SCHOOL OF
ECONOMICS AND
MANAGEMENT

LUND UNIVERSITY

Analysis of the Performance of ETFs

A study on the US market

By

Salome Tskhoidze

Vu Thi Kim Lien

May 2021

Master's Programme in Finance

Supervisor: Hossein Asgharian

Abstract

Exchange Traded Funds are known as a relatively recent financial innovation and have been gaining investors' interest in recent years. The performance of ETF in comparison to other benchmarks is still the central concern when investors make an investment decision. This thesis conducts empirical studies in the US market, using the Mean-Variance portfolio optimization, to construct optimal portfolios for both ETFs and underlying assets with and without short-selling constraints to compare these two portfolios based on historical and expected performance. The performance of individual ETFs and their underlying assets during the whole testing period and different downturns is analyzed before portfolios are constructed. The thesis then applies Fama and French three-factor model to assess the portfolios' risk-adjusted returns. Finally, different financial metrics are used to evaluate the portfolios' performance. The empirical results show that if not taking the cost-efficiency and high liquidity characteristics of ETF into account, the ETF portfolio often underperforms the underlying asset portfolio. Besides, with the short-selling constraint, the ETF portfolio also underperforms the market portfolio in the testing period. However, if those advantages of ETF are taken into consideration and if short-selling is allowed and used as an effective hedging tool, the ETF portfolio might outperform these benchmarks.

Key Words: ETF, portfolio optimization, Sharpe ratio, Mean-Variance, financial downturns, benchmark.

Acknowledgments

We first would like to express our gratitude to Professor Hossein Asgharian of the Department of Economics at Lund University. He always supported us enthusiastically via zoom calls and emails when we had questions, especially during this Covid-19 time. We then would like to thank the Educational Developer Garo Harwood of Academic Skills Services at Lund University for his useful feedback on the thesis. Thank you all for your support. This thesis would not have been possible without your help.

Contents

1. Introduction

This study particularly focuses on Exchange Traded Funds (ETFs) as a relatively recent financial innovation of passive portfolio management and seeks to compare the financial performance of an ETF portfolio with an underlying asset portfolio. Additionally, this paper aims to investigate ETF's financial performance in comparison with the market portfolio.

The Exchange Traded Fund (ETF) was initially introduced to the North American market in the early 1990s with the TIPs in Canada and SPDRs in the US. The ETF market then witnessed an effective boom in March 1999 with the launch of the Nasdaq-100 Index Tracking Stock (Zopounidis et al. 2010). After more than 30 years of development, the number of ETFs worldwide is 7,602 and the value of assets managed by ETFs globally is 7.74 trillion USD by the end of 2020 (Statista Research Department, 2021).

Due to the financial crisis and the bankruptcy of major investment banks in 2008, investors become more risk-averse and require the investment assets to be more transparent, inexpensive, and outperforming, leading to the explosion of the ETF industry. Also, the underperformance of actively managed funds and advantages of ETF, namely, lower cost, tax efficiency, transparency, and lower management fees make ETF more attractive as an investment strategy. These benefits have led to increasing demand for ETF that achieved growth of 2,650 percent between 2003 and 2020 (Statista Research Department, 2021). In August 2019, index-based equity mutual funds and ETF surpassed actively managed funds for the first time, reaching one of the biggest milestones in its modern history (Gittleson, 2019). That signifies a strong trend towards passive investing.

The increasing financial shocks and asymmetric information prevent investors from allocating assets efficiently. Investors have to run their businesses in an extremely risky environment, resulting in an increasing demand for a safe investment tool that can minimize the risk. That is one of reasons why ETF has been gaining the interest of investors in recent years. Along with the spread of ETF, more and more studies are carried out to research different aspects of ETF.

Many studies have recognized the effectiveness of ETFs. Tufano (2003) finds ETFs as a solution for market incompleteness, agency problems, avoidance of taxes and regulations, and high transaction costs. He argues that ETFs allow investors to invest in a portfolio that provides passive exposure to a variety of asset classes, risk factors, and security characteristics on top of intraday pricing and efficient tax management. French (2008) notes that the large cost of active investing encourages investors to change their investment policies in favor of passive investment strategies.

Several studies have been conducted about a comparison of active and passive portfolio management. Sharpe (1991) concludes that before deducting costs, on average, the return of actively managed and passively managed funds must be equal to each other, while after costs the return of the passively managed funds exceeds the return of the actively managed funds. This principle can be violated as a result of improper measurement. Pedersen (2018) contradicts Sharpe's equality which includes the assumption that the market portfolio never changes. He claims that equality breaks down because of this implicit assumption, which fails in the real world. He underlines the importance of active management as a tool of allocating resources efficiently, while passive investment's economic role is defined to create low-cost access to the market. French (2008) also compares the cost of active and passive investing from 1980 to 2006 in the US stock market. He calculates the cost of active investing 7.0, 30.5, and 101.8 billion dollars in 1980, 1993, and 2006 respectively, implying that the total cost grows with the market. He concludes that investors would increase their average annual returns by 67 basis points if they switched to a passive market portfolio. Fama and Litterman (2012) analyze the performance of actively managed mutual funds on the US market and explain the reason behind the fast-shifting to passive investing. Based on their study, before costs, return distributions for the universe of all mutual funds are centered at zero and their tails have the same size, meaning that average active managers cannot outperform the market. After subtracting costs, only the top 3% of managers could produce sufficient returns to cover only costs, indicating that top performers are expected to be as good as a low-cost performance of an index fund while bearing more costs.

As far as we know, there is limited research in the literature investigating the performance of ETF portfolios in comparison to corresponding underlying asset and market portfolios in the US market. Thus, this study focuses on the performance of ETF in the US market. The research question is whether the ETF portfolio can manage to beat the performance of its benchmarks. To answer this question, the thesis first compares each ETF with its corresponding underlying asset based on return and risk and analyzes the effect of financial downturns on their performance. Then, the thesis constructs optimal portfolios of ETF and underlying assets and compares the financial performance of the ETF portfolio with its benchmarks.

The target portfolios consist of three asset classes: equity, bond, and commodity. The underlying asset portfolio contains S&P 500 future, ten-year US treasury, and S&P GSCI index. Following this, ETF portfolio includes iShares Core S&P 500 UCITS ETF, iShares 7-10 Year Treasury Bond ETF, and iShares S&P GSCI Commodity Indexed Trust (GSG) indices. Within the portfolio construction framework, with and without short-selling constraints are applied. Since short-selling provides liquidity and increases the efficiency of the market, these two methods are used to measure the flexibility and effectiveness of short-selling on portfolio weights.

The main empirical results of this study show that ETFs and their underlying assets have parallel moving patterns in the whole testing period and different financial downturn periods. Results also demonstrate that the historical performance of ETF portfolio is always worse than that of the underlying asset portfolio both with and without short-selling constraints but the expected performance of ETF portfolio can better or worse than the performance of its benchmarks in different circumstances.

The remaining part of this study is structured as follows. Chapter 2 represents the literature review relating to the previous studies regarding ETF performance, market efficiency, and financial distress. Chapter 3 describes ETF background. Chapter 4 demonstrates data for empirical estimations and the applied methodology. Chapter 5 displays the empirical results and analytical interpretations, and Chapter 6 concludes the research findings.

2. Literature Review

This chapter presents a review of recent literature on ETFs. It is divided into three subsections: the literature on ETF performance in comparison with other types of funds and benchmarks, the literature on market efficiency of ETFs, and the literature on ETF performance in financial downturns.

2.1 ETFs Performance

Many papers confirm the outperformance of the ETFs compared with actively managed Open-Ended Funds (OEFs) and passively managed Closed-Ended Funds (CEFs). For example, Blakey (2007) makes a comparison between OEFs and ETFs. He analyzes that ETFs overcome the limitations of OEFs, coupled with reaping the benefits of low cost, diversification, and tax efficiency. Harper et al. (2006) conduct a study that compares the risk and return performance of ETFs for foreign markets and closed-ended country funds for 14 countries over the sample period from 1996 to 2001. They show that ETFs have a higher mean return and Sharpe ratio than foreign CEFs, while CEFs reveal negative alphas, indicating that passive investment strategies using ETFs surpass active investment strategies using closed-end mutual funds. Bello (2012) obtains evidence that equity-index ETFs outperform both equity-index mutual funds and S&P 500 index, even though mainstream findings suggested before that equity-index ETFs showed underperformance of both their benchmark indices and equity-index-mutual funds. The same research delivers that ETFs have lower expense ratios, lower portfolio turnover and smaller portfolio holdings compared with equity-index-mutual funds.

Much of previous research on ETFs have focused on the characteristics of ETFs that distinguish them from traditional active mutual funds. Gastineau (2001) gives special attention to the low expense ratios of ETFs and the ability of ETFs to avoid significant capital gains contributions. Fuhr (2001) highlights the flexibility of ETFs that makes it possible to trade with futures that, in turn, help both individual and institutional investors to increase or reduce their exposure to specific countries, sectors, industries, and styles. One of the most important features of ETFs is tax efficiency. With the help of the ‘redemption-in kind’ strategy, the distribution of realized capital gains is prevented, which makes ETFs predominant over traditional equity mutual funds (Poterba & Shoven, 2002). ETFs are mostly used for diversification purposes. Miffre (2007) emphasizes the advantages of country-specific ETFs for international diversification that can be achieved at low cost, with low tracking error, and in a tax-efficient manner. Investment in ETFs brings benefits for investors who are engaged in Sustainable Development Goals, particularly, for those who evince a special interest in achieving economic growth and improvement in portfolio performance compared with initial stock-bond portfolio (Miralles-Quirós et al. 2019).

