Mathematics of Leveraged ETFs, Beta Decay, and Long-Term Investment Viability

By Emil Perdue

April 2024

Undergraduate Honors Thesis
Computational and Applied Mathematics Major

University of Chicago
Thesis Advisor: Prof. Dylan Hall

Contents

1 Introduction **3**

2 About ETFs **4**
2.1 ETF Creation and Redemption 4
2.2 LETF Introduction 5

3 Literature Review **8**
3.1 Structural Slippage of Leveraged ETFs 9
3.2 Leveraged ETFs, Holding Periods and Investment Shortfalls 13
3.3 Long Term Performance of Leveraged ETFs 15
3.4 Long-Term Investing in Triple Leveraged Exchange Traded Funds 18
3.5 Portfolio Insurance Using Leveraged ETFs 20

4 Leveraged ETF Tracking Abilities **22**
4.1 Data Collection 22
4.2 Methodology 22
4.3 Statistical Testing and Results 23
4.4 Decomposing Tracking Error 26
4.5 Different Time-Frames 31

5 Leveraged ETF Investment Viability **33**
5.1 Buy and Hold Long-Term Performance 33
5.2 Portfolio Optimization Review 37
5.3 Markowitz Implementation 39

6 Conclusion **41**
6.1 Discussion of Short-Term Trading 41
6.2 Limitations and Recommendations 42

7 Appendix **43**

References **44**

1 Introduction

Leveraged Exchange-Traded Funds (LETFs), introduced in 2006, are among the more complex financial instruments in today's financial markets. These securities utilize financial derivatives to amplify the returns of their underlying indices. Despite a significant presence, with over 200 LETFs trading and nearly $40 billion in assets under management ("Leveraged ETFs", 2024), LETFs are perceived as less accessible and more complex than other derivatives, largely due to the uncertainties and volatility affecting their returns. Furthermore, there exists a gap in accessible data for evaluating LETFs' collective performance against their respective benchmarks, despite notable sources cautioning against their long-term investment without strong empirical evidence. Similarly, much of the academic literature on LETFs is relatively outdated and doesn't reflect the recent significant market fluctuations over the last 3 to 5 years. This thesis' objective is to present a concise, yet comprehensive report, accompanied by a rigorous mathematical treatment, to shed light on how LETFs are performing against their benchmarks to ultimately support or refute the claims made about their investment viability. This thesis aims to provide a comprehensive overview of LETFs, examine beta decay from a mathematical perspective, and finally conduct novel research into the longevity of LETFs today and begin to understand which LETFs track their underlying index better than others and why.

First, we will introduce the financial background necessary for understanding both ETFs and LETFs, followed by an examination of the ETF Creation and Redemption process. We will then conduct a comprehensive literature review, presenting and synthesizing the mathematical underpinnings of LETF beta decay and the challenges these products face in accurately tracking their underlying indices, due to rebalancing. Our study will then include derivations detailing the mathematics of rebalancing and how these products respond to stochastic price movements. After reviewing the relevant literature, we will conduct novel research into LETFs' potential as long-term investments. Specifically, Python will be utilized to analyze the performance of all 200+ LETFs listed on the U.S. stock exchange against their benchmarks, examining their tracking efficiency over time. Daily price action data will be sourced from Yahoo Finance and Bloomberg. Through this analysis, we look to understand why certain LETFs track their indices more closely than others by examining decay levels across entire price histories, specific lookback periods, and under varying market conditions (e.g., bear, bull, high volatility). This process will enable us to juxtapose modern-day findings with those in the examined academic papers, allowing us to assess the results of our literature review and establish realistic performance expectations for LETFs as investment options. Finally, these results will be utilized to explore potential portfolio allocations, concluding with a discussion on the trade-offs associated with these products and the considerations investors must account for.

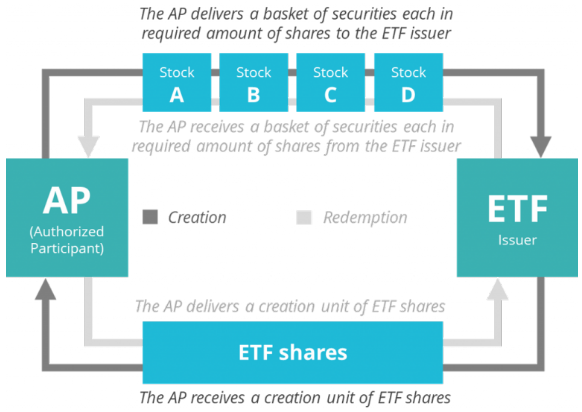
2 About ETFs

Exchange-traded funds (ETFs) are financial instruments that can be publicly traded and are designed to track an underlying index, commodity, or sector. They achieve this goal by owning a diverse array of financial assets including stocks, bonds, debts, currencies, and futures. Examples include VOO and SPY which replicate the S&P 500 Index, DIA which tracks the Dow Jones Industrial Average, and QQQ which matches the performance of the Nasdaq 100. ETFs provide the convenience of being able to follow an index without the need for rebalancing, leveraging, or borrowing. Additionally, unlike mutual funds, ETFs are priced continuously throughout market hours, enabling intraday trading. This flexibility is enhanced by ETF providers' mechanism of share creation and redemption whose management ensures that their fund's trading value remains aligned with their Net Asset Value (NAV).

2.1 ETF Creation and Redemption

The ETF creation process begins when a sponsor files a petition for a new ETF with the U.S. Securities and Exchange Commission (SEC). Upon approval, this sponsor works with an authorized participant, typically institutional investors or market makers, who can create or redeem shares. The authorized participant then acquires the necessary underlying stock shares, places them in a trust, and uses them to formulate ETF creation units, generally consisting of approximately 50,000 shares (Smith, 2024). The trust, in return, issues individual ETF shares to the authorized participant, which represent legal claims to the shares held. Following the receipt of the ETF shares, the authorized participant is then able to sell these shares to the public on the open market. The securities used to assemble the creation units remain in the trust account and are not impacted by market transactions that involve the ETF shares.

The creation process now facilitates the issuance of additional ETF shares to be put into circulation in return for the underlying asset, whereas redemption removes ETF shares from circulation for the underlying asset. For instance, when an investor wants to sell their ETF holding, they can either sell ETF shares directly on the open market or accumulate enough ETF shares to form a creation unit, which can be exchanged or redeemed for their underlying securities. This dual process of creation and redemption ensures liquidity and aligns the ETF's market price closely with its Net Asset Value (NAV), which is the sum of all underlying assets.



(David Abner, 2016)

Authorized participants are critical in providing liquidity to ETFs, enabling sponsors to transact in large blocks. For example, consider an ETF comprising assets A and B, each valued at \$1, setting the Net Asset Value (NAV) at \$2. If the ETF's hypothetical market price was \$2.02, the authorized participant would be incentivized to create more ETF shares, profiting by exchanging assets for ETF shares at a gain of 2 cents per share. Conversely, should the NAV exceed the ETF's market price, the authorized participant would seek to redeem their ETF to acquire the underlying assets.

This mechanism allows authorized participants to both inject liquidity into the ETF market and help align the prices of ETFs with their underlying assets, contributing to overall market efficiency. While individual investors are not able to create or redeem ETF shares themselves, they benefit indirectly from the liquidity and price alignment facilitated by authorized participants. This accessibility to ETFs enables investors to diversify their portfolios across a broad range of underlying assets, which would be difficult to invest in proportionally otherwise.

2.2 LETF Introduction

LETFs are specialized securities that utilize financial derivatives, including structured notes, index futures, equity swaps, index options, and debt to amplify the returns of an underlying index, typically by a factor of 2:1 or 3:1. For instance, Daily S&P 500 Bear 3X (SPXS) mimics the 3x inverse daily return of the S&P 500, and Proshares Ultrapro Short QQQ (SQQQ) aims to deliver the 3x inverse daily return of the Nasdaq 100. Typically, these 3x LETFs utilize a combination of index futures and call/put options to replicate a specified degree of leverage. While these products can result in significant gains, they are also prone to amplified losses and reduced performance over time due to rebalancing and tracking errors, leading them to not be considered sound for long-term holding, but rather be advertised as short-term investment vehicles.

One of LETF's primary downfalls as a product is *daily rebalancing*. ETFs need to maintain a constant leverage ratio (typically 2x or 3x), which brings complexity from the fact that interest, management fees, and fluctuations in the underlying price subsequently alter the value of the LETFs' assets, necessitating continuous adjustments by the fund manager in the total amount of index exposure. For instance, if a fund with \$100M in assets has \$200M in index exposure, a rise of 1% of the index would yield \$2M in profits and lead to the fund now having \$102M of assets, as well as $204M in index exposure. This attempt to maintain a constant leverage ratio is the essence of daily rebalancing and is what allows a fund to provide levered exposure, regardless of changes in the underlying index. However, this same process generates beta decay, a phenomenon that is illustrated in the following rebalancing example: if an index falls by 1% daily over four days, then gains 4.1% on the fifth day, there is effectively a 0% change. However, a 2x LETF would not return to its starting point, but rather be down .2% overall. This comes as a result of a proportionally more minuscule asset base inside the leveraged fund, therefore requiring a larger return of 8.42% to break even. Essentially, for four straight 2% drops, we need a multiplier of $1/(.98^4)$, or a return of 8.42% to return to 0 (Yates). Because a majority of LETFs reset to their underlying benchmark index daily to maintain a fixed leverage ratio, this effect of rebalancing can quickly compound leading to discrepancies between the underlying fund and its index exposure.

Beta decay in the markets is exemplified by the Direxion Daily Financial Bull 3X Shares ETF (FAS) which is designed as a 3x leveraged reflection of the Russell 1000. Specifically, on June 30th, 2021, while FAS's 5-year return was 34.68%, Russell 1000's return was 18.75%. FAS's prospectus attributes this inconsistency to daily rebalancing and the compounding of each day's returns over time (Investopedia, 2022). At the same time, however, there are significant exceptions, such as TQQQ, a 3x leveraged Nasdaq ETF, which has been successful in tracking the large up-moves of its underlying index (Glenn, 2011). Additionally, one of the most critical risks is that LETFs, due to their fixed leverage ratios, can face total collapse if the underlying index they track declines by a certain percentage in a single day. This threshold of decline is inversely related to their leverage ratio where, for example, an LETF with a 3x leverage ratio risks total collapse if its underlying index falls by more than 33% in one day.

We have compiled a list of all currently existing LETFs, as well as their respective underlying ETFs, leverage ratios, category (focusing on debt, currency, equity), market cap, and average volume. Below are the ten most popular LETFs, examples of the pairings we will further explore in the research section.

Table 1

Summary of top 10 LETFs of 186 studied sorted by average volume, as of March 31st, 2024

LETF Symbol	ETF Symbol	Description	Leverage Ratio	Category 1	Category 2	Category 3	Market Cap	Average Volume
TQQQ	NDX	ProShares UltraPro QQQ (3x) ETF	3.00	US Equity	Broad market	Large-Cap	12,774,888,000	164,492,850
SOXL	ICESEMIT	Direxion Semiconductor Bull (3x) ETF	3.00	US Equity	Single industry	Semiconductors	5,681,627,549	85,193,280
SQQQ	NDX	ProShares UltraPro Short QQQ (3x) ETF	-3.00	US Equity	Broad Market	Large-Cap	5,147,058,480	101,902,900
QLD	NDX	ProShares Ultra QQQ ETF	2.00	US Equity	Broad Market	Large-Cap	3,592,880,000	6,670,822
SSO	ŜPX	ProShares Ultra S&P500 ETF	2.00	US Equity	Broad Market	Large-Cap	3,163,116,000	7,466,952
SPXL	SPXT	Direxion Large Cap Bull 3x Shares ETF	3.00	US Equity	Broad Market	Large-Cap	2,435,373,565	18,292,720
SH	ŜPX	ProShares Short S&P 500 ETF	-1.00	US Equity	Broad Market	Large-Cap	2,371,719,640	36,657,871
UPRO	ŜPX	ProShares UltraPro S&P 500 (3x) ETF	3.00	US Equity	Broad Market	Large-Cap	2,219,210,000	14,067,479
TECL	IXTTR	Direxion Technology Sector (3x) ETF	3.00	US Equity	Single Sector	Information Tech	1,684,380,000	3,695,465
FAS	IXMTR	Direxion Financial Bull 3x Shares ETF	3.00	US Equity	Single Sector	Financials	1,486,019,056	2,794,789

3 Literature Review

We now aim to conduct a comprehensive literature review and provide an expository lens for the mathematical framework underlying LETFs. Specifically, we will present the formulas used by LETFs to track their corresponding ETFs and how these formulas can fail from beta decay, causing the performance of LETFs to diverge from their intended leverage ratio. Additionally, we will delve into the existing literature on the long-term investment performance of LETFs, consider the effects of beta decay, and whether any portfolio constructions exist that can mitigate risk while remaining appealing to investors.

