LUT UNIVERSITY

School of Business and Management

Strategic Finance and Business Analytics

Master’s thesis

Leveraged Exchange-Traded Funds: Expecting K-times Returns

Teemu Huotari

2019

1st examiner: Mikael Collan

2nd examiner: Azzurra Morreale

TIIVISTELMÄ

Tekijä:	Teemu Huotari
Otsikko:	Vivutetut ETF-rahastot: Olettama k-kertaisista tuotoista
Tiedekunta:	LUT School of Business and Management
Maisteriohjelma:	Strategic Finance and Business Analytics
Vuosi:	2019
Pro Gradu:	94 sivua, 13 taulukkoa, 32 kuvaajaa
Tarkastajat:	Professori Mikael Collan
	Tutkijatohtori Azzurra Morreale
Hakusanat:	vivutettu ETF, tuotto, riskikorjattu tuotto, k-kertaiset tuotot, naiivi oletus, pitkän aikavälin suoriutuminen, regressioanalyysi

Tämä Pro Gradu -tutkielma tarkastelee vivutettujen ETF-rahastojen riskiä ja tuottoa, testaamalla kuinka ne pitäytyvät oletetuissa k-kertaisissa tuotoissa eri sijoitusajanjaksoilla. Empiirinen analyysi on tehty tutkimalla 22 eri vivutetun ETF-rahaston tuottoja, verraten niitä niiden vivuttamattomiin vastinpareihin. Riskin ja tuoton suhdetta on tutkittu regressioanalyyseilla ja validoitu edelleen t-testien avulla. Lisäksi volatiliteetin vaikutusta on tutkittu korrelaatiotesteillä. Tämän työn tulokset vahvistavat aiempia löydöksiä ja teorioita siitä, että sijoittaja pystyy suhteellisen turvallisesti odottamaan k-kertoimen mukaisia tuottoja, enintään yhden kuukauden pituisilla sijoitushorisonteilla. Tulokset osoittavat, että tämä on vielä todennäköisempää positiivisesti vivutettujen ETF-rahastojen, kuin negatiivisesti vivutettujen tapauksessa. Lisäksi tulokset osoittavat, että 2x-vivutetut rahastot voivat tuottaa k-kertaisia tuottoja luotettavasti jopa kolmen kuukauden sijoitusajanjaksoilla, toisin kuin negatiivisesti vivutetut rahastot. Vivutettujen ETF-rahastojen keskeiset teoriat osoittavat, että volatiliteetti aiheuttaisi tuottojen heikentymisen pitkillä ajanjaksoilla. Vaikka tämän tutkimuksen tulokset osoittavat, että volatiliteetilla on negatiivinen korrelaatio rahaston tuottojen suhteen, kertyneiden tuottojen ja polkuriippuvuuden positiiviset vaikutukset herättivät huomiota erittäin pitkillä sijoitushorisonteilla, herättäen jatkokysymyksiä vivutettujen ETF-rahastojen sopivuudesta erittäin pitkän sijoitushorisontin instrumentiksi.

ABSTRACT

Author:	Teemu Huotari
Title:	Leveraged Exchange-Traded Funds: Expecting K-times Returns
Faculty:	LUT School of Business and Management
Master's Program:	Strategic Finance and Business Analytics
Year:	2019
Master's Thesis:	94 pages, 13 tables, 32 figures
Examiners:	Professor Mikael Collan
	Post-Doctoral Researcher Azzurra Morreale
Keywords:	leveraged ETF, returns, risk-adjusted returns, k-times returns, naïve expectation, long-term performance, regression analysis

This Master's thesis focuses on risk and return of leveraged exchange-traded funds by testing how they hold on to their expected k-multiplier returns on different investment periods. Empirical analysis is done by studying returns of 22 leveraged and inverse leveraged ETFs against their conventional benchmark ETFs, from last seven years. Risk-return relationship is tested with regression analyses, and further validated with t-tests. In addition, effect of volatility is tested with correlation tests. Results of this thesis strengthen previous findings and theories of an investor being able to relatively safely expect returns denoted by the leverage-multiplier, on investment periods up to 1-month. Results suggest that this is even more likely to occur on positively leveraged ETFs than when investing in inverse LETFs. In addition, results suggest that 2x-leveraged ETFs could provide k-times returns even on 3-month investment periods, unlike inverse LETFs. Theories behind LETFs suggest that volatility causes LETF returns to decay on long investment periods. Although results of this study also suggest that volatility has negative correlation to LETF returns, positive effects of compounding and path-dependence of LETFs are noted on very long time periods, awaking further questions about performance on ultra-long investment periods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Exchange traded funds (abbr. ETFs), having only been around in their current form from the beginning of 90's, are a very popular investment vehicle. ETFs, most often designed to track a certain market index, can provide an easy and cost-efficient diversification for a retail investor. Most important thing separating ETFs from mutual funds is their tradability. As the name suggests, unlike mutual funds, exchange traded funds are tradeable between investors in stock-exchange. Since the introduction of first ETFs in 1993, by July 2019 global ETF assets had reached total of $5.58 trillion (ETF Guide 2019, 6).

This thesis focuses on special type of ETFs, leveraged ETFs. Being fairly new investment vehicle, LETFs and issues related to their returns have however been studied a lot recently. Daily rebalancing, constant leverage trap and unwanted effects of volatility has made leveraged ETFs rather controversial investment vehicles and motivated academic world to study them, since they were first introduced in 2006. Controversy around these instruments has further risen by investors, as there has seemingly been misunderstanding about how these instruments actually work.

This thesis focuses on the risk and returns of leveraged ETFs, with an empirical section to test, how well leveraged ETFs can provide returns suggested by their leverage multiplier, on different investment periods. Although same type of research has been done in the academic field of leveraged ETFs, this thesis applies these methodologies to the latest 7-year period, including leveraged ETFs with different amount of leverages, and not only limiting to equity ETFs, but also including fixed-income ETFs to the study. Arising from the misunderstanding among investors, in this thesis returns of leveraged ETFs are studied in different holding periods to test, how well they hold on to the returns suggested by the leverage-multiplier, against their non-leveraged ETF equivalents, serving as benchmarks.

After this introductory chapter, focusing on theoretical framework, motivation, objectives and research questions, is provided a comprehensive literature review. Literature review starts with introduction to leveraged exchange-traded funds as an instrument. This is then followed by presenting problems and theories suggested by previous literature and academic research of the subject. In addition to focusing on the returns, literature review discusses the most important aspects of LETFs studied in the academic world. After literature review, dataset for the empirical section of this thesis, as well as the criteria for choosing the data, are presented. Chapter 4 presents the methodology, followed by chapter 5 presenting the empirical results from the tests presented in methodology chapter. Final chapter summarizes the results, discusses conclusions and limitations of the thesis, also providing ideas for further research.

1.1. Theoretical framework and motivation of study

Where conventional ETFs are often designed to track a particular index, leveraged ETFs target to capture the daily changes of the underlying index, with certain multiplier. For example, ProShares Ultra S&P500 (SSO) has a daily target of doubling the change of the underlying index S&P500: If underlying index S&P500 would change +0.5% in a single day, should leveraged SSO have a change of +1.0% at the end of that day.

In addition to these positively leveraged bullish ETFs, inverse, or negatively leveraged, bearish ETFs have been included into this research. Like positively leveraged ETFs, these inverse ETFs also have a target index to track, but the multiplier is negative, leading the price of the ETF to move in the opposite direction than the underlying index. In addition to equity ETFs, which track market indices, fixed-income ETFs are also included to this study, as they seem to be even scarcer in the academic research. Fixed-income leveraged ETFs are a very niche type of ETFs, and there is not much test subjects and data available. However, the motivation to add these fixed-income, or bond LETFs to this study, is based on the increased popularity among investing in fixed income. In addition, for example in Finland, direct investing in fixed-income can be somehow limited and bond ETFs seem to be rather convenient way to get fixed-income instruments in a portfolio.

Leveraged ETFs, especially their risk and return have been researched before in academic literature. However, regarding master theses from Finland, there seems to be an absence of thesis focusing mainly on leveraged ETFs, no matter that conventional, non-leveraged ETFs seem to be a popular subject among thesis authors.

The concept of risk and return of leveraged ETFs is very interesting from a point of an investor. If assuming that the daily returns of leveraged ETFs matching their targets on daily basis would imply same kind of double or triple performance also on long periods, investors could get rapid gains or losses with lesser capital than what would be needed with conventional ETFs. When adding the factor of risk to this assumption, an investor could double or triple their desired returns, by doubling or tripling their risk, using these instruments.

Considering these assumptions, they actually hold on 1-day investment periods. If not taking account of tracking error (in other words the fund not being able to track its benchmark index), for 1-day periods an investor can actually reliably seek for double or triple profits, with the same ratio of risk than with using conventional ETFs. This is simply because of how the instruments are built. By doubling the risk, one would end up double the profits. This is the base for studying the returns matching the leverage-multiplier returns on longer investment periods, because of being able to do so, the risk-ratio would stay static.

1.2. Objectives and research questions

Objectives of this thesis is to research risk and return of leveraged ETFs. This is done by studying and testing them against their non-leveraged counterparties, referred as conventional ETFs. Considering the articles discussed in the literature review, studies of leveraged ETFs can roughly be divided in three groups. First type of studies are empirical studies based on real data. Second type is simulation-based studies. As leveraged ETFs have been around for only a short period of time, especially the first studies were done by simulating the data in purpose of getting a longer span of observations. Thirdly, returns of LETFs have been studied in purely theoretical way, by observing the returns and for example effects of volatility in long periods with mathematical approach.

In the aspect of LETF returns, the object of this thesis is to provide latest empirical evidence from a selection of popular leveraged ETFs and reflect the results from previous research in the light of these results. In theory, leveraged ETFs can not reach their k-times returns (k being the leverage-multiplier) on a long term, as explained in the literature review of this thesis.

First research question of this thesis is approached from the aspect of k-times returns and is defined as:

Can leveraged ETFs provide k-times returns on longer than 1-day holding periods?

H0: K-times returns are provided on a given holding period

Considering daily returns of a leveraged ETF with a k-multiplier of 2x, that LETF would provide double the returns and risk on a 1-day investment period, compared to its non-leveraged benchmark ETF, assuming no tracking error on 1-day basis. With this assumption, risk-return-ratios of these leveraged and non-leveraged ETFs are the same. The null-hypothesis is based on this idea, that risk-return-ratio would stay static and then k-multiplier returns would also reliably occur on longer investment periods.

If the question for this answer is positive, it awakens further questions, such as for how long investment periods the k-multiplied returns can be matched, and is there difference between the groups of LETFs, such as positively leveraged and inverse leveraged ETFs, or equity LETFs and fixed-income LETFs. To assess the development and behavior of returns in different investment periods, second research question is defined as:

Are there significant differences in the behavior of returns between groups of LETFs?

H0: There are no significant differences in behavior of returns between groups

This question is based on the idea, that there would be recognizable characteristics in the returns of LETFs, so conclusions and distinction between groups could be made. For example, based on the theories presented in the literature review, it is expected that different behavior between for example positively and inverse leveraged ETFs will be observed.

Although risk-return ratio is assessed within the first research question, the context can be approached from another perspective. With additional methodology, presented in chapter 4.3. I assess risk-return ratio by directly comparing risk-adjusted returns of LETFs and their non-leveraged benchmarks. Third research question is then defined as:

Do leveraged ETFs provide risk and return in the same ratio as their non-leveraged counterparties?

H0: Risk-adjusted returns between leveraged and non-leveraged ETFs are the same

Based on the idea that if k-times returns would be achieved on any given period, the risk-return ratio would stay static, the third research question is fundamentally similar to the first research question, as both answer to the LETFs ability keep the risk-return ratio static. Third question is however separated to its own, to define additional null-hypothesis which also backs up the results achieved regarding the first two questions, as will be seen in results and conclusions of this thesis.