Some research effort has gone into investigating ETF's performance compared with the benchmark. Most of the studies indicate that ETFs underperform the benchmark. Harper et al. (2006) discover that ETFs experience negative sides relative to the benchmark. Gastineau (2001), Svetina and Wahal (2008) examine ETF’s performance against their relative benchmark and conventional index funds. They conclude that ETFs are underachieved as opposed to competitive conventional mutual funds and their benchmark indices. Shin and Soydemir (2010) evaluate tracking errors from 26 ETFs using three methods. They use Jensen's model to test ETF relative performance. Their findings suggest that investing in ETFs does not provide a substantial benefit compared to their benchmark returns. On the other hand, Ma et al. (2011) show the performance improvement of constructed ETF portfolios to the benchmark using an optimization algorithm under a Markov regime-switching framework. The objective is to develop innovative optimization algorithms that have less reliance on active management strategy. As a result, the regime-dependent strategy outperforms the benchmark strategy. Mean returns and Sharpe ratios are much higher in the bull market rather than in the transition market.

Overall, taken together, these studies support the notion that ETFs are superior to OEFs and CEFs. Nevertheless, ETFs do not provide best-in-class performance with benchmark indices. Researchers also point out positive characteristics of ETFs that make them an attractive strategy for investors.

2.2 Market Efficiency and ETFs

There is an active debate around market efficiency in the ETF market. Some researchers argue that ETFs provide improved price efficiency, while others think that ETFs bring detrimental effects to markets by rising non-fundamental return volatility, altering correlation patterns, and reducing the liquidity of securities (Ben-David et al. 2015). Rompotis (2011) tests the weak form of market efficiency on the ETFs market, in the sense that whether all the already publicly available information is mirrored in the prices of ETFs or whether an investor can obtain above-market returns based on the released information. As a result, the efficient market hypothesis holds for

ETFs that implies that the majority of US ETFs are efficiently priced. A recent study by Zhao et al. (2021) explores Chinese A-share market-listed companies from 2002 to 2019. They provide evidence that the pricing efficiency of ETF-holdings stocks is substantially higher in comparison with stocks not held by ETFs. The pricing efficiency of ETFs improves greatly after ETFs enter the underlying stocks and decreases when ETFs exit the underlying stocks. In contrast to ETFs, traditional mutual funds show detrimental effects on market pricing efficiency. Tiwari et al. (2017) look into the efficient market hypothesis for Dow Jones sector ETF indices regarding short-run and long-run horizons and before and after Global Financial Crisis (GFC). The authors claim that there can be seen remarkable discrepancies in terms of market efficiency, between the short-term and long-term horizons. In addition to this, the ETF market efficiency has considerably decreased after the GFC.

ETFs are designed such that their prices are very close to Net Asset Value (NAV), providing price efficiency. However, existing evidence indicates that the existing persistent premiums or discounts tend to make the prices not very close to NAV, especially when it comes to ETFs that are not based in the US. Engle and Sarkar (2006) examine the magnitude of premiums and discounts for domestic and international ETFs. As it seems, domestic ETFs are priced very close to their true NAVs, while international ETFs are less precisely priced. Other researchers put a lot of effort to investigate the reasons behind this. Ackert and Tian (2008) examine the performance of US and other country ETFs currently traded in the US. Based on their investigations, the US funds show that they are priced close to their NAV, while the country funds demonstrate distorted prices. The mispricing of country funds can be explained by the existence of momentum, illiquidity, and size effects. They also find out that the relationship between fund premiums and market illiquidity is shaped like an inverted U, which suggests more active trading results in lower mispricing given that a certain level of liquidity is reached. According to Elton et al. (2002), the deviation of price from NAV is restricted by investor’s ability to create or remove ETFs at the end of every trading day. They state that management fees and dividends on the underlying securities may cause mispricing of ETFs. Blitz et al. (2002) state that not only fund expenses but also taxes are contributing factors for mispricing, especially dividend taxes.

To summarize the literature on the market efficiency of ETFs, it documents mixed evidence in terms of decrease or increase market efficiency. On the one hand, some papers consider ETFs as a liquid instrument that provides effective price discovery. On the other hand, the group of researchers finds that ETFs are priced distortedly.

2.3 Financial Distress and ETFs

Stock market shock known as "Black Monday" in 1987 triggered the call for a product such as ETFs by the Security Exchange Commission (SEC). Based on the SEC report, index futures and program trading drive stock markets to crash during market turbulent periods. Therefore, SEC was interested in such an instrument instead that anyone should devise one, in other words, having the possibility for investors to trade with baskets of securities (Liebi 2020).

The impact of ETFs on financial downturns is controversial. Many authors have recognized that ETF products stayed remarkably stable during the financial turmoil. For example, Goltz and Schröder (2011) analyze financial crises from 2007 to 2008 and point out that ETFs experienced considerable growth in some market segments; investment in equity ETFs consisted of more than 75% of all participating institutional investors in 2008 to greater than 84% in 2009. Likewise, investment in government bond ETFs has become more popular during financial crises. However, there are alternative asset classes, such as hedge fund ETFs that suffered from a severe decline in both usage and satisfaction levels within 12 months. Fuhr (2009) also indicates that compared to other types of funds, ETFs were not hit by the financial crisis and collected USD 74 billion. By contrast, European mutual funds experienced USD 570 billion outflows in the same period.

Several studies have tested that ETFs improve the liquidity of their component stocks during non-turbulent market times. Richie and Madura (2007) identify that the liquidity of the underlying stocks is improved as a result of the Nasdaq 100 Index (QQQ) creation. The more improvement is, the less heavily weighted stocks are. Hegde and McDermott (2004) search for a tool to improve market liquidity. They find out that after the introduction of an ETF, liquidity has increased for DJIA 30 index stocks as a result of a reduction in the asymmetric information cost of trading. On the contrary, a previous study explored by Pan and Zeng (2017) has emphasized that the liquidity of ETFs can deteriorate during times of market turbulence. The market illiquidity hinders ETF arbitrage to work efficiently, leading to persistent relative mispricing.

During several episodes in recent years, the ETF is accused of exacerbating liquidity shocks in times of financial distress. The best illustrative example is The Flash Crash on May 6th, 2010 that was caused by spreading shocking news about The Greek debt crisis. Borkovec et al. (2010) provide a shred of evidence that the liquidity of ETFs has been drastically reduced both in absolute terms and relative to individual securities in the baskets tracked by the ETFs. Liebi (2020) points out that the 9% reduction in Dow Jones Industrial Average (DJIA) during period causes price dislocations and illiquidity of ETFs. Due to this fact, the SEC imposed a new rule about requiring specific securities and ETFs with high volatility, leading to stop some ETFs from trading under certain conditions. As a consequence, over 300 ETFs stopped being traded in 2015. Many authors investigate the causes of the above-mentioned fact. The breaking down of the arbitrage mechanism increases the liquidity of underlying securities highlighted by Pan and Zeng (2017). Analogously, Ben-David et al. (2012) indicate that arbitrageurs significantly reduce their trading activities during periods of market distress, leading to the deterioration of the liquidity and market situation.

Overall, there seems to be some evidence to indicate that ETFs show better performance compared with other types of funds in times of financial crisis. Besides, investment in ETFs enables investors to increase the liquidity of their component stocks during non-turbulent market times. On the contrary, ETFs can be no guarantee to protect investors from financial distress. The Flash Crash is the best illustration of this.

3. ETFs Background

This part provides general information about ETFs. It first describes the history of ETFs, then gives definition and types, the buy and sell mechanism with participants, and some characteristics of ETFs.

The consequences of the stock market crash which was marked on October 19th, 1987, well-known as Black Monday, pushed SEC to ask for a single investment product that could be traded as a whole basket of securities on stock exchanges. The first product meeting this requirement, S&P 500 SPDR (Standard & Poor's Depositary Receipt), was introduced by SPDR on January 23rd, 1993 to seek and track the S&P 500 index (Liebi, 2020). It is also the highest valued ETF in the world with a market capitalization of about 362.3 billion dollars as of April 28th, 2021 (Statista Research Department, 2021). After ETF became popular in the US, it has started to be officially traded in European stock exchanges in 2000.

According to Zopounidis et al. (2010), ETF is one of the main types of investment funds and is a kind of open-end fund. It is designed to replicate the performance of the underlying assets. Initially, ETFs are established to replicate only broad-based stock indices. To meet the increasing demand of the investment market, providers develop new kinds of ETFs, from equity indices to others like fixed-income instruments and commodities, from the whole industry to a specific one, from one country to a group of countries or the whole world, from original to leveraged ETFs, etc. Recently ETFs have increased strongly in the assets under management, market significance, and diversity.