We also want to emphasize that much of the literature on LETFs is fairly outdated, especially in light of recent market regime shifts such as those following the COVID-19 pandemic and the retail-driven volatility in 2021. Our research aims to address these gaps, analyzing the relevance and potential limitations of previous findings in the context of current market conditions. Given the scarcity of literature on LETFs, our objective is not only to reevaluate existing knowledge but to also apply new research methodologies to corroborate or challenge existing theories about LETF performance.

We have selected five prominent papers for our review, each of which dissects LETF return tracking and how they perform in long-term investment strategies. Notably, the third and fourth papers provide contrasting perspectives on long-term LETF performance and their suitability as investment vehicles for holding periods exceeding one month.

- Structural Slippage of Leveraged ETFs
- Leveraged ETFs, Holding Periods, and Investment Shortfalls
- Long-Term Performance of Leveraged ETFs
- Long-Term Investing in Triple Leveraged Exchange Traded Funds
- Portfolio Insurance Using Leveraged ETFs

3.1 Structural Slippage of Leveraged ETFs

To provide a deeper understanding of beta decay and the strategies employed by managers to navigate market fluctuations, we refer to "Structural Slippage of Leveraged ETFs" by Doris Dobi and Marco Avellaneda. This study examines the daily rebalancing techniques used by LETF managers, who utilize total return swaps to systematically adjust their index exposure near market close each day to maintain a desired leverage ratio. The authors highlight the inherent challenge in this process, as managers are forced to buy high and sell low, culminating in suboptimal tracking performance over time. This paper outlines LETFs' foundational flaw in their rebalancing and why these products struggle to track their underlying index over periods of greater than one month.

The authors further discuss the implications of daily rebalancing, including the potential for front-running, by which sophisticated investors anticipate and predict these adjustments. This awareness, which is due in part to LETFs' public mandates, can result in managers getting less favorable prices for the assets they need to trade. The study concludes that the cumulative costs of rebalancing become statistically significant for holding periods of as little as one day, resulting in LETF performance that underperforms the advertised leverage ratio. Similarly, this misalignment becomes exacerbated when taking into account high borrowing costs of up to 600 basis points (where LETF managers must borrow funds from brokerages to invest in assets/derivatives, incurring interest payments), longer holding periods, and periods of elevated volatility (Avellaneda and Dobi, 2012).

The paper provides a detailed walkthrough of a typical rebalancing procedure. First, the authors denote an LETF's leverage ratio with its ETF to be β. The authors define discrete times Δt as one day and examine the price of the LETF at date t. Theoretically, the LETF's return should mirror the underlying index's return multiplied by the leverage ratio β.

Buy High Sell Low: Long Example

Consider an LETF with $\beta = 2$ which has an underlying index with a price of \$100 and an AUM (Assets Under Management) of \$1B (billion). Now, consider its index values and corresponding AUM over a series of days and subsequent market moves.

- **Day 0:**
  - The manager begins with \$2000M of exposure in TRS (Total Return Swaps).
- **Day 1:**
  - When the underlying index moves down by 10% to \$90, the AUM drops by Initial AUM * (1 - β * Change) = \$1000M * (1 - 2 * .1) = \$800M.
  - The manager now needs β * AUM = 2 * \$800M = \$1600M in exposure.
  - However, the real notional value of the TRS from Day 0 is now AUM * (1 - Change) = \$2000M * (1 - .1) = \$1800M pre-adjustment.
  - As a result, the manager must **sell** \$200M of TRS for the actual exposure to meet the exposure needed (\$1800M to go to \$1600M).
- **Day 2:**
  - When the underlying index goes up by 10% to \$99, the AUM grows by Initial AUM * (1 + β * Change) = \$800M * (1 + 2 * .1) = \$960M.

– The manager now needs β * AUM = 2 * $960M = $1920M in exposure.

– However, the real notional value of the TRS after the change is AUM * (1 + Change) = $1600M * (1 + .1) = $1760M.

– To get the necessary exposure of $1920M, the manager must **buy**, going long $160M of TRS (which is the difference in actual vs. expected exposure, or $1920M - $1760M).

Table 1: $\beta = 2$, Initial Index Value = $100, Initial AUM = $1B

Day	Index Value($)	AUM($M)[1]	Exposure Needed($M)	Exposure Before Adjustment($M)	TRS Adjustments($M)
0	100	1000	2000	- - -	- - -
1	90	800	1600	1800	-200
2	99	960	1920	1760	+160

(Avellaneda and Dobi, 2012)

Buy High Sell Low: Short Example

Next, in the case of an inverse LETF, we look at an LETF with a $\beta = -2$. Once again, the theoretical AUM is $1B and the underlying index has an initial price of $100.

- **Day 0:**
  – The manager begins with -$2000M of exposure in TRS (Total Return Swaps).
- **Day 1:**
  – When the underlying index moves down by 10% to $90, the AUM grows by Initial AUM * (1 - β * Change) = $1000M * (1 + 2 * .1) = $1200M.
  – The manager now needs β * AUM = -2 * $1200M = -$2400M in exposure.
  – However, the real notional value of the TRS from Day 0 is now AUM * (1 - Change) = -$2000M * (1 - .1) = -$1800M pre-adjustment.
  – As a result, the manager must **sell** $600M of TRS for the actual exposure to meet the exposure needed ($1800M to go to $2400M).
- **Day 2:**
  – When the underlying index goes up by 10% to $99, the AUM drops by Initial AUM * (1 + β * Change) = $1200M * (1 - 2 * .1) = $960M.
  – The manager now needs β * AUM = -2 * $960M = -$1920M in exposure.
  – However, the real notional value of the TRS after the change is AUM * (1 + Change) = -$2400M * (1 + .1) = -$2640M.
  – To get the necessary exposure of $1920M, the manager must **buy**, going long $720M of TRS (which is the difference in actual vs. expected exposure, or $2640M - $1920M).

Table 2: $\beta = -2$, Initial Index Value = \$100, Initial AUM = \$1B

Day	Index Value(\$)	AUM(\$M)[1]	Exposure Needed(\$M)	Exposure Before Adjustment(\$M)	TRS Adjustments(\$M)
0	100	1000	-2000	- - -	- - -
1	90	1200	-2400	-1800	-600
2	99	960	-1920	-2640	+720

(Avellaneda and Dobi, 2012)

As demonstrated in the previous two examples, fund managers of both long and short LETFs are required to continually rebalance in the same direction as that of the market, leading to purchases during market uptrends and sales during downtrends. The authors then present a mathematical derivation explaining how the maintenance of a constant leverage ratio necessitates rebalancing in the same direction as the market, irrespective of the LETF's direction. They begin by defining S_t as the price of the underlying index at time t and E_t as the LETF's AUM at time t. Finally, the time step Δt represents one day and r is the daily risk-free interest rate. Using these definitions, they calculate the profit and loss (P&L) of one share of the underlying index to be:

$$\Delta S_t - rS_t\Delta t$$

By defining n_t to be the number of shares that the underlying index holds to hedge at time t, the P&L for the total number of shares is:

$$n_t(\Delta S_t - rS_t\Delta t)$$

To hold a consistent leverage ratio β, a manager must ensure that the following equation holds at the start of every defined trading period, which relates AUM with the number of shares held at time t:

$$\beta E_t = n_t S_t \Rightarrow E_t = \frac{n_t S_t}{\beta}$$

When rebalancing, one can define the P&L of the fund in terms of the change in the value of the TRS position:

$$\Delta E_t = n_t(\Delta S_t - rS_t\Delta t) + E_t r\Delta t$$

Combining these results allows us to verify that a manager is forced to buy high and sell low to maintain a constant leverage ratio, regardless of the underlying direction of an LETF. The following equation depicts how both long and short LETF must rebalance in the same direction as that of the market:

$$\frac{\Delta n_t}{n_t} = (\beta - 1)\left(1 - \frac{(1 + r\Delta t)S_t}{S_{t+1}}\right)$$

Lastly, this paper introduces a model for slippage, positing that LETFs have negative expected returns due to the market's anticipation of the size and direction of hedging required by managers to maintain a predetermined leverage ratio. The authors define μt as a positive constant and ϵt as a mean-zero noise term. They then denote slippage at time t to be $(\mu t + \epsilon t)$ when constructing the following portfolio equation that accounts for daily rebalancing slippage, negatively impacting the expected change in AUM (Avellaneda and Dobi, 2012).

$$\Delta E_t = \underbrace{n_t(\Delta S_t - S_t r \Delta t) + E_t r \Delta t}_{\text{change in AUM without impact}} - \underbrace{E_t(\mu_t + \epsilon_t)}_{\text{additional change caused by daily rebalancing}}$$

As a result, we can take the actual return of the fund subtracted by the expected return of the fund as $-\mu_t - \epsilon_t$, which represents how the slippage term ultimately causes the difference between expected and realized returns:

$$\underbrace{\frac{\Delta E_t}{E_t}}_{\text{actual return of fund}} - \underbrace{\left(\beta \frac{\Delta S_t}{S_t} + r\Delta t(1-\beta)\right)}_{\text{expected return of fund}} = -\mu_t - \epsilon_t$$

Using these results, the authors argue that by shorting both the long and short LETFs, one would be able to capture a portion of the remaining slippage, defined above. Similarly, one could also identify slippage by shorting the long LETF and going long $|\beta|$ of the underlying index, or by shorting an inverted LETF and shorting $|\beta|$ of the underlying index. In our research, while we plan to utilize a similar notation for how β relates an LETF to its underlying ETF, we intend to utilize a more empirical approach, rather than mathematical, for identifying slippage.

3.2 Leveraged ETFs, Holding Periods and Investment Shortfalls

In their paper, authors Ilan Guedj, Guohua Li, and Craig McCann examine the performance of LETFs compared to using margin accounts to achieve similar leverage levels over various holding periods. They also seek to estimate the average holding period for LETFs and assess their performance across daily and monthly rebalancing.

The authors employ two models to estimate the number of LETF shares traded by investors. Firstly, the Proportional Trader Model (PTM), proposes that each outstanding share has an equal probability of being traded. According to this model, the proportion of shares traded each day is drawn from a pool of recently traded and not recently traded shares, based on the sizes of these two groups. For example, if a stock trades 200 shares in a day with 1,000 shares outstanding, PTM estimates that every investor has sold 20% of their shares and has kept the remaining 80% they held from the previous day. If 100 shares are traded the next day, PTM assumes that all investors have sold 10% of their shares.

The second model, the Multiple Trader Model (MTM), operates similarly to the PTM but applies its trading dynamic logic separately to shares outstanding and daily trading volume. The results from applying PTM to both shares outstanding and daily trading volume are added to one another to estimate the number of shares bought or sold on that day, or the daily trading volume. In this study, the authors utilize the MTM as a means of isolating retail trader flow from that of market makers and high-frequency trading (HFT) institutions, to accurately calculate investment shortfall for the average investor investing in these products as opposed to utilizing a margin account (Guedj, Li, and McCann, 2010).

The authors analyze LETF tickers DPK (-3X), TYO (-3X), RHO (-2X), SBB (-1X), and UVG (2X), employing the MTM in conjunction with volume data to calculate the average holding period and daily turnover ratio, which measures the number of buyers and sellers in a specific stock. They then calculate the investment shortfalls of these LETFs based on an estimated investment period, rebalancing on either a monthly or daily basis. The returns of an LETF in this study are defined as follows:

$$(1 + R_T^{\text{L-ETF}}) = (1 + R_T^{\text{index}})^x \cdot e^{\frac{(x-x^2)\sigma^2 T}{2}}$$

Here, x is the leverage ratio, σ is the volatility of the index, and T is the time for which the investment is held (Guedj, Li, and McCann, 2010).

