Investigating long-term short pairing strategies for leveraged exchange-traded funds using machine learning techniques

Elaheh Nikbakht

A Thesis

In

The John Molson School of Business

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Administration (Finance) at

Concordia University

Montreal, Quebec, Canada

August 2021

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Elaheh Nikbakht

Entitled: Investigating long-term short pairing strategies for leveraged exchange-traded funds using machine learning techniques

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Finance)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair

______________________________ Examiner
Dr. Ravi Mateti

______________________________ Examiner
Dr. Erkan Yonder

______________________________ Thesis Supervisor(s)
Dr. Thomas Walker

______________________________ Thesis Supervisor(s)

Approved by ______________________________
Dr. Nilanjan Basu Graduate Program Director

____
Dr. Anne-Marie Croteau Dean of Faculty

August, 2021

Abstract

Investigating long-term short pairing strategies for leveraged exchange-traded funds using machine learning techniques

Elaheh Nikbakht

Although the literature on leveraged exchange-traded funds (LETFs) concurs with the idea that they are short-term investment tools, recent studies offer some investment strategies for them that are also profitable in the long term. These strategies, however, are typically only tested on a limited number of highly traded LETFs. This study uses different types of LETFs to examine various portfolios with different combinations of bull and bear LETFs, to find the best investment strategy in the long run. It then uses different machine learning techniques to analyze which factors define the best investment strategies, with portfolios being rebalanced on a quarterly and annual basis. The sample of this study consists of 44 pairs of LETFs from 2012 to 2020 that have different underlying assets and leverage levels. The results reveal that short-selling the combination of both bull and bear LETFs does not yield a significant positive return compared to the market, however, the return generated from short-selling a portfolio with only bear ETFs can significantly beat the market, especially when the market is bullish. The quarterly and annual results are consistent and show that short-selling a full bear portfolio is the winning strategy in both of these intervals. Moreover, the results show that as the correlation of LETFs with their underlying index increases, the return of short-selling both bull and bear LETFs decreases. At the same time, an increase in the net asset value of bull LETFs results in an increase in the return of short-selling bull LETFs and a decrease in return of short-selling the bear LETFs.

Keywords: Leveraged ETFs, Short pairing strategy, Machine learning, Exchange-traded funds

Acknowledgments

I would like to express my greatest thanks to my dear supervisor, Dr. Walker, for all of his help during the development of my thesis. If it was not for his advice, I would not be able to complete this thesis. It was such an honor to be his student and I learned a lot from him.

Also, I wish to thank the committee members, Dr. Mateti and Dr. Yonder. I am so grateful to have two of the best finance professors at John Molson School of Business as the committee members. I highly appreciate their advice and help during my studies and specially for my thesis.

Additionally, I would like to thank my dear professor, Dr. Nebebe, for all of his support during my studies at John Molson School of Business. I was so lucky to have the chance to be his student and work for him during my studies. I will never forget his kindness.

Also, I would like to thank Arbour Foundation and especially dear Mr. Philippe Arbour for all of their financial and emotional supports during my studies. I will do my best to be a good representative as a scholar of Arbour Foundation.

Lastly, I like to thank my parents, Farzaneh and Masoud, and my sister, Zahra, for all of the support and love they always gave me. Thank you so much for being always there for me when the rain starts to pour!

Table of Contents

List of Tables

List of Figures

1. Introduction

Uncertainty is an innate part of the financial markets. Every financial product has a certain level of return and risk which is measured by the fluctuation of its return. Investors always attempt to maximize their wealth by choosing the best financial products, those which offer the highest return compared to their risk. Among a wide variety of financial products, leveraged exchange-traded funds (LETFs) have recently become increasingly popular. LETFs[1] promise a return of -1, ±2, ±3 or ±4 times the return of a given asset and are thus riskier than standard ETFs (Wided, 2018).

There has been increasing interest in LETFs among market participants, making them an interesting topic of study, and recent studies reveal that LETFs have a significant effect on the market of their underlying assets. For instance, Yagi and Mizuta (2016) find that the amount of managed assets of LETFs can affect the volatility of the underlying asset and if its impact is higher than its ordinary volatility then the market will be destroyed. Thus, it is essential to conduct a comprehensive study of these LETFs and identify profitable investment strategies in order to avoid unnecessary volatility in both their market and their underlying asset market.

LETFs are characterized by two prominent features: path dependency and decay. Due to these features, LETFs may have a considerably lower long-term cumulative return compared to their underlying assets. Thus, LETFs are not appropriate products for long-term investment and investors should constantly rebalance their portfolios when they invest in LETFs (Trainor and Carroll, 2013).

---

[1] A leveraged ETF is a marketable security that leverages the changes of a specific underlying asset by using different types of assets like derivatives or debt. While regular ETFs mainly track a specific asset on a one-to-one basis, leveraged ETFs can provide 2, 3, or 4 times the return of underlying assets. These tools provide the opportunity for investors to amplify their bet on or against the market easily and without any need of having expertise on derivatives.

Leung and Ward (2015) and Adhikari et al. (2020) find that the return of gold LETFs deviates from the underlying asset as the holding period extends. Wided (2018) show that investing in LETFs can be highly profitable when the volatility of the underlying assets is low. Shum et al. (2016) also confirm this conclusion and assert that when the market is volatile, LETFs will have larger tracking errors due to higher balancing costs. Based on these articles, we can conclude that LETFs are suitable when the holding period is short, and the market is less volatile.

However, recent articles offer some investment strategies that allow for profitable long-term investments in LETFs. For instance, DiLellio et al. (2014) find that if LETFs are added to a portfolio along with stocks and bonds for the purpose of diversification, they can actually be very profitable, even over a long-term holding period. They also find that the profitability of the LETFs is highly dependent on the market conditions. Jiang and Peterburgsky (2017) find that short-selling bull and bear pairs of LETFs along with a long position in Treasury bills can yield a return higher than the market. Hessel et al. (2018) show that short-selling a portfolio including 50% bull and 50% bear variations of four frequently traded LETFs can yield higher than market returns in the long run.

The idea behind short-selling pairs of LETFs is based on the path dependency feature of these investment tools. Since it has been shown that keeping LETFs for the long term typically results in massive losses, if a portfolio is created with a combination of bull and bear ETFs and is kept for the long run, its return should be highly negative. Thus, short-selling this portfolio should inverse the total return and should earn a very high profit.

Although few studies have investigated different long-term investment strategies for LETFs, some ambiguous issues still require clarification. Firstly, the previously mentioned studies examine a limited number of LETFs which are based on popular indices such as the S&P500, making it impossible to compare different types of LETFs. Secondly, it is not clear whether a short pairing strategy is suitable in all different market conditions. Thirdly, the main factors impacting the success of a short pairing strategy need to be examined. Finally, it is interesting to explore how the return of a short pairing strategy varies when a portfolio is balanced quarterly and annually since both of these investment intervals are considered long-term in the literature of LETFs.

The main goal of this study is to fill these observed gaps. In addition, we aim to find the best long-term strategy to invest in LETFs while considering market conditions and underlying index features and analyze which factors affect this strategy using machine learning techniques. To address these questions, we use US-traded LETFs from 2012 to 2020 and construct seven portfolios including different combinations of bull and bear LETFs. We then calculate the short-selling return of these portfolios using different measures in each quarter/year to find the best portfolio for that interval. We rebalance the portfolios quarterly/annually. Lastly, after finding the best short pairing strategy, we investigate the main determinants that make this strategy profitable and determine how they impact the strategy. We use various machine learning techniques to analyze important factors in the short pairing strategy. We use the in-sample data (70% of total data) to train the machine learning and regression model and then the out-of-sample data (30% of total data) for getting the accuracy and validation of the model[2].

The remainder of this paper is structured as follows. Following this introductory chapter, chapter 2 explores the concepts, theories, and recent articles related to LETFs and short pairing strategies. Chapter 3 describes the main research questions, explains our sample construction process, identifies our variables, and outlines the main statistical and machine learning methods used to test

[2] In-sample data is part of the data that is used to build the model. On the other hand, the out-of-sample data is the part of data that is unseen for the model, and it is only used to evaluate the forecasting performance of the achieved model.

the research questions. Chapter 4 highlights the results of our study and chapter 5 discusses the importance of our findings and offers suggestions for future research.

2. Literature Review

2.1. ETF introduction and structure

The main goal of the major top exchange-traded funds is to replicate the performance of an index. The investment combination of these ETFs only changes in response to changes in their underlying asset. Unlike their underlying index, ETFs can be purchased and sold during a trading day like other stocks, making it possible for both individual and institutional investors to invest in any index or market easily and at a low cost (Rompotis, 2014).

ETFs are undoubtedly one of the most popular and fastest-growing sectors of global investments, this is likely due to their lower cost. Unlike mutual funds or other types of investment products with relatively high fees and commissions, ETFs have a very low expense ratio, this is due to the high level of competition between ETF providers. Since ETFs do not bear the cost of active portfolio management, they can offer lower fees compared to mutual funds (Tsalikis, 2020). Their growth also may be due to the fact that ETFs facilitate investing in various markets like precious metals, emerging market bonds, and currencies (which were so costly to invest in before ETFs). Thanks to ETFs, all investors (including even small and unsophisticated ones) can easily access any part of any market (Aggarwal and Schofield, 2014). Their increasing popularity may also be linked to transparency. In addition to the fact that all investors are aware of the components of indices, ETFs are obliged to publish their holdings on their websites on a daily basis, whereas mutual funds disclose their portfolios quarterly and hedge funds display their holdings yearly. ETFs can therefore be considered one of the most transparent investment products (Hill et al., 2015). The accelerated growth of ETFs could also be tied to liquidity. ETFs can be traded in the secondary market on a daily basis. They can be held on margin, optioned, and shorted like real stocks, making their liquidity of ETFs is similar to equities (Tsalikis and Papadopoulos, 2018).

Similar to mutual funds, ETFs are professionally managed funds that invest a mix of many different assets. Investors can buy their shares and become the proportional owner of the cumulative assets of the fund. In spite of these common features between the structures of ETFs and mutual funds, ETFs involve a unique process called creation and redemption, which makes their structure different from other investment products (Hill et al., 2015). Each ETF has a list of authorized participants (market makers) that are in charge of the creation and redemption process. Investors trade the ETF shares in the secondary market. Hypothetically, if the demand for shares of an ETF significantly increases so that the market price of its share exceeds the net value of ETF assets, an authorized participant will react to this arbitrage opportunity. They would do this by short selling the ETF shares and buying the underlying asset, creating profit from this difference. Then, they would deliver the basket of underlying securities to the given ETF and receive new ETF shares. Using these new ETF shares, authorized participants can exit their short positions, this is defined as the creation process (Meziani, 2016). However, if the demand of the ETF shares drops and causes a reduction in their market price so that it falls below the net value of the ETF assets, an authorized participant would still react. In this scenario, they would buy the ETF shares and short sell the underlying securities. Then, they would deliver the ETF shares to the given ETFs and receive the basket of the securities. They can then exit their short position using the basket of securities, this is defined as the redemption process. Through this mechanism, the market price of ETFs deviates slightly from the net value of the ETFs assets (Gastineau, 2008). The process is illustrated in figure 1.

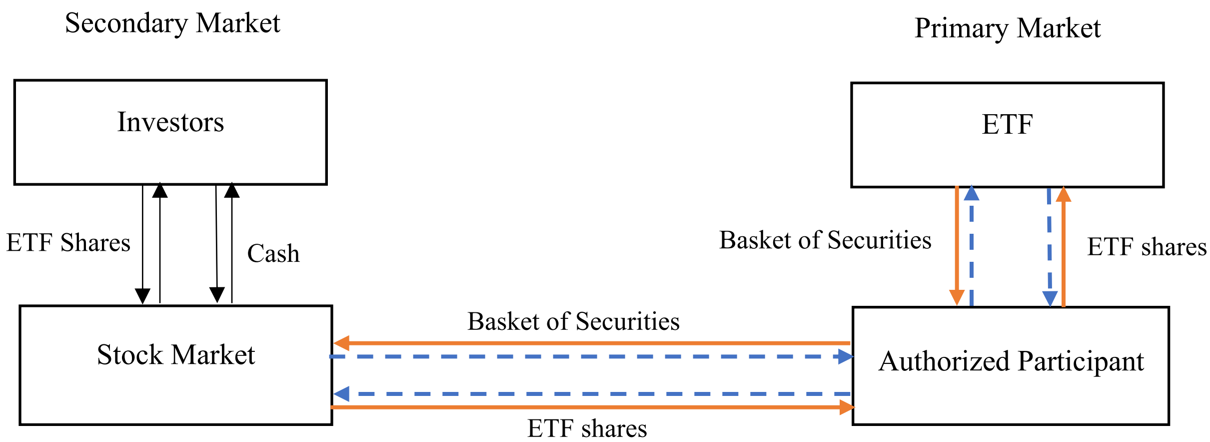

Figure 1. Creation and redemption process.
(Orange lines show the creation process and blue lines show the redemption process)

2.2. ETF categories

Due to the high investor interest in ETFs, these funds have been growing rapidly and offering various products. ETFs provide access to different markets globally and offer different modifications with respect to their underlying asset (Tsalikis, 2020).

2.2.1. Equity ETFs: Equity ETFs are the most popular form of ETFs. These are concentrated on the security market. They may invest in different indices related to the stock market or in stocks of the companies working in the same sector or producing the same products. They may also have a similar size stock selection, and the weight of each stock in the portfolio can significantly affect their volatility (Deville, 2008).

2.2.2. Fixed-income ETFs: These types of ETFs mainly invest in different bonds, including safe T-bills, or even riskier bonds such as junk bonds. Since bonds are typically traded over the counter, buying a small number can be very costly. These types of ETFs, however, provide a cost-effective way to access various types of bonds, in which investors only need to determine the level of credit risk that they are willing to take (Hill et al., 2015).

2.2.3. Commodity ETFs: Common products such as food, energy, and precious metals are the underlying assets of these ETFs. ETFs have made these assets much more affordable, and some of these commodities, such as precious metals, have low volatility making them a good option for diversifying portfolios (Gastineau, 2008).

2.2.4. Currency ETFs: These ETFs allow investors to include different currencies in their portfolios with a very low principal. Before ETFs, a separate account was needed to invest in currencies and the minimum amount of investment was relatively high. The first form of currency ETFs was "exchange-traded notes," which are debt obligations that promise payments based on the exchange rate of a specific currency to the US dollar on a given date. Now, these ETFs offer a wide variety of currencies such as emerging markets (Tsalikis, 2020).

2.3. Leveraged ETFs

The first leveraged ETFs (LETFs) were introduced to the U.S. market by Proshares in June of 2006. They are designed to amplify the daily return of the underlying index. There are two main types of LETFs: bull and bear. Bull, or "long ETFs", aim to achieve two, three, or four times the return of their underlying asset. While bear or "short ETFs" (inverse ETFs), try to match the performance of -1X, -2X, -3X the daily return of the underlying indices. To meet this goal, LETFs use various financial products, such as futures, swaps, and other derivatives (Rompotis, 2014).

LETFs have several advantages which have made them very popular among investors. First, by using LETFs, investors are able to increase their market exposure without needing to have expertise on derivative securities or an expensive margin account. Second, when the benchmark or underlying index is restricted for short selling, LETFs allow investors to short sell. Third, investors can use these securities to pursue hedge fund-like strategies with the liquidity and convenience of an ETF (Adhikari et al., 2020).

Similarly, to simple ETFs, LETFs can be classified into four asset categories: leveraged equity ETFs, leveraged bonds ETFs, leveraged commodity ETFs, and leveraged currency ETFs. LETFs have two main characteristics which are known as decay and path dependency (Hill et al., 2015).

2.4. Path dependence and decay

Although LETFs seem to be able to achieve their target return in the short-term (Trainor and Baryla, 2008), they fail to provide two or three times the return of benchmarks if the holding period is too long, an effect called "volatility decay" (Guo and Leung, 2015). Tsalikis and Papadopoulos (2018) investigate the performance of LETFs among American and European ETFs. They find that LETFs successfully deliver the intended return over holding periods of up to one week. However, as the holding period extends to one month, their performance deviates from the initially promised performance.

Ilan Guedj and McCann (2010) estimate that investors can lose 3% of their original investment by holding LETFs for several weeks. Rompotis (2014) suggests this may be due to the constant leverage traps and the lognormal nature of continuously compounded returns. The decay effect of compounded returns becomes more exacerbate as the holding period is extended. In the long run, not only LETFs do not double or triple their return but they also expose investors to two or three times the amount of risk (Trainor and Baryla, 2008).

