THREE ESSAYS ON EXCHANGE-TRADED FUNDS

by

DANIEL ELIJAH SHERRILL

DOUGLAS O. COOK, COMMITTEE CHAIR

SHANE E. UNDERWOOD
ROBERT W. MCLEOD
JUNSOO LEE
BRUCE E. BARRETT

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Economics, Finance, and Legal Studies
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2014

ABSTRACT

This dissertation consists of three essays on exchange-traded funds (ETFs). The dissertation research seeks to contribute to a deeper understanding of the impact of ETFs upon the financial markets, discover insights into the realm of performance persistence, and identify the factors leading to ETF liquidations.

The first essay investigates the impact that sector exchange-traded funds have upon stocks that they hold. We find that sector ETF ownership is associated with stock return comovement, especially with other industry stocks that are also held by sector ETFs. We show that sector ETF ownership is related to a muted abnormal return and trading volume reaction to earnings surprises. Even when considering other types of institutional investors, sector ETFs appear to be the main driver behind these findings.

The second essay documents the existence of ETF performance persistence. This calls into question interpretations used in the mutual fund literature suggesting performance persistence is evidence of manager skill. Given their passive nature, performance persistence should not exist amongst ETFs if the sole source of this persistence is manager skill. A decomposition of performance into stock composition and industry exposure sources reveals that this persistence is attributable predominately to a fund's industry exposure. Furthermore, the underlying source of the persistence is a flow-driven return effect where fund flows place price pressure on stocks leading to persistence in fund returns. An industry flow-based explanation

best accounts for positive persistence of winners while stock flow-based reasons better explain persistence of past losers.

The third essay studies the determinants of ETF liquidations. Investors are subject to tax, trading, and search costs as a result of holding a liquidated fund. I find that fund size and flows are essential to a fund's survival. Larger fund families are also more likely to produce funds that will avoid liquidation. Funds that are latecomers to a trending category that subsequently underperforms are less likely to survive. Finally, I find that the average investor holding a fund with an upcoming liquidation is best served to immediately sell the liquidating fund and purchase other funds in the same category.

DEDICATION

This dissertation is dedicated to my parents, Gracie and Randy Sherrill, and my grandmother, Betty Russell. Their loving support has been a constant source of encouragement, inspiration, and stability in my life. I am forever thankful for all they have done for me and know that their guidance has been instrumental in my success.

LIST OF ABBREVIATIONS AND SYMBOLS

α	Performance alpha
α_{Gross}	Performance alpha estimated using gross returns
α_{Industry}	Performance alpha estimated using gross returns that is attributable to a fund's industry exposure
α_{Stock}	Performance alpha estimated using gross returns that is attributable to a fund's stock composition
Ab. Vol.	Abnormal volume
β	Regression coefficient, beta
CAR	Cumulative abnormal return
CRSP	Center for Research in Security Prices
$\varepsilon_{i,t}$	Standard error of fund i at time t
E[FIT]	Expected flow-induced trading
EPS	Earnings per share
ETF	Exchange-traded fund
FE	Forecast error
HML	Fama and French high minus low factor
NAV	Net asset value
R^2	R-squared
$R_{f,t}$	Risk-free rate of return (proxied by 3 month Treasury bill rate)

$R_{Gross,\ t}$	Gross return at time *t*
$R_{i,t}$	Return of fund i at time t
$R_{Industry,\ t}$	Gross return at time *t* that is attributable to a fund's industry exposure
$R_{SIC,t}$	Value-weighted return of the standard industrial classification industry
$R_{SIC\ ETF\ Own,t}$	Value-weighted return of stocks in the standard industrial classification industry with sector ETF ownership
$R_{SIC\ ETF\ No\ Own,t}$	Value-weighted return of stocks in the standard industrial classification industry without sector ETF ownership
$RMRF_t$	Excess return on the market
SIC	Standard industrial classification
SMB	Fama and French small minus big factor
t	Computed value of *t* test
UMD	Carhart momentum factor
U.S.	United States of America
+	Additions
=	Equal to

ACKNOWLEDGMENTS

I would like to thank the many colleagues, friends, and faculty members who have supported me in the completion of this dissertation. I am most indebted to Douglas O. Cook, my dissertation chair, for sharing his research expertise and guidance throughout the dissertation process. I appreciate the vast amount of time, effort, and patience he exhibited toward me. His influence is apparent throughout this dissertation and has made me a much better researcher. I would also like to thank all of my committee members: Shane E. Underwood, Robert W. McLeod, Junsoo Lee, and Bruce E. Barrett for their invaluable comments and suggestions. Their ideas greatly shaped the essays included in this dissertation.

I am thankful to Binay Adhikari for his helpful recommendations and practical help during our time as officemates. I gratefully acknowledge my other colleagues and friends who encouraged me throughout my graduate school career and provided valuable insights and assistance. Specifically, I would like to thank the following individuals: Richard Arnatt, Brad Daughdrill, Collin Gilstrap, James Malm, Heather Rhodes, Kate Upton, Tony Via, and Ben Woodruff. I also extend my gratitude to the staff and faculty of the Department of Economics, Finance, and Legal Studies at The University of Alabama for their instruction, assistance, and kindness during my time in graduate school.

Finally, my heartfelt thanks go to all my family and friends who supported me during this endeavor and encouraged me to pursue my dreams. My parents, Randy and Gracie Sherrill, have made a tremendous impact in my life and never cease to demonstrate their love, patience,

encouragement, and support towards me. Words cannot express just how thankful I am for them and how important they are to me. To my grandma, Betty B. Russell, I thank for always brightening my days and being a steadfast source of encouragement. I gratefully recognize Dennis and Debra Russell, Nila and Jerry Sherrill, Janice Weber, and my late grandfather, Tom Russell, for their loving support. The prayers and words of encouragement by Kay Griffin were a source comfort during times of great trial. To my friends Rachael Albury, Daniel Burge, Daniel Ochocki, and Karly Wilson, I am thankful for making my final months in Alabama full of happy memories and reminding me of what is most important in life. I am also deeply thankful to my close friends who have stood by me throughout my time completing this process and who remind me to never let the little kid inside of you ever fully grow up. For bringing tremendous joy to my life, I thank Noelle Bondy, Lyndsay Brown, Whitney Crouch, Drew and Jenny Highsmith, Jin Kang, Rebekah Malcom, Chad Rowell, Brielle and Matt Shinall, Jenevieve Solomon, and David Webb. In particular, I thank Stephen Malcom for being my best friend and the closest thing to a brother that I will ever know. To God be all the glory, honor, and praise.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This dissertation consists of three essays on exchange-traded funds or, as they are commonly referred to, ETFs. ETFs have grown tremendously both in size and influence in recent years; however, the literature on the topic is still sparse. While there is a myriad of papers that study mutual funds, mutual funds and exchange-traded funds have numerous key differences that make both worthy of individual study. The research in this dissertation seeks to further develop our understanding of ETFs and their importance in the financial markets.

The first essay studies the impact of ETFs upon the financial markets. As ETFs have become a more popular investment choice, so to have the concerns that they negatively impact the assets that they hold. We find that ETFs, specifically ETFs that focus their holdings on a specific sector of the economy, have a significant impact upon stocks that they hold. First, they are strongly related with stock return comovement where stock returns follow the returns of their respective industry to a greater degree. We find that higher sector ETF ownership is strongly related to return comovement. In particular, stocks with higher sector ETF ownership tend to comove more with other industry stocks held by sector ETFs and less with industry stocks not held by sector ETFs. We further demonstrate this by showing a significant increase in comovement for stocks added to new ETFs for the first time that exceeds the change in comovement for a matched sample of stocks not held by sector ETFs. Furthermore, we document a dampened initial reaction to important information about firm fundamentals

associated with higher sector ETF ownership. Firms with earnings announcement surprises tend to have muted abnormal return and abnormal volume reactions when sector ETFs have a greater presence. This is consistent with the trading activities of ETFs which trade to track an index rather than use firm fundamentals to actively select stocks. Finally, we consider if other types of institutional holders are the driver behind these results. We find sector ETF ownership not only holds but also appears to be the main driver behind these results when also considering non-sector ETFs, mutual funds, and other institutional ownership. Overall, this essay illustrates that sector ETFs play an important role in the financial markets and impact the stocks that they hold.

In the second essay, I revisit a popular topic from the mutual fund literature of performance persistence and manager skill. Numerous mutual fund studies have used performance persistence as evidence of manager skill. However, I find the same performance persistence exists amongst passively managed ETFs where manager skill does not come into play. A decomposition of performance into that attributable to the industry exposure and stock composition demonstrates that a fund's industry exposure is the main driver behind both performance and performance persistence for ETFs. Furthermore, the underlying source is mainly attributable to an industry flow-driven return effect. Since investors chase returns, funds with high past performance generally see higher future inflows which they then reinvest into the same securities. Industries with high inflows from ETFs then may experience upward price-pressure. If an ETF holds many of these price-pressured industries, it tends to have higher short-term fund returns. This fully accounts for all the positive performance persistence found for ETFs as well as a majority of the negative performance persistence. The flow-induced return effect attributable to the specific composition of stocks within the fund is even better at explaining negative performance persistence but virtually none of the positive persistence. One

of the main implications of this essay is that factors other than manager skill can drive performance persistence and that caution should be used when inferring that performance persistence alone is evidence of manager skill.

In my third essay, I analyze the determinants of ETF liquidations. There has been a sharp increase in the number of ETFs liquidating since the late 2000's. Investors are subject to tax, trading, and search costs as a result of holding a liquidated fund. There are a number of factors investors should be aware of when considering the likelihood of it liquidating. Small funds with low flows are less likely to survive. Also, the health of the fund family is important to an individual fund's survival. Families with better health and a more established presence in the industry are more likely to produce funds that will survive. Funds that are latecomers to a trending category that subsequently underperforms are less likely to survive. Finally, I find that the average investor holding a fund with an upcoming liquidation is best served to immediately sell the liquidating fund and purchase other funds in the same category, especially if transaction costs are minimal.

CHAPTER 2[1]

THE IMPACT OF SECTOR ETFS ON STOCK RETURN COMOVEMENT AND ON STOCK PRICE MOVEMENTS

2.1 Introduction

Exchange-traded funds (ETFs) were first introduced in the United States in 1993 when State Street Global Advisors launched a fund that tracked the S&P 500 index. During the twenty subsequent years, ETFs have grown to over $1.5 trillion invested in all funds by the end of July 2013[2]. One of the most popular ETF fund types is the sector fund. State Street created the first nine sector funds in 1998 focusing on individual sectors such as health care, consumer staples, energy, and technology. Today, around one in every four ETFs focuses on a specific sector. Figure 2.1 illustrates the surge in growth that began around 2007. Since the beginning of 2007, the total assets managed by sector ETFs have tripled. Given the size of this market, this paper analyzes the influence of sector ETFs upon financial markets.

Today, there are a multitude of people who suggest that ETFs in general have a significant negative impact upon the market. For example, in October 2011, the U.S. Senate held a hearing to investigate whether trading in ETFs lead to an increase in market volatility. At the hearing, Harold Bradley, chief investment officer at Kauffman Foundation, claimed that ETFs were undermining the fundamental determinants of equity pricing, noting: "When individual common stocks increasingly behave as if they are derivatives of frequently traded and interlinked

[1] A working paper version of this chapter co-authored with Dr. Douglas O. Cook exists and is being circulated.
[2] Source: Investment Company Institute website

ETF baskets, then it is trading in the ETFs that is driving the prices of the underlying stocks rather than the other way around."[3]

ETFs have become a major part of the portfolios of many large institutions, including hedge funds. A November 2012 Goldman Sachs report finds that hedge funds have gross exposure of $112 billion to ETFs, over 8.5% of their total equity exposure[4]. One of the main attractions of ETFs for hedge funds is liquidity. Between 2008 and 2012, ETFs accounted for over 29% of overall monthly stock trading volume for stocks reported in the Center for Research in Security Prices (CRSP) stock database. Figure 2.2 illustrates the trading volume for ETFs and their respective fraction of overall market trading volume. The value of ETF liquidity is especially apparent in turbulent markets, as seen in Figure 2.2 at the end of 2008. More recently, following an announcement by Federal Reserve Chairman Ben Bernanke in June 2013, ETF trade volume exceeded $1 billion and accounted for nearly 40% of the overall stock-trading volume for the trading day[5]. Such an influential role in the market leads to concern over the impact of ETFs on the pricing of stocks.

Since ETFs mimic a chosen index, transactions are not motivated by fundamental or technical analysis. Furthermore, ETFs make large basket trades that involve numerous stocks and are frequently the vehicle of choice for large traders such as hedge funds that wish to avoid impacting stock prices and liquidity with large trades. If the trading price of the ETF deviates from net asset value, traders known as authorized participants can arbitrage the difference. To do this, they create or redeem ETF shares in large blocks that typically consist of 50,000-100,000 shares. This generates large, contemporaneous trading in all the underlying securities. Since

[3] http://www.banking.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=fedd383b-6946-403c-b43d-7ea9723f17f5

[4] Goldman Sachs Hedge Fund Trend Monitor dated November 19, 2012. Statistic includes both long and short positions.

[5] http://www.usatoday.com/story/money/personalfinance/2013/06/20/etf-volume-surges/2443651/

sector ETFs focus all their trading upon a specific sector, it is possible that the returns of the underlying stocks held by these funds may correlate more with the returns of the industry while firm fundamentals may correlate less. In this paper, we investigate the nature of this relationship by analyzing comovement and earnings surprises.

We find that comovement with a firm's industry is increasing in the fraction of firm stock shares outstanding held by sector ETFs (hereafter, ETF refers specifically to sector ETF unless otherwise stated in chapter 2). The comovement of ETF owned stocks with other stocks owned by ETFs is especially strong. Additionally, we find that ETF overweighting has a strong, positive correlation with comovement. Furthermore, overweighting is associated with increased comovement with other industry stocks held by ETFs but decreased comovement with industry stocks not held by ETFs. Together these results suggest that both the degree of ETF ownership and the relative stock weight within ETF portfolios are important factors related to comovement.

In the spirit of a quasi-natural experiment, we analyze the change in comovement for stocks that are added to new ETFs which had not been included in any prior ETF portfolios. Difference-in-difference tests indicate that there is a significant increase in return comovement, especially with other ETF-held stocks in the industry, when a stock is added to an ETF portfolio. This increase exceeds that for a comparison group matched on industry, size, and price-to-book ratio.

Finally, we find that ETF ownership is related to a muted response surrounding an earnings surprise announcement, both in terms of abnormal returns and abnormal trading volume, as well as an increased post-announcement drift. Thus it appears that stocks do not incorporate fundamental news as quickly nor is the immediate reaction as strong when ETFs are large stakeholders.

We consider the possibility that our results are driven by other types of institutional investors, such as mutual funds and non-sector ETFs. Even after controlling for ownership by these institutions, not only does the significant positive relation between sector ETF ownership and comovement hold, but sector ETFs appear to be the main driver of results. Although non-sector ETFs are related to comovement, this connection is weaker than that with sector ETFs. Furthermore, sector ETFs are the only institutional ownership group connected with a dampening effect upon abnormal returns around earnings surprises. Sector ETFs are also associated with a muted abnormal volume reaction around earnings surprises that is much stronger than that for other institutional owners. Overall, our results illustrate a clear relation between sector ETF ownership and stock return comovement with their underlying industry while there is less reaction to changes in firm fundamentals.

Our paper fits into a growing body of literature that studies comovement. Barberis and Shleiferr (2003) use a theoretical framework to show that style investing leads to comovement. Other work shows comovement based upon S&P 500 membership (Barberis, Shleiferr, and Wurgler, 2005), firm headquarters (Pirinsky and Wang, 2006), price level (Green and Hwang, 2009), index labels (Boyer, 2011), option listing (Agyei-Ampomah and Mazouz, 2011), and underwriter for equity offerings (Grullon, Weston, and Underwood, 2014). Greenwood (2008) finds that stocks whose weight in the Nikkei 225 stock index exceeds their value-weight exhibit increased comovement with the other stocks in the index but decreased comovement with those not in the index. Overall, the literature supports that factors other than fundamentals can significantly influence stock movements.

In addition, this paper is similar to the literature on mitigating factors related to earnings surprises and corresponding market responses. Several authors find that the market underreacts

to earnings announcements made after trading hours (Francis, Pagach, and Stephan; 1992), during low trading volume and down markets (Hou, Xiong, and Peng; 2009), Fridays (Dellavigna and Pollet, 2009), and when many companies make earnings announcements on the same day (Hirshleifer, Lim, and Teoh; 2009). Potter (1992) documents that institutional ownership is positively related to price variability around earnings announcements and Bartov, Radhakrishnan, and Krinsky (2000) find a positive relation between institutional ownership and post-earnings-announcement drift.

The rest of the paper is organized as follows. Section 2.2 describes the data and methodology. Empirical results are presented in Section 2.3 and Section 2.4 concludes.

2.2 Data and Methodology

We collect data on ETF stock holdings from Thomson Reuters Mutual Fund Holdings database (formerly known as CDA/Spectrum) and the CRSP Survivor-Bias Free Mutual Fund database. In order to identify which funds in the Thomson Reuters database are also ETFs, we hand-match these with those funds in CRSP know to be ETFs. We identify all ETFs in the CRSP database and obtain the fund name, inception date, and monthly total net assets. Then using MFLINKS, we identify funds in the Thomson Reuters database that are mutual funds so as to exclude them from consideration as potential ETFs. Notably this does not completely reduce our sample to all ETFs though as MFLINKS contains only a partial list of mutual fund matches between the CRSP and Thomson Reuters database. We then match each fund from the list of ETFs in the CRSP database with potential matches in the Thomson Reuters database first on name but then on size if there are multiple possible matches. In the process, we eliminate funds with holdings data prior to the ETF inception date. Thomson Reuters holdings data is provided

on a quarterly basis while the CRSP holdings data includes monthly observations. Whenever available, we use the monthly CRSP data in order to use the most timely holdings information.

A noteworthy point concerns Vanguard ETFs. Vanguard has a unique (and patented) structure in which their ETFs are treated as a share class of its mutual fund tracking the same index. As a consequence, the reported holdings data in both CRSP and Thomson Reuters databases are the combined holdings for both the mutual and exchange-traded funds. We use the total net assets reported by CRSP to adjust the holdings to only reflect those in the ETF.

We confirm matches of 959 equity ETFs, both sector and non-sector, with holdings data jointly in the two databases out of 1,167 total ETFs listed in the CRSP mutual fund summary database. In order to gauge the degree of coverage each year, we average the number of funds as well as the fraction of total net assets included in the sample relative to that reported in the CRSP mutual fund database of ETFs. We show the average quarterly level of sector ETF coverage between the CRSP and Thomson Reuters databases in Figure 2.3. The third quarter of 2003 marks the first time that the data consistently includes more than half of the quarterly ETF total net assets. This time period also corresponds to the beginning of ETF coverage in the CRSP holdings database; prior to this date, Thomson Reuters was the sole source of ETF holdings data. Additionally, Thomson Reuters began gathering ETF holding data from semi-annual N-CSR forms in 2003 as well as first and third quarter N-Q filings starting in 2004 (previously information came from N-30D forms). In order to avoid potential biases from the limited data coverage in the early years, we begin our sample in the third quarter of 2003. The sample ends in 2012.

To identify the comovement for each stock, we estimate the following regression:

$$R_{it} = \alpha_i + \beta_{SIC,i} R_{SIC,t} + \varepsilon_{it} \tag{1}$$

where R_{it} is the return of stock *i* on day *t* and $R_{SIC,t}$ is the value-weighted return on day *t* of stock *i*'s four-digit standard industrial classification (SIC) industry. Stock *i* is excluded from the industry return so as to avoid any mechanical correlation between the stock and industry returns. This is estimated using all daily returns within the quarter. This produces two measures of comovement: the R^2 value and the slope parameter, beta. Both of these measures of comovement are commonly used in the literature (Barberis, Shleifer, and Wurgler, 2005; Greenwood, 2008; Wahal and Yavuz, 2013). A higher value of these two measures indicates a greater degree of comovement with the industry.

Additionally, we compare the comovement between the industry group with sector ETF ownership and the industry group without sector ETF ownership:

$$R_{it} = \alpha_i + \beta_{SIC\ ETF\ Own,i} R_{SIC\ ETF\ Own,t} + \beta_{SIC\ ETF\ No\ Own,i} R_{SIC\ ETF\ No\ Own,t} + \varepsilon_{it} \qquad (2)$$

where $R_{SIC\ ETF\ Own,t}$ is the value-weighted return of all stocks in the SIC industry with the sector ETF ownership and $R_{SIC\ ETF\ No\ Own,t}$ is the value-weighted return of all stocks in the SIC industry without ownership by sector ETFs. Stock *i* is again excluded from the industry return calculation.

To test for the impact of stock ownership by ETFs, we use two measures. The first measure, *Fraction Owned*, is used to test whether the degree of ownership of a stock by ETFs impacts comovement. It is defined as the fraction of stock shares outstanding held by ETFs:

$$Fraction\ Owned_{it} = \frac{Shares\ held\ by\ ETFs_{it}}{Shares\ Outstanding_{it}}. \qquad (3)$$

The second measure, *Overweight*, draws upon the methodology in Greenwood (2008) to measure the degree that ETFs overweight a stock relative to its value-weight. It is defined as the natural log of one plus the stock weighting within the SIC industry by ETFs divided by the stock value-weight in the SIC industry:

$$Overweight_{it} = log\left(1+\frac{\text{Weight in SIC Amongst ETFs}_{it}}{\text{Weight in SIC Overall}_{it}}\right) = log\left(1+\frac{\text{\$ ETF holdings of stock / \$ ETF holdings in SIC}_{it}}{\text{Stock Market Cap /SIC Market Cap}_{it}}\right) \quad (4)$$

Together, these two measures provide a good indication of the impact of ETF ownership of stocks. First, *Fraction Owned* measures the importance of ETF ownership relative to other owners in the stock. Second, *Overweight* tests the relative importance of the stock weight within the ETF portfolio.

We next show the relation between the measures of comovement from Equation 1 and *Fraction Owned* using the following cross-sectional regressions:

$$R^2{}_{it} = a_i + b_t * Fraction\ Owned_{i,t} + u_{it}\ ,$$

and (5)

$$\beta_{it} = a_i + b_t * Fraction\ Owned_{i,t} + u_{it}\ .$$

Similar regressions use *Overweight* as the independent variable as well as $\beta_{SIC\ ETF\ Own,it}$ and $\beta_{SIC\ ETF\ No\ Own,it}$ as the dependent variable. Given past literature (Banz, 1981; Fama and French,1992) showing that size and book-to-market ratio contribute to commonality in average returns, Greenwood (2008) suggests adding the natural log of the price-to-book ratio and the natural log of market capitalization as controls:

$$R^2{}_{it} = a_i + b_t * Fraction\ Owned_{i,t} + c_t * log(P/B)_{i,t} + d_t * log(Market\ Cap)_{i,t} + u_{it}\ . \quad (6)$$

Similar regressions use betas as the dependent variable and *Overweight* as the independent variable. Compustat data is used to calculate the price-to-book ratio.

2.3 Empirical Results

2.3.1 Comovement

We begin by analyzing the comovement of stocks based upon the degree of ETF ownership. Table 2.1 compares the measures of comovement from Equation 1, *R*-squared and

beta, and Equation 2, beta with the ETF and non-ETF ownership group. Table 2.1 presents differences in comovement based on ETF ownership for stocks with zero ETF ownership versus those with ETF ownership. For summary statistics, ETF owned stocks are ranked into terciles based upon the fraction of ownership held by ETFs. The results show a significantly higher level of comovement for stocks with the highest amount of ETF ownership compared to those without ETF ownership. For example, using *R*-squared as the measure of comovement, the SIC industry returns explain 38% of stock returns for stocks in the top tercile of ETF ownership but only 19% of stock returns are explained by the industry returns for stocks without ETF ownership. Similarly, the slope coefficient, beta, for the top ownership tercile group is 0.89 compared to the beta of 0.53 for the zero ownership group[6]. The summary results for beta based on ETF ownership groups also show comovement is strongly influenced by a stock's ETF ownership status. The beta with industry stocks held by ETFs for the top ownership decile is more than quadruple that for stocks without ETF ownership. On the other hand, comovement with industry stocks not held by ETFs is significantly reduced for stocks with high ETF ownership. Overall the summary statistics provide initial evidence that ETF ownership is related to comovement with the industry and specifically industry members also held by ETFs.