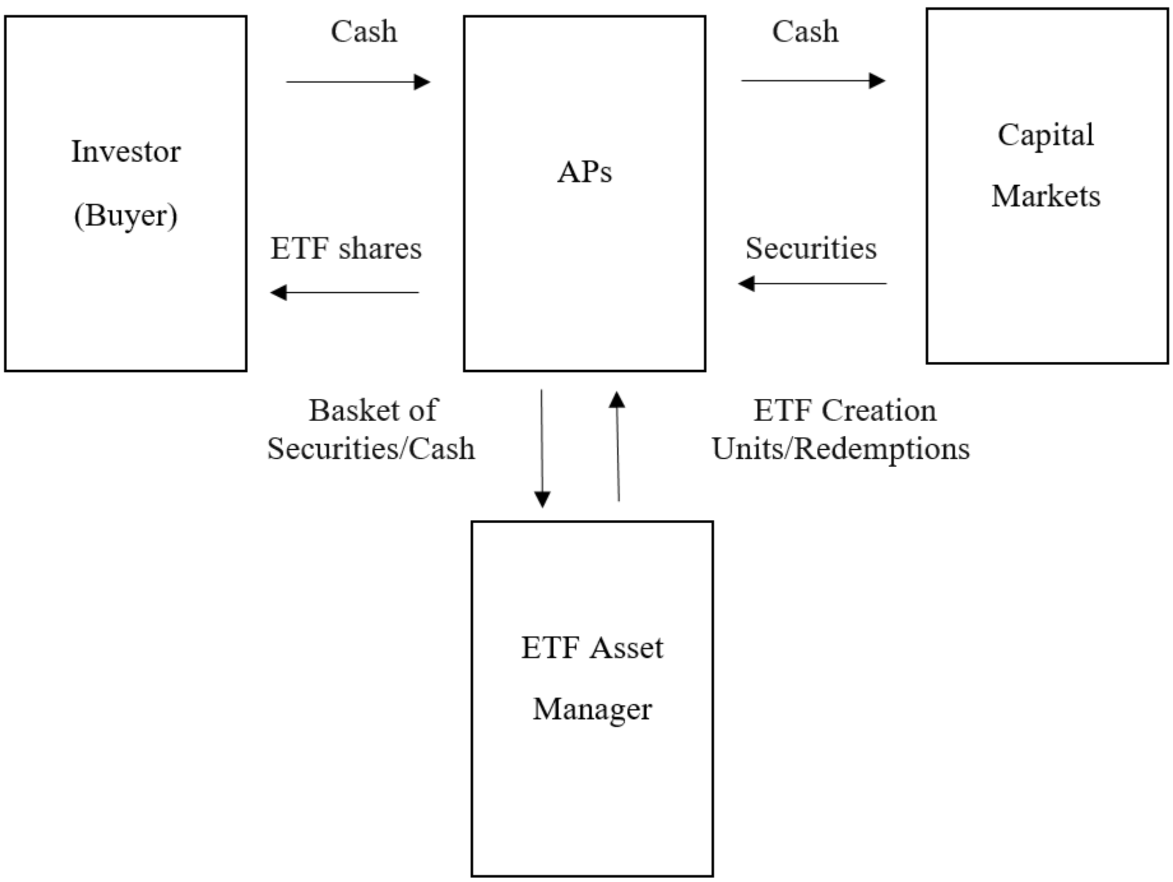
Nowadays, equity ETFs still account for the largest proportion of the ETF product assets but other asset classes like bond ETFs and commodity ETFs have also become more popular recently. Bond ETFs are not only portfolios of investment-grade and government bonds but are also created based on high-yield bonds and even bank loans. The reasons for the rapid growth of bond ETFs are diversified. First, they are traded intraday on stock exchanges with lower bid-ask spreads in comparison to their underlying bonds. Second, their bid and offer quotes are readily available thanks to transparency characteristics. Third, while fixed-income ETFs often have better liquidity and diversification, many individual bonds are traded infrequently and illiquid. Fourth, investors do not need to pay attention and trade bond ETFs constantly like their underlying assets. Commodity ETFs, which are often known as a hedge against inflation and a way of diversification, also gains the interest of investors despite their decreasing role since 2013 when prices of many commodities drop significantly. Most commodity ETFs must be invested indirectly via futures contracts, except for certain precious metals like gold. The reason is that the storage costs of commodities will force expenses of commodity ETFs to increase strongly. Since commodity ETFs provide exposure via futures contracts, they do not need to always reflect spot returns (Lettau & Madhavan, 2018).

ETF, unlike a traditional mutual fund, does not interact directly with capital markets. Instead, when ETF providers want to create new shares for their funds, they will sign a legal contract with one or several authority participants (APs) who can be market makers, specialists, or any other large financial institutions with buying power and who directly interact with the markets (Figure 1). ETF providers can buy or redeem ETF shares with APs in large blocks, which are called creation units. In turn, the APs exchange with ETF providers either a basket of underlying securities that the ETF tries to replicate or cash. The exchange takes place based on a fair value basis because the ETF shares and the basket of securities are equally valued at the NAV. This mechanism adjusts the number of ETF shares based on the relationship between supply and demand and it is known as the creation-redemption process, in which 'creations' mention an increase in the supply of ETF shares and 'redemptions' mention a decrease in the outstanding shares of the ETF. It also ensures the price to be strictly followed by the NAV (Lettau & Madhavan, 2018).

Both the current shares and the basket of underlying securities, which are accepted for creations or redemptions on the next business day, are published at the end of every trading day. Transactions between ETF providers and APs are either 'in-kind' or cash. Both sides gain benefit from this exchange since ETF providers have the underlying assets or ETFs they need and APs have ETF shares to sell or cash. Since ETFs can be bought and sold like a stock in stock exchanges during normal market hours, APs can buy or sell ETF shares in the secondary market. They can also purchase or redeem shares directly from ETF providers if they see opportunities to make a profit.

Figure 1: The creation-redemption process of ETFs

This figure describes how ETF providers transact with APs in a creation-redemption process and how different parties act in the ETF market. The figure is created based on research of Lettau and Madhavan in 2018.



As a hybrid instrument, ETFs take advantage of both open-ended funds and closed-ended funds. They can be not only traded intradaily in the stock exchanges, increasing their liquidity, but also allow creation and redemptions, avoiding big discounts or premiums due to excess offer or demand. Because of this specific creation and redemption feature, ETFs are often known as a more cost-effective investment tool in comparison to conventional mutual funds. According to Martin Lettau and Madhavan (2018), ETFs provide a better tax efficiency because of in-kind transfers to decrease capital gains distributions. They also have a lower cost with lower management expenses since their expenses are externalized, and a higher level of transparency since their holdings are published daily and investment strategies are specified in advance instead of quarterly compared to active mutual funds. Unlike traditional mutual fund investors, ETF investors can short-sell, lend shares, and buy on margins as stocks.

The ETF's total cost is known as the total expense ratio (TER) which measures total costs relating to the management and operation of an investment fund such as trading fees, legal fees, audit fees, marketing fees, and other operational expenses. TER is also considered as the after-reimbursement expense ratio since ETF investors use this ratio to determine if an ETF is a suitable investment after fees are taken into account. However, in the ETF world, TER does not show all costs of owning an ETF but the total cost of ownership (TCO) does. TCO contains TER, internal costs that are indispensable to run any investment fund and are missed by TER, and various external costs so that an ETF with the lowest TER is not necessarily the cheapest one. Internal costs consist of dealing fees, spreads, taxes, or swap fees which are incurred in the ETF's underlying holdings. External costs are platform charges, dealing fees and the bid-offer spread that investors must pay when trading ETFs. Unlike internal costs, external costs are relatively transparent because platform charges and dealing fees must be disclosed on the website of platforms and the bid-offer spread can be checked via prices of ETFs. Since there is no standard definition of TCO, investors cannot find TCO on any website or factsheet. Meanwhile, TER, the estimated annual cost of owning an ETF, can be easily obtained by a quote on a product's website or the Key Investor Information Document, resulting in the spread use of TER (Riedl, 2015).

ETFs also have some drawbacks. Martin points out that due to the lack of a common classification frame, some investors may not have enough financial knowledge to distinguish different types of ETFs. Intraday liquidity might also pose too much trading, leading to lower returns. Then, their liquidity might result in a similar effect as that of less-liquid mutual funds. Due to these disadvantages and the passive nature of ETF, some investors still prefer mutual funds. To meet the requirements of those investors, ETF providers introduced the first actively managed ETF in the US since 2008, which replaces the underlying market portfolio index with a specially designed one (The Balance, 2019). However, that tracking index often lacks the authoritative status and is still difficult to meet different needs of the market.

4. Data Selection and Methodology

The following chapter describes the data selection and the empirical methods used in the thesis. First, the data and its criteria are presented. Then, portfolio strategy and portfolio valuation are discussed. The strategy is to construct and evaluate two value-weighted portfolios, ETFs and underlying assets, under the Mean-Variance portfolio optimization with and without short-selling constraints, and then compare value-weighted ETF portfolio with the market portfolio using Fama and French three-factor model. The portfolio valuation uses four financial metrics, namely, Jensen's performance index, Sharpe ratio, Sortino ratio, and Information ratio.

4.1 Data Selection

The objective of this paper is to construct portfolios satisfying three main criteria, namely, high return in long-term growth, a safe tool during the recession, and protection during times of stagflation. Thus, the target portfolios consist of three asset classes: equity, bond, and commodity. The S&P 500 index is chosen for equity class since this thesis focuses on the US market to create a diversified portfolio with very liquid securities. Ten-year US treasury is chosen for bond class since this bond acts as a benchmark that guides other interest rates and it is also seen as a sign of investor sentiment about the economy. The S&P GSCI is chosen for the commodity class since there is no investable commodity index for the US market and the US commodity market also fluctuates along with the general trend of the world commodity market. Thus, the data used for the underlying asset portfolio includes three assets, namely, S&P 500 future, ten-year US treasury, and S&P GSCI index.

Given these three main criteria, the study sets up an ETF portfolio with three corresponding ETFs, namely, iShares Core S&P 500 UCITS ETF, iShares 7-10 Year Treasury Bond ETF, and iShares S&P GSCI Commodity Indexed Trust (GSG) indices. Since there is no ETF replicating only ten-year US treasury and this Bond ETF fluctuates in the same trend as the US treasury, 7-10 Year Treasury Bond ETF is used. The various qualitative and quantitative criteria are applied to choose corresponding ETFs. The most important criteria are:

1. The fund must have assets under management over 1 billion USD for high liquidity.

2. ETFs' age must be over 10 years to cover both upward and downward periods of the economy.
3. The total returns of ETFs which assume all distributions are reinvested and give the complete picture of ETF's performance must historically be less than returns of underlying assets at most 0.2% for equity and bond and 1% for the commodity over ten years.
4. ETFs must be very liquid with an average bid-ask spread within 60 days (the markup on the price) no more than 0.02% for equity and bond and 0.1% for commodity and average daily trading volume of more than 500 million dollars for equity and bond and 20 million dollars for commodity.
5. ETFs must strictly follow the underlying asset with a median tracking difference within 12 months of no more than 0.1% for equity and bond and 0.5% for commodity.
6. With all above-mentioned criteria being satisfied, the thesis chooses ETFs with the lowest TER.
7. The ETF funds must clearly present important information, policies, and investment strategy and show the financial stability throughout the researched period.

The empirical results are based on a big amount of data. All price series used for two portfolios are from a secondary source, namely, investing.com (US S&P 500 future index) and DataStream (other securities). Data on the factor portfolio from the three-factor model (risk-free rate, market return, value factor, and size factor) are collected from the data library of Kenneth R. French. Due to costs incurred from obtaining primary data and the length of data history, we must rely on the secondary data. Besides, data must be collected from different sources because no database provides all necessary data. Both DataStream and the data library of Kenneth R. French are highly reputed and reliable, which are frequently used by academic researchers to gather information for empirical studies.