Avellaneda and Zhang (2010) find similar results by examining 56 leveraged funds over quarterly horizons during 2008 and 2009. They conclude that over quarterly or annual investment horizons, LETFs do not replicate the corresponding multiple of benchmark returns. They suggest that LETFs may not be a suitable choice for passive investors using the buy-and-hold strategy, but that by using a dynamic rebalancing strategy, it is possible to replicate the underlying index return in the long term.

Similarly, Lu et al. (2009) examine Ultra ETFs (which have twice the positive return of the benchmark) and UltraShort ETFs (which have twice the negative return of the benchmark) from the ProShares family. They find that the returns of Ultra ETFs and UltraShort ETFs deviate from twice the returns of the underlying index when the holding periods increase to a quarter and a year. They further explain that the quadratic variation and the auto-variation during a long-term period negatively impacts the LETFs performance.

Shum and Kang (2013) decompose the deviation of the LETFs from their underlying asset into three parts: compounding, management factors, and trading premium or discount. They find that the management factor has the greatest effect on LETFs return deviation. They additionally find that returns of bear LETFs have a higher deviation from the underlying index compared to bull LETFs. Their results also suggest that financial crisis has a stronger effect on the volatility of bear LETFs rather than bull ones.

2.5. Pair investment strategy in LETFs

Based on the aforementioned research, as well as path dependency and decay in LETFs, it seems they are unsuitable for long term investments. However, considering the fact that bull and bear LETFs are inverse bids on the underlying assets, investing in both of them simultaneously may provide a long-term investment option (Jiang and Peterburgsky, 2017).

According to Hessel et al. (2018), we can define the compound return of the bull and bear LETF by equations 1 and 2 respectively.

$$R_{tn}^{L}=\prod_{i=1}^{n-1}(1+r_{t+i}^{L})-1 \quad (1)$$

$$R_{tn}^{I}=\prod_{i=1}^{n-1}(1+r_{t+i}^{I})-1 \quad (2)$$

Where r_{t+i}^{L} and r_{t+i}^{I} are daily returns of bull and bear LETFs respectively, *t* is the starting day of investment and *n* is the number of holding days.

The return of a portfolio created by shorting both bull and bear LETFs with the same weights (each 0.5) is defined by equation 3.

$$R_{tn}^{S}=-0.5R_{tn}^{L}-0.5\ R_{tn}^{I} \quad (3)$$

The daily return of bull and bear ETFs is equal to ±m times that of the underlying asset. For simplicity, we perform the shorting pair strategy for a two-day holding period, defining the return of the strategy by equations 1 and 2:

$$R_{t2}^{L}=(1+mr_{t}^{B})(1+mr_{t+1}^{B})-1=\ mr_{t}^{B}+\ mr_{t+1}^{B}+m^{2}r_{t}^{B}r_{t+1}^{B} \quad (4)$$

$$R_{t2}^{I}=(1-mr_{t}^{B})(1-mr_{t+1}^{B})-1=\ -mr_{t}^{B}-\ mr_{t+1}^{B}+m^{2}r_{t}^{B}r_{t+1}^{B} \quad (5)$$

Where r_{t}^{B} is the daily return of the underlying asset. We can then rewrite equation 3 as follows:

$$R_{tn}^{S}= -0.5(mr_{t}^{B}+\ mr_{t+1}^{B}+m^{2}r_{t}^{B}r_{t+1}^{B})\ -0.5(-mr_{t}^{B}-\ mr_{t+1}^{B}+m^{2}r_{t}^{B}r_{t+1}^{B})= -m^{2}r_{t}^{B}r_{t+1}^{B} \quad (6)$$

This equation highlights that the return of such a portfolio is profitable when the return of the underlying asset at times *t* and *t+1* have opposite signs. Hessel et al. (2018) demonstrate that this relationship is similar when assessing longer holding periods. This means that this strategy is suitable for a situation where there is a negative autocorrelation for the return of the underlying index, or in which there is high volatility in the return of the underlying index. They conclude that keeping both bull and bear LETF is like a bet on the mean reversion.

They support their findings with empirical data, constructing their shorting pair portfolios with equal weights on the monthly basis for six indices. They find that the shorting pair strategy has the highest return for LETFs and has high volatility and negative autocorrelation. They also find that the volatility of the underlying index is the most important predictor of the pairing strategy return. They find that the monthly returns of the shorting pair strategy are not linked to any of these known asset pricing factors and that there is a positive significant alpha associated with this strategy.

Jiang and Peterburgsky (2017) conduct a similar investigation to determine a long-term investment strategy for LETFs. They create portfolios with short positions in both bull and bear ETFs and a long position in T-bills. They also test different investment weights for them, including %100-%0, %75-%25, %66.7-%33.3, and %50-%50. Additionally, they consider different thresholds for rebalancing the portfolio. When the short balance is more or less than 5% (also 10% and 20%) of the portfolio value, then rebalancing occurs on a daily basis. The authors find that a portfolio with weights of %33.3 in bull LETF and %66.7 in bear LETF, with a threshold of 20% for rebalancing earns an annual return higher than the underlying index. Additionally, their results reveal that shorting only bear LETFs and longing treasuries can also yield a higher average annual year return, which is risk-adjusted, than the underlying index. However, the shorting pair strategy return is higher than investing in one of them. The authors also consider transaction costs and commission fees to make sure their results are robust, and they remain almost unchanged.

As a supplement to his above work, Peterburgsky (2020) tests the short pairing strategy with a long position in T-bills using various types of LETFs with different underlying assets. The results suggest that this strategy always yields a higher return than the underlying index. For LETFs based on the S&P500, this strategy earns the highest return when 3x LETFs are used, while those based on the Russell 2000 yield a higher return with 2x LEFTs. The author then uses a regression analysis and defines three regression lines for 3x, 2x, and 1x LETFs. In these regression lines, the difference between the short pairing return and the index return in quarter *t* is considered as the dependent variable. The independent variables are absolute returns on the index, its standard deviation, its skewness, and its kurtosis, all of these, in quarter *t*. He finds that for the 3x and 2x regression lines, absolute return and standard deviation are the most significant predictors. The author uses a 20%, 10%, and 5% rule for rebalancing the portfolios, and finds that as the rebalancing range decreases from 20% to 5%, the outperformance of the short pair strategy also decreases.

Saini (2019) examines the best times for the short pairing strategy. To meet this goal, the author develops a three-state regime-switching framework based on time-series momentum and volatility. The three states are the momentum regime, the variant regime, and the inert regime. When the reference index trends either upward or downward, it is classified as being part of the momentum regime. When it moves within certain range bounds, and is highly volatile at the same time, it is classified as belonging to the variant regime. When the underlying index changes in range bounds with low expected variance, its condition is classified as being that of the inert regime. After defining these different regimes, the author calculates the trend and volatility of each day by using the rolling moving averages of daily returns of the underlying index over 6-months, 3-months, 6-weeks, and 2-weeks. Additionally, the author considers the volatility of the S&P 500 index over 15 years as a benchmark for the intensity of volatility. The author uses absolute return, the Shapiro ratio, and the Sortino ratio as the measurements of the return. The author also compared the commodity and non-commodity indices for the short pairing strategies. Based on these three regimes, the author makes her investment decision and develops a portfolio. If the regime is identified as variant, then the investment strategy is to short both pairs. If the regime is either momentum or inert, the decision is to remain in cash.

In order to forecast the expected volatility of each day, the author uses various methods, including the seasonal random walk model, the implied volatility model, the simple moving average model, the generalized autoregressive conditional heteroskedasticity method (GARCH), the exponentially weighted moving average model, and the autoregressive integrated moving average (ARIMA). The author defines the commission cost as $0.005 per trade and slippage to be 0.30%. The cost of borrowing LETF is considered 6% per year. The daily return of this portfolio is compared to a passive portfolio using the z test and t-test.

The results show that the pair trading strategy in commodity LETFs outperforms a passive sell-and-hold trading strategy on a risk-adjusted basis according to the Sortino ratio. The author, however, is not able to find the same evidence using either Sharpe ratios or absolute returns.

2.6. Important variables affecting LETFs and machine learning techniques

Various articles investigate the main factors affecting ETFs and LETFs. The main focus of the literature, however, is on ETFs. But, since the concepts of these tools are similar, it is possible to argue that variables that affect ETFs are potentially important for LETFs too.

Liew and Mayster (2017) investigate the effect of different lags of past returns and volume for indices and ETFs, along with time dummy ETF on forecasting their returns using four different machine learning techniques. They find that using these variables, the input of machine learning techniques can significantly amplify the predictability of ETF returns.

Day and Lin (2019) also use previous data concerning the price, trend, and volatility of ETFs to predict Taiwan's ETFs returns and create an optimal portfolio using machine learning techniques.

They combine different techniques to come up with a Robo-advisor framework. They conclude that this framework can outperform the baseline ETF.

Brown et al. (2021) investigate different factors which can affect ETFs returns. They consider outstanding shares, trading volume, market capitalization, and bid/ask spread as the main explanatory variables in their model. They also examine the non-fundamental demand and possible arbitrage profit.

Bahadar et al. (2019) also study the LETF markets and its determinants. They find that market return, trading volume, and trading volatility are important factors affecting herding behavior in LETFs markets.

Deev and Linnertová (2014) examine the determinants of ETFs' short-selling returns. They consider trading volume, price stability, market capitalization, expense ratio, geographical focus, investment strategy, and availability of derivatives for the underlying index as the main important factors affecting the short-selling return of ETFs.

2.7. Market efficiency

According to this theory, in an efficient market, in an efficient market, all information related to a given stock (such as past performance and accounting variables) is immediately reflected in the prices due to information symmetry. The capital asset pricing model (CAPM) and Fama-French models are two popular models used to determine the fair price of a given asset based on its correlation with the market return. A positive alpha, in these models, means there is an abnormal return associated with a given stock. Some researchers consider this positive alpha as a violation of the market efficiency hypothesis (EMH). There are three forms of EMH including the weak, semi-strong, and strong forms (Saini, 2019).

The weak form of the efficient market hypothesis implies that all market trading data, including past price or trading volume, is reflected in stock prices. Because all of the investors can access past price trends as well as other market trading data and they know how to use this data, these signals lose their value. The weak form efficient market hypothesis is based on the "random walk theory," according to which stock price changes should be random and unpredictable. If new information is available, intelligent investors attempt to trade stocks based on that information before the rest of the market. As a result, stock prices immediately change based on the new information, it is therefore not possible to predict stock price changes (Prakash, 2012).

In order to test the weak form of EMH, we use serial autocorrelation. If there is a positive (negative) autocorrelation, this means that a positive return in the past will be followed by a positive (negative) return in the future—this effect is called the "momentum effect", or the reversal effect. Various researchers support this effect (Botha and Marx, 2017; Garg and Varshney, 2015; Prakash, 2012).

The semi-strong form of market efficiency is related to a firms' prospects and fundamental data. For instance, the earnings ratio, the book to market ratio, the size of the firm, and the P/E ratio seem to be able to predict abnormal risk-adjusted returns. There is rich literature showing that stock markets are not efficient in the semi-strong form (Jethwani, and Ramchandani, 2017; Mackey and Bacon, 2017; Prakash, 2012).

The strong form of market efficiency is related to insider information and data which are not publicly available. The profitability of insider traders in their own stock has been proven by various studies such as those of Leković (2018), Yadav (2016), and Bashir et al (2020). Although there are different laws in place limiting trades using insider information, we should not expect markets to be efficient in strong form.

2.8. Contributions

This paper contributes to different aspects of the literature on LETFs.

- There is a limited number of studies examining LETFs as a long-term investment tool, as these mostly consider LETFs which are frequently traded. Since these LETFs fall into one or two certain types in terms of their underlying index, it is not apparent if the short pairing strategy can also perform well for different types of these. This study, however, considers a sample of all LETFs with different types and leverage levels, which allows for more in-depth forms of comparisons.
- The market condition has not previously been specifically considered as an important factor affecting the short pairing strategy. This study's research period, spanning a period from 2012 to 2020, includes different market conditions, including that of the year 2020, which saw COVID-19 affect the whole market and economy. Thus, it is possible to compare different subsamples for market conditions like bullish or bearish markets and test the performance of the short pairing strategy in these.
- The determinant factors affecting LETFs are more or less investigated in the literature, and the important factors affecting the short pairing strategy have yet to be explored in depth. Our research collects a comprehensive list of potential variables, of which the most important are selected using various machine learning techniques. These methods are used to predict and determine the accuracy of the models. In fact, to the best of our knowledge, this study is the first to employ machine learning techniques to investigate the short pairing strategy in LETFs.

3. Data and methodology

3.1. Research questions and hypotheses

The main questions that this study aims to answer are as follows:

- What is the best combination of bull and bear LETFs which can be profitable in the long term?
- Is there a significant difference between the best portfolio that is selected for different leverage sizes or different types of LETFs?
- How do market conditions affect the profitability of the best portfolio?
- What are the most important factors affecting the performance of the best portfolio and how (positively or negatively) do they impact these?

3.2. Sample construction

In order to create the sample, we collect the names of all available ETFs from Morningstar, Bloomberg, and etfdb.com, we then use a keyword search in order to determine whether the ETF is leveraged. If the name of a given ETF includes one of these keywords, it is considered a LETF: 2x, 3x, leveraged, lvged, lvg, inverse, ultra, double, triple, enhanced, amplify, and short. We then check the list obtained manually to verify the selected ETFs are leveraged. Following these steps, a sample of 275 ETFs remains. The level of leverage for each of these is identified using the same keyword method. The sample of this study consists of all LETFs that meet the following criteria:

- Each LETF should have an inverse pair. The inverse pairs must have the same issuer, the same underlying asset, and the same leverage size.
- The LETF should be available for trading during the research period (from 2012 to 2020).

- The data for all research variables should be available during the research period for both pairs.
- The underlying asset or index should be opened before 2012 and used during the research period without any halt or disruption.
- The data of underlying assets should be available during the research period.
- The trading volume for each LETF should not be zero for more than 50 days.

Considering these conditions, a total of 88 LETFs (44 pairs) and 792 LETF-year observations remain for observation. The name and description of the selected pairs are summarized in table 1.

*** Insert Table 1 about here ***

According to table 1, all selected LETFs are double or triple since LETFs with other leverage sizes do not have appropriate inverse pairs. The underlying assets of the selected sample include US equity, emerging and foreign markets, commodities, currency, and real estate. The data for all research variables for the selected LETFs is obtained from the Bloomberg website. Additionally, the data of risk-free rates is needed for the calculation of LETFs' returns and is gathered from the treasury.gov website.

3.3. Variable definition

3.3.1. Return of LETFs

We use three methods to calculate the return of each LETF: cumulative return, the Sharpe ratio, and the Sordino ratio. The first method does not consider the risk of investment, while the others provide a risk-adjusted return for each LETF. To calculate these measures, primarily, the daily return is calculated as in equation 7:

$$DR_t^i = \frac{P_t^i - P_{t-1}^i}{P_{t-1}^i} \tag{7}$$

Where DR_t^i is the daily return of LETF i at day t and P is the closing price. Next, the Cumulative return is calculated using equation 8 (Fabozzi, 2008).

$$CR_t^i = \left(\prod_1^t 1 + DR_t^i \right) - 1 \tag{8}$$

The next method we employ is the Sharpe ratio which is a proxy for risk-adjusted return. The ratio considers the average of excess return per unit of volatility. Volatility is also defined as the standard deviation of all excess returns during a certain period. The Sharpe ratio is calculated as in equation 9 (Sharpe, 1994).

$$Sharpe\,{}_t^i = \frac{E_t^i(DR_t^i - rf_t)}{\sigma_t^i(DR_t^i - rf_t)} \tag{9}$$

Where rf_t is the risk-free rate at day t, E_t^i is the average at time t and σ_t^i is the standard deviation.

We then use the Sortino ratio to calculate return. This ratio is considered an extension of the Sharpe ratio. Like the Sharpe ratio, the average of the excess return is the proxy of return, however, the Sortino ratio assumes that investors mainly concern themselves with the downside of the risk or volatility of negative returns only, therefore, the volatility is defined as the downward risk (Sortino and Price, 1994). The Sortino ratio is calculated as in equation 10.

$$Sortino_t^i = \frac{E_t^i(DR_t^i - rf_t)}{DD_t^i} \tag{10}$$

Where DD_t^i is the downward risk for LETF i at time t. Downward risk is essentially considered as the standard deviation of all negative excess returns (Sortino and Price, 1994).