Additionally, Table 2.1 illustrates differences in the primary independent variables. Although the relationship is not monotonic, stocks with the highest ETF ownership tend to be much larger than stocks without ETF ownership. Stocks with the second highest degree of ETF ownership hold stocks that are on average $1.82 billion larger than those stocks with the most ETF ownership. Furthermore, ETFs tend to hold stocks with higher price-to-book ratios. The

[6] It is not surprising to have an average beta well below one. First, we use the four-digit SIC industry in calculating the beta. While several industries have well over 100 members, the average industry has 13.5 firms. As such, a firm's returns should not be as sensitive to its industry returns as much if a single, firm-specific event significantly impacts the industry return. Previous studies using value-weighting also observe average betas below one (e.g. Pirinsky and Wang, 2004; Wahal and Yavuz, 2013).

top ownership tercile of stocks has ETF ownership of 1.49% of shares outstanding and is overweighted by ETFs by 2.57 times the weight based on market capitalization, on average.

Following the methodology of Fama and MacBeth (1973) and using Newey and West (1987) autocorrelation and heteroskedasticity consistent standard errors, we formally test the relationship between ETF ownership and comovement. The results in Table 2.2 show comovement to be increasing in the fraction of shares outstanding held by ETFs. For example, in the second regression the coefficient on the fraction of shares owned by sector ETFs is 7.873 which is both statistically and economically significant. A one standard deviation increase in fractional ownership corresponds to a 6.53% increase in *R*-squared[7]. In other words, industry returns explain an extra 6.53% of stock return movements, on average, when there is a one standard deviation increase in ETF ownership. Results are similar using either measure of comovement as well as the inclusion of price-to-book and size controls. This means that more of a stock price's movements can be explained by changes in the returns of other members from the same SIC industry when ETFs control larger portions of the stock.

Table 2.3 compares comovement with industry firms containing ETF ownership versus those without such ownership. As ETF ownership increases, stocks comove more with other industry stocks also owned by ETFs but less with other industry stocks not held by ETFs. In economic terms, a one standard deviation increase in ETF ownership is related to an increase of 0.132 (a decrease of 0.0140) in the beta of industry stocks owned (not owned) by ETFs when controlling for size and price-to-book ratio[8]. Relative to the average betas of 0.346 and 0.351 for ETF owned and non-owned industry stocks, respectively, this is also economically significant.

---

[7] This is calculated by the multiplying the regression coefficient of 7.873 times the standard deviation of *Fraction Owned* of 0.0082963.

[8] These are calculated using the product of the regression coefficients, 15.929 and -1.694, and the standard deviation of *Fraction Owned* of 0.0082963.

We next test to see whether the relative weighting of the stock within ETF portfolios influences comovement. Since sector ETFs invest in numerous stocks from the same industry, we compare the market-capitalization weighting of the stock within the industry amongst ETF portfolios to that of the weighting of the stock within the industry. The greater this measure, *Overweight*, the greater the relative impact of ETF trading upon the stock. For example, say stock ABC (XYZ) has an industry weight of 1% (5%) based upon market cap but a 5% (5%) industry weight within ETF portfolios. Since sector ETFs trade both stocks jointly in a basket along with others in their respective industry, trading will influence both stocks. However, ABC should be influenced relatively more than XYZ given that ETFs trade ABC five times as often as would be expected based upon its value-weight. Consistent with these predictions, Table 2.4 demonstrates that comovement with the industry is increasing in overweighting. This finding is strongly significant using both comovement measures and controlling for size and price-to-book ratio. For example, the average sensitivity of the industry beta to overweighting is strongly significant at 0.275 in the third regression. The average R^2 from this cross-sectional regression is 0.067 suggesting that ETF overweighting explains an average of 6.7% of the daily comovement between a stock and its industry. Table 2.5 further illustrates that ETF overweighting is related to high comovement with ETF owned stocks in the industry and low comovement for industry stocks without ETF ownership. Together these results demonstrate that ETF ownership in a stock and the relative weighting of a stock within an ETF portfolio are both important and are related to a stock comoving with its industry.

For robustness, we also consider comovement using other industry classifications. In unreported results, we see similar findings using two-digit and three-digit SIC as well as Fama-

French 49 industry classification. Additionally, the results hold when using ordinary least squares regressions including fund and time fixed effects.

2.3.2 Stock Additions to New ETFs

In the spirit of a quasi -natural experiment, we analyze the change in comovement around ETF inceptions to further establish the connection between ETF ownership and comovement. We begin by collecting the stock holdings for each ETF at inception[9]. In order to test for the impact of the stock becoming ETF owned, we only consider stocks in ETF portfolios at inception that have not been included in previous ETF portfolios. We further require these stocks to continue to be held by the same ETF one year after inception. In order to clearly separate the time period before and after the ETF becomes part of the portfolio, we exclude the quarter of ETF inception. We consider the change in the comovement using daily returns one year before to one year after, weekly returns one and two years before and after, and monthly returns three years before and after. This produces a sample of 1,530 stocks during the sample time period.

Additionally, we use a difference-in-difference test to see the change in comovement for stocks added to new ETFs compared to stocks that are not held by ETFs. This addresses potential concerns that comovement changes are driven solely by overall market cointegration. We choose the matched sample based upon similar industry, size, and price-to-book ratio (P/B). First, matches are formed on the basis of the same four-digit SIC, market cap decile, and P/B tercile. If no match is found, we broaden the criteria to use the three-digit SIC, market cap decile, and P/B tercile. If these do not produce a match, we further loosen criteria by using a two-digit SIC and broaden the market cap decile up to +/- 3 deciles. When multiple stocks fit the

[9] We use the first holdings data available and assume the holdings are the same as at inception. Only funds with holdings data within two quarters of the inception date are included.

selection criteria, we use the stock with the closest market cap as the chosen stock for the matched sample.

Table 2.6 displays the results for the change in comovement of stocks added to new ETFs. The comovement measures in the first two columns show a strong, positive, and statistically significant increase in comovement for stocks after their addition to an ETF portfolio. For example, using *R*-squared as the measure of comovement estimated using daily returns, industry returns explain, on average, an additional 5.88% of a stock's return movements after being added to an ETF portfolio. This increase in comovement is consistent using daily, weekly, and monthly returns over a one to three year horizon. Additionally, we compare the comovement of stocks added to new ETF portfolios with those of other industry stocks based upon the ETF ownership status of the comparison group. Addition to an ETF portfolio is significantly related to a sharp increase in comovement with other industry stocks held by ETFs but related to a significant decline in comovement with industry stocks not held by ETFs.

Table 2.7 illustrates results from difference-in-difference tests to see whether this change in comovement is attributable to the addition to an ETF portfolio or if it is a trend common to other stocks not held by ETFs. Using both measures of comovement, there is a significant increase in return comovement for stocks added to new ETFs beyond that which is seen for stocks not held by ETFs, especially when using *R*-square as the comovement measure. Industry weekly returns over one year explain an additional 2.9% of a stock's return when added to a new ETF. Furthermore, these ETF-held stocks exhibit a significant increase in comovement with other ETF-held stocks in the industry that exceeds that common to stocks without ETF ownership. This result holds using daily and weekly returns over one and two year periods, although not with monthly returns over a three year period. When using a one year comparison

period, particularly with daily returns, there are some signs that ETF-held stocks decline in the degree of comovement with industry stocks not held by ETFs in excess of that seen for the matched sample.

For robustness, we also consider matched portfolios formed on Fama-French 49 and Fama-French 17 industries instead of SIC industries. The results are qualitatively similar and even stronger using Fama-French 17 than those using the SIC coding. Overall, these results provide further evidence that ETF ownership of stocks ties strongly with changes in return comovement.

2.3.3 Earnings Surprises

2.3.3.1 Methodology for Earnings Surprises

Thus far, we have shown a connection between sector ETF ownership of stocks and comovement. Next, we use earnings surprises to examine the relationship between ETF ownership, stock returns, and fundamentals. Many investors watch quarterly earnings carefully as they provide a measure for how well the company is performing. A number of sell-side analysts follow stocks and forecast quarterly earnings per share (EPS). Failing to meet these forecasts by even a small amount can lead to a significant negative stock price reaction. On the other hand, beating analyst expectations can provide a boost to the stock price. Either way, since earnings announcements give investors an update on fundamentals, markets react to this information. The scenario is different for ETFs, however, since trading activity is not driven by fundamentals but by the movement of an index. So while a manager who actively manages a mutual fund is likely to respond to an earnings surprise with increased trading, a passive ETF trades simply to track the return of the fund's index. Hence, one would expect stocks with high ETF ownership to react less to both positive and negative earnings surprises.

To test this hypothesis, we gather quarterly earnings announcement data from I/B/E/S during 2003 to 2012. Following Hirshleifer, Lim, and Teoh (2009), forecast errors (*FE*) are defined as the difference between the actual earnings, e_{iq}, and the median analyst forecast, F_{iq}, normalized by the end of quarter stock price, P_{iq}, for stock i in quarter q:

$$FE_{iq} = \frac{e_{iq} - F_{iq}}{P_{iq}}. \tag{7}$$

FE is then grouped based upon the direction of the forecast error and ranked on magnitude within those groups labeled *FE Quantile*, as in Dellavigna and Pollet (2009). Negative earnings surprises are ranked into quintiles where quantile 1 (5) includes the worst (smallest negative) surprises. All announcements with no forecast errors comprise quantile 6. Positive forecast errors are ranked in similar fashion to negative surprises where quantile 7 (11) includes the least positive (best) quintile of good surprises. *Top Group* is an indicator variable with value of 1 for surprises in quantile 10 or 11 and value of 0 for surprises in quantile 1 or 2. We use only the most recent consensus forecast reported in I/B/E/S prior to the announcement. To reduce potential data errors, we remove observations where the reported earnings or forecasts exceed the stock price, the stock price is under $1, or data is stale (more than 7 weeks old).

In order to calculate the cumulative abnormal return (CAR) surrounding an earnings announcement, we take the difference between the holding period return for the stock with an earnings announcement and the return for a matched portfolio of similar stocks based on size and book-to-market ratio:

$$\mathrm{CAR}[0,1]_{\mathrm{iq}} = \prod_{\mathrm{k=t}}^{\mathrm{t+1}}(1 + \mathrm{R_{ik}}) - \prod_{\mathrm{k=t}}^{\mathrm{t+1}}(1 + \mathrm{R_{pk}}) \tag{8}$$

where R_{ik} and R_{pk} are the returns for firm *i* and portfolio *p* on day *k* for the earnings in quarter *q* and where the date of the earnings announcement is day zero[10]. The matched portfolios are one of 25 value-weighted portfolios based on market capitalization and the book-to-market ratio. To form the matched portfolio, stocks are ranked into quintiles each June based upon market capitalization and then within each of these quintiles stocks are further ranked into quintiles based upon price-to-book ratio using the closing price of the previous year and prior year's book value of equity. Daily returns for these portfolios are available on Kenneth French's website[11].

Previous literature documents a post-earnings announcement drift where stocks tend to exhibit an upward drift following positive surprises and downward drift following negative surprises (e.g. Bernard and Thomas, 1989; Chan, Jegadeesh, and Lakonishok, 1996). To address the potential relation between ETF ownership and the drift, we calculate the drift using Equation 8 over the ensuing 20, 40, and 60 days.

We use the following OLS regression to test for the relation between ETF ownership and abnormal returns:

$$CAR = a_0 + a_1 FE\ Quantile + a_2 Fraction\ Owned + a_3(FE\ Quantile * Fraction\ Owned) + \sum_{i=1}^{n} b_i X_i + \sum_{i=1}^{n} c_i(FE\ Quantile * X_i) + \varepsilon. \quad (9)$$

where X_i are control variables. The dependent variable, CAR, covers days 0 to 1 for the immediate response and days 2 to 21, 2 to 41, and 2 to 61 for the post-announcement drift. Additional regressions use the subsample of the best two quantiles and worst two quantiles and *Top Group* replaces *FE Quantile* in Equation 9. There are a number of control variables included, similar to Hirshleifer, Lim, and Teoh (2009): book-to-market and size deciles, and, since the influence of expectations may be based on the extent of analyst coverage, the natural

[10] Results are similar using the period from one day prior to one day after the earnings announcement date.
[11] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

log of one plus the number of earnings estimates. *Reporting Lag* is the number of weeks from the quarter-end until the earnings announcement date. *Earnings Volatility* measures the variability in EPS from one year to the next and is calculated as the standard deviation of the difference between quarterly earnings and earnings one year prior; this is calculated for up to the past four years with a minimum of four observations and uses split-adjusted earnings. *Share Turnover* takes the one year average of monthly trading volume divided by the average shares outstanding. *Earnings Persistence* is the slope coefficient from the regression of current earnings on earnings from the prior period (minimum 4 observations over a 4 year window, earnings are split adjusted). Finally, indicator variables for year, month, and day of the week are included. Standard errors are robust to heteroskedasticity and clustered by day of the announcement.

Additionally, we analyze whether ETFs impact the trading volume around earnings announcements. Since ETFs do not trade in large quantities in reaction to an earnings announcement, it is possible that they may also lead to less abnormal trading. Abnormal trading volume is defined as in Hirshleifer, Lim, and Teoh (2009) as the average log dollar trading volume around the earnings announcement minus the average log dollar trading volume over days [-41,-11] relative to the earnings announcement:

$$\text{Ab. Vol. }[0,1] = \frac{1}{2}\sum_{j=t}^{t+1} \text{Log}\left(\text{DollarVol}_j + 1\right) - \frac{1}{30}\sum_{k=t-41}^{t-11} \text{Log}(\text{DollarVol}_k + 1). \qquad (10)$$

Regression estimates are as in Equation 9 with *Ab. Vol.[0,1]* as the dependent variable with the addition of market abnormal trading volume, calculated as the value-weighted average abnormal volume for all stocks in CRSP where abnormal trading volume of each stock is calculated as in (10).

2.3.3.2 Empirical Results for Earnings Surprises

Table 2.8 presents the descriptive statistics for the cumulative abnormal returns and abnormal trading volume surrounding quarterly earnings announcements. We present statistics for the full sample of negative and positive earnings surprises in Panels A and B, respectively. Panels C and D include only the biggest surprises, those in the bottom two quintiles of negative surprises and top two quintiles of positive surprises, respectively. To show the impact of ETFs, we further subdivide into a group with zero ownership by sector ETFs and stocks held by sector ETFs ranked into terciles based upon the fraction of shares outstanding held by ETFs.

The results for the full sample show signs of a dampened return reaction to positive earnings surprises in Panel A, although the same does not hold for negative surprises in Panel B. Both positive and negative surprises show a dampened abnormal volume reaction for stocks with the highest ETF ownership. The impact of ETF ownership is seen more clearly in Panels C and D when focusing on the best and worst surprises. For negative surprises, the initial abnormal return reaction is dampened but the post-announcement drift is greater. The CAR for the top ETF ownership tercile has an abnormal announcement return of 9 basis points less negative, on average, compared to stocks without ETF ownership. For positive surprises, the top tercile group has a CAR that is 57 basis points less than stocks not held by ETFs, on average. For negative and, especially, for positive surprises, stocks with the highest ETF ownership have a much lower average abnormal volume response to earnings announcements.

Table 2.9 exhibits the results from multivariate regressions to formally test for the relation between ETF ownership and the reaction to earnings surprises. The results provide strong evidence that the reaction to earnings surprises is significantly dampened for stocks with high ETF ownership. The coefficient on *FE Quantile* in the first regression indicates a one-step

increase in the forecast error quantile is associated with an increase of 50 basis points in the CAR in the initial two days. However, the interaction term between the FE quantile and ETF ownership shows a significant dampening effect upon the initial CAR as ETF ownership increases. A one standard deviation increase in ETF ownership is associated with a dampening effect of 6 basis points[12]. Similarly, the second regression shows a significant muting effect on the abnormal volume associated with ETF ownership. The last three regressions focus on the post-announcement drift over different horizons. If stocks with high ETF ownership initially underreact to information, then the market may subsequently react more in the ensuing weeks as it incorporates the information. Consistent with this conjecture, we show a larger post-announcement drift associated with ETF ownership, particularly in the 20 and 40 day windows. For example, in the 20 day window, a one-step increase in the FE quantile is related to a 20 basis point increase in the CAR while a one standard deviation increase in ETF ownership is related to a 4 basis point (=0.045*0.009369) larger CAR during that period.

In Table 2.10, we examine the largest good and bad surprises. We reduce the sample to include only the lowest two quintiles of negative surprises (which have a *Top Group* value of 0) and the best two quintiles of positive surprises (which have a *Top Group* value of 1). An advantage of using *Top Group* instead of *FE Quintile* is that it is easier to distinguish the impact of ETF ownership for positive versus negative surprises. For the initial CAR, the best surprises have a CAR that is 6.2% higher than those for the worst surprises, on average. The interaction term between *Top Group* and *Fraction Owned* indicates a significant dampening effect of ETF ownership for the best surprises. A one standard deviation increase in ETF ownership is

[12] This is calculated by multiplying the coefficient on the interaction term, -0.038, times the standard deviation of *Fraction Owned* for the sample used in Table 2.7 of 0.009369. We follow a similar procedure of taking the product of the regression coefficient on the interaction term and 0.009369 to calculate the impact of a one standard deviation increase in ETF ownership elsewhere in Table 2.7.

associated with a 57 basis point lower CAR[13]. On the other hand, for negative surprises, the coefficient on *Fraction Owned* indicates a one standard deviation increase in ETF ownership is related to a 22 basis point (=0.218*0.103211) increase in the CAR. Both of these show a significant dampening effect of ETF ownership, especially for positive surprises. In addition, there is a muted abnormal volume reaction for both good and bad surprises associated with ETF ownership that is significant at the one-percent level. Finally, we find that most of the initial under reaction is reversed within the next twenty days for positive surprises and is completely reversed for negative surprises. A one standard deviation increase in ETF ownership is related to a 49 basis point (=0.472*0.103211) greater positive drift in the ensuing 20 days. This captures around 86% (=0.472/0.552) of the initial dampened reaction. Overall, the results provide strong evidence to suggest that ETF ownership is related to a dampened initial reaction to earnings announcements that is eventually corrected in the ensuing weeks.

2.3.4 Other Institutional Owners

Thus far we have only considered the impact of sector ETFs. The rationale for this is that sector ETFs focus all their trades on a specific sector of the economy, while most other institutional holders hold more diversified portfolios. The joint trading of a narrow group of stocks by sector ETFs leads to greater influence of the group upon the individual stocks. Furthermore, the nature of ETFs tracking an index instead of having an actively managed portfolio, like most other institutional investors, leads to trading of stocks for reasons other than fundamentals. Unlike sector ETFs, an actively managed mutual fund will try to time trades using fundamental analysis and will generally invest in broad portfolios of stocks in different sectors.

[13] This is calculated by multiplying the coefficient on the interaction term, -0.552, times the standard deviation of *Fraction Owned* for the sample used in Table 2.8 of 0.0103211. A similar procedure is used elsewhere in Table 2.8.

Nevertheless, in Table 2.11 we consider the impact of other institutional owners. Using data from CRSP Mutual Fund, Thomson Reuters Institutional Holdings and Thomson Reuters Mutual Fund Holdings databases, we calculate the fraction of stock shares outstanding held by non-sector ETFs, mutual funds, and other institutional investors and add these to the regression specified in Equation 6. We present results for the full sample of stocks as well as the subsample which are owned by sector ETFs. Even after controlling for other institutional investors, sector ETFs continue to demonstrate a strong positive association with comovement. Although non-sector ETF ownership is also positively related to comovement and mutual funds show some weak association with comovement, the relative impact is larger for sector ETFs in each regression. In the full sample and using *R*-squared as the measure of comovement, a one standard deviation increase in sector ETF and non-sector ETF ownership is associated with a 6.41% and 3.7% increase in comovement, respectively[14]. The impact of sector ETF ownership is even more pronounced when looking at the subsample of sector ETF-held stocks. Sector ETFs are the only institutional owners with a statistically significant and positive relation with comovement using beta. When using *R*-square as the comovement measure, a one standard deviation increase in both sector and non-sector ETF ownership is related to a 7.83% and 2.92% increase in comovement, respectively[15]. Overall, the results indicate that while sector ETFs may not be the only source of comovement with a firm's industry, they appear to be one of the most important sources amongst institutional holders.

We provide further analysis of the roles of each type of institutional holding using earnings surprises in Table 2.12. We re-estimate regressions from Table 2.9 with the addition of

[14] These are calculated by multiplying the regression coefficients of 7.796 and 1.782 times the standard deviation of *Sector ETF Own* and *Non-Sector ETF Own* of 0.008223 and 0.0207478, respectively.
[15] These are calculated by multiplying the regression coefficients of 6.782 and 1.242 times the standard deviation of *Sector ETF Own* and *Non-Sector ETF Own* of 0.0115415 and 0.0235079, respectively, from the subsample of stock owned by sector ETFs.

the non-sector ETF, mutual fund, and other institutional ownership variables. Even with the inclusion of the other institutional variables, sector ETF ownership continues to show a strong dampening relationship with respect to both the cumulative abnormal return and abnormal trading volume around an earnings surprise. In the announcement period abnormal return regressions, sector ETFs are the only type of institutional fund associated with a significant dampening effect. The reversal effect on the post-announcement drift also holds for sector ETFs. Sector ETFs are not alone in exhibiting a dampening effect upon abnormal trading volume; however, the impact of sector ETF ownership is the largest amongst the types of institutional owners. Table 2.13 shows similar findings when using the subsample of the best and worst earnings surprises. Together these results provide further evidence that sector ETFs are linked to a dampened impact of fundamentals on stock price and volume around earnings announcements.

2.4 Conclusions

The impact of exchange-traded funds upon the marketplace has grown rapidly in recent years. With a large following by both institutional and retail investors, ETFs fill an important role in the financial markets. An increasing number of market participants and those in the press have questioned the potential negative impact that ETFs have on stocks they hold. Critics such as Harold Bradley (previously referenced in testimony before a U.S. Senate committee) suggest that the influence of ETFs is such that small cap stocks are "proverbial tiny boats being tossed around on the ETF ocean." Others suggest that ETFs do not negatively impact stocks.

In this paper, we shed further light upon this important question. We find that sector ETF ownership is strongly related to comovement of stocks with their industry. In particular, the comovement with stocks also held by ETFs is increasing in ETF ownership. However, ETF ownership is negatively related to comovement with stocks not held by ETFs. Higher

comovement also is observed when ETFs overweight a stock. In the spirit of a quasi-natural experiment, we apply difference-in-difference tests and find that the change in comovement for a stock added to a new ETF portfolio is significantly greater than that for a matched sample of stocks not owned by ETFs. Additionally, we show that higher sector ETF ownership is related to stocks having a dampened initial reaction to earnings surprises. Both the cumulative abnormal return and the abnormal trading volume are muted in the presence of higher ETF ownership.

Finally we show that other types of institutional holdings do not explain our results. Sector ETFs continue to show a strong positive relationship with comovement after controlling for non-sector ETFs, mutual funds, and other forms of institutional ownership. While we do not exclude the possibility that other institutional holdings influence comovement, the relation appears most strongly related to sector ETFs. Furthermore, only sector ETFs show a statistically significant relation to dampened CAR and abnormal trading volume immediately following the earning announcement. Overall, we find that industry returns explain an increasing amount of stock returns and changes in firm fundamentals have a decreased immediate influence upon stock returns, on average, as sector ETF ownership increases.

REFERENCES

Agyei-Ampomah, S., & Mazouz, K. (2011). The comovement of option listed stocks. *Journal of Banking & Finance*, 35(8), 2056-2069.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3-18.

Barberis, N., & Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, *68*(2), 161-199.

Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, *75*(2), 283-317.

Bartov, E., Radhakrishnan, S., & Krinsky, I. (2000). Investor sophistication and patterns in stock returns after earnings announcements. *Accounting Review*, *75*(1), 43-63.

Bernard, V. L., & Thomas, J. K. (1989). Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research*, *27*, 1-36.

Boyer, B. H. (2011). Style-related comovement: Fundamentals or labels? *Journal of Finance*, *66*(1), 307-332.

Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *Journal of Finance*, *51*(5), 1681-1713.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 607-636.

