The Problem with Passivity

How ETFs Affect the Volatility of Underlying Securities

Keaton Sullivan
Northwestern University
Mathematical Methods in the Social Sciences, Senior Thesis 2018
Advisor: Dimitris Papanikolaou
# Table of Contents

1. Introduction  
   pg. 3

2. Literature Review  
   pg. 11

3. Data & Methodology  
   pg. 17

4. Analysis  
   pg. 24

5. Results  
   pg. 29

6. Conclusion  
   pg. 36

7. References  
   pg. 38
Introduction

The creation of the exchange-traded fund, known colloquially as an “ETF”, was one of the most popular and impactful financial innovations of the past few decades. An ETF is an investment vehicle representing a basket of underlying securities, ranging from stocks to bonds to derivatives. Shares of an ETF trade on stock exchanges just as traditional stocks do. Buying a share of an ETF is equivalent to buying fractional shares of all of its underlying securities. In turn, ETF managers charge an annual expense ratio on the funds invested as a management fee. Most ETF expense ratios are under 1.00%, and many of the largest ETFs charging less than 0.10%. ETFs replicating more exotic strategies usually charge a higher premium.

ETFs are similar to mutual funds, an older type of investment vehicle. Both ETFs and mutual funds offer access to a targeted basket of diversified underlying securities. Yet, ETFs have exploded in popularity in recent years while mutual funds have largely stagnated. This divergence can be credited to the differences between the two. First, mutual funds often have minimum initial investments, a number that is disconnected from the intrinsic value of an investor’s share in the basket of securities. For ETFs, the minimum investment is simply the share price, often much lower than mutual fund investment minimums. Furthermore, ETFs offer different liquidity profiles than mutual funds. While mutual funds trade once a day at the end of the day, ETFs can be traded intraday like stocks since their shares are listed on various stock exchanges (“ETFs vs. Mutual Funds: A Comparison”).
The first ETF was launched in Canada in 1990. An ETF was not listed on a US stock exchange until 1993, and even then, it was used primarily by institutional investors looking to execute complicated trades. Yet, the growth of ETFs has been exponential since 1993. As of September 2017, ETFs managed $4.4 trillion in assets worldwide, with over half of those assets in US-listed ETFs. Funds managed by ETFs have grown at an unprecedented 21% compounded annual growth rate since 2005 (“What Is the History of ETFs?”). The number of different ETFs has also grown tremendously. In 2010, there were roughly 250 ETFs; by 2012, there were nearly 1,000; and by 2016, there were over 5,000 different ETFs listed on US stock exchanges. Conversely, the number of actual stocks listed on US stock markets has decreased. After peaking in 1997 with 7,487 different companies listed on US stock exchanges, as of 2016, there were only 4,342 (“There Are Now More Indexes Than Stocks”). Furthermore, the market for ETF providers is very concentrated. BlackRock’s iShares platform, State Street’s SPDR platform, and Vanguard account control nearly 80% of all ETFs assets in the market.

The earliest iterations of ETFs were designed to track pre-existing equity indices such as the Standard and Poor’s 500 index, Russell 2000 index, Nasdaq index, and Dow Jones Industrial Average index, amongst numerous others. For example, Boston asset manager State Street Global Advisors released the SPDR S&P 500 ETF, known by the ticker SPY, in January of 1993. Today, SPY is the largest ETF in the world, with $279 billion in assets under management as of January 2018. SPY has an average daily trading volume of 65.4 million shares which dwarfs Apple, the most widely traded stock in the US, which has an average daily trading volume of 28.6 million shares.
Furthermore, SPY frequently trades at a bid-ask spread (the difference between the market price a seller is willing to sell and the market price a buyer is willing to buy) within one cent, which is lower on average than the bid-ask spread of all of its constituent stocks (Narayanan).

Today, ETFs have evolved into a broader form of investment vehicle. There exist ETFs that track the stock market of nearly any country or region, ETFs that track specific sectors like technology, ETFs that track fixed-income products, ETFs that track commodities such as gold or oil, ETFs that invest according factors such as low-volatility or momentum, and many others. While most ETFs are passively managed, buying and selling according to fixed rules, some ETFs are actively managed, with fund managers selectively buying and selling stocks to target above-market returns or conform to a specific investment strategy. The advent of leveraged ETFs brought a new breed of ETF to the market. Leveraged ETFs use any combination of derivatives, debt, and traditional underlying securities to produce returns that are linearly related to an underlying benchmark. For example, Direxion Investments offers two ETFs, LABU and LABD, that give investors exposure to biotechnology stocks in the S&P 500 at a leverage factor of 300% and -300%, respectively. Thus, if the basket of underlying stocks gains 1%, LABU will gain 3% and LABD will fall 3%. Inverse leveraged ETFs, such as LABD, have created a simple, cost-effective way for retail investors to “short” the market, or take a position that profits when the value of the market declines. In the past, this strategy was reserved for sophisticated investors with the ability to borrow and lend shares.
ETFs fall within the broader category of exchange-traded products. While exchange-traded funds (ETFs) are by far the most popular, other products such as exchange-traded vehicles (ETVs) and exchange-traded notes/certificates (ETNs) also exist. The difference between these vehicles lies in the way the funds receive capital, invest it, and pay it out to investors.