2. LITERATURE REVIEW

This chapter presents the concepts handled in this thesis and makes an extensive look into the previous academic research of this topic. After a brief of leveraged ETFs as an instrument current state of research of LETFs is presented. Lastly, equations of 2x-leveraged ETF returns are presented, which is then applied to 3x-leveraged ETFs, to show how 3x-leveraged ETFs should be more prone to deviate from the expected k-times returns.

Majority of research articles presented in this chapter were discovered through LUT Finna's international e-materials search. Entries like "leveraged ETF", "leveraged ETF performance" and "leveraged ETF returns" provided relevant articles and already at very early stage, the most important papers concerning this subject were recognized. Similar results were achieved from Google Scholar -search, providing much of the same articles. Considering the amount of cross-referencing between articles, it was easy to recognize the pioneering studies of this field. In addition to the most relevant and referenced articles, publications from a number of financial journals were included, as LETF returns has also been discussed a lot outside the actual research papers.

2.1. Leveraged Exchange-Traded Funds

Leveraged ETFs (or LETFs), being fairly new instruments, were introduced first in 2006, when ProShares introduced six products with both bullish and bearish features. Two years later, Direxion introduced the first 3x leveraged ETFs. (ETFdb.com 2019a). In 2019, total leveraged ETF assets in the United States markets are $36.74 billion (etf.com 2019), which is only a fraction of all ETF assets in U.S. markets, being at around $4 trillion (ETF Guide 2019, 6). It is safe to say that leveraged ETFs are still a very niche type of ETF and their behavior, risk and return have been in great discussion among academic and investment world.

As explained earlier, leveraged ETFs are structured to track an underlying index with certain multiplier, usually 2x and 3x, or negative -2x and -3x. Leveraged ETFs are constructed by using debt and/or derivatives to generate the desired performance related to the benchmark index (The Motley Fool 2017). Based on the fact, that leveraged ETFs are designed to track the daily variance, rather than long-term returns of the underlying index, consensus among investors is that they are suitable only for short-term investing.

2.2. Returns of Leveraged ETFs

This chapter goes through previous research of LETF returns and discusses the theory and previous empirical findings behind them. Observing the theories behind LETF returns gives us a good basis to understand, why it makes sense to compare these LETF returns to the returns of their conventional counterparties.

2.2.1. K-Multiplier

As leveraged ETFs most often track some underlying index, they are set to replicate the daily variance of that particular index with a certain multiplier. For example, ProShares Ultra S&P500 (SSO), is leveraged to seek double the daily changes of underlying S&P500 index. In this case, we say that the k-multiplier of SSO is 2. In the case of inverse ETFs, for example ProShares UltraShort S&P500 (SDS), seeking to multiply daily changes to opposing direction, the multiplier is considered as -2. As with practically every leveraged ETF, the k-multiplier only represents the daily target variance in relation to its underlying index and does not indicate anything about returns on periods any longer than that single day.

The empirical section of this thesis, as lot of previous studies of this topic is based on the idea that a LETF could provide k-times returns on even longer than 1-day investment periods. If considering risk as daily volatility of a single ETF, its risk and return would always match its non-leveraged equivalent on 1-day investment period, if we do not take account on daily tracking errors. Tracking errors are shortly observed at the beginning of the empirical results of this thesis and based on that are assumed to be minimal. Cheng et al. (2016) studied tracking errors of LETFs on oil sector and found out that generally tracking errors are low and they correct themselves in matter of days. In the field of commodity-based LETFs, Murphy et al. (2010) have also empirically tested that tracking errors on 1-day basis are low, as 1-day returns did not statistically differ from expected 1-day returns. In this thesis, the expectation of k-times returns occurring on even longer than 1-day investment periods, providing risk and return in same relationship, is referred as naïve expectation of k-times returns.

2.2.2. K-times returns

Previous research has shown, that on a long-term, leveraged ETFs will not provide k-times the profit, k being the multiplier of the tracked index. Based on Lauricella's (2009) article, it seems that especially in the early days of leveraged ETFs, there seemed to be lot of misconception among the investors, regarding the presented multiplier. The article focuses on theme of reading the fine print on investment products, but also Lu et al. (2012) agree, that there is misperception among the investors about the long-term profits, referring to Lauricella's article as one evidence.

Still in 2014, 8 years after the launch of first leveraged ETFs, brokerages like Merrill Lynch and Morgan Stanley refused their advisers to trade leveraged ETFs on their platforms, because of concerns that individual investors wouldn't acknowledge the risks. Brad Stratton, an independent advisor interviewed in the article states, that the way the multiplier affects long-term returns is "still the biggest part the retail investor doesn't understand". (Lau 2014). It seems that the index-multiplier presented in the name of the is not very clear among retail-investors as they don't seem to be aware that it only guarantees the corresponding multiplying of daily, not long-term returns.

Lu et al. (2012) studied leveraged (and inverse) ETFs to find out, how they profit on different, longer than 1-day holding periods. They chose 4 ETFs tracking major indices as benchmark funds, and for each of these funds, they tested the performance of a 2x-leveraged and a -2x-inversed ETF. The results show that these ETFs leveraged to double (or inverse double) the tracked index, could provide the double profit up to one month of holding period. Although leveraged ETFs are set to reach the k-times returns only on a daily basis, the results suggest that an investor is not restricted to day-trading and can expect descent k-times tracking even on longer periods.

In addition to empirical studies based on historical or simulated data, there are studies that take on the risks and returns in theoretical way. Issues of the multiplier of leveraged ETF not directly transferring to profits, nor losses, is studied by Jarrow (2010) and seemingly one of the most cited articles in the field of LETFs, Cheng et al. (2009).

Zweig (2009) provides couple of examples, on how leveraged ETFs do not meet their k-multiplier returns, neither on positive, nor on inverse side. They provide evidence from the market, that on November 2008, there was a 17-day period, which during the Russell 1000 index had a notable decline of 25.6%. If referring to the common misconception of k-multiplier applying to any given period, Direxion Large Cap Bear 3X, an inverse fund tracking the Russell 1000 index, should have risen 3-times the realized price change, 76.8%. However, during that time, the inverse LETF went up over 4-times the corresponding index, 109.2%.

Zweig (2009) provides also an opposing example to this. On October of that same year, observing a 9-day period of trading, Chinese stock market fell a cumulative of 8%. If observing this situation with the misconception from previous example, a 2X inverse fund, ProShares UltraShort FTSE/Xinhua China, should have provided a 16% profit, inverse-doubling the 8% decline of the index. However, the realized profit of that LETF from the period was negative, -21.6%. Interestingly, Bansal & Marshall (2015) found out on their simulation-based research of 2x- and 3x-leveraged ETFs, that the tracking error of 3x-leveraged ETFs on long investment periods may actually be favorable, meaning that they would actually provide more return than the naïve expectation suggests. Based on the idea of overall economic growth on long-periods, and as seen in the data section of this thesis, this seems to be entirely reasonable. Very long investment periods of 3 and 6 years are also positively noticed by Širůček et al. (2018), as they studied 2x- and 3x-leveraged ETFs on 1-, 3-, and 6-year investment periods. Guedj et al. (2010) doubt investing on LETFs on any long periods, as they find out that daily rebalancing causing shortfall leads to some LETF investors to lose up to 3% of initial investment in just a few weeks.

These examples present, that although not effectively meeting k-times returns of the index, LETFs can also perform better on longer periods, than the multiplier implies. The key to this phenomenon is in volatility; LETF returns on long periods follow path-dependency. To better understand this concept, we need to observe the nature of profits and leveraging of these instruments.

2.2.3. Constant leverage trap and path-dependence

The nature of leveraged ETFs, and them not meeting the k-multiplier returns on longer periods, can be explained by observing the returns, their volatility and leveraging of these instruments.

As a concept, although having been around in the field of financial research before, the "constant leverage trap" got popular when first leveraged ETFs were introduced. In summary, the idea behind this concept is that maintaining a constant leverage ratio, by adjusting the leverage of portfolio to a certain degree at the end of every trading day, will cause the returns of assets to decay on a long investment period. (Yates 2007)

The problem of holding a constant leverage on a portfolio, applies the same way on leveraged ETFs. At the beginning of a trading day a portfolio, in this case LETF, holds equal proportion of equity and debt. When market goes up, the portfolio holds more equity than debt. If this is the case at the end of a trading day, the portfolio is rebalanced by buying more shares and vice versa. This daily rebalancing resembles a lot of "buy high-, sell low-strategy, which seems to be a bad investing strategy. (Yates 2007). However, the reason for rebalancing of leveraged ETFs is not something that is done to maximize profits. The reason is purely behind being able to meet the goal of following the daily variance of the underlying index, with a certain multiplier.

Previous studies have suggested that the profits of LETFs are tied to path-dependence. Liebowitz et al. (1999) define path-dependence generally as: "Where we go next depends not only on where we are now, but also upon where we have been." This seems to apply very well to many financial contexts, at least on some level, as future, or even current state of stocks or interest rates can hardly ever be estimated purely referring only to a certain equilibrium, rather than historical data. As Liebowitz et al. (1999) state, path-dependence battles against economic models that derive an equilibrium, as they do not take any historical information of the subject in account.

This all comes together, explaining the k-multiplier not applying on longer than 1-day holding periods. Effects of constant leverage trap and path-dependence can be illustrated with an example (Militaru et al. 2010) below (figure 1). These scenarios of three imaginary ETFs (1x, 2x, and 3x) show how the returns of initial investment of 100 behaves within these instruments, when for each step of time, the value of investment changes by 10 percentage on either way.

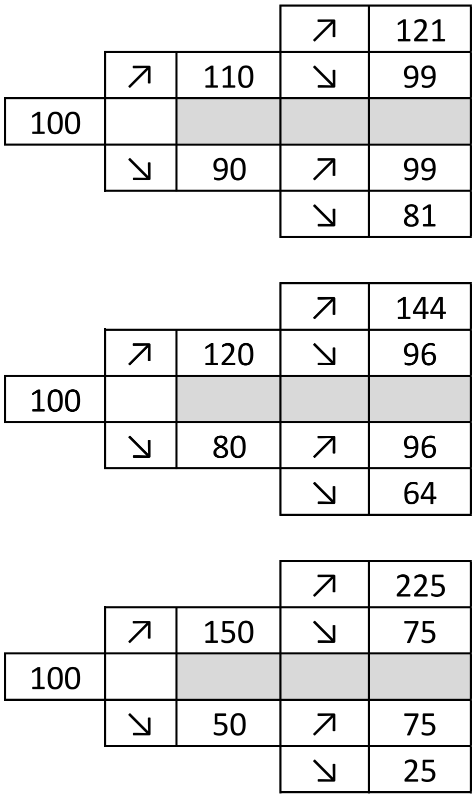

Figure 1: Binomial trees of 3 imaginary ETFs, after Militaru et al. (2010)

These trees of different scenarios with differently leveraged ETFs shows first the basic principal of compounding returns. Going from point A to B, and back to A, with 10% relative changes, will not return on the level we started off on *t+2*. When increasing daily variance, by entering a position on more leveraged instruments, the median diverges further away from the mean. The longer the holding period, and higher the leverage multiplier on the instrument, the stronger the effect is. The average return of these three differently leveraged ETFs on *t+2* is the same, 100. However, as the multiplier rises, the median return decreases. We can observe the scenarios and see, that with each of these instruments, only one of *t*+2 outcomes end up higher than the average. As the leverage multiplier increases, the median of *t+2* scenarios decrease (99, 96, 75).