3.3.2. Risk of LETFs

We use two measures to calculate the risk of each LETF: different types of value at risk or VaR (including analytical VaR, Expected VaR, historical VaR) and the standard deviation of all returns during the holding period. VaR is essentially defined as the smallest level of loss (or profit) such that the probability of losing more than this value is at most α%. Mathematically, VaR can be defined as follows: (Dionne, 2019):

$$VaR_{\alpha}(X) = -\inf \{x \in \mathbb{R} : F_X(x) > \alpha \} \tag{11}$$

We calculate all types of VaR with an α value of α=0.10. To practically calculate VaR for a LETF *i* at time *t*, we sort returns in ascending order. The historical VaR is the nth return where n is equal to the total number of days that the LETF is kept multiplied by α. Expected VaR is calculated as the average of all returns that are less than or equal to the nth return (Kochenderfer and Wheeler, 2019). Analytical VaR can be calculated as in equation 12.

AVaR= The average of all returns during time *t* – 1.65*St. dev. of all returns during time *t* (12)

In this equation, we obtain 1.65 is by using the inverse normal distribution for 95% probability (Duffie and Pan, 2001).

3.3.3. Regression variables

After finding the most effective investment strategy (or strategies), it is important to identify the circumstances in which it can perform well. More specifically, the next step is to find the factors impacting the performance of the winning strategy and determine whether it is going to be profitable for a given investment period. There is a large number of explanatory variables that could be considered as independent variables in this regression. They are described as follows:

- Price: the average closing price during the holding period for each LETF and underlying index.
- Daily return: the average daily return (from equation 7) during the holding period for each LETF and underlying index.
- Return trend: a dummy variable that is equal to 1 if the average return of the last t/4 , where t is the total number of days that LETF is held, is higher than the average return of the last t/2 days, otherwise it is equal to 0.
- The volatility of daily return: the standard deviation of all the daily returns during the holding period (Bahadar et al., 2019, Chovancová et al., 2019), as well as different types of VaR.
- The beta of price and the beta of return: the beta is the Pearson correlation coefficient of the price (or daily return) of each LETF with the price (daily return) of the corresponding underlying index during the investment period.
- Trading volume: the average trading volume during the holding period for each LETF and underlying index (Bahadar et al., 2019).
- Market capitalization: the average market cap during the holding period for each LETF. The market cap is defined as the number of shares outstanding multiplied by the price of LETF (Charupat and Miu 2013).
- Bid/ask spread: the average of the bid/ask spread during the holding period for each LETF. The bid/ask spread is the difference between the highest price that a buyer is willing to pay and the lowest price that a seller tends to take (Clifford et al., 2014, Baker et al., 2015, Chovancová et al., 2019, Charupat and Miu 2013).

- Turnover rate: the average turnover rate during the holding period for each LETF. The turnover rate highlights the replacement of holdings for a certain LETF and it is calculated as follows (Baker et al., 2015):

$$Turnover_i^t = \frac{Average\ dollar\ amount\ of\ holding\ purchased\ and\ sold\ during\ time\ t}{Total\ asset\ under\ management\ during\ time\ t} \quad (13)$$

- Net asset value (NAV): the average net asset value during the holding period for each LETF. Net asset value is the present value of all LETF holdings (Chovancová et al., 2019).

- Price to NAV ratio: the average value of the closing price over the NAV during the holding period for each LETF (Omar et al., 2021).

- Asset type dummy variables: a set of five dummy variables related to a certain type of underlying asset. For instance, the equity dummy is equal to 1 if the underlying asset for a given observation is US equity, otherwise it is equal to 0. Four other dummy variables are similarly defined for each type of underlying asset.

- Year dummy: a set of nine dummy variables related to different years. For instance, the *Year2012* dummy is equal to 1 if a given observation is related to the year 2012, otherwise it is equal to 0. Eight other dummy variables are defined every year of the observation period.

- Quarter dummy (for quarterly data): a set of four dummy variables related to different quarters. For instance, the *Quarter1* dummy is equal to 1 if a given observation is related to the first quarter, otherwise it equals 0. Similarly, three other dummy variables are defined for other quarters.

All of the calculations are done in both quarterly and annual horizons. The summary of descriptive statistics for annual data is mentioned in table 2.

*** Insert Table 2 about here ***

The dependent variable of this regression is the winning strategy. It is a portfolio whose risk-adjusted return, based on both the Sharpe ratio and the Sortino ratio, surpasses that of its underlying index. If at least one of these ratios, for all portfolios, is less than that of the index, then none of the portfolios are winners, which implies it is better to stay in cash. Based on this definition, two winning strategies are found. A dummy variable, named *winning strategy*, is defined as having the three following values:

- If none of the winning strategies can beat the market, then it is better to stay in cash than investing; in this situation, the *winning strategy* dummy is equal to 1.
- If the first strategy (defined as short-selling the full bull portfolio) is the winner, then the *winning strategy* dummy is equal to 2.
- If the second strategy (defined as short-selling the full bear portfolio) is the winner, then the *winning strategy* dummy is equal to 3.

When the *winning strategy* dummy variable for each quarter/year is determined, then one lag of explanatory variables is matched with it so that the information of the previous investment period can be used to anticipate the next investment period.

3.4. Methodology

3.4.1. Finding the best portfolio

We perform different types of tests in order to examine the hypothesis of this study. First, to find the winning long term short pairing strategy, we create seven portfolios. These include different combinations of paired bull and bear LETFs. The description of these portfolios is summarized in table 3. After constructing the portfolios, we test the short pairing strategy on a quarterly and annual basis. At the end of each investing period (a quarter/a year), we calculate and compare the short-selling return of all seven portfolios. This is done according to the measures described in section 3.3.1 and 3.3.2 and based on the methods of Hessel et al. (2018). The calculations of short-selling returns are detailed in table 3.

*** Insert Table 3 about here ***

Next, for each quarter/year, we compare the short-selling returns of portfolios with each other to find the most profitable portfolio. The portfolio that is selected as the winner in most quarters/years is considered the overall best short pairing strategy.

We additionally compare the short-selling returns of the portfolios with the return of the corresponding index from 2012 to 2021 to examine if they are able to beat the index and outperform the market. We use a two-sample t-test in order to perform this comparison.

The short-selling cost is one of the important variables that can affect short-selling returns and make the position less, or even not, profitable. After finding the best portfolio, different short-selling costs are considered in order to determine different draw-even points for it. We first transform the annual short-selling costs to the daily short-selling cost and then deduct this cost from daily short-selling returns to achieve the cost-adjusted daily short-selling return (defined by equation 14). Other calculations are obtained using this value.

$$SDR_t^i = (-1 * DR_t^i) - [\ (SCost_y + 1)^{1/\mathrm{DY}} - 1] \quad (14)$$

Where SDR_t^i is the cost-adjusted daily short-selling return for a certain LETF i at day t, $SCost_y$ is the annual short-selling cost for year y, and DY is the number of days in year y.

3.4.2. Comparing LETFs with different characteristics and in the different market situation

In order to investigate the short-selling returns further, and determine the best short pairing strategy, we divide the LETFs into two groups based on their leverage level including 2x and 3x LETFs. We once again compare the short-selling returns of the portfolios and compare these to the index using a two-sample t-test.

We then divide all LETFs into five categories based on the type of their underlying asset. These categories include US equity, emerging and foreign markets, commodity, currency, and real estate. We once again perform a two-sample t-test in order to compare their returns to the index.

In order to ensure the robustness of the results obtained and assess that the winning strategy is profitable in the different market situations, a new variable is defined which accounts for market situation. If the average of index returns for all three months in a quarter (or four quarters in a year) is negative, the market is bearish. Conversely, if the average of index returns for all three months in a quarter (or for four quarters in a year) is positive, the market is bullish, if neither applies, the market is natural. We then examine the return of all portfolios, especially that if the winning portfolio, and compare these with that of the index during the bullish and bearish market to make sure they can still yield a positive return and beat the market in the different situations. For this comparison, we once again employ a two-sample t-test.

3.4.3. Machine learning and regression analysis

After finding the winning strategy and making sure it is profitable during different conditions, we address which factors affect the performance of this pairing strategy. To achieve this, we gather an extensive list of potential affecting factors from the existing literature. From this, we establish the following variables: bull and bear price, index price, bull and bear daily return, index daily return, bull, bear, and index analytical VaR, bull, bear, and index historical VaR, bull, bear, and index expected VaR, bull, bear, and index volatility, bull and bear beta-price, bull and bear beta-return, bull and bear trading volume, bull and bear market capitalization, bull and bear bid/ask spread, bull and bear turnover, bull and bear NAV, bull and bear price to NAV ratio, underlying asset type dummies, year dummies, and quarter dummies (only for quarterly data). The dependent variable is a dummy variable, the winning strategy dummy, which holds three values: 1, 2, and 3.

Since the dependent variable is a nominal variable with k=3 outcomes, the best type of regression to use is multinomial logistic regression. The multinominal logistic regression is the generalized version of a logistic regression that is used to solve multiclass problems. Similar to the binary logistic regression, the multinomial logistic regression uses the maximum likelihood estimation and log odds ratios of outcomes to model the investigated relationship (Pampel, 2020). To get robust results while using the multinominal logistic regression, some assumptions must hold.

- The dependent variable should be nominal or ordinal
- Each observation should fall into no more than one category of the dependent variable
- There should be no relationship among the independent variables (no multicollinearity)
- The observations should not include any outliers (Strickland, 2017)

In addition to multinomial logistic regression, which calculates the probability of each outcome dependent on others, we use each of the possible outcomes (bull, bear, or none of them) in a simple logistic regression separately to find the independent coefficients for each one.

Although the multinominal logistic regression is an appropriate method to investigate the research question, if the relationship of the target variables is too complex, it may not lead to precise results. Thus, in addition to multinominal logistic regression, it is necessary to use another method that can tackle more sophisticated relationships. For this purpose, a random forest classifier is used. The random forest is a supervised learning algorithm in machine learning methods that builds many uncorrelated individual decision trees and aggregates their outcomes to obtain an accurate classification. A decision tree is a combination of possible conditions and their corresponding outcomes. After creating decision trees, a prediction is made by considering the outcome that is selected the most (Dangeti, 2017). The main hyperparameter that should be specified for this method is the number of trees that the procedure must construct. In this study, we consider 100 trees when creating the forest.

In order to achieve robust results by multinominal logistic regression and random forest algorithm, we must first perform preprocessing procedures. We must first detect multicollinearity issues, in order to achieve this, we calculate a pairwise Pearson correlation of all independent variables. A correlation coefficient higher than 80% is considered as a serious multicollinearity problem. All correlated variables are dropped from the model so that none of the remaining predictors are highly dependent on each other. We must then detect and winsorize the outliers. In order to achieve this, we consider observations that are higher/lower than 1.5 times of interquartile mean as outliers. Once we determine these outliers, we winsorize them using the Gaussian method. This method replaces the right/left tail outliers by G where is equal to the mean ± 3*standard deviation.

Winsorizing is only applied on the most important factors that are selected by the feature selection methods (bull net asset value, index volatility, and return beta for bear for quarterly model and bull net asset value and return beta for bear for annual model). Since the scales of explanatory variables are different, we must transform and scale them so that they are on the same scale. This step is necessary to obtain optimal results. We transform the explanatory variables by the yeo-johnson method and scale them according to the z distribution. The yeo-johnson transformation is a power transformer that supports zero and negative values and transforms observations by the best lambda. According to Yeo and Johnson (2000), transformation rules are shown as in equation 15.

$$y_i^{(\lambda)} = \begin{cases} ((y_i + 1)^{\lambda} - \frac{1}{\lambda} & if\ \lambda \neq 0, \mathrm{y} = 0 \\ \log(y_i + 1) & if\ \lambda = 0, \mathrm{y} \geq 0 \\ -\frac{[(-y_i + 1)^{2-\lambda} - 1]}{2 - \lambda} & if\ \lambda \neq 2, \mathrm{y} < 0 \\ -\log(-y_i + 1) & if\ \lambda = 2, \mathrm{y} < 0 \end{cases} \tag{15}$$

In the next step, we divide the sample into two parts including the training set or in-sample and test set or out-of-sample. 30% of all observations are randomly categorized as the out-of-sample and the remaining are used as the in-sample data. In-sample data is part of the data that is used to build the model, On the other hand, the out-of-sample data is the part of data that is unseen for the model, and it is only used to evaluate the forecasting performance of the achieved model. We estimate the coefficients of multinominal logistic regression using the training set. We use the same dataset to train the random forest algorithm. We use the test set to test the accuracy of the estimated models.

In order to avoid an overfitting problem, which can be caused by the high number of potential predictors, we must identify the most effective variables in the list of potential factors. In order to do this, we use different feature selection techniques in machine learning algorithms, including Recursive Feature Elimination, Exhaustive Feature Selector, Lasso Regression, Ridge Regression, and Elastic Net.

The Recursive Feature Elimination (RFE) first fits all predictors and then eliminates them one by one to achieve the desired number determined by the user. The hyperparameter that is mentioned for this method is the number of features that the algorithm should select (Kuhn and Johnson, 2013). For the purpose of this study, we consider this number to be five.

The Exhaustive Feature Selector examines all possible combinations of the selected features and their accuracy, then returns the set of features that yield the most precise results. The hyperparameter for this method is the maximum number of features that can be in the model (Kuhn and Johnson, 2013). For the sake of this study, this hyperparameter is equal to 15, which is the number of all quantitative features.

The Lasso Regression (L1 regularization) and Ridge Regression (L2 regularization) are two approaches that add a penalty term to the regression line to avoid overfitting. This penalty term forces the coefficients of redundant features to be very small, in case of L2, or zero, in case of L1. For L1 regularization, the penalty term is the sum of the absolute values of the regression coefficients, and for L2 regularization, the penalty term is the sum of the squared values of the coefficients. The Elastic Net is a combination of L1 and L2 regularization. In this method, the penalty terms of L1 and L2 are linearly combined and used as a new penalty term (Molin, 2019). The hyperparameter for these methods is alpha, which is the multiplier of the penalty term and shows the strength of regularization. In this study, alpha is equal to 0.2, 0.02, and 0.02 for L1, L2, and the Elastic Net methods respectively, for both quarterly and annual data. The solver method for L1 and L2 regularizations is liblinear, while for the Elastic Net, it is Saga.

These feature selection methods and preprocessing techniques are only performed on quantitative variables. After finding the best set of quantitative variables, we add the other dummy variables, and use these as the explanatory variables of multinomial logistic regression. We show the models that are estimated for quarterly data and annual data in equations 16 and 17.

$$\text{Q\_winner} = \alpha_0 + \alpha_1 *\text{bull\_NAV} + \alpha_2 *\text{bear\_BR} + \alpha_3 *\text{IR} + \alpha_4 *\text{IV} + \alpha_5 *\text{equity} + \alpha_6 *\text{foreign} + \alpha_7 *\text{commodity} + \alpha_8 *\text{currency} + \alpha_9 *\text{Q1} + \alpha_{10} *\text{Q2} + \alpha_{11} *\text{Q3} + \alpha_{12} *\text{Y2013} + \alpha_{13} *\text{Y2014} + \alpha_{14} *\text{Y2015} + \alpha_{15} *\text{Y2016} + \alpha_{16} * \text{Y2017} + \alpha_{17} *\text{Y2018} + \alpha_{18} * \text{Y2019} + \alpha_{19} * \text{Y2020} \qquad (16)$$

Where *Q_winner* is the dummy variable of the winning strategy in each quarter, *bull_NAV* is the average net asset value of bull LETFs, *bear_BR* is the correlation of bear LETFs return with their underlying index, *IR* is the average index return, *IV* is the average index volatility, *equity, foreign, commodity, and currency* are the dummy variables for their respective asset types, *Q1, Q2,Q3, and Q4* are the dummy variable for their respective quarters, Y2013 is the dummy for the year 2013, and the subsequent yearly variables follow the same pattern.

$$\text{Y\_winner} = \beta_0 + \beta_1 *\text{bull\_NAV} + \beta_2 *\text{bear\_BR} + \beta_3 *\text{IP} + \beta_4 *\text{equity} + \beta_5 *\text{foreign} + \beta_6 *\text{commodity} + \beta_7 *\text{currency} + \beta_8 *\text{Y2013} + \beta_9 *\text{Y2014} + \beta_{10} *\text{Y2015} + \beta_{11} *\text{Y2016} + \beta_{12} *\text{Y2017} + \beta_{13} *\text{Y2018} + \beta_{14} *\text{Y2019} + \beta_{15} *\text{Y2020} \qquad (17)$$

Where *Y_winner* is the dummy variable of the winning strategy in each year, and *IP* is the average index price. The other terms are as defined in equation 16.

We perform all of the aforementioned on a quarterly and annual basis and compare the results of these two investing intervals. We perform all operations and calculations using the Python[3] programming language and the Anaconda platform. Additionally, we use the Microsoft Excel software to restore and view the datasets and create the tables.