Name	Leverage Ratio	Inception Date	Investment Shortfall of ETF Rebalanced Daily	Investment Shortfall of ETF if Rebalanced Monthly
DPK	-3	12/17/2008	$ 1,412,489	$ 78,526
TYO	-3	4/15/2009	$ 535,768	$ (311,017)
RHO	-2	6/12/2008	$ 207,726	$ (12,389)
SBB	-1	1/25/2007	$ 1,573,060	$ (863,744)
UVG	2	2/22/2007	$ 464,699	$ 968,306

(Guedj, Li, and McCann, 2010)

Through the application of the MTM, the authors find the investment shortfall from investing in LETFs, as opposed to investing in margin accounts to be statistically significant. They discovered that by investing in LETFs, investors lose 3% of their total investment in less than three weeks, resulting in an annualized cost of nearly 50%. When examining this same effect across different rebalancing periods, they find that daily rebalancing exacerbates the investment shortfall for most LETFs, compared to monthly rebalancing. For example, when shifting from daily to monthly rebalancing, DPK's shortfall disappears, while TYU, RHO, and SBB's shortfalls become negative.

The findings suggest that even with reduced rebalancing frequency, the potential costs associated with long-term investing in LETFs cannot be fully mitigated when compared to using margin accounts. In our research, we plan to extend this examination to a wider range of LETFs and cover a more recent timeframe, without the usage of purely theoretical models, to assess the long-term performance of these investment vehicles.

3.3 Long Term Performance of Leveraged ETFs

In "Long-Term Performance of Leveraged ETFs", authors Lei Lu, Jun Wang, and Ge Zhang examine the long-term performance of four groups of Ultra Long/Short LETFs from the ProShare family (DIA, DDM, DXD, SPY, SSO, SDS, QQQQ, QLD, QID, IWM, UWM, and TWM) and the relationship between their respective benchmarks over a variety of holding periods. Specifically, the authors utilize two-day, one-week, one-month, one-quarter, and one-year intervals to analyze how the relationship between various LETFs and their benchmarks deviates over increasing periods, particularly in the case of inverse LETFs (Lu, Wang, and Zhang, 2009).

Through testing the eleven symbols selected, the authors find that for holding periods of one month or shorter, one can safely assume that all LETFs provide the returns of their underlying benchmark multiplied by their advertised leverage ratio. However, for holding periods of one quarter, this relationship begins to break down, specifically in the case of inverse LETFs. The authors find that the one-quarter performance of inverse 2x LETFs deviates extensively from the expected benchmark performance multiplied by a β of -2. Similarly, for holding periods of one year, the relationship between the traditional long 2x LETF and the benchmark deteriorates. They attribute this poor long-term tracking performance of LETFs to quadratic and auto variations during the examined periods, where auto-variation was the more dominant factor. In general, the authors conclude that LETFs are not reliable substitutes for long or short positions in their underlying indices over extended periods.

The authors employ a regression testing framework similar to the one we plan to employ in our research. They define ${r_t}^B$ as the daily return of the underlying index at time t, with ${r_t}^I$ and ${r_t}^D$ representing the daily returns of the inverse and long 2x LETF, respectively. They then set R^B_{tn} as the cumulative n-day return of the underlying index, and R^I_{tn} and R^D_{tn} as the cumulative returns of the 2x inverse and long 2x LETFs starting from date t. The author then proposes the following equation:

$$R^B_{tn} = \prod_{i=0}^{n-1}(1 + r^B_{t+i}) - 1,$$

$$R^D_{tn} = \prod_{i=0}^{n-1}(1 + r^D_{t+i}) - 1,$$

$$R^I_{tn} = \prod_{i=0}^{n-1}(1 + r^I_{t+i}) - 1.$$

To assess the tracking ability between the long-term holding of a LETF and the long-term holding of its underlying index, the authors first look at 2-day returns. They use the return formulas above when $n = 2$ to construct the following return series for the underlying index:

$$R^B_{t2} = (1 + r^B_t)(1 + r^B_{t+1}) - 1 = r^B_t + r^B_{t+1} + r^B_t r^B_{t+1}$$

From here, the author states that the 2x LETF has the following return over two days:

$$R^D_{t2} = (1 + 2r^B_t)(1 + 2r^B_{t+1}) - 1 = 2r^B_t + 2r^B_{t+1} + 4r^B_t r^B_{t+1}$$

This can also be written in terms of R^B_{t2}:

$$R_{t2}^{D} = 2R_{t2}^{B} + 2r_{t}^{B}r_{1+1}^{B}$$

Similarly, we can define the 2x inverse LETF return series to be:

$$R_{t2}^{I} = (1 - 2r_{t}^{B})(1 - 2r_{t+1}^{B}) - 1 = -2r_{t}^{B} - 2r_{t+1}^{B} + 4r_{t}^{B}r_{t+1}^{B}$$

This can also be written in terms of R_{t2}^{B} (Lu, Wang, and Zhang, 2009):

$$R_{t2}^{D} = -2R_{t2}^{B} + 6r_{t}^{B}r_{t+1}^{B}$$

As observed from the equations for the 2x long and inverse LETF returns, neither outcome delivers exactly 2x or -2x the performance of the underlying ETF. Notably, the inverse 2x LETF deviates even further from the underlying benchmark than that of the long 2x LETF. In the inverse case, we observe 2x the negative two-day return (which is ideal), but this is coupled with a 6x multiplier of the additional $r_{t}^{B}r_{t+1}^{B}$ term. In the case of the long LETF equation, while we obtain the correct 2x of the two-day benchmark return, this is similarly affected by the addition of the $r_{t}^{B}r_{t+1}^{B}$ term, albeit with a smaller coefficient than when compared to the short case.

The authors further suggest that long-term LETF performance is negatively related to quadratic variation, which can be computed by multiplying the number of days n and the variance estimate from the daily returns in the same n-day period. They then define S as the stock index price, where under standard Geometric Brownian Motion, μ is the return, σ is the volatility, and W denotes Brownian motion.

$$dS/S = \mu dt + \sigma dW$$

Next, assuming that the LETF tracks its underlying index perfectly, without interruption, the author writes the expected continuously compounded long return for the underlying ETF between $t = 0$ and $t = T$ as follows::

$$R_{T}^{B} = \mu T - \sigma^{2}T/2$$

Next, by defining Z to be the LETF price and multiple over its underlying ETF N, they derive:

$$dZ/Z = NdS/S = N\mu dt + N\sigma dW$$

Now, they define expected continuously compounded long-term returns of the LETF between $t = 0$ and T to be:

$$R_{T}^{L} = N\mu T - N^{2}\sigma^{2}T/2$$

Finally, by analyzing the N multiple (which represents the coefficient of the LETF versus its underlying) of R_{T}^{B} and R_{T}^{L}, the author yields (Lu, Wang, and Zhang, 2009):

$$R_{T}^{L} = NR_{T}^{B} - [(N^{2} - N)/2]\sigma^{2}T$$

In these equations, we find that the long-term return of an LETF compared to N times its underlying index differs by $-[(N^{2}-N)/2]\sigma^{2}T$. $-[(N^{2}-N)/2]$, representing the quadratic variation, is found to be negative and significant for all LETFs. Additionally, $\sigma^2 T$ represents the quadratic variation as it relates to the Brownian motion process component between $t = 0$ and $t = T$, exacerbating the tracking misalignment. In general, these two proofs demonstrate how LETFs, and in particular inverse LETFs, exhibit greater variation than N times the underlying index. In our research, we will look to verify the variability of LETF return tracking and whether the tracking ability between long and short LETFs is as asymmetric as presented in this paper.

3.4 Long-Term Investing in Triple Leveraged Exchange Traded Funds

Lewis Glenn's research on the long-term investment potential of LETFs presents a different perspective from those that were previously bearish on LETF performance. Glenn begins by acknowledging that as a result of daily rebalancing and beta slippage, the data has shown most 3x LETFs to perform quite poorly over longer periods. However, he goes on to identify an exception to this "rule", in the case of TQQQ, which has been shown to deliver a total return of over 10,000% over its 10-year lifespan. (Glenn, 2011). He does note the significant volatility associated with these returns, with monthly drawdowns reaching up to 49% and the largest daily drawdown nearing 70%. At the same time, patterns of large returns accompanied by large variations are seen in other 3x LETFs tracking Dow and the S&P 500. Given the high risks, Glenn proposes a strategy to limit these drawdowns by creating a portfolio that consists of an equal dollar amount of TQQQ and TMF, a 3x LETF tracking the 20+ year treasury bond, rebalanced on a bi-monthly basis (Glenn, 2011).

Glenn first defines the value of an underlying ETF to be $v(n)$ for day n, where N represents the number of trading days since the ETF's release, and the current day, n is bounded from $1 \leq n \leq N$. Next, one can normalize this function utilizing C as the cost basis where:

$$V(n) = v(n)/C$$

On day n, we can say that the entire return is: $V(n) - 1$. We next define our performance metrics, first with our Cumulative Annualized Growth Rate (CAGR) on day n in percentage terms (where the number of trading days is set to 250):

$$CAGR(n) = \left\{ \left[\exp\left(\frac{\log(V)}{n/250} \right) - 1 \right] \times 100 \right\}$$

Next, the daily drawdown is expressed as:

$$DDD(n) = \left\{ \frac{\max_{i=1}^{n}[V(i)] - V(n)}{\max_{i=1}^{N}[V(i)]}, \quad n = 1, \ldots, N \right\}$$

Using this, we obtain the max daily drawdown as (Glenn, 2011):

$$DDD_{\max}(n) = \max_{i=1}^{n} [DDD(i)]$$

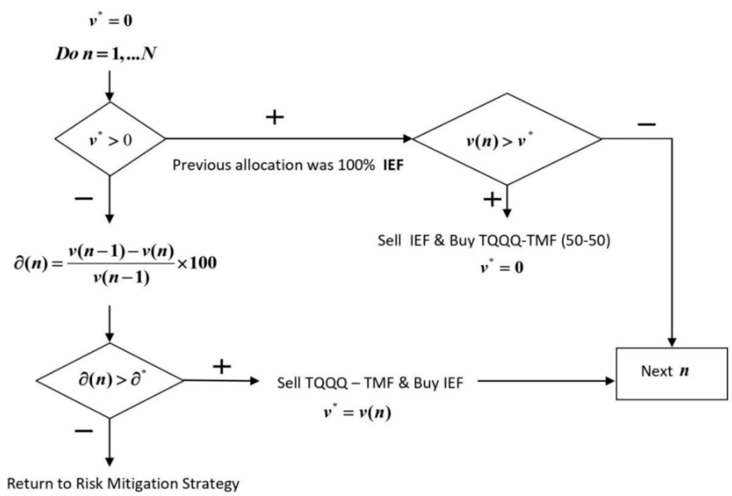
From here, the author implements his aforementioned portfolio strategy of equally weighting TQQQ and TMF, which, when balanced bimonthly, results in a portfolio with a CAGR of 45% and a max end-of-month drawdown of 24.5%, which is substantially reduced from that of TQQQ alone. However, because this strategy still has a max daily drawdown of 42.3%, as a result of a single-day downturn of over 20%, the author next presents a modification to the strategy. He first defines the daily negative return in TQQQ in percent terms:

$$\delta(n) = \frac{v(n-1) - v(n)}{v(n-1)} \times 100$$

In this modified strategy to the original equal dollar weighting approach, the author proposes switching to a total near-cash option when a large loss past a pre-set threshold occurs in TQQQ.

One would only return to the original strategy once the daily price of TQQQ matches or rises above a different set value. This new strategy ultimately results in better risk-adjusted returns than the original and a max daily drawdown of only 30%.

This strategy is outlined below where δ is defined as the value at which one moves their portfolio holdings to IEF (risk-free treasury ETF). Similarly, once TQQQ reaches the price level of v^* or exceeds it, the strategy reverts to its original, equally weighted form. Here, the author optimizes for the value of δ to ensure that the strategy's drawdown doesn't surpass 20%.



(Glenn, 2011)

While both the author's original and modified strategies have been backtested to perform well for LETFs tracking the S&P 500 (UPRO) and Dow Industrials (UDOW), the author states that this strategy's future success is contingent on a continuation in their performance since the publication of this paper (Glenn, 2011). We have included this trading strategy to identify the positive potential of long-term LETF holdings (such as TQQQ and UPRO) and how portfolios can be created to mitigate their risk potential. At the same time, however, this paper has demonstrated high levels of hindsight bias and overfitting in both the tickers it selects, and the parameters fine-tuned in its strategy. In our research, we hope to build on these findings on TQQQ as a long-term investment, as well as present potential portfolio strategies that are less grounded to fix for specified periods of past performance over predetermined sets of time.