All data are collected based on a daily basis in which the prices are adjusted-closed and denoted in US Dollar. Because securities have different starting days, the starting date for our data analysis is on 1 August 2006. The ending date is set to be on 26 Feb 2021 which is the latest data available when the study is conducted. Therefore, the data set contains 3644 daily observations, corresponding to 175 months. The whole period covers the financial crisis from 2007 to 2008, the European debt crisis from 2008 to 2012, and the Covid-19 downturn from 2020, providing the comparison of ETFs and their underlying assets during the downward period.

4.2 Methodology

In this section, the thesis first presents the method used for comparing the performance of ETFs with their underlying assets in the whole testing period and different financial distresses. The thesis then describes methods used for constructing the optimal portfolios and for evaluating performance of portfolios.

4.2.1 Risk Analysis of ETFs and Underlying Assets

In this part, the thesis shows the methodology that is used to capture the volatilities clustering patterns for individual ETFs and corresponding underlying assets. We compare time-varying variances of ETFs with their underlying assets to see how the risk of assets changes during normal and financial downturn periods. The data set includes the financial crisis from 2007 to 2008, the European debt crisis from 2008 to 2012, and the Covid-19 recession from 2020.

The mean equation of Garch (1, 1) is defined as:

$$r_t = \mu + \eta_t$$

$$\eta_t \sim N(0, \sigma_t^2)$$

Garch (1,1) conditional variance equation is defined as:

$$\sigma_t{}^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

The given Garch (1, 1) model describes the conditional variances based on the one lagged squared error and one lagged variance, with the sum of $\alpha + \beta < 1$ condition for stationary. Besides, ω, α, and β have to be positive to ensure the positive variance.

To obtain the maximum log-likelihood estimates of the Garch (1, 1) (ω, α, β, and μ) log-likelihood function needs to be maximized (Siaw et al. 2015).

$$L\,(\mu, \omega, \alpha, \beta) = \sum_{t=1}^{T}(-\frac{1}{2}ln(2\pi) - \frac{1}{2}ln(\sigma_t^2) - \frac{\eta_t^2}{2\sigma_t^2})$$

4.2.2 Portfolio Strategy

Under the Mean-Variance framework, Markowitz assumes that the preference of investment only depends on its mean and variance and that all investors are rational, risk-averse, and have homogenous behavior. Regarding these assumptions, he argues that all rational individuals will hold the market asset, the asset combination that yields Sharpe ratio, and the risk-free asset. Investors have unlimited access to borrow and lend money at a risk-free interest rate so that the concrete share depends on the individual degree of risk aversion. Since the combination of risky assets can be decided by an outsourced external portfolio manager, individual investors only have to add several risk-free assets that match their risk preferences (Bodie et al. 2009). Removing the risk-free assets from the portfolio returns also allows investors to better isolate the profits associated with risk-taking activities. Then the role of the portfolio manager is to find the combination of assets that yields the highest Sharpe ratio. To maximize the Sharpe ratio, the portfolios are set up as follows:

The portfolio expected excess returns are defined using the excess returns of portfolio components and their weights:

$$E[r_p] - r_f = w'm$$

Portfolio Variance:

$$\sigma_p^2 = w'\Omega\, w$$

w = a vector of weights (dimension: N × 1)

m = a vector of expected excess returns (dimension: N × 1)

Ω = the covariance symmetric matrix (dimension: N × N)

Then maximize Sharpe ratio to find the optimal combination of risky assets, given that the total weight of component assets equals one (Danthine et al., 2015):

$$S_{\,p}\,(w) = (w)\ \frac{w'm}{\sigma_p}$$

$$\text{s.t } w'1 = 1$$

The thesis uses a moving window method on a monthly basis to calculate the expected excess returns. First, the excess returns and covariance matrix of components of each portfolio are calculated monthly based on daily observations. Then, with and without short-selling constraints are applied when rebalancing the portfolio monthly to see the effect of both long and short positions on the rebalance of the portfolio.

Without short-selling constraint, the monthly optimal holdings are calculated as:

$$w_S = \frac{\Omega^{-1}m}{1'\Omega^{-1}m}$$

With the short-selling constraint, the weights of components of each portfolio are now subjected to be non-negative. To rebalance the portfolio monthly, the weights of components of every month are initially set equal (one-third) and then found out based on the numerical optimization method to maximize Sharpe ratio, given the above-mentioned conditions.

Given optimal holdings, in-sample and out-of-sample Sharpe ratios of each portfolio are calculated to compare the performance of the ETF portfolio and the underlying asset portfolio. In-sample portfolio excess returns and standard deviation are calculated by optimal weights with excess returns or covariance in the same month. Meanwhile, out-of-sample expected excess return and standard deviation are calculated by optimal weights this month with excess returns or covariance next month. The reason is that the in-sample Sharpe ratios of two portfolios represent their historical performance while out-of-sample Sharpe ratios show expected performance. Therefore, the thesis uses the in-sample Sharpe ratio to evaluate the past performance (ex-post) and uses expected portfolio performance and expected the risk-free rate to calculate an estimated out-of-sample Sharpe ratio to evaluate the expected performance (ex-ante).

4.2.3 Portfolio Evaluation

This section presents the methodologies used to evaluate the performance of the ETF portfolio and the market portfolio. First, the thesis describes the Fama and French three-factor model which is applied to analyze how much the market portfolio and other factors, namely size factor and value factor, can explain the ETF portfolio. Second, various financial metrics which are used to evaluate the performance of portfolios are shown.

4.2.3.1 Fama and French three-factor model

Fama and French three-factor model is the continuation of the capital asset pricing model (CAPM). The model improves the explanatory power of the multifactor model considerably. The study uses the model to decompose the total excess return of the ETF portfolio with and without short-selling constraints to achieve how much the return is explained by these factors. The portfolio's sensitivity is measured by three-factors: Excess Market return, Size factor, and Value factor. The estimated model is:

$$r_p - r_f = \alpha_p + \beta_{1,p}(r_m - r_f) + \beta_{2,p}SMB + \beta_{3,p}HML + \varepsilon_p$$

$r_p - r_f$ = the excess return of the ETF portfolio

α_p = the regression intercept to be referred as "Jensen's alpha".

$r_m - r_f$ = the excess return of the market portfolio

SMB (small minus big) = the size factor

HML (high minus low) = the value factor

4.2.3.2 Performance Metrics

To see the performance difference between the ETF portfolio and the market portfolio, appropriate evaluation metrics are used. Concerning the relationship between return of the ETF portfolio with return of the market portfolio, Jensen's Performance Index and Information Ratio are used. Considering the risk and return trade-off of portfolios, Sharpe ratio and Sortino ratio are calculated.

Jensen's Performance Index

Jensen's performance index, which is also known as Jensen's alpha, is a risk-adjusted performance measure that represents the average return on a portfolio or investment, above or below that predicted by CAPM, given the portfolio's or investment's beta and the average market return. This metric is also commonly referred to as alpha. It is the intercept, α, of the following model:

$$E[R_p] - R_f = \alpha_i + \beta_p(E[R_m] - R_f)$$

α_i is the value that captures overpricing or underpricing. We assume that α_p = 0 for the standard CAPM model. In general, to accurately analyze the performance of a portfolio, we must look at not only its overall return but also its risk to see if the investment's return compensates for its risk. Jensen's measure is one of the ways to determine if a portfolio is earning the proper return for its level of risk. Positive alpha means that the investment earns more than the risk-adjusted returns, while negative alpha indicates that the investment has not earned its required return.

Sharpe Ratio

The Sharpe ratio is a commonly used measure of risk-adjusted return and defined for portfolio p as:

$$SR_p = \frac{E[R_p] - R_f}{\sigma_p}$$

This ratio is the slope of the line from riskless asset to portfolio p and the tangency portfolio is the portfolio with the maximum Sharpe ratio. Sharpe ratio has become one of the most popular methods to calculate the risk-adjusted return. In general, the greater the value of Sharpe ratio is, the better the portfolio is. A high Sharpe ratio is good when compared to similar portfolios or funds with lower returns. Sharpe ratio is negative if the risk-free rate is greater than return of the portfolio or the portfolio's return is expected to be negative. A negative Sharpe ratio does not have any useful meaning.

Sortino Ratio

Sortino ratio measures the expected excess return per unit of downside risk (DR). The downside risk is a measure of the deviation of observations from the mean, only for values below the mean. It is calculated as the square root of semi-variance. Sortino ratio is defined for portfolio p as:

$$SO_p = \frac{E[R_p] - R_f}{DR_p}$$

The Sortino ratio improves upon the Sharpe ratio by isolating negative volatility from total volatility. It is a useful financial metric for investors to evaluate return of an investment for a level of bad risk. Because it only focuses on the downside risk, some investors think it brings a better view of the risk-adjusted performance of a portfolio since upside risk is considered as a benefit.

Information Ratio

The information ratio is a measurement of portfolio return beyond the return of a benchmark, usually an index, compared to the volatility of those returns. The benchmark used is typically an index that represents the market or a particular sector or industry. The information ratio is given by the following formula:

$$IR_p = \frac{E[\theta_{pt}]}{ste(\theta_{pt})} = \frac{\alpha_p}{\sigma_{\varepsilon_p}}$$

θ_{pt} = the residual returns for portfolio p.

Then Information ratio is the ratio between alpha and the standard error of the residual which is also called ‘tracking error’. The ‘tracking error’ can be calculated by taking the standard deviation of the difference between the portfolio returns and the index returns. Information ratio and Sharpe ratio both are indicators of risk-adjusted returns. However, while Information ratio measures the risk-adjusted return to a benchmark, Sharpe ratio measures the excess return between the asset's return and the risk-free rate of return. The Information ratio also measures the consistency of an investment's performance but Sharpe ratio measures how much an investment portfolio outperformed the risk-free rate of return on a risk-adjusted basis.

