher number of companies delisting. The trend suggests that the net benefit of being listed has dropped over time, and because the cost of listing has remained approximately the same, it is likely that the benefits of being listed have fallen.[8] However, why has there been a decline in the benefits of being listed? This chapter examines the role played by market microstructures, particularly market liquidity and the price increment at which stocks are quoted and traded (the tick size), as potential drivers of the decline in the number of US listed companies.

The conjecture that secondary market structure and market liquidity might be drivers of the decline has received significant attention from policymakers in recent years. For example, the US IPO Task Force (2011) has argued that changes in the US's capital market structure toward a low-cost frictionless environment has favored highly liquid large stocks at the expense of smaller companies. Much of the argument has centered on the impact of an important market structure change: a decrease in the minimum price increment (the "tick size") from one-eighth of a dollar ($0.125) to one-sixteenth of a dollar ($0.0625) and, finally, to one-hundredth of a dollar ($0.01) between 1996 and 2001. This market structure change coincided with a peak in the number of listed companies and was followed by two decades of declining numbers of listed companies. A Task Force report has noted that this decimalization of tick sizes has changed market maker incentives and profitability, thereby negatively impacting the liquidity of small stocks, making it less attractive for small companies to be listed. In response to these claims and the lack of consensus about the role of the tick size, the

[8] Changes in compliance and regulatory costs cannot explain the decrease in the number of stocks listed in the US (Coates and Srinivasan, 2014; Doidge, Karolyi, and Stulz, 2013; Gao, Ritter, and Zhu, 2013; Hanley, 2017).

US Securities and Exchange Commission (SEC) commenced the Tick Size Pilot in 2016. The pilot involved increasing the minimum tick size for selected stocks from one to five cents for a period of two years to investigate how the tick size affected market liquidity and volatility in small firms.

The US Tick Size Pilot shed light on some of the immediate and direct effects of tick size changes, such as the impacts of tick size on liquidity; for example, see Chung, Lee, and Rosch (2020) and Albuquerque, Song, and Yao (2020). However, it had a limited ability to draw conclusions about long-term impacts and less direct effects, such as capital raising and IPOs. This is because the pilot involved a temporary change to tick sizes over a limited time horizon and affected only select companies. Firms were unlikely to change how they raised capital on the basis of a temporary change, and unlisted firms were unlikely to undertake an IPO because of the pilot. Further, investors could switch to trading similar stocks, while brokers, liquidity providers, and investment bankers may not have fully committed their resources to changing their activities or dealing with pilot stocks because the program was only temporary.

This chapter therefore takes a different approach to examining whether changes in the tick size contributed to the decline in listed companies—it revisits the tick size reductions of the late 1990s and analyzes their medium- and long-term effects on companies, IPOs, and corporate manager decisions. Across a wide range of tests, it consistently finds that decimalization is unlikely to have caused or even contributed to the subsequent decline in the number of listed stocks and number of IPOs.

The main results are as follows. First, we find that secondary market liquidity for US stocks, including smaller ones, has improved significantly since the late 1990s across a number of measures. Firms' quoted bid-ask spreads have dropped by more than two-thirds, the Amihud's (2002) illiquidity measure (ILLIQ) has also dropped, and turnover has improved noticeably since the reductions in tick size. Better liquidity reduces stocks' liquidity premiums, allowing firms to raise capital at lower costs.

However, a challenge in measuring liquidity is that it has several dimensions (e.g., spread, depth, and execution time) that are difficult to capture with a single measurement. This chapter uses a variety of liquidity metrics to capture different aspects of liquidity. Bid-ask spreads help determine trading costs directly. Turnover and Amihud's illiquidity are ubiquitously used in many applications to measure trading frequency and quantity. Other dimensions of liquidity include full order-book depth and time which is difficult to capture given limited data. To overcome this limitation, this chapter therefore also examines the impacts on company valuations, which are likely to reflect the net effect of changes in different dimensions of liquidity as perceived by investors. Using a difference-in-differences approach, it finds that the valuations of small companies, measured by Peters and Taylor's (2016) Total Q, have improved relative to their larger counterparts since the decimalization. This evidence contradicts the conjecture that decimalization contributed to the decline in listed companies by harming small companies and making IPOs less attractive.

If smaller tick sizes were detrimental, especially for smaller stocks, corporate managers would have undone the decrease in tick size by undertaking stock splits to reduce their stock prices and restore the desired relative tick size (tick size divided by price).[9] However, the number of stock splits has declined since the decreases in tick size from 1997, suggesting that corporate managers have not sought to undo the tick size reductions. Similarly, when a company undertakes an IPO, it can effectively choose its relative tick size by setting a number of shares and thus offering prices accordingly. If the reductions in tick sizes were undesirable for newly listing companies, IPO offer prices would have likely decreased post-decimalization so that new companies would have had relative tick sizes in line with pre-decimalization levels. However, IPO offer prices have not decreased compared to pre-decimalization levels, suggesting that newly listing companies have not viewed smaller tick sizes as undesirable. Consequently, we find that stock price levels have remained practically unchanged after the decimalization because neither existing companies nor newly listing companies have sought to undo the tick size changes.

One concern about how measures of market liquidity have changed over time has been a potential selection bias in the sample of US stocks. There have been fewer firms listed on stock exchanges, and the listed firms have become much bigger (Doidge et al., 2017). Many listed companies have chosen to delist or have been acquired by others whilst small private firms have become reluctant to list on an exchange. This has potentially created a bias in comparing liquidity over time based on the stocks listed at a given point in time because the sample composition has changed over time. For the present study, we develop a simple

[9] Angel (1997) argues that a larger relative tick size incentivizes dealers to make markets.

approach to address this issue, which involves estimating what the time-series of market liquidity would be if the sample of stocks were held fixed over time. The selection-adjusted estimates confirm the improvements in liquidity since decimalization.

This research contributes to three strands of literature. The first strand explores the trends and drivers of the numbers of firms listed on stock markets. Doidge et al. (2017) document a drop in the number of US listed firms by almost 50% from its peak in 1996 due to a decrease in the number of new listings and an increase in the rate of delistings. Many policies have been put in place to assist small firms with IPOs, but the decline in the number of IPOs persists. Regulatory burdens and compliance costs cannot explain the decrease in the number of listed firms (Coates and Srinivasan, 2014; Gao et al., 2013; Hanley, 2017). Doidge et al. (2013) argue that financial market globalization has contributed to the declining trend. However, there is no evidence that small US IPOs have migrated to other markets or that small IPOs have increased in other countries (Lowry, Michaely, and Volkova, 2017). This section's contribution to this literature is to investigate one of the possible explanations (decimalization and secondary market liquidity) that has received significant policymaker attention.

The second strand of literature studies the relation between secondary market liquidity and firms' capital raising activities. Mauer and Senbet (1992) and Ellul and Pagano (2006) develop theoretical models in which secondary market illiquidity and adverse selection risk increase IPO underpricing. Stulz, Vagias, and Dijk (2013) show that market liquidity and equity issuance (IPOs and secondary equity offerings, SEOs) are positively correlated. After controlling for other firm and market factors, firms are more likely to utilize private than public equity issues and postpone public equity issues when market liquidity deteriorates. Other studies find that illiquidity attracts an expected return premium, which increases a firm's cost of capital (Amihud and Mendelson, 1986; Anthonisz and Putnins, 2017; Bao, Pan, and Wang, 2011; Chen, Lesmond, and Wei, 2007; Friewald, Jankowitsch, and Subrahmanyam, 2012). Therefore, if stocks' liquidity levels are low causing a high cost of capital, it might be sensible for those companies to delist and seek capital privately. However, consistent with Fang, Noe, and Tice (2009), this chapter shows that stock liquidity has improved since decimalization, suggesting a reduction in the liquidity premium.[10] The higher liquidity increases the benefits of being listed.

The third strand of literature investigates the impacts of changes in tick sizes (including decimalization) on stock liquidity. Overall, smaller tick sizes have decreased spreads, trade sizes, and long run volatility. Quoted depth has decreased since decimalization, but cumulative depth has not changed (Bacidore, Battalio, and Jennings, 2003; Bessembinder, 2003). However, it has appeared that the execution speed for institutional orders has declined since the decimalization, taking longer times and higher costs to fill institutional orders (Chakravarty, Panchapagesan, and Wood, 2005; Jones and Lipson, 2001). The short-term volatility of stock prices has increased post-decimalization while long-run volatility has decreased (Chakravarty, Wood, and Van Ness 2004; Ronen and Weaver, 2001). Interestingly, a higher tick size has improved liquidity for large orders but has deteriorated liquidity for small orders (Chung, Lee, and Rosch, 2020). A higher tick size also has a detrimental effect on stock prices because higher trading costs lead to higher costs of capital (Albuquerque, Song, and Yao, 2020). Given the changes in tick sizes affecting different dimensions of liquidity and market characteristics in different ways, it is not clear what the net effects will be on companies, especially smaller ones. This chapter provides evidence regarding the net effects by analyzing valuation effects and company decisions to undo (or not undo) the tick size changes.

2.2 Hypotheses

The public listing of companies on stock exchanges has a number of benefits, not only for the companies but also for the broader economy and society. For example, Ljungqvist and Tag (2016) provide a political economy model of delisting and show that “delisting can inadvertently impose an externality on the economy by reducing citizen-investors’ exposure to corporate profits and thereby undermining popular support for business-friendly policies. By facilitating companies’ departures from the stock market, private equity firms can trigger

[10] Fang, Noe, and Tice (2009) investigate firm valuations around decimalization, whereas this section compares firm valuations before the first reduction in tick size in 1997 and after decimalization.

a chain of events that may lead to long-term reductions in aggregate investment, productivity, and employment."

Despite the economy-wide benefits of publicly listing companies, the ultimate listing decisions are made by individual companies that weigh their private benefits and costs associated with going public. Direct costs consist of the initial costs of going public, such as IPO fees, underwriter hiring fees, and due diligence and disclosure costs. Additionally, the monetary and human capital costs of maintaining a listing are substantial. Public firms must produce quarterly and annual reports. Insiders, such as executives, are required to disclose their transactions in securities in their companies, which are also monitored by the companies by law. Many firms, such as insurance companies and banks, are regulated by many governmental organizations (for example, the SEC, the Federal Reserve, and the Federal Deposit Insurance Corporation), as well as the exchange, for which requirements can be duplicated, creating redundancies that increase compliance costs (Hanley, 2017). Indirect costs include the results of revealing competitive information and risking becoming a target of mergers and acquisitions. If the costs outweigh the benefits, being listed is not worthwhile for companies, causing them to delist. However, as aforementioned, changes in compliance and regulatory costs cannot explain the decrease in the number of stocks listed in the US (Coates and Srinivasan, 2014; Doidge, Karolyi, and Stulz, 2013; Gao, Ritter, and Zhu, 2013; Hanley, 2017).

Recently, US policymakers and some researchers have proposed that decimalization and Regulation National Market System (Reg NMS) lead to a loss of liquidity and aftermarket support for new-issue and small stocks—for example, the IPO Task Force Report (2011); Weild and Kim (2010); and Weild, Kim, and Newport (2013). They argue that as bid-ask spreads and commissions approach zero in today's efficient market (with the decimalization of tick sizes being a major step in this direction), there is less incentive for market making firms to commit capital in smaller stocks and less ability to profit from market making in such stock. The demise of market making profits in small stocks may have flow-on effects in that sell-side firms are less willing to provide analyst coverage of such stocks and promote them to investors. Therefore, changes in the secondary market structure that may have been beneficial for large stocks may have been harmful for small stocks, contributing to the decline in the number of listed companies.

Based on extant literature and the conjectures of the IPO Task Force (2011), this chapter proposes three hypotheses about the influence of decimalization on firms' listing decisions; these hypotheses correspond to the view that the tick size reductions have been harmful and have contributed to the decline in the number of listed companies. The alternative, if these hypotheses are not supported by the data, is that the tick size changes are not responsible for the decline in the number of listed companies.

Hypothesis 1: The secondary market liquidity for small firms has deteriorated due to decimalization and has caused firms to delist.

Next, we consider the effects of decimalization on companies' valuations. Becoming a public company increases liquidity for the equity of firms by increasing the firms' investor bases, which can improve company valuations through three different channels. First, it improves corporate decision making by enhancing price discovery, which allows managers to learn from market reactions and make better decisions (Foucault and Gehrig, 2008). Second, agency costs are reduced because more informative stock prices provide the market with signals of managerial performance. Liquid stocks allow firms to use more effective stock-based remuneration and align shareholders' and managers' interests more closely. Third, improved price discovery and stock liquidity diminish adverse selection risks and illiquidity premiums, which then lowers firms' costs of capital (Amihud and Mendelson, 2005; Barclay and Hendershott, 2004). If decimalization is harmful to small stocks, changes in valuations are more severely affected in small firms compared to large firms.

Hypothesis 2: *Decimalization decreases small firms' valuations relative to large firms' valuations.*

Angel (1997) proposes an optimal relative tick size hypothesis, which contends that firms adjust their stock prices to maximize their valuations. If a wide tick size is optimal, companies will undo the tick size change by lowering their stock prices through stock splits so that the relative tick size (tick size divided by stock price) moves back to the pre-decimalization level. Moreover, newly listed firms (IPOs) will prefer to offer their stocks at lower price ranges so that the relative tick size is at the optimal level. Consequently, if corporate managers perceive the tick size reductions to be harmful to their stocks' liquidity or valuation, a lower average price for all stocks can be expected after decimalization. We test the following hypotheses:

Hypothesis 3a: The number of stock splits increases post-decimalization.

Hypothesis 3b: IPO offer prices are lower post-decimalization.

Hypothesis 3c: Stock prices decrease post-decimalization.

2.3 Data

This study's sample includes all stocks listed on US stock markets (NYSE, AMEX, and Nasdaq). We obtain prices, returns, and trading volumes from the Center for Research in Securities Prices (CRSP) and accounting variables from Compustat. We exclude collective investment vehicles such as ETFs and REITs. We use daily data to calculate quarterly liquidity measures, and merge with the accounting variables that are reported every three months.

Table 2.1: Variable definitions

Subscripts $\boldsymbol{i}$ and $\boldsymbol{t}$ denote firms and quarters, respectively. All liquidity metrics are winsorized at the 1st and 99th percentiles for each stock and each date.

Variable	Description	Data source
Quoted spread	Average quarterly quoted spread using daily observations, calculated as $Quoted\ spread_{i,t} = \frac{1}{D}\sum_{d=1}^{D}\frac{Ask_d - Bid_d}{m_d}$, where $m_d = \frac{Ask_d + Bid_d}{2}$; Ask_d and Bid_d are the ask and bid quotes at the end of trading day d.	CRSP
ILLIQ	Average quarterly Amihud's (2002) illiquidity using daily observations, calculated as $ILLIQ_{i,t} = log\left[1 + \frac{10^5}{D}\sum_{d=1}^{D}\frac{\lvert r_{i,t}\rvert}{\$Volume_{i,t}}\right]$, where $r_{i,t}$ and $\$Volume_{i,t}$ are daily return and traded dollar volumes, respectively, for stock i during day d of quarter t.	CRSP
Turnover	$Turnover_{i,t} = \frac{\$Volume_{i,t}}{MktCap_{i,t}}$, where $\$Volume_{i,t}$ is the total dollar volume of traded shares in stock i during quarter t and $MktCap_{i,t}$ is the market capitalization.	CRSP
MktCap	The average market capitalization of stock calculated using the product of daily prices and the number of outstanding shares.	CRSP
Volatility	The standard deviation of daily stock returns.	CRSP
MB	$\frac{Total\ book\ value\ of\ stockholders'\ equity}{Outstanding\ shares_{i,t} \times Stock\ price\ at\ end\ of\ the\ financial\ year}$	Compustat

Leverage	$\frac{Total\ book\ value\ of\ long\ term\ debt\ and\ current\ debt}{Total\ book\ value\ of\ stockholders'equity}$	Compustat
ROE	$\frac{Net\ income}{Total\ market\ value\ of\ stockholders'equity}$	Compustat
DY	$\frac{Total\ dividends\ for\ the\ year}{Stock\ price\ at\ end\ of\ the\ financial\ year}$	Compustat

However, the main interest is at a firm level, so we use PERMCO as an identity variable. In instances wherein there are multiple PERMNOs per PERMCO, we choose those with higher market capitalization. The liquidity of all stocks is measured using quoted bid-ask spreads, ILLIQ (Amihud's (2002) illiquidity), and trading turnover, which is a dollar volume scaled by the total market cap of the stock. We obtain the GDP deflator factor from DataStream and IPO data from the SDC Platinum database, including IPO issue prices. Finally, for liquidity metrics, the study's period is from 1993 to 2016 due to the availability of spreads in CRSP, but for other parts of the analysis, the longer time frame of 1975 to 2016 is used. We winsorize liquidity measures at the 1st and 99th percentiles. Table 2.1 reports the definitions and data sources of the variables used in this study.

It is important to note that NASDAQ is a multi-dealer market where trades between dealers are recorded and therefore trading volumes were inflated. This might result in a structural impact on some of the metrics. One might suggest an exclusion of NASDAQ stocks as a robustness check. However, this exclusion would improve our results. NASDAQ has transitioned over time and their volumes have become comparable with NYSE. Some of our liquidity measures such as turnover and Amihud's Illiquidity use trading volume in which NASDAQ stocks would inflate the numbers in earlier years and impose a systematic decline over time. Thus, excluding NASDAQ stocks would improve our results as it removes the systematic decline in trading volumes in NASDAQ stocks.

2.4 Empirical analysis

2.4.1 Number of listed companies

Table 2.2: Listings, delistings, and IPO counts; total market capitalizations; and IPO proceeds over time

This table reports snapshots of the numbers of listed companies, total market capitalizations, numbers of IPOs and amounts raised, and numbers of delistings in the US from 1975 to 2015. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1993 to 2016 (CRSP share codes 10 and 11). The IPO information is from the SDC Platinum database and is accurate after 1980. The CRSP delist codes are used to segregate the numbers of delistings. Delistings (mergers) includes firms that delisted due to mergers and acquisitions. Delistings (other) includes firms that delisted due to liquidation, not meeting exchange/regulatory requirements, or voluntary decisions. The US population is obtained from the World Bank database.

Year	Number of listed companies	Listed companies per million capita	Total domestic stock market cap ($ billion)	Number of IPOs	IPO proceeds ($ billion)	Delistings (total)	Delistings (mergers)	Delistings (other)
1975	4,781	22.1	619.7		0.17	176	90	86
1980	4,896	21.5	1,062.8	142	1.37	290	185	105
1985	6,144	25.8	1,807.5	305	5.38	542	264	278
1990	6,071	24.3	2,690.1	192	4.72	511	195	316
1995	7,268	27.2	5,195.5	640	31.64	539	321	218
2000	6,982	24.7	14,966.0	575	64.47	860	578	282
2005	4,933	16.6	13,260.9	179	29.15	374	231	143
2010	4,090	13.2	12,245.3	120	32.24	319	194	125
2015	3,839	11.9	20,996.0	157	26.79	239	174	65

Table 2.2 reports the numbers of listed stocks, delistings, IPOs, and total amounts raised from IPOs from 1975 to 2016. The number of listed stocks rises from 4,781 in 1975 to a peak at 7,885 in 1996 and plummets to 3,742 by 2016. The numbers of IPOs and delistings exhibit similar trends. Both peak in the late 1990s and decrease dramatically over the next 20 years. However, despite the fall in the number of listed stocks, the total market capitalization for all US stocks increases significantly—more than thirty fold—over four decades, suggesting that the average US firm becomes considerably larger. The firms that are listed in each exchange are examined separately, and the trends are similar. Thus, this is not an exchange-specific phenomenon.

2.4.2 Changes in US listed company compositions

This subsection documents the changes in the numbers of large and small listed stocks in the US. Large, medium, and small firm groups are formed based on their market capitalizations. Starting with the 1975 sample, we sort firms into deciles. The companies in the lowest six deciles are classified as small whereas firms in the top decile (decile ten) are classified as large. The firms in deciles seven, eight, and nine are classified as medium firms. The thresholds for the small and large firms are obtained and inflated with the GDP deflator provided by the US Bureau of Economic Analysis. As illustrated in Figure 2.1, the total number of US listed stocks is almost 5,000 in 1975. This number rises for the next 20 years and reaches a peak in 1996 with about 8,000 stocks. During the next two decades, the number of listed firms consistently decreases, reaching just under 4,000 by 2015, or about as half as many as at the peak in 1996. There are both an increase in the number of firms that delist and a decrease in the number of IPOs that contribute to this dramatic drop (Doidge et. al., 2017).

Equally important are the changes in the numbers of large and small firms. By design, the number of small firms at the start of the sample is six times higher than the number of large firms. However, these two groups display opposite trends over the 40-year period. The number of large stocks increases steadily while the number of small stocks diminishes by almost 70%. In 2016, the number of large stocks is more than double the number of small stocks and higher than the number of medium firms as well. These trends demonstrate that while there are fewer firms being listed on stock exchanges, firms have become much larger.

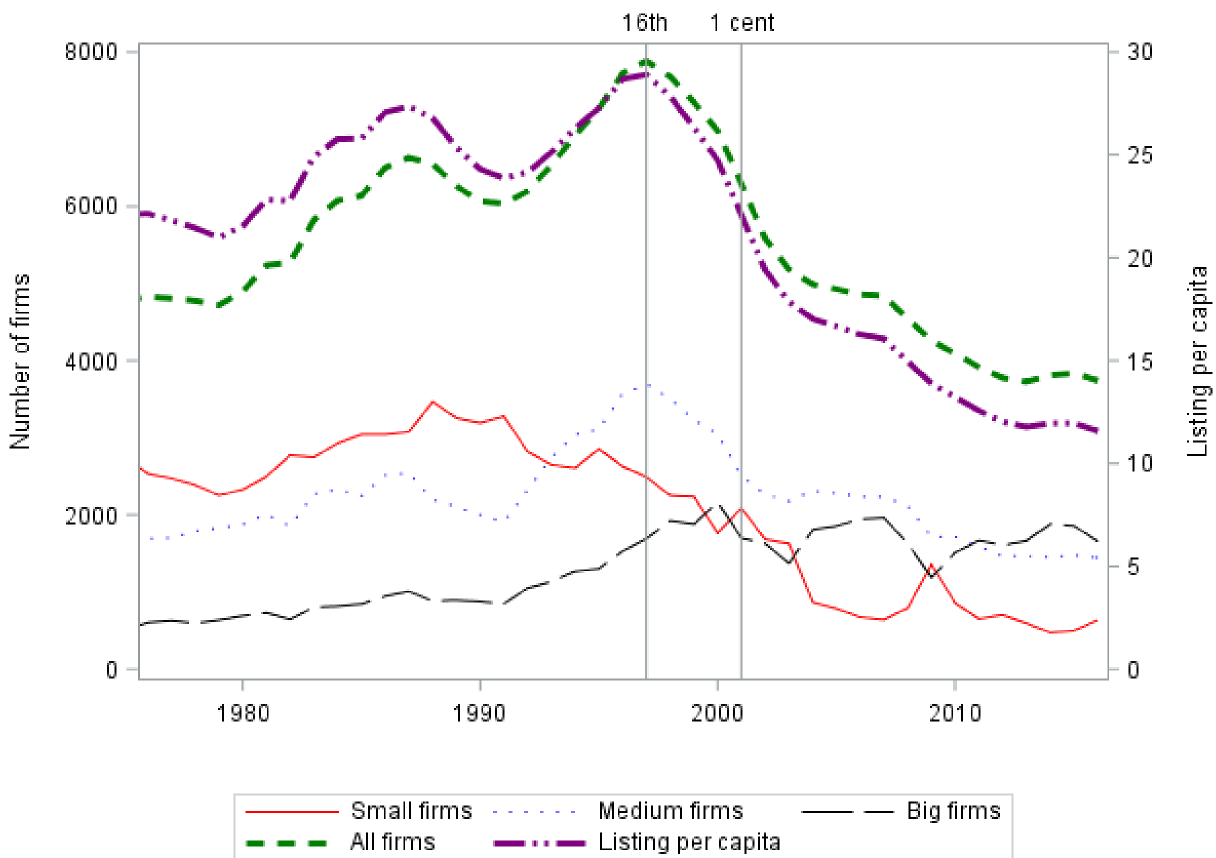


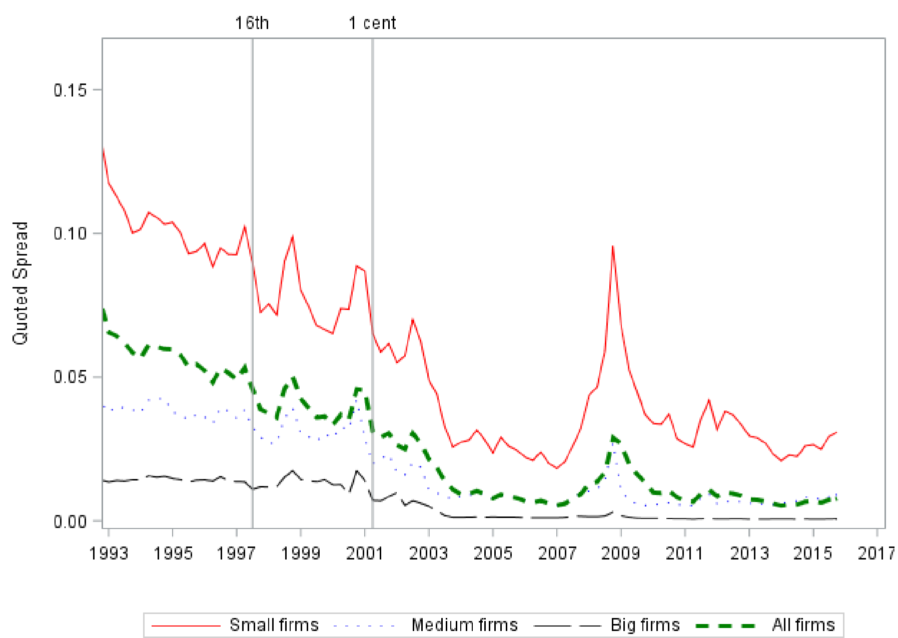
Figure 2.1: The numbers of small, medium, and large firms

This figure shows the trends in the numbers of small, medium, and large firms in the US from 1975 to 2016. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1993 to 2016 (CRSP share codes 10 and 11). Firms are sorted into deciles at the start of 1975. The first six deciles are the level below which firms are classified as small. The top decile is the level at which firms are classified as large. Firms in the sixth to ninth decile range are classified as medium. These thresholds are updated yearly using the GDP deflators provided by the US Bureau of Economic Analysis. The number of firms per capita is displayed on the secondary axis in units of million. The two vertical lines indicate the times when the tick size is reduced to a 16th of a dollar and 1 cent.

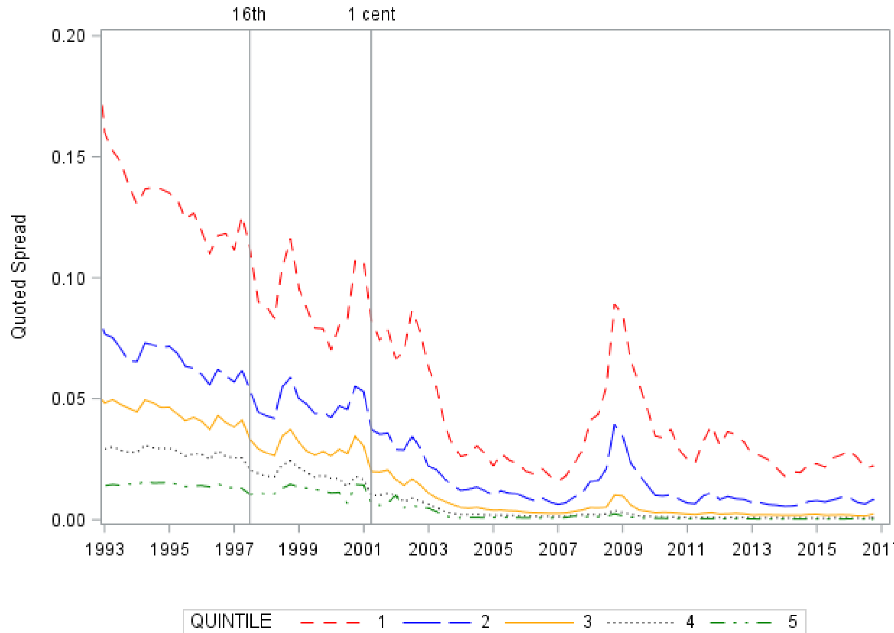
2.4.3 Stock liquidity over time

The IPO Task Force (2011) claims that developments in market microstructures toward a frictionless environment are harmful for small companies and cause fewer small firms to list on stock markets. We investigate this claim by assessing stocks' liquidity. Panels A and B of Figure 2.2 show the equally weighted average ILLIQ and quoted spreads for small, medium, and large firms. Overall, the spreads and ILLIQs for all the groups decrease, suggesting that all the stocks become more liquid on average. This is true whether firms are classified as small, medium, or large using market capitalization thresholds that are inflated over time (Panels A1 and B1) or whether quintiles are considered by market capitalization for each year. Although illiquidity and quoted spreads of small firms are higher compared to medium and large firms, they have improved undeniably since the decimalization.