3.5 Portfolio Insurance Using Leveraged ETFs

Author Jeffrey George explores the usage of LETFs as a hedging tool within the framework of Constant Proportion Portfolio Insurance (CPPI). CPPI is a strategy in which a portion of the overall portfolio is allocated to a risky asset, while the remainder is invested in a risk-free asset. If the value of the risky asset decreases, its proportion in the portfolio is reduced. Once this value declines past a certain threshold, the percentage allocated to this asset can eventually reach zero (George, 2017).

George proposes that incorporating LETFs into CPPI strategies can yield higher returns, even when considering financing costs, return decay, and high expense ratios. He suggests that rather than investing 50% of the portfolio in equities and allocating the remaining 50% to earn the risk-free rate, one could utilize a 2x LETF requiring a CPPI strategy to be only 25% invested in equities (through the LETF), allowing an additional 75% of the portfolio to now earn the risk-free rate. Similarly, the usage of a 3x LETF would require only a 16.67% allocation in the risky asset. In general, the improvement of this strategy stems from the ability to allocate a larger portion of the portfolio to a risk-free asset while effectively hedging the riskier asset.

For this study, George utilizes the 1-year Treasury as the risk-free asset and the S&P 500 as the underlying index when studying four different CPPIs which include CPPI S&P, CPPI 2x, CPPI 3x, and CPPI 4x. In terms of performance metrics, we can first define the Sharpe ratio as:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

where:

- R_p = Return of Portfolio
- R_f = Risk-Free Rate
- σ_p = Standard Deviation of the Portfolio's Excess Return

The Sharpe ratio is the primary metric by which investors measure risk-adjusted returns. This formula compares a fund's historical returns with its variability, returning a metric that allows one to compare differing strategies based on their risk-reward trade-off. Here, the risk-free rate refers to the risk premium, or how much one can make by investing in a "safe" asset, such as treasury bills (Fernando, 2024). In general, one always seeks to obtain a Sharpe ratio that is positive, and a Sharpe that is greater than one is typically considered quite strong in terms of performance, however, this threshold is quite subjective.

Next, the author defines daily returns for the various LETFs examined as follows:

$$R_L = LR_{S\&P} - R_{exp} - (L-1)R_f$$

In the above equation, $R_{S\&P}$ represents the S&P's daily returns, R_{exp} is the daily expense ratio, R_f is the borrowing rate, and R_L is the LETF's daily return with a leverage ratio of L. Next, the proportion allocated in the risky asset for the CPPI S&P at time t can be calculated to be:

$$\max\left\{\min\left[m(V_p - F), V_p\right], 0\right\}/V_p$$

Here, VP is the portfolio value, F is the floor, and m represents the multiplier. In this study, the multiplier is set to 5, indicating that 50% of the initial investment is in the risky asset. For a

2x LETF, the multiplier is set to 2.5 to maintain 50% exposure, and for a 3x LETF, the multiplier is set to 1.67. Additionally, the researchers decided to rebalance the portfolio only when the risky asset moves up or down by 2.5%, acknowledging that intra-month leverage decay can punish the accuracy of monthly rebalancing. Next, the formulas for the CPPI 2x and CPPI 3x risky asset proportions are defined similarly as follows (George, 2017):

$$\min\left[\max\left\{\min\left[m(V_p - F), V_p\right], 0\right\}/V_p, 0.5\right]$$
$$\min\left[\max\left\{\min\left[m(V_p - F), V_p\right], 0\right\}/V_p, 0.33\right]$$

George's analysis finds that in simulated scenarios from 1947 to 2016, when the risk-free rate is greater than 3%, LETF usage within CPPI portfolios yields annual returns that are 0.5% to 1.3% higher than those of traditional CPPI strategies. This increase in returns also leads to higher Sharpe ratios, primarily due to a reduction in Value at Risk (VaR) (George, 2017). However, George highlights that while CPPI strategies with 2x and 3x LETFs yield higher returns, they also possess greater risks as a result of larger standard deviations. Despite this, the improved Sharpe ratios point towards the CPPI LETF strategy being superior to the conventional 50% allocation strategy. For investors with higher risk tolerance, opting for more highly leveraged CPPI portfolios may be considered optimal. The one identified drawback is that from 2010 to 2017, LETFs in CPPI strategies underperformed as a result of the risk-free rate being below 1% (George, 2017).

While we do not seek to replicate the portfolio insurance techniques presented by George, understanding these results can provide further justification for investigating the various use cases of LETFs. Similarly, this paper provides a foundation for LETF portfolio construction and the various risk metrics we plan to study.

4 Leveraged ETF Tracking Abilities

We now will conduct data analysis on LETFs' potential viability as long-term investments to verify whether the advertised leverage ratios of the studied LETFs match up statistically with the beta between their historical returns vs. their benchmarks'. This process will enable us to compare modern-day findings with those in commonly cited academic papers, allowing us to assess the results of our literature review and develop realistic, modern performance expectations for LETFs to see if they are viable investment vehicles. For instance, while the WSJ has spoken quite positively to LETF's outperformance, such as in "Are Leveraged ETFs Worth the Costs and Risk?" (Horstmeyer, 2023), this article fails to include a wide variety of LETFs, look back periods beyond recent bull runs, or an outline of any rigorous statistical techniques. We hope to address several of the claims made in the literature review about tracking ability, long-term investment viability, and potential for portfolio diversification utilizing the most recent data in conjunction with different statistical methodologies.

4.1 Data Collection

From the information available online, we have found there to be 211 LETFs on the U.S. stock market in 2024. For the first step of the data collection process, Stock Market MBA's "List of Leveraged ETFs" was used to access a list of all current LETFs, which also included their respective leverage ratio, category, market cap, and average volume ("List of Leveraged ETFs", 2023). We then manually sourced each of these LETFs' underlying ETFs from websites including Direxion, ProShares, and others. From here, the Yahoo Finance API was used to pull daily stock price close data since inception for each of these LETFs and their respective underlying ETFs. Because the majority of these tickers did not have publicly available data on Yahoo, Bloomberg was employed to pull data on the rest of the tickers.

In terms of LETF data collection, we opted to not utilize any LETFs that had been delisted in the past year. Similarly, any LETF/ETF pairings with less than 2 years of available data were removed for the sake of robust regression results. This process enabled us to construct a dataset of the remaining 186 LETF/ETF pairings up through early 2024, with each of their daily closes since inception, giving us a much more comprehensive dataset than any of the other papers studied.

4.2 Methodology

Utilizing this dataset, we wanted to understand which LETFs track better than others and why. To measure this, a for-loop was developed that regressed each pairing of LETF onto its underlying ETF through its entire lookback period. In this function, the latter of the two start-dates between the LETF and its respective ETF was used in all regression outputs. As mentioned in our literature review, one would hope to see a beta coefficient in our regression that matches up with the advertised leverage ratio of the LETF. Specifically, from this regression one would expect that for an arbitrary LETF/ETF pairing, the daily returns of the LETF would be identical to β times the daily returns of the underlying ETF. After running 186 regressions, we plan to compare each of the obtained beta coefficients with the respective advertised leverage ratio. From here, we will then run hypothesis tests to find the proportion of LETF betas that differed significantly from their leverage ratios.

Beyond regressing just from the earliest data available for each of the LETF/ETF pairings, we will also regress these pairings across different market regimes, such as bear/bull market and periods of high volatility, to see if tracking remains the same or improves/decays under different market regimes. Our results will then be visualized to see whether factors such as the ‘Leverage Ratio’, ‘Category’, ‘Market Cap’, or ‘Average Volume’ have any correlation with how well an LETF tracks its underlying index, utilizing the R^2 from our regressions as a proxy.

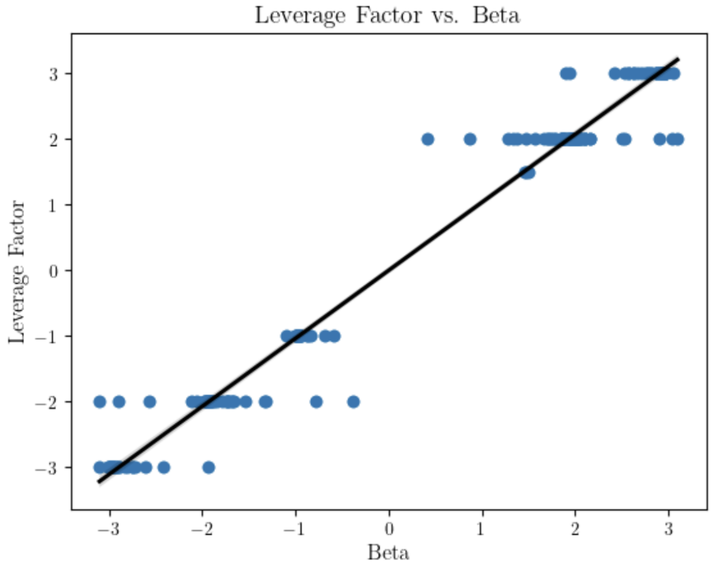
Finally, we will backtest LETF returns across various periods, in response to the cited WSJ article and our literature review, to see how promising the long-term returns are. Our research will also attempt a portfolio optimization approach (different from that in the portfolio insurance paper) to understand if LETFs can be utilized as hedging tools and if any combination of risky assets can result in a portfolio with dampened volatility.

4.3 Statistical Testing and Results

From the results of our regressions on each LETF/ETF pairing from inception through early 2024, 186 rows were obtained in our new dataset, copied below. This regression takes about 2 minutes to run. Moreover, all of these regressions have alpha values of near 0 (with only one regression having an alpha marginally higher than .001), indicating that nearly all of the performance in the LETF could be attributed to its tracking of the underlying ETF.

	LETF Symbol	ETF Symbol	Leverage Factor	Category 1	Category 2	Category 3	Market Cap	Average Volume	Alpha	Beta	Standard Error	Number of Observations	R^2	MSE
0	TQQQ	NDX	3.0	US Equity	Broad market	Large cap	1.277489e+10	164492850.0	-0.000017	2.925249	0.003	3469.0	0.996886	5.121809
1	SOXL	ICESEMIT	3.0	US Equity	Single industry	Semiconductors	5.681628e+09	85193280.0	-0.000127	2.938764	0.006	2809.0	0.988750	7.961999
2	SQQQ	NDX	-3.0	US Equity	Broad market	Large cap	5.147058e+09	101902900.0	-0.000096	-2.931665	0.003	3469.0	0.996117	5.144300
3	QLD	NDX	2.0	US Equity	Broad market	Large cap	3.592880e+09	6670822.0	-0.000023	1.939941	0.003	4386.0	0.990664	3.393206
4	SSO	^SPX	2.0	US Equity	Broad market	Large cap	3.163116e+09	7466952.0	0.000053	1.945431	0.003	4410.0	0.990301	2.655659

First, one can look at a plot of the advertised leverage ratio vs. the beta, where our betas represent the actualized leverage ratios of these LETFs. We can see that two generally align with a linear fit, but there are several deviations, primarily around the -2x and 2x leverage ratios, where slight deviations can be visually identified.