4. Results

4.1. The best short pairing portfolio

Following the methodology established in section 3.4, we construct seven portfolios and keep these for quarterly and annual investment intervals. Table 4 provides a summary of the short-selling return of these portfolios.

*** Insert Table 4 about here ***

Table 4 shows that, out of the seven portfolios constructed based on different combinations of bull and bear LETFs, the best investment strategy is to short-sell a portfolio with a full investment in bear LETFs for both quarterly and annual investment intervals. While the absolute value of the expected VaR for this portfolio is slightly above others, indicating a higher risk, this strategy remarkably surpasses other portfolios when taking into account other measurements. This strategy is superior in 70% of studied quarters (and 75% of studied years). Additionally, according to cumulative return, the Sharpe ratio, and the Sortino ratio, this strategy generates a return that is, on average, 4.65%, 0.71%, and 1.083% higher than the market during a quarterly period. Similarly, for an annual investing period, this strategy beats the market by 16.47%, 0.79%, and 0.9%. Thus, the strategy which short-sells the full bear LETF outperforms the market, and its achieved margin

[3] Source code: https://github.com/ElaheNikbakht/Investigating-long-term-short-pairing-strategy-for-Leveraged-Exchange-Traded-Funds-using-machine-lea

is statistically significant at a 90% confidence level for both quarterly and annual intervals, this is true for all tests except for those of the Sharpe ratio values for annual data. Therefore, this strategy is not only able to beat the market, but it can also perform better compared to the other created portfolios.

Portfolios including both bull and bear LETFs are never selected as the winner. This is likely mainly due to the fact that bear LETFs and bull LETFs move in opposite directions, and this relationship holds in both short and long run. Thus, during each investing period, the gains of the profitable pair will be canceled out by the loss of the other. For the duration of the period studied, for both quarterly and annual investment periods, the average absolute return of these portfolios is always less than the portfolios with full investment in one of the pairs and they did not outperform the market. The only exception is the cumulative return of the 25%bull – 75% bear (1:3) portfolio which significantly exceeds the market by 1.49% for quarterly intervals and by 4.88% for annual intervals, but its short-selling returns are still lower than the full bear portfolio.

Eventually, it seems that the quarters (and years) that the full bear strategy is not a winner, the full bull strategy performs better and is the winning strategy. Although the average short-selling return of this portfolio is highly negative according to all measurements, and it yields the highest loss among all portfolios in 20% of quarters (25% of years), it is selected as the best portfolio. Thus, for a passive investor, full investment in bear LETFs can generate a return higher than the market in the short and long run, while an active investor merely needs to choose between short-selling a full bull or bear LETF portfolio for the next investment period. It is worth mentioning that if an investor picks the best portfolio for short-selling each quarter (year), they could earn an average cumulative return of 17.79% (38.1%), a Sharpe ratio of 11% (6.44%), and a Sortino ratio of 12.45% (6.72%) which is significantly higher than both the market and a full bear portfolio.

We then consider different short-selling costs for their returns on the full bear portfolio to examine if this strategy can still outperform the market. We summarize the results of this in table 5.

*** Insert Table 5 about here ***

The results show that if the short-selling cost is 4% per annum or less, the full bear portfolio is able to beat the market and earn a higher cumulative return as well as a risk-adjusted return in both quarterly and annual investment periods. However, the marginal profit is small. As the short-selling cost falls below 1%, the marginal difference of short-selling the portfolio and the index increases and becomes statistically significant at the 90% and 95% confidence levels.

4.2. Comparing 2x and 3x LETFs

To check the robustness of the results, we divide all LETFs into two groups based on their leveraged size. All the calculations are done for both of these groups separately and we summarize the results in table 6.

*** Insert Table 6 about here ***

The results in table 6 suggest that the best investment strategy for both 2x and 3x LEFTs is the full bear portfolio. Compared to the index, this strategy, in each quarter (year) is able to gain on average a 3.39% (12.58%) higher cumulative return, a 0.71% (0.77%) higher Sharpe ratio, and a 1.25% (0.99%) Sortino ratio for 2x LETFs over 9 years. Compared to the index, this strategy, in each quarter (year) is able to gain on average a 7.04% (23.82%) higher cumulative return, a 0.72% (0.65%) higher Sharpe ratio, and a 0.76% (0.64%) Sortino ratio for 3x LETFs for the same period. Although these differences are not statistically significant for the case of the Sharpe ratio only, they are economically significant. The average cumulative return of 2x LETF is lower than the 3x

LETFs, however, when the return is adjusted by the risk, 2x LETFs are slightly superior to 3x LETFs.

4.3. Comparing different types of LETFs

We then separate the LETFs into five groups based on their underlying asset types. We construct similar portfolios for each category and calculate various measurements, we show the results of this in table 7.

*** Insert Table 7 about here ***

Interestingly, the best investment strategy varies across LETFs with different underlying asset types. For LETFs whose underlying asset is categorized as equity, foreign equity, or real estate, the best investment strategy is still to short-sell the full bear LETF portfolio. The cumulative return of it is significantly higher than the market for all three of these categories according to the Sharpe and Sortino ratios. However, the positive difference is only statistically significant for equity LETFs. The results of quarterly data and annual data are similar in this respect too.

Conversely, short-selling the full bear portfolio is the least profitable strategy for LETFs with a type of currency as their underlying asset. For these ones, the winning portfolio is that in which investors short-sell the full bull portfolio. The cumulative return of this portfolio for currency LETFs is significantly higher than the market. All portfolios significantly outperform the market, in terms of cumulative returns, when the underlying asset is a commodity. Though short-selling the full bull portfolio earns the highest return, the difference between this and other strategies is small.

Among all types of LETFs, equity LETFs earn the highest cumulative return and risk-adjusted return, while currency LETFs earn the lowest return.

4.4. Market conditions and the short pairing strategy

In order to better analyze the winning strategy, we examine the portfolios during bullish and bearish market conditions. Table 8 shows the summary of these results.

*** Insert Table 8 about here ***

The results in table 8 show that in a bullish market, short-selling a full bear portfolio is the winning strategy. It can obtain, quarterly (yearly), an average cumulative return which is 11.9% (20.21%) higher than the marker, as well as a higher Sharpe ratio of 0.23% (0.71%), and Sortino ratio of 1.41% (1.27%), this positive return is statistically significant too. In a bearish market, short-selling the full bull portfolio is the best investment strategy. It can beat the market on average by 52.37% (90.72%), 31.35% (13.68%) 33.64%, (13.95%) based on cumulative return, Sharpe ratio, and Sortino ratio respectively during each quarter (year). By comparing the quarterly and annual short-selling returns of the full bull portfolio in a bearish market, we notice an obvious path dependency characteristic, while the short-selling returns of a full bear portfolio in a bullish market does not highlight this same characteristic. This may be due to the higher volatility of the market during bearish periods, which shows a higher risk for the full bull portfolio when it is selected as the winner.

Table 8 also shows that the average cumulative return that short-selling a full bear portfolio can yield when the market is bullish is 23.36% (46.8%) while the potential loss of this portfolio in a bearish market is -36% (-57.3%) in each quarter (year). It therefore seems as though the potential loss of such a portfolio surpasses its return, and short-selling the full bull portfolio might be a better investment strategy. It seems as though these results contradict the previous findings, and that choosing a full bear portfolio would be a better long-term strategy. However, when the risk-adjusted measures (the Sharpe ratio and the Sortino ratio) are considered, it is obvious that the earnings of the short full bear portfolio in a bullish market overcome its loss in a bearish market. For the Sharpe and Sortino ratios, the difference of profit during the bull market and loss during the bearish market for this portfolio is 3.32% (5.42%) and 7.89% (5.95%) respectively for each quarter (year). Thus, in the long term, when some investment intervals are bullish and some are bearish, this strategy can still be profitable and earn a positive risk-adjusted return. Similarly, the risk-adjusted return of shorting the full bull portfolio in a bearish market is less than its loss during the bullish market. Thus, we can conclude that short-selling the full bear portfolio is the best strategy for the long term.

4.5. Machine learning techniques and regression analysis

Based on our analysis, we conclude that that short-selling a full bear portfolio, on average, beats the market in the long run. However, the results also suggest that there are some quarters/years in which this strategy may not perform well. To investigate this further, we obtain an extensive set of potential explanatory variables and use them in a multinomial logistic regression. We first remove the detected multicollinearity issue. Table 9 shows the results of the correlation test.

*** Insert Table 9 about here ***

This table relates to yearly data; however, quarterly results are similar. We consider a correlation higher than 80% to be a serious multicollinearity problem and drop one of the correlated variables (the one with more dependency to other variables) from the model. After removing the correlated variables, we choose 15 variables. These variables are average index price, average index return, index trend, index volatility, price beta of bear LETFs, return beta of bear LETFs, the average trading volume of bull and bear LETFs, the average market capitalization of bear LETFs, the average of bid/ask spread for bull and bear LETFs, the average of net asset value for bull and bear LETFs, the average of turnover ratio for bear LETFs, and the average of net asset value to price ratio for bear LETFs. We use these variables along with a set of five dummy variables which show the type of underlying asset for the LETFs, a set of nine dummy variables which represent the years, and a set of four dummy variables showing the quarters (only for quarterly data). For the chosen variables, we perform a yeo-johnson transformation, scaling, and winsorizing. Since the number of explanatory variables is still high, to avoid the overfitting problem, we use five feature selection methods. Only quantitative variables are entered as the input in these models. After selecting the best set, we add different dummy variables to the model. We summarize the results of these techniques in table 10.

*** Insert Table 10 about here ***

We calculate the accuracy of the selected models using logistic regression and random forest. The results in table 10 show that the feature set selected by the Ridge model can achieve the highest accuracy for quarterly balanced portfolios, while the LASSO selected features can generate the highest accuracy for annual balanced portfolios. The selected features for quarterly data are bull net asset value, index return, index volatility, and return beta of bear LETFs for quarterly data. For annual data, the selected features are bull net asset value, index price, and return beta of bear LETFs for annual data. These selected features are able to achieve an accuracy score of 63.45% (88.57%) in the logistic model and 73.74% (92.38%) in the random forest for quarterly (annually) balanced portfolios. We conclude that these features are the main determinants of the winning strategy since they are mostly repeated in other selection methods. While index volatility and index return appear to be more important on a quarterly basis, and index price in annual data, the rest of the important features are the same in these two intervals. It is worth mentioning that if all 15 variables, along with dummies, are used in models without any preprocessing, the accuracy score of logistic regression would be 54% (63.45%) and the accuracy of random forest would be 64% (73.74%) for quarterly data (annual data).

Using the selected features as the explanatory variables, along with different sets of dummy variables, we estimate a multinomial logistic regression model to investigate the impact of these factors on determining the winning strategy in each quarter/year. We detail the mathematical forms of these regression lines in equations 16 and 17 and show their results in table 11.

*** Insert Table 11 about here ***

Results for quarterly data show that the correlation of LETFs return with their underlying index is a good determinant of whether investing these is profitable in the coming quarter. A low return correlation increases the probability that short-selling LETFs will be profitable in the next quarter by 0.9 for bull LETFs and by 1.09 for bear LETFs. Additionally, bull net asset value and index return are significant determinants for investing in the short full bull portfolio. An increase in bull net asset value or index return increases the probability that short-selling a full bull portfolio will be profitable in the next quarter by 0.545 and 0.443 respectively, compared to the short-selling full bear strategy. On the other hand, increasing index volatility is a strong sign of staying in cash and not short-selling the LETFs (especially bull LETFs). We capture seasonality effects with quarter dummy variables. According to the results, the first quarter is the best time to decide whether to short-sell LETFs for the coming quarter. However, there is no significant difference between the second and fourth quarters' investment options. Our results indicate that short-selling bull LETFs whose underlying assets are foreign equity, commodity, or currency can earn significantly higher returns compared to real estate LETFs, while there is no significant difference between real estate and equity LETFs in this regard.

Similarly, to the quarterly results, annual data shows that an increase in bull net asset value has a positive impact on the probability of choosing the short-selling bull LETFs in the next year. An increase in correlation of LETFs return and their corresponding indices, on the other hand, is a sign that short-selling them will not be profitable in the coming year. More specifically, one unit increase in the return correlation decreases the probability of choosing the short bull portfolio by 0.87 and short bear portfolio by 1.255 compared to staying in the cash option. Similarly, an increase in index return is a sign to stay in cash for the coming year. In other words, there is a significant negative relationship between index return and probability of choosing one of LETFs as the best compared to staying in the cash option. The coefficients of comparing real estate LETFs with other types of LETFs are not significant for yearly data.

5. Conclusion

This study aims to conduct comprehensive research on LETFs as a long-term investment tool. Previous studies on the topic conclude that these investment tools are only favorable in the short-term, as keeping them too long would lead to the decay of they achieved returns. This characteristic of LETFs is called path dependency (Guo and Leung, 2015, Trainor and Baryla, 2008).

Recently, a few studies have offered different investment strategies to construct LETF portfolios which are profitable in the long-term. These strategies include combining them with other investment tools (Peterburgsky, 2020), only investing in certain LETFs considering the market condition (Saini, 2019), or using a combination of bull and bear LETFs in a portfolio (Hessel et al., 2018).

However, the sample of these studies was limited to a number of commonly traded LETFs. Additionally, these studies mainly investigate different portfolios to find the best investment long-term strategy, however, less attention is paid to the determinant factors affecting the winning strategy. This study aims to fill this gap in the corresponding literature and provide an extensive study on LETFs.

The main idea of this study is based on the path dependency feature of LETFs, and stems from the idea that short-selling a portfolio which is meant to have negative returns could result in a profitable outcome. We examine out hypothesis by collecting 44 LETF pairs. We then create seven portfolios with different combinations of bull and bear LETFs which are rebalanced on a quarterly and annual basis. Overall, the results for quarterly data and annual data are parallel and there is no remarkable difference between them.

We calculate the performance of these portfolios using various measurements including the cumulative return, the Sharpe ratio, the Sortino ratio, and the expected VaR. Our results reveal that the best investment strategy is the short-selling of the full bear portfolio. Interestingly, portfolios with a combination of both bull and bear LETFs do not earn a very high return. This may be because the return of profitable LETF pairs is canceled out by the loss of the other pair, so the overall return of the portfolio is low.

We then compare the LETFs with 2x and 3x leverage levels. For both of these LETFs, the best investment strategy is short-selling the full bear LETF. Additionally, according to the results, although the cumulative return of 3x LETFs is higher than 2x LETFs, the risk-adjusted return of 3x LETFs is lower than 2x LETFs due to the higher risk of 3x LETFs.

LETFs with different underlying asset types are then compared. The results for Equity LETFs, foreign equity LETFs, and real estate LETFs are similar to the aforementioned findings. However, for commodity LETFs and currency LETFs, the best investment strategy is short-selling the full bull portfolio. This difference may be due to the price trend of the underlying index of these LETFs as commodity and currency indices do not change as much as the others and have a stable trend compared to the upward trend of other indices

It is possible to argue that the short-selling of bear LETFs is the best investment strategy because, during the period in which the study was conducted, the market remained mainly bullish. To further investigate this issue, observations are divided into two groups, one for the bullish market and one for the mainly bearish market. All calculations are performed for each group. Results show that although short-selling the bear LETFs is not a winning strategy in the bearish market, the risk-adjusted return of this strategy during the bullish market is higher than its loss during the bearish market. Thus, in the long horizon, this strategy can have a positive return. While other strategies, more specifically, short-selling the bull LETFs, yield a higher loss compared to their profit when the market condition is unfavorable.

We use a multinomial logistic regression to find the main determinants of the winning strategy. In this step, we use various machine learning techniques. We first employ feature selection methods such as LASSO, Ridge, Elastic net, RFE, and EFS to find the best sets of explanatory variables from the potential list. Next, we calculate the accuracy of the selected variables by the logistic regression and random forest methods. The set of features with the highest score is selected as the main independent variables in the model. The selected variables for quarterly data are the return beta for bear, the index return, the index volatility, and the bull net asset value, which are the output of the Ridge model. For annual balanced portfolios, the selected variables are the bull net asset value, the return beta for bear, and the index price, which are provided by the LASSO model.

These variables, along with dummy variables for year, quarter, and asset type are used in a multinomial logistic regression as variables to explain and predict the winning strategy for the coming investment interval. The results show that on both a quarterly and an annual basis, there is a positive and significant relationship between the return of short-selling bull LETFs and the net asset value of the bull. On the other hand, a decrease in the correlation of LETFs return and their indices also has a positive impact on the return of both LETFs. The results also imply seasonality effects in the first and third quarters. Additionally, the difference of real estate LETFs and other asset types (except for equity) are significant for quarterly data, while they become insignificant in annually balanced portfolios.