We can say, that compounding returns are rather lognormally, than normally distributed. As we can see from the binomial trees, the average returns being higher than the median returns, the distribution is positively skewed. Trainor et al. (2008) researched leveraged ETFs with a Monte Carlo simulation model. They suggest that the lognormality of returns is a major reason for LETFs not reaching the k-times returns. Impact of compounding returns and path-dependence of LETF returns is also recognized by Avellaneda et al. (2010a), as they also studied the problem of k-multiplied returns on long investment periods. Furthermore, considering the volatility of returns, Hunter et al. (2013) suggest that volatility has more negative effect on inverse bear LETFs, than on positively leveraged LETFs. As well as several other studies, Trainor et al. (2013) present the decaying effect of volatility on LETF returns, but state that on times of low volatility, the decay could even be outweighed by significant trend. In my opinion, this backs up the idea of path-dependence of the returns, and possibility to gain good returns in low volatility environments. Lastly, Rompotis (2014) discusses leveraged and inverse leveraged ETFs in the light of previous studies and concludes that market volatility is the most important risk to affect returns of leveraged ETFs. In addition to volatility, also the length of the investment period is recognized as generally negative factor affecting the LETFs ability to track its index. Loehr et al. (2013) compare investing in leveraged ETFs to playing at a casino. The longer one would play (invest in LETFs), the more the odds would move in favor of the house.

Some research suggest that there is a significant impact from trading and rebalancing activities of an LETF to volatility of the underlying index. Chen et al. (2012) support this idea as they studied spillover and asymmetric-volatility effects of LETFs with EGARCH-M-ARMA model. Similarly, Shum et al. (2016) found out in their empirical research that rebalancing activities increased end-of-day volatility, while also suggesting that leveraged ETFs have larger tracking errors during volatile periods. Although not directly related to empirical section of this thesis, effects of rebalancing on volatility is important and popular subject among research of leveraged ETFs. Similar findings in this context have been presented by Bai et al. (2012), stating that act of LETF rebalancing significantly effects the price and increases the volatility of underlying stocks. Interestingly, Trainor (2010) could not find LETF rebalancing actions significantly affecting volatility of underlying S&P500 index, finding it rather coincidental than systematic.

Investment strategies for leveraged ETF in attempt to overcome the volatility decay have also been researched. For example, in the field of commodity LETFs, Guo et al. (2015) tested a common strategy where investor shorts a positive-inverse pair of LETFs. They found out that in times of high volatility, this strategy profits on average, but the returns are prone to suffer from periods of low volatility of the benchmark index. Hessel et al. (2018) came out with similar results, as they agreed that the returns on strategy of shorting LETF-IETF pairs is not dependent on the direction of the underlying benchmark, but rather benefits from the high volatility of underlying benchmark index. Co (2009) suggests that for long-investment periods, using only LETFs to employ leverage is not necessary the best strategy. Barnhorst et al. (2011) suggest that using exchange-traded notes ETNs in combination with future contracts or options would provide an alternative to employ leverage, but with lesser exposure to volatility than when investing purely on LETFs.

Although stating that LETFs have high probability of getting close to k-times returns on short investment periods with low volatility, Hill et al. (2009a) present a strategy to overcome constant leverage trap, by rebalancing the LETF portfolios in attempt to get even closer to k-times index returns. In addition, the strategy of rebalancing is considered to be most beneficial in times of high volatility (Hill et al. 2009b). The issue of rebalancing is also acknowledged by LETF providers, which have offered monthly leveraged funds (Trainor 2011)

Although lognormality and nature of continuous compounding suggest that on a long investment period the volatility would decay the returns, there is an opposing argument. As in the example presented by Zweig (2009) and binomial tree presented after Militaru et al. (2010), there is a way to make profit, and even beat the index over k-times, using leveraged ETFs. As these examples and theory suggests, it would demand a highly positive trend, consecutive days of gaining, with minimal volatility. This is why path-dependence and volatility are important concepts within LETFs, as we need to understand, that succeeding with them on a long term, would require constant good performance. With precise timing and correct length of holding period could provide good return on these instruments. To be more specific, investor should hit a holding period with a good ascending trend and minimal volatility.

2.3. Mathematical interpretation of LETF returns

In the same article as mentioned earlier, in addition to testing long-term performance of LETFs in an empirical way, Lu et al. (2012) provide a generally applicable model to express the behavior of LETF returns on different time periods. This mathematical interpretation of LETF returns explains, how longer holding periods should decrease the expected returns, and increase deviation. It shows the behavior of returns on two different ETFs and a benchmark index: benchmark index (R_{tn}^{B}), 2x leveraged ETF (R_{tn}^{D}) and -2x inverse ETF (R_{tn}^{I}). 2.3.1. presents the original equations by Lu et al. (2012), and 2.3.2. presents my own idea of how it adapts to 3x- and -3x-leveraged ETFs.

2.3.1. 2-day returns of leveraged ETFs

For these example ETFs, R_{tn} represents a n-day cumulative return starting at date t, which for the benchmark index and two ETFs can be represented as:

$$R_{tn}^{B} = \prod_{i=0}^{n-1} (1 + r_{t+1}^{B}) - 1 \quad (1)$$

$$R_{tn}^{D} = \prod_{i=0}^{n-1} (1 + r_{t+1}^{D}) - 1 \quad (2)$$

$$R_{tn}^{I} = \prod_{i=0}^{n-1} (1 + r_{t+1}^{I}) - 1 \quad (3)$$

Returns of leveraged ETFs in relation to the benchmark index can be expressed by presenting 2-day LETF returns as a function of returns of the benchmark index. For a two-day period, n being 2,

$$R_{t2}^{B} = (1 + r_t^{B})(1 + r_{t+1}^{B}) - 1 = r_t^{B} + r_{t+1}^{B} + r_t^{B} r_{t+1}^{B} \quad (4)$$

$$R_{t2}^{D} = (1 + 2r_t^{B})(1 + 2r_{t+1}^{B}) - 1 = 2r_t^{B} + 2r_{t+1}^{B} + 4r_t^{B} r_{t+1}^{B} \quad (5)$$

$$R_{t2}^{I} = (1 - 2r_t^{B})(1 - 2r_{t+1}^{B}) - 1 = -2r_t^{B} - 2r_{t+1}^{B} + 4r_t^{B} r_{t+1}^{B} \quad (6)$$

Double and inverse double ETFs can be presented in terms of $\boldsymbol{R_{t2}^{B}}$:

$$R_{t2}^{D} = 2R_{t2}^{B} + 2r_t^{B} r_{t+1}^{B} \quad (7)$$

$$R_{t2}^{I} = -2R_{t2}^{B} + 6r_t^{B} r_{t+1}^{B} \quad (8)$$

These last two equations provide a mathematical explanation to the problem occurring from the nature of compounding returns, when considering long-term returns of leveraged ETFs. If these ETFs would provide double the cumulative returns in contrast to the benchmark index, the equations would simply show as:

$$R_{t2}^{D} = 2R_{t2}^{B} \quad (9)$$

$$R_{t2}^{I} = 2R_{t2}^{B} \quad (10)$$

The main thing to consider here is the additional cross-product of $\boldsymbol{r_t^{B}}$ and $\boldsymbol{r_{t+1}^{B}}$, which deviates the expected returns from the equations above. As we can see from equation 8, the multiplier of the cross-product is even higher on the inverse ETF (*6*), than on the positively leveraged ETF (*2*), on equation 7. In addition, the multiplier of cross-product grows drastically, when the k-multiplier of the instrument rises to higher levels. In the following subchapter I have applied the Lu et al. (2012) equations of LETF returns to 3x leveraged and -3x inverse ETFs.

2.3.2. Application to 3x leveraged ETFs

Just as in previous chapter, let $\boldsymbol{R_{tn}}$ represent a n-day cumulative return starting at date $\boldsymbol{t}$. Now let's consider two additional ETFs, ***U*** with a positive leverage of 3x, and an inverse ***W***, with a multiplier of -3x. When applied to Lu et al. (2012) model:

$$R_{tn}^{U} = \prod_{i=0}^{n-1}(1 + r_{t+1}^{U}) - 1 \quad (11)$$

$$R_{tn}^{W} = \prod_{i=0}^{n-1}(1 + r_{t+1}^{W}) - 1 \quad (12)$$

Observing the 2-day cumulative returns:

$$R_{t2}^{U} = (1 + 3r_t^B)(1 + 3r_{t+1}^B) - 1 = 3r_t^B + 3r_{t+1}^B + 9r_t^B r_{t+1}^B \quad (13)$$

$$R_{t2}^{W} = (1 - 3r_t^B)(1 - 3r_{t+1}^B) - 1 = -3r_t^B - 3r_{t+1}^B + 9r_t^B r_{t+1}^B \quad (14)$$

And finally, in terms of $\boldsymbol{R_{t2}^{B}}$:

$$R_{t2}^{U} = 3R_{t2}^{B} + 6r_t^B r_{t+1}^B \quad (15)$$

$$R_{t2}^{W} = -3R_{t2}^{B} + 12r_t^B r_{t+1}^B \quad (16)$$

As we can see, there is dramatic increase in the multiplier of cross-product for both ETFs, when the leverage-multiplier was increased. For the positively leveraged ETF, the change of leveraging from 2 to 3 caused the cross-product multiplier to rise from 2 to 6 (equation 15). The respective change of the negative multiplier on inverse ETF caused the cross-product multiplier to change from 6 to 12 (equation 16). These equations would back up findings of previous studies like Lu et al. (2012) and Hunter et al. (2013), about volatility affecting inverse LETF more heavily than positively leveraged ETFs.

As the cross-product multiplier represents deviation from the naïve expectation of k-multiplier returns on long investment periods, its dramatical increasing when leverage-multiplier rises, is in great importance. Based on this mathematical interpretation, 3x-leveraged ETFs are supposed to be even more prone to volatile markets, and therefore probably even worse suitable for long-term investing, than 2x-leveraged ETFs.

3. DATA

Dataset of this study consists of daily returns of 30 selected LETFs. this chapter presents the selection criteria for the data and information about the selected data. In addition, compounding returns through the 7-year period of data are calculated and plotted to assess the overall performance of ETFs chosen for this study.

3.1. Datasets

For the empirical part of this research, a selection of leveraged and inverse ETFs is made, in addition to their conventional equivalents, which serve as benchmarks. Because of LETFs having been around for relatively short time, simulations based on the performance of underlying index has often been used in previous studies (for example Jiang et al. (2017)), to create a longer timeframe of data. This thesis however uses actual daily prices of the chosen LETFs. The dataset includes a total 30 ETFs and the daily prices are exported from Yahoo! Finance, using the daily adjusted closing prices. Count of different types of ETFs selected for this thesis can be seen in the following table 1.

Table 1: Count of ETFs grouped by type

	Positively leveraged	Inverse leveraged	Non-leveraged (benchmark)
Equity	8	8	5
Fixed-Income	3	3	3

3.1.1. Selection of equity ETFs

Selecting the equity ETFs for the study focused on three main criteria: size, popularity and availability. Size refers to the amount of ETF's total assets. Size would directly imply popularity, but in this criteria popularity refers to the relevance of the underlying index.

Considering the relevance and popularity of the underlying index, the available ETFs are narrowed down to fulfill the third criteria, availability. Criteria of availability refers to choosing the type of LETFs, that would also have an inverse pair available.

For some of the major indices like S&P500, and NASDAQ-100, both 2x and 3x leveraged ETFs are included, besides their inverse pairs. All the available equity LETFs were observed and chosen from a list consisting of 164 equity LETFs, from etfdb.com (ETFdb.com 2019b). Although some of the largest equity LETFs by total assets were exchange-traded notes (ETN), they were deliberately left out of the selection, to keep the variety of dataset as coherent as possible.

Observing the dataset, it has to be noticed that all of the selected leveraged and inverse ETFs are issued and managed by two companies, ProShares and Direxion. This is not intentional, and purely a result of making the selection using the three criteria defined. It should be noted that these two providers are the clearly the largest providers of leveraged and inverse ETFs as other providers seem to offer most of their leveraged vehicles as ETNs (ETFdb.com 2019b).

3.1.2. Selection of fixed-income ETFs

Three criteria of making the selection of equity ETFs could not be used with fixed-income ETFs. Reason for that is simply the small variety of leveraged fixed-income ETFs in the market. Although the count and variety of conventional fixed-income ETFs is quite large, 404 different products (ETFdb.com 2019c), the corresponding list for leveraged and inverse fixed-income ETFs is minuscule, 11 entries (ETFdb.com 2019d). In addition to this list of 11 entries, other sources were observed, but no other ETFs than those listed on ETFdb.com could be found.