Since ETFs are traded on stock exchanges, their prices are determined by supply and demand, rather than the value of the slice of underlying securities it represents, called the Net Asset Value ("NAV") of an ETF. The creation and redemption mechanism is what keeps the market price and NAV of an ETF in sync. The two main participants in the creation/redemption mechanism are the ETF issuer and an authorized participant ("AP"), usually a large financial institution. When the NAV of an ETF portfolio falls below its market value, an AP purchases underlying securities on the open market according to weights set by the ETF. The AP then delivers the securities to the ETF issuer in exchange for shares of the ETF equal to the same amount when valuing the ETF based on its NAV. The AP then sells the ETF shares on the open market at the market price, making a profit on the difference between the NAV and market price. Since the supply of shares increases, the market value of the ETF will eventually fall in line with the ETF’s NAV. When NAV exceeds the ETF’s market price, the AP does the same process in reverse, buying ETF shares on the open market and trading them to the ETF issuer for shares of the underlying securities. This process ensures that both parties are incentivized to transact in a manner that keeps the ETF’s market price in line with its NAV ("What Is The Creation/Redemption Mechanism?").
The explosive growth in the ETF market is not without good reason. Rather than building a diversified portfolio manually, ETFs offer an easy way to hold a diversified portfolio while technically owning shares in only one publicly traded stock ticker. ETFs also allow investors to target geographic markets or investment factors. Rather than sifting through all publicly traded stocks for those that qualify as value stocks, one can simply purchase an ETF that tracks what it deems to be value stocks in a particular index. Similarly, if an investor wants to invest in Chinese equity markets, they can simply purchase an index that tracks Chinese stock markets rather than buying individual Chinese stocks on their own. As more and more companies flock to offer their own ETFs, as well as create ETFs that invest according to new strategies, expense ratios have plummeted. Hence, ETFs are not only convenient, but also cost-effective. Coupled with the intraday liquidity that ETFs provide, ETFs have increasingly become a mainstay in the portfolios of institutional and retail investors alike.

Although ETFs appear to be a purely positive addition to the array of investment vehicles available, they may have unintended consequences on the underlying securities they own. As the assets managed by ETFs increases year after year, ETFs are forced to purchase more and more of their underlying securities according to the fixed rules in their prospectuses. Today, top ETF providers such as Vanguard, BlackRock, and State Street are the top shareholders for a majority of publicly traded stocks. As the proportion of stocks’ market capitalization owned by passive ETFs grows, less capital is allocated towards investments hoping to profit on potential mispricings in the market. With less money actively seeking to profit off of mispricings, the market as a
whole becomes less efficient. In turn, ETFs may create a multitude of problems such as co-movement in stocks, distorted volatility, or even flash crashes.

In this paper, I will explore the effect that traditional, long-only, equity ETFs have had on the volatility of their underlying securities. First, I study the relationship between the two by conducting a fixed-effects OLS regression of monthly volatility on ETF ownership for all stocks in the S&P 500 index, on a monthly basis, from 2010-2017. The regression indicates that ETF ownership is negatively correlated with monthly volatility, and is statistically significant at the 1% level. Specifically, if the proportion of a stock owned by ETFs increases by 1%, the regression indicates that the monthly volatility of the stock will fall by 0.029. Yet, different stocks have different characteristics that affect their volatility. Small companies are known to be more volatile than large companies, and illiquid stocks are known to be more volatile than liquid ones. When adding three additional variables to control for size and liquidity differences among S&P 500 stocks, we see that the effect is reduced but still statistically significant at the 1% level. After accounting for these variables, a 1% increase in the proportion of a stock owned by ETFs corresponds to a fall in monthly volatility by 0.012. Since high or low volatility tends to cluster, I added a final variable to the regression in order to account for a stock’s recent volatility. This updated regression indicates that an increase in ETF ownership by 1% leads to a decrease in monthly volatility by 0.010, which is also statistically significant at the 1% level.

ETFs control immense amounts of capital, and must invest according to fixed rules. In order to study the effect of ETFs on volatility further, I also studied how volatility has been affected for stocks included and excluded from the S&P 500 over time. When
a stock joins the S&P 500 constituency, all ETFs tracking the S&P 500 are forced to buy the stock at the same time. As ETFs have grown in popularity and size, the amount that ETFs must purchase of these stocks has grown with it. Conversely, when stocks leave the S&P 500, ETFs tracking the S&P 500 are forced to sell the stock at the same time and in greater amounts as ETFs become larger. For my second regression, I compiled a list of changes to the stocks comprising the S&P 500 from 2008-2017 due to endogenous factors (such as minor changes in market capitalization or replacing a company that was acquired). I then calculated the change in ETF ownership from before and after the index inclusion or exclusion, and regressed weekly volatility of the stock for the week following the change on this metric. This initial regression yielded results that were not statistically significant at the 10% level. In order to more accurately link the fund flows into ETFs to a change in volatility, I added an initial regression of the ETF ownership change on US equity ETF assets under management on an annual basis, then regressed weekly volatility on the fitted ETF ownership change. The coefficient of the augmented regression was statistically significant at the 1% level. Specifically, if the change in ETF ownership for a stock entering the S&P 500 increases by 1% (i.e. – if S&P 500 ETFs were previously forced to buy 2% of the stock’s market capitalization and now have to buy 3% due to the increase in assets managed by the ETF), the volatility for the week following the change increased by 82.8 basis points (a percent of a percent).

In summary, while the ETF was one of the largest financial innovations in recent history, ETFs have swelled to a size where they have begun distorting the volatility of their underlying securities. While volatility for stocks entering and exiting the S&P 500
has increased along with the size of ETFs, they also artificially subdue the volatility of stocks within the S&P 500. ETFs were, and largely still are, beneficial because they piggyback on the wisdom of the market as a whole. Active investors in the stock market invest in companies most likely to grow and against companies most likely to fail. When millions of active investors attempt to achieve this goal at the same time, the aggregate is the best approximation for which companies will indeed succeed and fail. Since the overwhelming majority of ETFs are passive, they allow the active money to make stock-level investment decisions and ride the overall increase in the value of the equity market. Yet, as ETFs have grown to control immense amounts of capital, their allocations no longer only ride the aggregate wisdom of the market, but also influence it, as shown by the statistically significant effects they have on stocks’ volatility. It is apparent that the rise in ETFs has created a market environment where the tail wags the dog, as opposed to the other way around.
Currently, there does not exist many academic papers focusing on the effects of the rise in ETFs on financial markets. The few studies that focus on ETFs usually focus on their effect on correlations between stocks or how they may inadvertently generate opportunities for above-market returns; even fewer studies focus on how ETFs affect volatility. Furthermore, the papers do not always focus on traditional ETFs, as a sizable portion of the papers study the effect of leveraged ETFs. This gap in academic literature provides a unique opportunity for my research to help fill a void in the existing body of knowledge surrounding ETFs. Yet, the existing literature helped aid the direction of my research and provide new ideas for approaching my own hypotheses.