To prove from a statistically significant lens, we run a hypothesis test for each of these pairings to see if the beta is statistically significantly different from the leverage ratio:

$$\text{Test Statistic} = \frac{\text{Sample Statistic} - \text{Value of the Population Parameter Under } H_0(\theta_0)}{\text{Standard Error of the Sample Statistic}}$$

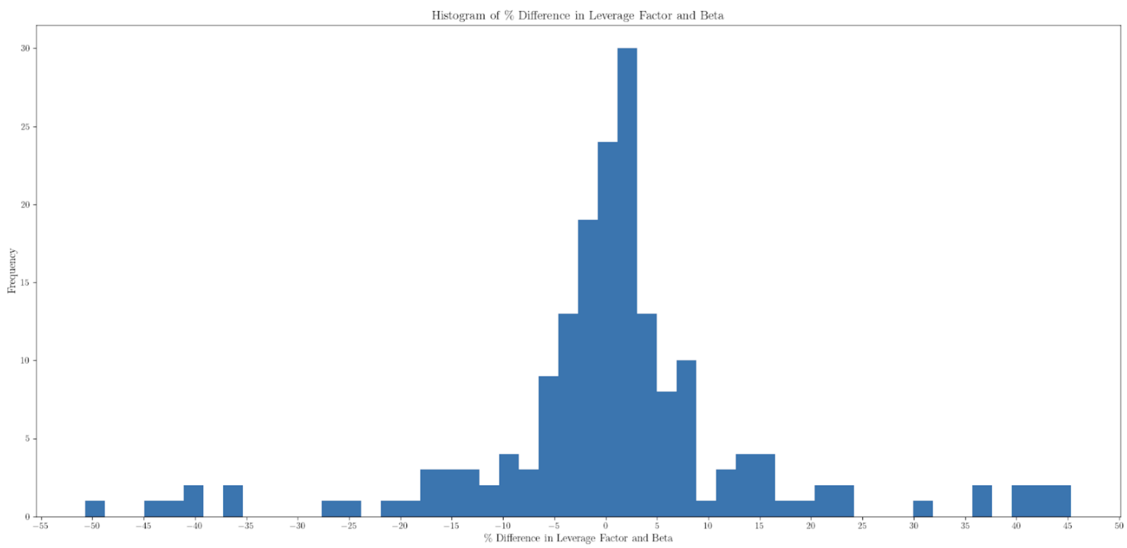
$$\text{Test Statistic} = \frac{\bar{X}_\mu - \theta_0}{s_{\bar{X}}} = \frac{\bar{X}_\mu - 0}{s/\sqrt{n}}$$

We run a two-tailed test where (Margenot and Granizo-Mackenzie, 2018):

$$\text{P-Value} = 2 \times (1 - CDF(|\text{Test Statistic}|))$$

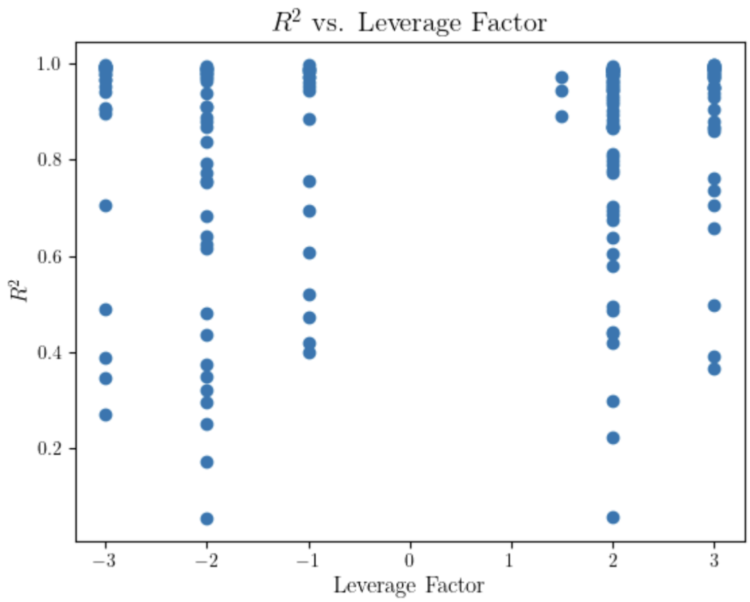
We ultimately find that from inception through the present, **12.36%** of LETFs tracked their underlying ETF to a statistically significant degree. Tracking ability over various periods is elaborated upon in the next section. We now check to see whether the advertised leverage ratios fall in or outside of the 95% confidence interval constructed for each of the betas. Similarly, a histogram is included below of the percent difference between the beta and the leverage ratio for all the LETFs studied. The percent difference is a metric utilized to measure how accurately an LETF's performance can match its advertised leverage ratio. LETFs that have near-perfect matching have a percent difference of 0%, and a percent difference of between -1% and 1% can be subjectively considered to have relatively strong accuracy. However, from our histogram below, we see a slightly skewed normal distribution, meaning that while most LETFs track accurately, it appears that there is a prevalence of betas that are misaligned with their actual leverage ratios.

Percent Difference	Count	Proportion of Counts
$[-100, -30]$	7	3.85%
$[-30, -20]$	3	1.65%
$[-20, -10]$	13	7.14%
$[-10, -5]$	13	7.14%
$[-5, -1]$	31	17.03%
$[-1, 1]$	23	12.64%
$[1, 5]$	47	25.82%
$[5, 10]$	18	9.89%
$[10, 20]$	14	7.69%
$[20, 30]$	4	2.20%
$[30, 100]$	9	4.95%



We have also included a table of our outputs and the respective count of each observation:

4.4 Decomposing Tracking Error



We first look to decompose LETF tracking error by analyzing variability across leverage ratios (-3x, -2x, -1x, 1.5x, 2x, or 3x). Specifically, we utilize R^2 to track how well the variability in the LETF is explained by the ETF. From our plot above, one can see that the R^2 values appear relatively constant across all leverage ratios, indicating that the specific leverage ratio of the LETF has little impact on its ability to track its underlying. Additionally, the leverage ratio and R^2 have a .12 correlation, which demonstrates a near insignificant relationship.

Next, as related to the third paper in our literature review, we wanted to see if short LETFs have significantly different tracking than long LETFs. We conducted a hypothesis test, splitting up our sample into long LETFs and short LETFs and then pulling the proportion of LETFs that had betas that differed significantly from their leverage ratios for both long and short groups. Next, we run a two-proportion z-test using the following proportions and hypotheses:

Long LEFs - Proportion Tracked Inaccurately: 0.89, SE: 0.31
Short LETFs - Proportion Tracked Inaccurately: 0.85, SE: 0.36

$$H_0 : p_1 = p_2$$

$$H_a : p_1 \neq p_2$$

We construct our sampled proportions and pooled sample proportions below:

$$\hat{p}_1 = \frac{x_1}{n_1}, \quad \hat{p}_2 = \frac{x_2}{n_2}$$

$$\hat{p}_{\text{pooled}} = \frac{x_1 + x_2}{n_1 + n_2}$$

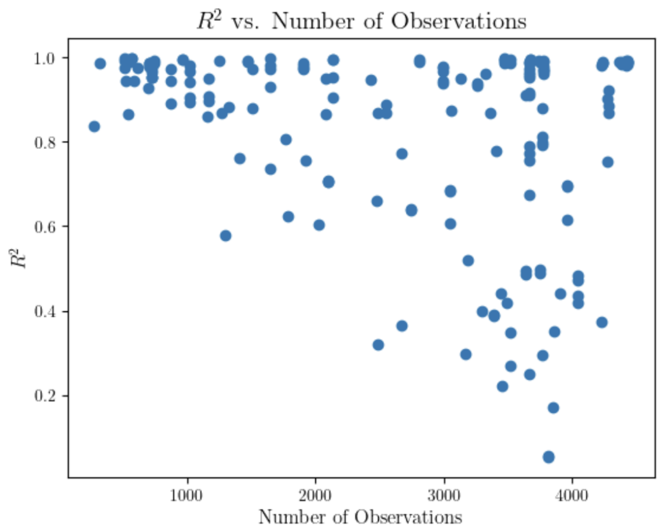
We then define our SE, Z-Statistic, and p-value as follows (Margenot and Granizo-Mackenzie, 2018):

$$SE = \sqrt{\hat{p}_{\text{pooled}} \cdot (1 - \hat{p}_{\text{pooled}}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

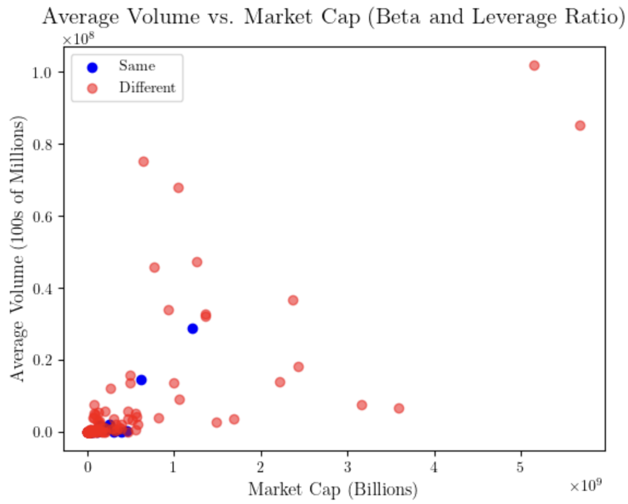
$$z = \frac{\hat{p}_1 - \hat{p}_2}{SE}$$

$$\text{P-Value} = 2 \times P(Z > |z|)$$

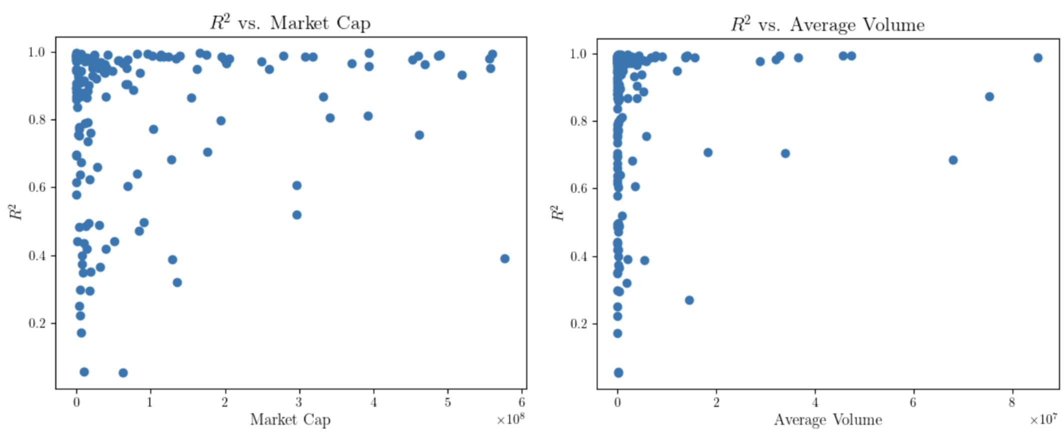
We eventually get a p-value of .37 which is greater than our alpha of .05, indicating that one cannot confidently conclude that long and short LETFs differ in their tracking abilities to a statistically significant extent.



Now, we want to see if there is any correlation between the R^2 and how many data points are available, which measures how long an LETF has been in existence. The graph above shows a slight negative correlation at -.29, meaning that an LETF that has existed longer may be more inaccurate in terms of its tracking ability, which aligns with our literature review, as there is more time for tracking errors to compound.



The graph above looks to depict whether there is a relationship between the average volume vs. the market cap of the LETF and whether or not its beta is the "same", statistically, as its leverage ratio. For this study, the market cap is defined as the total value of an LETF's outstanding shares, and the average volume is the average number of shares an LETF trades daily. However, from the plot above, one can find little discernible relationship between "Same" vs. "Different" points. Next, we examine market cap and average volume individually, and how they each correlate with the R^2 of the regression output.



In both graphs above, it appears that LETFs with larger market caps or larger average volumes generally have more robust tracking. However, there are only correlations of .15 and .12 in the two graphs, respectively. Similarly, this result may be confounded by survivorship bias where LETFs that have larger volume and are thus popular are simply those that have stood the test of time, whereas LETFs that had worse tracking may have been delisted over time.

We next look at tracking ability over the various categories of LETFs. Specifically, each LETF has three category breakdowns provided by "Stock Market MBA". Below is a table containing the overall percentage breakdown per category across all 186 LETFs for categories that had more than one entry. We then pulled the subsection of LETFs that possessed statistically significant tracking ability and included this corresponding category percentage breakdown. Finally, for each category, the categories are sorted by their average R^2 values in descending order, retaining categories that have more than one entry once again. In general, LETFs that track debt, fixed income, and commodities generally have stronger tracking abilities than equities at large, when controlling for their original percentage breakdown across LETFs.