Francis, J., Pagach, D., & Stephan, J. (1992). The stock market response to earnings announcements released during trading versus nontrading periods. *Journal of Accounting Research*, *30*(2), 165-184.

Green, T. C., & Hwang, B. H. (2009). Price-based return comovement. *Journal of Financial Economics*, *93*(1), 37-50.

Greenwood, R. (2008). Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. *Review of Financial Studies*, *21*(3), 1153-1186.

Grullon, G., J. Weston, & S. Underwood (2014). Comovement and investment banking networks. *Journal of Financial Economics*, forthcoming.

Hou, K., Xiong, W., & Peng, L. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum. Working paper, The Ohio State University.

Newey, W. K., & West, K. D. (1987). A Simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703-708.

Potter, G. (1992). Accounting earnings announcements, institutional investor concentration, and common stock returns. *Journal of Accounting Research*, 30(1), 146-155.

Pirinsky, C., & Wang, Q. (2004). Institutional investors and the comovement of equity prices. Working Paper, George Washington University.

Pirinsky, C., & Wang, Q. (2006). Does corporate headquarters location matter for stock returns? *Journal of Finance*, *61*(4), 1991-2015.

Wahal, S., & Yavuz, M. D. (2013). Style investing, comovement and return predictability. *Journal of Financial Economics*, *107*(1), 136-154.

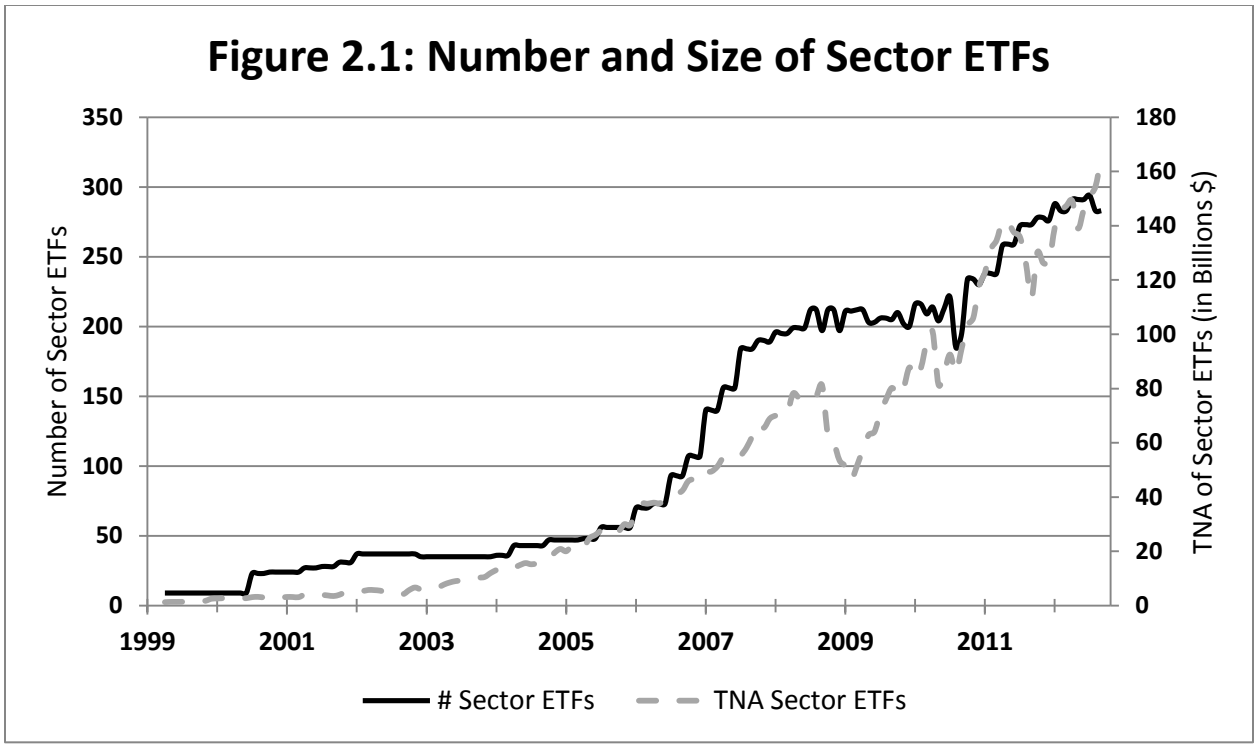
Figure 2.1: Number and Size of Sector ETFs
Number of Sector ETFs
TNA of Sector ETFs (in Billions $)
350
300
250
200
150
100
50
0
180
160
140
120
100
80
60
40
20
0
1999
2001
2003
2005
2007
2009
2011
Sector ETFs
TNA Sector ETFs

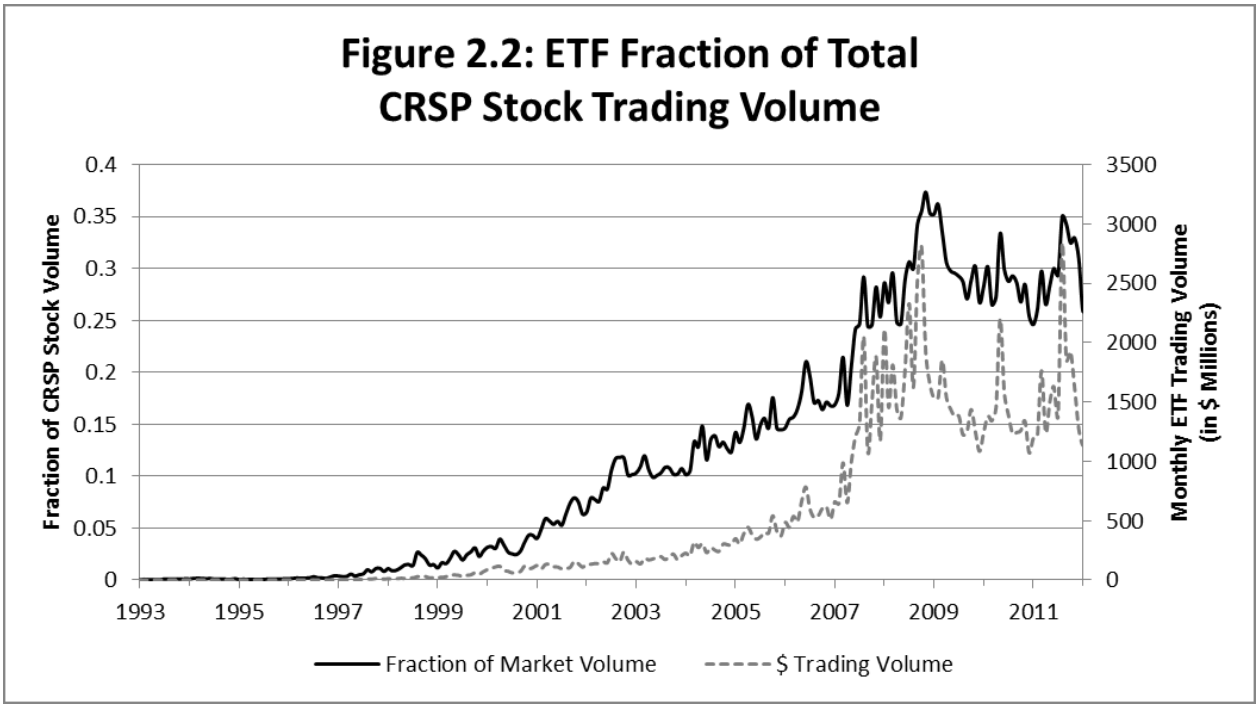
Figure 2.2: ETF Fraction of Total CRSP Stock Trading Volume
Fraction of CRSP Stock Volume
Monthly ETF Trading Volume (in $ Millions)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
3500
3000
2500
2000
1500
1000
500
0
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
Fraction of Market Volume
$ Trading Volume

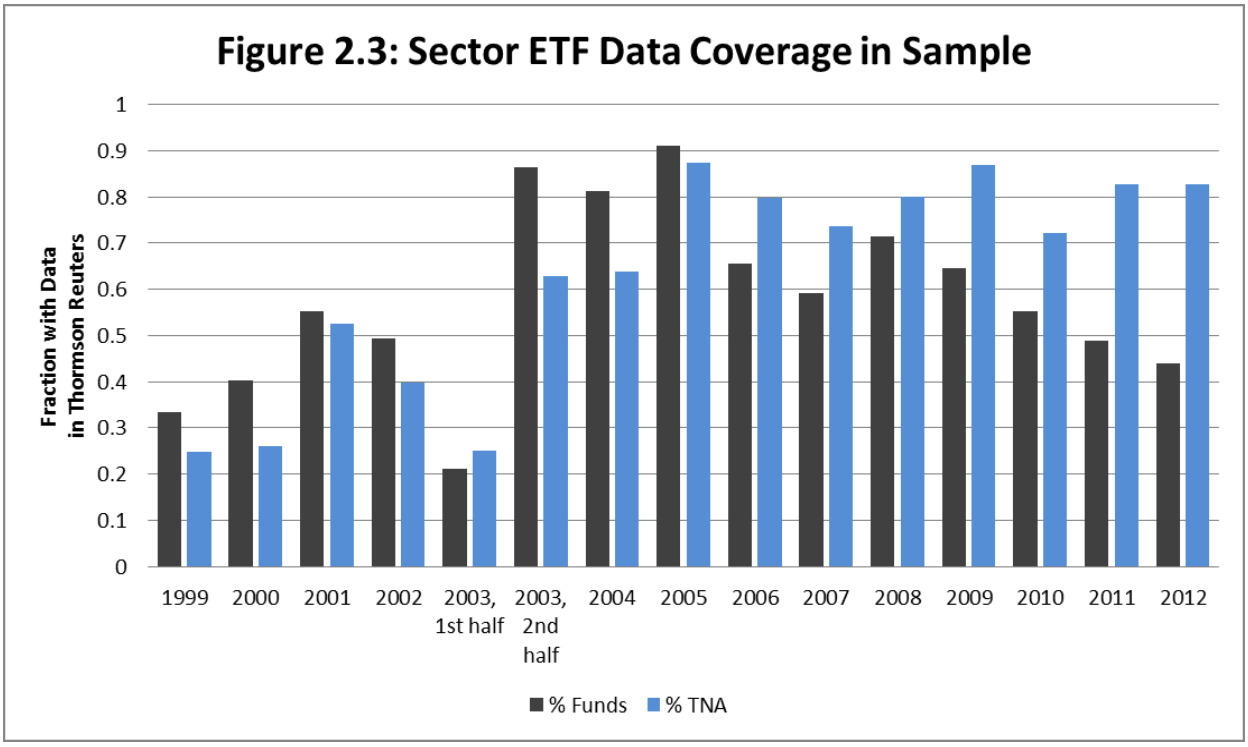
Figure 2.3: Sector ETF Data Coverage in Sample
Fraction with Data in Thormson Reuters
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1999
2000
2001
2002
2003, 1st half
2003, 2nd half
2004
2005
2006
2007
2008
2009
2010
2011
2012
% Funds
% TNA

Table 2.1
Summary Statistics by ETF Ownership Group

This table presents the univariate anaylsis for measures of comovement. The first two comovement measures are the R-squared and Beta values from a regression of stock returns for firm i on their value-weighted 4 digit SIC industry return: $R_{it}=\alpha_i + \beta_i * R_{SIC,t} + \varepsilon_{it}$. The final two comovement measures are the slope coefficients from a regression of firm i's returns on the value-weighted returns of firms in the same SIC industry with sector ETF ownership as well as the group without sector ETF ownership: $R_{it}=\alpha_i + \beta_{1,it} * R_{SIC\ ETF\ Owned,t} + \beta_{2,it} * R_{SIC\ ETf\ Not\ Owned,t} + \varepsilon_{it}$. *ETF weight ÷ Value weight* is the weight of the stock amongst sector ETF portfolios within the SIC industry divided by the CRSP value weight of the stock in the SIC industry. Mean and median statistics are shown for all stocks without any sector ETF ownership as well as the tercile ranking for stocks with sector ETF ownership. *** Denotes the difference between the coefficient of the ownership decile is statistically different than the zero ownership group at the 1% level.

		Zero Ownership	Ownership Tercile 1	Ownership Tercile 2	Ownership Tercile 3
R-squared	Mean	0.19	0.26***	0.31***	0.38***
	Median	0.06	0.22***	0.28***	0.36***
Beta	Mean	0.53	0.75***	0.74***	0.89***
	Median	0.43	0.72***	0.73***	0.89***
Beta of Industry Stocks Owned by ETFs	Mean	0.18	0.57***	0.6***	0.73***
	Median	0.09	0.54***	0.58***	0.73***
Beta of Industry Stocks Not Owned by ETFs	Mean	0.41	0.29***	0.22***	0.22***
	Median	0.28	0.2***	0.15***	0.15***
Market Cap (in billions $)	Mean	0.77	2.36***	8.88***	7.06***
	Median	0.12	0.72***	1.88***	1.95***
Price-to-Book	Mean	2.85	2.95**	3.39***	3.39***
	Median	1.56	1.92***	2.24***	2.19***
Percent ETF Ownership	Mean	0.000	0.062***	0.221***	1.485***
	Median	0.000	0.018***	0.148***	0.851***
ETF weight ÷ Value weight	Mean	0.00	1.39***	1.84***	2.57***
	Median	0.00	1.18***	1.9***	2.31***
# Observations		163,140	30,448	30,448	30,448

Table 2.2
Comovement and Sector ETF Ownership

This table presents the cross-sectional regression results following Fama and MacBeth (1973) with corrections for autocorrelation and heteroskedasticity using Newey and West (1987) standard errors. The dependent variable is one of two comovement measures, as listed at the heading of each column. The comovement measures are the *R*-squared and Beta values from a regression of daily stock returns over one quarter for firm i on their value-weighted 4-digit SIC industry return: $R_{it} = \alpha + \beta_i * R_{SIC,t} + \varepsilon_{it}$. *Fraction Owned* is defined as the natual log of one plus the fraction of stock shares outstanding held by sector ETFs. *Price-to-Book* and *Market Cap* are the natural log of the price-to-book ratio and market capitalization, respectively.

	R-Squared		Beta	
Fraction Owned	13.199***	7.873***	25.053***	16.321***
	(3.39)	(3.17)	(2.88)	(2.80)
Price-to-Book		-0.016***		-0.006
		(-6.45)		(-0.89)
Market Cap		0.061***		0.083***
		(17.50)		(19.50)
Intercept	0.192***	-0.579***	0.580***	-0.483***
	(10.87)	(-20.91)	(23.97)	(-6.58)
No. Obs	173164	173164	173164	173164
R^2	0.087	0.391	0.037	0.136

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.3
Comovement and Sector ETF Ownership, by Groups

This table presents the cross-sectional regression results following Fama and MacBeth (1973) with corrections for autocorrelation and heteroskedasticity using Newey and West (1987) standard errors. The dependent variable is one of two comovement measures, as listed at the heading of each column. The comovement measures are the slope coefficients from a regression of firm *i*'s daily returns over one quarter on the value-weighted returns of firms in the same 4-digit SIC industry with sector ETF ownership as well as the group without sector ETF ownership: $R_{it}=\alpha_i + \beta_{1,it}*R_{SIC\ ETF\ Owned,t} + \beta_{2,it}*R_{SIC\ ETF\ Not\ Owned,t} + \varepsilon_{it}$. *Fraction Owned* is defined as the natual log of one plus the fraction of stock shares outstanding held by sector ETFs. *Price-to-Book* and *Market Cap* are the natural log of the price-to-book ratio and market capitalization, respectively.

	Beta of ETF Ownership Group		Beta of Non-ETF Ownership Group	
Fraction Owned	24.210***	15.929***	-0.926	-1.694**
	(3.32)	(3.25)	(-1.02)	(-2.05)
Price-to-Book		-0.004		0.002
		(-1.38)		(0.53)
Market Cap		0.083***		0.001
		(41.26)		(0.35)
Intercept	0.400***	-0.663***	0.265***	0.245***
	(13.05)	(-13.95)	(22.66)	(5.52)
No. Obs	152705	152705	152705	152705
R^2	0.041	0.132	0.002	0.005

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.4
Comovement and Sector ETF Overweighting

This table presents the cross-sectional regression results following Fama and MacBeth (1973) with corrections for autocorrelation and heteroskedasticity using Newey and West (1987) standard errors. The dependent variable is one of two comovement measures, as listed at the heading of each column. The comovement measures are the *R*-squared and Beta values from a regression of daily stock returns over one quarter for firm *i* on their value-weighted 4-digit SIC industry return: $R_{it} = \alpha + \beta_i * R_{SIC,t} + \varepsilon_{it}$. *Overweight* is defined as the natural log of one plus the ratio of the weighting of the stock within the industry amongst ETF portfolios to the firm's CRSP industry value-weight. *Price-to-Book* and *Market Cap* are the natural log of the price-to-book ratio and market capitalization, respectively.

	R-Squared		Beta	
Overweight	0.161***	0.061***	0.275***	0.125***
	(17.60)	(11.69)	(36.70)	(13.23)
Price-to-Book		-0.016***		-0.007
		(-6.20)		(-0.99)
Market Cap		0.058***		0.077***
		(16.36)		(16.91)
Intercept	0.165***	-0.549***	0.531***	-0.411***
	(10.91)	(-18.96)	(111.79)	(-5.42)
No. Obs	172973	172973	172973	172973
R^2	0.147	0.381	0.067	0.134

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.5
Comovement and Sector ETF Overweighting, by Groups

This table presents the cross-sectional regression results following Fama and MacBeth (1973) with corrections for autocorrelation and heteroskedasticity using Newey and West (1987) standard errors. The dependent variable is one of two comovement measures, as listed at the heading of each column. The comovement measures are the slope coefficients from a regression of firm *i*'s daily returns over one quarter on the value-weighted returns of firms in the same 4-digit SIC industry with sector ETF ownership as well as the group without sector ETF ownership: $R_{it}=\alpha_i + \beta_{1,it}*R_{SIC\ ETF\ Owned,t} + \beta_{2,it}*R_{SIC\ ETf\ Not\ Owned,t} + \varepsilon_{it}$. *Overweight* is defined as the natural log of one plus the ratio of the weighting of the stock within the industry amongst ETF portfolios to the firm's CRSP industry value-weight. *Price-to-Book* and *Market Cap* are the natural log of the price-to-book ratio and market capitalization, respectively.

	Beta of ETF Ownership Group		Beta of Non-ETF Ownership Group	
Overweight	0.282***	0.149***	-0.027***	-0.034***
	(29.17)	(18.77)	(-4.23)	(-8.03)
Price-to-Book		-0.006		0.003
		(-1.66)		(0.59)
Market Cap		0.075***		0.004
		(39.49)		(0.86)
Intercept	0.353***	-0.572***	0.269***	0.219***
	(12.53)	(-12.57)	(23.70)	(4.44)
No. Obs	152533	152533	152533	152533
R^2	0.069	0.131	0.002	0.006

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.6
Change in Comovement for Stocks Added to New ETFs

This table presents the change in the measures of comovement for stocks included in a new ETF and which have never been included in an ETF before. Stocks are required to still be in the same ETF portfolio one year after the fund's inception. The frequency of returns and the time period used to esimate the measures of comovement before and after the stock' addition to an ETF is indicated in the panel header. The measure of comovement is indicated at the top of each column. The comovement measures R-squared and Beta come from a regression of stock returns for firm i on their value-weighted 4-digit SIC industry return: $R_{it}=\alpha_i + \beta_t * R_{SIC,t} + \varepsilon_{it}$. Additional comovement measures are the slope coefficients from a regression of firm i's daily returns on the value-weighted returns of firms in the same 4-digit SIC industry with sector ETF ownership as well as the group without sector ETF ownership: $R_{it}=\alpha_i + \beta_{1,it} * R_{SIC\ ETF\ Owned,t} + \beta_{2,it} * R_{SIC\ ETF\ Not\ Owned,t} + \varepsilon_{it}$.

	Univariate		Bivariate	
	Δ *R*-Squared	Δ Beta	Δ Beta of ETF Owned Group	Δ Beta of Non-ETF Owned Group
Panel A: Daily Returns (1 Year Period)				
Difference	0.0588	0.0849	0.2041	-0.1075
p-value	0.00	0.00	0.00	0.00
Panel B: Weekly Returns (1 Year Period)				
Difference	0.0300	0.0475	0.1775	-0.1005
p-value	0.00	0.00	0.00	0.00
Panel C: Weekly Returns (2 Year Period)				
Difference	0.0165	0.0698	0.2249	-0.0883
p-value	0.00	0.00	0.00	0.00
Panel D: Monthly Returns (3 Year Period)				
Difference	0.0261	0.0449	0.0747	-0.0149
p-value	0.00	0.04	0.02	0.74

Table 2.7
Difference-in-Difference Tests for Change in Comovement for Stocks Added to New ETFs

This table presents the difference in differences between the change in comovement for stocks added to a new ETF compared to a matched sample. To be considered, the event must be the first time the stock has been in any ETF portfolio and it must still be in the ETF portfolio one year after the fund's inception. The frequency of returns and the time period used to esimate the measures of comovement before and after the stock' addition to an ETF is indicated in the panel header. The measure of comovement is indicated at the top of each column. The comovement measures R-squared and Beta come from a regression of stock returns for firm i on their value-weighted 4-digit SIC industry return: $R_{it}=\alpha_i + \beta_t * R_{SIC,t} + \varepsilon_{it}$. Additional comovement measures are the slope coefficients from a regression of firm i's daily returns on the value-weighted returns of firms in the same 4-digit SIC industry with sector ETF ownership as well as the group without sector ETF ownership: $R_{it}=\alpha_i + \beta_{1,it} * R_{SIC\ ETF\ Owned,t} + \beta_{2,it} * R_{SIC\ ETF\ Not\ Owned,t} + \varepsilon_{it}$. The matched sample is formed based upon SIC industry, size, and price-to-book measures.

	Univariate		Bivariate	
	ΔΔ *R*-Squared	ΔΔ Beta	ΔΔ Beta of ETF Owned Group	ΔΔ Beta of Non-ETF Owned Group
Panel A: Daily Returns (1 Year Period)				
Difference	0.0264	0.0075	0.0552	-0.0505
p-value	0.00	0.47	0.00	0.00
Panel B: Weekly Returns (1 Year Period)				
Difference	0.0290	0.0405	0.0710	-0.0267
p-value	0.00	0.02	0.01	0.32
Panel C: Weekly Returns (2 Year Period)				
Difference	0.0231	0.0478	0.1102	0.0292
p-value	0.00	0.00	0.00	0.41
Panel D: Monthly Returns (3 Year Period)				
Difference	0.0253	0.0114	-0.0264	0.0469
p-value	0.01	0.71	0.54	0.43

Table 2.8
Summary Statistics of Response to Earnings Surprises

This table reports the average cumulative abnormal returns over the 2-day announcement period, the 20, 40, and 60-day post-announcement period, as well as the abnormal volume over the announcement period. Statistics are reported for the full sample of negative earnings surprises in Panel A and for positive earnings surprises in Panel B. Panel C (Panel D) reports statistics for the worst (best) two quintiles of negative (positive) surprises. Earnings surprises are measured by the forecast error, the difference betwen the actual earnings and median analyst forecast normalized by the end of quarter stock price. Abnormal returns are calculated as the difference in returns for the announcing firm and a comparison group matched on size and book-to-market ratio. Abnormal volume is calculated as the difference in the average log dollar trading volume over the announcement period and the pre-announcement period [-41,-11]. Mean statistics are shown for all stocks without any sector ETF ownership as well as the tercile ranking for stocks with sector ETF ownership.