Further avenues of research would include studies which examine LETFs in Europe or Asia and compare their short-selling portfolios and determinant factors to those determined by this study. Additionally, other machine learning methods, such as deep learning or long short-term memory (LSTM) can be used to test if they can increase the accuracy of the models. Additionally, other potential affecting factors like expense ratio or availability of derivatives for the underlying index can be added to the model. Finally, other types of LETFs, like fixed-income or alternative LETFs can be added to the model and compared with the other mentioned types.

Table 1. Information of selected pair LETFs in the sample

Bull Name	Bear Name	Index Name	Size	Underlying asset
Direxion Daily MSCI Emerging Markets Bull 3X Shares	Direxion Daily MSCI Emerging Markets Bear 3X Shares	MSCI Emerging Markets Index	3x	Emerging and foreign market
ProShares Ultra Basic Material	ProShares Ultra Short Basic Material	Dow Jones U.S. Basic Materials Index	2x	Equity
ProShares Ultra Nasdaq Biotech	ProShares UltraShort Nasdaq Biotech	the Nasdaq Biotechnology Index	2x	Equity
ProShares Ultra Bloomberg Natural Gas	ProShares UltraShort Bloomberg Natural Gas	Bloomberg Natural Gas Subindex	2x	Commodity
Direxion Daily China 3x Bull Shares	Direxion Daily China 3x Bear Shares	FTSE China A50 Index	3x	Emerging and foreign market
ProShares Ultra FTSE China 50	ProShares UltraShort China 50	FTSE China A50 Index	2x	Emerging and foreign market
ProShares Ultra Consumer Goods	ProShares Ultra Short Consumer Goods	Dow Jones U.S. Consumer Goods Index	2x	Equity
ProShares Ultra Consumer Services	ProShares Ultra Short Consumer Services	Dow Jones U.S. Consumer Services Index	2x	Equity
ProShares Ultra Dow30	ProShares UltraShort Dow30	Dow Jones Industrial Average Index	2x	Equity
ProShares UltraPro Dow30	ProShares UltraPro Short Dow30	Dow Jones Industrial Average Index	3x	Equity
ProShares Ultra MSCI EAFE	ProShares UltraShort MSCI EAFE	MSCI EAFE Index	2x	Equity
ProShares Ultra MSCI Emerging Market	ProShares UltraShort MSCI Emerging Market	MSCI Emerging Markets Index	2x	Emerging and foreign market
Direxion Daily Energy Bull 2X	Direxion Daily Energy Bear 2X	Energy Select Sector Index	3x	Commodity
ProShares Ultra Euro	ProShares UltraShort Euro	U.S. Dollar price of the Euro	2x	Currency
ProShares Ultra Europe	ProShares UltraShort Europe	FTSE Developed Europe All Cap Index	2x	Emerging and foreign market
Direxion Daily Financial Bull	Direxion Daily Financial Bear	Russell 1000 Financials Industry Index	3x	Equity
ProShares Ultra Financials	ProShares Ultra Short Financials	Dow Jones U.S. Financials Index	2x	Equity
DB Gold Double Long ETN	DB Gold Double Short ETN	Deutsche Bank Liquid Commodityindex	2x	Commodity
ProShares Ultra Gold	ProShares UltraShort Gold	Dow Jones-UBS Gold Subindex	2x	Commodity
Direxion Daily Gold Miners Bull	Direxion Daily Gold Miners Bear	NYSE Arca Gold Miners Index	3x	Commodity
ProShares Ultra Health Care	ProShares Ultra Short Health Care	Dow Jones U.S. Health Care Index	2x	Equity
ProShares Ultra Industrials	ProShares Ultra Short Industrials	Dow Jones U.S. Industrials Index	2x	Equity

Table 1 Cont. Information of selected pair LETFs in the sample

Bull Name	Bear Name	Index Name	Size	Underlying asset
ProShares Ultra MSCI Japan	ProShare UltraShort MSCI Japan	MSCI Japan Index	2x	Emerging and foreign market
ProShares Ultra Midcap 400	ProShares UltraShort Midcap 400	S&P 400	2x	Equity
ProShares UltraPro MidCap400	ProShares UltraPro Short MidCap400	S&P 400	3x	Equity
ProShares Ultra QQQ	ProShares UltraShort QQQ	NASDAQ-100	2x	Equity
ProShares UltraPro QQQ	ProShares UltraPro Short QQQ	NASDAQ-100	3x	Equity
ProShares Ultra Oil & Gas	ProShares Ultra Short Oil & Gas	Dow Jones U.S. Oil & Gas Index	2x	Commodity
ProShares Ultra Real Estate	ProShares Ultra Short Real Estate	Dow Jones U.S. Real Estate Index	2x	Real Estate
ProShares Ultra Russell 2000	ProShares Ultra Short Russell 2000	Russell 2000 Index	2x	Equity
ProShares UltraPro Russell 2000	ProShares UltraPro Short Russell 2000	Russell 2000 Index	3x	Equity
Direxion Daily S&P 500 Bull 3X	Direxion Daily S&P 500 Bear 3X	S&P 500	3x	Equity
ProShares Ultra S&P 500	ProShares UltraShort S&P 500	S&P 500	2x	Equity
ProShares UltraPro S&P 500	ProShares UltraPro Short S&P 500	S&P 500	3x	Equity
direxion daily semiconductor bull 3x shares	Direxion daily semiconductor bear 3x shares	PHLX Semiconductor Sector	3x	Equity
ProShares Ultra Semiconductors	ProShares Ultra Short Semiconductors	Dow Jones U.S. Semiconductors Index	2x	Equity
ProShares Ultra Silver	ProShares UltraShort Silver	Bloomberg Silver Subindex	2x	Commodity
Direxion Daily Small Cap Bull	Direxion Daily Small Cap Bear	Russell 2000 Index	3x	Equity
ProShares Ultra Small Cap 600	ProShares Ultra Short Small Cap 600	S&P 600	2x	Equity
ProShares Ultra Technology	ProShares Ultra Short Technology	Dow Jones U.S. Technology Index	2x	Equity
ProShares Ultra Utilities	ProShares Ultra Short Utilities	Dow Jones U.S. Utilities Index	2x	Equity
ProShares Ultra Yen	ProShares UltraShort Yen	U.S. dollar price of the Japanese yen	2x	Currency
VelocityShares 3x Long Gold ETN	VelocityShares 3x Inverse Gold ETN	S&P GSCI Gold	3x	Commodity
VelocityShares 3x Long Silver ETN	VelocityShares 3x Inverse Silver ETN	S&P GSCI Silver ER	3x	Commodity

Table 2. Descriptive statistics for research variables

Variable	Equity	Foreign Market	Commodity	Currency	Real estate	Total
	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)
Bull price	41.960 (26.601)	54.794 (28.638)	761.644 (2754.637)	41.487 (26.873)	55.803 (15.257)	185.39 (1253.91)
Bear price	2503.156 (21604.063)	122.347 (187.645)	1717.044 (7705.413)	48.732 (27.152)	45.086 (27.396)	1851.63 (17063.8)
Index price	3315.347 (5684.881)	4387.260 (5317.901)	386.052 (341.841)	0.601 (0.596)	302.051 (35.917)	2666.22 (5034.47)
Bull Cu. return	0.349 (0.449)	0.125 (0.407)	-0.108 (0.445)	-0.058 (0.148)	0.139 (0.282)	0.205 (0.47)
Bear Cu. return	-0.342 (0.243)	-0.245 (0.267)	-0.089 (0.478)	0.047 (0.166)	-0.227 (0.147)	-0.259 (0.325)
Index Cu. return	0.146 (0.156)	0.090 (0.205)	-0.031 (0.218)	-0.016 (0.081)	0.054 (0.116)	0.094 (0.189)
Bull Sharpe	0.062 (0.060)	0.030 (0.058)	0.000 (0.052)	-0.023 (0.058)	0.046 (0.055)	0.041 (0.0644)
Bear Sharpe	-0.069 (0.060)	-0.039 (0.059)	-0.007 (0.052)	0.015 (0.060)	-0.056 (0.057)	-0.048 (0.0648)
Index Sharpe	0.058 (0.058)	0.033 (0.071)	-0.002 (0.050)	-0.017 (0.061)	0.034 (0.052)	0.038 (0.0643)
Bull Sortino	0.063 (0.061)	0.032 (0.063)	0.000 (0.052)	-0.023 (0.060)	0.045 (0.054)	0.042 (0.066)
Bear Sortino	-0.072 (0.062)	-0.038 (0.060)	-0.006 (0.054)	0.018 (0.064)	-0.057 (0.058)	-0.05 (0.067)
Index Sortino	0.057 (0.058)	0.037 (0.075)	-0.001 (0.051)	-0.014 (0.061)	0.033 (0.050)	0.039 (0.064)
Bull A. VaR	-0.041 (0.021)	-0.047 (0.018)	-0.062 (0.033)	-0.017 (0.004)	-0.033 (0.019)	-0.045 (0.025)
Bear A. VaR	-0.043 (0.022)	-0.048 (0.018)	-0.062 (0.033)	-0.017 (0.004)	-0.034 (0.018)	-0.046 (0.025)
Index A. VaR	-0.017 (0.009)	-0.018 (0.006)	-0.026 (0.013)	-0.008 (0.002)	-0.017 (0.010)	-0.019 (0.0099)
Bull H. VaR	-0.042 (0.021)	-0.046 (0.017)	-0.061 (0.030)	-0.017 (0.004)	-0.033 (0.016)	-0.045 (0.024)
Bear H. VaR	-0.041 (0.019)	-0.046 (0.017)	-0.062 (0.034)	-0.017 (0.004)	-0.032 (0.016)	-0.045 (0.024)
Index H. VaR	-0.018 (0.008)	-0.018 (0.007)	-0.026 (0.012)	-0.008 (0.002)	-0.017 (0.009)	-0.019 (0.0094)
Bull E. VaR	-0.061 (0.035)	-0.066 (0.028)	-0.087 (0.047)	-0.023 (0.006)	-0.050 (0.033)	-0.065 (0.038)

Table 2 Cont. Descriptive statistics for research variables

Variable	Equity	Foreign Market	Commodity	Currency	Real estate	Total
	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)	Ave. (St. Dev.)
Bear E. VaR	-0.057 (0.031)	-0.062 (0.025)	-0.086 (0.046)	-0.023 (0.007)	-0.044 (0.027)	-0.061 (0.036)
Index E. VaR	-0.026 (0.014)	-0.026 (0.011)	-0.036 (0.018)	-0.011 (0.003)	-0.025 (0.017)	-0.027 (0.015)
Bull Beta - price	0.974 (0.071)	0.827 (0.297)	0.905 (0.301)	0.981 (0.028)	0.992 (0.011)	0.94 (0.189)
Bear Beta - price	-0.930 (0.090)	-0.781 (0.292)	-0.845 (0.287)	-0.981 (0.031)	-0.884 (0.152)	-0.894 (0.191)
Bull Beta - return	0.953 (0.101)	0.598 (0.145)	0.884 (0.222)	0.905 (0.063)	0.997 (0.002)	0.889 (0.182)
Bear Beta - return	-0.929 (0.132)	-0.594 (0.147)	-0.882 (0.220)	-0.969 (0.010)	-0.996 (0.002)	-0.877 (0.191)
Bull vol.	4314458.094 (11701182.002)	213823.889 (452194.023)	354855.091 (749259.426)	4889.222 (6324.828)	103881.222 (72678.241)	2677235.87 (9265282.53)
Bear vol.	488371.103 (2514443.859)	49345.778 (85112.616)	182066.481 (353673.750)	263159.944 (276985.645)	59729.667 (49493.577)	347543.56 (1958301.68)
Bull Market Cap	572.093 (951.186)	90.661 (107.162)	288.107 (335.816)	5.325 (2.867)	208.146 (88.888)	415.61 (777.86)
Bear Market Cap	173.422 (292.071)	42.802 (30.958)	49.789 (58.955)	287.032 (192.797)	39.809 (26.189)	133.30 (241.26)
Bull spread	0.181 (0.370)	1.053 (2.253)	1.938 (6.108)	1.977 (5.175)	0.063 (0.030)	0.4613 (2.235)
Bear spread	19.402 (163.563)	1.762 (6.885)	3.130 (13.589)	0.032 (0.029)	0.054 (0.038)	5.457 (49.487)
Bull Turnover	85068134.795 (211196255.66)	7961960.427 (13218151.6)	43621523.53 (99500804.5)	113295.438 (111320.056)	4938111.044 (2984731.67)	60564359.78 (172132485.1)
Bear Turnover	43051337.184 (95938331.469)	4081861.266 (7885544.96)	19727848.4 (54583723.7)	7612628.956 (6837193.14)	2339535.522 (1911589.39)	30539696.12 (79659197.24)
Bull NAV	41.949 (26.581)	54.793 (28.658)	759.252 (2757.832)	41.462 (26.816)	55.814 (15.246)	184.92 (1255.07)
Bear NAV	2505.545 (21613.453)	122.191 (186.520)	1717.010 (7712.351)	48.737 (27.155)	45.084 (27.368)	1853.031 (17071.56)
Bull P/NAV	1.000 (0.004)	1.000 (0.003)	1.010 (0.094)	1.000 (0.002)	1.000 (0.001)	1.002 (0.422)
Bear P/NAV	0.999 (0.006)	1.000 (0.002)	0.991 (0.071)	1.000 (0.000)	1.000 (0.001)	0.998 (0.0322)
Number of Obs.	232 (26 pairs)	56 (6 pairs)	77 (9 pairs)	18 (2 pairs)	9 (1 pairs)	392 (44 pairs)

All calculations are on annual basis

Table 3. Different portfolios with different combination of bull and bear pairs and return calculation.

Portfolio	Bull weight	Bear weight	Performance of short-selling the portfolio
full bull (1:0)	100%	0%	(-1) * bull performance
75%bull – 25% bear (3:1)	75%	25%	(-0.75) * bull performance+ (-0.25) * bear performance
67%bull – 33% bear (2:1)	67%	33%	(-0.67) * bull performance + (-0.33) * bear performance
50%bull – 50% bear (1:1)	50%	50%	(-0.50) * bull performance + (-0.50) * bear performance
33%bull – 67% bear (1:2)	33%	67%	(-0.33) * bull performance + (-0.67) * bear performance
25%bull – 75% bear (1:3)	25%	75%	(-0.25) * bull performance + (-0.75) * bear performance
full bear (0:1)	0%	100%	(-1) * bear performance

Table 4. Summary statistics of short-selling returns and risks for the research portfolios

	Measures	Short-selling Portfolios						
		1:0	3:1	2:1	1:1	1:2	1:3	0:1
Cumulative Return	Average %	-5.58 (-20.48)	-2.42 (-8.90)	-1.41 (-5.19)	0.74 (2.69)	2.88 (10.57)	3.89 (14.28)	7.05 (25.87)
	Standard deviation	0.238 (0.47)	0.131 (0.297)	0.099 (0.246)	0.052 (0.158)	0.085 (0.146)	0.116 (0.175)	0.222 (0.325)
	Frequency of being best	469 (99)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1099 (293)
	t value – vs index	-12.183 (-11.68)	-11.518 (-10.29)	-10.649 (-9.32)	-5.784 (-5.407)	1.454 (0.958)	3.848 (3.734)	7.541 (8.649)
	p-value - vs index	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.146 (0.338)	0.000 (0.000)	0.000 (0.000)
Sharpe ratio	Average %	-5.61 (-4.36)	-2.75 (-2.12)	-1.83 (-1.41)	0.11 (0.12)	2.05 (1.64)	2.97 (2.36)	5.82 (4.59)
	Standard deviation	0.122 (0.065)	0.062 (0.033)	0.043 (0.023)	0.010 (0.006)	0.042 (0.023)	0.061 (0.033)	0.121 (0.065)
	Frequency of being best	467 (99)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1101 (293)
	t value – vs index	-24.646 (-17.83)	-22.82 (-16.409)	-21.34 (-15.292)	-16.232 (-11.49)	-9.405 (-6.468)	-6.236 (-4.148)	1.697 (1.576)
	p-value - vs index	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.09 (0.115)
Sortino ratio	Average %	-4.94 (-4.41)	-1.94 (-2.11)	-0.98 (-1.37)	1.06 (0.19)	3.10 (1.76)	4.06 (2.5)	7.06 (4.8)
	Standard deviation	0.127 (0.067)	0.065 (0.034)	0.047 (0.024)	0.030 (0.008)	0.059 (0.025)	0.078 (0.035)	0.141 (0.068)
	Frequency of being best	466 (99)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1102 (293)
	t value – vs index	-23.706 (-17.71)	-21.468 (-16.33)	-19.818 (-15.21)	-14.517 (-11.31)	-7.933 (-6.184)	-4.98 (-3.842)	2.225 (1.84)
	p-value - vs index	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.026 (0.066)
Expected VaR	Average %	-5.774 (-6.09)	-5.834 (-6.18)	-5.853 (-6.21)	-5.894 (-6.27)	-5.934 (-6.33)	-5.953 (-6.36)	-6.031 (-6.45)
	Standard deviation	0.039 (0.036)	0.039 (0.036)	0.040 (0.036)	0.040 (0.037)	0.041 (0.037)	0.042 (0.037)	0.044 (0.038)
	Frequency of being best	855 (249)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	719 (143)
	t value – vs index	-29.95 (-17.233)	-30.266 (-17.5)	-30.274 (-17.57)	-30.151 (-17.7)	-29.853 (-17.8)	-29.661 (-17.84)	-28.88 (-17.9)
	p-value - vs index	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)

The values outside the parentheses are related to quarterly data and the values inside the parentheses are annual results.