Category 1 (Overall)		**Category 1 (Significant Tracking)**		**Category 1 (Average R^2 Descending)**	
US Equity	0.56	US Fixed Income	0.45	US Equity	0.91
Global Equity	0.24	US Equity	0.25	US Fixed Income	0.88
Commodities	0.10	Global Equity	0.15	Special Security Types	0.81
US Fixed Income	0.07	Commodities	0.10	Global Equity	0.67
Special Security Types	0.02	Special Security Types	0.05	Commodities	0.66

Category 2 (Overall)		**Category 2 (Significant Tracking)**		**Category 2 (Average R^2 Descending)**	
Single Industry	0.25			Broad Market	0.94
Broad Market	0.23			Single Industry	0.92
Single Sector	0.18	Inverse Debt	0.30	Inverse Debt	0.91
Single Country	0.09	Leveraged Debt	0.15	Single Sector	0.90
Inverse Debt	0.05	Broad Market	0.15	Leveraged Debt	0.83
Precious Metals	0.04	Single Industry	0.15	Energy	0.76
Energy	0.03	Energy	0.10	Precious Metals	0.67
Leveraged Debt	0.03	Single Sector	0.05	Europe	0.65
Emerging Markets	0.03	Single Country	0.05	Single Country	0.50
Currency	0.02	Closed-End Funds	0.05	Currency	0.49
Developed Markets ex-US	0.02			Developed Markets ex-US	0.48
Europe	0.02			Emerging Markets	0.47

Category 3 (Overall)		**Category 3 (Significant Tracking)**		**Category 3 (Average R^2 Descending)**	
Large Cap	0.18			Consumer Discretionary	0.98
Large and Mid-Cap	0.10			Biotechnology	0.97
Information Technology	0.06			Financials	0.97
Energy	0.05			Treasury Bond Index	0.97
Small Cap	0.05			U.S. Treasury Bonds	0.97
Mid-Cap	0.04			Information Technology	0.96
Treasury Bond Index	0.04			Small-Cap	0.96
Multi-Cap	0.04			REITs	0.95
Biotechnology	0.03			Energy	0.94
Gold	0.03	Treasury Bond Index	0.32	Real Estate	0.94
Gold Miners	0.03	U.S. Treasury Bonds	0.11	Banking	0.94
Financials	0.03	Oil	0.11	Utilities	0.94
Single Currency	0.02	Mid-Cap	0.11	Gold Miners	0.92
Oil	0.02	Information Technology	0.11	Retail	0.92
Real Estate	0.02	Large-Cap	0.05	Mid-Cap	0.88
Semiconductors	0.02	REITs	0.05	Industrials	0.88
U.S. Treasury Bonds	0.02	Large and Mid-Cap	0.05	Materials	0.86
Banking	0.02	Corporate High Yield	0.05	Semiconductors	0.83
Basic Services	0.02			Basic Services	0.83
Industrials	0.02			Oil	0.79
Utilities	0.02			Large-Cap	0.77
REITs	0.01			Multi-Cap	0.76
Silver	0.01			Gold	0.72
Natural Gas	0.01			Natural Gas	0.69
Materials	0.01			Large and Mid-Cap	0.57
Retail	0.01			Silver	0.55
Consumer Discretionary	0.01			Single Currency	0.49
Consumer Staples	0.01			Consumer Staples	0.46

4.5 Different Time-Frames

In addition to looking at LETF tracking abilities since their inception, we now want to see how LETF tracking accuracy compares across various periods. Specifically, we will look at how tracking improves or deteriorates over a variety of market conditions, which include historical bull and bear runs and periods of high and low volatility, respectively. Moreover, to supplement our literature review, LETF tracking accuracy is studied across more recent years, after the publication of the papers from the literature review.

We first identify prominent bull runs to fall between 2020 - 2019 and 2017 - 2020 and bear runs to be between 2007 - 2009 (2008 stock market crash) and 2020 - 2021 (COVID-19). Additionally, we also looked at more recent tracking accuracy, including ranges from 2020 - 2023 and 2022 - 2024. For each of these time frames, we include the percentage of betas that are the same as their leverage ratio, in terms of statistical significance, as well as the percentage of those that are within |5%| and |10%| of their leverage ratio. We also added our initial time frame which includes the inception of each of the LETFs through 2024 and indicated how many LETFs are included in each time frame, as a result of certain dates possessing less data.

Since Inception

Inception of LETFs - 2024 (186 Available)
% of Betas Same as Leverage Ratio: 0.12
% of Betas Within -5% / 5% of Leverage Ratio: 0.54
% of Betas Within -10% / 10% of Leverage Ratio: 0.71

Bull Runs

2010 - 2019 (161 Available)
% of Betas Same as Leverage Ratio: 0.10
% of Betas Within -5% / 5% of Leverage Ratio: 0.41
% of Betas Within -10% / 10% of Leverage Ratio: 0.57

2017 - 2020 (166 Available)
% of Betas Same as Leverage Ratio: 0.09
% of Betas Within -5% / 5% of Leverage Ratio: 0.43
% of Betas Within -10% / 10% of Leverage Ratio: 0.61

Bear Runs

2007 - 2009 (92 Available)
% of Betas Same as Leverage Ratio: 0.15
% of Betas Within -5% / 5% of Leverage Ratio: 0.34
% of Betas Within -10% / 10% of Leverage Ratio: 0.62

2020 - 2021 (185 Available)
% of Betas Same as Leverage Ratio: 0.23
% of Betas Within -5% / 5% of Leverage Ratio: 0.55
% of Betas Within -10% / 10% of Leverage Ratio: 0.74

Recent History

2020 - 2023 (186 Available)
% of Betas Same as Leverage Ratio: 0.17
% of Betas Within -5% / 5% of Leverage Ratio: 0.57
% of Betas Within -10% / 10% of Leverage Ratio: 0.71

2022 - 2024 (186 Available)
% of Betas Same as Leverage Ratio: 0.31
% of Betas Within -5% / 5% of Leverage Ratio: 0.70
% of Betas Within -10% / 10% of Leverage Ratio: 0.77

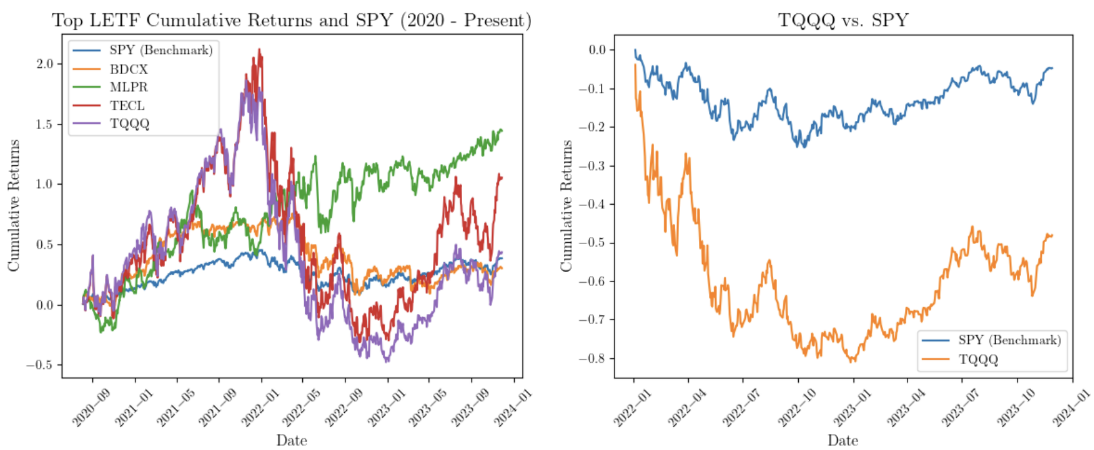
The results above show that generally shorter time frames result in improved LETF tracking abilities. Similarly, it appears that LETF tracking has been stronger in more recent years, although these periods are also typically on the shorter end. In general, there does not appear to be any clear relationship between tracking ability and market performance or volatility.

5 Leveraged ETF Investment Viability

Finally, we want to test the long-term investment viability of LETFs by both implementing Mean-Variance Portfolio Optimization and analyzing the results of a buy-and-hold strategy.

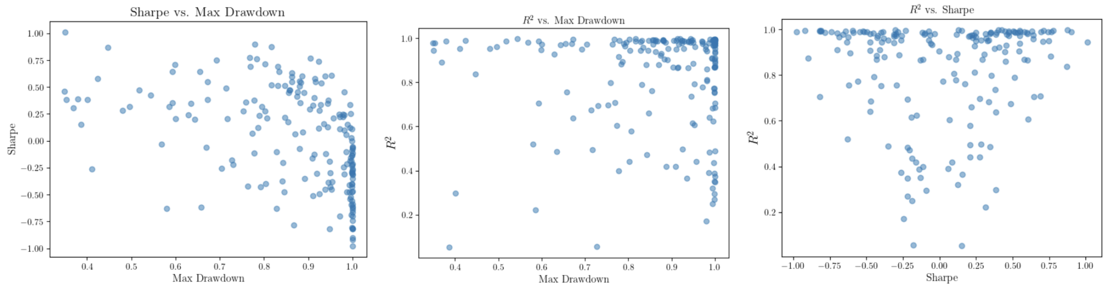
5.1 Buy and Hold Long-Term Performance

We now want to look at the long-term buy-and-hold performance of a variety of LETFs, separate from their tracking abilities. To standardize their returns, this study will look at the percentage of currently existing LETFs that have outperformed SPY, a benchmark for the S&P 500, across a variety of periods. Specifically, we will look at how LETFs performed in aggregate in terms of both their absolute returns and their risk-adjusted Sharpe ratio. The graphs below provide an initial snapshot of what top LETF performance has looked like since 2020, when compared to SPY, on an absolute basis. These graphs begin to visualize both the massive volatility and potential downswings these products can face, detracting from their high absolute returns, a consequence explored below.



(Top LETF Absolute Returns vs. Benchmark)

First, we construct a buy-and-hold strategy for all of the studied LETFs from their inception through the end of 2023. For each of these LETF holdings since inception, there exists a -.45 correlation between an LETF's Sharpe and its max drawdown. This result can potentially be driven by LETFs that have failed in the past, resulting in max drawdowns of 100% and deeply negative Sharpes. Next, there is a correlation of approximately 0 between tracking ability (as measured by R^2) and max drawdown. Finally, we find a **correlation of .03 between Sharpe and tracking ability**. These results lead us to believe that an LETF's ability to track doesn't necessarily imply it is a superior or inferior product.



(Top LETF Absolute Returns vs. Benchmark)

Next, we take a more nuanced view by analyzing the buy-and-hold returns between each LETF and its corresponding underlying ETF. These returns are run over a variety of three-year periods centered around 2020 to 2023 (2018 - 2021, 2019 - 2022, 2020 - 2023) and find that on average, while the annualized daily mean returns of LETFs generally align with their leverage ratio multiplied by the annualized daily mean returns of their underlying ETFs, LETFs always possess a lower Sharpe than these ETFs. This discrepancy can be attributed to amplified volatility, which also exacerbates max drawdown. For instance, from 2020 to 2023, TQQQ and QQQ demonstrated the following return statistics:

Product	Annualized Daily Mean Returns	Annualized Daily Standard Deviation	Max Drawdown	Sharpe
QQQ	.18	.21	.36	.89
TQQQ	.53	.61	.72	.86

A crucial result is that when we expand upon this research, looking at the performance of each of our 186 LETFs versus their underlying index over 7 three-year periods, **we were not able to come up with any result where the LETF produced a higher Sharpe than its underlying index, across any period**. As a result of LETFs never appearing to be able to outperform their underlying on a risk-adjusted basis, it is plausible to consider them as inferior assets from an expected value perspective and to be unfit for an economically rational investor. In fact, during these periods studied, it was found that the average LETF drawdown was 81.59% and the median was 75.83%, which was significantly larger than the S&P 500's max drawdown of 31.5% (Putnam, 2023). These results make intuitive sense from our literature review and prior results where LETFs, as a result of their fees and constant rebalancing, have no signal to outperform their underlying indices from a risk perspective.

Additionally, we find that LETFs' high absolute returns are not the norm and that on average, SPY beats the aggregate of LETFs in terms of both absolute returns and Sharpe. Specifically, below is a list of the percentage of LETFs that beat SPY over the key market periods previously identified: 2020-2021, 2010-2019, 2007-2009, 2020-2021, 2017-2020, 2020-2023, and 2022-2023. Across these periods, we identify how many of the 186 LETFs were available, the percentage of LETFs with higher Sharpe than SPY (which was typically quite small), the percentage with higher annualized means, and then the specific LETFs with higher Sharpes and with both higher annualized means and Sharpes than SPY.