Panel A: Negative Surprises

Group	CAR[0,1]	Ab. Vol [0,1]	CAR[2,21]	CAR[2,41]	CAR[2,61]
Zero Own	-0.0265	0.5166	0.0009	0.0062	0.0111
Tercile 1	-0.0337	0.6120	-0.0001	0.0014	-0.0022
Tercile 2	-0.0332	0.6628	0.0026	0.0010	0.0030
Tercile 3	-0.0270	0.4962	-0.0024	0.0004	0.0022

Panel B: Positive Surprises

Group	CAR[0,1]	Ab. Vol [0,1]	CAR[2,21]	CAR[2,41]	CAR[2,61]
Zero Own	0.0210	0.6381	0.0105	0.0181	0.0263
Tercile 1	0.0250	0.7317	0.0004	0.0001	0.0013
Tercile 2	0.0187	0.7160	0.0010	0.0018	0.0019
Tercile 3	0.0160	0.5898	0.0022	0.0033	0.0034

Panel C: Largest Two Groups of Negative Surprises

Ownership Group	CAR[0,1]	Ab. Vol [0,1]	CAR[2,21]	CAR[2,41]	CAR[2,61]
Zero Own	-0.0362	0.5569	-0.0014	0.0023	0.0076
Tercile 1	-0.0506	0.6285	-0.0075	-0.0077	-0.0093
Tercile 2	-0.0531	0.6754	0.0029	-0.0062	-0.0019
Tercile 3	-0.0353	0.4747	-0.0151	-0.0109	-0.0073

Panel D: Largest Two Groups of Positive Surprises

Ownership Group	CAR[0,1]	Ab. Vol [0,1]	CAR[2,21]	CAR[2,41]	CAR[2,61]
Zero Own	0.0278	0.7071	0.0141	0.0209	0.0323
Tercile 1	0.0383	0.7869	-0.0015	-0.0052	-0.0019
Tercile 2	0.0296	0.7668	-0.0025	-0.0050	-0.0034
Tercile 3	0.0221	0.5679	0.0011	-0.0033	-0.0036

Table 2.9
Abnormal Response Around Earnings Surprises

This table presents the results for the reaction to quarterly earnings announcements. The cumulative abnormal return (CAR) is shown over days 0 to 1, 2 to 21, 2 to 41, and 2 to 61 where day 0 is the earnings announcement day. Abnormal trading volume is presented over days 0 to 1. Abnormal returns are calculated as the difference in returns for the announcing firm and a comparison group matched on size and book-to-market ratio. Abnormal volume is calculated as the difference in the average log dollar trading volume over the announcement period and the pre-announcement period [-41,-11]. *FE Quantile* is the ranking of the forecast error from lowest to highest. Negative surprises are ranked into quintiles and given values from 1 to 5; zero surprises are given a value of 6; positive surprises are ranked into quintiles and given values from 6 to 11. *Fraction Owned* is the natural log of one plus the fraction of shares outstanding held by sector ETFs. Controls include book-to-market and size deciles, log(1+ # Analysts), *Reporting Lag* (in weeks), *Earnings Volatility* defined as the standard deviation of the difference in quarterly earnings in quarter *q* and *q-4, Share Turnover* calculated as the one year average of monthly trading volume divided by the average shares outstanding. *Earnings Persistence* defined as the slope coefficient from the regression of current earnings on earnings from the prior period, and indicator variables for year, month, and day of the week. *Market Abnormal Volume* is also included as a control variable for the abnormal volume regression and is calculated using all stocks in CRSP. Control variables are interacted with *FE Quantile* in the CAR regressions. Standard errors are robust to heteroskedasticity and clustered by announcement date.

	CAR [0,1]	Ab. Vol. [0,1]	CAR [2,21]	CAR [2,41]	CAR [2,61]
FE Quantile	0.005***	0.012**	0.002***	0.002**	0.004***
	(9.25)	(2.56)	(3.43)	(2.03)	(2.61)
Fraction Owned	0.400***	-4.051***	-0.342***	-0.442**	-0.257
	(4.66)	(-8.20)	(-2.82)	(-2.47)	(-1.09)
FE Quantile *Fraction Owned	-0.064***	-0.256***	0.045***	0.046*	0.039
	(-5.31)	(-3.61)	(2.65)	(1.95)	(1.29)
Intercept	-0.038***	0.486***	-0.049***	-0.070***	-0.085***
	(-8.23)	(11.17)	(-8.50)	(-7.33)	(-6.60)
& Interactions w/ FE Quantile	Yes	Yes	Yes	Yes	Yes
No. Obs	75422	75422	75422	75422	75422
R^2	0.113	0.075	0.016	0.024	0.019

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.10
Abnormal Response Around Earnings Surprises: Top and Bottom Groups

This table presents the results for the reaction to quarterly earnings announcements. The dependent variables are the announcement period cumulative abnormal return (CAR) , announcement period abnormal volume, and post-announcement CAR drift. *Top Group* takes a value of 1 (0) for the highest (lowest) two quintiles of positive (negative) earnings surprises. Additional variables are defined as in Table 2.9. Control variables are interacted with *Top Group* . Standard errors are robust to heteroskedasticity and clustered by announcement date.

	CAR [0,1]	Ab. Vol. [0,1]	CAR [2,21]	CAR [2,41]	CAR [2,61]
Top Group	0.062***	0.100**	0.019**	0.031**	0.048***
	(10.75)	(2.06)	(2.48)	(2.41)	(3.61)
Fraction Owned	0.218**	-4.798***	-0.486***	-0.674***	-0.561**
	(2.41)	(-9.34)	(-3.34)	(-3.49)	(-2.35)
Top Group*Fraction Owned	-0.552***	-2.291***	0.472**	0.489**	0.481
	(-4.37)	(-3.17)	(2.58)	(1.97)	(1.63)
Intercept	-0.034***	0.658***	-0.052***	-0.096***	-0.100***
	(-5.70)	(11.54)	(-6.97)	(-8.59)	(-7.12)
Controls Variables & Interactions w/ Top Group	Yes	Yes	Yes	Yes	Yes
No. Obs	25171	25171	25171	25171	25171
R^2	0.148	0.068	0.026	0.037	0.033

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.11
Comovement and Different Institutional Owners

This table presents the cross-sectional regression results following Fama and MacBeth (1973) with corrections for autocorrelation and heteroskedasticity using Newey and West (1987) standard errors. The dependent variable is one of two comovement measures, as listed at the heading of each column. The comovement measures are the *R*-squared and Beta values from a regression of daily stock returns over one quarter for firm i on their value-weighted 4-digit SIC industry return: $R_{it}=\alpha_i + \beta_t * R_{SIC,t} + \varepsilon_{it}$. *Sect ETF Own, Non-Sect ETF Own, Mutual Fund Own,* and *Other Institutional Own* is the natural log of one plus the fraction of shares outstanding held by sector ETFs, non-sector ETFs, mutual funds, and all other institutionanl owners, respectively. *Price-to-Book* and *Market Cap* are the natural log of the price-to-book ratio and market capitalization, respectively.

	Full Sample		Sample with Sector ETF Ownership	
	R-Squared	Beta	*R*-Squared	Beta
Sector ETF Own	7.796***	16.101**	6.782***	13.252**
	(3.01)	(2.63)	(2.99)	(2.45)
Non-Sector ETF Own	1.782***	5.336***	1.242***	0.728
	(4.49)	(3.37)	(3.83)	(1.41)
Mutual Fund Own	0.025*	-0.017	0.012	-0.026
	(1.83)	(-0.50)	(0.52)	(-0.75)
Other Institutional Own	-0.019	-0.002	-0.060***	-0.137***
	(-1.46)	(-0.06)	(-3.14)	(-2.86)
Price-to-Book	-0.016***	-0.007	-0.037***	-0.025***
	(-6.40)	(-0.98)	(-13.25)	(-3.29)
Market Cap	0.058***	0.075***	0.053***	-0.013***
	(20.35)	(14.04)	(21.48)	(-3.54)
Intercept	-0.555***	-0.415***	-0.420***	1.034***
	(-23.00)	(-5.21)	(-15.94)	(20.39)
No. Obs	173164	173164	78467	78467
R^2	0.407	0.153	0.177	0.053

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.12
Abnormal Response Around Earnings Surprises: Other Institutional Investors

This table presents the results for the reaction to quarterly earnings announcements. The dependent variables are the announcement period cumulative abnormal return (CAR) , post-announcement CAR drift, and announcement period abnormal volume. *Sector ETF Own, Non-Sector ETF Own, Mutual Fund Own,* and *Other Instutional Own* is the fraction of shares outstanding held by sector ETFs, non-sector ETFs, mutual funds, and all other institutionanl owners, respectively. Additional variables are defined as in Table 2.9. Standard errors are robust to heteroskedasticity and clustered by announcement date.

	CAR [0,1]	Ab. Vol. [0,1]	CAR [2,21]	CAR [2,41]	CAR [2,61]
FE Quantile	0.004***	0.014***	0.003***	0.003**	0.004***
	(6.92)	(2.72)	(3.46)	(2.29)	(2.92)
Sector ETF Own	0.489***	-4.035***	-0.335***	-0.467**	-0.294
	(5.59)	(-7.93)	(-2.71)	(-2.53)	(-1.20)
Sector ETF Own * FE Quantile	-0.078***	-0.296***	0.044**	0.050**	0.045
	(-6.42)	(-3.97)	(2.57)	(2.04)	(1.43)
Non-Sector ETF Own	-0.243***	-1.670***	-0.025	0.061	0.040
	(-6.00)	(-4.81)	(-0.49)	(0.83)	(0.42)
Non-Sector ETF Own * FE Quantile	0.039***	0.127***	0.004	-0.006	-0.012
	(6.42)	(2.92)	(0.59)	(-0.59)	(-0.97)
Mutual Fund Own	-0.044***	0.507***	-0.015	-0.017	-0.020
	(-5.10)	(7.34)	(-1.39)	(-1.13)	(-1.05)
Mutual Fund Own * FE Quantile	0.007***	-0.009	-0.000	-0.001	-0.001
	(6.14)	(-1.07)	(-0.12)	(-0.47)	(-0.60)
Other Institutional Own	-0.019***	0.249***	0.003	0.010	0.017**
	(-5.56)	(8.03)	(0.49)	(1.48)	(2.01)
Other Institutional Own * FE Quantile	0.003***	-0.011***	-0.000	-0.001	-0.001
	(6.00)	(-2.81)	(-0.42)	(-1.27)	(-1.26)
Intercept	-0.033***	0.381***	-0.042***	-0.058***	-0.066***
	(-7.36)	(8.04)	(-6.83)	(-5.73)	(-4.77)
Interactions w/ FE Quantile	Yes	Yes	Yes	Yes	Yes
No. Obs	75422	75422	75422	75422	75422
R^2	0.116	0.084	0.016	0.024	0.019

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 2.13
Abnormal Response Around Earnings Surprises: Top and Bottom Groups with Other Institutional Investors

This table presents the results for the reaction to quarterly earnings announcements. The dependent variables are the announcement period cumulative abnormal return (CAR) , announcement period abnormal volume, and post-announcement CAR drift. *Top Group* takes a value of 1 (0) for the highest (lowest) two quintiles of positive (negative) earnings surprises. *Sector ETF Own* , *Non-Sector ETF Own* , *Mutual Fund Own* , and *Other Instutional Own* is the fraction of shares outstanding held by sector ETFs, non-sector ETFs, mutual funds, and all other institutionanl owners, respectively. Additional variables are defined as in Table 2.9. Control variables are interacted with *Top Group* . Standard errors are robust to heteroskedasticity and clustered by announcement date.

	CAR [0,1]	Ab. Vol. [0,1]	CAR [2,21]	CAR [2,41]	CAR [2,61]
Top Group	0.052***	0.120**	0.018**	0.032**	0.052***
	(9.06)	(2.34)	(2.36)	(2.51)	(3.85)
Sector ETF Own	0.282***	-4.583***	-0.473***	-0.686***	-0.558**
	(3.02)	(-8.54)	(-3.21)	(-3.44)	(-2.25)
Sector ETF Own * Top Group	-0.666***	-2.824***	0.457**	0.510**	0.486
	(-5.18)	(-3.71)	(2.45)	(1.99)	(1.57)
Non-Sector ETF Own	-0.197***	-2.094***	-0.030	0.055	-0.085
	(-3.50)	(-4.30)	(-0.37)	(0.45)	(-0.57)
Non-Sector ETF Own * Top Group	0.318***	1.450***	0.052	-0.062	-0.049
	(4.53)	(2.72)	(0.61)	(-0.47)	(-0.30)
Mutual Fund Own	-0.017	0.531***	-0.026*	-0.031	-0.027
	(-1.51)	(5.61)	(-1.71)	(-1.46)	(-1.04)
Mutual Fund Own * Top Group	0.041***	-0.122	0.005	-0.005	-0.023
	(3.22)	(-1.25)	(0.31)	(-0.23)	(-0.80)
Other Institutional Own	-0.013***	0.224***	0.003	0.009	0.031***
	(-3.04)	(6.00)	(0.41)	(1.05)	(2.78)
Other Institutional Own * Top Group	0.021***	-0.109**	-0.000	-0.005	-0.015
	(3.94)	(-2.53)	(-0.01)	(-0.50)	(-1.12)
Intercept	-0.027***	0.656***	-0.055***	-0.102***	-0.111***
	(-4.50)	(10.92)	(-7.00)	(-8.73)	(-7.43)
Controls Variables & Interactions w/ Top Group	Yes	Yes	Yes	Yes	Yes
No. Obs	25171	25171	25171	25171	25171
R^2	0.151	0.074	0.026	0.037	0.034

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

CHAPTER 3

ETF RETURN PREDICTABILITY

3.1 Introduction

The value of active fund management is an important question that has long been the source of much debate. If managers are able to generate abnormal returns, then investors are well served to continue investing in actively managed mutual funds. On the other hand, if managers are not skilled or not skilled enough to make up for the added expense of active management, then investors are better served to invest in cheaper, passive funds. Not surprisingly, the literature extensively studies the topic; however, it is yet to arrive at a clear consensus. Previous studies often use the existence of performance persistence as evidence of manager skill. Several earlier studies find return persistence and attribute this to manager skill (e.g. Grinblatt and Titman, 1992; Hendricks, Patel, and Zeckhauser, 1993; Elton, Gruber, and Blake, 1996; Gruber, 1996). However, Carhart (1997) finds the "hot hands" effect largely disappears when controlling for momentum with the exception of persistence among poor performers.

More recently, there have been a host of papers with differing conclusions on performance persistence. Elton, Gruber, and Busse (2004) study S&P 500 index funds and find that past performance is predictive of future performance. Bollen and Busse (2005) find evidence of short-term mutual fund performance persistence even after controlling for momentum and conclude managers exhibit stock-picking ability over the short-term. Cohen,

Coval, and Pástor (2005), Avramov and Wermers (2006), Kosowski et al. (2006), and Cremers and Petajisto (2009) also find evidence of performance persistence. However, Barras, Scaillet, and Wermers (2010) find some evidence of persistence for mutual funds before 1996 but very few by the end of their sample in 2006. Busse and Tong (2012) conclude that industry selection and market timing account for one-third of mutual fund performance and are a primary driver of performance persistence. In contrast, Lou (2012) finds that performance persistence can be explained by factors other than manager skill. He uses expected fund flows into stocks to show that temporary price pressure placed upon stocks can lead to higher mutual fund returns for funds holding stocks with high expected flow-induced trading. Although there is a rich literature that studies the topic, the conclusions on manager skill remain mixed.

In this paper, I investigate fund performance persistence and its implications concerning manager skill using exchange-traded funds (ETFs). The structure and purpose of ETFs make them an ideal investment for comparison with mutual funds. Mutual funds and ETFs share many similarities, offer many of the same diversification benefits, have many of the same fund families, and often hold similar investments in their portfolios. ETFs also have several key differences from mutual funds that I make use of in this study. First, the vast majority of ETFs are passively managed funds that track an index[16]. This is in contrast to actively managed mutual funds which rely on manager skill for stock picking ability in an attempt to beat the benchmark index. Another useful difference between funds is that all ETF shares trade at market value and can be sold short while mutual funds trade at net asset value (NAV) and very few can be shorted. Gruber (1996) notes that if there is manager ability that is not priced (given that mutual funds trade at NAV instead of an exchange-set price), there should be performance

[16] In recent years, some actively managed ETFs have started but the vast majority of ETFs are passively managed. I focus exclusively on passively managed funds in this paper.

persistence. While manager skill should play no role in performance persistence of ETFs, if there are other factors related to the fund management or fund family leading to performance persistence, then these factors should be accounted for in the ETF price. This is also an advantage of ETFs over index mutual funds for comparison with actively managed mutual funds given that index mutual funds trade at NAV. So, if persistence among mutual funds is due entirely to manager skill, then there should be no performance persistence among ETFs. On the other hand, if there is performance persistence for ETFs, then this suggests there are factors other than manager skill which can lead to performance persistence.

I begin by testing for the existence of performance persistence amongst ETFs. Using daily returns, I estimate abnormal returns each quarter using Carhart's (1997) four-factor model and test for their predictive power over future abnormal returns. Similar to Bollen and Busse (2005), I find evidence of ETF performance persistence. This persistence is short lived, though, with the exception of the lowest performing funds which tend to continue underperforming. Using decile portfolios formed on the ranking of the previous period alpha, ETFs in the highest performance decile see an average quarterly abnormal return of 0.44% in the following quarter. By two periods after forming the portfolios, however, return persistence largely fades with the exception of the worst performing funds. Cross-sectional regression analysis also indicates significant performance persistence that lasts one quarter.

I next utilize the method of Busse and Tong (2012) to separate performance persistence into stock and industry components. Busse and Tong use these measures to infer manager skill in choosing individual stocks as well as market timing and choosing successful industries. I find that industry exposure rather than individual stock composition drives almost all of the ETF abnormal performance and the majority of performance persistence. Using estimated net fund

returns, I form portfolios based on the decile rankings of both industry and stock selection performance and calculate the total performance in the following period. Industry exposure accounts for over two-thirds of the spread between the top and bottom deciles. The predictive power of the industry exposure is short-lived, however, fading after one quarter. I further confirm the findings using cross-sectional regressions.

Another method used to explain performance persistence proposed by Lou (2012) is a flow-based return effect. The intuition behind this is fairly simple. Due to investors chasing returns, funds with high past returns are likely to see higher future flows while low performing funds are likely to see lower future flows. Fund inflows (outflows) lead to increased purchases (sales) of stocks by the fund. Such flow-induced trading aggregated amongst all funds can lead to stock price pressure and ultimately drives fund returns based upon the portfolio of these price pressured stocks. When applying Lou's methodology, I find that this flow-driven return effect explains nearly all of the performance persistence for ETFs. This effect is short lived, however, as the majority of the predictive power is lost after one quarter.

I next modify Lou's (2012) methodology to test the predictive power of a fund's industry exposure as well as stock composition. I first calculate the expected ETF flows into each two-digit SIC industry. I then find the ETF's weighted portfolio exposure to industries and their respective expected industry flows. I then rank funds based upon their exposure to industries likely to see price pressure from flow-induced trading. I find this fully accounts for the top decile of performance persistence and 65% of the bottom decile performance persistence. On the other hand, the flow-driven effect attributable to individual stocks in the portfolio fails to explain the performance persistence of past winners but explains nearly all of the negative performance persistence of past losers.

This paper has dual contributions, both for the mutual fund and exchange-traded fund literatures. This is the first study, to my knowledge, that shows ETFs exhibit performance persistence. I find that most of this is attributable to industry selection rather than stock selection. In other words, the average fund would experience similar persistence if it invested in other stocks from the same industry instead of the specific stocks included in the portfolio. Furthermore, it is an industry flow-driven return effect which drives a majority of the performance persistence, particularly any positive persistence. The stock flow-driven return effect additionally explains most of the negative fund persistence. A second contribution is this paper demonstrates that it is possible for a fund to exhibit performance persistence for reasons other than manager skill. Previous papers in the mutual fund literature use performance persistence to conclude mangers have stock-picking and market-timing abilities. While this paper's findings on performance persistence do not necessarily mean that manager skill does not lead to performance persistence amongst mutual funds, it does suggest that caution should be used in inferring manager skill from performance persistence. It is also possible that some results previously attributed to manager skill may be due to other factors.

The rest of the paper is organized as follows. Section 3.2 details the data and methodology. The empirical results on performance persistence, performance decomposition, and the flow-driven return effect appear in Section 3.3 and Section 3.4 concludes.

3.2 Data and Methodology

I obtain daily ETF and stock returns from the Center for Research in Security Prices (CRSP) stock database from July 2004 until June 2013. Additional information on fund characteristics comes from CRSP survivorship-bias-free mutual fund database. ETF portfolio holdings come from Thomson Reuters and CRSP mutual fund holdings databases. The holdings

sample consist of 959 equity ETFs- jointly in the two databases- out of 1,167 total equity ETFs listed in the CRSP mutual fund database.

Towards the end of the sample, a small number of actively managed ETFs begin. So as to preserve the comparison of actively managed mutual funds to passively managed ETFs, I exclude actively managed ETFs from the sample. I use Morningstar Direct to identify actively managed funds.

To measure quarterly performance, I use daily returns to estimate regression parameters for each fund. I use Carhart's (1997) four-factor model and add the lag of each factor, following Bollen and Busse (2005), to capture the effect of infrequent trading of individual stocks on fund returns:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}RMRF_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}UMD_t + \beta_{5,i}RMRF_{t-1} + \beta_{6,i}SMB_{t-1} + \beta_{7,i}HML_{t-1} + \beta_{8,i}UMD_{t-1} + \varepsilon_{i,t}\ , \quad (11)$$

where $R_{i,t} - R_{f,t}$ is the excess return for fund *i* over the risk-free rate on day *t*; $RMRF$, SMB, and HML are the Fama and French (1993) factors and UMD is Carhart's (1997) momentum factor[17].

3.3 Empirical Results

3.3.1 Performance Persistence

I begin by testing for performance persistence amongst ETFs in Table 3.1. In quarter zero, I rank funds into deciles based upon the alpha estimated in Equation 11 and then calculate the performance of each decile in the ensuing quarters. First, I average the returns within each decile each quarter and then average across quarters. The difference between the top decile (decile #10) and the bottom decile (decile #1) as well as the differences between the top five and

[17] Data for these factors are obtained from Ken French's website.

bottom five deciles is presented in the table along with the statistical significance calculated assuming unequal variances. The top decile earns an average abnormal return of 0.0070% per day while the bottom decile has an average abnormal return of -0.0262% per day. This corresponds to an average quarterly abnormal return of 0.44% and -1.63%, respectively[18]. By two quarters after forming the portfolio, most of the predictive power dissipates with the exception of the lowest performers which have strong persistence. Carhart (1997) also finds that funds with poor past performance persist strongly. The Spearman rank correlation between past alpha decile and future alpha is statistically significant at the 1% level, further indicating evidence of persistence.

While the performance of the top decile in quarter one is similar to Bollen and Busse's (2005) findings for actively managed mutual funds, the bottom decile is considerably more negative. To ensure the result is not driven by a handful of outliers, I winsorize at the bottom and top one-tenth of one percent. In untabulated results, I also repeat the same test for the full sample of actively managed mutual funds over the same period. Bollen and Busse's sample covers only 230 mutual funds from 1985 to 1995 and finds a spread between the top and bottom decile of 1.16% in quarterly terms. When I repeat the test for actively managed mutual funds over the ten year period beginning July 2004, I find a top minus bottom decile spread of 1.89% per quarter which is considerably closer to the spread found for ETFs.

To provide additional insights and further test for the relationship between past and future performance, Table 3.2 uses the intercept, alpha, from Equation 11 to regress future alpha on past alphas. I use various combinations of past alpha from time *t-1* to time *t-3* and look at their relationship with future alpha in times *t*, *t+1*, and *t+2*. The regressions are estimated following the methodology of Fama and MacBeth (1973) with Newey-West (1987) standard errors to

[18] To estimate quarterly returns from average daily abnormal returns, α, I compute $(1+\alpha)^{63}-1$.

correct for autocorrelation and heteroskedasticity. Similar to the findings in Table 3.1, the results show a significant positive relationship between past risk-adjusted returns and those in the following quarter. This relationship is robust to the inclusion of two additional alpha lags. Persistence is short lived, however, fading after one quarter as there is no statistically significant relationship between performance separated by more than one quarter apart. This is consistent with the findings in Table 3.1 where there is positive performance persistence for the top funds lasting only one quarter.

The existence of performance persistence for ETFs has important implications. The mutual fund literature uses the existence of performance persistence amongst actively managed funds as evidence to suggest mutual fund managers are skilled. However, manager skill does not play a role in passively managed ETFs and hence ETF performance persistence must be attributable to something else. If factors other than manager skill lead to performance persistence for ETFs, then at least part of mutual fund performance persistence may be due to reasons besides manager skill.