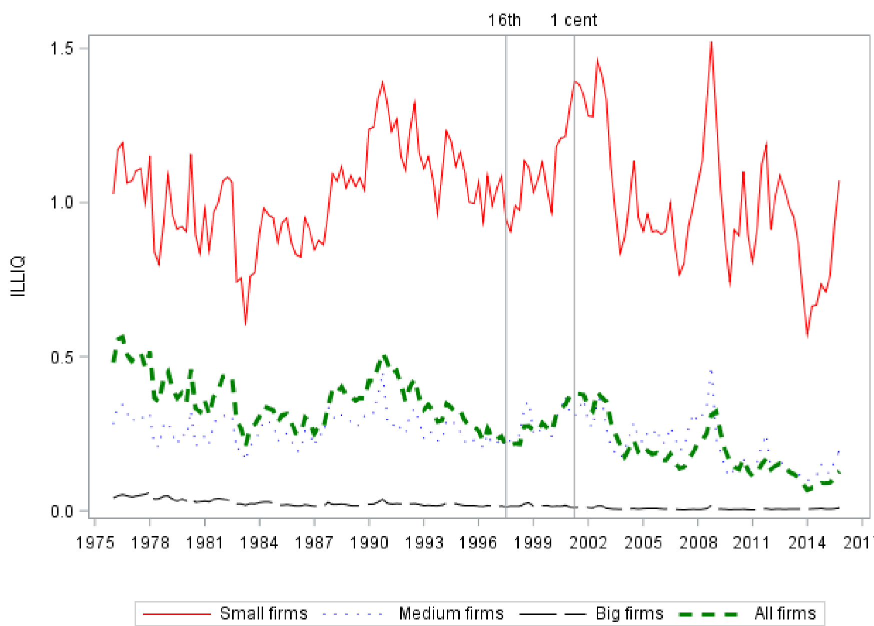
Panel A1: Quoted spread – by groups



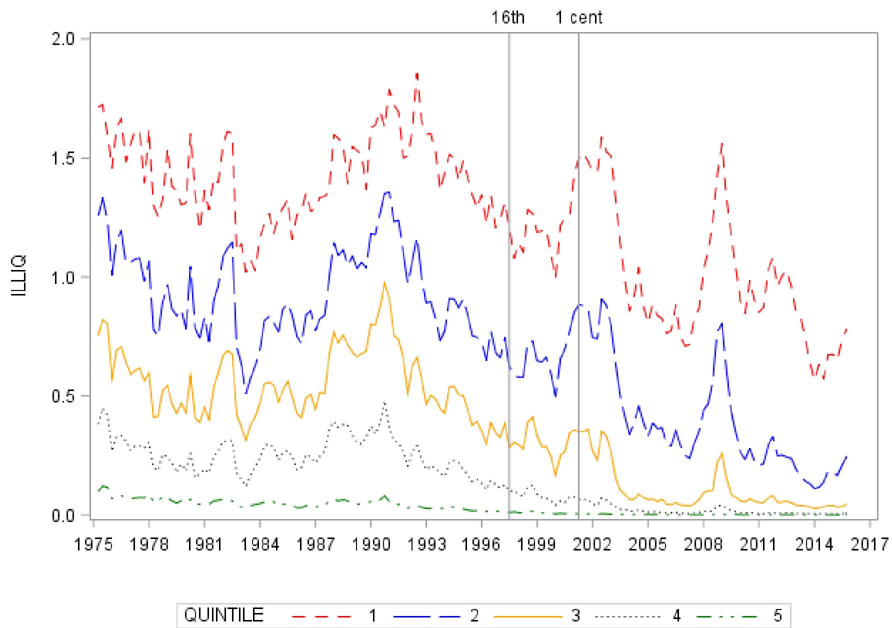
Panel A2: Quoted spread – by quintiles



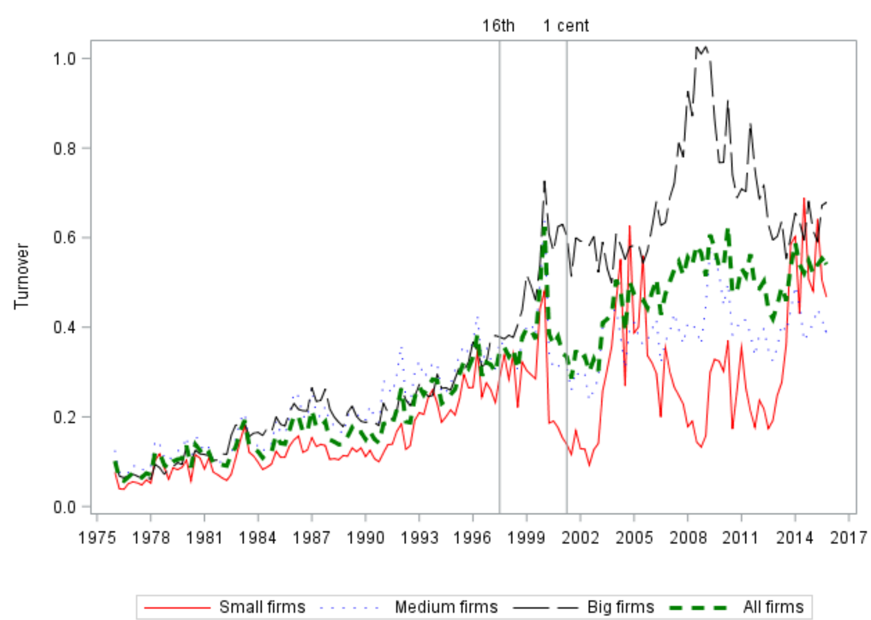
Panel B1: Amihud's 2002 illiquidity (ILLIQ) – by groups



Panel B2: Amihud's 2002 illiquidity (ILLIQ) – by quintiles



Panel C1: Turnover – by groups



Panel C2: Turnover – by quintiles

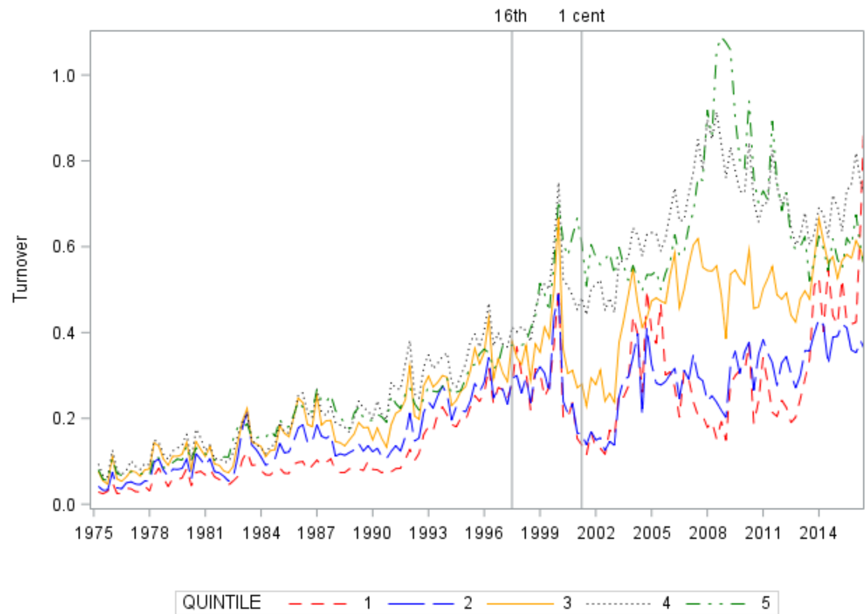


Figure 2.2: The liquidity measures for US stocks

This figure shows the equally weighted average percentage *Quoted spread* (bid-ask spread divided by midquote price), *IILIQ* (Amihud's (2002) illiquidity), and *Turnover* (a stock's total traded dollar volume during the quarter divided by its market capitalization) for small, medium, and large firms in the US and for firms in market capitalization quintiles formed quarterly. Quintile 1 has the smallest firms and quintile 5 has the largest firms with quintiles formed each quarter. The small, medium, and large stock groups are formed in three steps. First, we sort stocks into market capitalization deciles in the first quarter of 1975 and find the market capitalization thresholds that separate the deciles Second, we fill those thresholds forward in time by inflating them using the US GDP deflator. Third, for each quarter stocks are divided into three groups based on their market capitalizations compared to the thresholds (small firms have market capitalizations less than the threshold between the sixth and seventh deciles, large firms have market capitalizations greater than the threshold between the ninth and tenth deciles, and medium firms are the remainder. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1993 to 2016 (CRSP share codes 10 and 11). The sample period is 1975 to 2016. The two vertical lines indicate the times when the tick size is reduced to one 16^{th} of a dollar and 1 cent.

Turnover (traded dollar volume divided by market capitalization) for the small, medium, and large groups offers a slightly different picture. Turnover for the whole market has improved since early 90s but there is a big gap between the groups. Benefits from market developments appear to be strongest for large firms. Turnover of medium firms exhibits a slight improvement whereas that of small firms is very volatile. It rises almost to the levels of large firms in 2005 but plummets during periods of market turmoil such as the Dot Com period in 2000 and the Global Financial Crisis in 2008.

We formally test these trends by regressing liquidity metrics on a time trend variable, $Time$, which is the number of quarters since January 1993. We also include interaction terms between firm groups and the time trend variable to examine the differences between firm groups' liquidity changes through time. According to Table 2.3, the coefficients of the $Time$ variable and its interaction with dummy variables for small, medium, and big stocks are statistically significantly negative at a 1% level for quoted spreads and illiquidity. In contrast, the coefficients for these variables are statistically significantly positive at a 1% level for turnover. These findings suggest that liquidity has improved for all stocks through time.

The turnovers (traded dollar volume divided by market capitalization) for the small, medium, and large groups offer a slightly different perspective. The turnover for the whole market has improved since the early 1990s, but there is a significant gap between the groups. The benefits of market developments appear to be the strongest for the large firms. The turnovers of the medium firms exhibit a slight improvement whereas that of the small firms is very volatile. It rises almost to the levels of the large firms in 2005 but plummets during periods of market turmoil, such as the dot-com era in 2000 and the global financial crisis in 2008.

These trends are formally tested by regressing liquidity metrics on a time trend variable, $Time$, which is the number of quarters since January 1993. Interaction terms are included between firm groups and the time trend variable to examine the differences beween

Table 2.3: Regressions testing changes in liquidity over time

This table reports the results from the regressions of liquidity metrics on a $Time$ trend and interaction terms between $Time$ and dummy variables (D_{small}, D_{medium}, and D_{big}) for the firm groups (small, medium, and large). Firms are sorted into deciles at the start of 1975. The first six deciles are the level below which firms are classified as small. The top decile is the level at which firms are classified as large. Firms in the sixth to ninth decile range are classified as medium. These thresholds are updated yearly using the GDP deflators provided by the US Bureau of Economic Analysis. Liquidity metrics are the quarterly averages using daily observations. The liquidity metrics include *Quoted spread* (the bid-ask spread divided by the midquote price), *IILIQ* (Amihud's (2002) illiquidity), and *Turnover* (a stock's total traded dollar volume during the quarter divided by its market capitalization). $Time$ is the number of quarters since January 1993. The first three columns include stock fixed effects while the last three columns do not. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1993 to 2016 (CRSP share codes 10 and 11). T-statistics are based on double clustered standard errors by stock and time and reported in parentheses. ***, **, and * indicate statistical significance at the levels 1%, 5%, and 10%, respectively.

	Quoted spread (× 1,000)	*ILLIQ* (× 1,000)	*Turnover* (× 1,000)	*Quoted spread* (× 1,000)	*ILLIQ* (× 1,000)	*Turnover* (× 1,000)
Panel A: Cross-section						
Intercept	0.007	−1.780	0.000	56.781**	591.705***	274.190
	(0.01)	(−0.24)	(0.00)	(40.26)	(43.73)	(25.99)

$Time$	−0.435***	−2.192***	3.056***	−0.674***	−4.762***	3.413***
	(−19.90)	(−9.22)	(13.16)	(−23.43)	(−19.15)	(17.03)
R^2 (%)	9.8	1.7	0.9	18.0	4.5	1.7
Stock fixed effects	Yes	Yes	Yes	No	No	No
Panel B: Different size groups						
Intercept	0.037	−1.442	0.304	56.794***	591.035***	276.399***
	(0.06)	(−0.20)	(0.05)	(40.27)	(43.69)	(25.97)
$D_{small} \times Time$	−0.513***	−3.228***	3.911***	−0.696***	−4.271***	1.734***
	(−14.47)	(−6.41)	(9.47)	(−20.71)	(−8.38)	(3.80)
$D_{medium} \times Time$	−0.503***	−2.965***	2.617***	−0.661***	−4.283***	2.047***
	(−16.55)	(−8.26)	(8.42)	(−22.11)	(−12.09)	(7.11)
$D_{big} \times Time$	−0.411***	−1.915***	3.043***	−0.675***	−4.839***	3.644***
	(−19.07)	(−8.15)	(12.30)	(−23.40)	(−19.14)	(17.89)
R^2 (%)	9.9	1.8	0.9	18.1	4.5	1.8
Stock fixed effects	Yes	Yes	Yes	No	No	No

firm groups' liquidity changes over time. As shown in Table 2.3, the coefficients of the $Time$ variable and its interaction with dummy variables for small, medium, and large stocks are statistically and significantly negative at a 1% level for the quoted spreads and illiquidity. In contrast, the coefficients for these variables are statistically and significantly positive at a 1% level for turnovers. These findings suggest that liquidity improves for all the stocks over time.

Small stocks appear to benefit the most, in contrast to large stocks, since the magnitudes of the coefficients on the interaction terms for the small stocks are higher than those of large stocks when we control for firms' fixed effects. We also analyze these liquidity metrics over time for quintiles of stocks sorted by market capitalization, and the trends are similar. The outcomes indicate that the tick size reductions enhance both the large and small stocks' liquidity.

2.4.4 Adjusting liquidity measures for selection bias

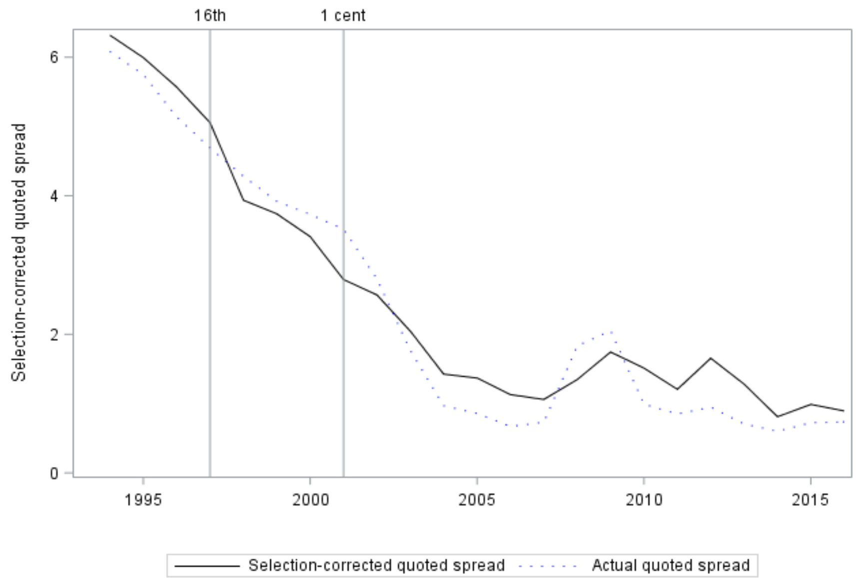
As shown above, the number of listed firms has dropped by one half while listed firms have become much larger over the past two decades. As a consequence, the liquidity metrics of recent years show improvements, perhaps not because of development in the market but because of changes in the composition of the sample toward large, liquid firms as smaller and less liquid forms tend to delist. However, this might give rise to potential sample selection issues causing the liquidity improvements to be overstated. We investigate this possibility using a method to correct for the changing sample composition. The details are in Appendix 2A. In short, it is hypothesized that if stock market developments have enhanced stock liquidity, the 8,000 stocks that existed twenty years ago would exhibit improved liquidity measures if they were traded in today's market environment.

For illustration purposes, consider two stocks, A and B, that are identical in all firm aspects such as size, volatility, and leverage. Stock A is traded in 1993 whereas stock B is traded in 2016. If the market liquidity has remained the same through the years, the liquidity measures for stocks A and B will be identical. If the market liquidity has improved over time, the liquidity measures for stock B will be better than the measures for stock A.

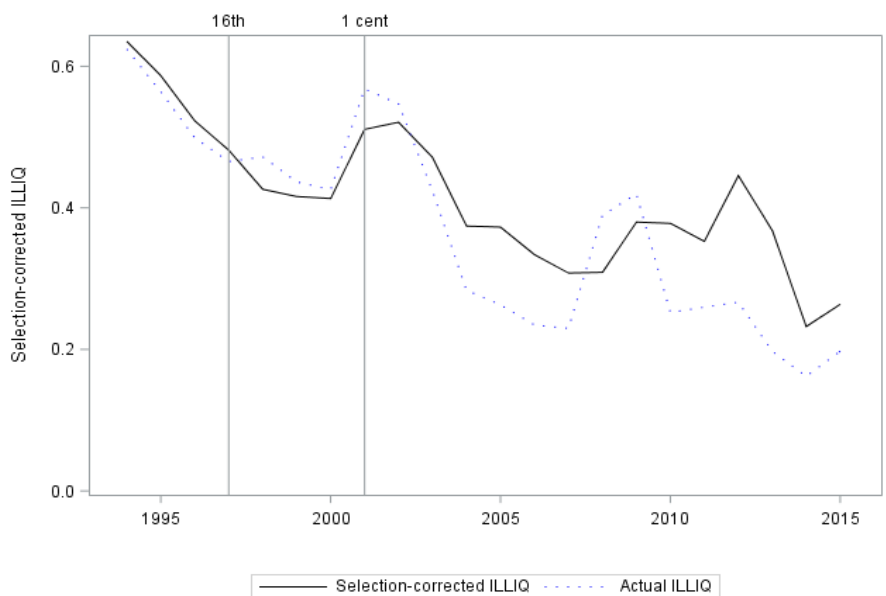
The process described in Appendix A captures the relation between a stock's idiosyncratic characteristics and its liquidity measures over time. These stock-specific characteristics have been shown in the literature to be strong determinants of stock liquidity. They include market capitalization, volatility, leverage, market-to-book ratio, dividend payout ratio, leverage ratio, and return on equity. Appendix A reports the coefficients from regressing stocks' quoted spreads, illiquidity, and turnover for these idiosyncratic characteristics. The estimated parameters of market capitalization and volatility are statistically significant throughout the whole period, suggesting that they are strong consistent factors in firms' liquidity. Larger firms have lower quoted spreads, illiquidity, and higher turnover. Volatility appears to increase not only quoted spreads and illiquidity but also turnover.

We obtain the coefficients from the regressions in Appendix 2A and multiply them with the idiosyncratic characteristics of stocks that existed in 1993 to estimate the hypothetical liquidity for these stocks as if they were being traded from 1994 to 2016. If a selection bias exists, a significant gap will occur between this selection-corrected liquidity and the actual overall market liquidity. Figure 2.3 represents the actual and selection-corrected liquidity metrics from 1994 to 2016. The actual and selection-corrected quoted spreads follow each other closely throughout the period. ILLIQ and turnover figures exhibit similar trends. Overall, the actual and selection-corrected values are not statistically and significantly different from each other. This suggests that there is no evidence of selection bias in the recent liquidity measures. Equity market microstructure developments, including reductions in tick size, appear to have improved liquidity for stocks.

Panel A: Quoted spread



Panel B: Amihud's 2002 illiquidity (ILLIQ



Panel C: Turnover

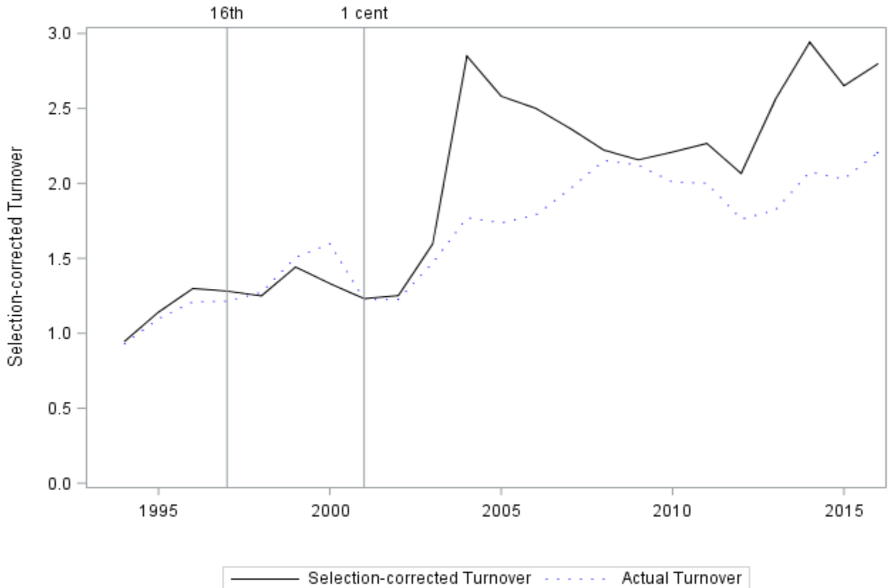


Figure 2.3: The actual and selection-corrected liquidity measures

This figure shows the actual and selection-corrected equally weighted average liquidity measures *Quoted spread* (the bid-ask spread divided by the midquote price), IILIQ (Amihud's (2002) illiquidity), and *Turnover* (a stock's total traded dollar volume during the quarter divided by its market capitalization) for US listed stocks. The actual liquidity values are calculated for all stocks in the given year whereas the selection-corrected measures are estimated for all listed stocks on NYSE, AMEX, and NASDAQ, which existed in 1993. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1993 to 2016 (CRSP share codes 10 and 11). *Quoted spread* is the bid-ask spread divided by the midquote price; *Turnover* is the stock's total traded dollar volume divided by its market capitalization; and, as aforementioned, *ILLIQ* is the Amihud's (2002) illiquidity measure. The two vertical lines indicate the times when the tick size is reduced to one 16^{th} of a dollar and 1 cent.

2.4.5 Impact on company valuations

There are several dimensions of liquidity, such as cost, depth, and time spent filling orders. While the previous section examined several liquidity measures capturing multiple dimensions of liquidity, data limitations prevent us from examining depth directly and there may be other features of liquidity not well captured by the existing measures. For example, investors with large orders, such as mutual funds, are interested in depth not only at the top of an order book but also at the next few levels and the price slippage through the course of trading a large "parent order" that is broken down into smaller "child orders". One remedy is to investigate how company valuations change as a result of the tick size reductions because changes in liquidity as perceived by market participants are expected to cause changes in required returns and thus valuations.

Secondary market liquidity influences firms' valuations through several channels. First, it affects investor required returns and therefore firms' costs of capital; illiquid stocks have higher liquidity premiums (Amihud and Mendelson, 1986). Second, liquidity enhances price discovery, allowing managers to learn from market reactions and make better decisions (Bond, Edmans, and Goldstein, 2012). Third, liquidity can reduce agency costs by making stock-based remuneration more effective or increasing the risk of activist interventions in companies (Edman, Fang, and Zur, 2013). Therefore, changes in market design that result in better liquidity should improve firms' valuations.

The decrease in the number of listed stocks started in 1996, which coincides with the tick size reduction in 1997. Weild, Kim, and Newport (2013) and the IPO Task Force (2011) argue that lower tick sizes caused a loss of secondary market liquidity for small firms. However, if decimalization were harmful to small firms, a decrease in firm valuation for small firms relative to their larger counterparts would be expected. Similarly, if tick size reductions are harmful, we should see larger falls in valuations for companies with a relatively larger change in the relative tick size (tick size divided by price).

We examine the valuation effects of the reductions in tick sizes using a difference-in-differences model. Unlike a standard model in which there are treated and control groups, both groups in this analysis are exposed to the reduction in the tick size, but importantly the impacts are expected to be different. The aim is to measure the difference in the impact of reductions in tick sizes on the valuations of companies with high/low relative tick-to-price levels. We sort firms into quartiles based on their relative tick sizes. The top quartile, which contains firms with high relative tick sizes, and the bottom quartile, which contains firms with low relative tick sizes, are included as a treated group and a control group, respectively. We estimate the following regression.

$$Y_{i,t} = \gamma_0 + \gamma_1 D_{high} + \gamma_2 D_{post} + \gamma_3 D_{high} D_{post} + e_{i,t}, \tag{2.1}$$

where $Y_{i,t}$ is firm i's Peters and Taylor (2016) Total Q, which is an improved version of the Tobin Q measure with intangible assets included.[11] Total Q is the ratio of a firm's market value of equity and debt to the replacement cost of its total assets. D_{high} is a dummy variable that takes a value of one if the firm has a high relative tick size (the top quartile) and a value

[11] Total Qs for firms are obtained from the Peters and Taylor Total Q database provided by WRDS.

of zero if the firm has a low relative tick size (the bottom quartile). To capture valuations before the tick size changes could be widely anticipated, we select the year 1995 as the "pre" period given the first of the tick size changes occurred in 1997, and as the "post" period we select the year 2001, which is when the tick sizes had been reduced to one cent. D_{post} is a dummy variable for the post period (2001). The parameter of interest is γ_3 as it measures the difference in the effect of decimalization on firms with high versus low relative tick sizes. Firms with missing Tobin Q values for either 1995 or 2001 are excluded. This leaves 1,464 observations each year.

Table 2.4 reports the results of the difference-in-differences regressions. Panel A reports the results for groups formed based on relative tick size while Panel B reports the results for groups formed based on market capitalization. All the coefficients are statistically significant at the 1% level. The negative coefficients on D_{high} indicate that firms with high relative tick sizes have lower Total Qs. This is reasonable since a majority of these firms have low stock prices and small market capitalizations. The negative coefficient on D_{post} implies that companies' valuations decrease post-decimalization. This might be partly due to the drop in overall market prices between the two periods. Interestingly, the positive coefficient on the interaction term $D_{high}D_{post}$ suggests that the valuation for firms with high relative tick sizes and small firms increases compared to firms with low relative tick sizes and big firms following decimalization. This result suggests that reductions in tick sizes are not harmful to company valuations of firms more affected by the tick size changes or small firms.

Table 2.4: The difference-in-differences tests of the valuation effects around the decimalization

This table reports the results from the difference-in-differences regressions. In Panel A, stock groups are formed based on relative tick sizes (tick size divided by price) whereas in Panel B, stock groups are formed based on market capitalizations. Stocks are sorted into quartiles based on their relative tick sizes (A) or market capitalizations (B). The top quartile contains the stocks with the highest sort values. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1991 to 2005 (CRSP share codes 10 and 11). The dependent variable is firm i's $Total\ Q$ obtained from the Peters and Taylor's (2016) Total Q database. The independent variables include D_{high}, a dummy variable for firms with high relative tick sizes (A); D_{small}, a dummy variable for firms with low market capitalizations (B); and D_{post}, a dummy variable for the periods after the decimalization (2001). In column 1, the pre- and post-decimalization periods are 1995 and 2001, respectively. In column 2, the pre- and post- decimalization periods are 1991–1995 and 2001–2005, respectively. ***, **, and * indicate statistical significance at the levels 1%, 5%, and 10%, respectively.

	(1)	(2)
	$Total\ Q$	$Total\ Q$

Panel A: groups formed by *relative tick size*		
Intercept	2.860***	2.290***
	(15.95)	(45.47)
D_{high}	−1.777***	−0.681***
	(−8.74)	(−9.67)
D_{post}	−1.315**	−0.934***
	(−6.49)	(−12.79)
$D_{high} \times D_{post}$	1.211***	0.444***
	(4.93)	(4.37)
R^2 (%)	0.83	1.26
Panel B: Groups formed by market capitalization		
Intercept	2.563***	1.251***
	(15.76)	(26.60)
D_{small}	−1.596***	−0.972***
	(−8.29)	(−14.47)
D_{post}	−1.025***	−0.289***
	(−5.47)	(−4.37)
$D_{small} \times D_{post}$	0.764***	0.564***
	(3.33)	(5.80)
R^2 (%)	4.32	1.61

This test is also conducted for large and small firms, and similar results are found in Panel B of Table 2.4. The reduction in the valuations of small companies is lower than that for large companies post-decimalization. In column 2, periods of five years before and after the changes in tick size (i.e., 1991–1995 as the years before and 2001–2005 as the years after) can be seen. The results are similar to those in column 1 and suggest that the reductions in tick size are beneficial for small companies.

2.4.6 Do corporate managers undo tick size changes through stock splits and IPO prices?

Angel (1997) argues that companies use stock splits to achieve optimal relative tick sizes (tick size divided by stock price) because a larger relative tick size incentivizes dealers to make markets. A larger relative tick size also makes investors inclined to place limit orders to supply liquidity. If the smaller tick size after decimalization is detrimental to stocks' liquidity and hence the benefits of being listed, corporate managers can undo the tick size change by lowering stock prices through stock splits as this will bring back the relative tick size to the pre-decimalization level. Additionally, companies undertaking IPOs can select lower price ranges for their offer prices after decimalization to also effectively undo the tick size reductions.

To examine these possibilities, we first analyze the number of stock splits through time. Angel (1997) and Schultz (2000) argue that stock splits increase percentage quoted spreads, which encourage brokers to endorse the stocks. Therefore, if decimalization reduces percentage quoted spreads, causing brokers to be less inclined to endorse the stocks, managers will attempt to split stocks to undo these effects.

Figure 2.4 shows that stock splits have become less frequent since 1998.[12] The number of splits fluctuates around more than 100 per quarter before 1999 but diminishes thereafter. From 2008 to 2016, the number of splits levels out to around 10 per quarter. This time series contradicts the hypothesis that managers undo tick size changes by reducing stock prices through stock splits.

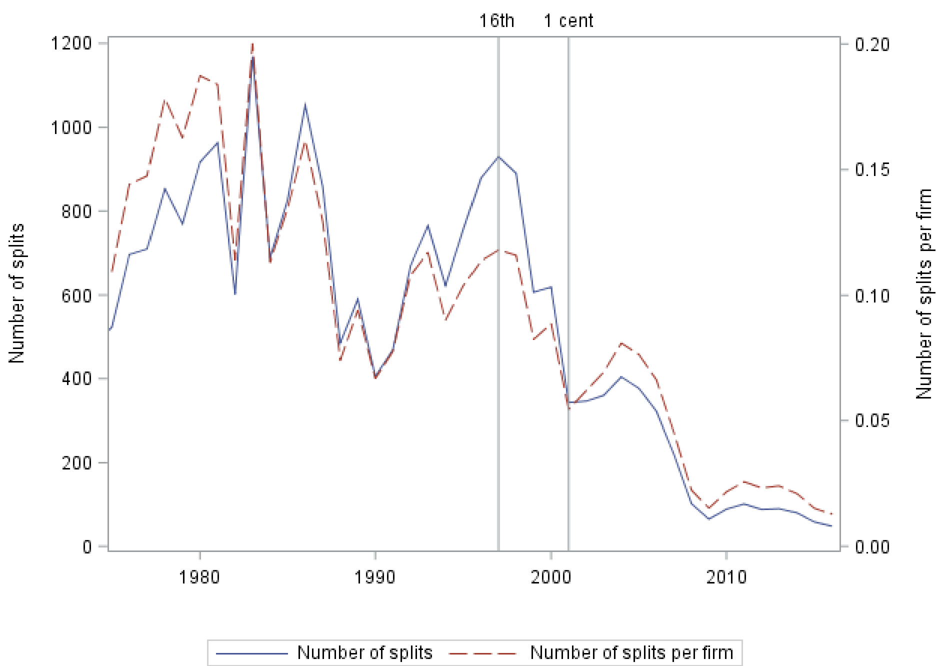


Figure 2.4: The number of stock splits

This figure shows the total number of stock splits per year and the yearly average number of stock splits per firm from 1975 to 2016. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE available

[12] Reverse splits are removed for the purposes of this analysis. They make up around 7% of the total number of splits from the last two decades. However, the inclusion of reverse splits changes the results negligibly.

from CRSP (CRSP share codes 10 and 11) with stock splits identified by the distribution code 5523. Reverse splits are not included.

There is another way for corporate managers to undo the tick size reduction; when conducting an IPO, they can choose a lower issue price (issuing a larger number of shares). Figure 2.5 shows the trend of the average IPO offer prices from 1990 to 2016.[13] The average IPO offer prices show an upward trend during the 1990s but remain stable beginning in 1998. This is consistent with a recent study by Lowry, Michaely, and Volkova (2017). Firms do not appear to lower their offer prices to reduce the relative tick size, implying that decimalization might not be as detrimental as the IPO Task Force (2011) claims.

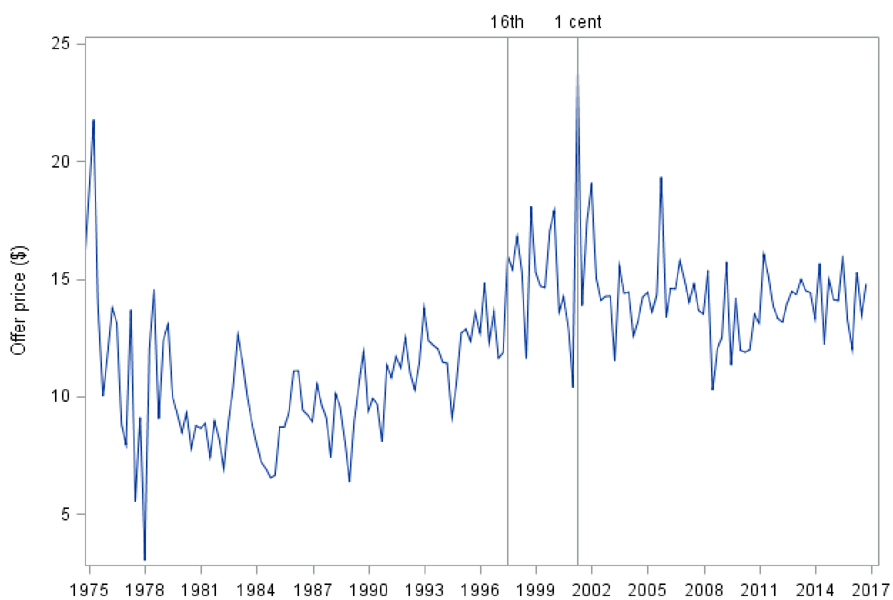


Figure 2.5: The average US IPO issue prices

This figure shows the average IPO issue prices in the US from 1990 to 2016. The sample includes all the IPOs of common stocks that list on US exchanges (NYSE, AMEX, and NASDAQ). Penny stocks, for which the offer price is less than $1.00, are not included. The two vertical lines indicate the times when the tick size is reduced to one 16th of a dollar and 1 cent.