Table 5. Summary statistics of different short-selling costs for the full bear portfolio.

Measures		The maximum cost at which the difference of short-selling return and index is		
		Significant at 95%	Significant at 90%	Insignificant positive
Cumulative return	Cost	14% (15%)	15% (16%)	20% (20%)
	Average difference	1.344% (4.34%)	1.128 (3.57%)	0.06% (0.45%)
	Test value	2.14 (2.088)	1.79 (1.705)	0.104 (0.228)
	p-value	0.032 (0.037)	0.073 (0.884)	0.917 (0.819)
Sharpe ratio	Cost	-* (-)	0.1% (-)	3% (4%)
	Average difference	- (-)	0.717% (-)	0.15% (0.02%)
	Test value	- (-)	1.65 (-)	0.346 (0.0466)
	p-value	- (-)	0.098 (-)	0.729 (0.962)
Sortino ratio	Cost	0.6% (-)	1% (0.5%)	5% (4%)
	Average difference	1.1% (-)	0.874% (0.793%)	0.011% (0.128%)
	Test value	1.988 (-)	1.806 (1.671)	0.023 (0.273)
	p-value	0.046 (-)	0.071 (0.095)	0.981 (0.784)

The values outside the parentheses are related to quarterly data and the values inside the parentheses are annual results

* Before considering any short-selling cost, the difference of short-selling return and index return is not significant at the corresponding level.

Table 6. Summary statistics of short-selling returns for the research portfolios for 2x and 3x LETFs

	Measures	Short-selling Portfolios						
		1:0	3:1	2:1	1:1	1:2	1:3	0:1
2x	Average Cu. Return %	-4.5 (-17.431)	-1.97 (-7.73)	-1.16 (-4.625)	0.56 (1.971)	2.28 (8.568)	3.09 (11.673)	5.62 (21.37)
	t value for Cu. Return	-10.271*** (-10.121***)	-9.773*** (-9.107***)	-9.056*** (-8.335***)	-5.280*** (-5.169***)	0.149 (-0.169)	2.069** (2.005**)	5.308*** (6.097***)
3x	Average Cu. Return %	-7.62 (-26.3)	-3.27 (-11.12)	-1.88 (-6.26)	1.07 (4.062)	4.03 (14.384)	5.42 (19.241)	9.76 (34.42)
	t value for Cu. Return	-7.229*** (-6.655***)	-6.683*** (-5.649***)	-6.098*** (-5.006***)	-2.853*** (-2.521**)	1.938* (1.576)	3.385*** (3.37***)	5.391*** (6.294***)
2x	Average Sharpe ratio %	-5.5 (-4.293)	-2.72 (-2.109)	-1.831 (-1.41)	0.057 (0.075)	1.946 (1.560)	2.835 (2.259)	5.612 (4.44)
	t value for Sharpe ratio	-19.253*** (-13.860***)	-17.788*** (-12.725***)	-16.617*** (-11.843***)	-12.618*** (-8.861***)	-7.301*** (-4.934***)	-4.829*** (-3.123***)	1.318 (1.334)
3x	Average Sharpe ratio %	-5.812 (-4.5)	-2.804 (-2.15)	-1.841 (-1.4)	0.205 (0.193)	2.25 (1.789)	3.213 (2.539)	6.22 (4.885)
	t value for Sharpe ratio	-15.446*** (-11.254***)	-14.370*** (-10.407***)	-13.458*** (-9.721***)	-10.271*** (-7.364***)	-5.973*** (-4.222***)	-3.973*** (-2.764***)	0.982 (0.841)
2x	Average Sortino ratio %	-4.808 (-4.33)	-1.857 (-2.06)	-0.912 (-1.34)	1.095 (0.203)	3.101 (1.742)	4.046 (2.467)	6.997 (4.73)
	t value for Sortino ratio	-18.415*** (-13.748***)	-16.609*** (-12.642***)	-15.282*** (-11.738***)	-11.041*** (-8.625***)	-5.837*** (-4.563***)	-3.527*** (-2.729***)	2.047** (1.667*)
3x	Average Sortino ratio %	-5.2 (-4.575)	-2.09 (-2.199)	-1.104 (-1.44)	1.001 (0.177)	3.105 (1.793)	4.096 (2.553)	7.191 (4.93)
	t value for Sortino ratio	-15.009*** (-11.198***)	-13.680*** (-10.382***)	-12.699*** (-9.712***)	-9.523*** (-7.388***)	-5.492*** (-4.257***)	-3.647*** (-2.798***)	0.950 (0.821)

The values outside the parentheses are related to quarterly data and the values inside the parentheses are annual results

*** Significant at 99%

** Significant at 95%

* Significant at 90%

Table 7. Summary statistics of short-selling returns for the research portfolios for LETFs with different underlying assets.

	Measures	Short-selling Portfolios						
		1:0	3:1	2:1	1:1	1:2	1:3	0:1
Equity	Average Cu. Return %	-8.88 (-34.9)	-4.18 (-17.6)	-2.7 (-12.09)	0.52 (-0.33)	3.714 (11.42)	5.218 (16.95)	9.92 (34.24)
	t value for Cu. Return	-16.2*** (-15.9***)	-15.6*** (-14.7***)	-14.9*** (-13.9***)	-9.7*** (-10.2***)	0.259 (-2.52**)	3.95*** (1.79*)	9.6*** (10.383***)
Foreign equity	Average Cu. Return %	-3.73 (-12.5)	-1.18 (-3.22)	-0.36 (-0.26)	1.37 (6.024)	3.106 (12.31)	3.923 (15.27)	6.47 (24.51)
	t value for Cu. Return	-3.720*** (-3.433***)	-3.239*** (-2.722***)	-2.849*** (-2.306**)	-1.213 (-0.906)	0.95 (1.032)	1.686 (1.855*)	2.832*** (3.348***)
Commodity	Average Cu. Return %	1.330 (10.78)	1.198 (10.32)	1.156 (10.17)	1.066 (9.85)	0.976 (9.52)	0.934 (9.37)	0.802 (8.9)
	t value for Cu. Return	0.950 (2.436**)	1.405 (3.430***)	1.594 (3.773***)	1.737* (4.043***)	1.303 (3.402***)	1.081 (2.993***)	0.628 (1.982*)
Currency	Average Cu. Return %	1.402 (5.79)	0.737 (3.166)	0.525 (2.326)	0.072 (0.542)	-0.38 (-1.24)	-0.59 (-2.08)	-1.25 (-4.7)
	t value for Cu. Return	1.555 (1.804*)	1.557 (1.828*)	1.470 (1.735*)	0.859 (1.069)	0.010 (0.143)	-0.262 (-0.166)	-0.708 (-0.694)
Real estate	Average Cu. Return %	-3.74 (-13.9)	-1.34 (-4.76)	-0.58 (-1.83)	1.051 (4.383)	2.68 (10.6)	3.446 (13.53)	5.84 (22.67)
	t value for Cu. Return	-1.881* (-1.793*)	-1.532 (-1.296)	-1.284 (-1.033)	-0.314 (-0.183)	0.895 (1.048)	1.278 (1.613)	1.816* (2.611**)
Equity	Average Sharpe ratio %	-8.2 (-6.445)	-4.1 (-3.166)	-2.73 (-2.12)	0.095 (0.112)	2.916 (2.341)	4.244 (3.39)	8.393 (6.67)
	t value for Sharpe ratio	-29.933*** (-22.249***)	-27.875*** (-20.694***)	-26.122*** (-19.367***)	-19.975*** (-14.704***)	-11.675*** (-8.418***)	-7.808*** (-5.505***)	1.781* (1.659*)
Foreign equity	Average Sharpe ratio %	-4.26 (-3.22)	-2.01 (-1.51)	-1.29 (-0.96)	0.240 (0.2)	1.772 (1.364)	2.492 (1.911)	4.744 (3.62)
	t value for Sharpe ratio	-7.647*** (-5.161***)	-6.910*** (-4.570***)	-6.430*** (-4.224***)	-4.950*** (-3.204***)	-3.066*** (-1.947*)	-2.175** (-1.354)	0.163 (0.220)
Commodity	Average Sharpe ratio %	-0.349 (-0.2)	-0.11 (-0.03)	-0.034 (0.02)	0.129 (0.137)	0.291 (0.252)	0.368 (0.306)	0.606 (0.48)
	t value for Sharpe ratio	-0.401 (-0.049)	-0.201 (0.196)	-0.107 (0.297)	0.126 (0.515)	0.346 (0.677)	0.426 (0.723)	0.580 (0.774)

Table 7 Cont. Summary statistics of short-selling returns for the research portfolios for LETFs with different underlying assets.

	Measures	Short-selling Portfolios						
		1:0	3:1	2:1	1:1	1:2	1:3	0:1
Currency	Average Sharpe ratio %	1.813 (1.721)	0.782 (0.752)	0.452 (0.442)	-0.25 (-0.22)	-0.95 (-0.88)	-1.28 (-1.19)	-2.3 (-2.16)
	t value for Sharpe ratio	1.529 (1.655)	1.355 (1.472)	1.240 (1.351)	0.884 (0.974)	0.445 (0.506)	0.246 (0.293)	-0.240 (-0.233)
Real estate	Average Sharpe ratio %	-5.96 (-4.87)	-2.88 (-2.34)	-1.9 (-1.534)	0.188 (0.188)	2.277 (1.91)	3.26 (2.72)	6.332 (5.25)
	t value for Sharpe ratio	-3.848*** (-3.068***)	-3.448*** (-2.75**)	-3.162*** (-2.515**)	-2.221** (-1.732)	-1.022 (-0.742)	-0.486 (-0.307)	0.783 (0.699)
Equity	Average Sortino ratio %	-7.76 (-6.59)	-3.35 (-3.21)	-1.94 (-2.13)	1.05 (0.171)	4.044 (2.47)	5.453 (3.552)	9.856 (6.93)
	t value for Sortino ratio	-29.228*** (-22.279***)	-26.025*** (-20.756***)	-23.775*** (-19.370***)	-16.958*** (-14.438***)	-9.064*** (-7.911***)	-5.643*** (-4.965***)	2.597*** (2.078**)
Foreign equity	Average Sortino ratio %	-3.67 (-3.14)	-1.371 (-1.4)	-0.64 (-0.85)	0.926 (0.33)	2.489 (1.511)	3.224 (2.067)	5.522 (3.8)
	t value for Sortino ratio	-7.601*** (-5.137***)	-6.859*** (-4.567***)	-6.393*** (-4.232***)	-4.982*** (-3.246***)	-3.207*** (-2.029**)	-2.363** (-1.454)	-0.111 (0.083)
Commodity	Average Sortino ratio %	0.73 (-0.135)	0.939 (0.034)	1.006 (0.087)	1.15 (0.202)	1.293 (0.317)	1.361 (0.371)	1.572 (0.54)
	t value for Sortino ratio	-0.272 (-0.020)	-0.091 (0.227)	-0.009 (0.328)	0.190 (0.548)	0.373 (0.710)	0.438 (0.754)	0.553 (0.794)
Currency	Average Sortino ratio %	3.891 (2.146)	2.643 (1.12)	2.244 (0.791)	1.395 (0.093)	0.547 (-0.6)	0.148 (-0.93)	-1.1 (-1.96)
	t value for Sortino ratio	1.641 (1.677)	1.552 (1.520)	1.471 (1.408)	1.149 (1.034)	0.637 (0.540)	0.381 (0.311)	-0.253 (-0.251)
Real estate	Average Sortino ratio %	-5.46 (-5.05)	-2.34 (-2.47)	-1.34 (-1.64)	0.785 (0.111)	2.909 (1.863)	3.908 (2.688)	7.03 (5.266)
	t value for Sortino ratio	-3.780*** (-3.085***)	-3.332*** (-2.8**)	-3.028*** (-2.573**)	-2.07** (-1.782)	-0.911 (-0.75)	-0.404 (-0.298)	0.79 (0.733)

he values outside the parentheses are related to quarterly data and the values inside the parentheses are annual results

*** Significant at 99%

** Significant at 95%

* Significant at 90%

Table 8. Summary statistics of short-selling returns for the research portfolios in different market situations.

	Measures	Short-selling Portfolios						
		1:0	3:1	2:1	1:1	1:2	1:3	0:1
Bull market	Average Cu. Return %	-26.7 (-70.9)	-14.2 (-41.5)	-10.2 (-32.1)	-1.66 (-12.1)	6.85 (7.943)	10.85 (17.36)	23.36 (46.8)
	t value for Cu. Return	-33.46*** (-24.29***)	-33.628*** (-24.11***)	-33.016*** (-23.846***)	-27.225*** (-22.059***)	-10.149*** (-14.798***)	-1.203 (-7.775***)	14.834*** (11.701***)
Bear market	Average Cu. Return %	36.18 (59.81)	18.1 (30.531)	12.3 (21.16)	0.00 (1.248)	-12.3 (-18.7)	-18.1 (-28.3)	-36.2 (-57.3)
	t value for Cu. Return	23.172*** (13.773***)	22.933*** (13.349***)	20.388*** (11.305***)	10.565*** (5.618***)	2.039** (1.581)	-0.651 (0.325)	-5.673*** (-2.106*)
Bull market	Average Sharpe ratio %	-18 (-11.6)	-8.91 (-5.78)	-6 (-3.905)	0.162 (0.085)	6.331 (4.075)	9.234 (5.953)	18.31 (11.82)
	t value for Sharpe ratio	-60.518*** (-39.159***)	-57.638*** (-37.253***)	-54.593*** (-35.256***)	-43.117*** (-27.653***)	-26.815*** (-16.764***)	-19.063*** (-11.594***)	0.383 (1.255)
Bear market	Average Sharpe ratio %	15.435 (6.98)	7.83 (3.636)	5.395 (2.566)	0.222 (0.291)	-4.95 (-1.98)	-7.38 (-3.05)	-14.99 (-6.4)
	t value for Sharpe ratio	32.037*** (15.902***)	30.95*** (18.362***)	29.441*** (18.952***)	23.472*** (17.137***)	14.943*** (10.052***)	10.924*** (6.78***)	0.927 (0.361)
Bull market	Average Sortino ratio %	-17 (-11.827)	-7.36 (-5.76)	-4.24 (-3.81)	2.309 (0.312)	8.882 (4.439)	11.975 (6.38)	21.64 (12.45)
	t value for Sortino ratio	-56.485*** (-39.640***)	-48.081*** (-36.812***)	-43.06*** (-34.182***)	-29.992*** (-25.265***)	-16.875*** (-14.27***)	-11.467*** (-9.47***)	1.559 (1.942*)
Bear market	Average Sortino ratio %	19.133 (7.19)	10.91 (3.765)	8.28 (2.668)	2.69 (0.339)	-2.9 (-1.991)	-5.532 (-3.1)	-13.75 (-6.5)
	t value for Sortino ratio	25.415*** (14.085***)	26.748*** (15.755***)	26.983*** (16.18***)	25.875*** (15.259***)	19.632*** (9.849***)	14.792*** (6.797***)	0.910 (0.294)

The values outside the parentheses are related to quarterly data and the values inside the parentheses are annual results.
*** Significant at 99%
** Significant at 95%
* Significant at 90%

Table 9. Summary of pairwise correlation coefficients for the potential explanatory variables on annual basis