The methodology employed in Tables 3.1 and 3.2 is used to test for stock selection skill in the mutual fund literature. However, in addition to being able to pick stocks, another goal of managers is market timing by interpreting macroeconomic environments to pick good asset and industry allocations. Bollen and Busse (2005) use the market timing models of Treynor and Mazuy (1966) and Henriksson and Merton (1981) to conclude there is performance persistence also attributable to market timing. These models are designed to capture the convexity of fund returns attributable to market timing and test for the sensitivity of fund returns to excess market returns in up and down markets. In untabulated results, I perform similar tests on the ETF sample but do not find evidence of performance persistence.

The distinction between stock-picking and market timing ability is an important one when considering mutual fund performance persistence. Busse and Tong (2012) note that the timing methods of Treynor and Mazuy (1966) and Henriksson and Merton (1981) emphasize managers switching between stocks and cash during different market conditions; however, since funds must stay fully invested (save for balanced and asset allocation funds), these may not fully capture market timing skill. Instead, Busse and Tong (2012) suggest a methodology that captures the ability to select industries well as a more realistic choice of the market timing skills of mutual fund managers. In the next section, I utilize this methodology to determine the driver behind ETF performance persistence.

3.3.2 Performance Decomposition

Given the finding that performance persistence exists for ETFs, I next seek to identify the source of the persistence. I utilize the methodology of Busse and Tong (2012) to decompose performance into two components which I label as industry-exposure and stock-composition skill. They find that a manager's choice of industry exposure accounts for one-third of fund performance and the remainder is due to stock-composition. Additionally, they find industry selection to be the driver behind performance persistence, something they conclude is a manager's ability to time the market. For the ETF scenario where manager skill is not a factor, this methodology identifies how much of performance persistence arises from a fund's exposure to different industries versus its portfolio of particular stocks.

I begin by collecting the stock holdings of ETFs from CRSP and Thomson Reuters mutual fund holdings databases. CRSP has monthly data while Thomson Reuters has quarterly data. Whenever possible, I use the monthly CRSP data so as to use the most up-to-date

information. Portfolio weights from the holdings data are then merged together with daily stock return in the CRSP daily stock file in order to calculate the gross daily fund return, $R_{Gross,t}$:

$$R_{Gross,t} = \sum_{i=1}^{n} r_{i,t} w_{i,m} \ , \tag{12}$$

where $r_{i,t}$ is the return of stock *i* on day *t* and $w_{i,m}$ is the weight of stock *i* in the ETF during month m. The gross return is then used to estimate alpha using the following regression:

$$R_{Gross,t} - R_{f,t} = \alpha_{Gross} + \beta_{1,i} RMRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} UMD_t + \beta_{5,i} RMRF_{t-1} + \beta_{6,i} SMB_{t-1} + \beta_{7,i} HML_{t-1} + \beta_{8,i} UMD_{t-1} + \varepsilon_{Gross,t} \ . \tag{13}$$

Next, I decompose fund performance into industry and stock components. First, I calculate the daily value-weighted return for each two-digit SIC industry. Then for each stock *i* within an ETF portfolio, I replace the daily stock return with the return of its respective industry, $r_{i,SIC,t}$[19]. For example, a fund holding Apple stock would have the stock's return replaced with the value-weighted return of all other stocks in Apple's two-digit SIC industry of 35. Each industry return receives the same portfolio weight as the stock it replaces. Then, the weighted daily return, $R_{Industry,t}$, is calculated for each fund by summing the product of the industry returns and respective weights:

$$R_{Industry,t} = \sum_{i=1}^{n} r_{i,SIC,t} w_{i,m} \ . \tag{14}$$

This new set of returns excludes the role of the individual stock components and instead captures the influence of the industry composition for fund returns. In other words, this can show if performance persistence is due to having the right industry exposure at the right time as opposed to holding the right stock within an industry.

I then use this set of returns to estimate the following regression:

[19] The specific stock is not included when calculating the industry return. This enables a cleaner measure of the potential return a fund could have made if it had chosen other stocks in the industry instead of the specific stock.

$$R_{Industry,t} - R_{f,t} = \alpha_{Industry} + \beta_{1,i}RMRF_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}UMD_t + \beta_{5,i}RMRF_{t-1} + \beta_{6,i}SMB_{t-1} + \beta_{7,i}HML_{t-1} + \beta_{8,i}UMD_{t-1} + \varepsilon_{Industry,t} \,. \quad (15)$$

I interpret the intercept, $\alpha_{Industry}$, as the performance attributable to a fund's industry exposure. Similar to Equation 11, I estimate the regression for each ETF each quarter using daily returns. Performance attributable to the individual stock composition, α_{Stock}, is then defined as the difference between α_{Gross} from Equation 13 and $\alpha_{Industry}$ from Equation 15.

While gross returns are important for assessing a fund's performance, ultimately it is a fund's net returns that are most significant. The gross return ignores transaction costs, fund expense ratio charges, non-U.S. equity holdings, and any portfolio rebalancing that occurs during the month. Given this, I also perform the analysis using estimated net fund returns. The ETF total net returns are easily obtainable from CRSP; however, no such data exists in CRSP for the industry and stock components. Hence, it is necessary to use the gross returns and then make a cost adjustment to perform the return decomposition analysis. To do this, I obtain the net monthly ETF returns from the CRSP stock files. I then subtract this from the gross monthly fund return and divide the difference by the number of trading days in the month. This difference captures the estimated daily impact of the expense ratio, transaction costs, nondomestic equity holdings, and index changes that occur during the month. Then, to obtain the estimated net total and industry returns, I subtract this daily gross-net difference from the daily gross total ($R_{Gross,t}$) and industry ($R_{Industry,t}$) returns. I use this new time series of returns to re-estimate Equations 13 and 15.

I begin by looking at the univariate statistics both with and without the cost adjustment in Table 3.3. Coefficients are averaged across all funds each quarter and then averaged across quarters. The *t*-statistic for alpha is calculated using Newey-West standard errors. This table

gives an initial estimate of the relative importance of the industry and individual stock roles in ETF performance. In Panel A, the average gross total alpha is 0.0288%, which is greater than zero, as expected, given there is no cost or expense adjustment. The key takeaway from this table is that the total alpha is entirely explained by the industry-exposure alpha. In fact, the average gross stock-composition alpha is negative. This gives an initial indication that ETF performance is driven by the industry composition of the fund rather than the makeup of the individual stocks. The cost-adjusted returns in Panel B further indicate that it is the stock components rather than the industry exposure, on average, that is responsible for a fund's negative performance.

Next, I test for the influence of the industry and stock components for performance persistence in Table 3.4. Similar to before, I rank the funds into deciles based upon performance (total gross alpha, industry-exposure alpha, or stock-composition alpha) at time zero and calculate the returns of each portfolio in the future. I first average the performance within each decile each quarter and then average across quarters. Notably, the results are much more positive than those found in Table 3.1. Given that Table 3.4 uses gross returns without cost adjustments, this difference is expected. First, when forming decile portfolios based upon past gross returns, returns show strong persistence in the following quarter. The Spearman rank correlation between past gross return decile and future gross returns is12% which is statistically significant at the 1% level. This persistence continues in the ensuing quarters to a smaller but still significant degree. Next I form portfolios based on industry exposure alpha in order to predict future gross returns. There is a strong degree of persistence in the quarter directly following the ranking quarter. The difference between the top and bottom deciles, the difference between the top and bottom halves, and the Spearman rank correlation test are all statistically

significant at the one-percent level. This persistence quickly fades and is largely gone within two quarters. Finally, I form the decile portfolios on the stock-composition alpha and find only weak indications of performance persistence.

The results using estimated net returns, which use a cost adjustment, are reported in Table 3.5. Similar to before, total alpha and industry-exposure alpha both exhibit predictive power over future total alpha. Using estimated net returns predicts persistence with a daily average abnormal return spread of 0.0384% between the top and bottom decile portfolios. Likewise, portfolios formed on past industry-exposure see an average 0.027% daily abnormal return when going long the top decile and short the bottom decile. However, this persistence is short lived and fades after just one quarter. Spearman rank correlation results reinforce this as the correlation is strongly significant in quarter 1 but statistically insignificant afterwards. Also, the performance attributable to the individual stocks has no statistically significant persistence when comparing the top and bottom deciles in quarter 1. Together with Table 3.4, this initially suggests that industry rather than individual stock components are the main driver behind ETF performance persistence.

I next use Fama-MacBeth (1973) cross-sectional regressions with Newey-West (1987) standard errors to further test for the driver of performance persistence in Table 3.6. I use past total, industry-exposure, and stock-composition alphas to predict future gross total alpha in Panel A and to predict future estimated net-alpha in Panel B. Consistent with the results found using decile rankings, performance is strongly persistent based upon past industry-exposure alpha using estimated net returns. In other words, funds which have stocks from the best performing industries tend to have higher future abnormal returns in the following period. This effect is short lived, however, with persistence only lasting one quarter. Interestingly, only industry-

exposure alpha positively predicts gross and net returns one quarter later; both total alpha and stock-composition alpha have a statistically insignificant coefficient. This suggests that industry exposure is the predominant driver behind performance persistence and that the individual selection of stocks has little, if anything, to do with ETF performance persistence. Given that ETFs are not designed to pick stocks, this result is plausible. It is possible that an ETF could simply be heavily weighted in well-performing industries, by chance, and that those industries continue to do well driving up fund returns. It is likely that such a trend would be short lived too, consistent with my findings.

3.3.3 Flow-Driven Return Effect

Thus far, the results show that ETFs exhibit performance persistence that is driven primarily by the fund's industry exposure. However, I consider an alternative approach to analyze the source of ETF performance persistence. Lou (2012) suggests that the effects of flows can explain mutual fund return predictability. Since investors chase returns, funds with high past performance tend to see higher inflows which they then reinvest in most of the same stocks already in the portfolio. Stocks which are held by many funds with these flow-induced trades are likely to see upward price-pressure associated with many funds purchasing additional shares of the stock. Fund portfolios holding these price-pressured stocks are, in turn, more likely to have better short-term performance as well. A similar explanation follows for funds with poor past performance having lower expected future flows, downward price-pressure on stocks, and ultimately lower fund returns. If this explanation is correct for mutual funds, then the same should hold true for ETFs. I next empirically test this theory for ETFs.

To test this, I follow the methodology of Lou (2012) and begin by calculating the expected flow-induced trading for each stock, $E_t[FIT_j]$. In other words, this measures the

predicted flows from all ETFs into individual stocks. $E_t[FIT_j]$ is calculated in each period *t* for stock j as:

$$E_t[FIT_j] = \frac{\sum_{i=1}^{n} shares_{i,j,t} * E_t[flow_i]}{\sum_{i}^{n} shares_{i,j,t}}, \tag{16}$$

where $E_t[flow_i]$ is the expected flow to ETF i and $shares_{i,j,t}$ is the number of j stock shares held by ETF i. In order to estimate expected flows, I use the lagged four-factor alpha in Equation 11. Alpha is a suitable choice to predict flows given the return chasing behavior of ETF investors (Clifford, Fulkerson, and Jordan, 2014).

An advantage of this methodology for ETFs over actively managed mutual funds is that ETF inflows and outflows do not change the portfolio weights of the underlying holdings since ETFs track an index. On the other hand, Lou finds that only 62% of mutual fund inflows are invested in the same securities and hence must account for this in the estimation procedure. This results in a simpler, more accurate calculation of the expected flow-induced trading.

Next, the value-weighted average $E_t[FIT_j]$ is calculated for each ETF based upon its portfolio of stocks. This estimates the price-pressure due to flow-induced trading of the average stock held by the ETF. Specifically, this is calculated as:

$$E_t[FIT_i^*] = \sum_{j=1}^{m} \left(E_t[FIT_j] * \omega_{i,j,t}\right) \tag{17}$$

where $\omega_{i,j,t}$ is the weight of stock j in ETF i's portfolio at time t. I then rank ETFs into deciles based on $E_t[FIT_i^*]$ and calculate the average daily abnormal return for each decile portfolio over the next three quarters. This provides a way to test if the average fund which holds stocks with the highest expected inflows and upward price-pressure performs better than the portfolio holding stocks with the highest expected outflows and downward price-pressure.

The results on flow-driven returns appear in Table 3.7. The abnormal returns show a positive relationship with the ranking of $E_t[FIT_i^*]$. The top and bottom deciles have an average daily abnormal return of 0.0063% and -0.0274%, respectively. This accounts for nearly all of the performance persistence in both the top and bottom deciles seen in Table 3.1. Also, similar to the findings with performance persistence, the ability to positively predict future performance of the top decile group using $E_t[FIT_i^*]$ quickly diminishes after one quarter while the bottom decile maintains negative persistence. Overall, the stock flow-driven return effect does a good job of explaining ETF performance persistence.

In light of the previous findings on the importance of industry selection, I next alter the flow-based test to instead use expected flows into an industry as a factor to predict fund performance. $E_t[FIT_k]$ is calculated in each period *t* for each two-digit SIC industry as:

$$E_t[FIT_k] = \frac{\sum_{i=1}^{n} Dollar\ Value_{i,k,t-1} * E_t[flow_i]}{\sum_{i=1}^{n} Dollar\ Value_{i,k,t-1}}, \tag{18}$$

where $Dollar\ Value_{i,k,t-1}$ is the market value of all stocks in two-digit SIC industry *k* held by ETF *i* at time *t-1*. Equation 18 differs from Equation 16 by calculating the expected flows by all ETFs into a particular industry instead of a particular stock. In the second step, I calculate the ETF portfolio value-weighted average $E_t[FIT_k]$ for each fund:

$$E_t[FIT_i^{**}] = \frac{\sum_{i=1}^{m} Dollar\ Value_{i,k,t-1} * E_t[FIT_k]}{\sum_{i=1}^{m} Dollar\ Value_{i,k,t-1}}. \tag{19}$$

I then rank ETFs into deciles based on $E_t[FIT_i^{**}]$ and calculate the average daily abnormal return of the decile portfolio over the next three quarters. Intuitively, this tests if funds holding stocks from industries that are expected to have high inflows from all ETFs tend to have higher future returns while those expected to have lower ETF inflows.

The results for the industry flow-based persistence appear in Table 3.8. They show strong persistence for one period that fades in the ensuing quarters. The top decile group accounts for all of the performance persistence documented for the top decile group in Table 3.1. The daily alpha average of the bottom decile group is -0.0169%, which accounts for 65% of the persistence found for the bottom decile group in Table 3.1. In general, the results suggest that a majority of the flow-induced trading found in Table 3.7 can be attributed to the industry. In other words, if the ETF portfolio had a different set of stocks but kept the same industry exposures, then a majority of performance persistence attributable to the flow-driven return effect would hold.

I next isolate the flow-driven return effect attributable solely to the individual stocks and not the industry exposure. Similar to the decomposition used for returns above, the expected flow-induced trading attributable to the stock composition, $E_t[FIT_i^{***}]$, is defined as the difference between $E_t[FIT_i^{*}]$ in Equation 17 and $E_t[FIT_i^{**}]$ in Equation 19. Intuitively, this isolates the impact on fund returns due to the flow-based price pressure of individual stocks in the portfolio that are not common to, or are in excessive of, other stocks in the same industry. The results appear in Table 3.9 and indicate two key findings. First, the positive persistence for the top decile group is driven by the industry and not the stock flow-effect. In fact, the top two deciles of $E_t[FIT_i^{***}]$ predict negative returns in quarter one. Second, the stock effect rather than the industry effect best explains the negative performance persistence of the bottom groups. For example, the performance of the bottom decile group formed on the stock flow-based effect explain 98% of the bottom decile performance persistence in Table 3.1 compared to the industry flow-based effect which only explains 65%. Both effects do have a significant negative impact on the returns of the lowest group, but the stock effect explains much more.

Overall, the results indicate that the flow-induced trading return effect does a good job of explaining ETF performance persistence. Specifically, the expected price pressures from ETF flows into industries are especially useful in explaining positive performance persistence of past winners. On the other hand, flow-based effects of individual stocks better explain the negative persistence of past losers. This is particularly helpful in understanding how a passively managed fund could exhibit performance persistence given that manager skill is not a factor for ETFs.

3.4 Conclusion

There has been a longstanding debate as to the value of active fund management. The answer to this question is one with profound implications for investors and fund families alike. Although an extensive literature studies the topic, there is still no clear consensus on the topic. Carhart (1997) and Barras, Scaillet, and Wermers (2010) find little to no evidence of performance persistence while others such as Bollen and Busse (2005) and Cremers and Petajisto (2009) do find persistence and value in active management. Still others find there is persistence that is attributable to other factors such as industry selection and market timing (Busse and Tong, 2012) and a flow-driven return effect (Lou, 2012).

In this paper, I seek to shed further light on the subject by investigating performance persistence amongst ETFs. In many respects, ETFs are similar to mutual funds as far as the types of assets they hold and the diversification benefits they provide. I utilize two key differences between the funds to glean insight into the implications of fund performance persistence. First, ETFs are passively managed and track an index. This eliminates manager ability as a possible explanation for ETF performance persistence. Second, ETFs trade at market prices and can be sold short. This is in contrast to mutual funds which trade only at net asset value and do not have the same ability to arbitrage away factors such as manager skill.

With these differences in mind, I first test for and find that ETFs demonstrate performance persistence similar to actively managed mutual funds. Prior quarter abnormal returns predict future performance with a spread of 2.07% quarterly abnormal returns between the highest and lowest past performance deciles. This return predictability quickly dissipates, except for the worst performing funds which show greater signs of persistently poor returns. A decomposition of net returns into industry exposure and stock composition components reveals that ETF performance and performance persistence is driven predominately by industry exposure. This suggests that funds can obtain similar abnormal returns, on average, by investing in other stocks from the same industry. Finally, I demonstrate that the underlying source of performance persistence links to flow-based trading. Popular ETFs attract flows which they then reinvest into the same stocks in the portfolio. Across all funds, this leads to price pressure on the individual stocks and funds holding portfolios of these stocks correspondingly see fund returns driven in a similar fashion. Furthermore, it is flows into the industry rather than to specific stocks that largely drives this effect. The stock flow-driven return effect does explains nearly all of the negative performance persistence while the industry flow-driven return effect fully accounts for positive performance persistence.

These findings have important implication for both ETFs and mutual funds. First, it documents that ETFs exhibit performance persistence and that this can be explained by flow-based reasons. Second, it demonstrates that it is possible for funds to exhibit performance persistence for reasons other than manager skill. Some in the mutual fund literature use performance persistence of actively managed mutual funds as evidence of manager skill. While the conclusions of this paper do not preclude the possibility that manager skill exists for mutual

funds, it does suggests that caution should be used when relating performance persistence to manager skill.

REFERENCES

Avramov, D., & Wermers, R. (2006). Investing in mutual funds when returns are predictable. *Journal of Financial Economics*, *81*(2), 339-377.

Barras, L., Scaillet, O., & Wermers, R. (2010). False discoveries in mutual fund performance: Measuring luck in estimated alphas. *Journal of Finance*, *65*(1), 179-216.

Bollen, N. P., & Busse, J. A. (2005). Short-term persistence in mutual fund performance. *Review of Financial Studies*, *18*(2), 569-597.

Busse, J. A., & Tong, Q. (2012). Mutual fund industry selection and persistence. *Review of Asset Pricing Studies*, *2*(2), 245-274.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57-82.

Chevalier, J., & Ellison, G. (1999). Career concerns of mutual fund managers. *Quarterly Journal of Economics*, *114*(2), 389-432.

Clifford, C. P., Fulkerson, J. A., & Jordan, B. D. (2014). What drives ETF flows?. *Financial Review*, forthcoming.

Cohen, R. B., Coval, J. D., & Pástor, Ľ. (2005). Judging fund managers by the company they keep. *Journal of Finance*, *60*(3), 1057-1096.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, *22*(9), 3329-3365.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivor bias and mutual fund performance. *Review of Financial Studies*, *9*(4), 1097-1120.

Elton, E. J., Gruber, M. J., & Busse, J. A. (2004). Are investors rational? Choices among index funds. *Journal of Finance*, *59*(1), 261-288.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 607-636.

Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance*, *47*(5), 1977-1984.

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *Journal of Finance*, *51*(3), 783-810.

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988. *Journal of Finance*, *48*(1), 93-130.

Kosowski, R., Timmermann, A., Wermers, R., & White, H. (2006). Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance*, *61*(6), 2551-2595.

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

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703-708.

Table 3.1
Short-term Performance Persistence

This table lists the average daily abnormal returns (in %) during the quarter where the portfolio is formed on the basis of the decile ranking of the quarter abnormal return at time 0. Abnormal return is calculated using the four factor model including lagged factors. Alpha is averaged within each decile portfolio each quarter and then averaged across quarters. Two-tailed *t*-test use unequal variances. The Spearman rank correlation between the decile ranking and future alpha is shown at the bottom of the table.

Decile	Quarter 1	Quarter 2	Quarter 3
Worst 1	-0.0262	-0.0507	-0.0413
2	-0.0112	-0.0247	-0.0326
3	-0.0082	-0.0139	-0.0182
4	-0.0081	-0.0031	-0.0155
5	-0.0045	-0.0078	-0.0031
6	-0.0001	-0.0046	-0.0095
7	0.0016	-0.0102	-0.0045
8	0.0006	-0.0107	-0.0067
9	0.0039	-0.0128	-0.0039
Best 10	0.0070	-0.0015	-0.0082
Decile 10 - 1	0.0331***	0.0492***	0.033***
Top 5 deciles - Bottom 5 deciles	0.0142***	0.012***	0.0155***
Spearman Corr.	0.069***	0.0506***	0.0857***

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.2
Short-term Performance Persistence Regressions

This table presents the results of the Fama-MacBeth (1973) cross-sectional regressions relating abnormal returns in quarters t-1 through t-3 to future abnormal returns in quarters t, t+1, and t+2, as denoted by the column header. Abnormal returns are calculated using the four factor model including lagged factors. Standard errors are adjusted for autocorrelation following Newey and West (1987).

	Alpha(t)			Alpha(t+1)			Alpha(t+2)		
Decile	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3
Alpha(t-1)	0.044**	0.037**	0.051***	0.026	0.026	0.026	0.045	0.010	-0.032
	(2.54)	(2.14)	(3.79)	(0.69)	(0.70)	(0.61)	(0.78)	(0.19)	(-0.55)
Alpha(t-2)		-0.005	-0.002		0.011	-0.011		0.026	0.048
		(-0.15)	(-0.07)		(0.29)	(-0.29)		(0.91)	(1.11)
Alpha(t-3)			-0.003			0.030			0.048
			(-0.08)			(1.43)			(0.95)
Intercept	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(-0.79)	(-0.56)	(-0.34)	(-1.23)	(-0.88)	(-0.75)	(-1.32)	(-1.36)	(-1.32)
Observations	23590	22228	20874	23330	21974	20648	21991	20663	19368
R-squared	0.020	0.063	0.090	0.040	0.067	0.097	0.048	0.076	0.101

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.3
Factor model estimates using daily data

This table reports statistics from the four-factor model regressions estimated over quarterly horizons. Coefficients are averaged across all funds each quarter and then averaged across quarters. Abnormal returns are calculated using the four factor model and their lags. Industry-exposure alpha uses the fund return generated using each stock's two-digit SIC industry value-weighted return instead of the stock's return for the regressand in the four factor model. Stock-composition alpha is the difference between toal alpha and indsutry-exposure alphas. Newey-West corrected standard errors are used to estimate *t*-statistics.

Panel A: Alpha calculated using gross returns and no cost adjustment

	α	*t*-statistic	β_m	β_{Smb}	β_{hml}	β_{umd}	*R*-square
Total	0.000288	13.15	1.0210	0.1487	0.0541	0.0115	0.840
Industry	0.000316	27.64	0.9794	0.0491	0.0382	0.0132	0.932
Stock	-0.000027	-3.12					

Panel B: Alpha calculated using gross returns and cost adjustment

	α	*t*-statistic	β_m	β_{Smb}	β_{hml}	β_{umd}	*R*-square
Total	-0.000026	-3.96	1.0222	0.1489	0.0541	0.0106	0.840
Industry	-0.000001	-1.87	0.9807	0.0488	0.0377	0.0124	0.927
Stock	-0.000025	-3.11					

Table 3.4
Performance Persistence As Measured by Gross Alpha and its Components

This table lists the average total gross alpha (in %) during the quarter based upon portfolios formed on the basis of the decile ranking of the performance in quarter 0. Portfolios are formed on either total gross, industry-exposure or stock-composition alpha as indicated by each column header. Abnormal returns are calculated using the four factor model including lagged factors. Industry-exposure alpha uses the fund return generated using each stock's two-digit SIC industry value-weighted return instead of the stock's return for the regressand in the four factor model. Stock-composition alpha is the difference between toal alpha and indsutry-exposure alphas. Alpha is averaged within each decile portfolio each quarter and then averaged across quarters. Two-tailed *t*-test use unequal variances. The Spearman rank correlation between the decile ranking and future alpha is shown at the bottom of the table.