From these 11 available fixed-income LETFs, 6 were chosen in addition to 3 benchmark fixed-income ETFs. Criteria for this selection was simply the availability of data in addition to availability to the benchmarking conventional ETF. Conciseness of this list encourage to research other sources for additional entries for the dataset, but suitable LETFs were nowhere to be found. There are few other fixed-income LETFs available, but the for the purpose of this study they were not suitable as the problem arises with finding the benchmarking conventional ETFs. For example, Direxion Daily 20-Year Treasury Bull 3X (TMF) tracks NYSE 20 Year Plus Treasury Bond Index, with a multiplier of 3. If we were to benchmark the particular LETF to the index, it could be chosen. However, as any conventional ETF tracking that index does not seem to exist, the LETF needs to be dropped out, as there would be no benchmark for it.

3.2. Equity ETFs

Following table presents the names, tickers, leverage-multipliers, benchmarks ands dates of introduction for equity ETFs selected for this thesis.

Table 2: Set of equity ETFs

Name	Ticker	Leverage	Benchmark	Date of Introduction
ProShares UltraPro QQQ	TQQQ	3x	QQQ	*2/2010*
ProShares Ultra QQQ	QLD	2x	QQQ	*6/2006*
ProShares UltraPro S&P500	UPRO	3x	IVV	*6/2009*
ProShares Ultra S&P500	SSO	2x	IVV	*6/2006*
ProShares UltraPro Financial Select Sector	FINU	3x	IYF	*7/2012*
ProShares Ultra Financials	UYG	2x	IYF	*1/2007*
Direxion Daily Small Cap Bull 3X Shares	TNA	3x	IWM	*11/2008*
Direxion Daily Technology Bull 3X Shares	TECL	3x	XLK	*12/2008*
ProShares UltraPro Short QQQ	SQQQ	-3x	QQQ	*2/2010*
ProShares UltraShort QQQ	QID	-2x	QQQ	*7/2006*
ProShares UltraPro Short S&P500	SPXU	-3x	IVV	*6/2009*
ProShares UltraShort S&P500	SDS	-2x	IVV	*6/2006*
Direxion Daily Technology Bear 3X Shares	TECS	-3x	XLK	*12/2008*
ProShares UltraPro Short Financial Select Sector	FINZ	-3x	IYF	*7/2012*
Direxion Daily Small Cap Bear 3X Shares	TZA	-3x	IWM	*11/2008*
ProShares UltraShort Financials	SKF	-2x	IYF	*1/2007*
Invesco QQQ Trust	QQQ	-	-	*3/1999*
iShares Core S&P 500 ETF	IVV	-	-	*5/2000*
iShares U.S. Financials ETF	IYF	-	-	*1/2007*
iShares Russell 2000 ETF	IWM	-	-	*5/2000*
Technology Select Sector SPDR Fund	XLK	-	-	*12/1998*

3.3. Fixed-income ETFs

Next table presents similar information of ETFs representing the fixed-income side, also known as bond ETFs.

Table 3: Set of fixed-income ETFs

Name	Ticker	Leverage	Benchmark	Date of Introduction
ProShares Ultra 20+ Year Treasury	UBT	2x	TLT	*1/2010*
ProShares Ultra 7-10 Year Treasury	UST	2x	IEF	*1/2010*
ProShares Ultra High Yield	UJB	2x	HYG	*4/2011*
ProShares UltraShort 20+ Year Treasury	TBT	-2x	TLT	*4/2008*
ProShares UltraShort 7-10 Year Treasury	PST	-2x	IEF	*4/2008*
ProShares UltraPro Short 20+ Year Treasury	TTT	-3x	TLT	*3/2012*
iShares 20+ Year Treasury Bond ETF	TLT	-	-	*7/2002*
iShares 7-10 Year Treasury Bond ETF	IEF	-	-	*7/2002*
iShares iBoxx $ High Yield Corporate Bond ETF	HYG	-	-	*4/2007*

3.4. Timespan of the dataset

The dataset consists of 1760 trading days from 22 October 2012 to 21 October 2019. As the selection of ETFs for this study were partly based on the availability of data, the timespan of the study is defined so that there is equal amount of days for each of the ETFs. The latest issued ETFs selected for this study are ProShares UltraPro Financial Select Sector (FINU) 3x-leveraged ETF and its inverse counterparty, ProShares UltraPro Short Financial Select Sector (FINZ), which both were introduced July 2012. The introduction date of these instruments limited the timespan to start earliest at July 2012. To get exactly 7 years of data, start date is rolled few months forward to October 2012 from which it was possible to get the exact 7 years. As presented in the methodology chapter, this dataset is used to construct overlapping investment periods of different lengths throughout the 7-year timespan.

3.5. Ultra-long-term returns of leveraged ETFs

As previous literature suggests, leveraged ETFs do not reach k-multiplied returns on longer than 1-day basis. Concluding the theories in literature review, this is based on constant leverage trap and the nature of compounding returns. The reason for these factors being problematic within leveraged ETFs, seems to be in the volatility and deviation of the returns.

The empirical section of the thesis focuses on providing empirical results of LETF returns from the last 7 years, 10/2012 – 10/2019. The main focus is to test with different holding periods, how well the selected leveraged and inverse ETFs could reach their k-multiplier returns. Before that, in the next subchapters we will take a look on the overall performance of selected ETFs during this period. At the same time, this observation and measurement of performance, provides us an empirical result of how much an investor would have earned, if they were to invest into these ETFs at the beginning of the timespan, and holding them until the end of it. In other words, each of the graphs can be also interpreted as a 7-year ultra-long investment period of ETFs presented in them.

For each ETF, compounding returns are calculated through the observing period, starting from value of 1000. Return *r* for each day *t* is calculated by market prices *p* by:

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \tag{17}$$

Leung et al. (2016, 7) present a general equation for calculating value of leveraged ETF against its benchmark as:

$$L_n = L_0 * \prod_{j=1}^{n}\left(1 + \beta R_j\right) \tag{18}$$

, where L_n denotes the value of LETF on day n, and β is the leverage-multiplier of that ETF, R_j being the daily return of benchmark. Calculating values of ETFs and LETFs on any given day t, the way the ETF portfolios are constructed in this thesis, can be expressed as:

$$V_t = V_0 * \prod_{j=1}^{t} (1 + r_j) \tag{19}$$

In the data chapter, value of 1000 is used as V_0 when calculating the values of ETFs. 1000 denominates the value of investment at the beginning of the period. In the empirical sections of methodology and results, V_0 is defined as 1 for more sophisticated application to regression analyses.

3.5.1. Equity ETFs

In this subchapter we will take a look on the overall performance of selected equity ETFs during the observing period of 10/2012 – 10/2019. Observations and graphs are grouped based on their benchmark ETFs. In other words, we will observe one benchmark ETF, and all related LETFs at a time. First, let’s take a look on the ETFs tracking S&P500 index:

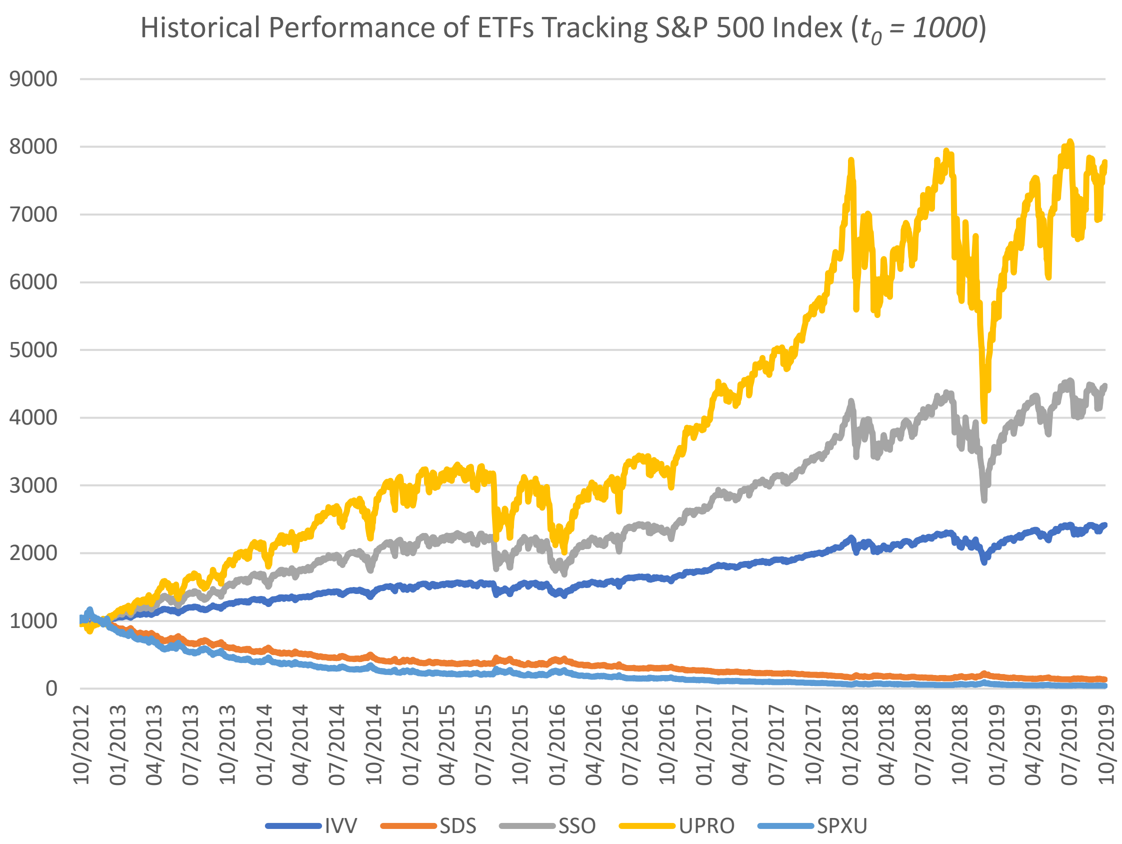

Figure 2: Historical performance of S&P500 ETFs

The graph represents the historical performance of the conventional ETF (IVV) tracking the underlying index S&P500, in addition to 2 positively leveraged and 2 inverse ETFs.

All ETFs start from point 1000. Only by a quick glance, we can already see that the inverse ETFs (SDS & SPXU) begin an unrecoverable downward path at the very beginning phases of the observed timespan. When considering this timespan, starting from 10/2012, we can clearly see that for this timespan, the positively leveraged ETFs have actually performed incredibly well. Starting from point 1000, on 22 October 2012, the underlying ETF “IVV”, has reached a value of 2416.88 on 21 October 2019. We can say that the overall performance of this benchmark ETF has been good.

Previous literature suggests that the returns of leveraged ETFs are prone to volatility on long investment periods. Figure 2 above practically represents a 7-year holding period.

Interestingly enough, even a simple graphical observation of this period shows that the ProShares' 2x-leveraged "SSO" has performed even better than the benchmark ETF, rising from 1000 points to 4471.26.

"UPRO", also a ProShares' product, represents the 3x-leveraged ETF. It is clearly visible that it has performed the best of the observed ETFs. The starting amount of 1000 would have increased to 7774.77 by the end of the period observed. Again, the performance on this particular 7-year period is exceptional, as this 3x-leveraged ETF would have actually gained more than the k-multiplier times the benchmark index.

Only getting into graphical observation of this time period, some questions and implication rise. Based on the graph, it seems that for the observed period, increasing leverage would provide better returns. However, the effects of volatility among the leveraged ETFs is also visible in the graph. Between 07/2015 and 10/2015, there is a just noticeable, sheer drop on the benchmark ETF "IVV". The drop is however much more dramatic for the leveraged ETFs. This emphasizes around 12/2018, where again, just barely visible but slightly larger drop occurs, the effects on 2x- and even more on 3x-leveraged ETFs are massive. From the end of September 2018 to end of December, just in 3 months, the return index falls from 7943.12 to 3949.00, before starting to increase again.