If corporate managers undo tick size reductions through any means of influencing their stock prices (e.g., splits, IPO price decisions, and other means) we should find considerably

[13] The sample of IPOs excludes REITs, depositories, close-ended funds, units, and penny stocks with offer prices lower than $1.00.

lower stock prices after decimalization as it will allow companies to get their relative tick sizes back to pre-decimalization levels. Therefore, we examine the time series of average stock prices. Figure 2.6 shows the average prices of stocks sorted into market capitalization deciles. In contrast to the proposition that firms lower stock prices to undo tick size changes, the average prices in all deciles increase over time. The average prices for the deciles in 2005 are statistically significantly greater than those in 1995. These results indicate that firms do not lower their prices to revert to pre-decimalization relative tick size levels.

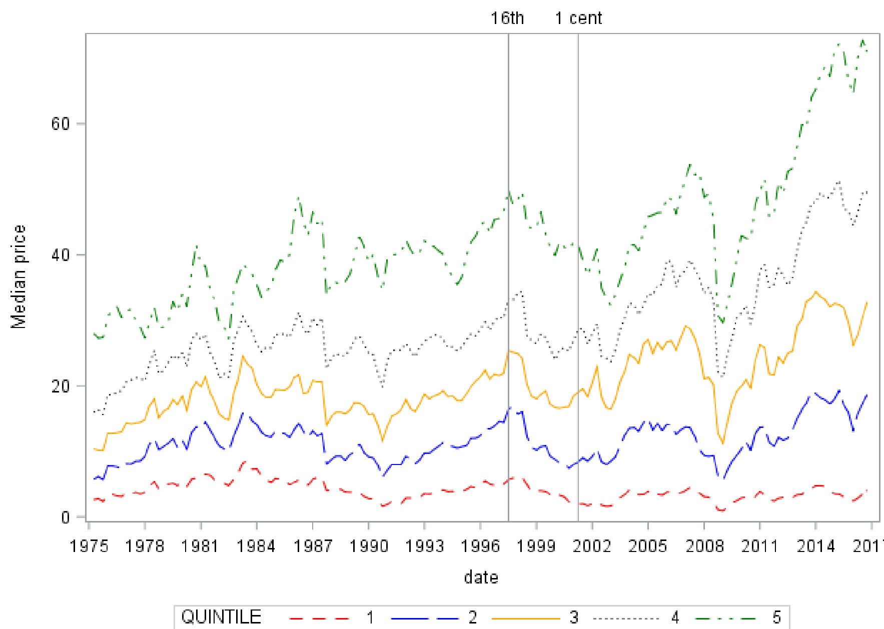


Figure 2.6: The average prices of US stocks by market capitalization deciles

This figure shows the quarterly median prices of US stocks by market capitalization quintiles from 1975 to 2016. The sample includes all ordinary stocks listed on AMEX, NASDAQ, and NYSE from 1975 to 2016 available from CRSP (CRSP share codes 10 and 11). The two vertical lines indicate the times when the tick size is reduced to one 16^{th} of a dollar and 1 cent.

2.4.7 Is delisting probability related to relative tick size?

If the reductions in tick size are harmful to stocks and reduce the benefits of being listed, we would expect a negative relation between relative tick size and delisting probability. To test this, we estimate a logit model. The dependent variable takes the value of one if the firm delists in a given year and zero otherwise. The key explanatory variable is firms' relative tick size and the median price in the previous year. We include year fixed effects. We control for prices falling prior to delisting.

Table 2.5 reports the results, which show that the higher the relative tick size, the higher the probability that a firm will delist. These results therefore also contradict the conjecture that reducing the tick size contributed to the tendency for firms to delist from stock markets. To further control for stock prices tending to decline prior to delisting causing an increase in relative tick size, we estimate similar regressions using one, two, and three lags of relative tick size as explanatory variables and to obtain similar results. The results reject the hypothesis that small relative tick sizes increase firms' probabilities of delisting.

Table 2.5: The relation between relative tick size and delisting probability

This table reports results from logit regressions estimated using the period 1993–2012. The binary variable that is modeled is one if a firm delists in that year and zero if the firm remains listed. The key independent variable *relative tick size* is computed as a stock's tick size divided by its median stock price that quarter. Models 2 and 4 include year fixed effects. *Price* is the median price the previous period. The sample includes all firms listed on AMEX, NASDAQ, and NYSE available from CRSP (the share codes 10 and 11). Delisting events are identified using CRSP delisting codes. ***, **, and * indicate statistical significance at the levels 1%, 5%, and 10%, respectively.

	Dependent variable = Delisting dummy variable			
	(1)	(2)	(3)	(4)
Intercept	−2.314***	−16.908	−2.266***	−13.908
	(−21.67)	(−0.22)	(−15.86)	(−0.81)
Relative tick size	7.134***	5.327***	6.869***	5.319***
	(33.22)	(9.80)	(31.44)	(9.67)
Price			−0.002***	−0.000
			(−4.91)	(−0.05)
Pseudo R^2 (%)	2.13	75.38	2.17	75.38
Year fixed effects	No	Yes	No	Yes

2.4.8 Discussion

All the results of the above tests indicate that reductions in tick size are not harmful to small firms. Liquidity, measured by percentage quoted spreads and ILLIQ, improves significantly for all stocks, including the small ones. Companies do not appear to undo the tick size changes by lowering their stock prices through stock splits nor do new IPOs occur at lower offer prices. Valuations of small firms do not decrease as much as those of their larger counterparts after the decimalization. These results are consistent with an increase in the benefits of firms being listed due to improved liquidity, as opposed to supporting the conjecture that the tick size reductions were harmful and contributed to delisting by small firms.

The results are consistent with recent studies on the US 2016 Tick Size Pilot in which minimum tick sizes for selected stocks increased from one cent to five cents. Song and Yao (2017) document the negative impacts of the pilot. The increase in tick size during the pilot increased quoted spreads, realized spreads, and price impacts while deteriorating liquidity, trading volume, and price efficiency. Prices for pilot firms also dropped, suggesting a decrease in their valuation. Similarly, Hansen, Li, Lunde, and Patton (2017) find that although quoted depths at the national best bid and offer prices increase, volatility increases by around 16% and traded volume decreases for pilot stocks. Griffith and Roseman (2017) obtain similar results and contend that an increase in tick size fails to improve market quality for small stocks.

These results from the tick size pilot are more limited in being able to identify the effects on company listing decisions than those for the long sample and permanent changes examined in this paper because the effects might not have shown up in the pilot for various reasons, mainly due to the fact that the Tick Size Pilot was temporary. Since it was implemented for only two years, market participants might not adjust their behavior or trading strategies. Investors might avoid trading the pilot stocks by switching to other similar stocks that were not involved in the pilot. Hansen, Li, Lunde, and Patton (2017) find a substitution effect wherein trading in stocks within a control group was influenced because investors substituted these stocks for stocks in test groups. In addition, market makers may have been aware that the tick size would drop after the pilot and might not have wanted to commit their resources to providing liquidity and research reports as the pilot anticipated.

2.5 Conclusion

This chapter investigates how reductions in the tick size influence liquidity, valuations, and corporate manager decisions including stock splits, IPO pricings, and listings. We find no evidence to support the conjecture that the tick size is a driver of the fall in the number of US listed companies. Reductions in the tick size improve firms' liquidity measured by quoted spreads, ILLIQ, and turnover. Improved secondary market liquidity tends to lower the cost of secondary offerings, making it cheaper for firms to raise capital from equity markets. This tends to incentivize rather than to disincentive firms to be listed.

This chapter has policy implications in light of a recent debate about the causes of the declining number of listed companies and the regulatory experiments, such as the SEC's Tick Size Pilot. The results suggest that increasing the tick size is unlikely to increase IPOs, bring back the number of listed stocks to the level in the late 1990s, or improve secondary market liquidity. The results also show that the tick size had nothing to do with the decline in the US listings. Future research should explore other potential causes of the declining number of listed stocks.

Appendix 2.A: Selection-corrected liquidity measurement

Recent samples of stocks might be rather liquid due to the relatively illiquid stocks dropping out of the sample. As illustrated in Figure 2.1, the number of small stocks has dropped significantly, whereas the number of large stocks have shown a noticeable increase. There were three times as many small stocks as large stocks in 1993, but the gap reduced in 2016. The number of medium firms also increased significantly and formed 40% of the total number of firms in 2016. These changes in stock sample composition might have inflated both equal- and value-weighted liquidity measures in recent years.

We develop a simple approach to address this issue. If stock market developments have enhanced stock liquidity, stocks that existed 20 years ago should exhibit better liquidity if they were to find themselves trading in today's market environment. We estimate the following cross-sectional regression to examine how firms' characteristics influenced their stock liquidity in 2016:

$$Y_i = \alpha_i + \beta_C C_i + e_i \tag{2.2}$$

where Y_i is one of firm i's liquidity measures (*Quoted spread*, Amihud's illiquidity (*ILLIQ*), and *Turnover*); C_i is a vector of firm-specific characteristics, and e_i is the error term. We obtain and apply these coefficients to the 1993 sample:

$$\widehat{Y_j} = \alpha_i + \widehat{\beta_C} C_j + e_j \tag{2.3}$$

where $\widehat{Y_j}$ is the estimated liquidity of firm j in 2016 and $\widehat{\beta_c}$ is a vector containing estimated parameters from Equation (1). The estimated 2016 liquidity of the 1993 sample is compared to the actual liquidity of the 2016 sample. If selection bias exists, these two figures would be statistically significantly different.

We extend this analysis and estimate the liquidity of stocks in the 1993 sample in every year between 1994 and 2016. We are interested in what the overall market liquidity would be if all the firms that existed in 1993 also existed in these years.

$$\widehat{Y_{J,t}} = \alpha_j + \widehat{\beta_{C,t}} C_{j,t} + e_{j,t} \tag{2.4}$$

We obtain one set of coefficients $\widehat{\beta_{C,t}}$ for each year from 1994 to 2016. This becomes a panel analysis where t denotes the year in which the coefficients are estimated and j represents firms that existed in 1993. The results are reported in the table below.

Table 2.A.1 Regressions of quoted spreads on stock characteristics.

This table reports the results obtained from regressing firms' liquidity measures (*Quoted spread, ILLIQ,* and *Turnover*) on firms' market capitalization (*MktCap*), stock return volatility (*Volatility*), market-to-book ratio (*MB*), return on equity (*ROE*), dividend yield (*DY*), and *Leverage* computed using COMPUSTAT annual data. The sample includes all firms listed on AMEX, NASDAQ, and NYSE from 1993 to 2016. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Year	Intercept	*MktCap* (×10,000)	*Volatility*	*MB* (×1000)	*ROE*	*DY* (×1000)	*Leverage* (×1000)	R^2 (%)
Panel A: Quoted spread								
1994	1.292***	-1.255***	108.648***	-3.584	-0.262***	-29.089	7.710*	57.9
1995	0.978***	-1.264***	119.877***	0.037	-0.361***	-0.131	2.837**	61.4
1996	1.068***	-1.074***	111.363***	-16.513***	0.014	11.218	35.033***	61.1
1997	0.175***	-0.682***	121.017***	-7.798***	0.125***	33.616*	13.338***	60.4
1998	0.322***	-0.446***	105.845***	-4.511***	-0.078*	-2.298	8.135**	54.8
1999	1.119***	-0.292***	63.682***	-9.152***	-0.043	-23.950	4.392*	35.7
2000	1.363***	-0.188***	53.532***	-10.206***	-0.165***	-17.142	-1.077	25.4
2001	1.628***	-0.197***	37.721***	-8.526***	-0.903***	-0.318	34.694***	16.6
2002	-0.101	-0.141***	64.482***	-7.391***	-0.011***	15.238*	10.074***	43.3
2003	-0.338***	-0.155***	64.677***	0.144	-0.018***	8.748	0.196	49.4
2004	0.016	-0.097***	45.120***	-0.443	-0.073***	0.323	0.037	44.2
2005	0.076**	-0.091***	29.933***	-0.350	-0.057*	6.826**	0.527	23.2
2006	-0.003	-0.077***	30.487***	-0.171	-0.026	5.022*	0.372	25.6
2007	0.001	-0.058***	24.927***	-2.882***	-0.112***	3.352*	5.417**	24.3
2008	0.083***	-0.045***	21.589***	-3.149***	-0.145***	0.683	2.469**	19.1
2009	-0.561***	-0.084***	42.489***	-2.024*	-0.084**	-6.740**	2.252	25.8
2010	0.726***	-0.171***	22.883***	0.192	-0.019***	-3.204**	0.598	20.2
2011	-0.445***	-0.048**	43.751***	-2.613**	0.039**	1.456	1.424	31.4
2012	-0.280***	-0.056***	32.850***	-1.568**	-0.107***	1.941	4.746**	21.1
2013	-0.377***	-0.030*	45.114***	-2.150**	-0.064***	1.046	1.333	30.1
2014	-0.053	-0.034***	29.702***	-2.189**	-0.032***	-0.215	2.764	22.6
2015	0.244***	-0.045***	12.354***	-0.422	-0.137***	1.542	0.122	11.9
2016	0.124***	-0.040***	19.014***	-0.338	-0.125***	0.648	-0.292	18.0

(Table 2.A.1 continued)

Year	Intercept	*MktCap (×10,000)*	*Volatility*	*MB (×1000)*	*ROE*	*DY (×1000)*	*Leverage (×1000)*	R^2 (%)
Panel B: ILLIQ								
1994	0.271***	-0.202***	7.093***	-0.694*	-0.011*	-9.384*	1.826***	25.7
1995	0.308***	-0.234***	7.823***	-0.125	0.022*	-9.711*	0.392*	21.8
1996	0.271***	-0.167***	7.509***	-3.382***	0.031***	-7.797*	6.355***	21.8
1997	0.173***	-0.118***	8.075***	-1.220***	0.020***	-2.129	1.541**	19.3
1998	0.134***	-0.059***	7.999***	-0.598***	0.022***	-8.974***	0.674	21.7
1999	0.175***	-0.039***	5.837***	-1.420***	0.006	-10.141***	0.679*	17.8
2000	0.187***	-0.026***	5.318***	-1.408***	-0.001	-9.767***	0.124	13.3
2001	0.312***	-0.033***	2.230***	-1.252***	-0.027**	-3.152***	5.402***	4.1
2002	0.255***	-0.037***	5.846***	-1.629***	-0.001*	-5.000***	2.405***	13.5
2003	0.177***	-0.046***	7.732***	-0.194	0.000	-2.157	0.673	17.8
2004	0.162***	-0.038***	6.930***	-0.126	-0.003	-0.766	0.012	16.2
2005	0.119***	-0.033***	5.710***	-0.137	-0.004	1.226	0.215	6.9
2006	0.085***	-0.033***	6.425***	-0.096	0.000	1.494	0.236	8.3
2007	0.103***	-0.030***	5.128***	-1.426***	-0.019*	1.239	3.709***	5.9
2008	0.085***	-0.020***	4.962***	-1.547***	-0.037**	0.159	1.361**	6.8
2009	0.001	-0.026***	6.849***	-0.486*	-0.029***	-1.450**	0.555	16.6
2010	0.226***	-0.044***	3.479***	0.057	-0.002*	-0.671**	0.136	13.8
2011	-0.072***	-0.017***	10.090***	-0.723**	0.017***	0.349	0.348	23.3
2012	-0.024	-0.021***	8.344***	-0.563**	-0.027***	0.476	1.655***	13.4
2013	-0.062***	-0.011**	11.307***	-0.630**	-0.008	0.165	0.326	21.3
2014	-0.021*	-0.010***	8.680***	-0.770**	0.003	-0.091	1.105	16.9
2015	0.055***	-0.012***	3.921***	-0.209	-0.023*	0.395	0.055	9.0
2016	0.044***	-0.012***	4.854***	-0.168	-0.028***	0.191	-0.058	12.0

(Table 2.A.1 continued)

Year	Intercept	*MktCap (×10,000)*	*Volatility*	*MB (×1000)*	*ROE*	*DY (×1000)*	*Leverage (×1000)*	R^2 (%)
Panel C: Turnover								
1994	1.073***	0.109	0.006	-0.600	0.023*	-52.818***	-2.254*	1.0
1995	0.935***	0.274***	0.460	0.230	0.003	-65.832***	-0.523	0.9
1996	1.057***	0.281***	2.134***	7.593***	0.051**	-64.794***	-17.310***	1.7
1997	0.965***	0.007	7.514***	3.687***	0.054***	-10.479**	-4.339**	3.1
1998	0.995***	0.102***	6.797***	2.399***	0.053**	-50.533***	-3.270**	2.9
1999	0.744***	0.097***	11.610***	5.693***	0.160***	-45.021***	-1.797	5.6
2000	0.495***	0.058**	21.002***	7.795***	0.102***	-23.858**	-1.801	10.5
2001	0.141***	0.075***	27.391***	3.901***	0.428***	-9.165***	-12.903***	17.9
2002	1.045***	0.098***	3.947***	6.809***	0.004*	-22.517***	-5.520***	1.8
2003	1.147***	0.118***	2.347***	0.702	0.003	-20.123***	-2.039	0.6
2004	1.213***	0.062**	8.507***	0.882**	0.023***	-8.342**	-0.098**	2.0
2005	-0.205**	0.159***	68.449***	0.434	0.316***	-17.837**	-0.635	13.6
2006	0.464***	0.160***	47.646***	0.390	0.320***	-20.118***	0.404	8.5
2007	0.888***	0.141***	36.045***	7.022***	0.221***	-17.393***	-8.116	8.0
2008	1.412***	0.097***	21.848***	6.802**	0.448***	-10.289***	-3.546	3.1
2009	2.267***	0.114***	-1.074	1.492	0.109***	3.318	-0.879	1.0
2010	2.155***	0.033	-0.073	-0.538	-0.001	9.724***	-0.251	1.7
2011	1.594***	0.161***	13.228***	3.522	-0.019	-7.541**	-0.595	2.1
2012	1.216***	0.145***	23.626***	2.527**	0.204***	-8.829**	-6.336**	3.8
2013	1.287***	0.072***	17.326***	2.431*	0.107***	-3.986	-0.146	2.5
2014	0.950***	0.051***	35.646***	2.903	0.048**	-1.468	-0.313	8.2
2015	1.031***	0.050	41.953***	0.302	-0.218	-7.068	-0.135	9.8
2016	0.919***	0.039*	37.727***	1.361	-0.287***	-2.551	-0.059	14.6

3. ETFs and anomalies

3.1 Introduction

The rapid growth in ETFs since the turn of the millennium has had a profound impact on the investment management landscape and the market for individual stocks. Assets under management (AUM) by ETFs in the US surpassed $5.4 trillion in 2020, and the number of ETFs continues to increase.[14] ETFs hold more than 10% of stock market capitalization and over 30% of trading volume in stocks is attributable to ETFs (Ben-David et al., 2017). Many market observers and policymakers fear that this significant shift in the investments landscape toward low-cost "passive" index-tracking vehicles has the potential to undermine market efficiency.[15] Opponents argue that ETFs are like parasites that free-ride on the price discovery produced by active fund managers, even going so far as to label ETFs as being "worse than Marxism" in that they impede the capital allocation role of efficient market prices.[16]

This chapter examines whether this is true by investigating the relation between ETF holdings of stocks and the profitability of well-known asset pricing anomalies as a measure of market efficiency. We show that in contrast to the claims of ETF opponents, ETFs positively contribute to market efficiency, attenuating the mispricing of individual stocks, and some of the effects are indirect—by relaxing short-selling constraints through the stock lending activities of ETFs.

There is a common perception that the majority of ETFs are passive investments that mechanically purchase stocks to track specific indices. Accordingly, many view these "passive" investment vehicles as unable to contribute to price discovery because they do not discriminate between stocks within the index—they buy or sell all stocks in the index according to index weights irrespective of whether a stock is over or underpriced. Given that ETFs attract many uninformed investors, they may siphon away liquidity from the underlying

---

[14] See Statista (https://www.statista.com/topics/2365/exchange-traded-funds/).

[15] For example, "the more assets flow into index and ETF strategies, the more likely it is that market mispricing will arise for active managers to exploit" (State Street Global Advisors, 2018) and "The rise of passive—in adding unthinking investors to the market—surely makes the process (of information being reflected in prices) take rather longer than in the past. Markets . . . are becoming increasingly inefficient" (Financial Times, Aug 2018).

[16] See Sanford-Bernstein Research Report (August 23, 2016): "The silent road to serfdom: Why passive investing is worse than Marxism."

stocks and decrease the incentives for informed traders and active investors to expend resources on firm-specific information; for example, see Israeli et al. (2017) and Bond and Garcia (2019). According to this view, ETFs may cause general deterioration in the pricing efficiency of underlying securities.

In contrast, there are reasons why ETFs might have the opposite effect and might improve informational efficiency, including incorporating stock-specific information into prices. One reason is that ETFs now cover many narrow and highly specialized exposures and indexes. Investors can use these ETFs as the building blocks of active portfolios (Easley, Michayluk, O'Hara, and Putnins, 2021) or the hedging instruments in active single-stock bets (Huang, O'Hara, and Zhong, 2021). Such active uses of passive instruments can contribute to informational efficiency. Further, smart beta ETFs can make it cheaper and easier to trade on factors, which can encourage more efficient pricings of factor information. Finally, to generate ancillary revenue that can help reduce net fees, ETFs can lend a substantial proportion of their holdings to traders who want to short-sell stocks (Blocher and Whaley, 2015). These stock lending activities have the potential to contribute to market efficiency by reducing the constraints and costs of short-selling.

This chapter sheds new light on these contrasting views by empirically analyzing the impacts of ETFs on efficiency as reflected by the profitability of well known–cross sectional asset pricing anomalies. The focus on anomalies is motivated by the fact that they typically require long and short positions to exploit and therefore may reflect the impacts of ETF stock lending activities. Furthermore, anomalies reflect a variety of sources of mispricing and are therefore not specific to particular information types (e.g., accounting information). They also reflect longer-horizon inefficiencies than some of the intraday effects that have been documented in previous studies.

First, we examine the returns of 15 well-known asset pricing anomalies. For concision, the focus is on the three anomalies that reflect the general effects: Mispricing, Idiosyncratic Volatility (IV), and Momentum. We choose the Mispricing anomaly in a study by Stambaugh, Yu, and Yuan (2015) because it combines 11 anomalies and hence reflects an aggregate mispricing factor. The results for the IV and momentum anomalies are also reported because these anomalies can be constructed easily using stock prices, require no additional accounting information, and have been shown to be consistent over time. Results from other anomalies tend to be consistent, and they are therefore reported in Internet Appendix 3A.

We find that the anomaly returns show a distinct decline around the time that ETFs become prolific. While the long-short portfolios of these anomalies tend to be profitable from 1980 to 2002, the returns are much smaller and not statistically significant thereafter. Breaking the anomaly profits into the contribution from the long positions and the short positions ("long leg" and "short leg"), we show that the disappearing profitability is largely due to the short leg of the anomalies no longer producing abnormal returns. For example, shorting stocks with the highest level of IV in the previous month yields an average benchmark-adjusted return of 13.5% per year before 2003, but this drops to only 6.33% subsequently and is no longer statistically significant. In fact, as is shown, the short leg of these anomalies has always been responsible for most of the anomaly profits, so when the short leg ceases to be profitable, so does the anomaly as a whole.

However, one may wonder whether ETFs contribute to this change in the profitability/efficiency of the short leg. Miller (1977) argues that short-selling constraints play a crucial role in preventing rational trades from exploiting overpricing, suggesting that the profitability of the short leg stems largely from short-selling constraints. Consistent with this study's hypothesis that ETFs contribute to the decline in profitability, the findings show that the ETF ownership of underlying stocks substantially decreases the level of short-selling constraints, as measured by short-selling fees and the supply of securities able to be borrowed and shorted. The reduction in short-selling constraints allows informed investors to take advantage of overpricing using short-selling.

If the conjecture is correct in that by relaxing short-selling constraints, ETFs allow short-sellers to incorporate more information into prices, thereby reducing mispricing, it should be seen that not only are anomaly returns (particularly short legs) lower in the era of ETFs but also there are cross-sectional variations in the extents to which anomaly returns attenuate. Stocks that are widely held by ETFs should experience larger attenuations of anomaly returns than stocks that are not as widely held by ETFs. This study's results from double sorting stocks (by anomaly factors and by ETF holdings) support this conjecture. We find that anomaly returns are strongest in the quintile of stocks with the lowest levels of ETF holdings and decrease monotonically with the level of ETF holdings. Among stocks that are least widely held by ETFs, anomaly returns are statistically significant and economically meaningful, while among stocks are the most widely held by ETFs, anomaly returns are not statistically distinguishable from zero. These results suggest that mispricing remains persistent in stocks that are not widely held by ETFs. We repeat these tests using market capitalization as the second sort variable and confirm that this phenomenon is not driven by firms' sizes.

Finally, to rule out other confounding factors and examine whether the results reflect a causal effect from ETFs, we exploit exogenous variations in ETF holdings using index inclusion. Following an empirical design similar to that of other studies that have examined causal effects of ETFs, we use a stock's inclusion in heavily tracked indices, such as the Russell 1000 and 2000 and the Standard and Poor's (S&P) large-, mid-, and small-cap indices, as instruments for how widely the stock is held by ETFs. These tests control for stock size and non-linear transformations of stock size to capture only the discontinuities in ETF holdings around index inclusion thresholds. We find that ETFs play a causal role in attenuating anomaly returns and that much of this effect occurs through the reduction of short-selling constraints. However, inclusion in a major index can impact returns indirectly, perhaps through an increase in information availability and a large amount of capital tracking these indices. We perform several robustness tests to confirm that these effects do not impact the results.

While some mutual funds also engage in stock lending, they are not as active in this practice as ETFs (Blocher and Whaley, 2015; Prado et al., 2016). There are two main reasons. First, many mutual funds' investment goals include outperforming the benchmark index rather than just tracking it. Usually, this involves overweighting some stocks while underweighting others in comparison to the underlying benchmark index. Intuitively, managers do not lend out stocks that they overweight because doing so can put a downward pressure on prices and cause the fund to underperform in the index. Second, many mutual funds are restricted from lending activities due to their investment philosophies or regulations (Evans et al., 2017; Prado et al., 2016). In contrast, ETFs generally do not aim to beat the benchmark index and thus are not concerned with price fluctuations stemming from lending activities, as long as the values of ETFs and the underlying stocks are aligned.

This chapter makes two main contributions to the existing literature. First, the results contribute to a growing body of research on the impacts of ETFs. One side of the debate argues that ETFs can harm market quality. Basak and Pavlova (2013) and Krause et al. (2014) propose theoretical models in which passive investing increases stock volatility. Using the reconstitutions of the Russell 1000 and 2000 as exogenous events, Ben-David et al. (2018) find that higher ETF ownership tends to increase stock volatility because arbitrage transmits shocks from the ETF to the underlying stocks. ETFs have also been linked to higher trading costs and commonality in liquidity and returns (Agarwal et al., 2018; Da and Shive, 2013; Israeli et al., 2017). Bhattacharya and O'Hara (2016) contend that some assets are not as liquid as the ETFs that hold them, so some of the price discovery process happens at the ETFs rather than just in the underlying assets. Madhavan (2012) shows market fragmentation makes markets more fragile and susceptible to flash crashes and that when a flash crash occurs, ETFs can experience the greatest price movements due to the breakdown of the arbitrage mechanism.

However, several researchers document the positive impacts of ETFs, including improved price discovery (Hasbrouck, 2003; Ivanov et al., 2013) and higher turnover and thus better liquidity (Ben-David et al., 2018). Coles et al. (2017) show that while index investing adds noise to prices, it does not impede long-term price efficiency. Further, Huang et al. (2021) show that ETFs can contribute to price discovery by allowing investors to more efficiently hedge risks and therefore take more aggressive positions. While many of these studies focus on the direct impacts of ETFs, this chapter also considers the indirect effects of ETFs on market efficiency through their effects on short-selling constraints. By analyzing anomaly returns, the analysis is not limited to a specific form of information, such as earnings or accounting information, and longer-horizon mispricing is captured than in studies that focus on the intraday effects of the arbitrage mechanisms inherent in ETFs.

Second, this chapter is also related to extant literature on short-selling and asset prices. Previous studies find that short-sellers are sophisticated and informed investors who are better at scrutinizing information and anticipating bad news (Christophe et al., 2004, 2010; Engelberg et al., 2012; Karpoff and Lou, 2010). Miller (1977) argues that divergence in opinions and impediments to short-selling causes prices to be overvalued and reflect only optimistic views. Stambaugh et al. (2012) study a number of well-known anomalies and conclude that profit is stronger during periods of high sentiments consistent with impediments to short-selling, allowing stocks to become overpriced. Similar findings are have been documented in studies by Beneish et al. (2015) and Drechsler and Drechsler (2014), who show that short-selling constraints are more severe in anomalies' losers (stocks that are selected to be shorted) and prevent arbitrageurs from trading away the mispricing. Moreover, recent papers by Huang et al. (2021) and Li and Zhu (2016) claim that investors take short positions in ETFs rather than the short-constrained underlying securities, and this has positive effects on market efficiency.

This section's findings also help understand one of the drivers of the disappearance of return anomalies in recent times—for example, see Mclean and Pontiff (2016). Stambaugh et al. (2012) contend that impediments to short-selling are the main causes of overpricing, which is the main driver of anomalies. This study's results support this view and suggest that ETFs play a role in attenuating these impediments and thereby reducing the prevalence of overpricing.

The results have implications for regulators and other market participants. First, there are concerns about stock lending activities by ETFs as this poses a risk of declines in security prices beyond fundamental values. We show that security lending activities can have the opposite effect. Increasing the supply of stocks that can be borrowed and shorted improves price efficiency, enabling information to be impounded into price more effectively and hence preventing significant negative price shocks. Second, we shed light on the view that the growth in ETFs leaves more "dumb money" on the table for active funds. The results suggest the opposite. ETFs help lower short-selling constraints and thereby improve market efficiency, making the discovery of mispriced stocks more difficult for active managers than it would be otherwise.

3.2 Data and portfolio formation

3.2.1 Data

We obtain data on stock prices and returns from CRSP for 1980–2018. The sample includes stocks that are listed on the NYSE, NASDAQ, and AMEX exchanges. We exclude stocks with prices of less than $1.00 in their month of portfolio formation. One exception is the Mispricing anomaly, wherein the data from Stambaugh's website is available for only stocks with prices higher than $5.00. Thus, when we analyze this anomaly, stocks with prices lower than $5.00 are excluded.

We obtain data for ETF holdings from several sources. We identify ETFs by using CRSP share code 73. Information on the amount of each stock held by each ETF is obtained from the Thomson Reuters Mutual Fund Holding database. These amounts are matched with stock data from CRSP to compute ETF holdings as a proportion of the total market capitalization of each stock. Although the first ETF was introduced in 1993, the ETF market remained small until around 2003. Figure 3.1 illustrates that the total amount of assets under management by ETFs in the US surpassed $200 billion in 2003. Thus, 2003 is used as a starting year for the "high ETF activity" period.