Many of the earliest papers focus on the effects of index design, a notion closely tied to ETFs as the largest ETFs are index funds, specifically searching for the potential existence of opportunities for above-market returns. “Do Demand Curves for Stocks Slope Down?” was written in 1986, and is one of the first papers to analyze the effect of index inclusion on stock returns. In it, Shleifer concludes that stocks newly included in the S&P 500 index exhibit abnormal positive returns for up to 10 days after index inclusion. The paper further argues that when stocks enter the S&P 500, usually due to an existing firm in the S&P 500 being acquired, index funds are forced to buy large portions of the stock, thus raising the demand for the stock and causing abnormal positive returns. Since the index inclusions are devoid of other “informational effects” that may cause a stock’s price to rise, the paper concludes that the demand curves for
stocks must slope down, in contrast with the efficient market hypothesis which states the demand curves for stocks should be horizontal in theory.

The Shleifer paper inspired other academics to study the effect of index inclusion and exclusion on stock returns. After 1986, Standard and Poor's, the company responsible for maintaining the S&P 500 index, changed their policy for index constitution changes, electing to announcing changes 5 days prior to the effective date. In 1996, Beneish and Whaley coined the term “S&P Game”, which is the act of buying stocks that have been announced to join the S&P 500 at the close of announcement day, and selling them at the close of the effective date. The authors conclude that this change in the rules of index constitution creates an opportunity for abnormal positive returns by “playing” the S&P Game, with stocks gaining 3.1% the night after the announcement, and an additional 4.1% by the close of the effective date, on average. Furthermore, they conclude that the returns largely persist after the effective inclusion date.

Madhavan applied a similar concept in “The Russell Reconstitution Effect”, rather focusing on the Russell 2000 and Russell 3000 indices and studying a “spread portfolio” of long index additions and short index removals. The Russell indices differ from the S&P 500 with respect to the fact that Russell indices reconstitute on the same date each year, June 1st, rather than irregular intervals throughout the year. Furthermore, since the only criterion for index inclusion is market capitalization as of May 31st, it is relatively easy to predict which stocks would be included or excluded. Madhavan calculates that the spread portfolio for the Russell 3000 index returns 14.94% in the month of June alone, with the Russell 2000 spread portfolio returning 10.78% for the
same period, on average. Madhavan further notes significant additional positive return on average for the spread portfolio during the months preceding index reconstitution, specifically March through May, yet shows that the same portfolios tend to have negative returns for July, suggesting potential mean-reversion in stock prices.

It took longer for academia to begin looking at the effects of ETFs on volatility. Yet, many volatility papers focused on the effect of leveraged ETFs, who are often forced to settle their derivatives contracts at the end of each day. In 2013, Boney-Dutra, Guirguis, and Mueller studied how the intraday trading behavior required of leveraged ETFs at the end of each trading day affects the volatility of REITs (real estate investment trusts) and the Dow Jones Industrial Average, another popular index. The paper concludes that the end-of-day rebalancing of these ETFs increased volatility in 75% of instances for REITs and in 69% of instances for the Dow Jones Industrial Average, offering a glimpse of how ETFs have affected the underlying characteristics of the stock market.

One year prior to the release of the previous paper, Curcio, Anderson, Guirguis, and Boney released a similar paper, focusing only on real estate stocks but also studying the effects of traditional ETFs on volatility in conjunction with the effects of leveraged ETFs they elaborated on in the later paper. They concluded that leveraged ETFs tied to the Dow Jones US Real Estate and Financial indices increased the volatility of underlying real estate stocks by 300% on average. Weaker but still significant, they also concluded that traditional ETFs tied to the indices caused a 70% increase in the volatility of underlying real estate stocks. This assertion was one of the first explicit analyses of traditional ETFs and the volatility of their underlying securities.
Perhaps the most explicit study of the effect of traditional ETFs on stocks’ underlying volatility was published in 2015, titled “Do ETFs Increase Volatility?” by Ben-David, Franzoni, and Moussawi. The paper asserts that a one standard deviation increase in ETF ownership corresponds to a statistically significant increase in volatility between 9-15% of a standard deviation for stocks in the S&P 500; they also note that this relation holds when expanded to the Russell 3000, but smaller by a factor of four. They claim that the relationship is indeed causal, rather than merely correlated, by analyzing the way Russell constitutes their Russell 2000 and 1000 indices. Lastly, the authors conclude that intraday arbitrage trading is the mechanism with which volatility is distorted. These results are important as they “[run] contrary to the belief that the rise in passive investing is unambiguously related to increased pricing efficiency”, a belief my paper will also call into question (Ben-David et al. 5).