Since Inception

Inception of LETFs - 2024 (186 Available)
Percent of LETFs with Higher Sharpe: 22.04%
Percent of LETFs with Higher Annualized Mean: 41.94%
LETFs Higher Sharpe: TQQQ, SOXL, QLD, SSO, SPXL, UPRO, TECL, FAS, TNA, UDOW, SARK, ROM, DDM, SVXY, USD, SPUU, URTY, CURE, NAIL, MVV, BIB, HIBL, MIDU, RETL, DRN, WANT, UMDD, TPOR, DUSL, UXI, UCC, UGE, MJIN, BDCX, FBGX, FEDL, FNGO, FNGU, NRGU, OILU, SCDL
LETFs Higher Annualized Mean and Sharpe: TQQQ, SOXL, QLD, SSO, SPXL, UPRO, TECL, FAS, TNA, UDOW, SARK, ROM, DDM, SVXY, USD, SPUU, URTY, CURE, NAIL, MVV, BIB, HIBL, MIDU, RETL, DRN, WANT, UMDD, TPOR, DUSL, UXI, UCC, UGE, MJIN, BDCX, FBGX, FEDL, FNGO, FNGU, NRGU, OILU, SCDL

Bull Runs

2010 - 2019 (161 Available))
Percent of LETFs with Higher Sharpe: 6.83%
Percent of LETFs with Higher Annualized Mean: 35.4%
LETFs Higher Sharpe: TQQQ, QLD, UDOW, DFEN, CURE, WEBL, HIBL, UCC, BNKU, HDLB, PFFL
LETFs Higher Annualized Mean and Sharpe: TQQQ, QLD, UDOW, DFEN, CURE, WEBL, HIBL, UCC, BNKU, HDLB

2017 - 2020 (166 Available)
Percent of LETFs with Higher Sharpe: 10.24%
Percent of LETFs with Higher Annualized Mean: 44.58%
LETFs Higher Sharpe: TQQQ, SOXL, QLD, TECL, ROM, USD, WEBL, WANT, UST, UCC, BDCX, CEFD, FBGX, FNGO, FNGU, GDXU, MVRL
LETFs Higher Annualized Mean and Sharpe: TQQQ, SOXL, QLD, TECL, ROM, USD, WEBL, WANT, UCC, BDCX, CEFD, FBGX, FNGO, FNGU, GDXU, MVRL

Bear Runs

2007 - 2009 (92 Available)
Percent of LETFs with Higher Sharpe: 53.26%
Percent of LETFs with Higher Annualized Mean: 50.0%
LETFs Higher Sharpe: QLD, SPXL, SH, UPRO, TECL, FAS, SDS, TNA, UCO, RWM, TMV, ROM, AGQ, ERX, DOG, TBF, USD, UGL, YANG, DIG, MVV, TWM, DXD, EDC, EFZ, MIDU, UYM, DRN, TYO, SAA, SKF, EET, SBB, MYY, XPP, EFO, YCL, ULE, UPW, EWV, UGE, EZJ, SDD, SDP, EFU, SCC, SIJ, SZK, DGP
LETFs Higher Annualized Mean and Sharpe: QLD, SPXL, SH, UPRO, TECL, FAS, SDS, TNA, RWM, TMV, ROM, AGQ, ERX, DOG, TBF, USD, UGL, YANG, DIG, MVV, TWM, DXD, EDC, EFZ, MIDU, UYM, DRN, TYO, SKF, EET, SBB, MYY, XPP, EFO, YCL, ULE, EWV, UGE, EZJ, SDD, SDP, EFU, SCC, SIJ, SZK, DGP

2020 - 2021 (185 Available)
Percent of LETFs with Higher Sharpe: 12.43%
Percent of LETFs with Higher Annualized Mean: 39.46%
LETFs Higher Sharpe: TQQQ, SOXL, QLD, TECL, SARK, ROM, USD, RETL, WANT, UGE, MJIN, BDCX, CEFD, FBGX, FEDL, FNGO, FNGU, IWDL, IWFL, MVRL, QULL, SCDL, USML
LETFs Higher Annualized Mean and Sharpe: TQQQ, SOXL, QLD, TECL, SARK, ROM, USD, RETL, WANT, UGE, MJIN, BDCX, CEFD, FBGX, FEDL, FNGO, FNGU, IWDL, IWFL, MVRL, QULL, SCDL, USML

Recent History

2020 - 2023 (186 Available)
Percent of LETFs with Higher Sharpe: 6.99%
Percent of LETFs with Higher Annualized Mean: 38.17%
LETFs Higher Sharpe: TQQQ, SOXL, QLD, TECL, ROM, USD, NAIL, YCS, MJIN, BDCX, FNGO, FNGU, OILU
LETFs Higher Annualized Mean and Sharpe: TQQQ, SOXL, QLD, TECL, ROM, USD, NAIL, YCS, MJIN, BDCX, FNGO, FNGU, OILU

2022 - 2024 (186 Available)
Percent of LETFs with Higher Sharpe: 29.57%
Percent of LETFs with Higher Annualized Mean: 39.25%
LETFs Higher Sharpe: SOXL, TECL, UCO, GUSH, TBT, TZA, RWM, TMV, SARK, ERX, TTT, SJB, SVXY, TBF, USD, DRV, SRTY, UGL, DFEN, NAIL, YANG, KOLD, DIG, TWM, BRZU, LABD, SRS, TBX, EUO, REK, EUM, PST, WEBS, SEF, EDZ, TYO, YCS, FXP, DUSL, MEXX, SBB, EMTY, EEV, YXI, UBR, SDP, SZK, MJIN, BDCX, DGP, FEDL, FNGO, FNGU, NRGU, OILU
LETFs Higher Annualized Mean and Sharpe: SOXL, TECL, UCO, GUSH, TBT, TZA, RWM, TMV, SARK, ERX, TTT, SVXY, TBF, USD, DRV, SRTY, UGL, DFEN, NAIL, YANG, KOLD, DIG, TWM, BRZU, LABD, SRS, TBX, EUO, REK, EUM, PST, WEBS, SEF, EDZ, TYO, YCS, FXP, DUSL, MEXX, SBB, EMTY, EEV, YXI, UBR, SDP, SZK, MJIN, BDCX, DGP, FEDL, FNGO, FNGU, NRGU, OILU

While a select few of these products can outperform their benchmark on an absolute basis, they fail to maximize expected value when compared to their underlying index. This leads to a debate on whether the benefit of not needing leverage for a similarly levered position outweighs the reduced Sharpe ratio, which ultimately depends on investor preference. Typically, retail customers without access to cheap leverage might opt for LETFs, but they are also the ones less equipped to handle significant drawdowns. In the case that an investor has a high conviction in a particular LETF, and it is more economically sound for them to utilize a derivative product than taking out leverage to maximize absolute returns, then there may be an argument for these products' usage. Otherwise, it can be concluded that a rational economic actor should most likely opt for investment in any LETF's underlying rather than the LETF itself.

5.2 Portfolio Optimization Review

Finally, we want to explore portfolio allocation strategies to determine if the benefits from LETFs' absolute results can be complemented by risk-reduction techniques, ideally resulting in improved Sharpe ratios. Specifically, we will utilize Markowitz optimization, which involves taking long/short positions with various weightings to create a diversified portfolio of assets.

To begin, one must first introduce the basics of portfolio construction and its hedging advantages. Modern portfolio theory, or mean-variance analysis, aims to determine the optimal asset weightings in a portfolio that maximizes expected returns for a given level of risk. Moreover, by diversifying across multiple assets, one can reduce the overall risk, and thus the standard deviation of an entire portfolio.

We first start with two assets S_1 and S_2, each of which has weights w_1 and w_2 (where $w_1 + w_2 = 1$). Next, consider a portfolio P that holds these two assets, each with their respective weightings, where the two assets have the following mean returns and standard deviation:

$$P = w_1 S_1 + w_2 S_2$$

$$\mu_1, \sigma_1 \text{ and } \mu_2, \sigma_2$$

Now, one can compute the expected returns and standard deviation of the portfolio by decomposing the returns of the portfolio into the expected return of each asset multiplied by their respective weights.

$$E[\mu_P] = E[\omega_1 \mu_1 + \omega_2 \mu_2] = \omega_1 E[\mu_1] + \omega_2 E[\mu_2]$$

We now decompose the variance of the portfolio where one can dictate the following to be true for the variance of the portfolio and the correlation between S_1 and S_2:

$$\sigma_p^2 = \text{VAR}[P]$$

$$\text{COR}[S_1, S_2] = \frac{\text{COV}[S_1, S_2]}{\sigma_1 \sigma_2} = \rho_{12}$$

Utilizing these relationships yields the variance decomposition formula below:

$$\begin{aligned}
\sigma_p^2 &= \text{VAR}[P] \\
&= \text{VAR}[\omega_1 S_1 + \omega_2 S_2] \\
&= \text{VAR}[\omega_1 S_1] + \text{VAR}[\omega_2 S_2] + \text{COV}[\omega_1 S_1, \omega_2 S_2] \\
&= \omega_1^2 \text{VAR}[S_1] + \omega_2^2 \text{VAR}[S_2] + 2\omega_1 \omega_2 \text{COV}[S_1, S_2] \\
&= \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\rho_{12} \omega_1 \omega_2 \sigma_1 \sigma_2
\end{aligned}$$

Finally, for a portfolio that holds n assets, one gets a variance constructed as follows (Margenot and Granizo-Mackenzie, 2018):

$$\sigma_p^2 = \sum_i \omega_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} \omega_i \omega_j \sigma_i \sigma_j \rho_{ij}, \quad i, j \in \{1, \ldots, n\}$$

One can see that for portfolios made up of n assets, our risk, measured in variance, continuously goes down with the addition of each asset. Additionally, through modern portfolio theory, an investor can reduce their risk even further by ensuring that the assets in their portfolio are uncorrelated (thereby possessing a ρ_{ij} of near 0, reducing the number of terms in the equation above). Therein lies the premise of Markowitz portfolio optimization, as it constructs a portfolio that incorporates the risk-free rate to maximize expectancy for the specified risk an investor is willing to put on. We seek to apply this portfolio methodology to LETFs, which has not been previously applied in other research papers.

Applying Markowitz portfolio optimization calls for an assessment of its assumption. Specifically, it asks that investors are rational, risk-averse, can handle bearable loss, and have unlimited access to market change information. It next assumes that markets are efficient and that only two types of assets exist, those with low returns and those with high returns. Markowitz portfolio also emphasizes the importance of diversification and that the only way to effectively minimize one's portfolio risk is by reducing the variability of returns.

With this in mind, one can now construct an efficient frontier by first considering a 2D graph where the risk (measured in standard deviation) is on the x-axis and the returns are on the y-axis. We then construct all possible portfolio combinations of our assets, creating an area of various points along these axes. The efficient frontier refers to the portfolios that have means and standard deviations that lie along the upper boundary of all of the portfolios on this graph, meaning that they maximize the returns for a given risk threshold. This leads to the formulae for the Capital Market Line, which represents the point at which one can most efficiently maximize their portfolio value while minimizing risk by utilizing a combination of risky and risk-free assets ("Markowitz Model", 2024).

$$R_P = \text{Portfolio Expected Returns}$$

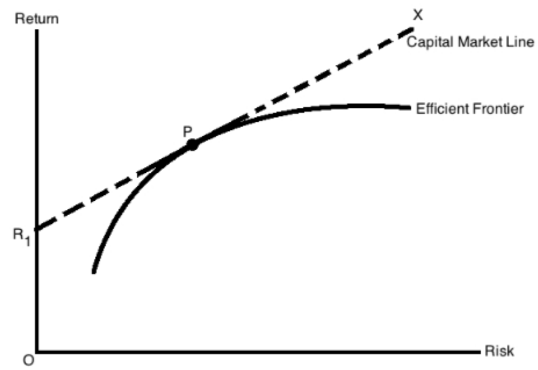
$$R_M = \text{Market Portfolio Expected Returns}$$

$$I_{RF} = \text{Risk-Free Rate}$$

$$\sigma_M = \text{Standard Deviation of Market Portfolio}$$

$$\sigma_P = \text{Standard Deviation of Portfolio}$$

$$R_P = I_{RF} + (R_M - I_{RF})\frac{\sigma_P}{\sigma_M}$$



("Markowitz Model", 2024)

Based on the above equation, $(R_M - I_{RF})$ measures the risk premium, which tells us what the investor gains from holding the risky portfolio as opposed to the risk-free portfolio. Additionally, $\left(\frac{R_M - I_{RF}}{\sigma_M}\right)$ is the slope of the Capital Market Line, which provides the reward per each unit of market risk we take on. The tangency point between the Capital Market Line and the Efficient Frontier is denoted as P, which can be considered the point at which one achieves the Market Portfolio, which is the most diversified portfolio attainable that is purely long and short in its underlying securities, with all positions adding up to 1. Moving to the right of P on the Capital Market Line requires purchasing risky assets borrowing at the risk-free rate, and to move to the left, one must incorporate a combination of risky and risk-free assets. One can note that the Capital Market Line is always sloped upwards as risk must be priced positively, where investors must be compensated for taking on more risk ("Markowitz Model", 2024).