5. Results and Analysis

The following chapter presents the empirical findings of the thesis and analyses the results. The main results are presented in three sections. First, the descriptive statistics for ETFs and underlying assets are analyzed and discussed. Second, the results of the portfolio strategy are shown. Finally, the results of the portfolio evaluation using the Fama and French three-factor model and other financial metrics are described and analyzed.

5.1 The comparative analysis of ETFs and Underlying Assets

Table 1 below describes two main characteristics of financial assets: mean and standard deviation. We primarily compare ETFs with corresponding underlying assets individually according to return and risk and then examine their performance during the time of financial distress.

Table 1: Performance Metrics for individual ETFs and underlying assets

This table reports the performance metrics: mean and standard deviation of the individual ETF stocks and underlying assets based on daily observations as a whole period, as well as during the financial downturn periods. The whole period includes the financial data from August 2006 to February 2021. The financial crisis starts from January 2007 to December 2009, the European debt crisis from January 2008 to December 2012, and the Covid-19 recession from January 2020 to February 2021.

	Period	**SP**	**SP ETF**	**Bond**	**Bond ETF**	**GSCI**	**GSCI ETF**
Mean	Whole Period	7.56%	7.58%	2.49%	2.47%	-0.30%	-8.79%
	Financial Crisis	-23.28%	-22.81%	10.40%	8.99%	-10.94%	-16.98%
	European Debt	-0.80%	-0.50%	4.66%	4.26%	1.17%	-9.59%
	Covid-19	13.18%	13.23%	5.75%	4.65%	5.56%	-13.55%
Standard Deviation	Whole Period	20.70%	20.31%	7.63%	6.70%	23.81%	24.03%
	Financial Crisis	32.23%	30.13%	9.35%	8.20%	32.34%	33.43%
	European Debt	26.68%	25.75%	9.64%	8.37%	28.61%	29.64%
	Covid-19	32.32%	32.88%	8.24%	7.06%	32.62%	31.50%

There can be seen different tendencies with regard to asset classes. To start with S&P and S&P ETF, it is noticeable that they are quite similar to each other in which S&P ETF outperforms S&P

slightly, both in terms of return and risk for all periods, except for the time of Covid-19 when the risk for S&P ETF is 0.5 basis point higher than that for S&P. Bond and Bond ETF expose the traditional risk-return tradeoff in which the higher the return is, the higher the risk is. In this case, the Bond has a greater return with higher risk, while the Bond ETF promises lower risk. It is noteworthy that small numerical differences are observed between performance measures of Bond and Bond ETF similar to the S&P and S&P ETF case. GSCI and GSCI ETF attract special attention when it comes to their performance metrics since both show very poor performance, negative return, and high risk. This result can be motivated given the fact that commodities belong to an asset class that is negatively correlated with other types of asset class, such as stocks and bonds signifying that when stocks and bonds increase in value, commodities will decrease in value, and vice versa. Another possible explanation is that since 2013, prices of many commodities fell strongly, resulting in negative returns in the test period. From Table 1, it appears that GSCI ETF reveals much higher negative returns than GSCI in some periods. Unlike S&P 500 and Bond ETF, GSCI ETF is not a standard ETF. Its shares are not subject to the same regulatory requirements as mutual funds. Investments in this ETF are speculative and involve a high degree of risk. Therefore, GSCI ETF is affected more strongly by crises than GSCI which tracks commodity, a protection during stagflation.

The second analysis is conducted based on financial distress periods, which are also represented in Table 1. As it seems, the financial crisis has the most severe effect (higher negative return) on the performance of S&P and S&P ETF. This finding on S&P ETF contradicts the argument from the literature review, that ETFs have not been hit during the financial crisis (Fuhr 2009). S&P ETF tracks S&P, thus such a trend of S&P ETF is not surprising. On the other hand, Bond and Bond ETF remain remarkably stable during the financial crisis period. Even more, they gain quite high returns compared with other classes of funds. This result is reasonable as long as they belong to risk-free securities and bonds are known as the hedging tool during recessions. GSCI and GSCI ETF also experience negative returns and a high level of risk in that period but relatively less negative return compared to S&P and S&P ETF. The European Debt Crisis brings unfavorable outcomes to the performance of stocks' negative returns and high risks. During that period GSCI outperforms GSCI ETF regarding positive return and relatively smaller risk. The recent crisis of Covid-19 results in higher volatilities of funds for S&P as well as GSCI indices.

The performance of the underlying assets and their corresponding ETF can be examined from the perspective of their prices. Figure A1 and Figure A2 show that S&P ETF and Bond ETF replicate underlying assets S&P and Bond respectively. Their performance is almost the same which is consistent with the Mean-Standard deviation analysis discussed above. Thus, investors can choose either underlying assets or corresponding ETFs since they produce a similar performance. However, the lower cost and tax effectiveness of ETFs emphasizes the superiority of ETFs compared with underlying assets. Figure A3 demonstrates that the price fluctuation of GSCI and GSCI ETF follows the same trend, albeit there can be noticed the underperformance of GSCI ETF in comparison to GSCI. This finding calls into question the choice of GSCI ETF by the investors as an individual fund but having it in the portfolio offers a hedge and can be useful.

To capture clustering patterns of high volatility periods, the Garch (1,1) model is used. We obtain the parameters for time-varying variances for the individual ETFs and the underlying assets to see how the risk of assets changes during normal and financial downturn periods.

Table 2: Garch (1,1) Parameters for individual ETFs and underlying assets

The table presents optimal parameters (ω, α, β, and μ) of Garch (1,1) model for individual ETFs and underlying assets. The coefficient of α represents the effects of the new information on stock volatility. The coefficient of β describes the persistence of the volatility. The coefficient α + β shows the overall measurement of persistence of volatility. Parameters are obtained through the Maximum Likelihood estimator.

Assets	**$\mu * 10^{-2}$**	**ω**	**α**	**β**	**α + β**	**Log *L***
S&P	0.08	0.00	0.19	0.80	0.99	15,166.66
S&P_ETF	0.08	0.00	0.17	0.82	0.99	15,198.59
Bond	0.08	0.00	0.06	0.93	0.99	17,969.49
Bond_ETF	0.08	0.00	0.07	0.92	0.99	18,432.89
GSCI	0.08	0.00	0.07	0.92	0.99	14,015.85
GSCI_ETF	0.08	0.00	0.07	0.92	0.99	13,964.01

The values for constants are very close to zero in the case of both ETFs and underlying assets. Alpha and beta parameters describe the effect of past shocks and variances on the future conditional variance. For the daily data alpha usually ranges between about 0.05 to 0.1. The closer the parameter is to 0.1, the more volatile the market is. Alpha parameters for S&P and S&P ETF are greater than 0.1, indicating an extremely fluctuating, unstable market. For other funds, alpha parameters are between 0.06 and 0.07. Parameter beta defines the smoother of conditional variance, higher beta results in higher persistence and smoother conditional variance series. Parameter beta typically ranges from 0.85 to 0.98. S&P and S&P ETF have values of betas less than 0.85. The lower betas result in lower persistence of past observations, less smooth series. Summing up two coefficients for each fund individually, we obtain a persistence level for all funds, each is equal to around 0.99. These results indicate that conditional variance is highly persistent long, in other words there exist lasting periods of high volatility.

In addition to evaluating the parameters of Garch (1,1), there are provided three figures depicting volatility clustering graphically. As stated above, S&P and S&P ETF have intensely non-smoothing series illustrating in Figure A4. There also can be seen the effects of financial downturns mentioned above. It's also noteworthy to highlight the fact that S&P ETF fluctuates with less magnitude compared to S&P itself. The volatility clustering for bond and bond ETF is relatively smoother, however, there also can be seen the effects of the Financial crisis, European Debt, and Covid-19 (Figure A5). GSCI and GSCI ETF reveal a similar pattern of volatilities clustering with comparatively smooth series (Figure A6).

5.2 Portfolio strategy

With regard to the historical performance, it can be seen from Table 3 that all portfolios have high in-sample Sharpe ratios at more than 2, especially the underlying asset portfolio without short-selling, showing that these portfolios perform well in the testing period and offer high excess returns relative to their volatility. Specifically, ETF portfolio always underperforms the underlying asset portfolio with and without short-selling constraints. This is not unexpected because ETF gains investors' interest thanks to its cost-efficiency and high liquidity characteristics which the thesis does not take into account. Table 3 also shows that Sharpe ratios of ETF and underlying asset portfolios are both smaller when short-selling is allowed. A possible explanation is that more than half of the testing period is financial downturns and this leverage tool exacerbates high volatility of risky assets, leading to lower Sharpe ratios in these periods and then the whole testing period.

Table 3: In-sample and out-of-sample Sharpe ratios of ETF and underlying asset portfolio

This table shows in-sample and out-of-sample Sharpe ratios of the ETF portfolio and the underlying asset portfolio with and without short-selling for the testing period from August 2006 to February 2021 on a yearly basis.

	Condition	ETF portfolio	Underlying asset portfolio
In sample	With short-selling	2.17	2.62
	Without short-selling	2.81	3.43
Out of sample	With short-selling	1.61	1.34
	Without short-selling	-0.23	1.64

Considering the expected performance via out-of-sample Sharpe ratios, it is noticeable that ETF portfolio outperforms the underlying asset portfolios when short-selling is allowed but underperforms in the rest case. The outperformance of ETF portfolio is mainly because of using short-selling reasonably as an efficient hedging tool. Specifically, when the market has a downward trend and price of stock goes down, selling stock today and buying it in the future make it profitable for investors. Short-selling in this case increases profit and reduces loss for investors, leading to higher excess return than the case that does not use short-selling. Since more than half of the testing time is downturn periods, the short-selling strategy helps the ETF portfolio with short-selling have higher excess return and then higher Sharpe ratio than a portfolio without short-selling. This outperformance also points out that despite not taking the cost-efficiency and high liquidity advantages into account, ETF portfolio can still outweigh the underlying asset portfolio if a reasonable portfolio strategy is applied.