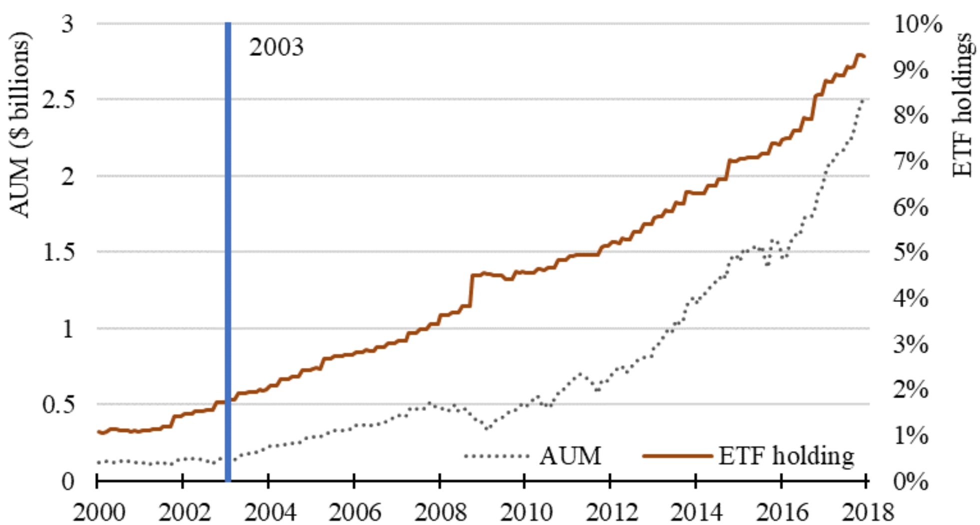


Figure 3.1: The growth of ETFs

This figure shows the growth of ETFs from 2000 to 2018. The dotted line shows ETFs' AUMs, which are the aggregate asset values held by ETFs in billions of dollars. The solid line shows value-weighted average ETF holdings, which are the proportions of stocks' outstanding shares held by ETFs. The vertical line indicates the start of the high-ETF growth period, which is January 1, 2003. The sample consists of AMEX, NASDAQ, and NYSE stocks and ETFs.

We obtain short selling and stock lending data from Markit Securities Finance. This database is commonly used in short-selling research. It covers almost 90% of the OTC security lending market (Ben-David et al., 2018). We obtain the supply of lendable shares for each stock ($Short\ supply$) and the fees for borrowing the stocks to short-sell them ($Shorting\ fee$). $Shorting\ fee$ is the annualized cost of borrowing the stock. The higher the fee, the harder it is to borrow the stock, perhaps due to a deficiency in the lendable supply and other restrictions. $Short\ supply$ is the active lendable number of shares as a proportion of the total outstanding shares. Thus, $1 - Short\ supply$ is a proxy for short-selling constraints. As the short-selling data are available beginning in 2006, the sample for the regression analysis starts in 2006.

Table 3.1 reports the descriptive statistics of the main variables used in this study. The average market capitalization of stock in the sample is \$4.6 billion, and its return is approximately 1% per month with an IV level of 2%. An average firm has 4.5% of its outstanding shares held by ETFs, whereas the highest ETF ownership in a firm is 15.5% in this sample. The average ETF holding has grown over time, from just under 1% in 2003 to almost 9% in 2016. The average cost of borrowing a stock is 1.5% per year. However, the distribution of fees is heavily skewed to the right. Hence, the log of one plus fee for the analysis. Additionally, approximately 15% of outstanding shares are available for borrowing, of which one-fifth are utilized (i.e., on loan). We winsorize the key variables at the 1st and 99th percentiles of the whole sample. An exception is the monthly return variable, which we winsorize at the 0.5 and 99.5 percentiles.

Table 3.1: Descriptive statistics

This table reports descriptive statistics of the main variables. The sample comprises ordinary stocks in the CRSP universe, which have short-selling and ETF holding data for 2003–2017. $Short\ supply$ is the lendable supply of a stock divided by its total number of shares, Fee is the indicative fee in the Markit database (log of one plus the borrowing cost of stock at the end of each month). P25 and P75 represent the 25th and 75th percentiles, respectively.

	Mean	STD	Min	P25	Median	P75	Max
Return	0.010	0.172	−0.356	−0.054	0.006	0.068	0.467
Market capitalization (\$ million)	4,647.62	20,433.57	0.148	134.08	512.28	2,062.19	882,331.55
$ETF\ holdings$	0.045	0.036	0.000	0.017	0.036	0.065	0.155
$1 - Short\ supply$	0.849	0.124	0.577	0.743	0.863	0.974	1.000
Fee	0.015	0.037	0.003	0.004	0.004	0.006	0.303

3.2.2 Anomalies

As aforementioned, to keep the presentation of results in this paper concise, the focus is on three well-known anomalies, namely the Mispricing factor, Momentum, and Idiosyncratic volatility. We choose these three anomalies for several reasons. The Mispricing factor parsimoniously captures the degree of mispricing across stocks by combining a variety of mispricing indicators, including accounting variables. Stambaugh, Yu, and Yuan (2015) show that this anomaly performs better than the individual anomalies, so one can view it as an aggregate anomaly. Further, the latter two anomalies require only stock prices, making them simple and easy to replicate; yet, as is shown, despite their simplicity, they still produce anomaly returns in the early parts of the sample and in the stocks with low ETF holdings. Additionally, these three anomalies tend to outperform other anomalies from over the past 40 years.

The Mispricing factor, introduced by Stambaugh, Yu, and Yuan (2015), combines 11 anomalies related to stock mispricing, many of which are based on accounting data. A higher value means a higher level of overpricing. Exploiting the anomaly involves selling stocks with the highest values of the mispricing factor and buying stocks with the lowest values. We obtain data on this factor from Robert Stambaugh's website.[17]

The return momentum anomaly, popularized by Jegadeesh and Titman (1993), refers to the observation that returns tend to be persistent, so high past returns tend to predict high future returns. While many horizons and ways of constructing momentum strategies have emerged in extant literature, we construct the momentum strategy by sorting stocks into deciles each month based on their six-month cumulative returns (month $t-6$ to month $t-1$), skipping the current month to avoid reversals and forming a long-short portfolio by buying the stocks with the highest past returns (the top decile) and selling the stocks with the lowest returns (the bottom decile).

The IV anomaly refers to the phenomenon that stocks with high past IVs tend to perform poorly in the future. We compute monthly IV for each stock following the approach used in Campbell et al. (2001): the sum of squared daily returns of stock in excess of the value-weighted CRSP index. We sort stocks into deciles based on IV and buy stocks that are

[17] http://finance.wharton.upenn.edu/~stambaug/

placed in the bottom decile while short stocks that are placed in the top decile. The holding period for all three anomalies is one month.

We also study 15 other prominent anomalies, namely Total accruals (Sloan, 1996), Asset growth (Cooper, Gulen, and Schill, 2008), Book-to-Market (Fama and French, 1992), Composite stock issue (Pontiff and Woodgate, 2008), Financial distress (Campbell, Hilscher, and Szilagyi, 2008), Dispersion in analyst forecasts (Diether, Malloy, and Scherbina, 2002), Gross profitability premium (Novy-Marx, 2010), Investment to assets (Titman, Wei, and Xie, 2004), Maximum daily return (Bali, Cakici, and Whitelaw, 2011), Net operating assets (Hirshleifer, Hou, Teoh and Zhang, 2004), Net stock issue (Loughran and Ritter, 1995), Ohlson's O-score (Ohlson, 1980), Fundamental health score (Piotroski, 2000), Return on assets (Wang and Yu, 2010), and Unexpected earnings (Rendelman, Jones, and Latane, 1982). We report key results for these anomalies in the Internet Appendix 3A.

3.3 Empirical analysis

3.3.1 Anomalies before and after the proliferation of ETFs

This study first investigates the trend in anomaly returns over time. The sample is divided into two periods: 1980–2002 (the pre-proliferation period) and 2003–2016 (the post-proliferation period). The year 2003 is chosen as the breakpoint because ETF activity and growth were minimal before that and picked up noticeably beginning in 2003. Additionally, the total equity managed by ETFs exceeded \$500 billion for the first time in 2003.

Table 3.2 reports the performance of the long-short portfolios that exploit the three anomalies. Panel B shows the average annualized returns in excess of the market. It is evident that the long-short returns diminish in the latter sample period for all the anomalies. Take the momentum anomaly, for example. Winners (stocks in the long leg) have an average 13.7% excess return per year before 2003. This number drops by more than 50% to 5.2% after 2003 and is not statistically different from zero in the later period. Losers (stocks in the short leg) have a similar trend, with excess returns almost zero in the post-2003 period. Consequently, there is a sharp decline in the profits of the long-short momentum strategy from an average of 13.6% per year before 2003 to 3.3% after 2003. The Mispricing and IV anomalies have similar trends. While the profitability of the Mispricing anomaly is still significantly different from zero in the later period, its magnitude decreases by almost 75%.

In Panel B of Table 3.2, we control for other risk factors such as the Fama-French (1993) market, size, and value factors, by regressing each anomaly's monthly excess returns on the factors and measuring the alpha (intercept). This procedure gives benchmark-adjusted returns. The same trends in anomaly profitability are apparent in the Fama-French alphas. The declines are not as big as in excess returns but are still noticeable. Before 2003, the Mispricing anomaly has an average benchmark-adjusted return of 27% per annum (p.a). This return is halved to about 10.5% after 2003.

While the returns in Table 3.2 show a clear decline in portfolio returns in the post-proliferation period, we are interested in whether this decline is statistically significant. Excess and benchmark-adjusted returns are regressed on D_{Post}, a dummy variable that takes the value of one if the observation is after January 1, 2003 and takes the value of zero otherwise. The coefficient estimates for this variable allow one to conclude whether the returns before and after 2003 are statistically and significantly different from each other. The results are reported in the last three columns of Table 3.2. The coefficients are statistically and significantly different from zero and negative for the long-short portfolios of the Mispricing and Momentum anomalies. The excess and benchmark-adjusted returns of the Mispricing anomaly are at least 17.3% lower than in the period before the proliferation of ETFs. This figure is similar for the Momentum anomaly, with a drop of about 16% in the long-short returns. Contributing to these drops is the lower returns of the long legs and higher returns of the short legs. This trend is present for the other 15 anomalies. A similar table for these anomalies is included in Internet Appendix 3A.

Figure 3.2 illustrates the cumulative returns of portfolios over time. We use the most conservative approach to calculate excess returns. In particular, we sum portfolio returns year by year—instead of allowing them to compound—to demonstrate the revolution of anomaly profit. It is clear that before 2003, the three anomalies have consistent returns. The Mispricing anomaly appears to have the best performance with a consistent and stable performance over time. For the Mispricing and Momentum anomalies, the long legs show an upward trend while the short legs show a gradual downward trend. This results in a steady increase in the long-short portfolio. The IV anomaly also exhibits an upward but volatile trend for the long-short strategy. However, the majority of these returns are attributable to the short legs. As the graphs illustrate, the long legs barely outperform the market. In contrast, the short legs show a consistent downward trend and contribute to the majority of the excess returns of the long-short portfolios. Interestingly, all the portfolios for the anomalies display plateaus after 2003. These findings suggest that the returns from the short legs are the primary drivers of anomalous returns but diminish after 2003.

Table 3.2: The anomaly returns before and after the ETF boom

This table reports the annualized returns (in %) of the long-short portfolio of each anomaly. The stocks are sorted into deciles based on the anomaly variable. We go long and short the stocks in the bottom or the top decile depending on the anomaly. The anomaly variables are the Mispricing factor from Stambaugh, Yu, and Yuan (2015); Idiosyncratic Volatility from Campbell, Lettau, Malkiel, and Xu (2001); and Momentum, which is the cumulative return from month $t-6$ to $t-1$. *Excess returns* (Panel A) are equally weighted average returns in excess of market returns. *Three factor alphas* (Panel B) are the intercepts from regressing portfolios' excess returns (over the risk-free rate) on the Fama-French's (1993) market, size, and value factors. The first three columns report the results for the period of 1980–2002 (pre-proliferation) whereas the next three columns report the results for the period of 2003–2016 (post-proliferation). The last three columns report the differences in performance between the pre- and post-proliferation periods. The differences are measured as the coefficient estimates of regressing portfolios' returns on D_{Post}. In Panel A, *Excess returns* on D_{Post} are regressed. In Panel B, the same regression is performed in the first six columns, and then D_{Post} is added to the set of three Fama-French's (1993) factors. The sample consists of stocks listed on AMEX, NASDAQ, and NYSE. Returns are reported in annualized percentages, and t-statistics are in parentheses using standard Newey-West standard errors with six lags. *, **, and *** indicate significance at the 10%, 5%, and 1%, respectively.

	1980–2002 (pre-proliferation)			2003–2016 (post-proliferation)			Difference (D_{Post})		
	Mispricing	IV	Momentum	Mispricing	IV	Momentum	Mispricing	IV	Momentum
Panel A: Excess returns (% p.a.)									
Long leg	16.381***	7.485***	13.738***	4.579***	3.351**	5.177	−11.802***	−4.134	−8.561
	(8.46)	(2.77)	(3.63)	(3.46)	(2.08)	(1.37)	(−4.94)	(−1.28)	(−1.61)
Short leg	−9.377***	−13.509***	−8.465**	−2.143	−6.330	−0.459	7.234*	7.180	8.006
	(−3.52)	(−2.87)	(−2.06)	(−0.66)	(−1.16)	(−0.08)	(1.74)	(1.00)	(1.14)
Long-short	25.758***	20.994***	13.630***	6.722**	9.681	3.345	−19.036***	−11.314	−16.567***
	(10.69)	(3.54)	(2.99)	(2.12)	(1.51)	(0.59)	(−4.76)	(−1.3)	(−2.71)
Panel B: Three factor alphas (% p.a.)									
Long leg	15.538***	7.678***	13.191***	4.203***	6.889***	2.140	−11.983***	−2.138	−11.279***
	(14.17)	(4.83)	(6.20)	(5.11)	(7.68)	(0.9)	(−7.88)	(−1.12)	(−3.26)
Short leg	−11.396***	−13.465***	−10.616***	−6.366***	−12.522***	−7.376**	5.293**	3.677	5.577
	(−5.76)	(−4.83)	(−3.56)	(−3.66)	(−4.45)	(−2.28)	(2.07)	(0.82)	(1.15)
Long-short	26.934***	21.143***	15.876***	10.569***	19.411***	8.831*	−17.276***	−5.814	−16.856***
	(12.20)	(5.86)	(3.25)	(5.17)	(6.72)	(1.93)	(−5.79)	(−1.12)	(−2.73)

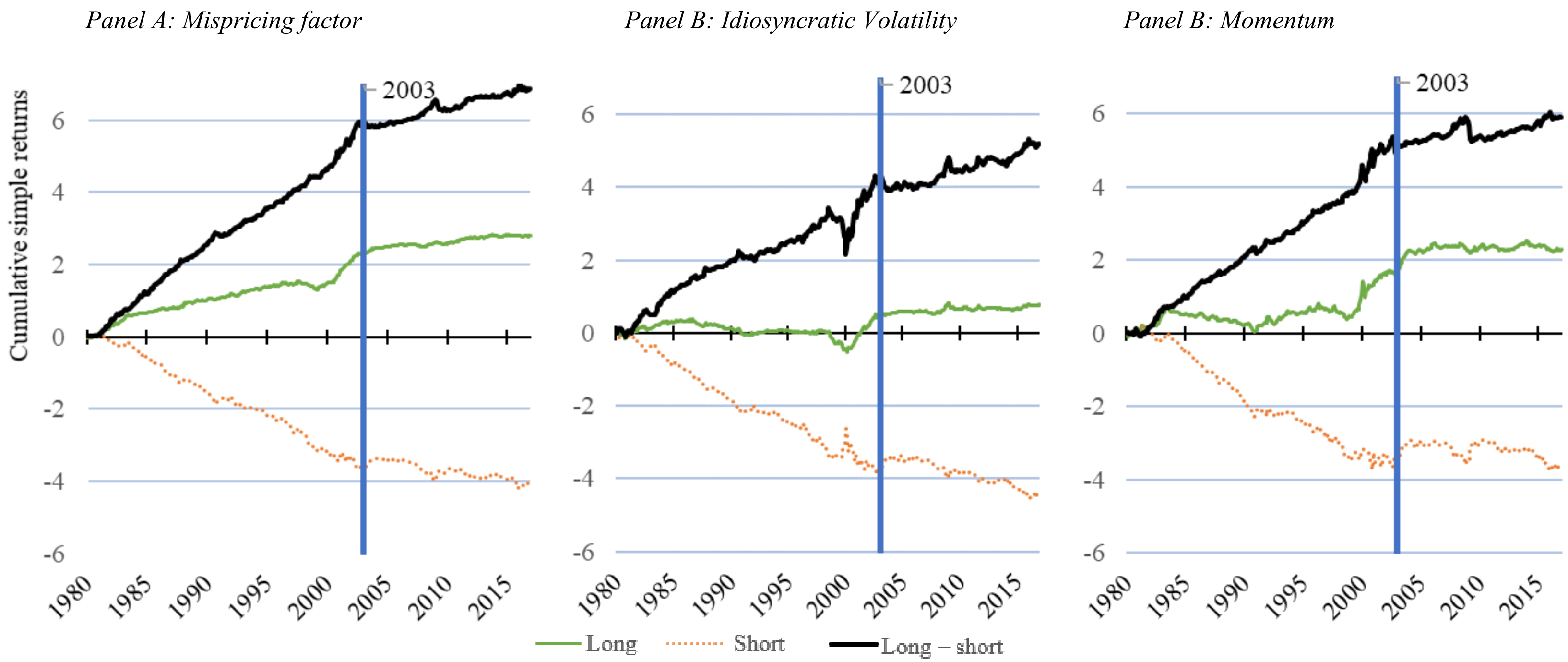


Figure 3.2: The cumulative returns of the anomaly portfolios

This figure shows the cumulative simple returns (in decimals and not compounded [i.e., equivalent to cumulative log returns]) of the long leg, short leg, and long-short portfolios in excess of market returns. The portfolio returns are monthly equally weighted average returns. The stocks are sorted into deciles based on three anomaly variables: the Mispricing factor (Panel A), which is from Stambaugh, Yu, and Yuan (2015); Idiosyncratic Volatility (Panel B), which is computed as in Campbell, Lettau, Malkiel, and Xu (2001); and Momentum (Panel C), which is the cumulative return from month $t-6$ to $t-1$. For the Mispricing and Idiosyncratic Volatility anomalies, the long leg involves buying stocks in the lowest decile and the short leg involves buying stocks in the highest decile, respectively. It is the opposite for the Momentum anomaly. The long-short portfolio is the difference in the returns of the long and the short legs. The vertical line indicates the start of the high-ETF growth period, which is January 1, 2003. The sample consists of AMEX, NASDAQ, and NYSE stocks from 1980 to 2016.

3.3.2 *Two-way sort*

To investigate the effects of ETF ownership through short-selling, we utilize two-way sorts. At the end of each month, we sort stocks into quintiles based on the above anomaly factors (Mispricing, Momentum, and IV). For each quintile in each month, stocks are further

Table 3.3: The long-short anomaly returns by ETF holdings using two-way sorts

This table reports the annualized percentage excess returns and the three-factor alpha for the long-short portfolios. The portfolios are formed each month by sorting stocks into quintiles based on an anomaly factor, and, within each quintile, sorting stocks into quintiles by ETF holdings, which is the fraction of outstanding shares held by ETFs. For each ETF quintile, long stocks are in the bottom anomaly quintile and short stocks are in the top anomaly quintile. Excess returns are equally weighted average returns in excess of T-bills. The three-factor alpha (α) is the intercept from regressing portfolio excess returns on the Fama-French (1993) market, size, and value factors. Panels A, B, and C report the long-short portfolio returns for the mispricing factor from Stambaugh, Yu, and Yuan (2015); the IV from Campbell, Lettau, Malkiel, and Xu (2001); and momentum, which is the cumulative return from month $t-6$ to $t-1$, respectively. $Size$ is the average market capitalization of stocks in each ETF quintile. Fee is the log 1 + borrowing cost of stock. Returns are reported as annualized percentages and t-statistics are in parentheses using Newey-West standard errors with six lags. The sample consists of stocks listed on AMEX, NASDAQ, and NYSE and covers the period of 1980−2016. *, **, and *** indicate significance at the levels 10%, 5%, and 1%, respectively.

	ETF quintile				
	(1)	(2)	(3)	(4)	(5)
Panel A: Mispricing					
Excess returns (% p.a.)	7.29***	1.57	1.46	0.92	0.95
	(2.64)	(0.49)	(0.41)	(0.36)	(0.36)
Three-factor alpha (% p.a.)	9.44***	5.33**	6.11***	3.97*	3.47
	(4.08)	(2.54)	(2.7)	(1.86)	(1.49)
Panel B: Idiosyncratic Volatility					
Excess returns (% p.a.)	8.03	−3.33	−0.59	−3.70	−4.12
	(1.33)	(−0.55)	(−0.10)	(−0.62)	(−0.70)
Three-factor alpha (% p.a.)	14.86***	3.77	8.63***	6.32**	5.54
	(4.22)	(1.17)	(3.30)	(2.30)	(1.47)
Panel C: Momentum					
Excess returns (% p.a.)	3.04	2.16	−1.19	−1.22	−5.15
	(0.82)	(0.44)	(−0.24)	(−0.24)	(−1.00)
Three-factor alpha (% p.a.)	5.38*	4.74	2.36	3.43	−0.74
	(1.66)	(1.19)	(0.60)	(0.88)	(−0.17)
Panel D: Stock characteristics					
$Size$ ($ million)	900.39	3,366.54	5,074.10	7,580.90	6,310.14
ETF holdings (%)	0.47	1.44	2.85	5.17	9.77
Fee (bps)	249.31	152.13	112.55	78.77	65.10

sorted into quintiles based on their ETF ownership measures, with the top quintile consisting of the stocks with the highest ETF holdings and the bottom quintile consisting of the stocks with the lowest ETF holdings. This procedure gives 25 portfolios each month. The reported portfolio returns are equally weighted. The value-weighted returns are also examined and reported in the Internet Appendix. The main focus is the stocks in the top and bottom quintiles of the first sort based on the anomaly factors because they belong to the long and short portfolios. Unlike the one-way sort procedure above, the sample period for this procedure is from 2003 to 2016 because of the availability of ETF data. Additionally, the main interest is the period with high ETF ownership that starts in 2003. Panels A, B, and C in Table 3.3 show the results for the mispricing, IV, and momentum anomalies while Panel D shows the statistics of stocks in quintiles sorted by ETF holdings.

ETF quintile one contains stocks with the lowest level of ETF holdings whereas ETF quintile five contains stocks with the highest level of ETF holdings. On average, about ten percent of stocks in the top quintile are held by ETFs while this figure is only a half of a percent for stocks in the bottom ETF quintile. Interestingly, the average market capitalization of each quintile does not change monotonically. The stocks with the highest levels of ETF holdings are not the largest stocks. ETFs appear to hold more medium to large stocks. One possible explanation is that there might be more ETF capital tracking medium-sized companies. Another possible reason is that the same amount of invested capital might make up a larger proportion of medium stocks than large stocks.

Consistent with the regression analysis in the following section, the stocks in the bottom ETF quintiles have the highest borrowing fees. The average cost of borrowing stocks in the bottom ETF quintile is 250 bps, five times higher than that of the top quintile. This figure decreases monotonically when the average ETF holding increases.

There appears to be a negative relation between ETF holdings and anomaly profitability. Excess returns and three-factor alphas increase monotonically from the bottom quintiles to the top quintiles sorted by ETF holdings. Take the Mispricing anomaly, for example; buying and selling stocks in the bottom ETF quintiles produces statistically significantly positive excess returns. The same strategy yields very low and insignificant returns for stocks with higher ETF holdings. When we consider Fama-French's (2003) risk factors, the returns become higher, but the trend remains similar in that the three-factor alphas decrease when ETF holdings increase. Overall, the trading of stocks with low ETF holdings generates the best returns. There are almost no abnormal returns for stocks that are heavily held by ETFs. These figures are similar to the Momentum and IV anomalies. This indicates high levels of mispricing among low ETF ownership stocks.

However, one might argue that this phenomenon might be due to size. Accordingly, ETFs tend to hold larger stocks because they must track indexes and the weights of these stocks are more important. While controlling for the Fama-French three factors might address this concern, we also tested this supposition by using size as a second sort instead of ETF ownership. However, we do not find a distinct difference in the long-short returns among size quintiles.

3.3.3 ETFs and short-selling constraints

ETFs can influence anomaly returns through different channels. In this section, we apply multivariate analysis to study the effects of ETF holdings on short-selling constraints and anomaly returns. Many factors can influence stocks' price discovery processes and arbitrage opportunities. For example, an increase in profitability or improvement in a stock credit rating can result in higher information coverage and a reduction of short-sale constraints. This, in turn, can impact the stock's price discovery and efficiency. To isolate the causality of ETF holdings on short sale constraints and their price efficiencies, we use a two-stage regression with an instrument variable approach. The instrument variable is the inclusion into the Russell 1000, Russell 2000, S&P 500, S&P 400, and S&P 600. Because these are the largest indices that many ETFs track, inclusion in one of these indices can cause an increase in ETF holdings in a stock. Russell 1000 and Russell 2000 are constructed purely by the market capitalization of companies and rebalanced once a year. While the S&P indices are constituted largely based on market capitalization, the index selection committee decides the constituents based on other characteristics, such as liquidity.

The first stage regression is:

$$ETF_{it} = \alpha + \beta_{Rus1000} D_{it}^{Rus1000} + \beta_{Rus2000} D_{it}^{Rus2000} + \beta_{SP500} D_{it}^{SP500} + \beta_{SP400} D_{it}^{SP400} + \beta_{SP600} D_{it}^{SP600} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it}, \quad (3.1)$$

where $D_{it}^{Rus1000}$, $D_{it}^{Rus2000}$, D_{it}^{SP500}, D_{it}^{SP400}, and D_{it}^{SP600} are dummy variables that take a value of one if stock i is included in the Russell 1000, Russell 2000, S&P 500, S&P 400, or S&P 600 at time t, respectively. $Controls_{it}$ includes a log of the market capitalization of firm i at time t and its square and cube, Fama and French's (1993) three factors, and LIQ, which is minus the log of one plus Amihud's (2002) illiquidity (ILLIQ). Since the focus is how ETFs influence stock returns conditional on the level of mispricing, anomaly variables are included in the second-stage regression. Thus, they are included in the first-stage regression as control variables to avoid the issues of inconsistent and biased error terms. There are two versions. In the first version, the anomaly variable is the actual value of the anomaly variable. In the second version, it is D_{it}^{L} and D_{it}^{S}, which are dummy variables that take the value one if a stock is included in the long leg or short leg of the anomaly at time t, respectively, and take the value zero otherwise. This allows for the examination of the distinct effects of ETFs on stocks in the long and short legs. Time fixed effects are also controlled for, which are denoted by δ_t.

Table 3.4 reports the results from Regression 3.1 for the Mispricing anomaly. The results for the other anomalies are excluded from the table because they are nearly identical. All the coefficients on the dummy variables are positive and statistically significant at the 1% level. This suggests that being included in these widely followed indices increases the proportion of outstanding shares held by ETFs. For example, inclusion in the Russell 2000 index causes a 1.6% rise in ETF ownership on average. Similarly, a stock in the S&P 600 index has a 3.1% higher ETF ownership compared to an average stock that is not part of this index. However, the figure for the S&P 500 is 2.4%, which is lower than that of the S&P 600. This might seem counterintuitive because the S&P 500 consists of the largest stocks. However, due to the difference in size, the same amount of ETF capital invested in an S&P 600 stock significantly intensifies the ETF ownership but marginally affects ETF ownership in an S&P 500 stock.

An interesting outcome is that the stocks in the short legs of the anomalies exhibit statistically significantly higher levels of ETF holdings. On average, ETFs hold approximately 1% more of these stocks. This might be due to factors and industry ETFs that hold these stocks to exploit mispricing. In contrast, stocks in the long leg show different numbers depending on anomaly. ETFs tend to hold more stocks with low levels of IV and fewer stocks with high past six-month returns.

Table 3.4: The first-stage regressions from the 2SLS models

This table reports the coefficient estimates from Regression 3.1, where i and t denote stock and year–month, respectively. ETF_{it} is the fraction of outstanding shares of company i held by ETFs. Instrumental variables are index inclusion dummy variables ($D_{it}^{Rus1000}$, $D_{it}^{Rus2000}$, D_{it}^{SP500}, D_{it}^{SP400}, and D_{it}^{SP60}) that equal one if stock i is included in the Russell 1000, Russell 2000, S&P 500, S&P 400, and S&P 600, respectively, in year–month t and that equal zero otherwise. $Controls_{it}$ are control variables that include [$\ln(mktCap)_{it}$] (the log of the market capitalization of firm i in year–month t) and its square and cube, Liq_{it} (the log [1 + ILLIQ]). Model 1 includes $Misp_{it}$, which is the value of the mispricing factor, while model 2 includes D_{it}^{L} (D_{it}^{S}), which is a dummy variable that equals one if stock i is included in the long (short) leg of the mispricing anomaly (these variables are included in second-stage regressions and therefore also included in the first stage). δ_t are time fixed effects. The sample consists of AMEX, NASDAQ, and NYSE stocks between 2003 and 2016. t-statistics are in parentheses using standard Newey-West errors with six lags. *, **, and *** indicate significance at the levels 10%, 5%, and 1%, respectively.

	(1)	(2)
$D_{it}^{Rus1000}$	0.003**	0.003**
	(2.33)	(2.34)
$D_{it}^{Rus2000}$	0.016***	0.016***
	(23.71)	(23.70)
D_{it}^{SP500}	0.024***	0.024***
	(22.47)	(22.48)
D_{it}^{SP400}	0.030***	0.030***
	(34.10)	(34.14)
D_{it}^{SP600}	0.031***	0.031***
	(56.88)	(56.99)
$Misp_{it}(\times 10^3)$	0.038***	
	(2.82)	
$D_{it}^{L}(\times 10)$		−0.005
		(−1.53)
$D_{it}^{S}(\times 10)$		0.011***
		(3.40)
Liq_{it}	0.002**	0.002**
	(2.29)	(2.26)
$\ln(mktCap)_{it}$	0.001	0.001
	(0.17)	(0.18)
$\ln(mktCap)_{it}^2$	0.001**	0.001**
	(2.28)	(2.28)
$\ln(mktCap)_{it}^3$ $(\times 10)$	−0.001***	−0.001***
	(−4.23)	(−4.23)
Time fixed-effects	Yes	Yes
R^2(%)	56.00	56.01
N	5,725	5,725

The short-selling constraint variables are also regressed on these dummy variables to understand the impacts of index inclusion on short-selling, but the results are not reported here. Short sale constraints for stocks in these large indices are lower. An average stock in our sample has a borrowing cost of 5.9% per year and virtually no share is available for short sales, while a stock in the S&P 500 index that has a borrowing cost of 4.7% and 5% of its outstanding shares can be lent out. Additionally, a negative relation between size and fee indicates that larger firms have lower borrowing fees and a higher proportion of active lendable shares, which can be lent out.

Next, we investigate the effects of ETF holdings on short-sell constraints. We use the fitted values of ETF holdings ($\widehat{ETF}_{it}$) from Regression (3.1) in the second-stage regression to isolate the effects of ETF holdings.

$$ShS_{it+1} = \alpha + \beta_{\widehat{ETF}} \widehat{ETF}_{it} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it}, \tag{3.2}$$

where ShS_{it} is a short-selling constraint variable for firm i at time t. This includes $Shorting\ Fee$ and $Short\ supply$. Because the distribution of fees is highly skewed to the right, their log values are used in regressions. $Short\ supply$ is a constraint variable that equals one minus the proportion of outstanding shares that are made available for short-selling. A higher value of this variable means a higher proportion of outstanding shares that are not available for borrowing.