	Total Alpha			Industry-Exposure Alpha			Stock-Composition Alpha		
Decile	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3
Worst 1	-0.0003	-0.0023	0.0110	0.0022	0.0122	0.0158	0.0137	0.0130	0.0064
2	0.0147	0.0212	0.0207	0.0165	0.0238	0.0332	0.0162	0.0210	0.0188
3	0.0270	0.0351	0.0245	0.0229	0.0276	0.0265	0.0183	0.0263	0.0230
4	0.0241	0.0268	0.0297	0.0297	0.0358	0.0305	0.0267	0.0247	0.0309
5	0.0303	0.0302	0.0301	0.0338	0.0282	0.0338	0.0274	0.0334	0.0338
6	0.0354	0.0366	0.0303	0.0363	0.0350	0.0334	0.0322	0.0312	0.0281
7	0.0346	0.0333	0.0385	0.0378	0.0315	0.0320	0.0383	0.0333	0.0319
8	0.0439	0.0377	0.0421	0.0391	0.0307	0.0398	0.0471	0.0331	0.0433
9	0.0476	0.0313	0.0396	0.0374	0.0330	0.0375	0.0442	0.0327	0.0434
Best 10	0.0316	0.0285	0.0262	0.0328	0.0213	0.0110	0.0244	0.0305	0.0333
Decile 10 - 1	0.0318***	0.0307***	0.0152	0.0305***	0.009	-0.0047	0.0106	0.0174*	0.0268**
Top 5 deciles - Bottom 5 deciles	0.0193***	0.0112***	0.0121***	0.0156***	0.0047**	0.0027	0.0167***	0.0084***	0.0134***
Spearman Corr.	0.12***	0.0679***	0.0946***	0.0931***	0.0248***	0.0349***	0.0905***	0.0623***	0.0891***

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.5
Performance Persistence As Measured by Estimated Net Alpha and its Components

This table lists the average estimated total estimated net-alpha (in %) during the quarter based upon portfolios formed on the basis of the decile ranking of the performance in quarter 0. Portfolios are formed on either total, industry-exposure or stock-composition alpha as indicated by each column header. Abnormal returns are calculated using the four factor model including lagged factors. Industry-exposure alpha uses the fund return generated using each stock's two-digit SIC industry value-weighted return instead of the stock's return for the regressand in the four factor model. Stock-composition alpha is the difference between toal alpha and indsutry-exposure alphas. Performance is adjusted for cost by subtracting the difference between gross and net returns. Alpha is averaged within each decile portfolio each quarter and then averaged across quarters. Two-tailed *t*-test use unequal variances. The Spearman rank correlation between the decile ranking and future alpha is shown at the bottom of the table.

	Total Alpha			Industry Exposure Alpha			Stock Composition Alpha		
Decile	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3
Worst 1	-0.0220	-0.0111	-0.0200	-0.0127	-0.0021	-0.0159	-0.0107	-0.0138	-0.0125
2	-0.0148	-0.0034	-0.0087	-0.0275	-0.0028	-0.0031	-0.0107	-0.0088	-0.0126
3	-0.0108	-0.0052	-0.0092	-0.0111	-0.0117	-0.0053	-0.0184	-0.0037	-0.0099
4	-0.0082	0.0000	-0.0016	-0.0031	0.0004	-0.0036	-0.0052	-0.0044	-0.0002
5	-0.0064	0.0074	-0.0050	-0.0064	-0.0009	-0.0030	-0.0020	-0.0005	0.0009
6	-0.0042	-0.0039	-0.0029	-0.0057	-0.0041	-0.0032	0.0031	-0.0043	-0.0049
7	-0.0046	-0.0075	-0.0025	0.0005	-0.0062	-0.0031	0.0025	0.0000	-0.0036
8	0.0051	-0.0085	0.0010	0.0037	0.0004	-0.0007	0.0037	-0.0081	-0.0004
9	0.0082	-0.0072	0.0049	0.0068	-0.0111	-0.0051	0.0081	-0.0061	0.0001
Best 10	0.0165	-0.0087	-0.0034	0.0143	-0.0096	-0.0035	-0.0118	0.0007	-0.0033
Decile 10 - 1	0.0384***	0.0024	0.0166	0.027**	-0.0074	0.0123	-0.001	0.0145	0.0092
Top 5 deciles - Bottom 5 deciles	0.0166***	-0.0046	0.0083***	0.0161***	-0.0026	0.003	0.0104***	0.0026*	0.0044**
Spearman Corr.	0.0817***	0.005	0.0639***	0.0575***	0.0031	0.0439***	0.0594***	0.0197**	0.0368***

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.6
Short-term Performance Persistence Regressions and its Components

This table presents the results of the Fama-MacBeth (1973) cross-sectional regressions for performance persistence. Abnormal returns are calculated using the four-factor model including lagged factors. Total alpha, industry-exposure alpha, and stock-composition alpha are each estimated in quarter 0 and used as indicated by each column header to predict total alpha in quarters 1 to 3. Abnormal returns are calculated using the four factor model and their lags. Industry-exposure alpha uses the fund return generated using each stock's two-digit SIC industry value-weighted return instead of the stock's return for the regressand in the four factor model. Stock-composition alpha is the difference between toal alpha and indsutry-exposure alphas. Panel A uses gross returns while Panel B uses estimated net returns which are adjusted for cost by subtracting the difference between gross and net returns. Standard errors are adjusted for autocorrelation following Newey and West (1987).

Panel A: Gross Returns

	Total Alpha			Industry-Exposure Alpha			Stock-Composition Alpha		
	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3
Performance(t-1)	0.0361	0.0632***	0.0464	0.1439	0.1113	-0.0684	0.0130	0.0407	0.0625
	(0.68)	(2.87)	(0.89)	(1.31)	(1.37)	(-0.79)	(0.25)	(1.39)	(1.30)
Intercept	0.0003***	0.0003***	0.0003***	0.0002***	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***
	(3.81)	(3.62)	(3.23)	(3.50)	(3.14)	(3.07)	(3.60)	(3.84)	(3.66)
Observations	11583	10413	9560	11583	10413	9560	11583	10413	9560
R-squared	0.0524	0.0590	0.0381	0.0313	0.0497	0.0368	0.0475	0.0486	0.0389

Panel B: Estimated Net Returns

	Total Alpha			Industry-Exposure Alpha			Stock-Composition Alpha		
	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3	Quarter 1	Quarter 2	Quarter 3
Performance(t-1)	0.0400	0.0246	0.0491*	0.1008***	-0.0091	0.0257	-0.0318	0.0380	0.0345
	(1.27)	(0.72)	(1.80)	(2.84)	(-0.26)	(0.80)	(-0.83)	(1.64)	(1.05)
Intercept	-0.0000	-0.0000	-0.0000	-0.0001	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(-0.82)	(-0.87)	(-0.75)	(-1.06)	(-0.94)	(-0.86)	(-0.91)	(-0.74)	(-0.76)
Observations	11583	10413	9560	11583	10413	9560	11583	10413	9560
R-squared	0.0443	0.0602	0.0535	0.0452	0.0458	0.0522	0.0374	0.0394	0.0276

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.7
Short-term Total Flow-Induced Price Effect

This table reports the average four-factor alpha (in %) for portfolios based on the decile ranking of *E[FIT*]* , the average expected flow-induced trading for the ETF portfolio. The expected flow-induced trading, *E[FIT]* , is calculated for each stocks based upon the aggregate expected flows by ETFs, estimated using lagged alpha. Alpha is averaged within each decile portfolio each quarter and then averaged across quarters. Two-tailed t-test use unequal variances. The Spearman rank correlation between the decile ranking and future alpha is shown at the bottom of the table.

Decile	Quarter 1	Quarter 2	Quarter 3
Worst 1	-0.0274	-0.0205	-0.0291
2	-0.0172	-0.0091	-0.0095
3	-0.0047	-0.0006	-0.0048
4	0.0015	0.0015	-0.0013
5	-0.0010	-0.0034	-0.0030
6	-0.0018	-0.0006	0.0023
7	-0.0044	0.0009	0.0002
8	0.0016	-0.0004	0.0008
9	0.0065	-0.0069	0.0000
Best 10	0.0063	-0.0041	-0.0161
Top decile - Bottom decile	0.0336***	0.0163***	0.0129***
Top 5 deciles - Bottom 5 deciles	0.0131***	0.0047***	0.0085***
Spearman Corr.	0.1073***	0.0228**	0.0521***

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.8
Short-term Industry Flow-Induced Price Effect

This table reports the average four-factor alpha (in %) for portfolios based on the decile ranking of *E[FIT*]* , the average expected flow-induced trading for the ETF portfolio. The expected flow-induced trading, *E[FIT]* , is calculated for each two-digit SIC industry based upon the aggregate expected flows by ETFs, estimated using lagged alpha. Alpha is averaged within each decile portfolio each quarter and then averaged across quarters. Two-tailed t-test use unequal variances. The Spearman rank correlation between the decile ranking and future alpha is shown at the bottom of the table.

Decile	Quarter 1	Quarter 2	Quarter 3
Worst 1	-0.0169	-0.0072	-0.0066
2	-0.0129	-0.0108	-0.0186
3	-0.0084	0.0009	-0.0114
4	0.0025	-0.0009	-0.0019
5	-0.0085	-0.0002	-0.0017
6	-0.0011	-0.0010	0.0033
7	-0.0040	-0.0019	0.0004
8	-0.0040	-0.0072	0.0024
9	-0.0026	-0.0086	-0.0040
Best 10	0.0167	-0.0074	-0.0229
Top decile - Bottom decile	0.0336***	-0.0001	-0.0162
Top 5 deciles - Bottom 5 deciles	0.0083***	0	0.0058***
Spearman Corr.	0.0826***	-0.0174*	0.0233***

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 3.9
Short-term Stock Flow-Induced Price Effect

This table reports the average four-factor alpha (in %) for portfolios based on the decile ranking of *E[FIT*]* , the average expected flow-induced trading for the ETF portfolio. The expected flow-induced trading, *E[FIT]* , is calculated for each two-digit SIC industry based upon the aggregate expected flows by ETFs, estimated using lagged alpha. Alpha is averaged within each decile portfolio each quarter and then averaged across quarters. Two-tailed t-test use unequal variances. The Spearman rank correlation between the decile ranking and future alpha is shown at the bottom of the table.

Decile	Quarter 1	Quarter 2	Quarter 3
Worst 1	-0.0257	-0.0211	-0.0321
2	-0.0171	-0.0100	-0.0194
3	-0.0074	-0.0061	-0.0014
4	0.0023	0.0002	0.0081
5	-0.0005	-0.0001	-0.0019
6	0.0047	-0.0019	-0.0035
7	0.0000	0.0014	-0.0001
8	0.0062	-0.0033	0.0030
9	-0.0015	0.0054	0.0007
Best 10	-0.0014	-0.0075	-0.0139
Top decile - Bottom decile	0.0243***	0.0136***	0.0181***
Top 5 deciles - Bottom 5 deciles	0.0132***	0.0082***	0.0085***
Spearman Corr.	0.0877***	0.044***	0.0653***

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

CHAPTER 4

ETF LIQUIDATION DETERMINANTS

4.1 Introduction

When State Street Global Advisors launched the first exchange-traded fund (ETF) to track the S&P 500 index in January 1993, it sparked the beginning of an important new type of investment. Although similar to both open-end mutual funds and close-end funds, ETFs enjoy numerous advantages over both types of funds. The tax advantages for many customers and the ability to trade intra-day at costs lower than many mutual funds has led to the total assets in U.S.-listed ETFs amassing $1.3 trillion by the end of 2012. Also, unlike the mutual fund industry, institutional investors as well as retail investors are major players in the ETF industry. With the number of ETFs created now over 1,400, not all funds have experienced success. In fact, the year 2012 saw a record breaking 73 funds liquidated through the end of September and over 17% of all ETFs ever created have liquidated.

Given the sharp increase in fund failures in recent years, investors should be aware of the risk that an ETF may close. There are several concerns surrounding fund liquidations for investors. ETFs boast that they are more tax efficient than mutual funds[20]. This is an attractive feature of ETFs and it draws many tax conscious investors. Under normal circumstances, an ETF investor incurs a capital gains tax liability when they sell the fund (assuming the fund is

[20] This is due to ETF's use of the Investment Company Act of 1940's provision allowing for "in kind" redemption. When an arbitrageur seeks to redeem shares of an ETF, the trustee distributes the underlying securities instead of cash. Since the trustee can choose to distribute shares with the highest capital gains, this allows ETFs to have lower capital gains distributions than mutual funds which do not take advantage of this redemption strategy.

held in a taxable account). The timing over this is something that investors have control over. However, when an ETF liquidates, investors are faced with unexpected tax consequences. Additionally, an investor incurs search costs for new investments when an ETF liquidates. In a market with well over a thousand ETFs to choose between, this can be a time consuming task and potentially present challenges in finding a suitable substitute, especially if the ETF is used for hedging purposes. Furthermore, investors also incur additional trading costs from needing to purchase new investments. These are all important factors impacting investors.

While this is the first paper, to my knowledge, studying the determinants of ETF liquidations, the mutual fund literature has several related studies on mutual fund mergers and liquidations. Brown and Goetzmann (1995) and Elton, Gruber, and Blake (1996) find that mutual fund survival is less likely for funds with poor performance. Zhao (2005) looks at mutual fund mergers and liquidations and finds that funds with smaller sizes and lower inflows are more likely to exit. Liquidations predominately occur for the youngest, smallest funds while within-family mergers are more likely for older funds from larger families. English, Demiralp, and Dukes (2011) find that within-family mergers are more likely than liquidations or across-family mergers when a fund has a higher management fee. They also note that funds with a 12b-1 fee are less likely to be liquidated, tend to be merged sooner than those without 12b-1 fees, and tend to take longer to liquidate than funds without 12b-1 fees.

While it is possible that mutual fund and ETF liquidations have similar determinants, differences between the two funds types may result in differences in liquidation decisions. First, exit decisions in the mutual fund industry are threefold: liquidation, within-family merger, and across-family merger. Zhao (2005) finds that 41% of mutual fund exits in his sample are liquidations. ETFs, on the other hand, either continue to operate or liquidate but do not merge.

Zhao finds key differences between the three types of mutual fund exits. Whether ETF liquidations act more similarly to mutual fund liquidations, mergers, or neither is an empirical question. Additionally, ETFs have trading concerns that mutual funds do not have due to ETFs trading throughout the day. ETFs are subject to a bid-ask spread, can trade at a premium or discount from their net asset value, and are influenced by trading volume. Furthermore, while the majority of mutual funds are actively managed, nearly all ETFs are passively managed. A consequence of this is that the ETF industry is extremely competitive on expense ratios. Other differences may arise from a much larger presence of institutional investors in ETFs than mutual funds. Also, the ETF industry is one that has greatly grown in recent years but is still relatively young compared to the mutual fund industry. It is unclear if these differences play into the decision to continue operating or close a fund.

In this paper, I seek to fill this gap in the literature and determine the fund specific, fund family, and category factors that lead to ETF liquidations. Using a sample of 1,412 ETFs from 1998 until 2012, I find 239 ETFs in my sample are liquidated by the fund family. Small fund size is a major factor leading to liquidation. The median living fund has $116.2 million in total net assets, while the median liquidated fund has only $4.2 million in the month prior to exit. The ETF market is a highly competitive one and, since most funds have very low expense ratios, they often have razor thin profit margins. A fund that is not able to attract enough dollars will fail to cover the costs for a fund family to operate it. Flows are also highly predictive of a decision to liquidate. The lower the growth in fund flows, the higher the probability of fund closure. Funds with more competitors in the same category, those from smaller families, and funds trading at a discount also exhibit greater likelihood of liquidation.

Next, I find that the decision to liquidate is also made in the context of the fund family's overall success. In multiple cases, a fund family chooses to liquidate a large portion of their ETFs or completely exit the ETF industry. Such situations mean that an individual fund could be performing relatively well but still terminate due to the fund family's weak success overall. Comparing families liquidating half or more of their funds with those with only limited liquidations, I find that ETFs included in a mass family liquidation have families that are younger, smaller, and have lower flows and returns, on average.

Given that over half of fund deaths occur within the first two years after inception, I next seek to determine if there are any key differences between young funds that live and die. I find that the average fund that dies starts smaller and grows slower in the first year of life than funds that live. Nearly all funds that liquidate are among the smallest in terms of total net assets at inception. Also, funds that die are more likely to have negative returns, lower dollar and percentage flows, and higher expense ratios in the first year of life. Additionally, trends among the Lipper category within the first year of life appear influential on the likelihood of success. Funds that liquidate tend to start in larger, more crowded categories that become increasingly crowded in the first year of life. Dead funds often come from categories that had performed well when the fund began but then perform poorly in the first year of life. One way to interpret this is that unsuccessful funds tend to come in at the tail end of trending areas. Also, there are noticeable differences in family characteristics between the average successful and unsuccessful fund. Funds that come from larger, greater flow generating, older families -especially those that were one of the first to enter the ETF market- are far more likely to survive than younger, smaller families, on average.

Last, I consider the best course of action for an investor who holds a fund with a recent liquidation announcement. Since families typically give at least a couple of weeks' notice before liquidating a fund, an investor must choose either to sell their fund early or wait until after a fund is liquidated to receive the fund's net asset value and then reallocate their funds. Selling immediately after an upcoming liquidation is announced incurs additional trading costs for an investor; however, this comes at the cost of an average negative 6 basis point daily abnormal return. Not counting transaction costs, these results suggest that investors who hold a fund that has just announced a liquidation are best served to immediately sell the fund and invest in other funds in the category.

The remainder of the paper proceeds as follows. Section 4.2 details the data, methods, and summary statistics. The results are presented in Section 4.3, followed by concluding remarks in Section 4.4.

4.2 Data, methodology, and summary statistics

Exchange-traded fund data comes from the Center for Research in Security Prices (CRSP) as well as Morningstar Direct. CRSP Survivor-Bias-Free database of U.S. mutual funds provides return, net asset value (NAV), total net assets, expense ratio, and other fund characteristic data. Trading prices, volume, and bid-ask spread data come from CRSP daily stock data. Additional fund characteristic and daily shares outstanding data come from Morningstar Direct. The sample covers 1,412 ETFs from 83 different Lipper categories from 1998 until 2012. This time period is chosen as CRSP mutual fund summary data begins in 1998 (although return data is available for the entire ETF history which began in 1993). Given their

differences in structure from traditional ETFs, exchange-traded notes and Holding Company Depository Receipts, commonly referred to as HOLDRS, are not included in my sample[21].

Table 4.1 reports the number of ETF liquidations by year during the sample[22]. Panel A shows liquidations based on their broad category. The overwhelming majority of liquidations are equity funds, with 95% of all liquidations. Although bond funds account for over ten percent of ETFs, only 12 bond funds die during the sample. Next, Panel B shows liquidations by Lipper category for categories with three or more liquidations during the sample. Growth funds have the largest number of liquidations with 43. Overall, there are 239 liquidations in my sample with a significant spike in the number of liquidations beginning in 2008. The year 2012 has the highest number of liquidations in any given year.

Table 4.2 presents summary statistics of the median values of fund size, age, returns, flows, expenses, bid-ask spread, premium-to-NAV, and volume. Table 4.2 shows median values for the full sample of living and liquidated funds as well as at specific time intervals prior to fund closure for liquidated funds. The *p*-values for the difference between the living fund sample and the liquidated fund samples are reported in parentheses. Liquidated funds are significantly smaller and younger than the sample of living funds. While the median fund size is over $116 million for living funds, it is only $5.6 million for liquidated funds. Next, both quarter and year raw returns as well as returns in excess of the Lipper category average are significantly lower for liquidated funds than living funds. For example, the median year return is 4.7% for living funds but a negative 3.5% for liquidated funds.

[21] ETNs are technically a debt security with a maturity date and backing based on the credit of the issuer. HOLDRS have several differences from ETFs. For example, HOLDRS allow investor retainer of voting rights of stock holdings, require round lots when trading, have greater concentration of holdings and fewer typical holdings, and have a fixed custody fee for each round lot. Additionally, once HOLDRS are started, portfolio composition has weightings based on the original shares at inception and hence the value weightings can change substantially.
[22] Years without any ETF liquidations are not reported in the table, for brevity.

Next, I investigate differences in flows. Following the mutual fund literature (Sirri and Tufano, 1998; Nanda, Wang, and Zheng, 2004) on fund flows, I define dollar flows as:

$$Dollar\ Flow_{i,t} = TNA_{i,t} - TNA_{i,t-1} * \left(1 + R_{i,t}\right), \tag{20}$$

and flow growth as:

$$Flow_{i,t} = Dollar\ Flow_{i,t}/TNA_{i,t-1}, \tag{21}$$

where $TNA_{i,t}$ is the total net assets for fund *i* in period *t* and $R_{i,t}$ is the one-period fund return for fund *i* in period *t*. Flows for liquidated funds are significantly lower than those for living funds. While living funds grow by a median of over 21% each year, liquidated funds shrink in size by over one percent on median in their final year.

The bid-ask spread for liquidated funds is nearly double that for the median living fund. Given the smaller fund size and smaller amount of trading that occurs for funds that are liquidated, it is not surprising to see a larger spread for liquidated funds. This, along with higher expense ratios, results in higher costs both to trade and hold funds that liquidate in the future. The spread measure is calculated as the average closing bid-ask spread over a given time period.

Liquidated funds typically trade at a discount while living funds tend to trade at a slight premium. The premium/discount is calculated by taking the share price in the CRSP daily stock file and subtracting the NAV in the CRSP mutual fund daily returns file. This difference is then divided by the NAV to show the percentage deviation from the underlying value of the assets at which the fund trades. Finally, funds that liquidate have median trading volumes that are significantly lower than those for funds that continue to operate. The median liquidated fund trades over 19 times the dollar trading volume and over 13 times the number of shares of a fund that liquidates in a quarter.

In Figures 4.1 and 4.2, I look for trends in liquidations based on fund age. Figure 4.1 shows the vast majority of funds that are liquidated are less than two years old. After the second year, there is a sharp downward decline in the number of funds liquidating. In Figure 4.2, I narrow the focus to funds that liquidate during the first three years and now look at age in months instead of years. Funds that are 10 months old have the highest death count with 30 funds. Seventy-two percent of all ETF deaths occur within the first two years and 89 percent of ETF deaths occur within the first three years. These figures illustrate the significance of a fund's initial years for its survival.

4.3 Results

4.3.1 Fund performance and characteristics

The summary statistics in Table 4.2 suggest important differences exist between living funds and those that liquidate. Next, I estimate logistic regressions to determine various factors' impact upon the fund's probability of liquidating within one quarter in Tables 4.3 and 4.4. The dependent variable is an indicator variable taking a value of one if the fund dies within one quarter, and zero otherwise. It is important to note that the fund family makes the ultimate decision as to if a fund stays open or closes. Since a family's decision criteria for closing down a fund are likely to be similar across all of the family's funds, I cluster by fund family. This assumes independence across fund families but not within fund families. Additionally, I include quarter dummies in each regression.

Table 4.3 looks at factors that are fund specific. Given the findings in Figure 4.1 that the majority of funds that die do so in the first two years, I separately estimate the subsample of funds age two or less ("young funds") in models 4, 5, and 6. Since flows and returns may be related (Warther, 1995; Ben-Rephael, Kandel, and Wohl, 2012), I include these in the logistic

regressions both separately and jointly. Fund size is the most influential liquidation determinant. The smaller a fund, measured by total net assets, the greater the likelihood is that it will liquidate. The sample of young funds also shows that ETFs with lower quarter flows and those trading at a discount are less likely to survive. These results are intuitive since the profitability of a family operating a fund is based upon the fund size.