Perhaps the most comprehensive study on ETFs and their ability to distort market fundamentals was conducted by Bradley and Litan in 2010. The authors claim that ETFs “are now undermining the traditional price discovery role of exchanges”, which can give rise to a litany of problems (4). They argue that ETFs increase the co-movement of stocks in the market as well as increasing systematic volatility of the market as a whole. They call upon tangible events as evidence for the dangers in the rise of ETFs, namely the “Flash Crash” of May 6, 2010. Perhaps the main argument of the paper, the authors conclude that ETFs are adversely affecting small companies, both those that are publicly traded and those looking to go public. For publicly traded companies with small market capitalizations, they argue that ETFs, rather than the intrinsic pulls between supply and demand for shares, have begun setting the market
price for shares of the company. Due to these distortions, the paper asserts that companies are less inclined to go public through an initial public offering ("IPO"), where a portion of the company’s shares are released publicly for the first time. They call upon the rising number of ETFs and declining number of publicly traded companies as evidence for this trend. The authors of the paper were so convinced of the ways ETFs may adversely distort the market, they testified a condensed version of the paper titled “ETFs and the Present Danger to Capital Formation” in front of Congress in October 2011.

In summary, the few academic papers focused primarily on ETFs tend to avoid exploring the effects of traditional ETFs on the volatility of underlying securities. “Do ETFs Increase Volatility” is perhaps the only paper to directly address this relationship, and I used their methods of obtaining data to drive my own data collection and calculations. Even though many of the other papers focused on returns associated with index constitution, I used their ideas to generate a framework to study volatility. Specifically, I adapted the concept of the “S&P Game” coined by Beneish and Whaley in “An Anatomy of the ‘S&P Game’ – The Effects of Changing the Rules” as well as the “spread portfolio” studied by Madhavan in “The Russell Reconstitution Effect”. The former paper studies the returns of stocks that are about to be added to the S&P 500, while the latter studies the returns of a portfolio of buying stocks entering the Russell 3000 and shorting stocks exiting the Russell 3000. Rather than studying the effects of these index constitution changes on returns, I studied how the changes affected the volatility of stocks after inclusion or exclusion. Therefore, while little academic research
applies directly to my research question, I used ideas from existing papers to formulate my own ways to study the effects of ETFs on volatility.
Data & Methodology

First and foremost, I selected the S&P 500 as the index to base my analysis on. This selection was for a variety of reasons. The S&P 500 is the world’s most popular index and is the benchmark for tracking the performance of trillions of dollars of assets. The S&P 500 represents roughly 80% of all investable assets in the US by market capitalization. Therefore, the S&P 500 spans a subset of the domestic equity universe that is robust enough to apply to the market on a broader scale.

Central to my research question was the establishment of a metric for ETF ownership. Through the Wharton Research Data Services ("WRDS") subscription available to me through the Kellogg School of Management, I was able to access University of Chicago’s Center for Research in Security Prices ("CRSP"). CRSP is one of the largest and most robust data sets for financial data, and was instrumental in obtaining accurate data.

The first step in creating a metric for ETF ownership was compiling a list of ETFs with which to include in the metric. As described before, certain ETFs use derivatives rather than physically owning the underlying security. Since these ETFs would not affect the shares outstanding of a stock, I restricted my search to a smaller subset of ETFs that met certain criteria. I sought ETFs that were domestic, equity, long-only, non-leveraged ETFs that have traded publicly between 2010-2017.

In order to compile this list of ETFs, I utilized the Mutual Fund Database within CRSP, specifically searching within the Fund Summary section. Using CRSP objective codes, the query returned results for equity-only domestic funds. Furthermore, I used
the more specific Lipper objective codes to narrow the selection further. The query
returned funds that identify as S&P 500 index objective funds (SP), mid-cap funds (MC),
small-cap funds (SG), micro-cap funds (MR), capital appreciation funds (CA), growth
funds (G), growth & income funds (GI), equity income funds (EI), basic materials funds
(BM), consumer goods funds (CG), consumer services funds (CS), health/biotechnology
funds (FS), industrials funds (ID), natural resources funds (NR), real estate funds RE),
science & technology funds (TK), telecommunications funds (TL), specialty funds (S), or
utilities funds (UT). In order to ensure no non-target funds slipped through, the query
filtered out funds in the natural resource space that do not own equities (CME, CMG,
CMM, CMP, or CMS), as well as equity leveraged (DL), dedicated short bias (DSB), and
long/short equity funds (LSE). Lastly, the query filtered out all funds that were not
categorized as “F” for the ETF Fund Flag variable (“F” indicates the fund is an ETF, “N”
indicates the fund is an ETN, and nothing indicates it is neither).

The result of the query was 892 unique domestic, equity, long-only, non-
leveraged ETFs. The average assets under management for the ETFs compiled was
$2.2 billion as of December 2017, ranging from funds with under $1.0 million in assets
to SPY with $251.4 billion in assets. Descriptive statistics for the list of ETFs used can
be found in Table 1. The market for this subset of ETFs is very segmented, with a
handful of index funds controlling the broad majority of assets under management, as
can be seen in Chart 1.
Like mutual funds, ETFs are required by the Investment Company Act to disclose their positions on a monthly basis. Using our list of ETFs and their monthly filings, I was able to build an ETF ownership metric for stocks in the S&P 500. Within CRSP’s Mutual Fund Database, now searching within the Portfolio Holdings tab, I searched for monthly holdings reports for the 892 ETFs I previously compiled by their unique ETF Fund Numbers (Fundno). I iterated this search for every year from 2010-2017. I selected 2010 as the start of my analysis because CRSP’s data is incomplete for certain ETFs before 2010. The result of the search yielded all of the positions of the 892 ETFs (including tickers, CUSIPs, and shares owned by the fund), on a monthly basis for 2010-2017.
Using Stata, I was then able to build a do-file that sums the total shares of all S&P 500 stocks as of December 31, 2017 owned by my target ETFs, on a monthly basis. I then iterated the do-file on each of the 8 years of data from 2010-2017. After then appending the data into one data set, I had effectively calculated the number of shares of S&P 500 stocks held by domestic, equity, long-only, non-leveraged ETFs, for each month, from 2010-2017.