We hypothesize that a negative relation exists between short-selling constraints and ETF ownership. The results from Table 3.5 confirm this hypothesis. The last two columns use the fitted value for ETF holdings from the first version of the first-stage regression where the actual values of the anomaly variable are used as control variables. In the last two columns, dummy variables for the long and short legs are used as control variables to estimate the fitted ETF holdings. The results from the two versions do not differ from each other much. Accordingly, all the coefficients on $\widehat{ETF}$ are negative and significant at a 1% level. Specifically, if ETF holdings increase by 1%, $Shorting\ Fee$ appears to decrease by almost 8% and $Short\ supply$ seems to rise by 2.59%. This is economically meaningful since the average value for $Short\ supply$ is about 15%. Larger and more liquid stocks appear to have lower short-selling constraints.

Table 3.5: The second-stage regressions for how ETF holdings impact short-selling constraints

This table reports the coefficient estimates from the following regressions:

$$ShS_{it+1} = \alpha + \beta_{\widehat{ETF}} \widehat{ETF}_{it} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it},$$

where i and t denote stock and year–month, respectively. The dependent variable is a short-selling constraint measure, ShS_{it+1}, which is either $Short\ supply_{it+1}$ (one minus the lendable supply of stock i divided by its total number of shares) or Fee_{it+1} (the log [1 + the borrowing cost of stock i]) in the year–month $t+1$. In the columns (1) and (2), ETF_{it} is the actual ETF holdings (OLS regressions). In columns (3) and (4), $\widehat{ETF}_{it}$ is the fitted value of the ETF holdings from the first-stage regressions (2SLS regressions). D_{it}^{L} (D_{it}^{S}) is a dummy variable that equals one if stock i is included in the long (short) portfolio of the Mmispricing anomaly in year–-month t and that equals zero otherwise. $Controls_{it}$ are control variables that include $[\ln(mktCap)_{it}]^{1\dots N}$(the log of the market capitalization of firm i in year–month t), its square and cube; and Liq_{it} $(-(1+\text{Illiq}))$. δ_t are time fixed effects. The sample consists of AMEX, NASDAQ, and NYSE stocks between 2003 and 2016. t-statistics are in parentheses using standard Newey-West standard errors with six lags. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) Fee_{it+1}	(2) $Short\ supply_{it+1}$	(3) Fee_{it+1}	(4) $Short\ supply_{it+1}$
ETF_{it}	−0.099***	−2.150***		
	(−9.21)	(−46.44)		
$\widehat{ETF}_{it}$			−0.079***	−2.580***
			(−7.12)	(−39.63)
$D_{it}^{L}(\times 10)$	−0.004*	−0.091***	−0.004*	−0.079***
	(−1.87)	(−7.43)	(−1.80)	(−5.48)
$D_{it}^{S}(\times 10)$	0.044***	0.115***	0.044***	0.093***
	(10.30)	(10.08)	(10.28)	(7.32)
Liq_{it}	−0.003**	−0.023***	−0.003**	−0.023***
	(−2.27)	(−9.48)	(−2.27)	(−9.50)
$\ln(mktCap)_{it}$	−0.007*	−0.036***	−0.007	−0.037***
	(−1.67)	(−3.20)	(−1.64)	(−3.20)
$\ln(mktCap)_{it}^{2}$	0.001	0.004**	0.001	0.004**
	(1.33)	(2.48)	(1.31)	(2.49)
$\ln(mktCap)_{it}^{3}$ $(\times 10)$	0.000	−0.001**	0.000	0.001**
	(−1.25)	(−2.17)	(−1.25)	(−2.18)
Time fixed effects	Yes	Yes	Yes	Yes
R^2(%)	4.81	53.03	4.10	43.03
N	5,416	5,402	5,416	5,402

So far, the results confirm our hypothesis that ETF holdings decrease the level of stocks' short-sell constraints due to their lending activities. Ultimately, the focus is the impact of ETFs on the efficiency of stock prices. An indicator of market inefficiency is an anomaly profit. An efficient market requires that stocks are correctly priced, and it is extremely difficult to predict stock returns and outperform the market. In contrast, if an investor can consistently make positive returns over market returns using a simple long-short strategy based on a variable, the market is inefficient.

3.3.4 ETFs and stock returns

Thus far, this chapter has investigated the effects of ETFs on short-selling constraints and the profitability of anomalies. In this section, we examine how ETFs impact stock returns. First, we regress stocks' next month returns on the anomaly variables to confirm the relation between the anomaly variables and stock returns. Next, we let an ETF variable to interact with these anomaly factors. Specifically, we run the following regression:

$$r_{it+1} = \alpha + \beta_1 A_{it} + \beta_2 \widehat{ETF}_{it} + \beta_3 A_{it}\widehat{ETF}_{it} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it}, \tag{3.3}$$

where A_{it} is the anomaly variable and $\widehat{ETF}_{it}$ is the fitted ETF holding from the first-stage regression. $Controls_{it}$ is a set of control variables that is used in the first-stage regression, including a measure for liquidity and market capitalization.

Table 3.6 reports the results from running Regression 3.3 on the three anomalies. In the first column of each anomaly, monthly returns are regressed on the anomaly variable and the set of control variables. The coefficient estimates of A_{it} are statistically and significantly negative as expected. For example, stocks with high IV during the month are expected to perform poorly the following month or stocks with high Mispricing factor values indicate overpricing and hence produce negative returns the following month. To ensure consistency and ease of interpretation with the other anomalies, the sign of the momentum factor is flipped so that the bottom decile is the long leg, and the top decile is the short leg. Thus, a negative sign of A_{it} is interpreted as stocks with high past six-month returns will continue to perform well the following month and vice versa.

Next, we added the fitted values of ETF holdings and their interactions with the anomaly variables in the regressions. For the IV anomaly, stocks with higher ETF holdings appear to have higher average returns. On average, a 1% increase in ETF holdings gives rise to about 0.54 % in the following month's returns. This increase is doubled when we analyze the Momentum anomaly. Interestingly, ETF holdings have negative impacts on stock

Table 3.6: The second-stage regressions for how ETF holdings impact anomaly returns

This table reports the coefficient estimates from the following regressions:

$$r_{it+1} = \alpha + \beta_1 A_{it} + \beta_2 \widehat{ETF}_{it} + \beta_3 A_{it}\widehat{ETF}_{it} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it},$$

where i and t denote stock and year–month, respectively. r_{it+1} is the monthly return and A_{it} is the value of the anomaly variable. $\widehat{ETF}_{it}$ is the fitted value of ETF holdings from the first-stage regressions (2SLS regressions). $Controls_{it}$ are control variables that include $[\ln(mktCap)_{it}]^{1\cdots N}$ (the log of the market capitalization of firm i in year-month t); its square and cube; and Liq_{it} (minus the log [1 + ILLIQ]). δ_t are time fixed effects. The anomalies are the mispricing factor (misp), which is from Stambaugh, Yu, and Yuan (2015); IV, which is computed as it is in a study Campbell, Lettau, Malkiel, and Xu (2001); and momentum (mom), which is the cumulative return from month $t-6$ to $t-1$. The sample consists of AMEX, NASDAQ, and NYSE stocks between 2003 and 2016. t-statistics are in parentheses using standard Newey-West errors with six lags. *, **, and *** indicate significance at the levels 10%, 5%, and 1%, respectively.

	Anomaly					
	Mispricing	Mispricing	IV	IV	Momentum	Momentum
A_{it} (× 10)	−0.001***	−0.002***	−0.200*	−0.769***	−0.017*	−0.114***
	(−6.04)	(−5.47)	(−1.75)	(−4.46)	(−1.81)	(−6.73)
$\widehat{ETF}_{it}$		−0.050*		0.054***		0.101***
		(−1.87)		(4.31)		(7.54)
$A_{it}\,\widehat{ETF}_{it}$		0.002***		1.760***		0.257***
		(3.58)		(4.27)		(7.83)
Liq_{it}	0.004***	0.004***	0.008***	0.007***	0.008***	0.007***
	(5.42)	(5.34)	(9.89)	(9.31)	(9.31)	(8.59)
$ln(mktCap)_{it}$	−0.002	−0.005	−0.022***	−0.025***	−0.021***	−0.023***
	(−0.72)	(−1.42)	(−6.54)	(−7.37)	(−6.20)	(−6.69)
$ln(mktCap)^2_{it}$	0.000	0.001	0.003***	0.003***	0.003***	0.003***
	(0.59)	(1.11)	(6.00)	(6.50)	(5.69)	(5.83)
$ln(mktCap)^3_{it}$ (× 10)	0.000	0.000	−0.001***	−0.001***	−0.001***	−0.001***
	(−0.65)	(−0.99)	(−5.66)	(−5.83)	(−5.37)	(−5.16)
Time fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (%)	0.02	0.03	0.03	0.03	0.01	0.05
N	5,722	5,722	6,927	6,927	6,859	6,859

returns when the regressions are run for the mispricing anomaly. On average, stocks with higher ETF holdings have lower returns of about 50 bps the following month. However, the statistical significance of this figure is only 10%.

The main focus is the interaction of $\widehat{ETF}_{it}$ and A_{it}. All coefficient estimates of this interaction are statistically and significantly positive. This suggests that ETFs help attenuate anomaly returns. In other words, a higher level of ETF holdings mitigates the negative relation between the anomaly variable and stock returns. For example, if a stock is in the top decile sorted by IV and only 1% of its shares are held by ETFs, it is expected to have a negative 2% return the following month. However, a 0.5% increase in ETF holdings reduces this negative return by about a quarter.

However, the analysis above might capture the attenuation effect of ETFs, but this effect has been lessened because of stocks in the middle of the anomaly variable distribution. We suspect that the impacts of ETFs on stock returns through the short-selling channel are strongest in the short leg of the anomalies. To investigate this, we use a similar procedure with a focus on stocks in the long and short legs. Particularly, we run the following regressions:

$$r_{it+1} = \alpha + \beta_1 D_{it}^L + \beta_2 D_{it}^S + \beta_3 \widehat{ETF}_{it} + \beta_4 D_{it}^L \widehat{ETF}_{it} + \beta_5 D_{it}^S \widehat{ETF}_{it} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it}, \tag{3.4}$$

where D_{it}^L and D_{it}^S are dummy variables that take a value of one if stock i is included in the long or short leg in month t, respectively. Other variables remain the same as in regression 3.3. We include the long and short leg dummy variables to confirm the presence of abnormal returns and inspect the sensitivity of stock returns to anomaly factors (i.e., Momentum and Idiosyncratic volatility). We then include an interaction term between these dummy variables and the fitted ETF holdings from the first stage to investigate the relation between ETF holdings and stock returns conditional on the anomaly. If ETF holdings help mitigate anomalous returns, it is expected that the coefficient on this interaction variable will have a sign opposite that of the dummy variable. This allows for the examination of the different effects of ETFs on stocks in the long and short legs.

Table 3.7 reports the results from running Regression 3.4. We separate our regression into different anomalies and only report the models with time fixed effects. The positive (negative) and significant coefficient estimates for D_{it}^L (D_{it}^S) confirm the presence of anomaly returns. Accordingly, stocks in the short legs exhibit significantly lower returns than those in the long legs. The stocks in the short leg of the Mispricing anomaly earn about 4% per year lower than the other stocks on average. The coefficient estimate is negative for the long leg dummy, but it is small and statistically not different from zero. Interestingly, the long portfolio of the IV anomaly experiences small but statistically significant negative returns. The short leg exhibits a larger negative return that equates to about −3.25% per annum. Overall, except for the Momentum anomaly, wherein returns of stocks in the long and short legs are not much different from other stocks, larger negative returns from the short legs indicate that the majority of the anomaly profits come from the short legs.

Next, the fitted value of ETF and its interactions with D_{it}^{L} and D_{it}^{S} are added. The coefficient estimates for $\widehat{ETF}_{it}$ are statistically and significantly positive. This suggests that being held by ETFs increases stocks' next month returns. Take the IV anomaly as an average; a 1% increase in ETF holdings causes a 0.78% improvement in returns per year. The magnitude of this impact depends on which anomaly is being studied. This figure is almost doubled when we analyze the Momentum anomaly. The increase in returns might be caused by buying pressures from ETFs that cause prices to surge. An inclusion in ETFs might also give rise to a stock's popularity and information coverage, which might in turn reduce its risk and improve returns. Overall, stocks with better liquidity earn better returns while bigger stocks appear to earn slightly lower returns.

ETFs appear to have different impacts on stocks in the long and short legs of anomalies. The coefficient estimates of the interactions between D_{it}^{S}, D_{it}^{L}, and $\widehat{ETF}_{it}$ are positive and statistically significant except for the IV anomaly in which the t-stat is high but not high enough to reject the null hypothesis of a coefficient of zero. These coefficient estimates have a sign opposite to those of D_{it}^{S}. This suggests that ETFs help attenuate anomalous returns. Among the stocks in the short legs, the ones with higher ETF holdings appear to have lower negative returns than the others with lower ETF holdings. Take the mispricing anomaly, for example. On average, a stock in the short leg with no ETF holdings offers a return of 0.54% per month lower than an average stock. This approximates to about −6.5 % per year. A one standard deviation change in ETF holdings of 3.57 % improves the stock's returns by 3.4% per year. If 5% of the outstanding share of this stock is held by ETFs, its average return is not statistically different from that of an average stock. This suggests that ETFs help eradicate mispricing and anomalous returns.

The momentum anomaly has similar statistics. A stock that is not held by any ETFs in the short leg earns approximately 2.5% per year lower than an average stock with a similar size and liquidity. Remarkably, a 1% increase in ETF holdings can eliminate these abnormal negative returns. In contrast, a stock with no ETF holdings in the long leg earns a 6.5% extra return per year. A one standard deviation increase in ETF holdings staggeringly wipes out three-quarters of the excess returns. If the stock has the average ETF holding across all stocks of 4.5%, there are no abnormal returns left to be made, unlike an average stock in the market.

Table 3.7: The second-stage regressions for how ETF holdings impact long-short portfolio returns

This table reports the coefficient estimates from the following regressions:

$$r_{it+1} = \alpha + \beta_1 D_{it}^{L} + \beta_1 D_{it}^{S} + \beta_1 \widehat{ETF}_{it} + \beta_1 D_{it}^{L}\widehat{ETF}_{it} + \beta_1 D_{it}^{S}\widehat{ETF}_{it} + \beta_C Controls_{it} + \delta_t + \varepsilon_{it},$$

where i and t denote stock and year–month, respectively. r_{it+1} is the monthly return and D_{it}^{L} (D_{it}^{S}) is a dummy variable that equals one if stock i is included in the long (short) portfolio of the anomaly and that equals zero otherwise. $\widehat{ETF}_{it}$ is the fitted value of ETF holdings from the first-stage regressions. $Controls_{it}$ are control variables that include $[ln(mktCap)_{it}]^{1 \cdots N}$ (the log of the market capitalization of firm i in year–month t); its square and cube; and Liq_{it} (minus the log [1 + ILLIQ]). δ_t are time fixed effects. The anomalies are the mispricing factor (misp), which is from Stambaugh, Yu, and Yuan (2015); IV, which is computed as it is in a study by Campbell, Lettau, Malkiel, and Xu (2001); and momentum (mom), which is the cumulative return from month $t-6$ to $t-1$. The sample consists of AMEX, NASDAQ, and NYSE stocks between 2003 and 2016. t-statistics are in parentheses using standard Newey-West errors with six lags. *, **, and *** indicate significance at the levels 10%, 5%, and 1%, respectively.

	Anomaly					
	Mispricing	Mispricing	IV	IV	Momentum	Momentum
$D_{it}^{L}(\times 10)$	−0.005	0.002	−0.008***	−0.022***	0.004	0.054***
	(−1.39)	(0.22)	(−2.63)	(−3.22)	(0.92)	(5.25)
$D_{it}^{S}(\times 10)$	−0.033***	−0.054***	−0.027***	−0.037***	−0.001	−0.020*
	(−6.96)	(−5.79)	(−4.38)	(−3.64)	(−0.26)	(−1.80)
$\widehat{ETF}_{it}$		0.033**		0.064***		0.097***
		(2.42)		(5.09)		(7.14)
$D_{it}^{L}\widehat{ETF}_{it}$		−0.014		0.030**		−0.110***
		(−1.07)		(2.56)		(−6.26)
$D_{it}^{S}\widehat{ETF}_{it}$		0.047***		0.032		0.043**
		(2.80)		(1.46)		(2.18)
Liq_{it}	0.004***	0.004***	0.008***	0.007***	0.008***	0.007***
	(5.60)	(5.54)	(9.92)	(9.39)	(9.27)	(8.67)
$ln(mktCap)_{it}$	−0.003	−0.005	−0.022***	−0.022***	−0.021***	−0.023***
	(−0.74)	(−1.44)	(−6.55)	(−6.61)	(−6.16)	(−6.46)
$ln(mktCap)_{it}^{2}$	0.000	0.001	0.003***	0.003***	0.003***	0.003***
	(0.58)	(1.10)	(5.96)	(5.68)	(5.65)	(5.61)
$ln(mktCap)_{it}^{3}$ $(\times 10)$	0.000	0.000	−0.001***	−0.001***	−0.001***	−0.001***
	(−0.61)	(−0.93)	(−5.60)	(−4.99)	(−5.34)	(−4.95)
Time fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
$R^2(\%)$	0.02	0.03	0.03	0.03	0.01	0.03
N	5,722	5,722	6,927	6,927	6,859	6,859

While the coefficient of the interaction term between $\widehat{ETF}_{it}$ and D_{it}^{S} is not statistically significant for the IV anomaly, its direction is similar. Its sign is the opposite of the coefficient of D_{it}^{S}. The coefficient of the interaction between $\widehat{ETF}_{it}$ and D_{it}^{L} also has a sign opposite that of D_{it}^{L} and is statistically significant. This implies that ETF reduces the effects of being in the long and short legs of the anomaly.

There are several channels through which ETFs are able to influence anomaly returns. One channel is smart beta ETFs, the number of which has increased significantly with a growth rate of 30% and an AUM of almost $1 trillion in 2017.[18] Smart Beta ETFs are investable products that track an index or a group of stocks with the aim of harvesting returns from risk factors or anomalies, such as size, quality, IV, and Momentum. These ETFs include single-factor and multi-factor ETFs that are widely used by institutional investors.[19] While single-factor ETFs help facilitate factor timing and reduce factor returns, it has been claimed that multi-factor ETFs are less sensitive to cycles and are more stable for picking up factor premiums. A consequence of this growth is that anomaly profits might disappear quicker as investors find more efficient and cheaper ways to exploit mispricing.

Another channel through which ETFs help eliminate market inefficiencies that take the forms of anomalies is providing an efficient way to hedge risks. Huang, O'Hara, and Zhong (2021) document that investors use industry ETFs to hedge against their bets on stocks. Accordingly, industry ETFs have higher short interests than other ETFs and there is a negative relation between industry ETF membership and market reactions to positive earnings news. This implies that more information is incorporated into prices before earnings announcements. Without these industry ETFs, such risky bets by investors might not happen. This can cause a delay in the incorporation of information into prices and an increase in mispricing and anomalies.

3.3.5 Robustness tests

This section performs several robustness tests. First, one-year lag values of ETF holdings are used as instrumental variables to avoid a simultaneity bias. This method is used by Spilimbergo (2009) and Hayo, Kutan, and Neuenkirch (2010). ETF holdings for the

[18] https://www.etfstream.com/news/smart-beta-etf-growth-extraordinary-claims-morgan-stanley-research/
[19] The Economist, Financial Times, and Morningstar.

previous year can predict 90% of the current year's holdings. There is a statistically significant and negative relation between ETF holdings and short-sell constraints. When ETF holdings in a stock increase, the supply for lendable shares of that stock also rises and its shorting fee decreases. We then include ETF holdings and their interaction terms with anomaly factors in the regression of stock returns on the anomaly factors. These regressions are similar to the main set of regressions in the previous section, except that we use anomaly factors (past six-month returns and idiosyncratic volatility) instead of dummy variables the long and short legs. The results are consistent with the analysis above, confirming the hypothesis that ETFs attenuate anomaly returns through short-selling constraints.

However, there might be concerns that the ETF effects that are being studied might be primarily caused by the index-inclusion effect. Stocks that are included in indices might receive more attention and scrutiny and hence improved pricing efficiencies. To address this concern, we tested different versions of the first-stage regression. We include interaction terms between the index inclusion dummy variables in Regression 3.1 above with a dummy variable, D_{it}^{Post}, that takes a value of one if the observation is after 2010 and a value of zero otherwise. We choose 2010 because it is the middle of the main sample period. The later period has a higher level of ETF holdings on average than the earlier half. If the effects are mainly related to the fact that stocks are included in the indices, we expect statistically insignificant coefficients for the interaction terms. However, we find the opposite. All the coefficients are statistically significant and positive. This suggests that inclusion in these indices does indeed increase stocks' ETF ownership. Additionally, we let the index inclusion dummy variables to interact with the average ETF holdings across stocks in that month similar to Da and Shive (2018). The results are consistent with the above versions and our hypothesis.

Next, we test the effects of ETF holdings on short-sale constraints and anomaly returns at the portfolio level rather than the stock level. For each portfolio, we calculate the simple average values for the variables of interest (ETF holdings and short-sale constraints) and control variables. We perform a similar analysis and the results are consistent with our hypotheses. The average values for short-sale constraints are negatively related to the ETF holding levels of portfolios. In particular, quintile portfolios with high ETF holdings exhibit lower levels of short-sell fees and a higher supply of lendable shares. This relation is strongest in the short legs of the anomalies studied here.

3.4 Conclusion

Short-selling constraints impede the abilities of arbitrageurs to exploit overpricing; the constraints therefore harm market efficiency. This chapter demonstrates an unintended consequence of the rise in ETFs: by lending stocks, ETFs diminish impediments to short-selling, which attenuates anomalies and thereby increases market inefficiency. The profits from the long-short strategy that exploits mispriced stocks shrink in line with the growth of ETFs. Abnormal returns persist for only stocks with low ETF ownership levels. However, short-selling is not the only channel through which ETFs can affect anomaly returns. There are other channels that play important roles in enabling ETFs to impact market efficiency as well. Further research on these channels will add insights to our understanding of ETFs on the overall market quality.

4. Passive price pressure

4.1 Introduction

ETFs have grown significantly over the last 20 years. In 2000, the entire ETF ecosystem comprised fewer than 100 ETFs and approximately $63 million assets under management. By 2020, there were 2,200 ETFs managing $5.5 billion worth of stocks.[20] This is equivalent to an average annual growth rate of 50% over the last 20 years. Considerable financial capital has moved out of active mutual funds and into ETFs and passive index funds.

The large dollar amounts now passively managed and closely tied to specific indexes have raised substantial concerns regarding the impact of this trend on price discovery and market efficiency. In particular, as ETFs become large, so do the magnitudes of their inflows and outflows, which are passed on as buying or selling of the underlying stocks. Unlike traditional mutual funds, the secondary market liquidity that ETFs offer their investors makes them more susceptible to large inflows and outflows; therefore, their effects on the underlying stocks in their portfolios are also likely to be more substantial. Furthermore, when a stock is added or deleted from an index, the ETFs tracking that index are forced to buy or sell large quantities of the particular stocks in a short time due to their mandates to mechanically replicate the index constituents. Such events heavily strain the liquidity of the underlying stocks and can lead to substantial distortions or dislocations in the prices of the underlying stocks. The larger ETFs become, the greater are the potential distortions. This chapter investigates these concerns by analyzing how changes in ETF holdings of underlying stocks impact stock prices.

As an example, when the heavily followed Russell 1000 and 2000 indices were rebalanced in June 2019, $170 billion worth of stocks were traded within a single day.[21] Such trading might have created significant order and liquidity imbalances and caused stock prices to temporarily divert from their fundamental values, known as "price pressure." As described in Chapter 1, the rebalancing of the Russell indices caused the returns of many inclusions and exclusions to change by more than 5% within a few minutes.

---

[20] https://www.statista.com/

[21] https://www.wsj.com/articles/russell-rebalancing-brings-frenzy-to-a-summer-friday-11561636806

Petajisto (2011) shows an index premium effect when a stock is added or removed from the S&P 500 index. Stocks that get added to this index enjoy a premium of approximately 8%, whereas those that are removed suffer a discount of approximately 14% just a few days prior to the reconstitution. The premium and discount are partly attributable to the buying and selling pressure from funds that track the index. Because ETFs have been growing considerably over the past decade, they might amplify the price pressure around index rebalancing events.

We measure ETF trading by using the change in the number of shares held by ETFs for each stock. This change is further partitioned into two components: "flow" and "rebalancing." "Flow" is caused by investment flows into and out of an ETF. Authorized participants (APs) ensure the balance between the ETF size and the underlying basket of stocks through the creation and redemption process. When an ETF receives capital (inflow), it (or the APs) buys the underlying assets. The opposite is true when an ETF experiences a large outflow. In contrast, "rebalance" results from the rebalancing of an ETF's underlying index. When a stock is added or removed from an index, an ETF tracking the index must buy or sell the stock. Both these mechanical trading principles can create short-term price pressure due to short-term order and liquidity imbalances.

This chapter shows that changes in ETF holdings of stocks create price pressure; however, the strongest effects are not observed around index rebalancing events but rather around large unexpected in/outflows. The price pressure associated with ETF in/outflows has both temporary and permanent impacts, indicating that flows have information and that they temporarily distort prices due to imperfect liquidity. In contrast, changes caused by rebalancing events contain little information and do not significantly distort stock prices, contrary to some concerns about ETFs.

We present two potential reasons behind the disappearance of index premium and discount. First, ETFs now hold the cross-section of stocks more uniformly. Stocks that get added to a widely followed index are often removed from another popular index and do not necessarily yield an increase in passive holdings. For example, stocks that are moved from the S&P 400 index (mid-cap) to S&P 500 index (large-cap) and those that are moved from the Russell 2000 index (small-cap) to Russell 1000 index (large-cap) experience a large drop in ETF holdings due to a fall in their weights. These stocks experience selling pressure by

ETFs, which decreases their prices. This effect is the opposite of that incurred by the index premium, which inflates prices when the stocks are included in the S&P 500 and Russell 1000 indices. As a result, the two forces cancel out each other, leaving little distortions in prices. The opposite is true for the index discount. Second, index rebalancing events are significantly anticipated. With an increase in index arbitrage and forecasting of index changes, the potential distortionary effects of passive trading are attenuated by arbitrageurs, which effectively act as liquidity providers around rebalancing events.

To illustrate passive price pressure, we conduct an event study of large changes in ETF holdings. Specifically, we examine the cumulative simple returns in excess of market returns around the events, which are the days of extreme changes. The results show that stocks experience a sharp increase (decrease) in returns of approximately 2% within three days around the event day. A potential explanation for this is the price pressure created by an unexpected surge in demand (supply) for ETF stocks. This distortion starts to revert after the event and takes approximately 20 to 40 days to stabilize. However, the stock prices do not revert to the previous level, which suggests that changes in ETF holdings contain information. We also show a positive flow–performance relation between stock returns and changes in ETF holdings. In particular, there is an increase in ETF holdings following a stock's positive returns and a decrease in ETF holdings following a stock's negative returns. This finding is consistent with those obtained by the existing literature on the relation between fund flows and the performance of mutual funds and ETFs (Easley, Michayluk, O'Hara, and Putnins, 2021; Lou, 2012).

To illustrate how economically meaningful the price pressure is around ETF flows, we form long-short portfolios based on the daily changes in ETF holdings of stocks. Each day, we compute the change in the proportion of the stocks' outstanding shares held by ETFs (ETF holding). The long leg includes stocks that suffer an outflow of more than 0.1%, whereas the short leg includes stocks that experience an inflow exceeding 0.1% of the outstanding shares. The holding period is two months. On average, each leg of this strategy contains approximately 200–300 stocks daily. This simple long-short strategy offers an average of 5.3% equally weighted annualized returns after controlling for the three factors proposed by Fama and French (1993) (i.e., market size, and value). This abnormal return is statistically significant at the 1% level. At the end of the study period, this strategy earns approximately

50% more than the overall market, which suggests that ETF flows can impose significant material costs for ETF investors.

This chapter makes several contributions to the literature. First, it contributes to the on-going debate on the impacts of ETFs on market quality. The results show that ETFs help attenuate the temporary distortions in prices created by index premiums. If the presence of the index premium indicates market inefficiency, its disappearance suggests a positive impact of ETFs, which helps to remove market friction. This result contradicts the claim made by market commentators that ETFs can exacerbate the impact of indexing on stock returns. Second, the findings indicate that ETF flow contains information that fundamentally affects stock prices. Many investors nowadays gain exposure to stocks by investing in the ETFs that hold the stocks rather than the stocks directly. Thus, ETFs are a channel through which investors express their valuation and contribute to stock price formation. Finally, the findings indicate the material cost borne by investors in ETFs and the underlying stocks when ETFs receive large in/outflows. Therefore, investors should be aware of the subsequent passive price pressure and strategically adjust their portfolios.

4.2 Related literature and hypotheses

This chapter is related to two strands of literature: price pressure and ETFs. It discusses each of these bodies of literature in turn and formulates hypotheses to guide the empirical analysis.

4.2.1 ETFs

The first body of literature is the growing number of studies on ETFs. The competition for funds and performance has led investors to seek low-cost products, such as index funds and ETFs. There is an ongoing debate on the effect of ETFs on market efficiency.

There is evidence that ETFs can improve market quality. Madhavan (2012), Lettau and Madhavan (2016), and Madhavan and Sobczyk (2016) show that ETFs are a cost-effective method for investors to make directional bets on market-wide or index-related information. The application of this strategy would have been difficult if investors had to trade the underlying assets individually. The creation and redemption mechanism, in which APs and arbitrageurs ensure that the prices of ETFs and the values of their underlying assets are aligned, helps to incorporate this information into the underlying asset prices. Huang, O'Hara, and Zhong (2018) show that investors use industry ETFs to hedge against their bets on stocks. Without these industry ETFs, such risky bets by investors are not possible. This impediment can cause a delay in the incorporation of information into prices and exacerbate mispricing.

Anecdotal evidence suggests that ETFs engage in security-lending activities to lower the management fees and that short-selling plays a vital role in the price discovery process. In addition, extensive literature has contended that short-sale constraints can have a material impact on asset prices. Short-selling constraints restrict the ability of investors to incorporate unfavorable news into prices and leave them overvalued [see Miller (1977), Harrison and Kreps (1978), and Diamond and Verrecchia (1987) for theoretical models; see Diether, Lee, and Werner (2009), Boehmer, Jones, and Zhang (2013), Beber and Pagano (2013), and Stambaugh, Yu, and Yuan (2012) for empirical evidence]. The previous chapter showed that ETFs do indeed lend out securities and help reduce short-sale constraints, which, in turn, improves market efficiency by reducing stock mispricing.

In contrast, other researchers have asserted that ETFs are harmful to the market. Trading in ETFs can introduce nonfundamental volatility to asset prices because some types of information disproportionally affect stocks of an index to their weight. Thus, trading at the ETF level can severely influence the constituents' prices and reduce the informativeness of stock prices. Bhattacharya and O'Hara (2018) contend that some assets are not as liquid as ETF holdings; thus, the price discovery process occurs at the ETF level. Market makers, who have imperfect learning about the underlying assets' fundamentals, infer this information from ETF holdings. The outcomes are propagation of shocks, herding behaviors, and market instability.

In addition, ETFs can increase stock volatility. A liquidity shock at the ETF level can disseminate to the underlying constituents and add noise to prices. In the models proposed by Greenwood (2005) and Gromb and Vayanos (2010), arbitrageurs buy and sell ETF shares and hedge their position by trading the underlying securities. Although the underlying value might remain constant, this trading can cause prices to fluctuate and diverge from their fundamental values. To the best of our knowledge, Ben-David, Franzoni, and Moussawi (2018) are the only ones to have directly studied the relation between ETF ownership and stock volatility. Their paper uses a regression discontinuity design that exploits the mechanical index rebalancing of the Russell 1000 and 2000 indices and confirms that one standard deviation increase in ETF ownership results in a 16% increase in daily stock volatility. Brown, Davies, and Ringgenberg (2020) show that ETF flows create a nonfundamental demand that severely influences stock prices.