While the out-of-sample Sharpe ratios of ETF and underlying asset portfolios are quite close to each other in case short-selling is used, those of the rest case have a big difference. Specifically, due to high negative returns in most of the months in the testing period, the Sharpe ratio of ETF portfolio without short-selling is negative at -0.23 which does not convey any useful meaning, and Sharpe ratio of underlying asset portfolio without short-selling is positive at 1.64. A possible explanation of this difference is that the data is changing every day, especially in downturns, and the turmoil within the entire economy makes securities fluctuate dramatically every day. Therefore, rebalancing the portfolio next month based on the optimal holding this month poses a bigger loss for the ETF portfolio but a higher profit for the underlying asset portfolio in downturn periods. This also explains why out-of-sample Sharpe ratios of two portfolios are lower than in-sample Sharpe ratios.

5.3 Portfolio evaluation

5.3.1 Fama and French three-factor model

After the construction of ETF portfolios, with short selling and without short selling, we are comparing their performance with the market portfolio using Fama and French three-factor model. Another motivation for using Fama and French model is to determine ETFs' ability to capture factors' premiums. For the comprehensive analysis, it would be better to investigate individual ETFs ' relationship with the market first. The explanatory variables are calculated monthly. The parameter estimates from the regressions are presented in Table 4 below.

Table 4: Regression output (Fama and French three-factor model)

The table shows regression outputs for individual ETFs and ETF portfolios with short-selling and without short-selling. β_{mkt-rf} is the factor coefficient of excess return on the market portfolio. β_{hml} is the factor coefficient of high book-to-market stocks minus low book-to-market stocks. β_{smb} is the factor coefficient of return of small stocks minus the return of big stocks. R^2 describes how well the regression model fits the observed data. α is the component of fund returns that cannot be explained by factors exposure. The table also includes p (significance level of coefficients) values which are given in parenthesis.

	S&P ETF	Bond ETF	GSCI ETF	ETF Portfolio Short-selling	ETF Portfolio No Short-selling
α	0.00	0.00	-0.01	0.02	-0.01
	(0.80)	(0.16)	(0.03)	(0.27)	(0.12)
β_{mkt-rf}	0.84	-0.07	0.65	-0.06	0.38
	(0.00)	(0.00)	(0.00)	(0.83)	(0.00)
β_{hml}	-0.10	-0.07	0.27	0.58	0.16
	(0.07)	(0.14)	(0.21)	(0.34)	(0.19)
β_{smb}	0.12	-0.13	0.29	-0.78	-0.05
	(0.00)	(0.00)	(0.07)	(0.08)	(0.58)
R^2	0.87	0.18	0.28	0.02	0.24

The thesis starts the discussion with ETF's exposure to the market. S&P ETF reveals greater exposure to the market with a beta value of 0.84 that measures the risk of an S&P ETF compared to the systematic risk of the entire market. A negative beta value in the case of Bond ETF is not surprising since it is risk-free security. GSCI ETF also delivers positive beta which also seems reasonable due to the riskiness of commodity. However, all beta values are less than one, indicating all securities volatile less than the market itself. The coefficients are highly significant.

On the side of portfolio return, the constructed ETF portfolio without short-selling exposes lower beta values than individual S&P ETF and GSCI ETF indicating owning ETF portfolio results in lower risk rather than individual ETF securities. This finding is meaningful and statistically significant. ETF portfolio with short-selling has even lower beta but it is insignificant showing that the market portfolio cannot explain the ETF portfolio with short-selling constraint. The beta parameters for value factor (HML) are almost zero for S&P ETF and Bond ETF with the evidence of insignificance level. GSCI ETF has a relatively higher beta coefficient but still not significant. None of the ETF portfolios, with and without short-selling, manifest robust parameters for value factors. The sensitiveness of the individual ETFs to the size factor (SMB) is not powerful due to lower beta values. SMB ETF portfolio with short selling reveals a negative correlation with size factors which is significant at 10%. The explanation behind these results is that selected ETFs securities are large-cap stocks, consequently they show underperformance which is consistent with Fama and French findings. By contrast, an ETF portfolio without short selling is very close to zero and insignificant. The final row in the tables shows that all R^2 are very low except S&P ETF, which suggests that the standard three-factor model does not fit the data well and can't explain most of the total return variability. R^2 for S&P ETF is 0.873, which is an acceptable level. It's noteworthy, only S&P ETF has all three factors significant at 10% which is not surprising since it is one of the most commonly followed equity indices in the US market.

5.3.2 Jensen's Performance Index

There are presented the differential return for all individual ETFs and ETF portfolios in Table 5. Jensen's alpha is the component of fund returns that cannot be explained by factors exposure. It may have come from fees, costs, security selection, market timing, or exposure to a factor that is not included in the three-factor model. A positive value of Jensen's alpha indicates that a fund can generate a higher value for the investor relative to the benchmark index on a risk-adjusted basis.

S&P ETF and Bond ETF generate zero alpha but not statistically significant. GSCI ETF underperforms the benchmark and it's significant at 5%. The second hypothesis of this study is to examine the relationship between ETF portfolio and the market. As a result, an ETF portfolio with short selling outperforms the benchmark, though it is not significant. ETF portfolio without short-selling underperforms the benchmark and this is not significant either. The present results are not very encouraging since this study cannot provide exact answers to the research question.

5.3.3 Sharpe Ratio

In general, all financial metrics of ETF portfolio with the short-selling constraint are negative due to negative excess return, resulting in its underperformance in comparison to the market portfolio. It can be seen from Table 5 that during the testing period, the market portfolio has a very good Sharpe ratio and also good expected excess return per unit of volatility, approximately 3, showing that the market portfolio very outperforms the risk-free rate. Sharpe ratio of the market portfolio is shown to be better than those of ETF portfolios with and without short-selling constraints. This is not unexpected since some of current literature also shows that the ETF portfolio often underperforms its benchmarks.

Table 5: Financial Metrics of ETF portfolio and the market portfolio

This table presents financial metrics, namely, Sharpe ratio, Sortino ratio, and Information ratio, of optimal ETF portfolio with and without short-selling and the market portfolio for the testing period from August 2006 to February 2021 on a yearly basis.

Metrics	ETF portfolio		Market portfolio
	With short-selling	Without short-selling	
Sharpe ratio	1.61	-0.23	2.90
Sortino ratio	0.48	-0.02	0.28
Information Ratio	0.77	-3.30	NaN

5.3.4 Sortino Ratio

Table 5 also shows that Sortino ratios of all the three portfolios have smaller absolute value than their Sharpe ratios. The reason is that Sortino ratio only considers the standard deviation of the downside risk rather than the entire risk and because the downside risk fluctuates more dramatically than the upside risk most of the time in the testing period due to financial downturns.

This difference addresses one weakness of Sharpe ratio which is the use of total risk since upside volatility is beneficial to investors and is not a factor that most investors worry about. All the three portfolios have very low Sortino ratios, suggesting that the ETF and market portfolio perform much lower than expected during the testing time and investors are not rewarded for taking on additional risk.

Unlike Sharpe ratio, Sortino ratio of the market portfolio is better than that of the ETF portfolio without short-selling but worse than that of ETF portfolio with short-selling, reinforcing the claim that the leverage tool can help the ETF portfolio perform better than its benchmark if a reasonable portfolio strategy is applied. The empirical results of Sharpe ratios and Sortino ratios in the thesis might lead to a different conclusion for the comparison between the performance of ETF portfolio and the market portfolio, which is not unusual in the real world. Using which ratio depends on whether investors want to focus on total risk or just downside deviation.

5.3.5 Information Ratio

Similarity, when it comes to Information ratio, ETF portfolio without the short-selling constraint outperforms that with short-selling constraint thanks to the efficiency of leverage. To be more specific, Information ratio of the first is considered to be good at 0.77, showing that the portfolio has high return potential and has exceeded the market portfolio about 0.77 basis point on a yearly basis, and leading ETFs to be strongly favored by investors. This high Information ratio also points out the desired level of consistency in which the ETF portfolio tries to track the performance of the market portfolio. Meanwhile, the negative Information ratio of ETF portfolio with short-selling constraint indicates that the ETF portfolio is unable to produce any excess returns at all and then that investors should eliminate this portfolio from contention.

Due to weaknesses of financial metrics that they can be interpreted differently by different investors depending on their needs, goals, and risk tolerance levels, it is better to look at many financial metrics than only one single ratio to make a more comprehensive and informed investment decision. Based on the above-mentioned financial ratios and the current literature, the ETF portfolio often underperforms the market portfolio but it can also outperform the latter if investors can make use of the leverage characteristics of short-selling and use it as an effective hedging tool.

6. Conclusion

This thesis aims to investigate the performance of ETFs in the US market. First, its objective is to analyze and obtain the performance of each ETF with its underlying asset during the whole testing period and in different financial downturns. Then, the thesis compares the ETF portfolio with benchmarks, namely, the underlying asset portfolio and the market portfolio, to research if the ETF portfolio can outperform its benchmarks.

As discussed in Chapter 5, ETFs and their underlying assets have a similar fluctuation trend in the whole testing period as well as financial distress periods in which the return and risk of two securities in the same asset class are close to each other, except for the return of commodity. The commodity securities also have the worst performance in comparison to other asset classes. When comparing the performance of the ETF portfolio with the underlying asset portfolio, the empirical results show that both portfolios perform well in the testing period with a high Sharpe ratio. While the historical performance of the underlying asset portfolio is always better than that of the ETF portfolio, the expected performance of the first is worse than the latter without a short-selling constraint but better with a short-selling constraint. The study cannot provide a straight answer for the question associated with comparing the performance of the ETF portfolio with the market portfolio since some of the empirical results of the Fama and French three-factor model are not statistically significant. However, according to previous studies discussed in Chapter 2 and based on other mentioned financial metrics, namely, Sharpe ratio, Sortino ratio, and Information ratio, the thesis could argue that although ETF portfolio often underperforms the market portfolio, the first can still outperform the latter in case the short-selling strategy is used as an effective hedging tool. All comparisons are conducted without considering the cost-efficiency and high liquidity advantages of ETF. Therefore, it is possible for the ETF portfolio to outperform its benchmarks if taking these advantages into account and a reasonable strategy is used.