The last step in creating my ownership metric was to compare the number of shares owned by ETFs to the total number of shares outstanding of the S&P 500 stocks. For this step, I utilized the Bloomberg Terminals also provided through the Kellogg School of Management. Using Bloomberg, I was able to pull the free-float shares outstanding for each of the S&P 500 stocks on a monthly basis, for 2010-2017. I use free-float shares outstanding rather than total shares outstanding because the non-free-float shares will have no impact on the free-float stocks’ volatility. To finalize the ownership metric, I divided the number of shares held by ETFs by the total number of free-float shares, yielding a percentage according to the following formula:

$$ETF	ext{ Ownership}_{i,t} = \frac{\sum_{n=1}^{892} \text{shares of } i \text{ owned by the } n^{th} \text{ ETF}_{i,t,n}}{\text{shares outstanding}_{i,t}}$$

$i$ = stocks in S&P 500, $n$ = ETF, $t$ = time (monthly)

Descriptive statistics for the ETF ownership metric can be found in Table 2.

The ETF ownership metric would become the independent variable in my first regression, with monthly volatility of each stock as the dependent variable. Using
Bloomberg, I obtained daily price data for each of the S&P 500 stocks from 2010-2017. From this, I calculated logarithmic daily returns as:

\[
daily\ return_{i,t} = \log \left( \frac{price_{i,t}}{price_{i,t-1}} \right)
\]

\(i = \text{stocks in S&P 500, } t = \text{time (daily)}\)

From this, I calculated monthly volatility as the standard deviation of logarithmic daily returns:

\[
monthly\ volatility_{i,t} = \sigma \left( (daily\ return_{i,t})^{m} \right)
\]

\(i = \text{stocks in S&P 500, } t = \text{time (month), } m = \text{number of trading days in month } t\)

The last part of my data collection involved calculating a change in ETF ownership metric for stocks that enter or leave the S&P 500. I obtained data from Siblis Research listing all of the changes to the constituent list of the S&P 500 from 2000-2018. First, I reduced this sample to the years 2008-2017 in order to capture large growth in ETFs over time and maintain the accuracy of CRSP and Bloomberg data. My goal was to only include examples where stock inclusion/exclusion was endogenous, merely determined by a small change in market cap with all else equal, rather than causes from outside forces. Furthermore, I wanted to study how volatility surrounding the change was affected, so I chose to only include companies with valid data for the two months prior to and after the change. Thus, I filtered out additions to the S&P 500 that were due to a recent IPO of a large private company, companies absorbed through mergers and acquisition (“M&A”) activity, as well as other stocks with incomplete price
(and hence volatility) data immediately surrounding the inclusion or exclusion from the S&P 500.

The result was 284 instances of stocks entering or exiting the S&P 500 due to endogenous factors, specifically 187 inclusions and 97 exclusions. We note that there are more inclusions than exclusions; this difference is due to the fact that S&P 500 companies are occasionally acquired by other S&P 500 companies, thus opening a spot for a company just below the market capitalization threshold to enter the index’s constitution.

Using the same methodology as the stocks in the first regression, I calculated the ETF ownership metric for the 284 stocks for the month prior to and after the index inclusion or exclusion. For example, since EQT Corp (ticker: EQT) joined the S&P 500 in December 2008, I computed the ETF ownership metric for the months ending 11/30/2008 and 12/31/2008.

The goal of the second regression was to calculate how the rise in ETF popularity, specifically in terms of assets managed by ETFs, affect the volatility of stocks entering and exiting the S&P 500. Thus, we calculated our independent variable as the change in the ownership metric for the month before and after the change in index constitution. Descriptive statistics for the change in ownership metric can be found in Table 3. A specific example of how the metric was calculated for a stock included in the S&P 500 can be found in Table 5a, while the same can be found for a stock excluded from the S&P 500 in Table 5b.

\[
Ownership\ Change_{i,t} = ETF\ Ownership_{i,t} - ETF\ Ownership_{i,t-1}
\]

\[i = \text{stock}, \ t = \text{month of index addition/exclusion}\]
In order to study the effect of a change in ownership on a change in volatility, we computed the volatility of each stock for the week following the index inclusion or exclusion. Weekly volatility was computed as the standard deviation of the daily logarithmic returns:

\[
\text{weekly volatility}_{i,t} = \sigma \left( \left( \text{daily return}_{i,t} \right)_{t=1}^{7} \right)
\]

\(i = \text{stocks in S&P 500}, \ t = \text{time (week)}\)

Lastly, in order to specifically determine the effect of ETFs on ownership change, and thus volatility, I compiled ETF fund flow data from Statista on the annual level. This metric was used to calculate the assets under management of the broader ETF market on a yearly basis, and was restricted to only include US equity ETFs, keeping the data on par with the ownership data previously pulled from the 892 ETFs.
Analysis

The goal of my analysis was twofold: to measure how ETF ownership affects the volatilities of underlying securities in the S&P 500 generally, and to study how ETF constitution rules affect the volatility of stocks entering or exiting the S&P 500 over time, thus capturing the large rise in assets managed by ETFs. In order to study both aspects of ETFs’ effect on volatility, I conducted two main regressions.

First, I sought to measure the impact of ETF ownership on volatility generally. Through the data collected through CRSP, I calculated the percent of a stock’s free-floating shares outstanding owned by target ETFs, what I refer to as the ETF ownership metric. I used this as my independent variable, with volatility at the monthly level as a dependent variable. Since the ETF ownership metric and monthly volatility calculations were for the all stocks in the S&P 500 for every month across 8 years (2010-2017), both my independent and dependent variables were in the form of panel data. Thus, I conducted an OLS panel data regression with fixed effects of monthly stock volatility on ETF ownership.