	Bull Ave Price	Bear Ave Price	Index Ave Price	Bull Ave Return	Bear Ave Return	Index Ave Return	Bull Analytical VaR Long	Bear Analytical VaR Long	Index Analytical VaR	Bull VaR Long	Bear VaR Long	Index VaR	Bull Expected VaR Long	Bull Expected VaR Long	Index EVaR	Bull Trend	Bear Trend	Index Trend	Bull Volatility	Bear Volatility	Index Volatility	Bull Beta Price	Bear Beta Price	Bull Beta Return	Bear Beta Return	Bull Ave Volume	Bear Ave Volume	Bull Ave MCap	Bear Ave MCap	Bull Ave Spread	Bear Ave Spread	Bull Ave Turnover	Bear Ave Turnover	Bull Ave PTNAV	Bear Ave PTNAV	Bull NAV	Bear NAV	Bull Ave NAV	Bear Ave NAV
Bull Ave Price	1	0.06	0.05	0.18	0.16	0.16	0.3	0.29	0.32	0.29	0.31	0.31	0.25	0.25	0.26	0.09	0.09	0.08	0.3	0.3	0.32	0.01	0.06	0.06	0.07	0.05	0.03	0.04	0.05	0.8	0.01	0.02	0.03	0.01	0.01	0.98	0.07	1	0.06
Bear Ave Price	0.06	1	0.04	0.04	0.05	0.02	0.17	0.18	0.1	0.18	0.19	0.09	0.15	0.15	0.07	0.08	0.08	0.08	0.18	0.18	0.1	0.02	0.06	0.07	0.07	0.2	0.04	0.02	0.02	0.02	0.37	0.01	0.01	0	0	0.04	0.99	0.06	1
Index Ave Price	0.05	0.04	1	0.1	0.09	0.1	0.02	0	0.06	0.02	0.02	0.05	0	0	0.03	0.07	0.07	0.03	0.01	0.01	0.05	0.11	0.09	0.05	0.03	0.03	0.14	0.1	0.08	0.09	0.03	0.07	0.08	0.02	0.01	0.04	0.04	0.05	0.04
Bull Ave Return	0.18	0.04	0.1	1	0.99	0.93	0.33	0.17	0.43	0.31	0.18	0.4	0.28	0.28	0.38	0.28	0.29	0.3	0.25	0.25	0.35	0.01	0.08	0.08	0.06	0.23	0.01	0.14	0.12	0.22	0.07	0.18	0.16	0.01	0	0.11	0.01	0.18	0.04
Bear Ave Return	0.16	0.05	0.09	0.99	1	0.93	0.28	0.13	0.39	0.27	0.13	0.36	0.24	0.24	0.34	0.28	0.29	0.3	0.21	0.21	0.31	0.01	0.07	0.1	0.07	0.24	0.02	0.15	0.12	0.2	0.07	0.19	0.17	0.02	0	0.1	0	0.16	0.05
Index Ave Return	0.16	0.02	0.1	0.93	0.93	1	0.33	0.18	0.4	0.32	0.18	0.37	0.3	0.3	0.35	0.31	0.31	0.31	0.26	0.26	0.32	0.03	0.16	0.09	0.06	0.17	0	0.13	0.09	0.21	0.08	0.13	0.12	0.03	0.03	0.1	0.02	0.16	0.02
Bull Analytical VaR Long	0.3	0.17	0.02	0.33	0.28	0.33	1	0.99	0.86	0.98	0.97	0.84	0.98	0.98	0.83	0.09	0.08	0.09	1	1	0.85	0.02	0.2	0.06	0.09	0.05	0.06	0.12	0.1	0.23	0.04	0.21	0.16	0.05	0.02	0.25	0.18	0.3	0.17
Bear Analytical VaR Long	0.29	0.18	0	0.17	0.13	0.18	0.99	1	0.82	0.97	0.98	0.81	0.97	0.97	0.8	0.04	0.03	0.04	1	1	0.83	0.02	0.19	0.08	0.11	0.1	0.06	0.15	0.08	0.2	0.06	0.25	0.2	0.05	0.01	0.24	0.19	0.29	0.18
Index Analytical VaR	0.32	0.1	0.06	0.43	0.39	0.4	0.86	0.82	1	0.82	0.8	0.98	0.83	0.83	0.98	0.1	0.08	0.08	0.84	0.84	1	0.06	0.22	0.03	0.05	0.06	0.05	0.01	0.17	0.31	0.04	0.04	0.01	0.02	0.01	0.27	0.11	0.32	0.1
Bull VaR Long	0.29	0.18	0.02	0.31	0.27	0.32	0.98	0.97	0.82	1	0.95	0.84	0.98	0.98	0.81	0.08	0.07	0.08	0.98	0.98	0.82	0.01	0.17	0.09	0.11	0.06	0.05	0.14	0.1	0.21	0.04	0.23	0.18	0.05	0.01	0.24	0.18	0.29	0.18
Bear VaR Long	0.31	0.19	0.02	0.18	0.13	0.18	0.97	0.98	0.8	0.95	1	0.79	0.93	0.93	0.77	0.04	0.03	0.04	0.98	0.98	0.81	0.02	0.2	0.07	0.09	0.09	0.08	0.12	0.08	0.22	0.06	0.23	0.19	0.05	0.02	0.26	0.2	0.31	0.19
Index VaR	0.31	0.09	0.05	0.4	0.36	0.37	0.84	0.81	0.98	0.84	0.79	1	0.83	0.83	0.98	0.08	0.06	0.06	0.83	0.83	0.98	0.03	0.18	0.07	0.08	0.05	0.04	0.03	0.17	0.3	0.04	0.06	0.03	0.04	0.02	0.26	0.1	0.31	0.09
Bull Expected VaR Long	0.25	0.15	0	0.28	0.24	0.3	0.98	0.97	0.83	0.98	0.93	0.83	1	1	0.83	0.08	0.06	0.09	0.98	0.98	0.83	0.05	0.2	0.07	0.1	0.07	0.04	0.16	0.08	0.17	0.03	0.24	0.19	0.01	0	0.2	0.15	0.25	0.15
Bull Expected VaR Long	0.25	0.15	0	0.28	0.24	0.3	0.98	0.97	0.83	0.98	0.93	0.83	1	1	0.83	0.08	0.06	0.09	0.98	0.98	0.83	0.05	0.2	0.07	0.1	0.07	0.04	0.16	0.08	0.17	0.03	0.24	0.19	0.01	0	0.2	0.15	0.25	0.15
Index EVaR	0.26	0.07	0.03	0.38	0.34	0.35	0.83	0.8	0.98	0.81	0.77	0.98	0.83	0.83	1	0.1	0.08	0.08	0.82	0.82	0.98	0.04	0.17	0.05	0.06	0.05	0.03	0.05	0.16	0.25	0.02	0.06	0.02	0.02	0	0.22	0.08	0.26	0.07
Bull Trend	0.09	0.08	0.07	0.28	0.28	0.31	0.09	0.04	0.1	0.08	0.04	0.08	0.08	0.08	0.1	1	0.95	0.87	0.07	0.06	0.07	0.03	0.07	0.08	0.08	0	0.01	0.01	0.03	0.12	0.07	0.01	0.01	0.11	0.07	0.07	0.06	0.09	0.08
Bear Trend	0.09	0.08	0.07	0.29	0.29	0.31	0.08	0.03	0.08	0.07	0.03	0.06	0.06	0.06	0.08	0.95	1	0.86	0.05	0.05	0.05	0.02	0.05	0.1	0.1	0	0.02	0.01	0.02	0.12	0.07	0.01	0.01	0.11	0.08	0.07	0.07	0.09	0.08
Index Trend	0.08	0.08	0.03	0.3	0.3	0.31	0.09	0.04	0.08	0.08	0.04	0.06	0.09	0.09	0.08	0.87	0.86	1	0.07	0.07	0.05	0	0.01	0.05	0.06	0.03	0.03	0.01	0.01	0.12	0.08	0	0.01	0.01	0.05	0.07	0.07	0.08	0.08
Bull Volatility	0.3	0.18	0.01	0.25	0.21	0.26	1	1	0.84	0.98	0.98	0.83	0.98	0.98	0.82	0.07	0.05	0.07	1	1	0.85	0.02	0.2	0.07	0.1	0.07	0.06	0.13	0.09	0.21	0.05	0.23	0.18	0.06	0.03	0.24	0.19	0.3	0.18
Bear Volatility	0.3	0.18	0.01	0.25	0.21	0.26	1	1	0.84	0.98	0.98	0.83	0.98	0.98	0.82	0.06	0.05	0.07	1	1	0.85	0.02	0.19	0.08	0.1	0.07	0.06	0.13	0.09	0.21	0.05	0.23	0.18	0.05	0.01	0.25	0.19	0.3	0.18
Index Volatility	0.32	0.1	0.05	0.35	0.31	0.32	0.85	0.83	1	0.82	0.81	0.98	0.83	0.83	0.98	0.07	0.05	0.05	0.85	0.85	1	0.06	0.21	0.04	0.06	0.05	0.05	0.02	0.17	0.3	0.04	0.06	0.02	0.03	0.01	0.26	0.12	0.32	0.1
Bull Beta Price	0.01	0.02	0.11	0.01	0.01	0.03	0.02	0.02	0.06	0.01	0.02	0.03	0.05	0.05	0.04	0.03	0.02	0	0.02	0.02	0.06	1	0.85	0.61	0.59	0.07	0.06	0.1	0.1	0.02	0.02	0.08	0.09	0.06	0.3	0.01	0.02	0.01	0.02
Bear Beta Price	0.06	0.06	0.09	0.08	0.07	0.16	0.2	0.19	0.22	0.17	0.2	0.18	0.2	0.2	0.17	0.07	0.05	0.01	0.2	0.19	0.21	0.85	1	0.53	0.49	0.06	0.07	0.08	0.09	0.05	0	0.04	0.05	0.07	0.33	0.05	0.07	0.06	0.06
Bull Beta Return	0.06	0.07	0.05	0.08	0.1	0.09	0.06	0.08	0.03	0.09	0.07	0.07	0.07	0.07	0.05	0.08	0.1	0.05	0.07	0.08	0.04	0.61	0.53	1	0.92	0.2	0.13	0.3	0.23	0.02	0.04	0.24	0.26	0.03	0.15	0.06	0.07	0.06	0.07
Bear Beta Return	0.07	0.07	0.03	0.06	0.07	0.06	0.09	0.11	0.05	0.11	0.09	0.08	0.1	0.1	0.06	0.08	0.1	0.06	0.1	0.1	0.06	0.59	0.49	0.92	1	0.22	0.16	0.31	0.28	0.01	0.03	0.26	0.28	0.04	0.09	0.07	0.07	0.07	0.07
Bull Ave Volume	0.05	0.2	0.03	0.23	0.24	0.17	0.05	0.1	0.06	0.06	0.09	0.05	0.07	0.07	0.05	0	0	0.03	0.07	0.07	0.05	0.07	0.06	0.2	0.22	1	0.14	0.48	0.53	0.08	0.04	0.72	0.7	0	0	0.05	0.19	0.05	0.2
Bear Ave Volume	0.03	0.04	0.14	0.01	0.02	0	0.06	0.06	0.05	0.05	0.08	0.04	0.04	0.04	0.03	0.01	0.02	0.03	0.06	0.06	0.05	0.06	0.07	0.13	0.16	0.14	1	0.6	0.66	0.03	0.03	0.38	0.5	0.01	0.01	0.02	0.04	0.03	0.04
Bull Ave MCap	0.04	0.02	0.1	0.14	0.15	0.13	0.12	0.15	0.01	0.14	0.12	0.03	0.16	0.16	0.05	0.01	0.01	0.01	0.13	0.13	0.02	0.1	0.08	0.3	0.31	0.48	0.6	1	0.69	0.11	0.05	0.8	0.78	0.02	0.01	0.04	0.02	0.04	0.02
Bear Ave MCap	0.05	0.02	0.08	0.12	0.12	0.09	0.1	0.08	0.17	0.1	0.08	0.17	0.08	0.08	0.16	0.03	0.02	0.01	0.09	0.09	0.17	0.1	0.09	0.23	0.28	0.53	0.66	0.69	1	0.07	0.05	0.71	0.82	0.01	0.01	0.05	0.02	0.05	0.02
Bull Ave Spread	0.8	0.02	0.09	0.22	0.2	0.21	0.23	0.2	0.31	0.21	0.22	0.3	0.17	0.17	0.25	0.12	0.12	0.12	0.21	0.21	0.3	0.02	0.05	0.02	0.01	0.08	0.03	0.11	0.07	1	0.01	0.07	0.08	0	0	0.76	0.03	0.8	0.02
Bear Ave Spread	0.01	0.37	0.03	0.07	0.07	0.08	0.04	0.06	0.04	0.04	0.06	0.04	0.03	0.03	0.02	0.07	0.07	0.08	0.05	0.05	0.04	0.02	0	0.04	0.03	0.04	0.03	0.05	0.05	0.01	1	0.03	0.04	0	0.01	0.01	0.37	0.01	0.37
Bull Ave Turnover	0.02	0.01	0.07	0.18	0.19	0.13	0.21	0.25	0.04	0.23	0.23	0.06	0.24	0.24	0.06	0.01	0.01	0	0.23	0.23	0.06	0.08	0.04	0.24	0.26	0.72	0.38	0.8	0.71	0.07	0.03	1	0.94	0.01	0.01	0.02	0.01	0.02	0.01
Bear Ave Turnover	0.03	0.01	0.08	0.16	0.17	0.12	0.16	0.2	0.01	0.18	0.19	0.03	0.19	0.19	0.02	0.01	0.01	0.01	0.18	0.18	0.02	0.09	0.05	0.26	0.28	0.7	0.5	0.78	0.82	0.08	0.04	0.94	1	0.01	0.01	0.04	0.01	0.03	0.01
Bull Ave PTNAV	0.01	0	0.02	0.01	0.02	0.03	0.05	0.05	0.02	0.05	0.05	0.04	0.01	0.01	0.02	0.11	0.11	0.01	0.06	0.05	0.03	0.06	0.07	0.03	0.04	0	0.01	0.02	0.01	0	0	0.01	0.01	1	0.8	0.01	0	0.01	0
Bear Ave PTNAV	0.01	0	0.01	0	0	0.03	0.02	0.01	0.01	0.01	0.02	0.02	0	0	0	0.07	0.08	0.05	0.03	0.01	0.01	0.3	0.33	0.15	0.09	0	0.01	0.01	0.01	0	0.01	0.01	0.01	0.8	1	0	0.01	0.01	0
Bull NAV	0.98	0.04	0.04	0.11	0.1	0.1	0.25	0.24	0.27	0.24	0.26	0.26	0.2	0.2	0.22	0.07	0.07	0.07	0.24	0.25	0.26	0.01	0.05	0.06	0.07	0.05	0.02	0.04	0.05	0.76	0.01	0.02	0.04	0.01	0	1	0.04	0.98	0.04
Bear NAV	0.07	0.99	0.04	0.01	0	0.02	0.18	0.19	0.11	0.18	0.2	0.1	0.15	0.15	0.08	0.06	0.07	0.07	0.19	0.19	0.12	0.02	0.07	0.07	0.07	0.19	0.04	0.02	0.02	0.03	0.37	0.01	0.01	0	0.01	0.04	1	0.07	0.99
Bull Ave NAV	1	0.06	0.05	0.18	0.16	0.16	0.3	0.29	0.32	0.29	0.31	0.31	0.25	0.25	0.26	0.09	0.09	0.08	0.3	0.3	0.32	0.01	0.06	0.06	0.07	0.05	0.03	0.04	0.05	0.8	0.01	0.02	0.03	0.01	0.01	0.98	0.07	1	0.06
Bear Ave NAV	0.06	1	0.04	0.04	0.05	0.02	0.17	0.18	0.1	0.18	0.19	0.09	0.15	0.15	0.07	0.08	0.08	0.08	0.18	0.18	0.1	0.02	0.06	0.07	0.07	0.2	0.04	0.02	0.02	0.02	0.37	0.01	0.01	0	0	0.04	0.99	0.06	1

Table 10. Feature selection methods and their accuracy scores

	Quarter			Year		
Feature selection model	Selected features	Logistic scores	Random forest score	Selected features	Logistic scores	Random forest score
Recursive Feature Elimination	Bear net asset value Return beta for bear Bear trading volume Index volatility Bear turnover ratio	60.61%	66.73%	Bull net asset value Bear net asset value Bull trading volume Return beta for bear Bull bid/ask spread	87.61%	88.57%
Exhaustive Feature Selector	Bull net asset value Ave. Index price Index volatility Bear turnover ratio Bear price/turnover	61.92%	67.17%	Bull net asset value Index volatility Return beta for bear Bear turnover ratio	87.62%	88.57%
LASSO regression	Bull net asset value Ave. Index price Index volatility Return beta for bear	61.49%	71.99%	Bull net asset value Index price Return beta for bear	88.57%	92.38%
Ridge regression	Bull net asset value Index return Index volatility Return beta for bear	63.45%	73.74%	Bull net asset value Return beta for bear	87.62%	83.8%
Elastic net	Bull net asset value Ave. Index price Index return Index volatility Return beta for bear	63.01%	73.08%	Bull net asset value Return beta for bear	87.62%	83.8%