There is also a growing body of research on the relation between ETFs and corporate finance. While active managers play an important part in a company's policy, investment decision, and corporate governance due to their primary objective of delivering great returns, it is unclear whether passive managers (including ETFs) act similarly. Passive managers' primary objective is to closely track the underlying index, and they are not much worried about the performance of the underlying stocks. Thus, they have little motivation to spend time and resources in the management of the underlying companies.

Several authors have found a positive effect of passive managers on corporate stewardship. Boone and White (2015) exploit the exogenous change in institutional ownership when the Russell 1000 and 2000 indices are rebalanced. They find that as index fund ownership increases, companies improve their reporting transparency. Using the same natural experiment, Appel, Gormley, and Keim (2016) show that passive investors are not passive owners. Rather, they are active in promoting strong corporate governance in companies that they hold through their voting blocs. Accordingly, firms with a higher number of passive investors have more independent directors, removal of poison pills and restrictions on shareholders' ability to call special meetings, and fewer dual-class share structures.

4.2.2 Price pressure

Buying or selling flows can create price pressure that causes a stock's market price to deviate from its fundamental value. According to the price pressure hypothesis of Scholes (1972), a price can depart from its prevailing value (assumed to be the efficient price) after a large buying or selling order because of a temporary shortage of liquidity. This temporary change in prices is the compensation for liquidity provision by the counterparty. If large orders are purely liquidity-motivated, as funds need to rebalance their portfolios, prices will revert to the levels immediately before the trade. Harris and Gurel (1986) and Shleifer (1986) support this hypothesis by providing evidence of sudden price increase following the S&P 500 inclusion, but this increase is completely reversed after some days. Mitchell, Pulvino, and

Stafford (2004) show that uninformed shifts in excess demand cause significant price pressure around mergers, which are quickly reverted.

However, it remains unclear how flows from ETFs influence stock prices. A change in a stock's price after a large ETF flow can be initiated by either an adjustment to its fundamentals, a buying/selling pressure from large orders, or both. One can claim that these changes originating from index rebalancing have little relation with a firm's fundamental outlook. For example, the Russell 1000 index contains the largest 1000 companies, whereas the Russell 2000 index includes the next largest 2000 companies based on market capitalization. These indices are rebalanced once a year. Although companies at the bottom of the Russell 1000 index are larger than those at the top of the Russell 2000 index, their stocks exhibit much lower levels of ETF holding. This is because the former group forms a small fraction of the Russell 1000 index, whereas the latter group forms a large proportion of the Russell 2000 index.

Let us consider a situation in which after the Russell reconstitution, some stocks at the bottom of the Russell 1000 index drop out and become the top stocks in the Russell 2000 index. These stocks can experience large ETF inflows because they present larger proportions of the index that many ETFs track. If large buying from ETFs creates price pressure, these stocks' prices will increase. However, the change in ETF membership contains little information about their fundamental values. If prices are assumed to reflect information, their fundamental values deteriorate as their sizes decrease.

In contrast, being included or excluded in an index can signal changes in their fundamental values. For example, inclusion in the S&P 500 index can signal the future potentials of the stock, which would attract fund flows. This increase in institutional ownership can provide the company with more capital and draw more analyst coverage, which can increase its valuation.[22] Thus, ETF flows can generate permanent changes in prices. While mutual fund and ETF flows differ in many aspects, they have few similar characteristics. They are large orders that might contain information or affect the prospect of stocks and certainly influence stock prices, at least in the short term.

[22] Because increase in analyst coverage provides more information about the firms and reduces uncertainty and risks.

Generally, these large buying or selling orders by ETFs can be triggered by changes in the fundamentals. However, it is difficult to separate the effect of liquidity-induced price change (price pressure) from the original cause of price change. Coval and Stafford (2007) show that mutual fund flows are preceded by changes in firms' fundamental values and create significant price changes, which are followed by a reversal.

For illustration purposes, consider a stock traded at $15 and a large selling order by mutual funds. The behavior of the stock price depends on the motivation behind the order. If this is purely for liquidity purposes, a drop in the stock price is expected, but it would only exist temporarily.[23] In contrast, if the selling order is motivated by some information about a future decline in the firm's fundamentals, a permanent fall in price is probable. This fall comprises two parts: a temporary change due to selling pressure and a permanent change due to the new fundamental value. Suppose the new fundamental value of the stock is $12. Theoretically, trading in the market will decrease the price to $12. However, in practice, market overreaction can decrease the price to $10 before reverting to the efficient level of $12. Therefore, the $2 decrease, from $12 to $10, is a temporary price pressure effect, whereas the $3 decrease, from $15 to $12, is a permanent price change.

4.2.3 Hypotheses

This chapter investigates the impact of changes in ETF holdings on the underlying stock price. Formally, the following hypotheses are tested:

Hypothesis 1: Changes in ETF holdings of a stock have temporary and permanent impacts on stock prices.

The price pressure created by mutual funds is well-documented. Because ETFs have grown significantly and share many similar characteristics with mutual funds, we expect a similar impact following a change in stock holdings. Large trades by ETFs can create short-term order and liquidity imbalance, which can produce temporary price pressure. Additionally, many investors use ETFs to gain exposure to specific companies or industries,

[23] For example, an investor decides to withdraw their investment from the fund. This can cause a large sell of the fund positions to meet the investor's redemption.

or hedge their bets. Thus, being included and excluded by ETFs contains information and can cause permanent changes in stock prices.

Hypothesis 2: The sources of changes in ETF holdings impact stock returns differently. ETF flows create price pressure, whereas index rebalancing events do not influence stock prices.

Different sources of changes in ETF holdings can have different impacts. We divide these changes into two parts: flow and rebalancing. Empirically, the market has become efficient; therefore, any distortion in prices can be anticipated or removed quickly and arbitrageurs can immediately take advantage of the mispricing opportunities. Thus, price distortions caused by index rebalancing events should not persist (Hypothesis 2a). However, unexpected changes in ETF holdings are not anticipated and can create price pressure because of sudden surges in order–liquidity imbalance (Hypothesis 2b).

This chapter aims to test a recent claim made by Gabaix and Koijen (2020), according to which inflow into stocks creates a five-fold increase in stock returns. The closest study related to this chapter is that of Brown, Davies, and Ringgenberg (2020), who show that ETF flows generate a nonfundamental demand shock that increases the stock prices temporarily and reverts these changes after six months. We make two contributions to their findings. First, our daily holding data allow us to investigate the price pressure more closely because we can capture the exact days of changes in ETF holdings. Second, we contend that the changes made in ETF holdings are not completely nonfundamental. Many investors nowadays gain exposure to a certain stock by purchasing the ETFs that hold the stock rather than buying the stocks directly. Thus, ETF flow contains information.

4.3 Institutional details and the process by which ETF holdings change

This section describes the creation and redemption processes of ETF shares, as well as their settlement procedure. Understanding the potential misalignments between trades and the changes in holdings reported by ETFs is crucial. For most of this sample, the settlement cycle is $t + 3$, except from September 5, 2017, where the cycle is shortened to $t + 2$. The National Securities Clearing Corporation (NSCC), a subsidiary of the Depository Trust and

Clearing Corporation, provides clearing, settlement, and book-keeping services for most securities traded in the US, including ETF shares. NSCC acts as the middleman between an AP and an ETF sponsor to guarantee the transaction. If one party fails to deliver the security, the NSCC will cover the trade. Thus, NSCC members, including APs, are required to post collateral. This study focuses only on US domestic equity ETFs. Other ETFs that hold other securities, such as bonds, derivatives, or international equities, work on similar principles. As the ETF distributor and ETF agent act on behalf of the ETF sponsor, they are treated as the same entity for ease of reading.

APs can help ascertain the ETF share price close to its net asset value (NAV) through the in-kind creation and redemption process. The AP must deliver (receive) a basket of securities to (from) the ETF sponsor in exchange for ETF shares.

At the end of day $t-1$, the ETF submits the portfolio composition file to the NSCC, which entails the underlying stocks and their quantities. This file also reflects the new weights of the constituents if the underlying index rebalances on day t. This list becomes available to the AP around noon the next day (time t) so that the AP can trade the underlying securities.

Let us consider that at time t—when there is excessive demand to buy (sell), which makes the price of the ETF shares higher (lower) than their NAV—the APs create (redeem) ETF shares to accommodate the excess buying (selling) of the ETF, thereby minimizing the deviation between the ETF price and its NAV. During day t, the AP sells (buys) ETF shares and trades the corresponding underlying securities: when the AP sells the ETF, they will tend to buy the underlying stocks and vice versa.[24] At the end of the trading day (around 4:00 p.m.), the AP sends a creation/redemption order to the ETF sponsor, who will then check and send an instruction to the NSCC by 8:00 p.m. to create/redeem ETF shares. This instruction indicates the AP, ETF share, and number of units to be created or redeemed. The NSCC validates everything and sends a report to the ETF sponsor and AP. This report identifies the closing price at time t for each underlying security. At this point, a contract between the AP and ETF sponsor for the creation/redemption order is locked in.

At time $t+1$, the AP and ETF sponsor validate the information on the report sent by the NSCC. Any error in the price or quantity of the underlying securities, as well as the ETF

[24] At this point, the new ETF shares do not exist in the NSCC system or the ETF sponsor.

shares, is notified to the other involved parties and corrected. At midnight of day $t + 1$, the NSCC guarantees the settlement of all locked-in contracts.

At time $t + 2$, the NSCC sends a consolidated summary report to the ETF sponsor and AP, which shows the netted number of securities required to be settled the next day (i.e., $t + 3$). The NSCC also sends the Depository Trust Company (DTC) information and settlement instruction on individual securities so that the DTC can confirm the availability of securities in the AP's and ETF sponsor's accounts.

Late in the afternoon of $t + 3$, after the availability of securities is confirmed, the DTC simultaneously transfers the ownership of the underlying securities from the AP to the ETF and that of the ETF shares from the ETF to the AP. The cases in which the ETF shares or the underlying securities are not settled by $t + 3$ are reported as "failure-to-deliver" to the SEC. The APs, which act as market makers, have an additional three days to settle trades if their failure to deliver is a result of bona-fide market making.

So, what happens when the index provider changes the weightings of the ETF components (underlying stocks) (i.e., a rebalancing event)? The index provider notifies the ETF sponsors tracking the index about the inclusions and exclusions in advance so that they can manage their trading. Typically, ETF managers are the ones who trade the inclusions and exclusions, not the APs. They often trade at the closing auction of day $t - 1$ so that the positions on day t reflect the new inclusions and exclusions of the underlying index.

4.4 Data

We collect data from different sources. The sample period is 2012–2019. Data on stock returns and Fama–French factors are obtained from CRSP. Morningstar is the primary source for obtaining the number of outstanding shares for ETFs, which we use to compute ETF flows. The sample does not include stocks with a price lower than $5 at the start of each year.

ETF holding data are primarily obtained from Morningstar. In cases where Morningstar does not report the holdings, we collect data from Bloomberg and ETF Global. To validate the numbers, we verify the Bloomberg and ETF Global data against the Morningstar data three days before and after the dates with missing data. If these holdings do not match, we do not fill the missing values. For example, consider the missing holdings of an ETF on the 18th, 19th, and 20th of October 2016 from Morningstar, but not from Bloomberg. We first compare the holdings reported by Morningstar and Bloomberg for the 12th, 13th, and 14th (three days before the period of missing values) and 20th, 21st, and 24th (three days after the period of missing values). If these holdings are the same, we use Bloomberg data to fill in the missing data in Morningstar. If the holdings in one of these dates from these two sources do not match, we do not fill in the data and keep them as missing values.

After the above procedure, there are still cases in which the ETF holding data are missing. We then use ETF flows, which indicate the change in the number of ETF shares, to work out the holdings. We presume that without a rebalancing event, the number of shares held by an ETF changes proportionally with the flow into and out of the ETF. For example, consider an ETF holding of 1000 shares of a stock at $t-1$, which is missing for time t. If the ETF experiences a fund inflow of 0.5% at time t, we can compute the holding of that stock as 1005 ($1000 \times 0.5\%$). We limit the number of consecutive missing values to be filled to three. We check if the computed holdings match the next available numbers. In our example, we calculate the holding at $t+3$ and verify this against the actual holding from the data. If these two numbers are the same, we fill the missing data from time t to $t+2$. If they do not match, we do not fill in these missing values.

Following the literature (e.g., Brown, Davies, and Ringgenberg, 2020), we focus only on ETFs that are both (i) domiciled in the US and (ii) invest primarily in US stocks. Due to the nature of this analysis, the sample includes only ETFs that report holdings daily. Consequently, the sample contains 863 ETFs. However, some large ETFs, such as Vanguard's ETFs, are excluded because they only report their holdings monthly.

A crucial requirement of this study is the matching alignment of ETF flow and changes in ETF holding of the stocks. Following Brown, Davies, and Ringgenberg (2020), we measure ETF flows as:

$$flow_{f,t} = \frac{Shares\ outstanding_{f,t}}{Shares\ outstanding_{f,t-1}} - 1 \tag{4.1}$$

where $Shares\ outstanding_{f,t}$ is the number of outstanding shares of ETF f on day t. Most ETFs are structured as open-ended funds. When an ETF receives a fund flow, its total outstanding shares will change through a creation or redemption event. Therefore, the number of shares of the underlying stocks held by the ETF will have to change by the same fraction (assuming that there is no rebalancing event).

Another essential requirement of this study is the alignment of the timing of a recorded position change and when the trading of that position change occurs. Although Morningstar confirms that the holdings are reported when they are traded (i.e., the timing of these two changes should match), we still find inconsistencies. We make two adjustments to the data.

First, we adjust the timing of an ETF's holdings and its total outstanding shares based on whichever occurs first. For this purpose, we compute the cross-correlation between (i) the average changes in the holdings and (ii) the changes in the total outstanding shares for each ETF. If these numbers are aligned accurately, the value of the cross-correlation for $lag = 0$ will be close to 1, and we make no adjustment to the data. In contrast, if the cross-correlation for $lag = 0$ is lower than the value for $lag = 1$ or $lag = -1$, we interpret this as a misalignment in the data and adjust the holdings.

For example, if we observe that, at time t, an ETF's outstanding shares increase by 2% while its holdings remain the same but change by the same magnitude at time $t + 1$, we adjust this by one day to account for the difference between when trading occurs and when a position change is recorded. While some ETFs' total outstanding shares are either leading or lagging by one day for the whole sample, there are occasions where the misalignment exists for a small period (e.g., a few months). For these instances, we only make the adjustments for this period. Overall, the adjustments affect approximately 20% of the observations.

Second, we adjust the timing of the changes in the holdings by two days because of the $t + 3$ and $t + 2$ settlements, as discussed in the previous section. Thus, for the sample before (after) September 5, 2017, we move the ETF holding variable forward by three (two) days.

4.5 Empirical analysis

Our analysis focuses on stock returns around days with extreme changes in ETF holdings. We first show the stock return dynamics in an event study. Next, we perform a multivariate analysis to investigate the relation between stock returns and changes in ETF holdings. We then show how the passive price pressure created by changes in ETF holdings affects stock performance. Finally, we investigate the temporary and permanent impacts of these changes.

4.5.1 Descriptive statistics

Table 4.1 summarizes the statistics of the main variables used in this study. The average proportion of the underlying stocks that ETFs hold is 3.5% across stocks. This number is lower than the actual ETF ownership because, in this study, we only consider a subset of ETFs that report daily holding. For example, ETFs in the Vanguard fund family, which form approximately 20% of the total ssset under management (AUM) by ETFs, are excluded. On average, ETFs change their stock holding by only 0.17 bps per day, with a standard deviation of 5.46 bps. The number of changes in ETF holdings caused by fund flows and those caused by rebalancing are similar. Only a small subset of ETFs receives daily flow, and their underlying indices are not frequently rebalanced.

Table 4.1: Descriptive statistics

This table reports summary statistics for the main variables constructed using a sample of AMEX, NASDAQ, and NYSE stocks and ETFs between 2012 and 2019. The unit of observation is a stock day. Stocks with a price of less than \$5 at the beginning of each year are excluded. $ETFholdings$ is the proportion of the stock's outstanding shares held by ETFs that are traded in the US. $ETFNetBuy$ is the daily change in $ETFholdings$. $ETFNetBuy^{flow}$ is the daily change in $ETFholdings$ that is caused by the flow of funds into and out of ETFs. $ETFNetBuy^{rebal}$ is the daily change in $ETFholdings$ that is caused by the rebalancing of the underlying index. r_{excess} is the daily returns of stocks in excess of the market return. mktCap is the total market capitalization of the stocks.

Variable	Mean	Std	P25	Median	P75	N
r_{excess} (bps)	0.458	249.870	−90.660	−2.060	87.440	5,325,043
mktCap (\$ million)	8,188	32,334	383	1,249	4,245	5,325,193
$ETFholding$ (decimal)	0.035	0.027	0.019	0.031	0.046	5,313,375
$ETFNetBuy$(x10,000)	0.168	5.460	−0.563	0.018	0.874	5,187,394
$ETFNetBuy^{flow}$ (x10,000)	0.191	4.920	−0.593	0.016	0.848	5,233,722
$ETFNetBuy^{rebal}$(x10,000)	−0.014	4.610	−0.002	0.000	0.002	5,286,163

4.5.2 Passive price pressure

We begin by visualizing the return dynamics around ETF changes in holdings to determine whether the ETF-related trading of underlying stocks impacts the stock returns. In particular, we investigate the dynamics of stock returns around extreme changes in ETF holdings by using an event study analysis where the event is an extreme change in ETF holdings. We use a one-way sorting procedure based on stocks' daily changes in ETF holdings. We compute the change in ETF holdings for each stock i at time t as follows:

$$ETFNetBuy_t^i = \sum_{f=1}^{N} ETFNetBuy_t^{i,f} = \sum_{f=1}^{N} \left(\frac{Q_t^{i,f} - Q_{t-1}^{i,f}}{Shares\ outstanding_{t-1}^i}\right) \tag{4.2}$$

where $ETFNetBuy_t^{i,f}$ is the change in ETF f's position in stock i, which is scaled by the stock's total outstanding shares; $Q_t^{i,f}$ is the number of stocks i held by ETF f at time t; $Shares\ outstanding_{t-1}^i$ is the total number of outstanding shares of stock i at $t-1$; and N is the total number of ETFs in the sample. First, we compute the changes at the fund level and then take the sum across all funds to avoid missing data, which can bias the numbers.[25] We then sort stock-day observations into 20 groups (ventiles) based on the daily changes in ETF holdings. Because ETFs do not change their stock holdings often, most of the observations have zero value for the sorting variable $ETFNetBuy_t^i$. Thus, this analysis focuses on the top and bottom groups (20th and 1st ventiles) because they contain extreme values in both tails (i.e., significant increases and decreases in ETF holdings). Price pressure is likely to be the strongest when there is a large flow from ETFs.

Figure 4.1 shows the return dynamic of stocks that experience a significant increase in ETF holdings (red line) and a substantial decrease in ETF holdings (blue line). The returns are cumulative simple excess returns over the market. We subtract the equally weighted returns of the CRSP universe stocks to isolate individual stock returns from the market trend. The study window extends from 20 days before to 40 days after the event days ($t = 0$), which is when the stocks experience a sharp increase or decrease in the number of their shares held by ETFs.

[25] If we take the sum of the ETF holdings and compute the changes at time t, the numbers can be incorrect if there is a missing value for time $t-1$.

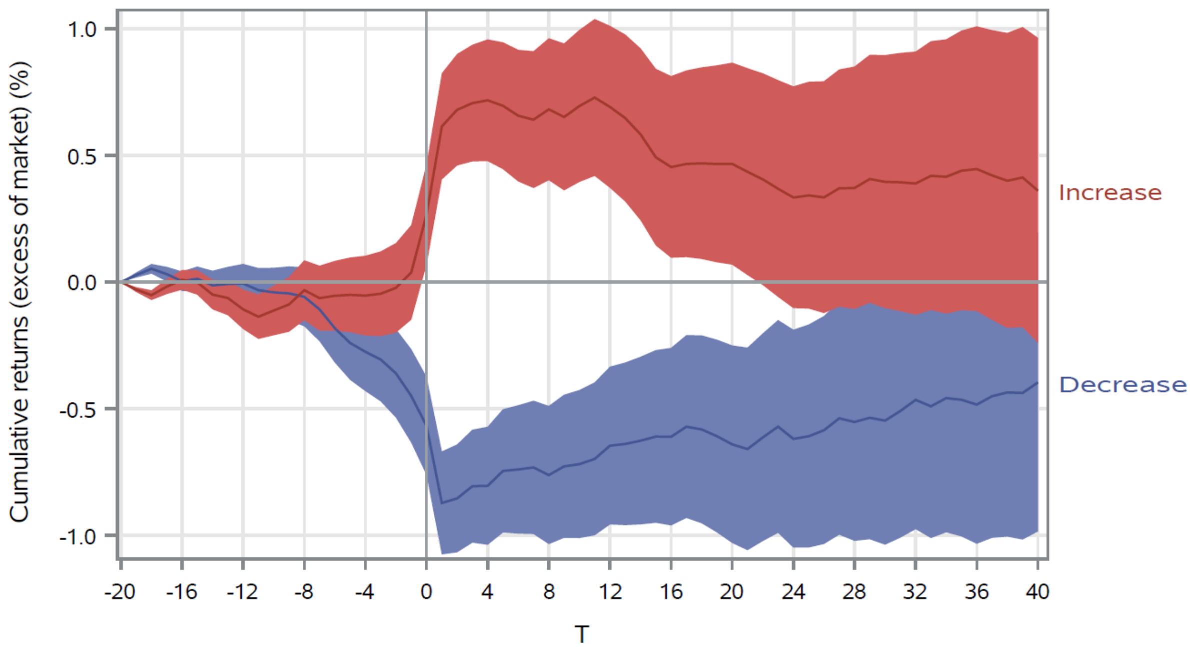


Figure 4.1: Passive price pressure

This figure plots cumulative excess returns (excess of the market) of stocks from 20 days before and 40 days after the event days ($t = 0$) (i.e., days on which stocks experience extreme changes in $ETFholding$). $ETFholding$ is the proportion of the stock's outstanding shares held by ETFs traded in the US. Stocks are sorted into ventiles based on their daily change in ETF holding. The red line (labeled "Increase") represents the top group, which contains stocks that experience the largest increase in ETF holdings, whereas the blue line (labeled "Decrease") represents the bottom group, which contains stocks that experience the largest decrease in ETF holdings. The sample comprises AMEX, NASDAQ, and NYSE stocks and ETFs between 2012 and 2019, excluding stocks with a price of less than $5 at the beginning of each year.

There are two major findings. First, the stocks added to ETF holdings vs. those removed from ETF holdings diverge around 10 days before the day on which there is an extreme increase or decrease in ETF holdings of the stocks ($t = 0$). The two ventiles do not differ from each other prior to $t = -10$. We also look at 60 and 40 days before $t = 0$ and confirm this indifference. One possible reason for this trend is the flow–performance relation. In particular, the performances of a stock and the ETFs that hold the stock are highly correlated, especially if the holdings are proportionally large. If a stock performs well, an ETF holding the stock would perform well and attract fund inflow. The opposite is true when a stock experiences a negative return, which causes the ETFs holding the stock to suffer from an outflow. A bad or good performance can also cause an index to exclude or include a stock.

The second finding from this analysis is that stocks in the top and bottom ventiles experience sharp changes around two days before the day on which ETFs significantly increase or decrease their stock holdings. Stocks in the top ventile experience a cumulative return of approximately 1.5% over a three-day window. Their counterparts in the bottom ventile also suffer a significant negative return of approximately 1%. However, the returns of these stocks appear to revert soon after the event and take approximately one month to stabilize. Interestingly, prices of stocks that experience an increase in ETF holdings do not fully revert to the previous level. This partial reversion suggests an asymmetric effect of changes in ETF ownership. Buying pressure from ETFs has both temporary and permanent impact on stock prices, whereas the selling pressure of ETFs has only a temporary impact. This finding supports our first hypothesis that ETFs create a price pressure that is not entirely temporary. Changes in ETF holdings contain valuable information.

4.5.3 Sources of change: flow vs. rebalancing

The analysis from the previous subsection confirms the generation of price pressure by changes in ETF holdings. In this section, we investigate whether the sources of these changes matter. We disentangle the changes in ETF holdings of stocks into two components: changes caused by fund flows to ETFs and those caused by index rebalancing. We compute the first component, ETF flows, using Eq. (4.1). This formula helps capture the number of shares of the underlying stocks bought or sold due to ETF flow. The difference between this change and the actual change in the number of shares held by the ETF is attributed to rebalancing. On a day when there is no rebalancing event, the percentage change in the number of shares of an underlying stock would be identical to the growth in the ETF's outstanding shares. Formally, the changes in ETF holdings can be divided into the following two components:

$$\sum_{f=1}^{N} \frac{Q_t^{i,f,total}}{Shares\ outstanding_{t-1}^{i}} = \sum_{f=1}^{N} \frac{Q_t^{i,f,flow}}{Shares\ outstanding_{t-1}^{i}} + \sum_{f=1}^{N} \frac{Q_t^{i,f,rebal}}{Shares\ outstanding_{t-1}^{i}} \tag{4.3}$$

where $Q_t^{i,f,total}$, which equals $Q_t^{i,f} - Q_{t-1}^{i,f}$, is the total change in the number of shares of stock i held by ETF f at time t. $Q_t^{i,f,flow}$ and $Q_t^{i,f,rebal}$ are the two components of the change in ETF f's holdings of stock i. The former is the change caused by the ETF inflow/outflow from ETF investors, which is computed as $flow_{f,t} \times Q^{i}_{f,t-1}$, where $Q^{i}_{f,t-1}$ is the number of shares in stock i held by ETF f at time $t-1$. The latter component (i.e., change caused by a rebalancing event) is the difference between the total change in the ETF's holdings of stock i and the change caused by the investor inflows/outflows. In the absence of rebalancing events, this component $Q^{i,rebal}_{f,t}$ is 0. Rebalancing events do not occur frequently, compared to ETF flows; 92.3% of the stock-day observations have zero value for $Q^{i,rebal}_{f,t}$, whereas this value is 52.9% for $Q^{i,flow}_{f,t}$. Finally, for each stock, we sum these quantity changes across all ETFs in our sample for day t and scale them by the total number of outstanding shares for the stock.

Consider an ETF that holds 1,000 shares of Apple stock at time t. The next day, it receives an inflow equivalent to 1% of its total outstanding shares. As a result, the ETF issues an additional 1% of its total outstanding shares and simultaneously increases its position in the Apple stock by 1% (i.e., 10 shares). If there is no rebalancing event, the number of Apple shares held by this ETF is 1,010. If the data show a different number (e.g., 1,050 Apple shares), the 40 additional shares that cannot be explained by inflows/outflows to the ETF are attributable to a rebalancing event. The change in ETF shares and the holdings of the underlying stocks can occur simultaneously because of the creation and redemption process. When the APs deliver a basket of stocks to the ETF, the reported holdings show additional shares, and at the same time, the ETF supplies the APs with its newly issued shares.

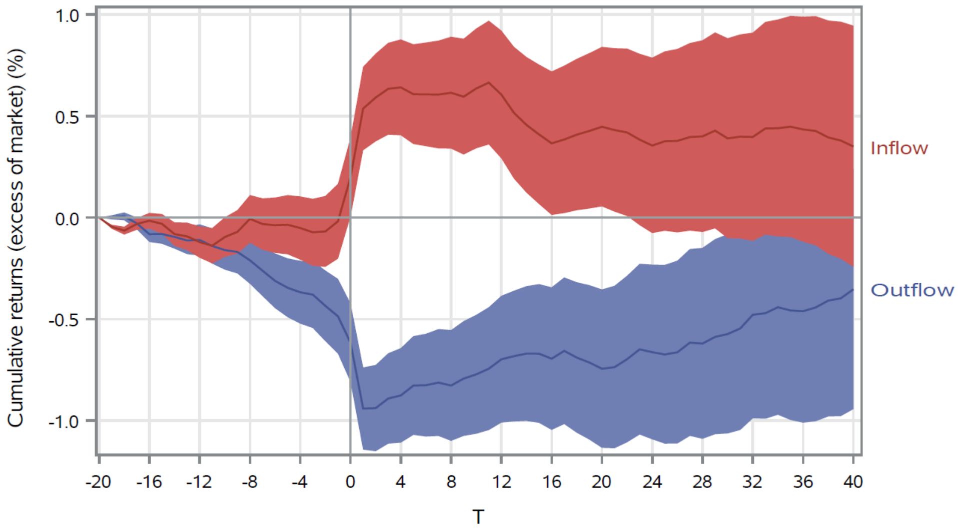
Note that the actual change in the number of shares of stocks might be slightly different from the change assumed according to the change in the number of outstanding ETF shares, perhaps due to rounding numbers or the ETF manager's discretion. Thus, if the difference between these two figures is smaller than 10 shares or 0.05% of the number of shares of the stock, whichever is higher, they are interpreted as changes caused by flows. Another case in which we categorize the change as a rebalancing component is the first or last time a stock appears in the ETF.

For each of the two components of changes in ETF holdings, we perform a separate univariate sorting procedure. For the first component, which is changes caused by ETF flow, we sort stocks into ventiles on the basis of the change in ETF holdings caused by the flow. The top ventile contains stocks whose ETFs receive large investment inflow, whereas the bottom ventile contains stocks that experience a large sell-off by ETFs due to outflow. The second source of change in ETF holdings originates from rebalancing events, which do not occur frequently. Most observations have a value of zero for this variable. For this reason, stocks are sorted into percentiles. Only the top and bottom percentiles are investigated. The top group contains stocks that experience large purchases by ETFs, whereas the bottom group comprises stocks that are heavily sold by ETFs.

Figure 4.2 illustrates different impacts on stock returns of changes in ETF holdings caused by flow and rebalancing. Panel A shows that changes in ETF holdings that are caused by ETF flows create significant price pressure. There is no initial drift 10 days before the extreme changes. The considerable jump and drop in returns start the day before the event ($t = 0$). On average, stocks in the top ventile show a 75-bps jump, whereas their counterparts in the bottom ventile experience a 100-bps drop in excess of market returns over the three days around the event. The stock returns in the bottom group (significant decreases in ETF holding caused by ETF outflows) steadily bounce back after 40 days, whereas those in the top group (significant increases in ETF holding caused by ETF inflows) remain level. This finding suggests that being held more by ETFs is perceived as positive news and creates a permanent price impact. In contrast, being held less by ETFs is negative news, but the impact is only temporary.

Another point worth noting is that there is no noticeable trend in returns for stocks in the top ventile, but there is a downward drift for stocks in the bottom ventile two weeks before the extreme changes in ETF holding. This indicates that a bad performance of stocks causes an outflow of funds from the ETFs holding these stocks. However, there is no similar movement among stocks in the top ventiles. One possible explanation for this is that observations in the top ventile might include fund inflows from other markets, such as active funds, bonds, and commodities. These inflows might not be related to stock performance.