**Regression 1a:** \[ Vol_{i,t} = \alpha_i + \beta_1ETF\text{Ownership}_{i,t} + u_{i,t} \]

Many other factors can have an impact on a stock’s monthly volatility. While the S&P 500 represents the largest 500 companies by market capitalization in the US, it contains companies of many different sizes and characteristics. For example, smaller companies tend to exhibit greater variance in the returns on their stock price than larger, better established companies. Thus, the size of a company is a factor which may skew
the effect of ETF ownership in my original regression. Furthermore, stocks have varying degrees of liquidity. Liquidity is the ease of which a security can be readily bought and sold. It is well documented that stocks with lower liquidity exhibit higher volatility, since with fewer counterparties willing to take the opposite end of a trade, each individual tick in the stock price becomes more pronounced. This heightened volatility can influence volatility at the monthly level. Therefore, it became apparent that I needed to account for differences in stock size and liquidity when determining the effect of ETF ownership on monthly volatility.

In order to disentangle how ETF ownership affects a stock’s volatility from its intrinsic characteristics such as size and volatility, I added three further independent variables to my regression. First, I computed the log of the stock’s market capitalization (equal to total shares * price of each share), for each stock for every month in 2010-2017. Second, I computed the inverse of the stock’s price for each stock for every month in 2010-2017. These two extra variables helped account for intrinsic differences in the size of the company for which the stock represents. Third, I computed the average bid-ask spread for each stock on a monthly basis. The bid-ask spread is the difference between the price a seller is willing to sell a stock from the price a buyer is willing to buy the stock. When a stock is traded frequently, the increased number of agents buying and selling the stock pushes the bid-ask spread lower. Conversely, when a stock is traded infrequently, this difference in willingness to pay between buyers and sellers increases. Hence, the bid-ask spread is an common proxy for the liquidity of a stock, and is used frequently in academic literature to represent it. This extra variable helped account for the intrinsic differences in the liquidity of each stock. When regressing ETF ownership
on monthly volatility while accounting for size and liquidity differences, I lagged each of
the three new variables by one month.

Regression 1b:

\[ \text{Vol}_{i,t} = \alpha_i + \beta_1 \text{ETF Ownership}_{i,t} + \beta_2 \ln(\text{market cap}_{i,t-1}) + \beta_3(1/\text{price}_{i,t-1}) + \beta_4(\text{bid} - \text{ask spread}_{i,t-1}) + u_{i,t} \]

While a stock’s size and liquidity can affect its volatility, other factors may persist
in affecting the data, and distacing my regression results from the true effect of ETF
ownership on volatility. Volatility can cluster, with periods of high volatility surrounding
other periods of high volatility, and periods of low volatility surrounding other periods of
low volatility. This clustering effect can be due to larger market forces such as
geopolitical events or unpredictable changes in monetary policy, but can also be due to
fundamental changes within the company such as unexpected company
announcements or increased media coverage. The fixed-effects model I implemented in
my OLS regressions accounted for trends in overall market volatility, but failed to
account for persistent periods of similar volatility at the stock level. Thus, I added
volatility lagged by one month to the list of extra independent variables to account for
these fundamental trends in volatility of a particular stock.

Regression 1c:

\[ \text{Vol}_{i,t} = \alpha_i + \beta_1 \text{ETF Ownership}_{i,t} + \beta_2 \ln(\text{market cap}_{i,t-1}) + \beta_3(1/\text{price}_{i,t-1}) + \beta_4(\text{bid} - \text{ask spread}_{i,t-1}) + \text{Vol}_{i,t-1} + u_{i,t} \]
In order to tackle my second objective, to measure how ETF constitution has affected stock volatility over time, I conducted a second main regression. As described earlier, I calculated the change in ETF ownership of a stock as the difference between a stock’s ETF ownership metric both before and after being included or excluded from the S&P 500. Intuitively, I hypothesized that stocks entering the S&P 500 would exhibit an increase in their ETF ownership metric, since most of the largest ETFs must purchase a stock that has just been added to the S&P 500 in massive quantities. Similarly, I hypothesized that stocks exiting the S&P 500 would show a decrease in their respective ownership metric. I assumed these changes to be symmetric, with a stock entering the S&P 500 having roughly the negative change in ETF ownership of a stock leaving the S&P 500, if in a similar time frame. Since my data accounted for both inclusions and exclusions to the S&P 500, in order to align the change in ETF ownership across these two types of events, I multiplied the ownership change of stocks exiting the S&P 500 by -1, thus correcting the ETF ownership change metric to treat inclusion and exclusion events in the same manner.

My goal was to analyze how the ownership change metric affected volatility over time, attempting to capture the effect of the growing pool of capital managed by ETFs throughout the past decade. Since I was no longer considering all stocks across all time frames, but rather a select few stocks over select time frames, I could no longer use panel data econometric methods. Instead, I averaged the change in ownership across each year in 2008-2017. Similarly, I averaged the volatility of the stock for the week following the index inclusion or exclusion for each year. With average ownership change acting as my independent variable and average volatility for the week following the
change in index constitution as my dependent variable, I conducted a time-series OLS regression.

**Regression 2a:**  \[ Vol_i = \alpha_i + \beta_i Ownership\ Change_i + u_i \]

One problem with the previous regression is that it did not account for explicit data measuring the flow of assets into ETFs. Since the assets managed by ETFs have increased monotonically with time, one could hypothesize that the ETF fund flows were responsible for a larger average ownership change metric year over year, and thus a larger effect on stock volatility if the regression returned a positive coefficient. Yet, this intuitive argument needed to be tested with explicit data; thus, I augmented the regression to include a prior step. For the second version of the regression, I used annual US equity ETF assets under management data as an independent variable, with the average annual ownership change metric serving as the dependent variable. Then, I regressed volatility of the stock for the week following the index constitution change on the fitted values of the ownership change metric for each year. This additional step, helped more accurately account for differences in ETF assets under management for each year, and thus more accurately represents how the flow of capital into ETFs have affected the volatility of underlying securities over time.