The topic of the performance of ETF is broad and complex. This thesis has just scratched the surface of the topic and still has outstanding questions that need more researches to answer. First, would the results be the same or different in other stock markets like the emerging markets and in other periods? Second, since the Covid-19 time has not finished, the empirical results in this thesis might just investigate the partial effect of Covid-19 period. How would the results look like in the whole Covid-19 time? Third, after taking the advantages such as cost-efficiency, tax efficiency, and liquidity into consideration, would the ETF portfolio outperform its benchmarks or not? Taking into account all these problems, a more comprehensive result about the performance of ETF could be obtained.

References:

Ackert, L.F., & Tian, Y.S. (2008). Arbitrage, Liquidity, and the Valuation of Exchange Traded Funds, *Journal of Financial Markets, Institutions, and Instruments,* vol.17, no.5, pp.331-362.

Blakey, P. (2007). Exchange Traded Funds, *Microwave Magazine*, vol.8, no.1, pp.22-30.

Bello, Z.Y. (2012). The Relative Performance of Equity Index ETFs and Equity Index Mutual Funds, *European Journal of Finance and Banking Research*, vol.5, no.5.

Ben-David, I., Franzoni, F., & Moussawi, R. (2015). Do ETFs Increase Volatility?, *Fisher College of Business*, Working Paper Series 20071.

Blitz, D., Huij, J. & Swinkels, L. (2012). The Performance of European Index Funds and Exchange-Traded Funds, *European Financial Management*, vol.18, no.4, pp.649-662.

Borkovec, M., Domowitz, I. & Serbin,V.(2010). Liquidity and Price Discovery in Exchange-Traded Funds: One of several possible lessons from the flash crash, *The Journal of Index Investing*, vol.1, no.2, pp.24-42

Danthine, J.P., Donaldson, J. B., New York, L., & Diego San Francisco, S. (2015). Intermediate Financial Theory.

Elton, E., Gruber, M., Comer, G., & Li, K. (2002). Spiders: Where Are the Bugs?, *Journal of Business*, vol.75, no.3, pp.453-472.

Engle, R., & D. Sarkar, (2006). Premiums-Discounts and Exchange Traded Funds, *Journal of Derivatives*, vol.13, no.4, pp.27-45.

Fama, E., & Litterman, R. (2012). An Experienced View on Markets and Investing, *Financial Analysts Journal*, vol.68, no.6, pp.15-19.

French, K. (2008). Presidential Address: The cost of active investing, *The Journal of Finance*, vol.53, no 4, pp.1538-1573.

Fuhr, D. (2001). Exchange-Traded Funds: A primer, *Journal of Asset Management,* vol.2, no.3, pp.260-273.

Fuhr, D. (2009). ETF Industry Review: Barclays global investors, *Working Paper*.

Gastineau, G. (2001). Exchange-Traded Funds, *The Journal of Portfolio Management*, vol.27, no.3, pp.88-96.

Goltz, F. & Schröder, D. (2011). Passive Investing Before and After the Crisis: Investors' views on exchange-traded funds and competing index products. *Bankers, Markets & Investors (110)*.

Gittleson, J. (2019). End of Era: Passive equity funds surpass active in epic shift | Bloomberg, Available: https://www.bloomberg.com/news/articles/2019-09-11/passive-u-s-equity-funds-eclipse -active-in-epic-industry-shift, [Accessed: 16 May 2021].

Hegde, S.P. & McDermott, J.B. (2004). The Market Liquidity of DIAMONDS, Q's, and their Underlying Stocks, *Journal of Banking and Finance*, vol.28, no.5, pp.1043-1067.

Kennedy, M. (2019). Actively Managed ETFs | Thebalance, Avaliable: https://www.thebalance.com/ list-of-actively-managed-etfs, [Accessed: 17 May 2021].

Liebi, L.J. (2020). The effect of ETFs on Financial Markets: A literature review, *Financial Markets and Portfolio Management*, vol.34, no.2, pp.165-178.

Lettau, M. & Madhava, A. (2018). Exchange-Traded Funds 101 for Economists, *The Journal of Economic Perspectives*, vol.32, no.1, pp. 135-154.

Ma, Y., Maclean, L., Xu, K., & Zhao, Y. (2011). A Portfolio Optimization Model with Regime-Switching Risk Factors for Sector Exchange Traded Funds, *Pac J Optim*, vol.7, no.2, pp.281-296.

Martin, L., & Ananth, M. (2018). Exchange-Traded Funds 101 for Economists, *The Journal of Economic Perspectives*, pp.135-146.

Miffre, J. (2007). Country Specific ETFs: An efficient approach to global asset allocation, *Journal of Asset Management*, vol. 8, no.2, pp.112-122.

Miralles-Quiros, J.L., Miralles-Quiros, M.M., & Nogueira, J.M. (2019). Diversification Benefits of Using Exchange-Traded Funds in Compliance to the Sustainable Development Goals, *Business Strategy and the Environment*, vol.28, no.1, pp.244-255.

Pan, K. & Zeng, Y. (2017). ETF Arbitrage under Liquidity Mismatch, *SSRN Electronic Journal*.

Pedersen, L.H. (2018). Sharpening The Arithmetic of Active Management, *Financial Analysts Journal*, vol.74, no.1, pp.21-36.

Poterba, J.M. & Shoven, J.B. (2002). Exchange-Traded Funds: A new investment option for taxable investors, *American Economic Review*, vol.92, no.2, pp.422-427.

Riedl, P. (2015). Cost of ETFs: Total Expense Ratio (TER) vs. Total Cost of Ownership (TCO) | justETF, Available:https://www.justetf.com/uk/news/etf/cost-of-etfs-total-expense-ratio-ter-vs-total-cost-of-ownership-tco.html, [Accessed: 15 May 2021].

Richie, N. & Madura, J. (2007). Impact of the QQQ on Liquidity and Risk of the Underlying Stocks, *Quarterly Review of Economics and Finance,* vol.47, no.3, pp.411-421.

Rompotis, G.G. (2011). Testing Weak-form Efficiency of Exchange Traded Funds Market. Aestimatio, *The Ieb International Journal of Finance*, vol.2, no.2, pp.2-33.

Shin, S., & Soydemir, G. (2010). Exchange-Traded Funds Persistence in Tracking Errors and Information Dissemination, *Journal of Multinational Financial Management*, vol.20, no.45, pp.214-234.

Siaw, R. O., Dei Ofosu-Hene, E., & Evans, T. (2015). Investment Portfolio Optimization with Garch Models, *Elk Asia Pacific Journal of Finance and Risk,* vol.2, issue.2, SSN.2349-2325.

Statista Research Department (2021). Number of ETFs Worldwide 2020 | Statista, Available: https://www.statista.com/statistics/278249/global-number-of-etfs/, [Accessed: 18 April 2021].

Statista Research Department (2021). ETFs-Statistics & Facts | Statista, Available: https://www.statista.com/topics/2365/exchange-traded-funds/ [Accessed: 18 April 2021].

Svetina, M., & Wahal, S. (2008). Exchange Traded Funds: Performance and competition, *SSRN Electronic Journal*.

Smimou, K., Bector, C. R., & Jacoby, G. (2008). Portfolio Selection Subject to Experts' Judgments. *International Review of Financial Analysis*, vol.17, no.5, pp.1036-1054.

Harper, J. T, Madura, J., & Schnusenberg, O. (2006). Performance Comparison between Exchange-Traded Funds and Closed-end Country Funds, *Journal of International Financial Markets Institutions and Money,* vo16, no.5, pp.104-122.

Tiwari, A., Albulescu, C., & Yoon, S. (2017). A Multifractal Detrended Fluctuation Analysis of Financial Market Efficiency: Comparison using Dow Jones sector ETF indices, *Physica A: Statistical Mechanics and its Applications,* vol.483, pp.182-192.

Tufano, P. (2003). Financial Innovation: Handbook of the economics of finance, Chapter 6, pp.307-335.

Zhao, X., Ran, G., Shen, B., & Li., X (2021). Do ETFs Improve the Pricing Efficiency of the A-share Market Examining ETF Holdings of Individual Stocks, *Applied Economics*.

Zopounidis, C., Doumpos, M., & Pardalos, P (2010). Handbook of Financial Engineering, vol.18, pp. 67-97.

Appendix A: Tables and Graphs

Figure A1: The prices of S&P and S&P ETF

The figure represents the daily price fluctuation of S&P 500 future and iShares Core S&P 500 UCITS ETF from August 2006 to February 2021.

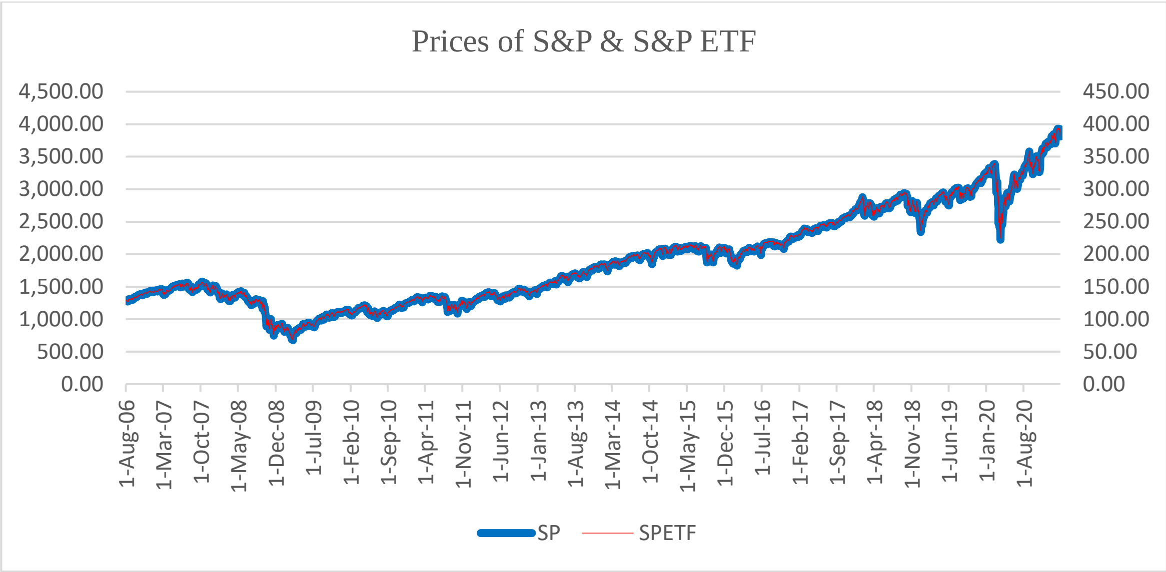


Figure A2: The prices of Bond and Bond ETF

The figure represents the daily price fluctuation ten-year US treasury and iShares 7-10 Year Treasury Bond ETF from August 2006 to February 2021.