Panel A: Changes in ETF holdings caused by fund flows



Panel B: Changes in ETF holdings caused by rebalancing

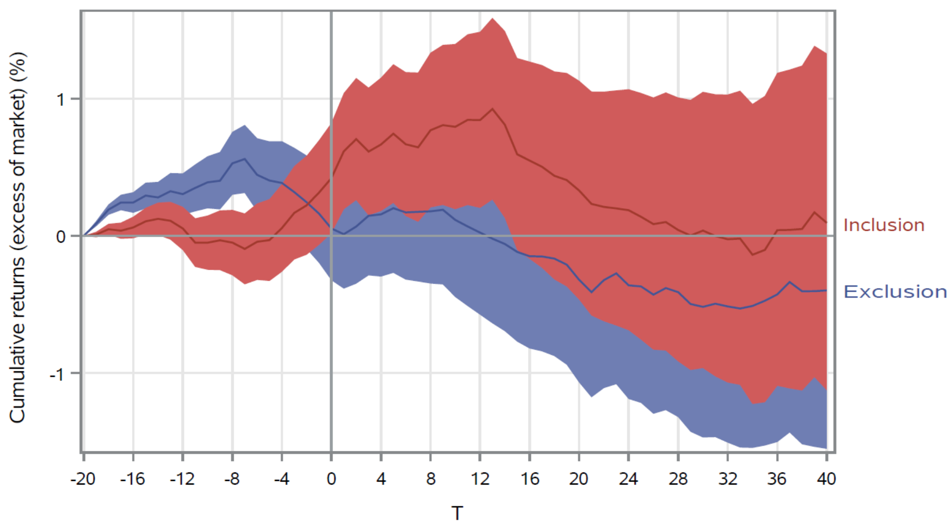


Figure 4.2: Passive price pressure: flow vs. rebalancing

This figure plots cumulative excess returns (excess of the market) of stocks from 20 days before and 40 days after the event days ($t = 0$) (i.e., days on which stocks experience extreme changes in ETF holdings). ETF holding is the proportion of the stock's outstanding shares held by ETFs traded in the US. Changes in ETF holdings are decomposed into two components: flows (Panel A) and rebalancing (Panel B). Stocks are sorted into ventiles (percentiles) based on their daily change in ETF holding that is caused by flows (rebalancing). The red line (labeled "Inflow" in Panel A and "Inclusion" in Panel B) represents the top group, which contains stocks experiencing the most significant increase in ETF holdings, whereas the blue line (labeled "Outflow" in Panel A and "Exclusion" in Panel B) represents the bottom group, which includes stocks experiencing the most significant decrease in ETF holdings. The sample comprises AMEX, NASDAQ, and NYSE stocks and ETFs between 2012 and 2019, excluding stocks with a price of less than $5 at the beginning of each year.

In addition, rebalancing events do not create significant price pressure. Panel B of Figure 4.2 shows that stocks that experience an increase in ETF holdings due to rebalancing have slightly higher returns than their counterparts. However, the performance of stocks in the top and bottom ventiles sorted on changes in ETF holdings caused by rebalancing events are not statistically different from each other, because the confidence bounds of these two groups overlap.

The results are consistent with the view that changes in ETF holdings caused by investment flows create price pressure (i.e., a temporary deviation of prices due to heavy buying or selling) because these events are unexpected, in contrast to index rebalancing events, which are relatively predictable and usually announced in advance of the actual rebalance dates. Market participants, including market makers, APs, and ETF managers, are less able to anticipate investor flows. As a result, these flows create order–liquidity imbalances, which lead to temporary price impacts. In contrast, index rebalancing events are usually known in advance, allowing arbitrageurs to effectively act as liquidity providers in trading against the rebalancing trades of passive funds, such as ETFs. An example is the Russell indices. Market participants can accurately forecast which stocks will go into and out of the Russell 1000 and 2000 indices based on their market capitalization. Additionally, market makers and ETF managers are notified in advance by index providers about which stocks will be added or removed from the index. For these reasons, rebalancing events do not have the "surprise" component.

4.5.4 A multivariate analysis

The previous findings confirm the generation of passive price pressure by ETF-related trading. However, momentum or reversal in returns could have driven these results. To control for these factors, we perform a multivariate analysis on the relation between stock returns and changes in ETF holdings. We regress stocks' cumulative returns on several explanatory variables, focusing on the variables that measure the changes in ETF holdings while controlling for momentum and reversals in returns.

We decompose ETF changes into two components: changes that are caused by investment flows ($ETFNetBuy_{t-3,t}^{i,flow}$) and those caused by index rebalancing ($ETFNetBuy_{t-3,t}^{i,rebal}$). We compute these changes by using the same approach as that proposed in the previous subsection. However, we calculate the changes over a three-day window because some ETFs might buy or sell in advance of the rebalancing or because of the possible delay in the reporting of holdings.

We test several windows of cumulative returns but report only two windows. The results for other windows are similar. Particularly, for each stock-date observation, we compute the cumulative returns over the next 40 ($R_{t,t+40}$) and 90 ($R_{t,t+90}$) days. The changes in the ETF holdings are captured by first aggregating the number of shares of stock i bought and sold by all ETFs in the sample and scaling this aggregate change in the holdings by the number of outstanding shares in stock i at date t.

To control the potential momentum and reversals in returns, we include two explanatory variables that capture the cumulative returns of stocks before the event date. The first variable, $R^i_{t-3,t}$, is the cumulative return from three days before the event ($t-3$) to the event date (t). The second variable, $R^i_{t-a,t-4}$, is the cumulative return from date $t-a$ to $t-4$, where a is the length of the return window of the dependent variables (i.e., $R^i_{t,t+40}$ and $R^i_{t,t+90}$) (40 or 90 days, respectively). Other control variables include stock market capitalization and total turnover (volume divided by total outstanding shares) from $t-3$ to t. We also control for other time-series effects by using a time-fixed effect. The regression can be expressed as follows:

$$R^i_{t,t+a} = ETFNetBuy^{i,flow}_{t-3,t} + ETFNetBuy^{i,rebal}_{t-3,t} + \delta_t + Controls^i_t \qquad (4.4)$$

Table 4.2 reports the estimates obtained from this regression. The coefficients for the ETF net buying variables are negative and statistically significant. These results indicate a negative relation between changes in ETF holdings and future returns, consistent with reversals of the temporary price impact caused by ETF buying or selling. When ETFs purchase shares, they tend to increase the prices of the underlying shares, and if there is a temporary component of that price impact, then the stock returns are expected to fall over the next 40 or 90 days. Conversely, stocks that experience selling by ETF might experience short-term negative returns but recover over the next two or three months. Therefore, a temporary negative impact from ETF selling is expected to be followed by a positive future abnormal return. These results suggest the presence of a temporary price impact and are consistent with the findings in the previous subsections, where we observe a consistent downward (upward) drift for stocks following extreme increases (decreases) in ETF holdings. Thus, the results support our first hypothesis that changes in ETF holdings create price pressure, which are temporary distortions in prices that are quickly reverted.

Table 4.2: Regression of stock returns on ETF changes

This table reports the results obtained from the following regression:

$$R^i_{t,t+a} = ETFNetBuy^{i,flow}_{t-3,t} + ETFNetBuy^{i,rebal}_{t-3,t} + \delta_t + Controls^i_t$$

where $R^i_{t,t+a}$ indicates the cumulative returns of stock i from day t to $t+a$, which takes a value of 40 (first column) or 90 (last column); $ETFNetBuy^{i,flow}_{t-5,t}$ is the cumulative net buying or selling (as a proportion of outstanding shares) by ETFs due to flow into and out of ETFs from day $t = -3$ to 0; and $ETFNetBuy^{i,rebal}_{t-5,t}$ is the cumulative net buying or selling (as a proportion of outstanding shares) by ETFs due to ETF rebalancing from day $t = -3$ to 0. $Controls^i_t$ includes market capitalization of stock i at time t; $Turn_{-3,0}$, which is the cumulative turnover of stock i from $t = -3$ to t; $R^i_{t-3,t}$, which is the cumulative returns of stock i from $t = -3$ to t; and $R^i_{t-a,t-4}$, which is the cumulative returns of stock i from day $t-a$ to $t-4$, where a is either 40 or 90. t-statistics are reported in parentheses. Standard errors are clustered by stock. The sample comprises AMEX, NASDAQ, and NYSE stocks and ETFs between 2012 and 2019, excluding stocks with a price of less than \$5 at the beginning of each year. The symbols *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

Variable	$R^i_{t,t+40}$	$R^i_{t,t+40}$
$ETFNetBuy^{i,flow}_{t-3,t}$	−2.040***	−2.650***
	(−9.31)	(−8.39)
$ETFNetBuy^{i,rebal}_{t-3,t}$	−0.309**	−0.670***
	(−2.08)	(−2.83)
$R^i_{t-3,t}$	0.294***	0.291***
	−48.57	−29.42
$R^i_{t-a,t-4}$	0.003	0.016***
	−0.92	−2.96
$mktCap^i_t$ (× 100)	0.001***	0.001***
	−3.05	−2.67
$Turn^i_{t-3,t}$	−0.036***	−0.095***
	(−3.41)	(−2.38)
Time FE	Yes	Yes
R^2	9.28	3.52

The two sources of changes in ETF holdings have different effects on stock returns. According to Table 4.2, changes that originate from ETF investor flows have a more substantial impact. The coefficients of the $ETFNetBuy^{i,flow}_{t-3,t}$ variable are more than five times those of the $ETFNetBuy^{i,rebal}_{t-3,t}$ variable. Using the 40-day window, when ETFs purchase 1% of a stock's total outstanding shares, the stock, on average, has an abnormal return of approximately –2% over the next 40 days, or equivalently approximately –12% annualized return.

In contrast, when an index is rebalanced, which causes ETFs to hold more of that stock, the stock return also tends to decrease following the change in ETF holdings but with a small magnitude; this is consistent with a lesser price impact of ETF trading. The same 1% increase in ETF ownership in the case of a rebalancing event is associated with a −0.31% subsequent abnormal return, or less than −2% annualized return over the next 40 days.

The coefficients in the second regression, which uses a 90-day window after the event, are just slightly higher than those in the first regression. Thus, much of the reversion in returns occurs during the first 40 days following the events.

The difference in the impacts of different sources of changes in ETF holdings supports our second hypothesis that the origins of changes matter. This difference is attributed to the "surprise" factor associated with the fund flows into or out of an ETF. In contrast, index reconstitution events are greatly anticipated. Many indices have a clear set of publicly available rules, and investors can predict these events with high levels of certainty and trade. Many large inclusions and exclusions are covered in the media a few weeks before their occurrence. The Russell 1000 and 2000 indices are reconstituted in June based on the stock market capitalization in May. This time gap leaves the market considerable time to buy or sell stocks before the ETFs that are tracking these indices compulsorily trade on the reconstitution dates. Even for indices that do not widely disclose the rules for addition and deletion, ETF managers and market makers know these changes in advance. This allows them to trade without much price impact.

In contrast, ETF managers and their APs cannot predict when the next fund flow will occur. When they receive an inflow or outflow, they must buy or sell the underlying stocks to keep the ETF's assets under management close to its market capitalization, i.e., to minimize the tracking errors. This mechanical trading and the unexpected fund flow can generate order–liquidity imbalance, which creates price pressure.

4.5.5 Disappearance of index premium

Given the evidence in Subsection 4.5.3 and 4.5.4 that index rebalancing events do not create much price pressure as opposed to ETF flow, it is natural to ask whether the returns associated with index inclusion or deletion events have changed over time with the rapid growth of ETFs. In this section, we analyze how ETFs influence the index premium, which has already been defined in the literature (Petajisto, 2011).

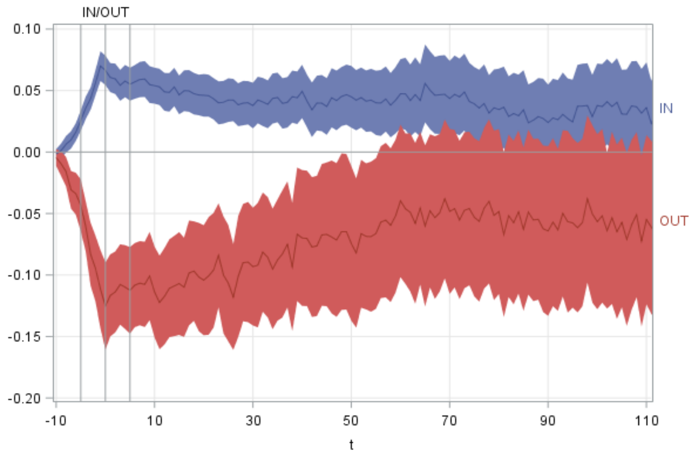
Petajisto (2011) shows that stocks are traded at a premium for index inclusion and a discount for index exclusion. In particular, a stock earns 8.8% when added to the S&P 500 index and experiences a 15.1% loss when dropped out of the same index. Petajisto also claims that this premium and discount has been growing over time. However, this study yields the opposite result. Figure 4.3 illustrates the disappearance of this phenomenon.

We start by replicating the main result of Petajisto (2011), as shown in Panel A of Figure 4.3. The top blue line shows the cumulative returns of stocks that are included into the S&P 500 index during the 1990–2005 period 10 days before and 110 days after the inclusion. These stocks experience a sharp increase of approximately 8% in returns before $t = 0$. In contrast, stocks that are dropped out of the S&P 500 index exhibit a more severe loss of approximately 15%. The gap between these two groups of stocks is clear in Panel A. While there is a slight reversion of these two groups 100 days after the event, the two lines do not intersect each other.

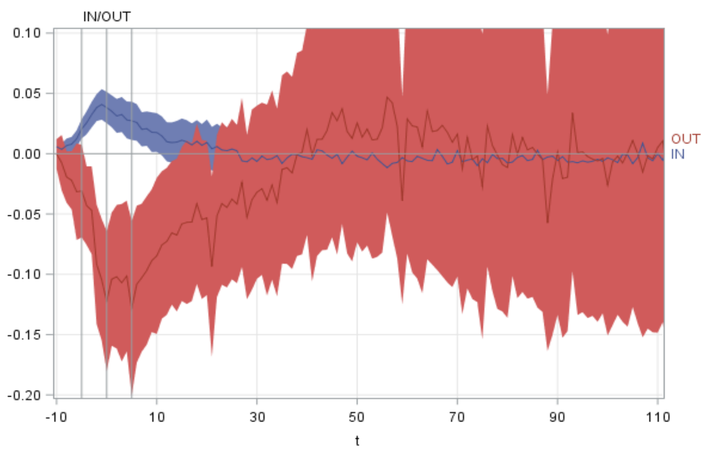
We then perform a similar analysis for stocks added and removed from the S&P 500 index for the period after Petajisto's (2011) study period. We break this follow-up period into two intervals, 2006–2011 and 2012–2019, the latter being this study's main sample period.

The gap between the cumulative returns of the inclusions and exclusions has reduced over time and completely disappeared recently. According to Panel B of Figure 4.3, the cumulative returns of inclusions during the 2006–2011 period do not have as large abnormal returns as their equivalents in the 1990–2005 period. The premium for being included in the S&P 500 index is only approximately 4%, which is half of that for the previous period. Similarly, exclusions do not suffer much discount compared to the period before.

Panel A: Period 1990–2005



Panel B: Period 2006–2011



Panel C: Period 2012–2019

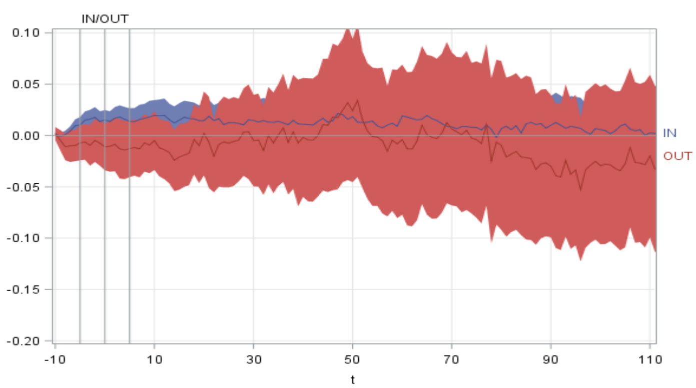
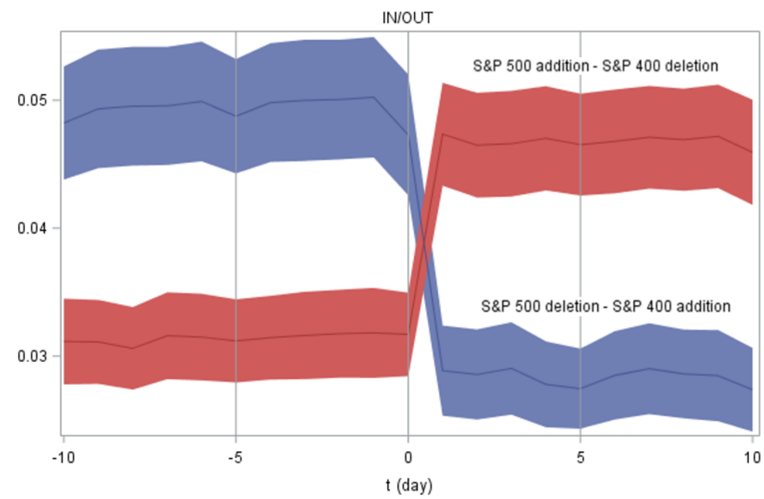


Figure 4.3: S&P 500 inclusion and exclusion

This figure presents the cumulative average market-adjusted returns in event time for stocks included and excluded from the S&P 500 index, as well as the 95% confidence intervals. Index changes are effective at $t = 0$. Panels A, B, and C present the sample periods 1990–2005, 2006–2011, and 2012–2019, respectively. The sample selection and method are similar to those used by Petajisto (2011).

In addition, the returns of these two groups converge and become identical after about two months since the reconstitution. This is different from what we see in Panel A of Figure 4.3, where there is no full convergence of the two groups and the gap in returns remains significant after more than three months since the reconstitution.

Panel A: S&P 500



Panel A: Russell 1000

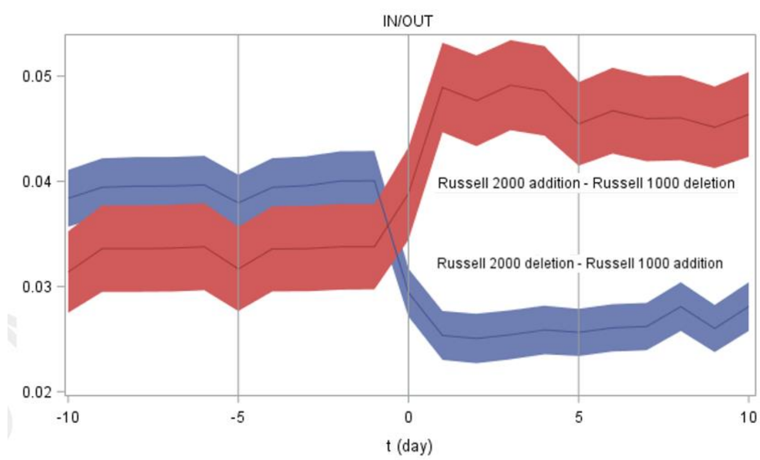


Figure 4.4: ETF holdings and rebalancing events

This figure presents the average $ETFholding$ of stocks (in decimal) that get added to (removed from) large-cap indices and removed from (added to) mid-cap indices. Panel A shows this for the S&P 500 index, whereas Panel B shows this for the Russell 1000 index. $ETFholding$ is the proportion of the stock's outstanding shares held by ETFs traded in the US. The sample comprises AMEX, NASDAQ, and NYSE stocks between 2012 and 2019, excluding stocks with a price of less than $5 at the beginning of each year. Confidence intervals of 95% are shown in the color bands of the lines.

The index premium and discount have completely disappeared in the recent 2012–2019 period, as shown in Panel C of Figure 4.3. The two lines representing the additions and deletions are not different from each other. There is no jump or drop in stock returns around the rebalancing dates, as in other previous periods. This finding contrasts with that of Petajisto (2011), who asserts that the index premium phenomenon increases over time.

We consider several explanations for this disappearance of index inclusion effects. ETF growth has made the holdings of stocks by passive funds more uniform. Stocks that drop out of an index often join another index. Additionally, the increased arbitrage activities around index rebalancing events have caused the index inclusion effects to disappear. Investors might be aware of the premium and discount of index reconstitution and act in advance; that is, they can buy the inclusions and sell the exclusions before the rebalancing day.

Another possible explanation for the disappearance of the index premium is the price pressure created by ETFs. When a stock is added to the S&P 500 or Russell 1000 index, its ETF holdings decrease. This is because the stock loses its index-weight importance and experiences a large sell out by ETFs. Previous sections showed that a large drop in ETF holdings causes downward price pressure and reduces stock prices. The opposite effect is observed for the index premium. These two forces can cancel out each other, and as a result, the index premium disappears. The opposite is true for stocks that are removed from the index. This mechanism also explains why there is not much price pressure created by rebalancing events as opposed to ETF flows.

Figure 4.4 illustrates the jumps (drops) in ETF holdings of stocks that are excluded (included) from the S&P 500 and Russell 1000 indices. Let us consider the large-cap S&P 500 index, as shown in Panel A of Figure 4.4. A stock that has risen in size and is included in the S&P 500 index is typically a large component of the mid-cap S&P 400 index, another widely followed index. Due to its size, the stock's weight in the S&P 400 index is expected to be high, and the ETF tracking the index should be able to hold this stock to minimize the tracking error. In contrast, when the stock is added to the S&P 500 index, its rank will be low in the index's weight ladder, and ETFs that track the S&P 500 index will not hold a large proportion of the stock. This is consistent with the results obtained by Appel, Gormley, and Keim (2016), who find a similar pattern with the Russell 1000 and 2000 indices, as shown in Panel B of Figure 4.4.

4.5.6 Economic impact on stock performance

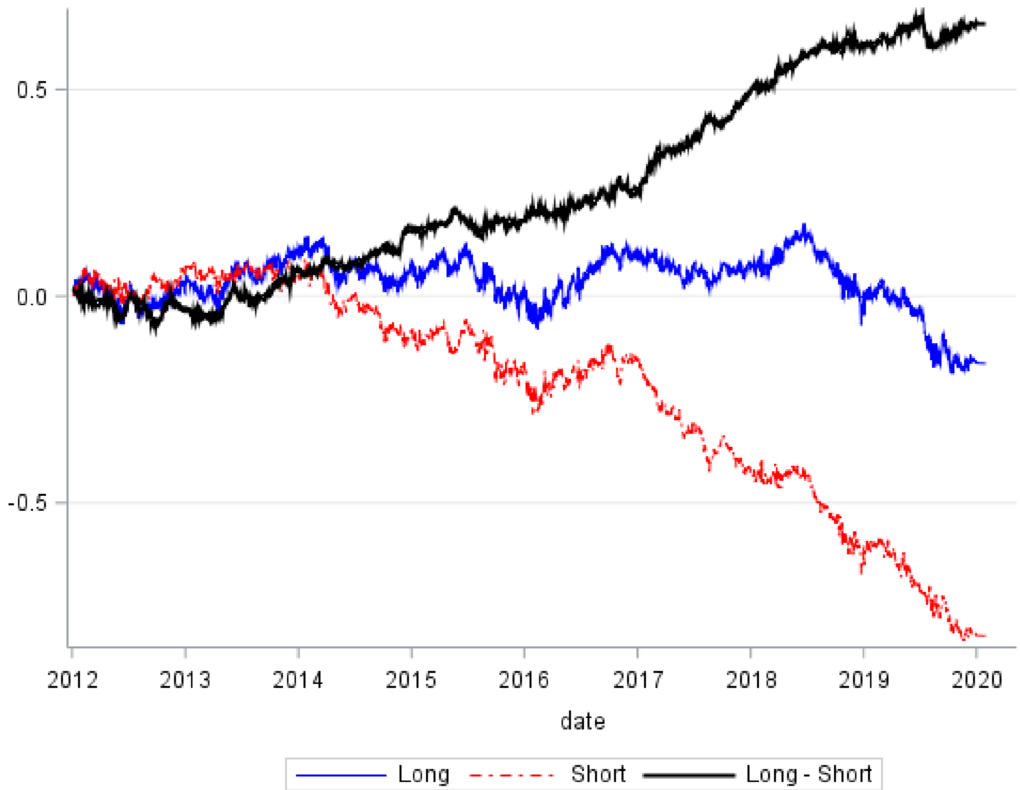


Figure 4.5: Cumulative returns of anomaly portfolios

This figure shows the cumulative returns (in decimal) of the long leg, short leg, and long-short strategy in excess of the market returns. Portfolio returns are equal-weighted average returns. Stocks that experience a decrease (increase) of more than 0.1% in ETF ownership are added to the long (short) portfolio. The holding period is 40 days. The long-short portfolio is the difference in returns of the long and short legs. The sample comprises AMEX, NASDAQ, and NYSE stocks between 2012 and 2019, excluding stocks with a price of less than $5 at the beginning of each year.

If ETF flows create price pressure that will partially revert, how does this price pressure affect ETF performance? In this subsection, we investigate the practical implications of the passive price pressure created by ETFs. We form a simple long-short portfolio based on the expected price distortions created by ETF-related trading and track the returns of this portfolio, as is common in asset pricing studies. The strategy buys stocks that are heavily sold by ETFs and short-sells stocks that are heavily bought by ETFs. For simplicity, the threshold for extreme buying and selling is set to 0.1% of the stock's total outstanding shares. This threshold is based on the cut-off points of the sorting procedure described in Subsection 5.2.

We also test the dynamic thresholds based on the values of the groups sorted using historical ETF changes, and the results are found to be similar.

On each day, we sort stocks based on the proportion of their shares held by ETFs. If ETFs sell more than 0.1% of the company's total outstanding shares, the long-short portfolio buys the stocks. Similarly, the strategy sells stocks when ETFs purchase more than 0.1% of their shares. Thus, each long-short position is held for 40 days. Additionally, if a stock is added to the long leg (short leg) of this strategy at time t, and then included to the short leg (long leg) later at $t + a$, which is before the 40-day holding period expires (i.e., $a \leq 40$), it is removed from the portfolio at $t + a$.

Figure 4.5 illustrates the performance of this long-short portfolio strategy. The blue solid line and red dashed line represent the cumulative simple returns in excess of the market return of the long and short legs, respectively. The bold black line represents the total return of this strategy. All returns are equally weighted. Accordingly, the overall portfolio exhibits significant positive returns over the study period. At the end of the study period, this strategy earns an extra 60% over the market returns. This is equivalent to an average abnormal return of approximately 9% per year. Most of the strategy returns are obtained from the short leg. Stocks that experience an extreme increase in ETF holdings exhibit consistently poor performance compared to the market.

Table 4.3 presents the results for this portfolio, including both equally and value-weighted returns. Panel A shows the returns in excess of the risk-free rate, whereas Panel B shows the returns after adjusting for the three Fama–French factors. For ease of interpretation, we report annualized returns. The equally weighted returns of the strategy are statistically significant and positive. The strategy offers, on average, more than 5% returns per annum even after controlling for the risk factors.

The value-weighted returns are statistically significantly negative, which suggests that the price pressure created by ETF activities is stronger among smaller stocks. In contrast, larger stocks might have better liquidity, which helps absorb unexpected buying and selling from ETF flows. These results are consistent with those obtained by Brown, Davies, and Ringgenberg (2020), who also find a positive performance for a similar strategy. However, they form portfolios based on ETF flows, whereas we consider rebalancing events; moreover, their holding period is six months.

Table 4.3: Long-short portfolio

This table reports annualized percentage returns of the long-short portfolio based on changes in $ETFholdings_t^i$, which is the proportion of the stock's outstanding shares held by ETFs that are traded in the US. Stocks that experience less than −0.02% change in $ETFholdings_t^i$ are long, whereas those that experience more than 0.1% in $ETFholdings_t^i$ are short. The holding period is 40 days. Excess returns are average returns in excess of market returns. Three-factor alpha is the intercept from regressing portfolio excess returns on Fama and French's (1993) market, size, and value factors. The sample comprises AMEX, NASDAQ, and NYSE stocks and ETFs, excluding stocks with a price of less than $5 at the beginning of each year. Returns are reported in annualized percentages, and the t-statistics are presented in parentheses using Newey–West standard errors with six lags. The symbols *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

Portfolio	Equally weighted	Value weighted
Panel A: Excess returns		
Long	−0.408	7.816**
	(−0.09)	(2.31)
Short	−5.728	4.819
	(−1.21)	(1.4)
Long−short	5.32***	−5.227**
	(3.64)	(−1.98)
Panel B: Three-factor alpha		
Long	12.459*	20.753***
	(1.92)	(4.26)
Short	7.119	17.815***
	(1.08)	(3.62)
Long-short	5.34***	−5.355**
	(3.72)	(−2.05)

4.5.7 Temporary vs. permanent price impact

The relation between ETF flows and stock returns illustrated in previous subsections suggests both temporary (which reverse subsequently) and permanent (do not reverse) components of the price impacts. To disentangle these components more formally, we adopt vector-autoregressive regressions (VAR), which is widely used in the market microstructure and asset pricing literature. We use a VAR model based on Hasbrouck's (1923) decomposition model, which analyzes the informational content of a trade. Hasbrouck's seminal paper uses a VAR model of order flow and stock returns to separate the temporary and permanent effects of a trade caused by an unanticipated shock. A change in the stock fundamentals can affect the long-run price of the stock, whereas a short-term, temporary effect (such as liquidity dry-up) can cause a temporary change in the stock price, which can revert to the prevailing level. We apply this method to our analysis to separate the temporary and permanent effects of changes in ETF holdings on stock returns. The temporary returns are noise and distortions caused by passive price pressure, whereas the permanent returns are valuable information impounded into prices.

Using daily stock returns and the two previous measures of changes in ETF holdings due to flows and due to rebalancing within the ETF portfolio, the VAR model in its structural form can be expressed as:

$$
\begin{aligned}
ETFNetBuy_t^{i,flow} &= A_0 + \sum_{n=1}^{10} A_{1n} ETFNetBuy_{t-n}^{i,flow} + \sum_{n=1}^{10} A_{2n} ETFNetBuy_{t-n}^{i,rebal} + \sum_{n=1}^{10} A_{3n} R_{t-n}^{i} + \varepsilon_t^A \\
ETFNetBuy_t^{i,rebal} &= B_0 + \sum_{n=0}^{10} B_{1n} ETFNetBuy_{t-n}^{i,flow} + \sum_{n=1}^{10} B_{2n} ETFNetBuy_{t-n}^{i,rebal} + \sum_{n=1}^{10} B_{3n} R_{t-n}^{i} + \varepsilon_t^B \\
R_t^i &= C_0 + \sum_{n=0}^{10} C_{1n} ETFNetBuy_{t-n}^{i,flow} + \sum_{n=0}^{10} C_{2n} ETFNetBuy_{t-n}^{i,rebal} + \sum_{n=1}^{10} C_{3n} R_{t-n}^{i} + \varepsilon_t^C
\end{aligned}
\tag{4.5}
$$

where i indexes stocks and t indexes days. $ETFNetBuy_t^{i,flow}$ and $ETFNetBuy_t^{i,rebal}$ are the changes in ETF holdings that are triggered by ETF flows and rebalancing events, respectively. We compute these measures by using the procedure discussed in previous subsections. R_t^i is the daily return of stock i at time t in excess of the market. The error terms (ε_t^A, ε_t^B, and ε_t^C) are the unanticipated changes in the dependent variables that are not captured by the independent variables. For each equation, we estimate the coefficients on 10 lags of each variable. We perform a separate estimation for each stock by using the entire sample. We then compute the impulse responses for a shock to each of the sources of changes in ETF holdings, keeping all other variables equal to their unconditional means. We assume that the contemporaneous causality starts with ETF flows. Let us consider a stock with no idiosyncratic news. A fund flow into an ETF holding the stock will create buying pressure that affects the stock price. In contrast, a change in the stock price (that might originate from some news about the stocks) will be less likely to produce a simultaneous change in its ETF holdings. Additionally, we contend that ETF flows affect stock prices, which will influence future rebalancing events.