Table 11. Coefficient estimation for the selected predictors

Variables	Quarter								
	Individual coef.			Bull VS None		Bear VS None		Bear VS Bull	
	None	Bull	Bear	Coef.	Test	Coef.	Test	Coef.	Test
Intercept	-0.83	0.012	0.823	1.334	0.004	2.15	0.006***	0.815	8.814***
Bull net asset value	-0.08	0.312	-0.22	0.403	2.583***	-0.14	-0.98	-0.545	-4.59***
Index volatility	0.372	-0.44	0.072	-0.83	-5.41***	-0.3	-2.13**	0.523	4.56***
Return beta for bear	0.66	-0.23	-0.42	-0.9	-5.89***	-1.09	-7.66***	-0.185	-1.712*
Index return	-0.21	0.327	-0.11	0.548	4.552***	0.102	1.036	-0.445	-4.323***
Equity	0.308	-0.02	-0.29	-9.23	-0.001	-9.52	-0.002	-0.294	-1.184
Foreign equity	0.222	0.12	-0.34	-6.38	-0.001	-6.57	-0.002	-0.484	-2.353**
Commodity	0.304	0.237	-0.54	-7.38	-0.001	-8.18	-0.002	-0.802	-3.611***
Currency	0.528	-0.02	-0.5	-4.07	-0.002	-4.59	-0.002	-0.494	-3.59***
quarter 1	-0.22	0.04	0.185	0.272	1.935*	0.419	3.118***	0.147	1.454
quarter 2	-0.01	-0.05	0.065	-0.03	-0.25	0.085	0.68	0.118	1.141
quarter 3	-0.09	-0.02	0.115	0.075	0.568	0.214	1.69*	0.139	1.378

Table 11 Cont. Coefficient estimation for the selected predictors

Variables	Year								
	Individual coef.			Bull VS None		Bear VS None		Bear VS Bull	
	None	Bull	Bear	Coef.	Test	Coef.	Test	Coef.	Test
Intercept	-0.967	-0.31	1.279	0.541	0.000	2.692	0.001	2.89	0.001
Bull net asset value	-0.039	0.737	-0.7	0.881	1.728*	-0.74	-1.571	-1.621	-3.832***
Return beta for bear	0.596	-0.09	-0.5	-0.87	-1.921*	-1.255	-3.104***	-0.385	-1.164
Index price	0.548	-0.23	-0.32	-0.856	-1.972**	-0.938	-2.57***	-0.082	-0.228
Equity	-0.333	0.261	0.073	3.567	0.00	-10.378	-0.000	-12.079	-0.000
Foreign equity	0.008	0.233	-0.24	2.568	0.00	-7.931	-0.000	-9.154	-0.000
Commodity	0.323	0.16	-0.48	2.315	0.00	-9.428	-0.000	-10.247	-0.000
Currency	0.475	0.502	-0.98	1.248	0.00	-6.86	-0.000	-7.225	-0.000

The fourth quarter and real estate dummy are eliminated from the models and considered as the benchmark for the rest of corresponding dummy variables. The estimated regression line also includes dummy variables for each year of study (2012-2020) and year 2012 is considered as the benchmark year. The coefficients of years are not mentioned for simplicity of the table, but they are available by request.

*** Significant at 99%
** Significant at 95%
* Significant at 90%

References

Adhikari, R., Panta, H., & Hussan, M. K. (2020). Performance of ProShares Triple-Leveraged Equity ETFs. *Capital Markets Review*, 28(2), 1-18. https://www.mfa.com.my/wp-content/uploads/2020/09/v28_i2_a1_pg1-18.pdf

Aggarwal, R., & Schofield, L. (2014). The growth of global ETFs and regulatory challenges. In Advances in financial economics. *Emerald Group Publishing Limited*, 16, 77-102. https://www.emerald.com/insight/content/doi/10.1108/S1569-3732(2013)0000016003

Avellaneda, M., & Zhang, S. (2010). Path-dependence of leveraged ETF returns. *SIAM Journal on Financial Mathematics*, 1(1), 586-603.https://nyuscholars.nyu.edu/en/publications/path-dependence-of-leveraged-etf-returns

Bahadar, S., Mahmood, H., & Zaman, R. (2019). The herds of bulls and bears in leveraged ETF market. *Journal of Behavioral Finance*, 20(4), 408-423. https://www.researchgate.net/publication/340341067_The_Herds_of_Bulls_and_Bears_in_Leveraged_ETF_market

Baker, H. K., Filbeck, G., & Kiymaz, H. (Eds.). (2015). *Mutual funds and exchange-traded funds: building blocks to wealth*. Oxford University Press. https://www.amazon.com/Mutual-Funds-Exchange-Traded-Financial-Investments/dp/0190207434

Bashir, T., Khan, K. I., & Urooge, S. (2020). Assumptions of Making a Good Deal with Bad Person: Empirical Evidence on Strong Form Market Efficiency. *Global Social Sciences Review*, 1, 154-162. https://www.researchgate.net/publication/341050655_Assumptions_of_Making_a_Good_Deal_with_Bad_Person_Empirical_Evidence_on_Strong_Form_Market_Efficiency

Botha, I., & Marx, C. (2017). The relative efficacy of momentum strategies in the South African equity market. *Journal of Economic and Financial Sciences*, 10(3), 522-540. https://journals.co.za/doi/abs/10.10520/EJC-abf8bbd61

Brown, D. C., Davies, S. W., & Ringgenberg, M. C. (2021). ETF arbitrage, non-fundamental demand, and return predictability. *Review of Finance*, 25(4), 937-972. https://academic.oup.com/rof/article/25/4/937/5919085?login=true

Charupat, N., & Miu, P. (2013). The pricing efficiency of leveraged exchange-traded funds: evidence from the US markets. *Journal of Financial Research*, 36(2), 253-278. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6803.2013.12010.x

Chovancová, B., Dorocáková, M., & Linnertová, D. (2019). Two investment options for bearish etf investors: Inverse etf and shorting etf. *International Journal of Financial Studies*, 7(2), 31. https://www.mdpi.com/2227-7072/7/2/31

Clifford, C. P., Fulkerson, J. A., & Jordan, B. D. (2014). What drives ETF flows? *Financial Review*, 49(3), 619-642. https://onlinelibrary.wiley.com/doi/pdf/10.1111/fire.12049?casa_token=2eZzaomf0oY

AAAAA:Nd72mHE-nhX57LyXXUksdD-jO7f6pOp7_zfU6AI6oGcRsSe7Lg0a_a2SZ3d2qXL8slm0HItXtrLB1_di

Dangeti, P. (2017). *Statistics for machine learning*. Packt Publishing Ltd. https://www.oreilly.com/library/view/statistics-for-machine/9781788295758/

Day, M. Y., & Lin, J. T. (2019). Artificial intelligence for ETF market prediction and portfolio optimization. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 1026-1033. https://dl.acm.org/doi/abs/10.1145/3341161.3344822?casa_token=Sy-ObQGoT60AAAAA:qqALTY8t087Ap0zs9eLoDEd30O93p4dFtr3X4vPw7DL9TV8pzgdVTk-QUcrhFkZo98bIpWhqK3-LvQ

Deev, O., & Linnertová, D. (2014). The determinants of ETFs short-selling activity. *Procedia-Social and Behavioral Sciences*, 109, 669-673. https://www.sciencedirect.com/science/article/pii/S1877042813051653

Deville, L. (2008). *Exchange-traded funds: History, trading, and research. In Handbook of financial engineering*. Springer, Boston, MA. https://link.springer.com/chapter/10.1007/978-0-387-76682-9_4

DiLellio, J. A., Hessea, R., & Stanley, D. J. (2014). Portfolio performance with inverse and leveraged ETFs. *Financial Services Review*, 23(2). 123-149 http://web.b.ebscohost.com/ehost/detail/detail?vid=0&sid=96e6c870-7b55-493a-b154-e975e60a0068%40pdc-v-sessmgr03&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=97329883&db=bth

Dionne, G. (2019). *Corporate risk management: Theories and applications*. John Wiley & Sons.https://www.amazon.com/Corporate-Risk-Management-Theories-Applications/dp/1119583128

Duffie, D., & Pan, J. (2001). Analytical value-at-risk with jumps and credit risk. *Finance and Stochastics*, *5*(2), 155-180. https://link.springer.com/article/10.1007/PL00013531

Fabozzi, F. J. (2008). *Handbook of Finance, Financial Markets and Instruments (Vol. 1)*. John Wiley & Sons. https://www.amazon.com/Handbook-Finance-Financial-Markets-Instruments/dp/0470078146

Garg, A. K., & Varshney, P. (2015). Momentum effect in Indian stock market: A sectoral study. *Global Business Review*, 16(3), 494-510. https://journals.sagepub.com/doi/10.1177/0972150915569940

Gastineau, G. L. (2010). *The exchange-traded funds manual* (Vol. 186). John Wiley & Sons. https://www.amazon.com/Exchange-Traded-Funds-Manual-Gary-Gastineau/dp/0470482338

Guo, K., & Leung, T. (2015). Understanding the tracking errors of commodity leveraged ETFs. In Commodities, energy and environmental finance. *Springer*, 74, 39-63. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2389411

Hessel, C., Nam, J., Wang, J., Xing, C., & Zhang, G. (2018). Shorting leveraged ETF pairs. *The Journal of Trading*, 13(2), 69-79. https://jot.pm-research.com/content/13/2/69

Hill, J. M., Nadig, D., & Hougan, M. (2015). *A comprehensive guide to exchange-traded funds (ETFs)*. CFA Institute Research Foundation. https://www.amazon.com/Comprehensive-Guide-Exchange-Traded-Funds-ETFs/dp/1934667854

Jethwani, D., & Ramchandani, K. (2017). Semi Strong Form of Efficiency of Stock Market: A Review of Literature. *Int. Multidiscip. Res. 4*, 2349-7637. https://rhimrj.co.in/wp-content/uploads/2020/09/RHIMRJ20170412007.pdf

Ilan Guedj, G. L., & McCann, C. (2010). Leveraged ETFs, Holding Periods and Investment Shortfalls. *The Journal of Index Investing* , 1(3), 45-57 https://www.researchgate.net/publication/270124896_Leveraged_and_Inverse_ETFs_Holding_Periods_and_Investment_Shortfalls

Jiang, X., & Peterburgsky, S. (2017). Investment performance of shorted leveraged ETF pairs. *Applied Economics*, 49(44), 4410-4427. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2211387

Kochenderfer, M. J., & Wheeler, T. A. (2019). *Algorithms for optimization*. Mit Press. https://algorithmsbook.com/optimization/files/optimization.pdf

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. https://www.springer.com/gp/book/9781461468486

Leković, M. (2018). Evidence for and against the validity of efficient market hypothesis. *Economic Themes*, 56(3), 369-387. https://sciendo.com/article/10.2478/ethemes-2018-0022

Leung, T., & Ward, B. (2015). The golden target: Analyzing the tracking performance of leveraged gold ETFs. *Studies in Economics and Finance*, 32(3), 278-297. https://www.emerald.com/insight/content/doi/10.1108/SEF-01-2015-0009/full/html

Liew, J. K. S., & Mayster, B. (2017). Forecasting etfs with machine learning algorithms. *The Journal of Alternative Investments*, 20(3), 58-78. https://jai.pm-research.com/content/20/3/58

Lu, L., Wang, J., & Zhang, G. (2009). Long term performance of leveraged ETFs. *Available at SSRN 1344133*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1344133

Mackey, A. M., & Bacon, F. W. (2017). Signaling with stock issues and repurchases: A test of semi-strong form market efficiency. *Journal of Business and Behavioral Sciences*, 29(1), 34-44. https://www.researchgate.net/publication/312977642_Signaling_with_Stock_Issues_and_Repurchases_A_Test_of_Semi-Strong_Form_Market_Efficiency

Molin, S. (2019). *Hands-On Data Analysis with Pandas: Efficiently perform data collection, wrangling, analysis, and visualization using Python*. Packt Publishing Ltd. https://www.amazon.com/Hands-Data-Analysis-Pandas-visualization/dp/1789615321

Omar, M. F., Muhamat, A. A., & Yahya, N. C. (2021). Global Exchange-Traded Funds (ETFs): Emerging Markets Potentials. *Empirical Economics Letters,* 20(2), 37-48. https://www.researchgate.net/profile/Amirul-Afif-Muhamat/publication/352039759_Global_Exchange-

Traded_Funds_ETFs_Emerging_Markets_Potentials/links/60b6cf8792851cde884a6fe8/Global-Exchange-Traded-Funds-ETFs-Emerging-Markets-Potentials.pdf

Pampel, F. C. (2020). *Logistic regression: A primer*. Sage publications. https://www.amazon.ca/Logistic-Regression-Primer-Fred-Pampel/dp/0761920102

Peterburgsky, S. (2018). Leveraged ETF pairs: An empirical evaluation of portfolio performance. *Available at SSRN 3264865*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3264865

Prakash, J. (2012). A study of weak, semi-strong and strong forms of market efficiency: review of literature. *Journal of global research & analysis*, 1, 98-105. https://www.geetalawcollege.in/wp-content/uploads/2017/05/December-2012-Issue.pdf#page=101

Rompotis, G. G. (2014). On leveraged and inverse leveraged exchange-traded funds. *Aestimatio*, (9), 150-181. https://www.ieb.es/wp-content/uploads/2014/07/1.pdf

Saini, N. (2019). Regime-switching advantage in statistical arbitrage strategies conditioned on time series momentum and volatility in leveraged exchange-traded funds: theory and evidence. *Doctoral dissertation, University of Wisconsin--Whitewater*. https://minds.wisconsin.edu/handle/1793/79421

Sharpe, W. F. (1994). The sharpe ratio. *Journal of portfolio management*, *21*(1), 49-58. https://jpm.pm-research.com/content/21/1/49

Shum, P., Hejazi, W., Haryanto, E., & Rodier, A. (2016). Intraday share price volatility and leveraged ETF rebalancing. *Review of Finance*, 20(6), 2379-2409. https://academic.oup.com/rof/article-abstract/20/6/2379/2418138

Shum, P., & Kang, J. (2013). Leveraged and inverse ETF performance during the financial crisis. *Managerial Finance,* 39(5), 476-508. https://www.emerald.com/insight/content/doi/10.1108/03074351311313825/full/html

Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *the Journal of Investing*, 3(3), 59-64. https://joi.pm-research.com/content/3/3/59

Strickland, J. (2017). *Logistic Regression Inside and Out*. Lulu Press. https://www.amazon.com/Logistic-Regression-Inside-Jeffrey-Strickland/dp/1365819159

Trainor Jr, W. J., & Carroll, M. G. (2013). Forecasting holding periods for leveraged ETFs using decay thresholds: theory and applications. *Journal of Financial Studies & Research*, 2013 (2013), 1-12. https://www.researchgate.net/publication/269641212_Forecasting_Holding_Periods_for_Leveraged_ETFs_Using_Decay_Thresholds_Theory_and_Applications

Trainor Jr, W. J., & Baryla Jr, E. A. (2008). Leveraged ETFs: A risky double that doesn't multiply by two. *Journal of Financial Planning*, 21(5), 48-55.

https://www.researchgate.net/publication/284699463_Leveraged_ETFs_A_risky_double_that_doesn%27t_multiply_by_two

Tsalikis, G. (2020). Exchange-traded funds in Europe and the US: performance and tracking errors. *Doctoral dissertation, University of Macedonia*. https://dspace.lib.uom.gr/handle/2159/24064

Tsalikis, G., & Papadopoulos, S. (2018). Assessing the performance of American and European leveraged exchange-traded funds. *Investment Management & Financial*, 15(2), 165-182. https://www.researchgate.net/publication/325349089_Assessing_the_performance_of_American_and_European_Leveraged_Exchange_Traded_Funds

Wided, K. (2018). On the Properties of Leveraged ETFs. *Finance Bulletin*, 1(2), 50-62. https://financebulletin.org/article/view/2314

Yadav, Y. (2016). Insider trading and market structure. *UCLA L. Rev.*, *63*, 968-1033. https://www.uclalawreview.org/insider-trading-market-structure/

Yagi, I., & Mizuta, T. (2016). Analysis of the impact of leveraged etf rebalancing trades on the underlying asset market using artificial market simulation. In *12th Artificial Economics Conference*, 1-11. https://www.researchgate.net/publication/307168438_Analysis_of_the_Impact_of_Leveraged_ETF_Rebalancing_Trades_on_the_Underlying_Asset_Market_Using_Artificial_Market_Simulation

Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, 87(4), 954-959. https://www.jstor.org/stable/2673623