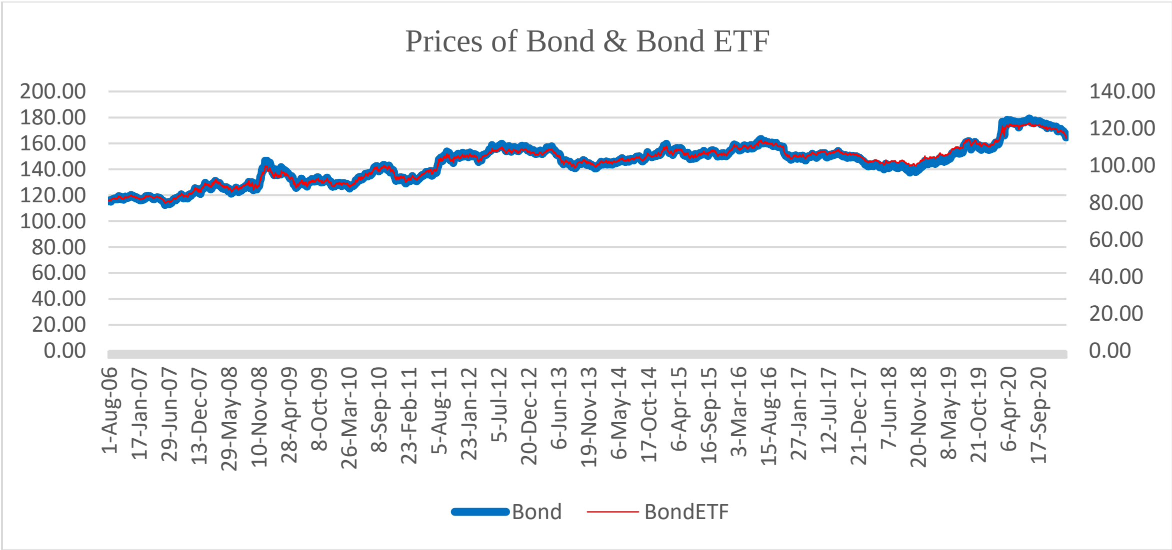


Figure A3: The prices of GSCI and GSCI ETF

The figure represents the daily price fluctuation of GSCI index and GSCI Commodity Indexed Trust (GSG) ETF from August 2006 to February 2021.

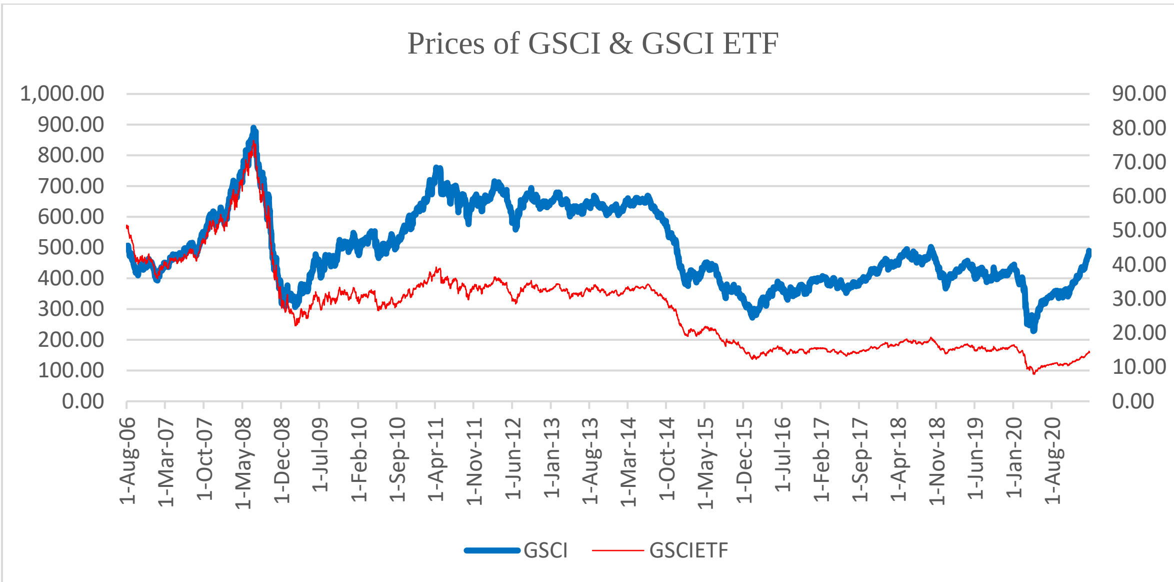


Figure A4: Garch (1,1) Model: Volatility of S&P and S&P ETF

This figure shows time-varying variance S&P 500 future and iShares Core S&P 500 UCITS ETF based on Garch(1,1) model from August 2006 to February 2021.

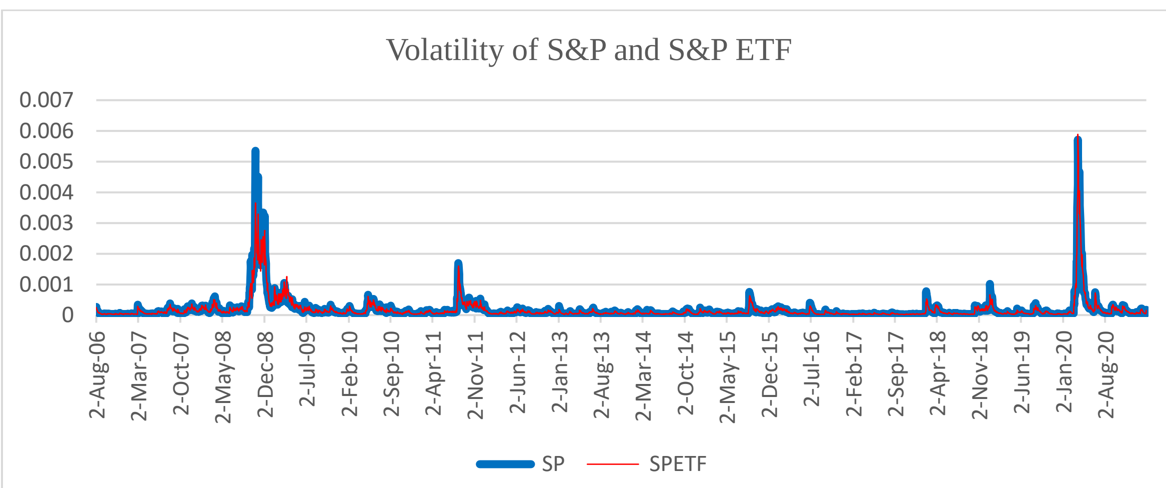


Figure A5: Garch (1,1) Model: Volatility of Bond and Bond ETF

This figure shows time-varying variance of 10-year US treasury and iShares 7-10 Year Treasury Bond ETF based on Garch(1,1) model from August 2006 to February 2021.

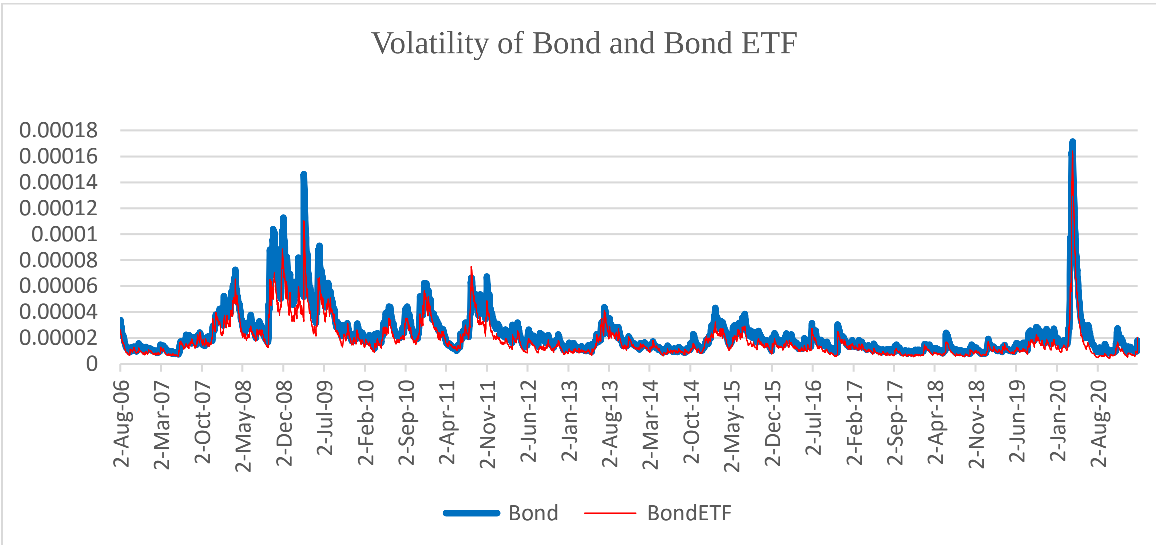


Figure A6: Garch (1,1) Model: Volatility of GSCI and GSCI ETF

This figure shows time-varying variance of GSCI index and GSCI Commodity Indexed Trust (GSG) index based on Garch (1,1) model from August 2006 to February 2021.