These visible spikes remind of the danger of volatility and leverage on not so bullish periods. The exceptional performance of these two positively leveraged ETFs, and its divergence of the consensus from previous literature, give a good motivation and reason to test for different holding periods during this timespan. Before that, let's observe the performance of rest of the selected ETFs. Below is presented the performance of selected ETFs following NASDAQ-100 index.

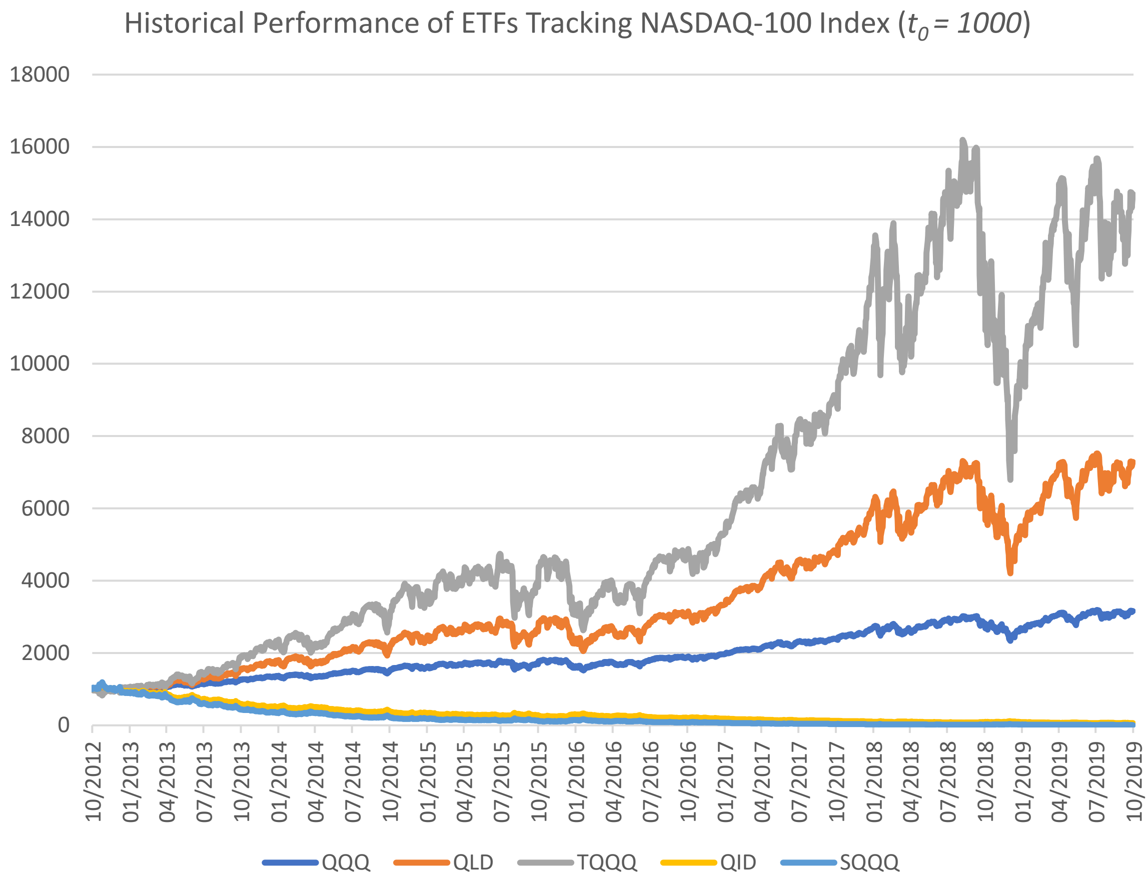

Figure 3: Historical performance of NASDAQ-100 ETFs

The graph of NASDAQ-100 LETFs looks very similar to the corresponding graph of S&P500 LETFs and it seems like these two indices are very heavily correlated. Keeping in mind the great performance of positively leveraged ETFs tracking S&P500 index, the corresponding 2x- and 3x-leveraged ETFs of NASDAQ-100 have performed even better, when considering solely this 7-year period. For benchmark ETF "QQQ", and its 2x- 3x-leveraged companions, ProShares' "QLD" and "TQQQ", the 1000-points based portfolios at the end of observing period are at 3163.98, 7283.75 and 14607.59 respectively. In addition to these great returns, both of positively leveraged indices, "QLD" and "TQQQ", have easily exceeded their naïve expectations of k-multiplier returns. Following graph presents the third set of equity ETFs, those which track the Dow Jones U.S. Financials Index.

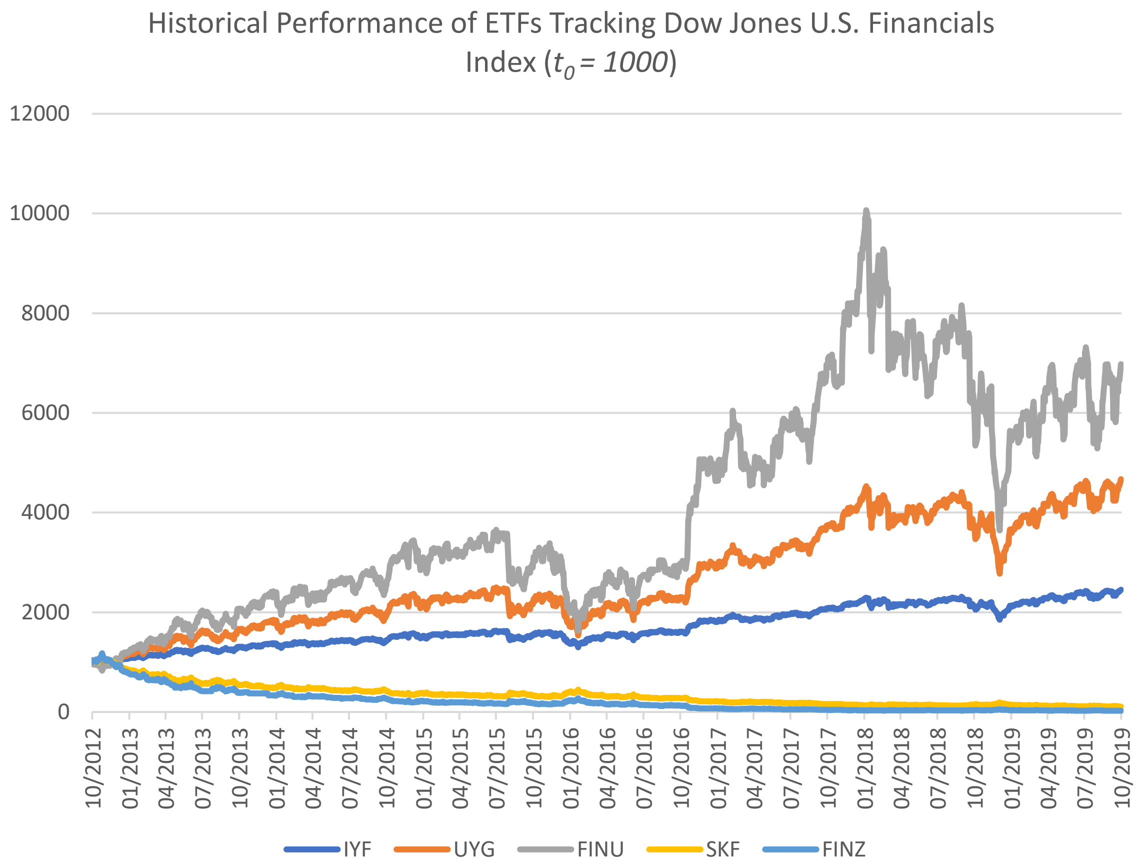

Figure 4: Historical performance of Dow Jones U.S. Financials ETFs

As with previously observed ETFs, price of the benchmark "IYF" has been in an overall positive trend throughout the 7-year period. Again, higher leverage would have provided better returns. Compared to the exceptional exceeding of k-multiplier returns within ETFs tracking NASDAQ-100, the Dow Jones Financials -tracking ETFs could not make as dramatic exceeding of k-multiplier returns at the end of the period. As benchmark ETF "IYF" and its leveraged companions, ProShares' "UYG" and "FINU" ended up at 2452.97, 4669.80 and 6973.66 respectively, they were able to exceed the naïve expectations of k-times returns as seen in table 4. However, these observations have to be considered with caution, as they only represent the returns at a single point in time and will not give reliable picture of the whole timespan. Performance of these ETFs on different holding periods, and ability to reach naïve expectations, are presented further in this thesis.

Of the three sets of ETFs presented above, each track some of the largest indices in the U.S. market. It needs to be noted here, that these three sets of ETFs were the only ones, where both 2x- and 3x-leveraged ETFs on both positive and inverse side were available. In addition to these sets, ETFs with a pair of either 2x- or 3x-leveraged ETFs are included. The following single LETFs are either 2x- or 3x-leveraged, but they are accompanied with a corresponding inverse ETF and the benchmark ETF. Again, the criteria for selecting these ETFs is presented in the “Data”-chapter of this thesis. Next, let’s observe iShares Russell 2000 ETF “IWM”, and corresponding 3x positive and inverse LETFs.

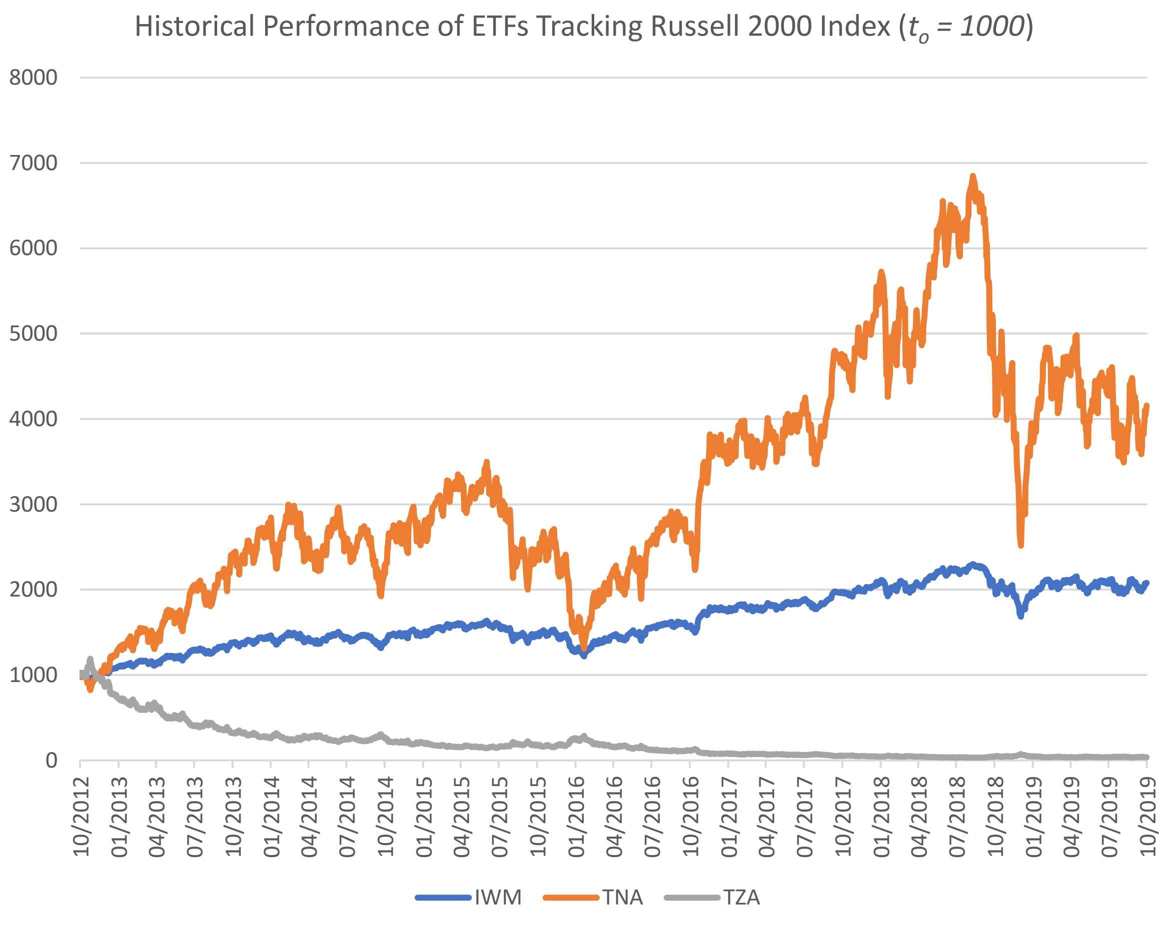

Figure 5: Historical performance of Russell 2000 ETFs

Again, on this ultra-long 7-year period, the positively leveraged ETF beats the benchmark ETF. Slightly after 01/2016 we can see a point where the lines almost meet, but after all the leveraged ETF stays above the benchmark. When observing the k-multiplier return at the end of this 7-year period, the 3x-leveraged ETF finishes at 4159.14, benchmark being at 2080.41. The return of leveraged “TNA” is around 2.9-times the returns of benchmark “IWM”, rather than the naïve expectation of tripling the return. But as discussed previously, this is only on the assumption of 7-year holding period. As we can see, the value of portfolio of “TNA” has been nearly 3-times compared to value of “IWM” slightly after 07/2018.