Additional factors related to the fund family and Lipper category are added to the models in Table 4.4. Similar to Table 4.3, fund size, flows, and premium/discount are all significant determinants of liquidations. In the full sample, liquidation is more likely when a family's weighted excess return amongst funds is lower and when a family operates fewer funds, suggesting a family's size and reputation can influence a fund's survival. The sample of young funds shows an ETF is less likely to survive as competition increases within a category. This is logical given the increased pressure new funds cause. For example, Svetina (2010) finds a permanent five percent decline in shares outstanding after a new ETF enters into the same asset class.

4.3.2 Liquidations and the Fund Family

As noted previously, the decision to liquidate a fund is made by the fund family. It is possible that the family is influenced by factors beyond an individual fund when deciding to terminate a fund. For example, a family might decide to narrow their focus and only keep funds in a few categories. Or, another potential situation is a family could decide to exit the ETF business altogether and terminate all their funds at the same time. Anecdotal evidence of such actions comes from Russell Investments. On October 16, 2012, Russell liquidated all 25 of its passive ETFs, leaving only its actively managed ETFs open. Russell cited challenging equity market conditions as a factor leading to the decision.

In Figure 4.3, I look at each month that a fund family has one or more liquidations and compute the fraction of funds in the family that die. There are fifty times when fund families experience at least one fund death in a month. I find 13 times that a fund family decides to liquidate all of their exchange-traded funds. Most of the remaining family deaths encompass thirty percent or less of the number of funds in the family. In Figure 4.4, I calculate the number of closures per fund family within a month. The majority of the time a family will only liquidate a handful of funds in the same month. However, there are some times when a family will liquidate 15, 17, or even 25 funds in the same month.

With these factors in mind, I compare variables when a family liquidates half or more of its funds, a "mass family liquidation," with a "limited family liquidation" where the fund liquidated less than half of its funds. Table 4.5 analyzes fund family variable means and medians for mass versus limited family liquidations. The *p*-value for the difference in the two groups is also reported. Mass family liquidations tend to occur for families that have fewer funds, smaller family size, lower family market share, lower family flows, lower family returns, younger age, and serve fewer Lipper objective categories than families with limited family liquidations. Funds in more crowded Lipper categories are more likely to be part of a mass liquidation. Additionally, funds that are part of a limited liquidation tend to be older and have lower year returns than mass liquidated funds. These factors further emphasize the importance not only of an ETF's individual fund performance and characteristics for survival, but also the importance of the fund family's performance and influence in the market. The worse the family's overall performance and success, the greater the chance a fund will be one of many casualties in the family.

4.3.3 ETF Success Near Inception

The previous results indicate the importance of success early in a fund's life. As Figure 4.1 shows, if a fund is able to survive its infancy, there is a high probability it will continue to live for many years. Given this, I next seek to determine the factors of success within the first year of a fund's life.

Table 4.6 presents univariate analysis comparing funds that eventually die versus those that continue to live in the first year after inception. Funds that die tend to start and stay smaller, have higher expense ratios, and earn lower raw and excess returns in the first year. Funds that continue to live are much more successful at attracting higher dollar flows and growing the size of the fund.

Next, I investigate differences based on category variables. Funds that die tend to begin in more crowded categories that get even more crowded in the year after inception. This is true when analyzing competition from both ETFs and mutual funds. Liquidated funds have a tendency to be in a category that had high returns at the time of birth but subsequently performed poorly in the following year. This suggests unsuccessful funds try to ride a recent wave but enter the market too late. Surviving funds come from categories which grow at significantly higher rates than those for liquidating funds.

Finally, I analyze fund family characteristics in the opening year of funds. Liquidated funds are likely to originate from families that are significantly smaller than funds that live in terms of asset size, number of ETFs, and market share. Surviving funds have families that are typically able to attract significantly higher dollar flows and have larger quarter flow growth. Families with lower average weighted expense ratios are more likely to start funds that will survive. Family experience also is influential in producing funds that will last. Living funds are

more likely to come from older funds that entered the market an average of four years earlier than funds that die. Liquidated funds have a family that was the 20th family to enter the ETF market compared to 11th for living funds, on average. Funds from one of the top three most popular families- Vanguard, Black Rock, and State Street Global Advisors- have a much higher life expectancy.

Figure 4.5 provides further insight into the initial size difference for dead versus living funds noted in Table 4.6 to see if some funds are doomed to fail from inception based on small initial size. Dead funds are among the smallest funds started, however, there are numerous funds that live that start with similar sizes. This suggests that funds that begin with a larger size, for instance $7 million or more, have a much higher likelihood of surviving but a small inception size is not, in and of itself, an insurmountable obstacle. In untabulated results, I re-estimate the regressions in Tables 4.3 and 4.4 using the natural log of the fund's size at inception instead of lagged TNA. The results show the probability of fund liquidation is significantly higher as the inception size decreases. This illustrates the importance of a fund's size at inception for its survival.

4.3.4 Trading Strategy Upon Liquidation Announcement

Next, I look at ETFs which have announced an upcoming liquidation and seek to determine the best course of action for an investor holding the fund. When a family decides to liquidate a fund, it announces the final day of trading for the fund. The announcement normally comes between two weeks and one month prior to the final trading day. During this time, investors can sell their shares or they can wait until the fund's assets are liquidated and receive a cash payment in the amount of the net asset value (NAV). After the final day of trading, a fund will be delisted from the stock exchange and the remaining assets will be wound down over the

next few days. Once all the assets are sold, cash distributions are made to investors still holding shares of the fund.

There are several factors for investors to consider when deciding whether or not to sell their shares early or wait until the fund is liquidated. If investors wait for the fund to liquidate, they receive cash for the amount of NAV. Notably, this is not NAV on the final day of trading but rather the NAV resulting from when fund assets are sold in the day(s) following the close of trading. Brokers will generally wave the commission charged for redeeming shares and the ETF issuer normally bears any legal costs and other expenses associated with the fund's closure. The strategy of not selling may result in a better price if the spread is large and/or the fund is trading at a large discount from its net asset value. However, this may come at the cost of continuing to invest in a fund that may perform poorly and incur additional tax burdens in the form of dividends and capital gains distributions when the fund liquidates. A second trading strategy is to sell immediately after the liquidation announcement and invest in another fund with a similar investment style. This strategy incurs an additional trading commission[23], but can be beneficial if the new investment receives a higher return than the liquidating fund and it may avoid incurring some tax liabilities associated with the fund liquidation. One final trading strategy I consider is selling the ETF prior to liquidation but not immediately after the liquidation announcement. If there is a temporary negative abnormal return surrounding the liquidation announcement that is subsequently reversed, then this strategy may prove dominant.

I begin by presenting the descriptive statistics for funds following the liquidation announcement in Table 4.7. For the purposes of developing a trading strategy, I focus on traditional equity funds and hence exclude bond funds as well as funds in Long-Short Bias

[23] I assume that an investor will purchase a new fund to replace the liquidating ETF and hence the commission for purchasing the replacement fund(s) will be identical regardless of which trading strategy I propose.

Lipper categories. I search for press releases giving the announcement date, liquidation date, and other details using Factiva and the New York Stock Exchange's website. First, I calculate the holding period return if a fund is held from the date of the liquidation announcement until the fund is liquidated. The mean and median raw return are both positive at 0.323% and 0.95%, respectively. The excess return, calculated as the difference between the fund return and the value-weighted return on all other non-liquidating funds in the same Lipper category, is negative with an average of -0.775%. Notably, the time period between the announcement and final liquidation date varies, but the average is around 17 trading days. The average daily raw and excess returns are -0.11% and -0.17%, respectively. The average end of day spread is 34 cents (1.6% of the fund trading price) while the average liquidating fund trades 5.4 cents below its NAV or 0.219% below NAV. Last, I look at the trading volume in a fund's final days. On average, 0.026% of shares outstanding trade in a day.

To formally test for an optimal trading strategy, I use the four-factor model to estimate daily abnormal returns:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}RMRF_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}UMD_t \qquad (22)$$

where $R_{i,t} - R_{f,t}$ is the excess return for fund i over the risk-free rate on day t, and $RMRF_t$, SMB_t, and HML_t are the Fama and French (1993) factors and UMD_t is Carhart's (1997) momentum factor[24]. The results in Table 4.8 show a negative six basis point average abnormal return from holding an ETF following a liquidation announcement. While it is not possible to do this strategy with any single liquidating ETF for more than a month, normally, this still amounts to 1.00% in excess return, on average, if the fund is held for the average of 16.7 days between liquidation announcement and execution.

[24] Data for these factors are obtained from Ken French's website.

One potential concern is that an ETF will experience a significant negative price reaction following the liquidation announcement. If there is a temporary drop in price that is subsequently reversed, an investor looking to sell may find it optimal to temporarily delay trading. Alternatively, given the previous results, if there is no significant price reaction following the liquidation announcement then it may be optimal to sell immediately. I use event study methodology to examine these alternatives. As liquidation announcement press releases are nearly always made public after the close of trading, I consider the first trading day after the announcement date as day zero. Market model parameters are estimated over the period from day -210 to day -11 and include the Fama-French and momentum factors. Table 4.9 shows the average abnormal returns from day zero to day five in Panel A and the cumulative abnormal returns (CARs) in Panel B. The abnormal returns and the CARs immediately following the liquidation announcement are negative but statistically insignificant. For example, there is a negative 0.29% mean CAR from day 0 to day 1. The ten day CAR prior to the liquidation announcement is also negative at -0.62%. These results do not provide statistically significant evidence to suggest a sudden drop in ETF price following liquidation announcement.

On average, these results suggest that investors' are typically better off if they immediately sell a fund once it announces a liquidation, as opposed to waiting for a liquidation payment of net asset value. However, this may not hold if the transaction cost for selling the fund or the bid-ask spread is too high.

4.4 Conclusion

Although they once gained only minor attention, ETFs now are a major competitor for mutual funds and ETFs have grown substantially in both size and importance. With the sharp increase in the number of exchange-traded funds in the 2000's, so too did the number of

liquidations sharply increase. In fact, the year 2012 saw a record number of liquidations and nearly 17% of all ETFs ever created have been liquidated.

In this paper, I study what factors impact a fund family's choice to liquidate a fund. Small size and low flow growth prove catastrophic for funds, especially for young funds. Liquidated funds are more likely to have higher expense ratios and be younger than funds that survive. Additionally, I show that the fund family's overall standing is influential upon fund exit. Oftentimes multiple funds within a family will close at the same times. In fact, there are thirteen times when a fund family decides to completely exit the ETF business. This means that investors concerned about a fund terminating should consider not only the individual fund but also the fund family situation. Fund families that are smaller and entered the ETF market late are less likely to produce funds that last as opposed to larger, more established families. When comparing funds that live and die during their first year of life, the performance and success of a fund's overall category also is influential. Funds that die show a trend of entering into a crowded category that becomes increasingly crowded during the first year of life. Furthermore, funds that die have a tendency to be in a category that had performed well when the fund began, but then performs poorly during the first year of life.

Next, I analyze the best response for an investor who holds an ETF that has recently announced an upcoming liquidation. I find the average investor earns a significant negative abnormal return of six basis points daily for holding an ETF that is soon to liquidate. Also, I do not find a significant negative abnormal return surrounding the liquidation announcement. From this, I conclude the average investor benefits most from immediately selling their shares in the liquidating ETF after a liquidation announcement is made, assuming minimal transaction costs.

Overall, these results provide important factors for investors to watch when concerned with a fund's chances of liquidation. Liquidations can lead to additional tax burdens, trading costs, and search costs. Also, this paper provides insights for investors facing upcoming fund liquidations. As the ETF market continues to change, investors must continue to be aware as additional liquidations are likely in the future.

REFERENCES

Ben-Rephael, A., Kandel, S., & Wohl, A. (2012). Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics*, *104*(2), 363-382.

Brown, S. J., & Goetzmann, W. N. (1995). Performance persistence. *Journal of Finance*, *50*(2), 679-698.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57-82.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivor bias and mutual fund performance. *Review of Financial Studies*, *9*(4), 1097-1120.

English II, P. C., Demiralp, I., & Dukes, W. P. (2011). Mutual fund exit and mutual fund fees. *Journal of Law and Economics*, *54*(3), 723-749.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Nanda, V., Wang, Z. J., & Zheng, L. (2004). Family values and the star phenomenon: Strategies of mutual fund families. *Review of Financial Studies*, *17*(3), 667-698.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Finance*, *53*(5), 1589-1622.

Svetina, M. (2010). Exchange traded funds: Performance and competition. *Journal of Applied Finance*, *20*(2), 130-145.

Warther, V. A. (1995). Aggregate mutual fund flows and security returns. *Journal of Financial Economics*, *39*(2), 209-235.

Zhao, X. (2005). Exit decisions in the US mutual fund industry. *Journal of Business*, *78*(4), 1365-1402.

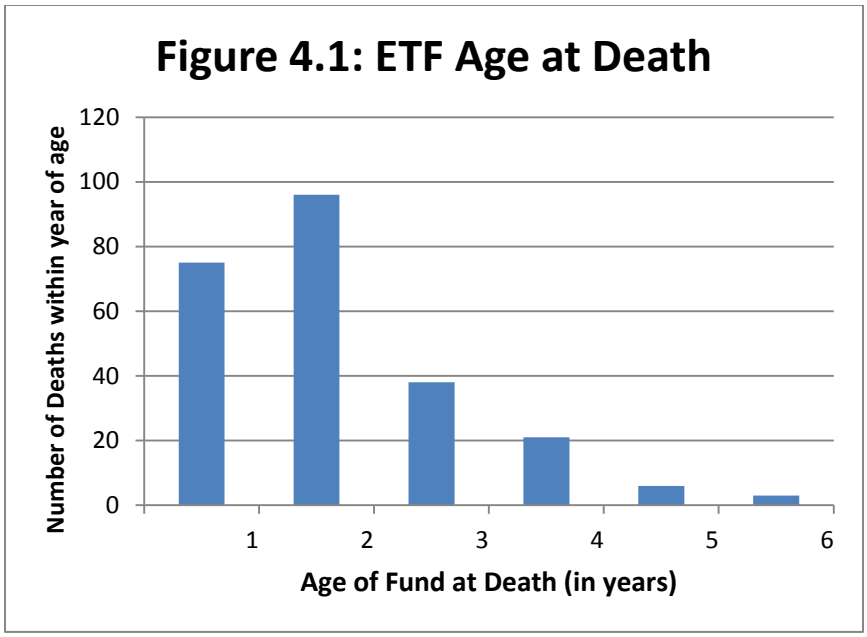
Figure 4.1: ETF Age at Death
Number of Deaths within year of age
120
100
80
60
40
20
0
1
2
3
4
5
6
Age of Fund at Death (in years)

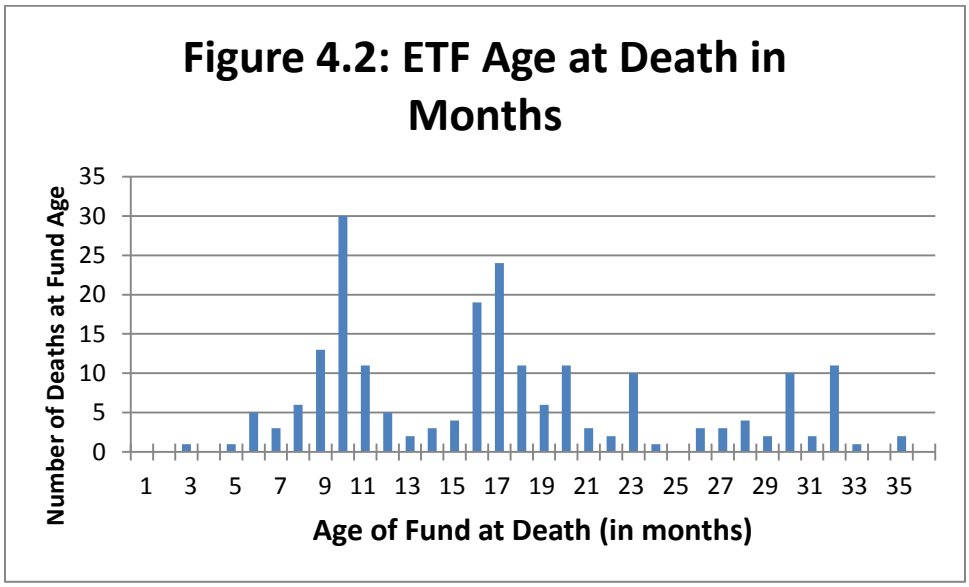
Figure 4.2: ETF Age at Death in Months
Number of Deaths at Fund Age
35
30
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35
Age of Fund at Death (in months)

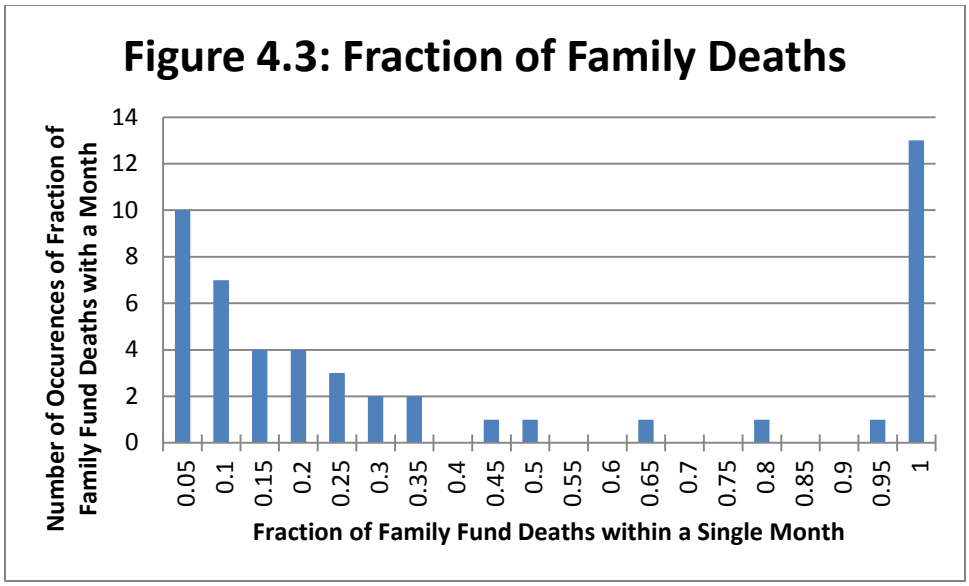
Figure 4.3: Fraction of Family Deaths
Number of Occurences of Fraction of Family Fund Deaths with a Month
14
12
10
8
6
4
2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
Fraction of Family Fund Deaths within a Single Month

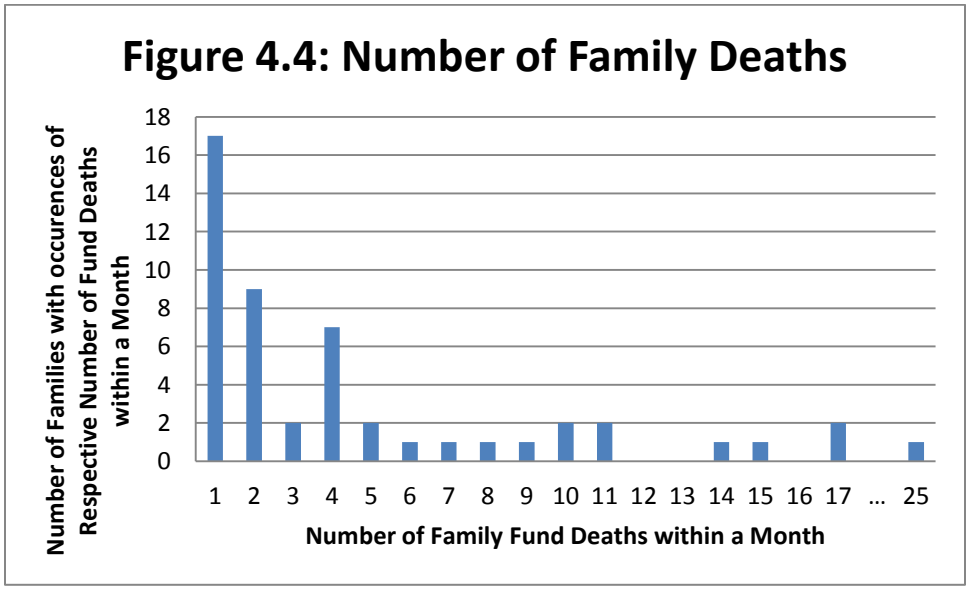
Figure 4.4: Number of Family Deaths
Number of Families with occurences of Respective Number of Fund Deaths within a Month
18
16
14
12
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 … 25
Number of Family Fund Deaths within a Month

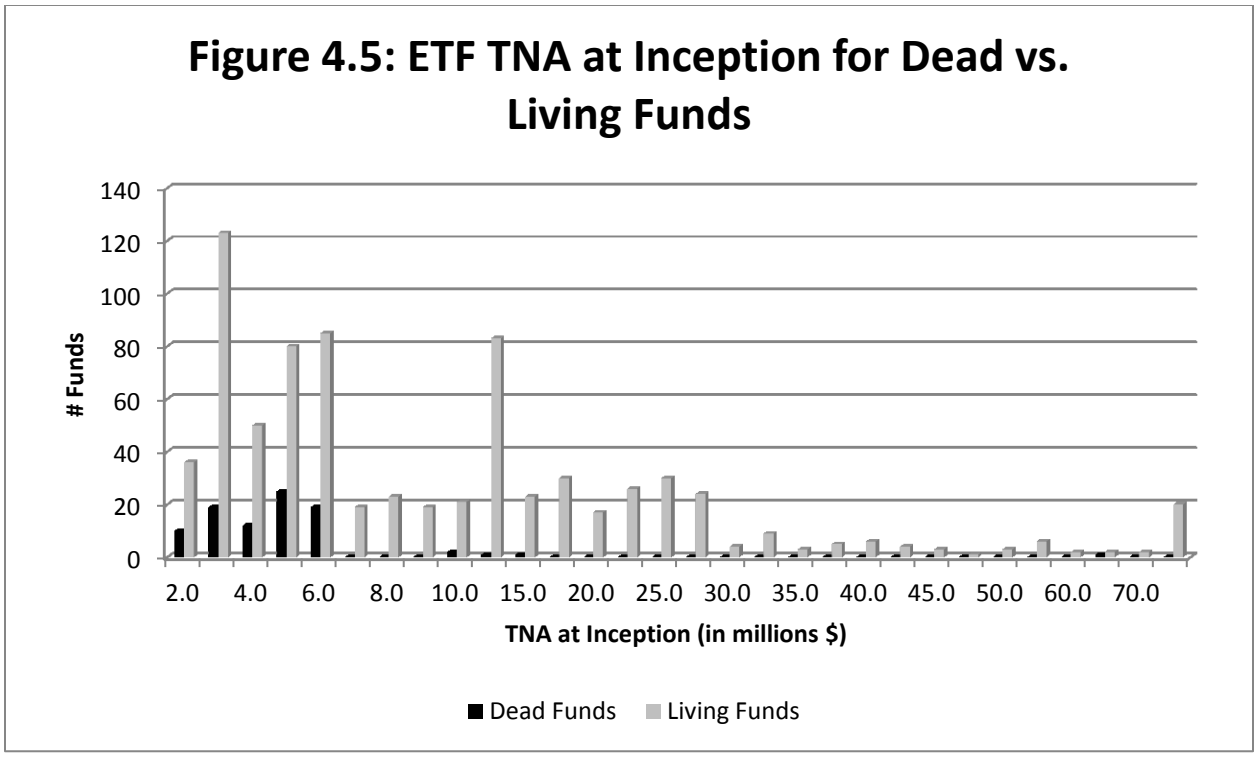
Figure 4.5: ETF TNA at Inception for Dead vs. Living Funds
140
120
100
80
60
40
20
0
Funds
2.0
4.0
6.0
8.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
50.0
60.0
70.0
TNA at Inception (in millions $)
Dead Funds
Living Funds

Table 4.1
ETF Liquidations by Year and Lipper Category

This table presents the number of liquidations occurring each year for each Lipper category. Years without any liquidation are omitted for brevity. Only categories with 3 or more total liquidation amongst all years are shown. The far right column shows the percent of all ETFs that are in the Lipper category as of September 30, 2012.