**Regression 2b:**

\[ Ownership\ Change_i = \alpha_i + \beta_i ETF\ Fund\ Flow_i + u_i \]

\[ Vol_i = \alpha_i + \beta_i Ownership\ Change_i + u_i \]
## Results

### Table 1
Descriptive Statistics for Assets Under Management of Target ETFs

<table>
<thead>
<tr>
<th>Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>892</td>
<td>2189.54</td>
<td>11651.33</td>
<td>0.72</td>
<td>251411.34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10th Percentile</th>
<th>25th Percentile</th>
<th>50th Percentile / Median</th>
<th>75th Percentile</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.97</td>
<td>13.07</td>
<td>102.41</td>
<td>714.69</td>
<td>3300.97</td>
</tr>
</tbody>
</table>

All values are in millions USD

### Table 2
Descriptive Statistics for ETF Ownership Metric

<table>
<thead>
<tr>
<th>Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>45,665</td>
<td>6.18%</td>
<td>4.10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10th Percentile</th>
<th>25th Percentile</th>
<th>50th Percentile / Median</th>
<th>75th Percentile</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.44%</td>
<td>2.78%</td>
<td>5.87%</td>
<td>8.67%</td>
<td>11.44%</td>
</tr>
</tbody>
</table>

### Table 3
Descriptive Statistics for Change in ETF Ownership Metric

<table>
<thead>
<tr>
<th>Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>284</td>
<td>1.34%</td>
<td>2.70%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10th Percentile</th>
<th>25th Percentile</th>
<th>50th Percentile / Median</th>
<th>75th Percentile</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.5%</td>
<td>0.4%</td>
<td>1.1%</td>
<td>2.5%</td>
<td>4.8%</td>
</tr>
</tbody>
</table>
### Table 4
Regressions 1a, 1b, and 1c

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Monthly stock volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S&amp;P 500</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>ETF ownership</td>
<td>-0.029*** (0.001)</td>
</tr>
<tr>
<td></td>
<td>-0.012*** (0.001)</td>
</tr>
<tr>
<td></td>
<td>-0.010*** (0.001)</td>
</tr>
<tr>
<td>ln(Market cap (t-1))</td>
<td>-0.351*** (0.000)</td>
</tr>
<tr>
<td></td>
<td>-0.225*** (0.000)</td>
</tr>
<tr>
<td>1/Price(t-1)</td>
<td>-5.367** (0.018)</td>
</tr>
<tr>
<td></td>
<td>-6.038*** (0.017)</td>
</tr>
<tr>
<td>Bid-ask spread (t-1)</td>
<td>-4.890*** (0.009)</td>
</tr>
<tr>
<td></td>
<td>-5.496*** (0.009)</td>
</tr>
<tr>
<td>Volatility (t-1)</td>
<td>33.417*** (0.005)</td>
</tr>
<tr>
<td>Constant</td>
<td>16.834*** (0.000)</td>
</tr>
<tr>
<td></td>
<td>10.189*** (0.002)</td>
</tr>
<tr>
<td></td>
<td>6.737*** (0.002)</td>
</tr>
<tr>
<td>Observations</td>
<td>44,334</td>
</tr>
<tr>
<td></td>
<td>43733</td>
</tr>
<tr>
<td></td>
<td>43726</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>0.234</td>
</tr>
</tbody>
</table>

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively
Fitted values represent the graph of the line from Regression 1a
### Table 5a
**Example: Inclusion**
EQT Corp (ticker: EQT) joined the S&P 500 in December 2008

<table>
<thead>
<tr>
<th>Date</th>
<th>ETF Ownership</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/31/2008</td>
<td>4.70%</td>
<td>1.80%</td>
</tr>
<tr>
<td>11/30/2008</td>
<td>2.90%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5b
**Example: Exclusion**
Meredith Corp (ticker: MDP) left the S&P 500 in January 2011

<table>
<thead>
<tr>
<th>Date</th>
<th>ETF Ownership</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>31-Jan-11</td>
<td>2.70%</td>
<td>-2.40%</td>
</tr>
<tr>
<td>31-Dec-10</td>
<td>5.10%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6
**Regressions 2a and 2b**

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Weekly stock volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample:</td>
<td>284 changes to S&amp;P 500 constitution</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
</tbody>
</table>

|                        |                        |                        |
| ETF Ownership Change   | 24.248                  | 82.833***               |
|                        | (14.572)                | (20.598)                |
| Constant              | 0.186                   | -0.615*                 |
|                        | (.219)                  | (0.289)                 |
| R-squared             | 0.260                   | 0.669                   |

*Note: ****, ***, and * denote statistical significance at the 1%, 5%, and 10% level, respectively*
For Regression 1, we see that the results that are highly statistically significant. Each coefficient for Regression 1a, 1b, and 1c is significant at the 1% level, except for the lagged inverse price variable in Regression 1b, which is statistically significant at the 5% level. Yet, we see that the R-squared for Regression 1a is quite low, due to the size of the data set and noise in the ETF ownership metric. We can see how the data relates to the results of the regression in Chart 2, which graphs the fitted values of Regression 1a on a scatter plot of the data used. We also note that the R-squared goes up significantly when adding size, liquidity, and then recent volatility controls. While this rise may indicate that the effect of ETF ownership on volatility is stronger, it is to be expected when adding more variables to the regression, so we cannot conclusively attribute it to a more accurate model.