Next set of ETFs track Technology Select Sector Index, which consists of large and mid-cap technology stocks. As with Russell 2000 index, one 3x-leveraged and one -3x inverse ETF are included in addition to the benchmark ETF.

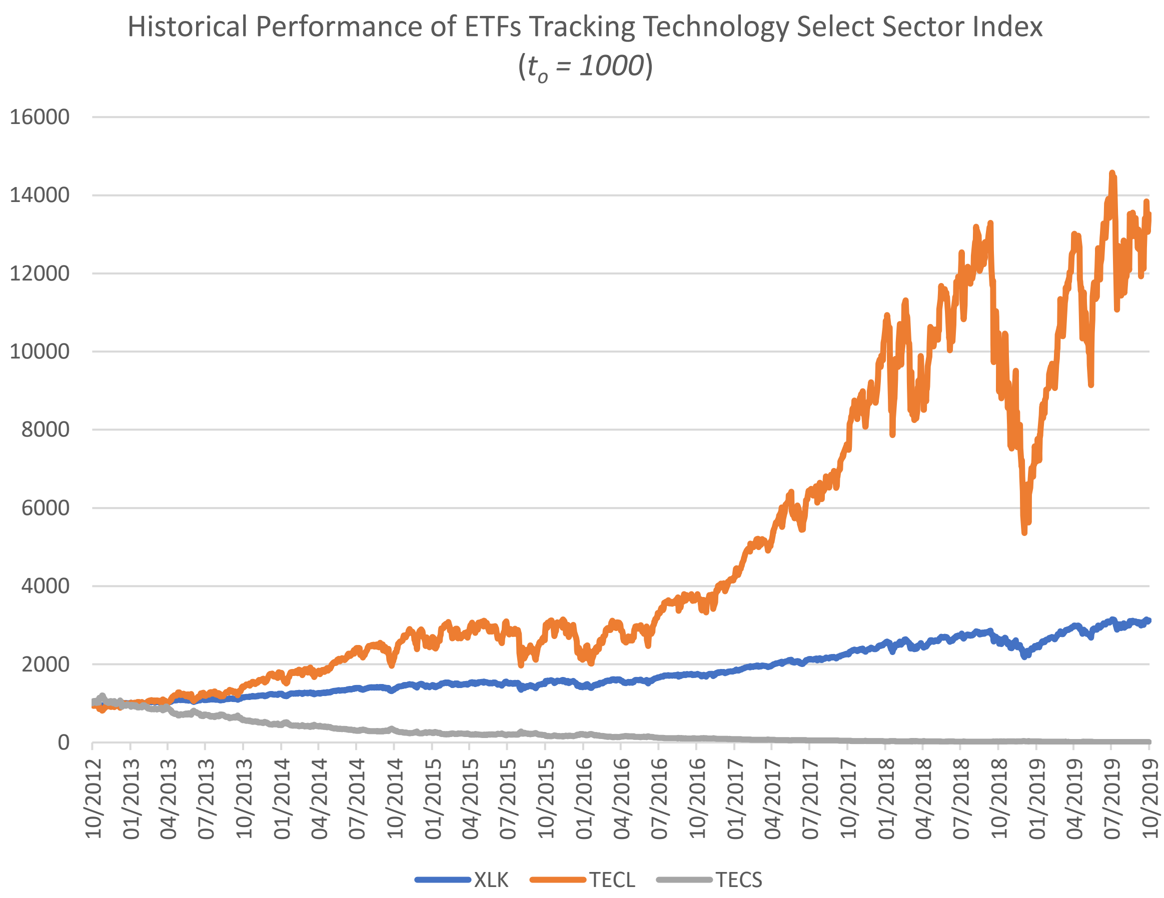

Figure 6: Historical performance of Technology Select Sector ETFs

The performance is very similar to the others observed above. Again, the 3x-leveraged ETF beats the conventional benchmark by ease. By the end of the 7-year period, investing in the 3x-leveraged “TECL” in the 2012 would have provided over 4 times the returns than when investing in the non-leveraged “XLK”.

It is clear that there is a common trend among the sets of ETFs observed above. For the ultra-long period, the ETFs with more leverage would always beat the ETFs with lesser or no leverage. It seems like these 7 year would have been exceptionally good to long-term invest on leveraged ETFs. This is interesting, since no previous literature would suggest investment periods this long, in context of leveraged ETFs. It is now very important to test different investment periods from this timespan to find the desirable holding periods, where one would earn k-multiplier returns with statistical reliability.

The other major trend, repeatedly observable in the presented ETFs, is the performance of inverse ETFs. Every one of the observed inverse ETFs, start to decline towards zero, and could provide absolutely no returns on investment period this long. It is totally reasonable and expected, as the underlying index on each set has been in upward trend. Also, when considering from the point of long-term investing, it is often justified by overall economic growth and how that would positively affect stock prices. Going short on any index, especially a large market index for a long investment period would make no sense from this point of view.

3.5.2. Fixed-income ETFs

This subchapter focuses on observing performance of leveraged fixed-income ETFs in contrast to their non-leveraged counterparties. The 7-year performance is measured and observed the same way as with equity ETFs. First, we will observe ETFs tracking Barclays Capital U.S. 20+ Year Treasury Bond Index.

Treasury indices, like Barclays Capital U.S. 20+ Year Treasury Bond Index, track the performance of corresponding securities. For example, this particular index is tied to track the performance of United States Treasury securities that have a remaining maturity of at least 20 years (ETFdb.com 2019e). Let's first observe the non-leveraged ETF tracking this index, iShares "TLT", to get a better understanding of overall performance of the underlying index during the 7-year observation period.

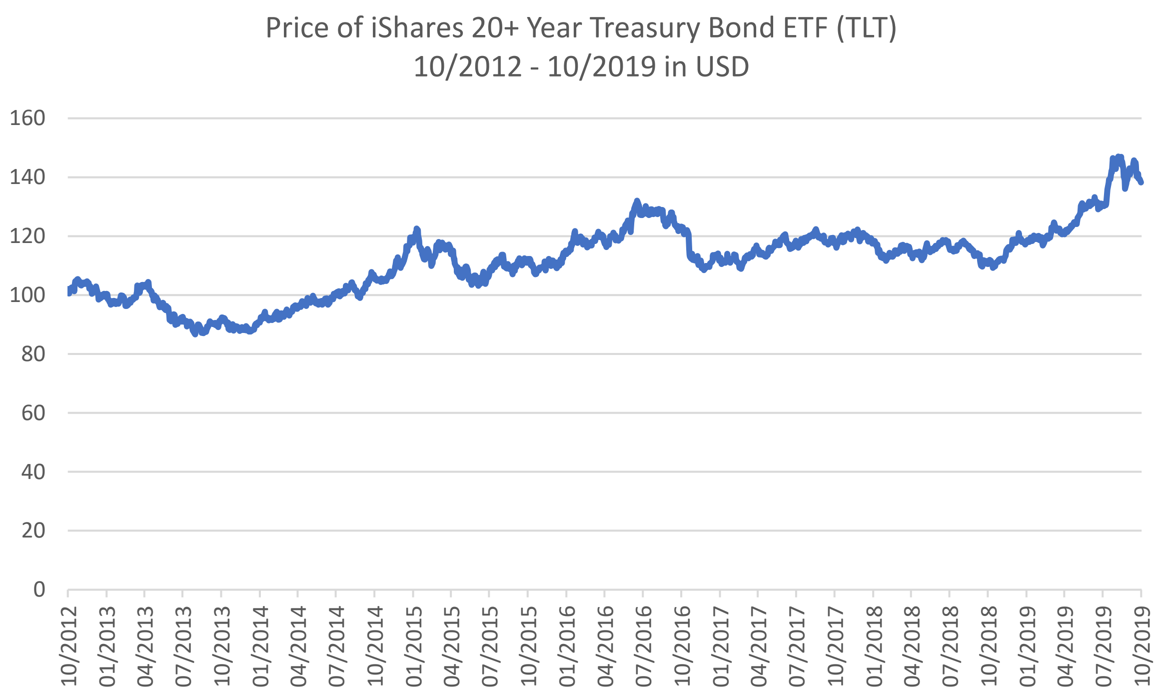

Figure 7: iShares 20+ Year Treasury Bond ETF

By graphical observation, we can say that despite alternating upward and downward trends, price of "TLT" is higher at the end of the period than at the beginning. During this 7-year period the price of "TLT" has risen from 100.60 to 138.29 USD. Now, let's consider the daily returns and add in the leveraged companions, just like is done with the equity ETFs in previous subchapter.

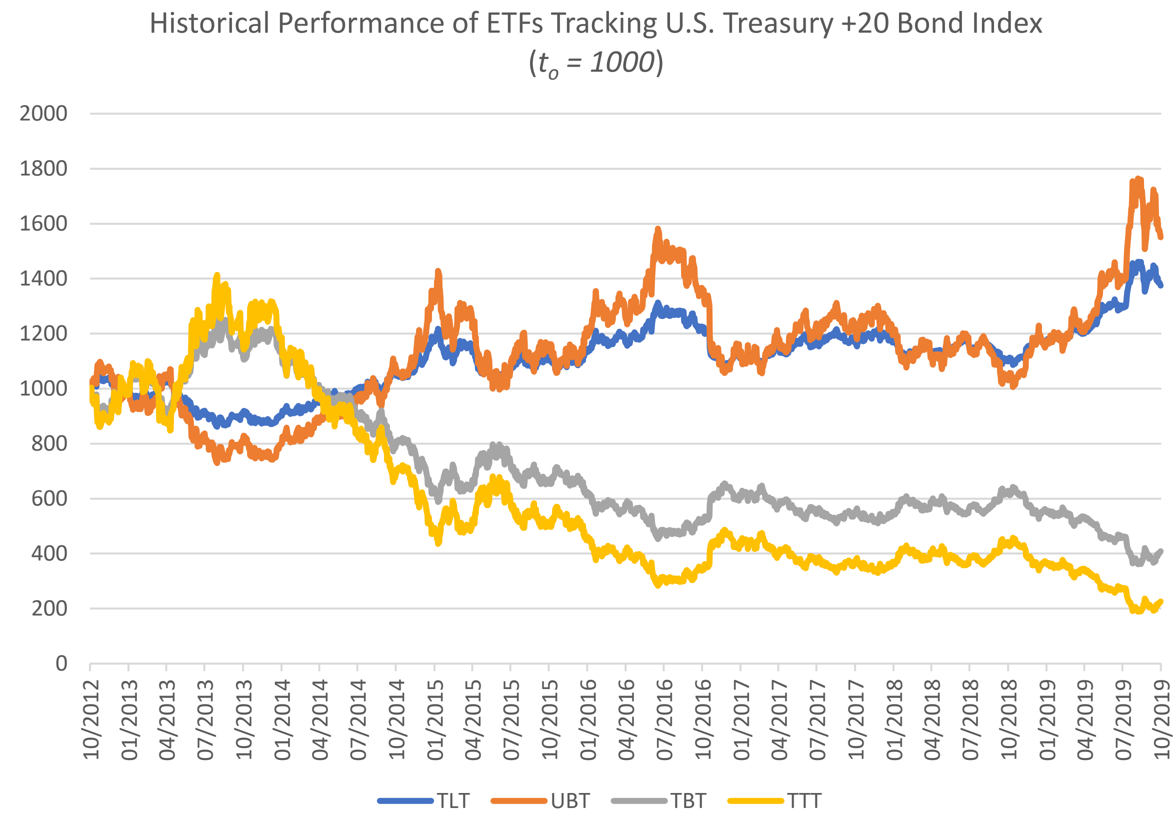

Figure 8: Historical performance of U.S Treasury +20 Bond Index ETFs

In the graph above, benchmark ETF "TLT" is accompanied with 2x-leveraged "UBT", and -2x and -3x inverse ETFs "TBT" and "TTT". The alternating trends visible in price graph above now reflect to this graph with leveraged products. The overall look of the graph is a lot different compared to ones with equity ETFs.