Panel A: Number of Liquidations in Year by Broad Asset Type Category

Broad Category	2002	2003	2006	2008	2009	2010	2011	2012	Total	Cat. % of All Liquidations	Cat. % of All Funds
All Equity Funds	3	2	1	44	47	46	14	70	227	95.0%	84.0%
All Tax-Exempt Bond Funds	0	0	0	0	0	0	0	1	1	0.4%	2.0%
All Taxable Bond Funds	0	0	0	6	0	2	1	2	11	4.6%	13.80%

Panel B: Number of Liquidations in Year by Lipper Category

Lipper Category	2002	2003	2006	2008	2009	2010	2011	2012	Total	Cat. % of All Liquidations	Cat. % of All Funds
Basic Materials	0	0	0	0	1	0	1	2	4	1.6%	1.9%
Canadian Funds	0	0	0	0	2	0	2	0	4	1.6%	1.3%
Commodities	0	0	0	1	1	2	0	2	6	2.5%	1.6%
Consumer Goods	0	0	0	0	1	0	0	2	3	1.2%	1.4%
Dedicated Short Bias	0	0	0	1	0	7	0	7	15	6.2%	6.0%
Emerging Markets	0	0	0	0	1	2	0	2	5	2.0%	4.4%
Equity Income	0	0	0	0	1	1	0	3	5	2.0%	1.3%
European Region	0	0	0	0	10	2	0	0	12	5.0%	2.8%
Financial Services	0	0	0	0	1	2	0	1	4	1.6%	2.2%
Global Financial	0	0	0	0	1	2	0	0	3	1.2%	1.0%
Global	0	0	0	1	1	1	0	1	4	1.6%	1.2%
Global Health/Biotechnology	0	0	0	11	0	1	0	0	12	5.0%	1.1%
Growth	0	0	1	3	6	3	8	22	43	17.9%	10.0%
Health/Biotech.	0	0	0	9	1	3	0	1	14	5.8%	2.5%
Industrials	0	0	0	0	1	3	2	3	9	3.7%	2.8%
International	1	0	0	0	1	3	0	5	10	4.1%	4.0%
Japanese	0	0	0	0	1	1	1	0	3	1.2%	0.9%
Mid-Cap	0	0	0	1	1	1	0	0	3	1.2%	2.7%
Natural Resources	0	0	0	1	1	0	0	1	3	1.2%	2.2%
Pacific Ex Japan	0	0	0	0	1	1	0	1	3	1.2%	1.4%
Pacific Region	0	0	0	0	3	0	0	0	3	1.2%	0.6%
Real Estate	0	0	0	7	0	0	0	1	8	3.3%	1.6%
Science & Technology	1	2	0	1	3	2	0	1	10	4.1%	2.6%
Short U.S. Treasury	0	0	0	2	0	1	0	0	3	1.2%	0.7%
Small-Cap	0	0	0	2	1	3	0	4	10	4.1%	3.3%
Specialty/Misc.	1	0	0	5	2	0	0	3	11	4.6%	1.7%
Telecommunication	0	0	0	0	0	2	0	1	3	1.2%	1.0%
Utility	0	0	0	0	1	1	0	1	3	1.2%	1.3%
All Other Lipper Categories	0	0	0	5	4	4	1	9	23	9.6%	34.5%
All Liquidations	3	2	1	50	47	48	15	73	239		16.9%

Table 4.2
Characteristics of Funds Near Liquidation

This table presents the median total net-assets (in millions of dollars), fund age (in months), raw return and returns in excess of the Lipper category average, flow growth, expense ratio, bid-ask spread, premium-to-net asset value, and trading volume . 'All Living Funds' refers to the median values for all funds currently active as of September 30, 2012. 'All Liq. Funds' refers to the median values for all funds that eventually liquidate, as of September 30, 2012. Other columns present the median values for liquidated funds at the specified number of months before liquidation. The *p*-values (in parentheses) represent the significance of the difference of the ETF in the months preceding the liquidation and the sample of all living funds using the Wilcoxon signed rank test.

	All Living Funds	All Liq. Funds	Liquidated Fund: Months Relative to Liquidation					
			-1	-3	-6	-9	-12	-24
TNA	116.20	5.60	4.20	4.30	4.80	5.20	5.40	8.75
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age	35	12	16	14	11	11	10	9
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Quarter Return	1.899%	-0.872%	-0.536%	-3.104%	-3.886%	1.342%	-2.939%	-0.333%
		(0.00)	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)
Year Return	4.719%	-3.466%	-5.803%	-6.440%	-3.592%	-5.571%	-6.498%	-2.480%
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Qtr. Excess Return	-0.003%	-0.512%	-0.761%	-0.927%	-0.737%	-0.181%	-0.931%	-0.778%
		(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.32)
Year Excess Return	-0.099%	-2.033%	-3.468%	-3.683%	-2.637%	-0.125%	-0.091%	-2.970%
		(0.00)	(0.00)	(0.00)	(0.22)	(0.22)	(0.64)	(0.01)
Quarter Flow	2.310%	-0.210%	-0.743%	-0.273%	-0.246%	-0.099%	0.018%	-0.134%
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)
Year Flow	21.312%	-1.119%	-1.118%	-1.456%	-0.406%	-1.164%	-2.704%	18.217%
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.23)
Expense Ratio	0.500%	0.700%	0.700%	0.700%	0.710%	0.750%	0.710%	0.650%
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Qtr. Spread	0.064	0.121	0.156	0.140	0.149	0.123	0.119	0.116
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Qtr. Premium-to-NAV	0.010%	-0.006%	-0.048%	-0.022%	0.006%	-0.006%	0.005%	0.001%
		(0.00)	(0.00)	(0.00)	(0.02)	(0.02)	(0.37)	(0.37)
Qtr. Share Volume	43,880	3,195	1,721	2,317	2,443	3,270	2,460	4,672
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Qtr. $ Trading Volume	1,624,041	84,986	42,540	49,193	60,607	82,570	67,901	153,019
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 4.3
Determinants of ETF Liquidations, Fund Variables

This table presents the logistic regression results for the determinants of fund liquidations. The dependent variable takes a value of one if the fund is liquidated within one quarter of the observation and zero otherwise. Time intervals are in quarters. Quarter dummies are included in all model specifications. Standard errors are clustered by fund family.

	All Funds			Funds Age 2 or Less		
Variable	1	2	3	4	5	6
LogTNA (t-1)	-1.216***	-1.239***	-1.217***	-0.976***	-1.058***	-0.976***
	(-9.46)	(-9.63)	(-9.51)	(-4.16)	(-4.29)	(-4.11)
Age	-0.017	-0.013	-0.017	-0.044	-0.019	-0.044
	(-0.85)	(-0.74)	(-0.85)	(-0.72)	(-0.31)	(-0.71)
Exp Ratio (t-1)	0.312	0.321	0.289	0.395	0.415	0.380
	(0.58)	(0.62)	(0.55)	(0.68)	(0.73)	(0.66)
Qtr. Spread (t-1)	-24.061	-28.275	-24.272	9.243	10.100	9.512
	(-1.01)	(-1.10)	(-1.03)	(0.34)	(0.37)	(0.35)
Qtr. Prem. (t-1)	0.046	-0.257	-0.117	-80.781*	-92.398**	-81.136*
	(0.01)	(-0.05)	(-0.02)	(-1.73)	(-2.09)	(-1.75)
6 mo. St. Dev. (t-1)	-7.515	-8.402	-7.674	-3.263	-3.420	-3.331
	(-0.48)	(-0.55)	(-0.48)	(-0.42)	(-0.44)	(-0.43)
Qtr Flow (t-1)	-0.218		-0.201	-0.764**		-0.738**
	(-0.61)		(-0.53)	(-2.06)		(-2.05)
Qtr Flow (t-2)	-0.355		-0.377	-0.518*		-0.532*
	(-1.23)		(-1.21)	(-1.75)		(-1.68)
Qtr Ret (t-1)		-1.301	-1.216		-1.252	-1.082
		(-0.83)	(-0.73)		(-0.50)	(-0.45)
Qtr Ret (t-2)		1.586	1.643		1.208	1.088
		(0.86)	(0.85)		(0.58)	(0.55)
Intercept	-0.967	-1.120	-0.973	-0.097	-0.662	-0.131
	(-0.63)	(-0.73)	(-0.62)	(-0.04)	(-0.28)	(-0.05)
No. Obs	33158	33675	33152	5851	6044	5851
R^2	0.337	0.334	0.339	0.281	0.271	0.282

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 4.4

Determinants of ETF Liquidations: Fund, Family, & Category Variables

This table presents the logistic regression results for the determinants of fund liquidations for funds two years of age or younger. The dependent variable takes a value of one if the fund is liquidated within one quarter of the observation and zero otherwise. Time intervals are in quarters. Quarter dummies are included in all model specifications. Standard errors are clustered by fund family.

	All Funds			Funds Age 2 or Less		
Variable	1	2	3	4	5	6
LogTNA (t-1)	-1.256***	-1.278***	-1.253***	-1.021***	-1.068***	-0.984***
	(-21.93)	(-23.10)	(-21.71)	(-4.43)	(-4.62)	(-4.22)
Age	-0.011**	-0.007	-0.011**	0.016	0.026	0.007
	(-2.21)	(-1.60)	(-2.12)	(0.23)	(0.36)	(0.10)
Exp Ratio (t-1)	0.026	-0.046	-0.031	-0.045	-0.100	-0.090
	(0.18)	(-0.34)	(-0.23)	(-0.09)	(-0.21)	(-0.20)
Qtr. Spread (t-1)	-8.328	-11.058	-8.443	18.742	15.169	18.312
	(-0.90)	(-1.22)	(-0.91)	(0.54)	(0.45)	(0.55)
Qtr. Prem. (t-1)	-5.902	-12.181	-5.709	-79.852**	-92.188**	-77.474**
	(-0.40)	(-0.81)	(-0.38)	(-2.00)	(-2.35)	(-1.99)
6 mo. St. Dev. (t-1)	0.762	0.355	0.511	-4.938	-4.570	-4.650
	(0.39)	(0.18)	(0.26)	(-0.71)	(-0.72)	(-0.72)
Qtr Flow (t-1)	-0.255		-0.183	-0.797*		-0.728*
	(-1.32)		(-0.93)	(-1.88)		(-1.78)
Qtr Flow (t-2)	-0.386**		-0.439***	-0.589**		-0.695**
	(-2.34)		(-2.59)	(-2.03)		(-2.39)
Qtr Excess Ret (t-1)		-1.313	-1.019		-0.762	-0.643
		(-1.38)	(-1.05)		(-0.24)	(-0.22)
Qtr Excess Ret (t-2)		0.877	1.128		-0.675	-0.440
		(0.97)	(1.22)		(-0.40)	(-0.28)
Cat. Qtr Flow (t-1)	-0.585		-0.601	-0.398		-0.522
	(-1.57)		(-1.56)	(-0.57)		(-0.71)
Cat. Qtr Flow (t-2)	-0.264		-0.339	-0.972		-1.000
	(-1.07)		(-1.31)	(-1.54)		(-1.36)
Cat. Qtr Ret (t-1)		0.097	-0.154		1.163	0.530
		(0.17)	(-0.26)		(0.51)	(0.23)
Cat. Qtr Ret (t-2)		-0.574	-0.711		-1.343	-1.826
		(-1.13)	(-1.33)		(-0.97)	(-1.26)
Fam. Qtr Ex. Ret (t-1)		-4.648**	-4.835**		-5.449	-4.730
		(-2.02)	(-1.99)		(-0.64)	(-0.55)
Fam. Qtr Ex. Ret (t-2)		1.316	1.977		5.197	7.010
		(0.56)	(0.81)		(0.64)	(0.89)
Category Number ETFs	0.002	0.002	0.001	0.010*	0.012**	0.010**
	(0.79)	(1.06)	(0.57)	(1.95)	(2.32)	(1.96)
Family Number ETFs	-0.012***	-0.012***	-0.012***	-0.036	-0.035	-0.035
	(-8.50)	(-8.59)	(-8.48)	(-1.61)	(-1.43)	(-1.50)
Intercept	-0.122	-0.335	-0.148	0.582	0.025	0.588
	(-0.35)	(-0.94)	(-0.40)	(0.29)	(0.01)	(0.29)
No. Obs	33158	33488	33152	5851	6002	5851
R2	0.364	0.360	0.366	0.393	0.386	0.400

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 4.5
Determinants of ETF Liquidations Within Family and Objective Category Variables

This table presents the univariate statistics for funds one month prior to liquidation. Funds are divided into two groups: those that are liquidated in the same month along with at least half of all ETFs in the fund family, "Mass Family Liquidation," or those that are liquidated in the same month along with less than half of all ETFs in the fund family, "Limited Family Liquidation."

	Mean Values			Median Values		
Variable	Limited Family Liquidation	Mass Family Liquidation	*p*-value	Limited Family Liquidation	Mass Family Liquidation	*p*-value
Fund Variables						
TNA (in millions $)	5.747	5.494	0.81	3.800	4.350	0.86
Age (in mnths)	26.406	13.081	**0.00**	27.000	16.000	**0.00**
Expense Ratio	0.007	0.008	0.27	0.007	0.005	**0.03**
Flow Qtr	-0.068	-0.039	0.37	-0.011	-0.008	0.12
Flow Yr	0.019	0.022	0.98	-0.042	-0.011	0.21
Return Qtr	0.001	-0.013	0.50	0.014	-0.011	0.41
Return Yr	-0.181	0.036	**0.00**	-0.201	0.052	**0.00**
Excess Return Qtr	0.004	-0.003	0.54	-0.004	-0.006	0.75
Excess Return Yr	-0.039	-0.031	0.78	-0.021	-0.027	0.63
Return Standard Deviation (6 mo.)	0.070	0.068	0.85	0.071	0.036	0.29
Bid-Ask Spread (Qtr)	0.226	0.427	**0.00**	0.116	0.241	**0.00**
Premium to NAV (Qtr)	-0.029	-0.018	0.51	-0.013	-0.013	0.91
Fund Family Variables						
# ETFs in Family	59.4	16.7	**0.00**	42.0	17.0	**0.00**
TNA (in millions $)	16283	140	**0.00**	6010	184	**0.00**
ETF Market Share	0.032	0.000	**0.00**	0.007	0.000	**0.00**
Age (in mnths)	70.820	22.892	**0.00**	46.000	19.000	**0.00**
Expense Ratio	0.007	0.007	0.38	0.006	0.005	**0.00**
Flow Qtr	0.038	-0.026	**0.00**	0.015	0.000	**0.00**
Flow Yr	0.353	-0.029	**0.00**	0.090	0.015	**0.00**
Return Qtr	0.037	0.003	**0.00**	0.037	-0.002	**0.00**
Return Yr	0.268	0.084	**0.00**	0.015	0.106	0.21
Excess Return Qtr	0.008	0.005	0.47	-0.001	-0.001	0.31
Excess Return Yr	0.063	-0.017	**0.00**	0.023	-0.012	**0.00**
# Objectives	22.578	5.396	**0.00**	17.000	5.000	**0.00**
Category Variables						
# ETFs in Category	32.1	45.2	**0.00**	22.0	29.0	**0.02**
TNA (in millions $)	19441	44758	**0.00**	8066	13740	**0.00**
Flow Qtr	0.105	0.071	0.25	0.046	0.032	0.77
Flow Yr	0.487	0.463	0.92	0.262	0.227	0.20
Return Qtr	0.011	-0.010	0.12	0.028	0.002	**0.07**
Return Yr	-0.024	-0.013	0.75	-0.020	0.021	0.36

Table 4.6
Univariate Statistics Across the Samples of Living and Dead Funds

This table presents the mean values for each variable for ETFs one, six, and twelve months after inception, respectively. The averages for funds currently active funds as of September 30, 2012 are compared with those that are no longer active. The *p*-values represent the *t*-test significance of the difference in means of living versus dead funds assuming unequal variance. Significance at the 10% (1%) level are in bold (underlined).

	Inception			Age 6 months			Age 12 months		
	Living	Dead	*p*-value	Living	Dead	*p*-value	Living	Dead	*p*-value
TNA (millions $)	16.99	4.78	**0.00**	90.89	8.45	**0.00**	165.15	11.86	**0.00**
Qtr Ret				0.008	-0.061	**0.00**	0.009	-0.016	0.11
Yr Ret							0.022	-0.076	**0.00**
Excess Qtr. Ret.				-0.001	-0.014	**0.02**	-0.004	-0.006	0.79
Excess Yr. Ret.							-0.008	-0.047	**0.00**
Qtr. Flow %				0.524	0.107	**0.00**	0.292	0.087	**0.00**
Yr. Flow %							4.546	0.766	**0.00**
Qtr. Flow $				30.63	0.51	**0.00**	37.33	0.54	**0.00**
Yr. Flow $							141.21	3.57	**0.00**
Expense Ratio	0.0058	0.0079	**0.00**	0.0057	0.0079	**0.00**	0.0057	0.0080	**0.00**
Spread Qtr. Ave.				0.270	0.465	**0.10**	0.299	0.317	0.81
Premium Qtr. Ave				0.095	0.024	0.36	0.136	-0.002	0.16
Qtr St. Dev.				0.066	0.070	0.43	0.062	0.072	0.15
Cat. TNA (millions $)	16,673	25,679	**0.00**	18,401	27,553	**0.00**	18,851	30,895	**0.00**
Cat. Qtr Ret.	0.013	0.024	**0.03**	0.009	-0.046	**0.00**	0.013	-0.010	**0.04**
Cat. Yr. Ret.	0.091	0.118	0.10	0.037	-0.009	**0.01**	0.028	-0.029	**0.00**
Cat. Qtr Flow %	1.135	0.107	0.27	0.708	0.075	**0.09**	1.333	0.119	0.14
Cat. Yr. Flow %	2.045	1.001	**0.09**	5.607	0.626	**0.05**	8.351	1.387	**0.02**
Cat. Qtr Flow $	637	488	0.22	869	516	**0.02**	587	644	0.82
Cat. Yr. Flow $	2,208	2,405	0.57	2,942	2,447	0.15	2,637	2,653	0.97
Cat. # of ETFs	16.35	26.92	**0.00**	21.02	35.33	**0.00**	22.69	39.32	**0.00**
Cat. # of Mutual Funds	512.71	784.29	**0.00**	526.79	873.41	**0.00**	548.07	947.62	**0.00**
Fam. TNA (millions $)	64,934	3,924	**0.00**	74,472	5,465	**0.00**	68,031	8,369	**0.00**
Fam. Qtr. Ret.	0.057	0.015	0.31	0.075	-0.012	**0.05**	0.080	0.022	0.24
Fam. Yr. Ret.	0.468	0.311	0.53	0.273	0.514	**0.07**	0.130	0.111	0.65
Fam Ex. Qtr Ret	0.012	-0.003	0.25	0.015	0.001	0.29	0.013	-0.004	0.23
Fam Ex. Yr. Ret	0.057	0.057	0.99	0.051	0.195	**0.01**	-0.002	0.024	0.10
Fam. Qtr. Flow %	0.104	0.050	**0.00**	0.133	0.091	**0.00**	0.100	0.049	**0.00**
Fam. Yr. Flow %	0.370	0.435	0.29	0.561	0.531	0.69	0.394	0.359	0.45
Fam. Qtr. Flow $	2,718	226	**0.00**	2,568	329	**0.00**	2,283	247	**0.00**
Fam. Yr. Flow $	8,428	657	**0.00**	10,357	1,428	**0.00**	9,685	1,107	**0.00**
Fam. Ave. Exp. Ratio	0.0059	0.0063	**0.02**	0.0059	0.0066	**0.00**	0.0057	0.0073	**0.00**
Fam. Age (Months)	71.79	24.20	**0.00**	80.00	35.49	**0.00**	81.31	48.80	**0.00**
Fam. Age Rank	11.43	20.35	**0.00**	11.37	20.16	**0.00**	11.38	17.74	**0.00**
Fam. Birth <2007	0.82	0.38	**0.00**	0.82	0.40	**0.00**	0.82	0.49	**0.00**
Fam. One of 1st 5 Started	0.48	0.16	**0.00**	0.48	0.17	**0.00**	0.48	0.23	**0.00**
Fam. One of 1st 10 Started	0.59	0.22	**0.00**	0.59	0.23	**0.00**	0.59	0.32	**0.00**
Fam. # of ETFs	54.48	17.30	**0.00**	67.58	31.63	**0.00**	67.40	40.42	**0.00**
Fam. # of Objectives	17.86	6.64	**0.00**	21.39	11.13	**0.00**	21.79	14.54	**0.00**
Top 3 Family	0.37	0.03	**0.00**	0.37	0.03	**0.00**	0.36	0.04	**0.00**
Fam Mkt. Sh.	0.11	0.01	**0.00**	0.13	0.01	**0.00**	0.13	0.02	**0.00**

Table 4.7
Characteristics of Funds Following Liquidation Announcement

This table presents descriptive statistics for ETFs from the time of their liquidation is announced until they liquidate. The first two variables give the return for an investor holding the fund from the liquidation announcement date until the final liquidation payment.

	Mean	Median	1st Quartile	3rd Quartile
Post-announcement holding period return for ETF	0.323%	0.950%	-2.764%	5.036%
Post-announcement holding period return for ETF minus Lipper category	-0.775%	-0.245%	-3.028%	1.679%
Average daily raw return following announcement	-0.11%	0.06%	-0.20%	0.32%
Average daily excess return following announcement	-0.17%	-0.02%	-0.17%	0.09%
Spread ($)	0.340	0.140	0.060	0.310
Spread/Price	1.557%	0.599%	0.261%	1.619%
Premium ($)	-0.081	-0.056	-0.176	0.020
Premium/NAV	-0.361%	-0.235%	-0.741%	0.001%
% Volume of shares	0.026%	0.006%	0.001%	0.023%
Trading days between announcement and final trade date	16.7	11.0	8.0	19.0

Table 4.8

ETF Liquidation-Related Trading Profits

This table reports the risk-adjusted returns for ETFs between the date of a liquidation announcement and the final trading day. Results are presented using both Fama and French's (1993) 3-factor model and Carhart's (1997) 4-factor model. The sample excludes non-equity funds and those in the long-short bias Lipper category.

Variable	3-factor	4-factor
Alpha	-0.0005 *	-0.0006 *
	(-1.67)	(-1.75)
Excess Return	0.815 ***	0.807 ***
	(32.74)	(31.84)
Small-Minus-Big	0.251 ***	0.251 ***
	(4.84)	(4.84)
High-Minus-Low	0.036	-0.022
	(0.82)	(-0.39)
Momentum		-0.054 *
		(-1.72)
Observations	2,931	2,931
Adj. R^2	39.5%	39.5%

*,**,*** represent significance at the 10%, 5%, and 1% level, respectively.

Table 4.9
Cumulative Abnormal Returns Near Liquidation Announcement

This table presents mean abnormal returns and cumulative abnormal returns (CARs) for liquidating ETFs. Day 0 is the first trading day after the announcement of the liquidation. Market model parameters are estimated over the period from day -210 to day -11. Abnormal returns are estimated using the Carhart four-factor model using an equally weighted index.

Panel A: Abnormal Returns

Day	Mean Abnormal Return	t
0	-0.13%	-0.67
1	-0.16%	-0.87
2	-0.04%	-0.23
3	-0.14%	-0.77
4	0.25%	1.31
5	-0.10%	-0.55

Panel B: Cumulative Abnormal Returns

Period	Mean CAR	t
[0,1]	-0.29%	-1.09
[2,4]	0.05%	0.15
[0,5]	-0.34%	0.43

CHAPTER 5

CONCLUSION

This dissertation studies three important areas related to exchange-traded funds. The first essay analyzes the relationship between sector ETF ownership with both return comovement and the reaction to new information on firm fundamentals. The results demonstrate a strong increase in return comovement with other industry stocks, especially those stocks held by sector ETFs, as sector ETF ownership increases. Furthermore, sector ETF ownership is shown to be strongly related to a dampened initial reaction to earnings announcement surprises. The second essay documents ETF performance persistence that is primarily attributable to a fund's industry exposure. The presence of performance persistence for passively managed funds suggests that caution should be used when concluding performance persistence demonstrates manager skill amongst actively managed mutual funds, as multiple past studies do. The third essays documents factors related to ETF liquidations. Funds with small sizes and low flows as well as those from struggling fund families who are latecomers to the ETF industry are less likely to survive. Also, funds that enter the market after a popular category peaks have a greater probability of liquidating. Finally, the average investor holding a fund with an upcoming liquidation is best served to immediately sell the liquidating fund and purchase other funds in the same category when transaction costs are of minimal concern.