Perhaps the most important revelation from Regression 1 is the sign of the ETF ownership coefficient. We see that even after controlling for size and liquidity differences, and then after adding a control for recent volatility, the coefficients are all negative and statistically significant at the 1% level. This result is surprising as it contradicts the results of Ben-David, Franzoni, and Moussawi, which asserts that traditional ETF ownership is positively correlated with volatility of stocks. We do note that the magnitude of the coefficient is halved when accounting for size and liquidity controls. When recent volatility is added to the list of controls, we see the magnitude fall to roughly one-third the level of the baseline regression. This trend indicates that the majority of the effect seen in Regression 1a is due to other factors besides ETF ownership, but cannot be ignored as the statistical significance is maintained.
There are several reasons that may explain the difference in direction for the effect of ETF ownership on volatility between my research and the earlier paper. My data was collected from the same data sources, but significantly updated to represent the current state of the ETF market. The earlier paper uses data from 2000-2012, while my own research uses data from 2008-2017. The ETF market has grown at an astonishing speed over the past decade, and the extra five years of data adds significant new data points. According to ETF fund flow data from Statista, the market for US equity ETFs increased nearly 240% from the end of 2012 to the end of 2017. Furthermore, this rise in assets under management correlates highly to the growth in the ETF ownership metric for the most recent years. I hypothesize that in the early-2000s, the effect of a new investment vehicle gaining popularity in the market exacerbated volatility for underlying securities, but as the size of the ETF market grew, stocks reached a “tipping point” where ETFs began inducing reduced volatility due to their inherently passive nature. While in the early 2000s, ETF ownership was large enough to impact a stock’s intrinsic trading characteristics, it was not until recently that ETF ownership was large enough to subdue the daily price movements of a stock. This tipping point hypothesis would explain the different results we observe, but would require further research to determine its validity or discern what level of ETF ownership is the tipping point between higher and lower volatility.

Although my research indicates that ETF ownership has reduced volatility for stocks in the S&P 500, we also see instances where they can cause heightened volatility. For Regression 2a and 2b, we note that the coefficients for change in ETF ownership are both positive. This result indicates that when a stock enters the S&P 500
(due to exogenous reasons like small market capitalization changes or firms already in the S&P 500 getting acquired), and thus receives a bump in its level of ETF ownership, volatility rises. Yet, we note that none of the coefficients in Regression 2a are statistically significant at the 10% level, obscuring the assessment of whether a correlation indeed exists. Regression 2b uses extra data on ETFs to further assess whether a correlation exists. By first regressing the ETF ownership change metric on traditional US equity ETF assets under management, then regressing weekly volatility on the fitted ETF ownership change values, we see that the effect becomes statistically significant at the 1% level. This shift indicates that when accounting for the actual fund flows into ETFs over the past 10 years and its corresponding impact on ETF ownership, there does exist a statistically significant impact on volatility. We also see the magnitude of the effect increase by over 300%, signaling that the effect is not only more statistically significant, but stronger than the baseline regression.
Conclusion

My research suggests that ETFs have a statistically significant effect on the volatility of underlying securities. For stocks already in the S&P 500, ETFs tend to artificially reduce their volatility. Yet, when stocks enter (leave) the S&P 500, we see that the capital flowing in (out) of the stocks causes volatility to rise. Hence, the size of the ETF market lends it the ability to distort the volatilities of underlying securities, sometimes higher and sometimes lower.

As the growth in ETFs is likely to persist, I predict these effects will magnify. If ETFs continue to grow at the speed with which they have for the past decade, the portion of each company in the S&P 500 owned by ETFs will increase to new highs. This growth will further reduce the pool of investment capital seeking to profit off of stock mispricings, and inhibit the efficient price-seeking behavior of the stock market, causing volatility to fall for equities across the board. With more and more money tied up in ETFs that invest according to fixed rules (such as the necessity of S&P 500 index funds to own whichever stocks are in the S&P 500), the amount of money that would need to pour into stocks when changes occur will continue to rise. The increased severity of these one-time shifts in capital will cause stocks to swing more violently.

Neither increased nor decreased volatility is inherently more desirable when seeking a market to operate more efficiently. Many investors, such as pension funds, seek stability in their investments; they would thus prefer lower volatility. The same applies to companies with publicly floating shares, desiring stability for their shareholders’ investments. Yet other market participants such as arbitrage traders,
market makers, and certain hedge funds tend to profit during periods of high volatility. In theory, the aggregate behaviors of all market participants should move a stock to its intrinsic value based on all publicly available information (a main tenet of the efficient market hypothesis). This aggregate should also cause the trading behavior of stocks, particularly a stock’s volatility, to reflect the fundamental uncertainty and cyclicality of the companies they represent. When the stock market operates efficiently, poorly performing companies perish while well performing companies thrive, a clearly desirable outcome. Under this scenario, the stock market and economy as a whole can grow at a rate which is sustainable. If markets operate inefficiently, giving rise to speculative bubbles or prolonged, fear-driven recessions, the economy as a whole suffers in the long-term. Therefore, neither low nor high volatility should take preference, but rather volatility that efficiently represents the intrinsic nature of the company.

The growth of the ETF market has begun impeding the equity market’s ability to efficiently determine intrinsic characteristics of a company’s stock. In both the reduced volatility observed for stocks within the S&P 500 as well as heightened volatility for stocks entering and exiting the S&P 500, “efficient volatility” is increasingly obscured. As the market strays further from efficiency, the undesirable byproducts that accompany it such as recessions may occur more often and more aggressively. While the invention of the exchange-traded fund was a groundbreaking financial innovation that brought low cost, liquid, and diversified investments to the masses, the unintended distortions it created cannot be ignored.
References