First off, we can see that the leveraged ETF "UBT" is not able to greatly divert from the benchmark ETF, as which was the case within all observed equity LETFs. Again, based on the previous literature, this is most likely because of larger amount of volatility throughout this period. Naïve expectations of k-multiplier returns are not reached as benchmark "TLT" and 2x-leveraged "UBT" end up in points 1374.62 and 1550.73 respectively. However, after this ultra-long investment period of 7 years, leveraged ETF performed better than the benchmark.

Another notable difference compared to equity ETFs, is the performance of inverse ETFs. On all cases of equity ETFs, inverse ETFs could provide absolutely no yield on the ultra-long period. As we can see from figure 8, even over 1 year after initial investment, both -2x and -3x inverse ETFs are profitable. When considering the whole 7-year period, the returns start to diminish towards zero, just like with inverse equity ETFs.

However, the decline is not as dramatic as with equity ETFs. The reason for this is likely behind the performance of underlying index, as there was very positive trend among the equity indices, and this fixed-income treasury index had more downward trends also. Next, let's observe the performance of ETFs, which track a very similar index to the previous one. Instead of 20+ year treasury bonds, the following index tracks treasury bonds with maturities within 7 to 10 years. Conventional ETF and both 2x and -2x leveraged ETFs tracking this index are presented in the following graph.

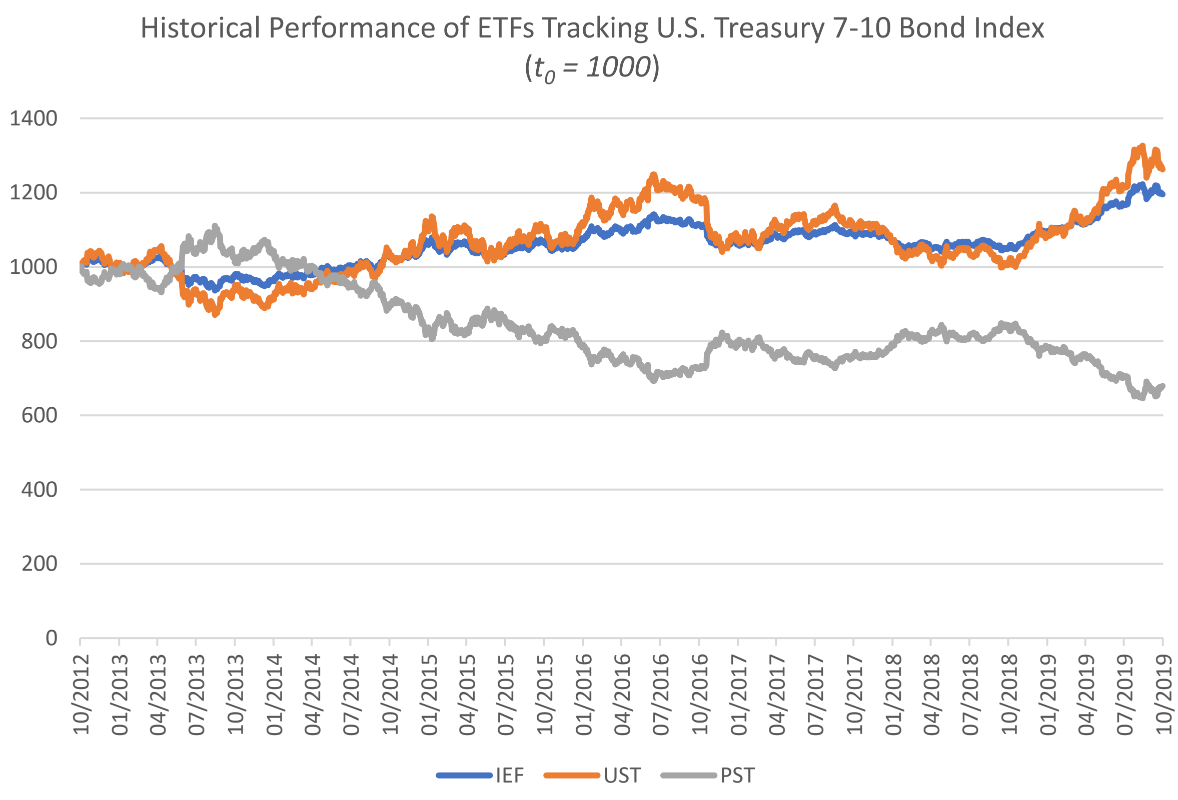

Figure 9: Historical performance of U.S. Treasury 7-10 Bond Index ETFs

Although missing a -3x inverse ETF like the 20+ index has, the overall shape of the graph is very similar, as the trends are visible in both graphs. Overall performance of the conventional ETF “IEF” is not as good as “TLT”, which tracks the performance of 20+ year maturity bonds. This is reasonable as bonds with longer maturity are supposed to provide more return, as they also hold more risk.

Findings from this graph are no different from the previous ones as the leveraged ETF is able to outperform the benchmark ETF at the ultra-long period. However, unlike with equity ETFs, among these fixed-income ETFs we can see several occasions of benchmark and leveraged ETF overlapping each other. This supports the idea, that when testing with different holding-periods throughout the timespan, the leveraged ETF may not actually be able to constantly outperform the benchmark.

Lastly is presented iShares iBoxx $ High Yield Corporate Bond ETF (HYG), and a 2x-leveraged ETF “UJB”, which both track iBoxx $ Liquid High Yield Index. As the name suggests, this index consists of liquid high yield corporate bonds for sale in the United States (ETFdb.com 2019f). “UJB” was the only leveraged ETF tracking an index consisting of corporate bonds, and unfortunately there was no inverse pair available. However, we can still study the performance between the 2x-leveraged and benchmark ETF.

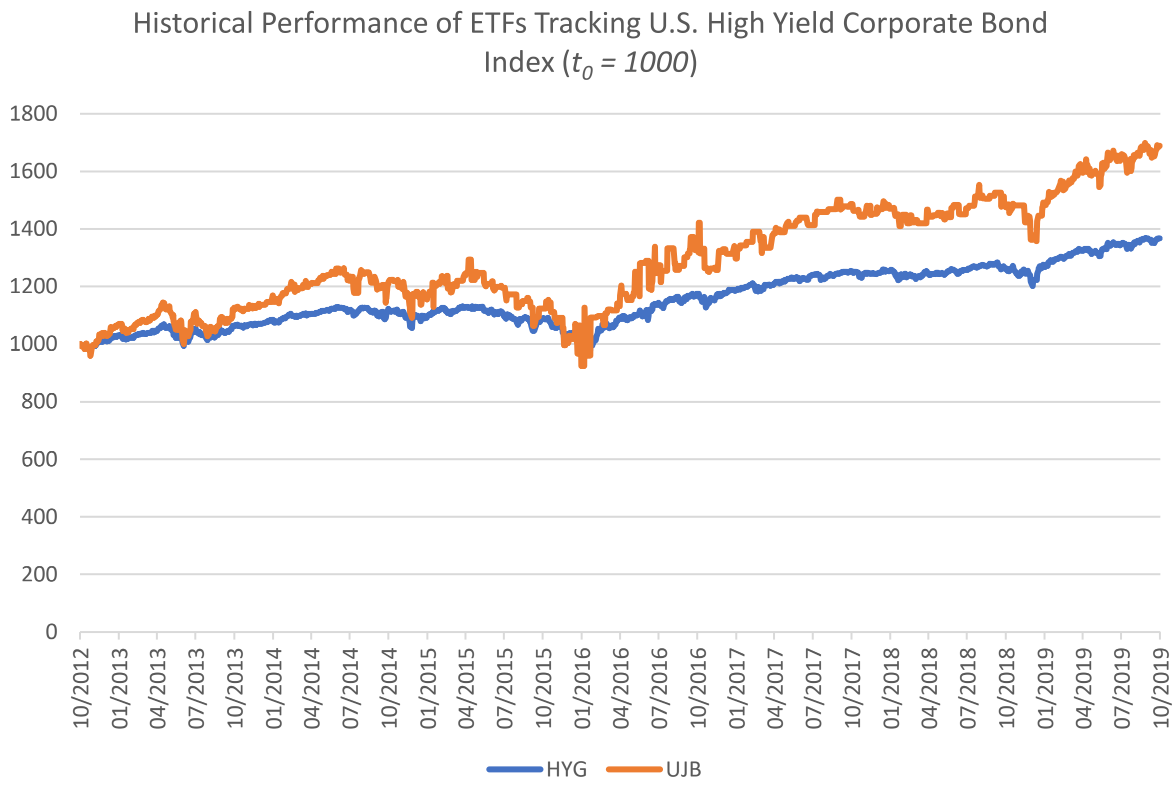

Figure 10: Historical performance of U.S. High Yield Corporate Bond Index ETFs

Even the high yield bond index “HYG” and its leveraged companion “UJB” have both been profitable on the ultra-long period. At the end of the period, non-leveraged “HYG” is at 1367.55, and leveraged “UJB” at 1688.65 points.

3.5.3. Summary of findings

When considering the previous studies of long-term performance of leveraged ETFs, the 7-year performance seen in this chapter is rather surprising. Performance of 8 benchmark indices were measured in addition to 11 positively leveraged and 11 inversely leveraged ETFs. Findings from this particular 7-year ultra-long period are clear: All of the leveraged ETFs and none of the inverse ETFs were able to beat the benchmarking ETFs. Below is presented summarized table of performance during the ultra-long period.

Table 4: Summary of 7-year performance

Type	Leverage	Ticker	Benchmark Return	LETF Return	Multiplier	Exceeded
Equity	3x	TQQQ	2163,98	13697,6	6,329814	1
Equity	3x	UPRO	1416,88	6774,77	4,781471	1
Equity	3x	FINU	1452,97	5973,65	4,111337	1
Equity	3x	TNA	1080,41	3159,14	2,92402	0
Equity	3x	TECL	2123,21	12510,3	5,892182	1
Equity	2x	QLD	2163,98	6283,75	2,903793	1
Equity	2x	SSO	1416,88	3471,27	2,449939	1
Equity	2x	UYG	1452,97	3669,8	2,525723	1
Equity (i)	-3x	SQQQ	2163,98	-988,12	-0,45662	0
Equity (i)	-3x	SPXU	1416,88	-958,31	-0,67635	0
Equity (i)	-3x	FINZ	1452,97	-976,07	-0,67178	0
Equity (i)	-3x	TZA	1080,41	-963,95	-0,89221	0
Equity (i)	-3x	TECS	2123,21	-988,65	-0,46564	0
Equity (i)	-2x	QID	2163,98	-937,8	-0,43337	0
Equity (i)	-2x	SDS	1416,88	-864,92	-0,61044	0
Equity (i)	-2x	SKF	1452,97	-891,35	-0,61347	0
Fixed-Income	2x	UBT	374,62	550,73	1,470103	0
Fixed-Income	2x	UST	195,34	262,38	1,343196	0
Fixed-Income	2x	UJB	367,55	688,65	1,873623	0
Fixed-Income (i)	-3x	TTT	374,62	-774,51	-2,06746	0
Fixed-Income (i)	-2x	TBT	374,62	-591,83	-1,57981	0
Fixed-Income (i)	-2x	PST	195,34	-320,56	-1,64104	0

On table 4 is presented the returns of each leveraged and inverse ETF, their benchmark returns, and the ratio between these two. In the last column is presented a binomial, telling us whether the particular LETF has exceeded its benchmark returns multiplied by the leverage-multiplier. In other words, it is telling us, if the naïve expectation of k-times returns on this particular 7-year period is exceeded. On the table 5 below, observed LETFs are divided in four groups. This table tells us the percentage of ETFs on that group that were able to beat the benchmark by the k-multiplier.