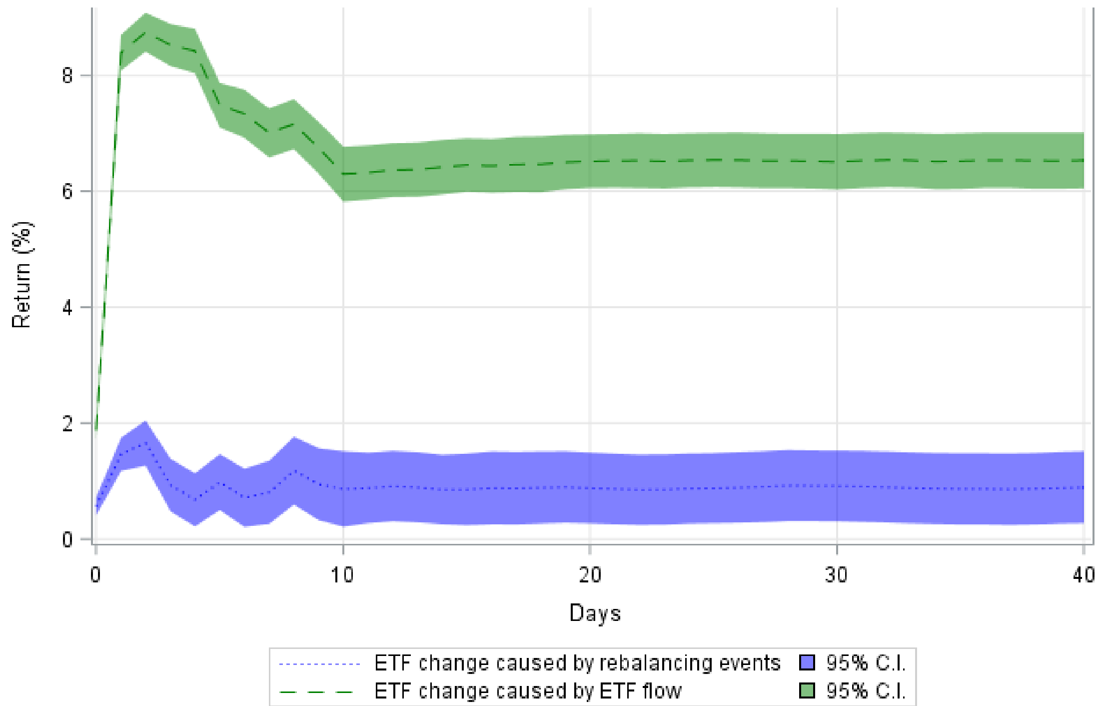


Figure 4.6: Cumulative impulse response of return to 1% change in ETF holdings caused by ETF flow and rebalancing

This figure shows the cumulative impulse response of stocks' excess return (in decimal) to a 1% increase in ETF holding. The VAR model with 10 lags is presented in Eq. (5). The VAR model is estimated separately for each stock by using daily observations. Each model coefficient is then averaged across stocks. The sample comprises AMEX, NASDAQ, and NYSE stocks between 2012 and 2019, excluding stocks with a price of less than $5 at the beginning of each year. Confidence intervals of 95% are shown in the color bands of the lines.

Figure 4.6 illustrates the cumulative impulse response function for the abovementioned VAR model in Equation (5). It depicts the response of stock returns to a 1% shock in $ETFNetBuy_t^{i,flow}$ and $ETFNetBuy_t^{i,rebal}$. The effect of changes in ETF holdings caused by ETF flow (green dashed line) is stronger than that caused by a rebalancing event (blue dotted line). When a stock experiences a 1% increase in $ETFNetBuy_t^{i,flow}$, its return tends to increase significantly on the first day. However, the return quickly drops, becomes negative, and stabilizes at approximately −1.7%. This is consistent with the passive price pressure indicated in Figure 4.1, which shows a downward drift in stock returns following an increase in ETF holdings. The results indicate that ETF-related trading contains valuable information. This is consistent with the results reported by Easley et al. (2021), who contend that investors use ETFs as liquid building blocks of active investment portfolios. ETFs allow investors to trade on industry or market-wide information or hedge against bets on specific stocks (Huang, Hara, and Zhong, 2018).

Changes in ETF holdings caused by rebalancing events show little effect on stock returns. The blue line in the figure shows that stock returns quickly become steady after remaining at the zero level for only two or three days. This indicates that rebalancing events do not have any permanent impact on stock prices. Conversely, the effect of changes caused by ETF flows appears to be permanent. This finding is also consistent with the results obtained in the previous sections and partly consistent with those obtained by Brown, Davies, and Ringgenberg (2020), who contend that ETF flows are nonfundamental demand shocks. However, we show that ETF flows indeed create a permanent impact.

Table 4.4 reports the coefficient estimates obtained from the VAR model and confirms the above results. Most coefficients of the $ETFNetBuy_t^{i,flow}$ terms in the third equation, where the dependent variable is stock return, are statistically significant at 1% significance level and negative. These coefficients indicate a negative relation between stock returns and ETF holding changes caused by ETF flows. In contrast, only four of the ten coefficients of the $ETFNetBuy_t^{i,rebal}$ terms are statistically significant at the 5% level. The coefficients for the higher lags are not statistically significant, and the magnitudes are much smaller compared to those of $ETFNetBuy_t^{i,flow}$. This finding suggests that the impact of rebalancing is short-lived.

Table 4.4: VAR regressions

This table reports the mean-coefficient-estimate VAR model, as shown in Eq. (5). The VAR model is estimated separately for each stock using daily observations. Each model coefficient is then averaged across stocks and reported in the table. The columns $flow_{n,t}^{i}$,$rebal_{n,t}^{i}$, and $R_{n,t}^{i}$ in each panel correspond to the independent variables $ETFNetBuy_{n,t}^{i,flow}$, $ETFNetBuy_{n,t}^{i,rebal}$, and $R_{n,t}^{i}$, respectively. The sample comprises AMEX, NASDAQ, and NYSE stocks and ETFs between 2012 and 2019, excluding stocks with a price of less than $5 at the beginning of each year. The symbols *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

	Panel A: Flow equation			*Panel B: Rebal equation*			*Panel C: Return equation*		
Lag	$flow_{n,t}^{i}$	$rebal_{n,t}^{i}$	$R_{n,t}^{i}$	$flow_{n,t}^{i}$	$rebal_{n,t}^{i}$	$R_{n,t}^{i}$	$flow_{n,t}^{i}$	$rebal_{n,t}^{i}$	$R_{n,t}^{i}$
1	2.474***	0.010	0.200***	−0.375	2.603***	0.037***	−33.852***	−77.031**	−2.014***
2	−1.109***	−1.577***	0.060***	−2.484***	−0.234	0.022***	−22.695***	−49.156***	−0.494***
3	−1.693***	−0.199	0.030***	−0.427***	0.02	0.018***	−2.245	−13.586	−0.288***
4	−2.502***	1.273***	0.033***	−1.147***	0.176*	0.014***	−54.78***	15.008	−1.231***
5	−0.891***	−0.117	0.027***	−0.438	0.434***	0.013***	−37.274***	−22.178***	−0.661***
6	−1.438***	0.363***	0.012***	0.623***	−0.072	0.014***	−1.911	23.823***	0.002
7	−0.544***	0.539***	0.008***	−0.75***	−0.197***	0.007***	−28.536***	18.861	−0.305***
8	−2.235***	0.182	0.011***	−0.538***	−0.691***	0.019***	−48.566***	−4.874	0.089
9	0.593***	−0.300***	0.027***	−0.525***	−0.15*	−0.003*	−72.167***	−18.926*	−0.189***
10	0.061	0.368***	0.023***	−0.671***	−0.199**	0.002	−68.418***	6.891	0.208***

4.6 Conclusion

This chapter investigates the impact of ETF-related trading on the underlying stock returns, testing concerns raised by market observers and regulators about potential dislocations in prices when considerable passively managed funds are rebalanced. It shows that changes in ETF holdings, which imply ETF-related trading of the underlying stocks, significantly impact the stock returns, but not necessarily in the way that market observers expect.

We decompose the changes in ETF holdings into two components: (i) changes caused by investor inflows and outflows into the ETF, which cause the ETF holdings of all stocks in its portfolio to expand and contract, and (ii) changes caused by index rebalancing, which involves an ETF substituting one stock for another stock in its portfolio (due to changes in the index composition) or changing the weights of the stocks it holds. While ETF flows create significant price pressure, rebalancing events do not impact stock returns as much as that believed previously. ETF flows are unexpected and generate sudden order–liquidity imbalances. In contrast, index rebalancing events are anticipated, and arbitrage activities can eliminate most of the index inclusion effects. In addition, the significant growth of ETFs has made holdings of stocks by passive funds more uniform and eradicated the index premium and discount.

The findings add insights into the concern that ETFs exacerbate the distorting effect of index rebalancing. ETFs not only create price pressure around rebalancing events but also help diminish the index premium and discount. We also show the temporary and permanent impacts of ETF flows. The former component suggests that ETFs create price pressure, which is a short-term distortion in prices that is quickly reverted. The latter component indicates that ETF-related trading contains information that fundamentally influences stock prices. This is consistent with the evidence that investors use ETFs as part of their active investment portfolios (Easley et al., 2021).

This chapter also reveals several limitations of the present study. First, the sample does not include all ETFs that are listed in the US and international ETFs that hold ETF stocks. Due to the nature of this study, ETFs are required to report their holdings daily. Thus, several large ETFs, such as Vanguard's, are not included. These funds form more than one-fifth of the total AUM of ETF. Their flow is large and can influence the results reported in this chapter. However, as many of these funds track the same indices as those of the ETFs included in our sample, the results would not be severely affected. Having more complete data on all ETFs might even make some results stronger. Another limitation of this study is the time of reporting trades. While many funds report changes in holdings as they occur, others report when the changes are settled. This misalignment in reporting time can make the price pressure appear weaker than it actually is because the effects are spread over several days. With an accurate reporting time of trades, the impact would concentrate on when the trades occur, and the results would be stronger.

5. Conclusion

This chapter summarizes the findings of this thesis, which are presented in the form of the following questions.

(i) How have tick size reductions in US stock markets impacted the liquidity of small stocks?

(ii) How do the tick size reductions influence a firm's propensity to go public?

(iii) How does ETF growth affect short-selling constraints?

(iv) How do ETFs impact market efficiency, as reflected in the returns of asset pricing anomalies?

(v) How do ETFs impact stock prices when they rebalance their portfolios or receive in/outflows?

This chapter ends with suggestions for future research.

5.1 How have the tick size reductions affected liquidity for small stocks?

Financial markets have come a long way since their inception. Technology and regulation have tremendously improved the overall market quality. The costs of trading stocks have decreased significantly, and liquidity for stocks has been enhanced considerably. However, these improvements might not be shared equally among stocks, leaving some stocks neglected. The existing literature lacks evidence regarding this inequality.

Chapter 2 of this thesis investigates the change in liquidity for small, medium, and large stocks in the US over the past two decades, particularly around changes in the tick size, using different measures of liquidity. Secondary market liquidity for all stocks, including small ones, has improved significantly since the late 1990s, when tick sizes were reduced. Quoted spreads have dropped by more than two-thirds, Amihud's (2002) illiquidity measure has decreased, and turnover has improved noticeably. Because some liquidity dimensions can be difficult to measure, company valuation provides an alternative overarching measure of the net effect of changes in different dimensions of liquidity. The results show that the reductions in tick size are associated with improved valuations of small stocks relative to their larger counterparts.

5.2 How does the frequency of stock splits affect IPOs?

If the reductions in tick size are bad for companies, corporate managers can effectively undo them using stock splits and by choosing initial public offering (IPO) prices that increase the relative tick size to predecimalization levels. The results from Chapter 2 show that the opposite occurs. The number of stock splits has diminished since the reductions in the tick size, IPO prices do not fall, and the distribution of stock price levels remains approximately unchanged. The results also indicate that liquidity and valuation of small stocks have improved after decimalization. This collection of results suggests that increasing the tick size is unlikely to encourage public listing of companies and reverse the trend in the number of listed companies, in contrast to policymaker proposals.

5.3 How does ETF growth affect short-selling constraints?

Anecdotal evidence suggests that ETFs lend out a substantial proportion of their underlying stocks to lower the management fees. Chapter 3 shows a negative relation between ETF ownership and short-sale constraints consistent with ETFs being active stock lenders. Through their security-lending activities, ETFs increase the supply of stocks that can be borrowed and reduce the borrowing fees, collectively reducing the short-selling constraints. The significance of this finding is that short-selling plays an important role in markets, helping to correct the pricing of stocks once they become overvalued. Consequently, through their stock-lending activities, ETFs are likely to inadvertently or indirectly contribute to market efficiency by reducing short-selling constraints.

5.4 How do ETFs impact market efficiency as reflected in returns of asset pricing anomalies?

Asset pricing anomalies, such as momentum, are an indicator of market inefficiency. They involve buying underpriced (long-leg) and selling overpriced (short-leg) stocks. Because it is easy to buy, the underpricing in stocks in the long leg tends to quickly dissipate. In contrast, it is much harder to borrow and short-sell stocks. The current literature argues that impediments to short-selling discourage investors from short-selling overpriced stocks, which prevents the incorporation of bad news into stock prices and allows mispricing in stocks to persist.

This thesis examines numerous anomalies that have been shown in the literature and analyzes the impact of ETFs on anomaly returns. Many anomalies have disappeared in the period following string growth in ETFs. A large proportion of the profit from these anomalies has originated from the short leg. The thesis shows a negative relation between ETF ownership and anomaly profit using methods that identify causality. Stocks with a high level of ETF ownership appear to have a lower level of mispricing. This effect is stronger among overpriced stocks (i.e., ones that are in the short leg), which we attribute to the previous finding that ETFs help reduce short-selling constraints. The evidence in this thesis suggests that ETFs positively contribute to market efficiency.

5.5 How do ETFs impact stock prices when they rebalance their portfolios or receive in/outflows?

This thesis shows that changes in ETF holdings of stocks create price pressure that distorts prices. There are two main sources of changes in ETF holdings. The first is flows from investors into and out of an ETF, which triggers trading of the underlying stocks through the creation/redemption mechanism. The second source is rebalancing the ETF portfolio when the constituents or weights in the index change. In such events, an ETF has to buy or sell stocks that are added or removed from the underlying index. The second source does not occur frequently because indices infrequently change their composition. In contrast, ETF flows occur frequently.

These two sources of changes in ETF holdings impact stock prices differently. This thesis shows that ETF flows create significant distortions in stock prices, whereas index rebalancing events do not. One of the reasons for this is that while ETF flows are not very predictable, index rebalancing events are. Thus, a large ETF flow creates a large order imbalance and price pressure, which distorts stock prices temporarily. In contrast, index rebalancing events are more anticipated and thus have less impact on stock prices. Moreover, the index premium phenomenon has diminished over time and virtually disappeared recently.

5.6 Avenues for future research

The decline in the number of listed stocks continues to be a concerning trend, as reflected in policymaker reports. While this thesis finds that changes in secondary markets are not to be blamed for the decline, investigating other causes of these trends is a worthwhile direction for future research. Further insights into the cause of the trend will not only enhance our understanding of financial markets but also help to revert this trend.

ETFs continue to become an even larger part of the investment landscape and affect financial markets in many ways. Future research can investigate the impact of ETFs using different measures of market quality, including their contribution to systemic risk, fragility of liquidity, competition between firms, and stock price synchronicity. Additionally, ETFs are not a homogeneous group. There are many types of ETFs, each having a different influence. Studying these different groups will enhance our knowledge about ETFs.

Finally, given the continuing growth of passive investing, will active investment disappear or become obsolete in the future? Who will carry out the research and information acquisition role and drive price discovery? These too are valuable lines of future research given the capital allocation role played by financial markets.

6. References

Albuquerque, R., Song, S., Yao, C., 2020. The price effects of liquidity shocks: A study of the SEC's tick size experiment. Journal of Financial Economics 138, 700–724.

Admati, A.R., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3–40.

Agarwal, V., Hanouna, P., Moussawi, R., Stahel, C., Robinson, J.M., 2018. Do ETFs increase the commonality in liquidity of underlying stocks? Unpublished working paper.

Aggarwal, R., Saffi, P.A.C., Stugess, J., 2015. The role of institutional investors in voting: Evidence from the securities lending market. Journal of Finance 70, 2309–2346.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223–249.

Amihud, Y., Mendelson, H., 2005. The liquidity route to a lower cost of capital. Applied Corporate Finance 12, 8–25.

Angel, J.J., 1997. Tick size, share prices, and stock splits. Journal of Finance 52, 655–681.

Anthonisz, S. A., and Tālis J. P., 2017. Asset pricing with downside liquidity risks. Management Science 63, 2549–2572.

Appel, I.R., Gormley, T.A., Keim, D.B., 2016. Passive investors, not passive owners. Journal of Financial Economics 121, 111–141.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T.J., Pedersen, L.H., 2018. Size matters, if you control your junk. Journal of Financial Economics 129, 479–509.

Bacidore, J., Battalio, R.H., Jennings, R.H., 2003. Order submission strategies, liquidity supply, and trading in pennies on the New York Stock Exchange. Journal of Financial Markets 6, 337–362.

Bakke, T. E., Whited, T.M., 2010. Which firms follow the market? An analysis of corporate investment decisions. Review of Financial Studies 23, 1941–1980.

Barth, M.E., Landsman, W.R., Taylor, D.J., 2017. The JOBS Act and information uncertainty in IPO firms. Accounting Review 92, 25–47.

Bartram, S.M., Brown, G., Stulz, R.M., 2012. Why are US stocks more volatile? Journal of Finance 67, 1329–1370.

Basak, S., Pavlova, A., 2013. Asset prices and institutional investors. American Economic Review 103, 1728–1758.

Beber, A., Pagano, M., 2013. Short-selling bans around the world: Evidence from the 2007–09 crisis. Journal of Finance 68, 343–381.

Ben-David, I., Franzoni, F., Moussawi, R., 2017. Exchange-traded funds (ETFs). Annual Review of Financial Economics 9, 169–189.

Ben-David, I., Franzoni, F., Moussawi, R., 2018. Do ETFs increase volatility? Journal of Finance 73, 2471–2535.

Beneish, M.D., Lee, C.M.C., Nichols, D.C., 2015. In short supply: Short-sellers and stock returns. Journal of Accounting and Economics 60, 33–57.

Bessembinder, H., 2003. Trade execution costs and market quality after decimalization. Journal of Financial and Quantitative Analysis 38, 747–777.

Bharath, S.T., Dittmar, A.K., 2010. Why do firms use private equity to opt out of public markets? Review of Financial Studies 23, 1771–1818.

Bhattacharya, A., O'Hara, M., 2018. Can ETFs increase market fragility? Effect of information linkages in ETF markets. Unpublished working paper.

Biais, B., Foucault, T., 2014. HFT and market quality. Bankers, Markets, and Investors 128, 5–19.

Blocher, J., Whaley, R.E., 2015. Passive investing: The role of securities lending. Research paper number 2474904. Vanderbilt Owen Graduate School of Management.

Boehmer, B., Boehmer, E., 2003. Trading your neighbor's ETFs: Competition or fragmentation? Journal of Banking and Finance 27, 1667–1703.

Boehmer, E., Fong, K.Y., Wu, J.J., 2013. Algorithmic trading and changes in firms' equity capital. Unpublished working paper.

Boehmer, E., Jones, C.M., Zhang, X., 2008. Which shorts are informed? Journal of Finance 63, 491–528.

Boehmer, E., Jones, C.M., Zhang, X., 2013. Shackling short sellers: The 2008 shorting ban. Review of Financial Studies 26, 1363–1400.

Boone, A.L., White, J.T., 2015. The effect of institutional ownership on firm transparency and information production. Journal of Financial Economics 117, 508–533.

Brown, D.C., Davies, S., Ringgenberg, M., 2020. ETF arbitrage, non-fundamental demand, and return predictability. Review of Finance, forthcoming.

Busse, J.A., Tong, Q., 2012. Mutual fund industry selection and persistence. Review of Asset Pricing Studies 2, 245–274.

Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance 56, 1–43.

Chakravarty, S., Wood, R.A., Van Ness, R.A., 2004. Decimals and liquidity: A study of the NYSE. Journal of Financial Research 27, 75–94.

Chan, K., Hameed, A., 2006. Stock price synchronicity and analyst coverage in emerging markets. Journal of Financial Economics 80, 115–147.

Chen, Q., Goldstein, I., Jiang, W., 2010. Payoff complementarities and financial fragility: Evidence from mutual fund outflows. Journal of Financial Economics 97, 239–262.

Christophe, S.E., Ferri, M.G., Angel, J.J., 2004. Short-selling prior to earnings announcements. Journal of Finance 59, 1845–1875.

Christophe, S.E., Ferri, M.G., Hsieh, J., 2010. Informed trading before analyst downgrades: Evidence from short sellers. Journal of Financial Economics 95, 85–106.

Chung, K.H., Lee, A.J., Rosch, D., 2020. Tick size, liquidity for small and large orders, and price informativeness: Evidence from the tick size pilot program. Journal of Financial Economics 136, 879–899.

Coates, J.C., Srinivasan, S., 2014. SOX after ten years: A multidisciplinary review. Accounting Horizons 28, 627–671.

Cohen, L., Diether, K.B., Malloy, C.J., 2007. Supply and demand shifts in the shorting market. Journal of Finance 62, 2061–2096.

Coles, J.L., Heath, D., Ringgenberg, M.C., 2017. On index investing. Unpublished working paper.

Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. Journal of Financial Economics 86(2), 479–512.

Cremers, K.M., Petajisto, A., 2009. How active is your fund manager? A new measure that predicts performance. Review of Financial Studies 22, 3329–3365.

Da, Z., Shive, S., 2017. Exchange traded funds and asset return correlations. European Financial Management 24, 136–168.

Dambra, M., Field, L.C., Gustafson, M.T., 2015. The JOBS Act and IPO volume: Evidence that disclosure costs affect the IPO decision. Journal of Financial Economics 116, 121–143.

Davis, R.L., Maslar, D.A., Roseman, B.S., 2017. Secondary market trading and the cost of new debt issuance. Unpublished working paper.

Decker, R.A., Haltiwanger, J., Jarmin, R.S., Miranda, J., 2016. Where has all the skewness gone? The decline in high-growth (young) firms in the US. European Economic Review 86, 4–23.

Diamond, D.W., Verrecchia, R.E., 1987. Constraints on short-selling and asset price adjustment to private information. Journal of Financial Economics 18, 277–311.

Diether, K.B., Lee, K.H., Werner, I.M., 2009. Short-sale strategies and return predictability. Review of Financial Studies 22, 575–607.

Doidge, C., Karolyi, G.A., Stulz, R.M., 2017. The US listing gap. Journal of Financial Economics 123, 464–487.

Drechsler, I., Drechsler, Q.F., 2014. The shorting premium and asset pricing anomalies. NBER working paper (number 20282).

Durnev, A., Morck, R., Yeung, B., Zarowin, P., 2003. Does greater firm-specific return variation mean more or less informed stock pricing? Journal of Accounting Research 41, 797–836.

Easley, D., Michayluk, D., O'Hara, M., Putnins, T.J., 2021. The active world of passive investing. Review of Finance, forthcoming.

Edmans, A., 2009. Blockholder trading, market efficiency, and managerial myopia. Journal of Finance 64, 2481–2513.

Edmans, A., Goldstein, I., Jiang, W., 2012. The real effects of financial markets: The impact of prices on takeovers. Journal of Finance 67, 933–971.

Ellul, A., Pagano, M., 2006. IPO underpricing and after-market liquidity. Review of Financial Studies 19, 381–421.

Engelberg, J.E., Reed, A.V., Ringgenberg, M.C., 2012. How are shorts informed?: Short sellers, news, and information processing. Journal of Financial Economics 105, 260–278.

Evans, R., Ferreira, M.A., Prado, M.P., 2017. Fund performance and equity lending: Why lend what you can sell? Review of Finance 21, 1093–1121.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Foucault, T., Gehrig, T., 2008. Stock price informativeness, cross-listings, and investment decisions. Journal of Financial Economics 88, 146–168.

Gabaix, X., Koijen, R.S., 2020. In search of the origins of financial fluctuations: The inelastic markets hypothesis. Unpublished working paper.

Gao, X., Ritter, J.R., Zhu, Z., 2013. Where have all the IPOs gone? Journal of Financial and Quantitative Analysis 48, 1663–1692.

Glosten, L.R., Milgrom, P.R., 1985. Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.

Goldstein, I., Ozdenoren, E., Yuan, K., 2013. Trading frenzies and their impact on real investment. Journal of Financial Economics 109, 566–582.

Greenwood, R., 2005. Short- and long-term demand curves for stocks: Theory and evidence on the dynamics of arbitrage. Journal of Financial Economics 75, 607–649.

Griffith, T., Roseman, B.S., 2019. Making cents of tick sizes: The effect of the 2016 US SEC tick size pilot on limit order book liquidity. Journal of Banking and Finance 101, 104–121.

Gromb, D., Dimitri V., 2010. Limits of arbitrage: The state of the theory. Annual Review of Financial Economics 2, 251–275.

Grullon, G., Larkin, Y., Michaely, R., 2019. Are US industries becoming more concentrated? Review of Finance 23, 697–743.

Gustafson, M.T., Iliev, P., 2017. The effects of removing barriers to equity issuance. Journal of Financial Economics 124, 580–598.

Hameed, A., Kang, W., Viswanathan, S., 2010. Stock market declines and liquidity. Journal of Finance 65(1), 257–293.

Hanley, K.W., 2017. The economics of primary markets. Unpublished working paper.

Hansen, P., Li, Y., Lunde, A., Patton, A., 2017. Mind the gap: An early empirical analysis of SEC's "Tick size pilot program." Unpublished working paper.

Harris, L., Gurel, E., 1986. Price and volume effects associated with changes in the S&P 500: New evidence for the existence of price pressures. Journal of Finance 41(4), 815–829.

Harrison, J.M., Kreps, D.M., 1978. Speculative investor behavior in a stock market with heterogeneous expectations. Quarterly Journal of Economics 92, 323–336.

Hasbrouck, J., 2003. Intraday price formation in US equity index markets. Journal of Finance 58, 2375–2399.

Hayo, B., Kutan, A.M., Neuenkirch, M., 2010. The impact of US central bank communication on European and Pacific equity markets. Economics Letters 108, 172–174.

Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does algorithmic trading improve liquidity? Journal of Finance 66, 1–33.

Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors over-value firms with bloated balance sheets? Journal of Accounting and Economics 38, 297–331.

Huang, S., O'Hara, M., Zhong, Z., 2021. Innovation and informed trading: Evidence from industry ETFs. Review of Financial Studies 34, 1280–1316.

Hutton, A.P., Marcus, A.J., Tehranian, H., 2009. Opaque financial reports, R2, and crash risk. Journal of Financial Economics 94, 67–86.

Iliev, P., Lowry, M., 2015. Are mutual funds active voters? Review of Financial Studies 28, 446–485.

Israeli, D., Lee, C.M.C., Sridharan, S.A., 2017. Is there a dark side to exchange-traded funds? An information perspective. Review of Accounting Studies 22, 1048–1083.

Ivanov, S.I., Jones, F.J., Zaima, J.K., 2013. Analysis of DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 ETFs and their influence on price discovery. Global Finance Journal 24, 171–187.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65–91.

Jin, L., Myers, S.C., 2006. R^2 around the world: New theory and new tests. Journal of Financial Economics 79, 257–292.

Jones, C.M., 2013. What do we know about high-frequency trading? Unpublished working paper.

Karpoff, J.M., Lou, X., 2010. Short sellers and financial misconduct. Journal of Finance 65, 1879–1913.

Krause, T., Ehsani, S., Lien, D., 2014. Exchange-traded funds, liquidity, and market volatility. Applied Financial Economics 24, 1617–1630.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Lehmann, B.N., 2002. Some desiderata for the measurement of price discovery across markets. Journal of Financial Markets 5, 259–276.

Lettau, M., Madhavan, A., 2016. ETF 101 for economists. Journal of Economic Perspectives 32, 135–154.

Li, F.W., Zhu, Q., 2016. Synthetic shorting with ETFs. Unpublished working paper.

Linnainmaa, J.T., Roberts, M.R., 2018. The history of the cross-section of stock returns. Review of Financial Studies 31, 2606–2649.

Ljungqvist, A., Persson, L., Tåg, J., 2016. Private equity's unintended dark side: On the economic consequences of excessive delistings. NBER working paper (number 21909).

Lou, D., 2012. A flow-based explanation for return predictability. Review of Financial Studies 25(12), 3457–3489.

Loughran, T., Ritter, J.R., 1995. The new issues puzzle. Journal of Finance 50, 23–51.

Lowry, M., Michaely, R., Volkova, E., 2017. Initial public offering: A synthesis of the literature and directions for future research. Forthcoming Foundations and Trends in Finance.

Luo, Y., 2005. Do insiders learn from outsiders? Evidence from mergers and acquisitions. Journal of Finance 60, 1951–1982.

Madhavan, A., 2012. Exchange-traded funds, market structure, and the flash crash. Financial Analysts Journal 68, 20–35.

Madhavan, A., Sobczyk, A., 2016. Price dynamics and liquidity of exchange-traded funds. Journal of Investment Management 14, 1–17.

Malkiel, B., 2003. Passive investment strategies and efficient markets. European Financial Management 9, 1–10.

Markowitz, H., 1952. The utility of wealth. Journal of political Economy 60, 151–158.

Massa, M., Xu, M., 2013. The value of (stock) liquidity in the M&A market. Journal of Financial and Quantitative Analysis 48, 1463–1497.

Mauboussin, M.J., Callahan, D., Majd, D., 2017. The incredible shrinking universe of stocks. Unpublished working paper.

Mauer, D.C., Senbet, L.W., 1992. The effect of the secondary market on the pricing of initial public offerings: Theory and evidence. Journal of Financial and Quantitative Analysis 27, 55–79.

Mclean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? Journal of Finance 71, 5–32.

Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. Journal of Finance 32, 1151–1168.

Mitchell, M., Pulvino, T., Stafford, E., 2004. Price pressure around mergers. Journal of Finance 59, 31–63.

Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: Why do emerging markets have synchronous stock price movements? Journal of Financial Economics 58, 215–260.

Petajisto, A., 2011. The index premium and its hidden cost for index funds. Journal of Empirical Finance 18, 271–288.

Prado, M.P., Saffi, P.A.C., Sturgess, J., 2016. Ownership structure, limits to arbitrage, and stock returns: Evidence from equity lending markets. Review of Financial Studies 29, 3211–3244.

Roll, R., 1988. The international crash of October 1987. Financial Analyst Journal 44, 19–35.

Scholes, M.S., 1972. The market for corporate securities: Substitution versus price pressure and the effects of information on stock prices. Journal of Business 45(2), 179–211.

Schultz, P., 2000. Stock splits, tick size, and sponsorship. Journal of Finance 55, 429–450.

Sharpe, W.F., 1991. The arithmetic of active management. Financial Analysts Journal 47, 7–9.

Shleifer, A., 1986. Do demand curves for stocks slope down? Journal of Finance 41(3), 579–590.

Song, S., Yao, C., 2017. The real costs of natural experiments. Unpublished working paper.

Spilimbergo, A., 2009. Democracy and foreign education. American Economic Review 99, 528–543.

Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288−302.

Stulz, R.M., Vagias, D., Van Dijk, M.A., 2013. Do firms issue more equity when markets are more liquid? NBER working paper no. 19229.

US Securities Exchange Commission, 2012. Report to congress on decimalization. Washington, DC.

Weild, D., Kim, E., 2010. Market structure is causing the IPO crisis—and more. Grant Thornton Capital Markets Series.

Weild, D., Kim, E., Newport, L., 2013. Making stock markets work to support economic growth: Implications for governments, regulators, stock exchanges, corporate issuers and their investors. OECD Corporate governance working paper no.10

Wu, Y., 2019. Closing auction, passive investing, and stock prices. Unpublished working paper.

Wurgler, J., 2000. Financial markets and the allocation of capital. Journal of Financial Economics 58, 187−214.