5.3 Markowitz Implementation

We next implement Markowitz Implementation utilizing a backtesting approach that pulls from the various hedging and portfolio allocation ideas from our literature review. Specifically, we will be running Markowitz Optimization on the LETFs with the top ten highest Sharpes over rolling two-year periods and tracking their results on the subsequent years through to the present. For reference, the Markowitz portfolio optimization code is attached in the appendix. For this calculation, the following returns metrics are defined:

$$\text{Annual Returns Daily Return Mean} \times 252$$
$$\text{Annual Return Standard Deviation} = \text{Daily Standard Deviation} \times \sqrt{252}$$

Top 10 LETF Sharpes 2016 - 2018, Markowitz Performance on 2019 - Present
Tickers: [BRZU, DDM, QLD, ROM, SOXL, TECL, UBR, UDOW, UPW, USD]
Weights: [0.04, 0.17, 0.11, 0.12, 0.07, 0.08, 0.06, 0.12, 0.14, 0.09]
Annualized Returns: 0.40

Annualized Volatility: 0.53
Sharpe Ratio: 0.76

Top 10 LETF Sharpes 2017 - 2019, Markowitz Performance on 2020 - Present
Tickers: [BNKU, FBGX, HDLB, HIBL, PFFL, QLD, ROM, TECL, TQQQ, WEBL]
Weights: [0.03, 0.16, 0.11, 0.03, 0.26, 0.12, 0.10, 0.07, 0.08, 0.03]
Annualized Returns: 0.30
Annualized Volatility: 0.54
Sharpe Ratio: 0.56

Top 10 LETF Sharpes 2018 - 2020, Markowitz Performance on 2021 - Present
Tickers: [BDCX, CEFD, FNGO, FNGU, GDXU, MVRL, QLD, ROM, TQQQ, WEBL]
Weights: [-0.01, 0.58, 0.09, 0.06, 0.01, 0.05, 0.08, 0.06, 0.05, 0.02]
Annualized Returns: 0.12
Annualized Volatility: 0.34
Sharpe Ratio: 0.36

Top 10 LETF Sharpes 2019 - 2021, Markowitz Performance on 2022 - Present
Tickers: [BDCX, CEFD, FEDL, FNGU, IWDL, MJIN, QULL, SARK, SCDL, USML]
Weights: [-0.09, -0.34, 0.06, -0.05, 0.09, 0.16, 0.06, 0.29, 0.33, 0.49]
Annualized Returns: 0.11
Annualized Volatility: 0.24
Sharpe Ratio: 0.45

Top 10 LETF Sharpes 2020-2022, Markowitz Performance on 2023 - Present
Tickers: [BDCX, BERZ, DIG, GDXU, MJIN, NRGU, OILU, SARK, SCDL, YCS]
Weights: [-0.10, 0.08, 0.15, 0.01, 0.09, 0.09, 0.08, 0.18, 0.00, 0.42]
Annualized Returns: -0.18
Annualized Volatility: 0.23
Sharpe Ratio: -0.76

Top 10 LETF Sharpes 2021 - 2023, Markowitz Weights for Future Use
Tickers: [DIG, ERX, GUSH, MEXX, MJIN, NRGU, SVXY, TBX, UCO, YCS]
Weights: [0.05, 0.06, -0.04, 0.23, 0.13, 0.00, 0.17, -0.57, 0.01, 0.96]

It appears that Markowitz has not been able to effectively bolster LETF Sharpe ratios in aggregate and that this rolling backtest strategy simply follows a long-term momentum strategy of sorts. Similarly, Markowitz portfolio optimization is known to break down during periods of high volatility and under testing periods with market regimes that differ from training, which may explain the poor backtesting results from 2020 - 2022 and testing on 2023 - present, correlating with the COVID-19 period. We opt to not utilize any other non-traditional forms of portfolio optimization for the risk of overfitting on any particularly strong LETF.

6 Conclusion

As a result of our literature review, we became acquainted with the inherent disadvantages of the LETF rebalancing structure and how fees can compound to result in tracking errors. These results were confirmed when it was found that while a majority of the LETFs were able to track their underlying within a few percentage points, those tracking to a significant degree were few and far between. Similarly, in terms of long-term investing, it was discovered that **every single LETF** had a decayed Sharpe ratio when compared to its underlying. We also found no correlation between tracking ability and risk-adjusted returns and close to no correlation between tracking ability and leverage ratio. Moreover, when thinking about investing in LETFs long-term, it is difficult to predict which LETFs will end up tracking better than others, as there were few statistically significant results when measuring tracking among a variety of metrics, besides category, where debt products typically fare better. Finally, we were also able to empirically conclude that tracking ability across LETFs is typically bolstered over shorter time frames.

While we were initially intrigued by these products and the leverage they provide, this analysis has shown that capital may be better suited to be invested in the underlying index due to LETFs' decayed Sharpe and massive drawdowns. When approaching any product, it is essential to conduct due diligence across holding periods and to not be swayed by absolute returns without considering risk. It is our hope that more research continues to be done on LETFs to track performance in the coming years, over a variety of new market regimes.

6.1 Discussion of Short-Term Trading

While our paper has focused on the shortcomings of LETFs from a long-term investment perspective, these products are traditionally viewed as instruments for short-term trading. Specifically, by holding LETFs for only a few days, one can mitigate most of the negative effects of beta decay that we previously explored, which are typically felt during long-term holding. Thus, it is feasible that these products' popularity stems from their ability to allow investors to easily take on leveraged bets ahead of, during, and after high-volatility events. For instance, by analyzing the net volume traded, or flow funds, using the ETF Fund Flows Tool on ETF.com, one can observe notable flow in LETFs around events such as the COVID-19 shutdown and the Russian invasion, among other recent significant macro events. As a case study, consider the Silicon Valley Bank collapse on March 10th, 2023. TQQQ trading exhibited strong negative flow during the three months surrounding the bank's collapse with net outflows of \$1.403 billion, compared to less negative outflows in the preceding and subsequent three months. One can also note the deeply contrasting flow between QQQ and TQQQ, where it appears that TQQQ's flow may more closely resemble macro sentiment, rather than an aggregate of investors' long-term bullishness in the stock market ("ETF Flow Funds", 2024).

Flow Numbers	QQQ	TQQQ
11/1/2022 - 11/1/2023	\$5,887M	-\$2,335M
11/1/2022 - 2/1/2023	-\$2,908M	-\$716M
2/1/2023 - 5/1/2023	\$2,965M	-\$1,403M
5/1/2023 - 8/1/2023	\$6,142M	-\$966M

In terms of utility, short-term LETF payoff profiles mimic those of options, which are frequently used by retail traders. According to an SEC study, the holding period and leverage of an LETF amplify the skew of the payoff distribution, decreasing the expected payoff per volatility unit. In this case, skew refers to how in most instances, expected returns are near-zero, but in unlikely cases, there is a massive right tail, representing large, infrequent upside returns. In fact, as the leverage multiplier increases, the distribution of LETF returns becomes increasingly skewed, represented by an increased likelihood of experiencing losses from long-term investments, while the magnitude of large, unlikely gains, increases as well ("Economics Note: The Distribution of Leveraged ETF Returns", 2019). This form of distribution serves as an analogy to an out-of-the-money option, which exhibits similar skew characteristics.

In general, as discerned from the total flow, it can be noted that the popularity of short term LETF trading may be a result of investors' desire to capitalize on large market swings with quick exits. We recommend further empirical research into the short-term trading of LETFs and the difference in volume between LETF and ETF trading during periods of high volatility.

6.2 Limitations and Recommendations

While our research began with the names of 211 LETFs, we ended up not including those that were delisted at any point from their inception until the present. Additionally, while we opted not to include these tickers because they may have had functional flaws leading to their delisting, this may have inadvertently resulted in our data containing survivorship bias. Similarly, LETF/ETF combinations that possessed data less than two years in length were also removed, possibly resulting in a systematic alteration of our dataset's composition.

In the future, we recommend applying more portfolio approaches with an emphasis on avoiding overfitting, a common issue we observed in the strategies explored during our literature review. It is also recommended to utilize a different approach when testing investment viability and tracking accuracy across various periods. Finally, our research points towards there being potential for research into profitable trading strategies by looking at LETF return variability when compared to their underlying.

7 Appendix

Markowitz Mean-Variance Optimization Code

```
df = pd.DataFrame(columns = returns.columns)

def create_percent(asset_prices):
    returns = asset_prices.pct_change().iloc[1:,:]
    return returns

def mean_variance(data):
    mu = data.mean()
    Sigma = data.cov()

    # Diagonalizing
    sig_d = np.zeros(Sigma.shape)
    np.fill_diagonal(sig_d,Sigma.to_numpy().diagonal())
    Sigma = sig_d

    Sigma_inv = np.linalg.inv(Sigma)
    weights = mu @ Sigma_inv
    weights = weights / weights.sum()
    wts_tan = pd.Series(weights, index = mu.index)

    return wts_tan

def allocate_portfolio(asset_prices):
    df = asset_prices
    wts =[]
    if (len(df) > 2):
        wts = mean_variance(returns)
        wts = wts.to_numpy()
    return wts

output_weights = allocate_portfolio(returns)
output_weights
```

References

[1] Abner, David. The ETF Handbook. Wiley, 2016.

[2] Avellaneda, Marco, and Doris Dobi. "Structural Slippage of Leveraged ETFs." SSRN Electronic Journal, 2012. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2127738

[3] "Creation & Redemption." Premia Partners, 2024. https://www.premia-partners.com/education/creation-redemption

[4] Division of Economic and Risk Analysis. "Economics Note: The Distribution of Leveraged ETF Returns." November 2019.

[5] ETF.com. "ETF Fund Flows Tool." https://www.etf.com/tools/etf-fund-flows-tool

[6] George, Jeffrey, and William Trainor. "Portfolio Insurance Using Leveraged ETFs." *SSRN Electronic Journal*, 2017. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3055199

[7] Glenn, Lewis A. "Long-Term Investing in Triple Leveraged Exchange Traded Funds." *SSRN Electronic Journal*, 2020. https://doi.org/10.2139/ssrn.3722318

[8] Guedj, Ilan, Guohua Li, and Craig McCann. "Leveraged and Inverse ETFs, Holding Periods, and Investment Shortfalls." *SSRN Electronic Journal*, 2010. https://www.researchgate.net/publication/270124896_Leveraged_and_Inverse_ETFs_Holding_Periods_and_Investment_Shortfalls

[9] "List of Leveraged ETFs." *Stock Market MBA*, 2023. https://stockmarketmba.com/listofleveragedetfs.php

[10] Lu, Lei, Jun Wang, and Ge Zhang. "Long Term Performance of Leveraged ETFs." *SSRN Electronic Journal*, 2009. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1344133

[11] Margenot, Maxwell, and Delaney Granizo-Mackenzie. "Position Concentration Risk." *Quantopian Research*, 2014. https://github.com/quantopian/research_public/blob/master/notebooks/lectures/Position_Concentration_Risk/notebook.ipynb

[12] "Markowitz Model." *Wikipedia*, 2024. https://en.wikipedia.org/wiki/Markowitz_model

[13] "Sharpe Ratio." *Investopedia*, 2024. https://www.investopedia.com/terms/s/sharperatio.asp

[14] "The Basics of Mutual Funds." *Investopedia*, 2024. https://www.investopedia.com/articles/mutualfund/05/062705.asp

[15] Team, Putnam Investments. "S&P 500 Index Earnings and P/E Multiple Performance During U.S. Recessions." *Putnam Investment*, 2024. https://www.putnam.com/dcio/content/perspectives/9535-sp-500-index-earnings-and-p-e-multiple-performance-during-u-s-recessions

[16] Team, The Investopedia. "Are Triple Leveraged ETFs a Good Idea?" *Investopedia*, September 15, 2022. https://www.investopedia.com/stock-analysis/2012/are-triple-leveraged-etfs-a-good-idea-fas-faz-tyh-typ0430.aspx

[17] "What Is a Leveraged ETF? Everything You Need to Know." *ETF.com*, 2024. https://www.etf.com/topics/leveraged

[18] Yates, Tristan. "The Dangers of Leveraged ETFs." *Investopedia*, 2024. https://www.investopedia.com/articles/exchangetradedfunds/07/leveraged-etf.asp

[19] Horstmeyer, Derek. "Leveraged ETFs: Are They Right for You?" The Wall Street Journal, 2023. https://www.wsj.com/finance/investing/leveraged-etfs-risk-cb95693f.