Table 5: Fraction of LETFs reaching 7-year multiplier by groups

	Leveraged	Inverse
Equity	0,875	0
Fixed-income	0	0

7 out of 8, or 87,5% of positively leveraged equity ETFs were able to beat the multiplier on this period. Out of the four groups, these were the only type of LETFs to reach it.

As stated before, no conclusions can be done from these measurements because of the possibility that this is only a single 1760-day period, that has been applied to different products. On their own, findings about performance of LETFs during this period are however unexpected, when considering the previous literature and the consensus of leveraged ETFs not being suitable for long-term investments. These observations of data suggest that there may be some possibilities to use leveraged ETFs for long period investing purposes. Bilello (2018) observed 3x long “TQQQ” against its benchmark “QQQ”, discussing the negative effects of volatility decay, but also noticing the tremendous gains that would have occurred if one had bought and held 3x-leveraged “TQQQ” for 9 years before 2018.

In overall, the purpose of this chapter was to give a good view of the data used in this study and strengthen the research gap and need to empirically revise previous literature of the subject. By only observing the data as a 7-year investment period, we would run to statistically invalid results, as a single holding period would be considered as biased. To tackle the issue of biased timespan, I’m conducting a test of different holding periods throughout this 1760-day set to see, whether there are suitable long-term holding periods for these instruments.

4. METHODOLOGY

In this chapter is presented the methodology for testing the LETF performance on different timespans. Ordinary Least Squares (OLS) -method of regression analysis is used to estimate the k-multiplier performance of subject LETFs against their benchmarks, on various timespans. Further, t-tests and correlation analyses are done to get more insight and understanding of the results.

4.1. Choosing the methodology

When going through previous literature and research of the subject, it turned out that Lu et al. (2012) provide a very applicable framework for the empirical section of this thesis. First of all, when considering the most suitable way to measure long-term performance, I had a very similar idea of rolling the investment period through the timespan of the data, which will be presented further in this chapter. When going through previous empirical studies of the subject, I found out about Lu et al. (2012) methodology of measuring long-term performance and found it very applicable on this study also. The selection of using regression analysis methodology was further reinforced by other studies, as seen in the 4.2.

Using this type of methodology seems reasonable. First of all, Lu et al. (2012) present a general formula of 2-day returns of 2x-leveraged and inverse ETFs. As I applied the equations to theoretical 3x-leveraged and inverse ETFs in chapter 2.3.2., we could see that increasing leverage should distance the expected 2-day return further away from the naïve expectation of k-multiplier returns. This creates a research gap to empirically test with 3x-leveraged ETFs also, as these were not included in the original study.

In their original dataset, Lu et al. (2012) have included 4 leveraged (2x) ETFs, 4 inverse (-2x) ETFs and 4 benchmarking indices. Dataset of this study is a lot different: in addition to adding 3x-leveraged and -3x-inverse ETFs, I have also included fixed-income ETFs into the study. It also has to be noted, that in the original study, the actual indices were used as benchmark. This study uses the conventional non-leveraged ETFs, as the focus is on comparing actual investment products to each other. From the point of suitability of the methodology, this should make absolutely no concerns. Tracking error of the benchmarks, in other words, the ETFs not being able to strictly follow its underlying index, is not a problem in this context, performance of leveraged ETFs is not compared to their benchmarking index. The focus is on researching the differences between the leveraged ETFs and non-leveraged ETFs tracking the same indices, and research and comparison being done between these subjects, it does not matter what happens in the underlying index, related to the benchmark ETFs.

In addition to studying wider set of leveraged ETFs, the data within the timespan is very different. As we will see from the next chapters, the overall performance of each of the conventional benchmark ETFs selected for this study, has been good during this period. This is emphasized more within equity ETFs, which have been following a very positive trend during these 7 years of observation. All the tests presented in the following subchapters are done using Real Statistics Add-In for Excel (2019), which provided all the functions needed for the tests conducted in this thesis.

4.2. Modeling k-times returns

For each timespan of 3-days, 1-month, 3-month, 1-year and 2-years, the performances of each leveraged and inversed ETFs are regressed against the performances of their benchmark portfolios. Rolling start-date performances are calculated by using overlapping data. The start date is rolled forward day-by-day to calculate historical performance for the given period.

Univariate or single variable version of regression model is used in this thesis. This is because of I am explaining the performance of leveraged ETFs solely with the performance of its benchmark ETF. When considering only these two portfolios and their relation to each other, there is no need for additional variables, as the interest is only in finding out, how the benchmark explains the observed LETF.

The following formula presents the regression used to estimate constant α and slope β. Variable R_L represents the performance of leveraged ETF, while variable R_B represents the performance of underlying benchmark ETF, during that same investment period.

$$R_L = \alpha + \beta R_B + e_t \tag{20}$$

By leaving out the constant term α, we could assess the k-multiplied returns of observed investment periods. The estimated β would then represent the assumed multiplier of the benchmarking ETF in relation to the leveraged ETF. However, when conducting regression analyses, it came out that in addition to β the constant term α is also an interesting factor in this study, in the context of theory behind long-term performance of LETFs, as we will see in the results.

Charupat et al. (2014) present similar formula which is traditionally used in several studies in the field of ETFs and leveraged ETFs. In the context of conventional ETFs, it is often used to measure tracking error between the ETF and its underlying index. Previous studies suggest, that it is also applicable to measuring performance of leveraged ETFs, where tracking error is considered to be the inability to match the k-times returns between the underlying benchmark ETF and the leveraged ETF: in addition to their own previous article of The pricing and performance of leveraged exchange-traded funds (Charupat et al. 2010), similar approach is used in Lu et al. (2012) and Shum et al. (2012), which both focus on measuring LETF performance. Charupat et al. (2014) present the equation as:

$$r_{t,t+N} = a + b * \left[\prod_{j=1}^{N} \left(1 + i_{t+j-1,t+j}\right) - 1 \right] + e_t \tag{21}$$

Comparing to the simplified equation 20, we can see that the part of the equation inside brackets on equation 21, equals the R_B in the first equation (20). Although Charupat et al. (2014) define *r* in the formula as return based on the net asset value (NAV) of the ETFs, the formula can equally be applied to calculate performance by daily prices also, as is done in this thesis.

Charupat et al. (2014) also present some criticism towards this classic model of measuring tracking error, since it is difficult to interpret from the results, as the coefficients of alpha and beta cannot really answer why they deviate from their expected values of 0 and *k*. The model is however well suitable to make comparative analysis between the test subjects. The main thing considered here, is an attempt to find out the lengths of investment periods, where investor can reliably assume that the leveraged ETF will provide k-multiplied returns. When the linear relationship between returns of a LETF and its underlying benchmark ETF break, the returns will not be reliably k-multiplied.

Equation 21 is used for the regression analyses in the empirical section of this thesis. As previous studies of LETF performance suggest, using overlapping data is often problematic because of possible heteroscedasticity and autocorrelation. Newey-West standard errors (Newey et al. 1987) are used to partly overcome these problems and get more reliable standard errors. Although having received criticism (Harri et al. 1998), it is still common way to handle standard errors when using overlapping data.

Issues of heteroscedasticity and autocorrelation violates the Gauss-Markov theorem of “BLUE”-assumptions of ordinary least squares regression analysis, not providing the “Best Linear Unbiased Estimators”. Despite this, as mentioned, similar datasets have been studied previously with regression analyses similar to 21. As we will see in the following chapters, the focus is on the timeframe where the expected k-slope linear relationship between the LETF and benchmark ETF breaks. Although the estimators not being precise, their magnitude is reliably observable.

As every single combination of a leveraged ETF and a timeframe needs a separate regressions analysis to be done, the ETFs related to the Russell 2000 - and Technology Select Sector Index are scrapped at this point. In addition to reducing the large amount of regression analyses needed, it is reasonable because of the very similar performance among all the chosen leveraged equity ETFs. Also, now all the benchmark equity ETFs have equally every type of leveraged and inverse (2x, 3x, -2x, -3x) available. The high yield fixed-income ETF is also left out at this point to get a good set of reasonably comparable sets of ETFs.

4.3. Risk-adjusted returns – T-test and Spearman Rank Correlation

This subchapter presents the additional methodology of testing the hypothesis of risk-adjusted returns – how the ratio of risk and return is supposed to be equal between leveraged and non-leveraged ETFs, as presented in the null-hypothesis. This can be concerned rather additional and notably different approach to the risk-return ratio than measuring k-multiplied returns. Methodology of calculating the risk-adjusted returns, as well as testing the statistical difference of those returns between leveraged and conventional ETFs is presented here.

Sharpe ratio is well-known meter for risk-adjusted returns. In equation of Sharpe ratio, excess return of an instrument over a risk-free rate is divided by volatility of that investment (Sharpe 1966). This equation returns a ratio, which is comparable between investments. Based on Sharpe ratio, historical volatility is used to measure risk of the investments, and the risk-free ratio is left out of the equation as I did not find a need to calculate the actual Sharpe ratios for comparison, and only wanted to adjust the returns with historical volatility. This keeps the comparison between LETFs in line with the regression analyses, which also did not take account on risk-free rate. Historical volatility is calculated by taking standard deviations of the daily returns for the observed period:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(r_i - \bar{r})^2}{N-1}} \tag{22}$$

Cumulative returns, are then divided by the standard deviation from that specific timeframe, leading to a value of risk-adjusted returns. Based on the null-hypothesis, risk-adjusted returns between leveraged ETFs and their non-leveraged benchmarks should be equal. Statistical difference between these groups is tested with t-tests.

As we will see in the results, there was an additional interest to test the effects of the volatility on abnormal risk-adjusted returns of leveraged ETFs. In this methodology, abnormal risk-adjusted returns are considered as the deflection of a LETF's risk-adjusted returns from its benchmark ETF's risk-adjusted returns, on a given period. Effect of volatility to the abnormal returns of leveraged ETFs is tested with Spearman Rank Correlation. From the previous literature, it should be expected that the volatility would have definitely a negative effect to abnormal returns of leveraged ETFs.

5. RESULTS

This chapter presents the results of this thesis. First, results of the regression analyses conducted in attempt to find out adequate investment periods for leveraged ETFs are presented. After that, there is an additional methodology of t-testing, to further study the risk-adjusted returns. Lastly, effects of volatility on the abnormal returns of leveraged ETFs presented by previous literature are empirically tested.

5.1. Performance of leveraged ETFs against benchmarks

Following subchapters contain the main methodology and core of the empirical side of this thesis. Regression analyses are done to find out investment periods, where investor can safely assume that their investments would provide k-times returns.

Let’s start by graphically observing the performance of LETFs on 1-day investment period. In the figures below, returns of leveraged ETFs are plotted against returns of the benchmarking conventional ETFs. Below is presented 1-day returns of these instruments. The purpose of this is to evaluate whether there is significant tracking error between the leveraged instrument and the benchmark. As discussed before, LETFs being constructed to track daily changes of the underlying index with a certain multiplier, there should not be significant tracking error on 1-day basis. Below, in figure 11 is presented graphs of 1-day returns of leveraged and inverse ETFs of S&P500-index, plotted against their benchmarks.

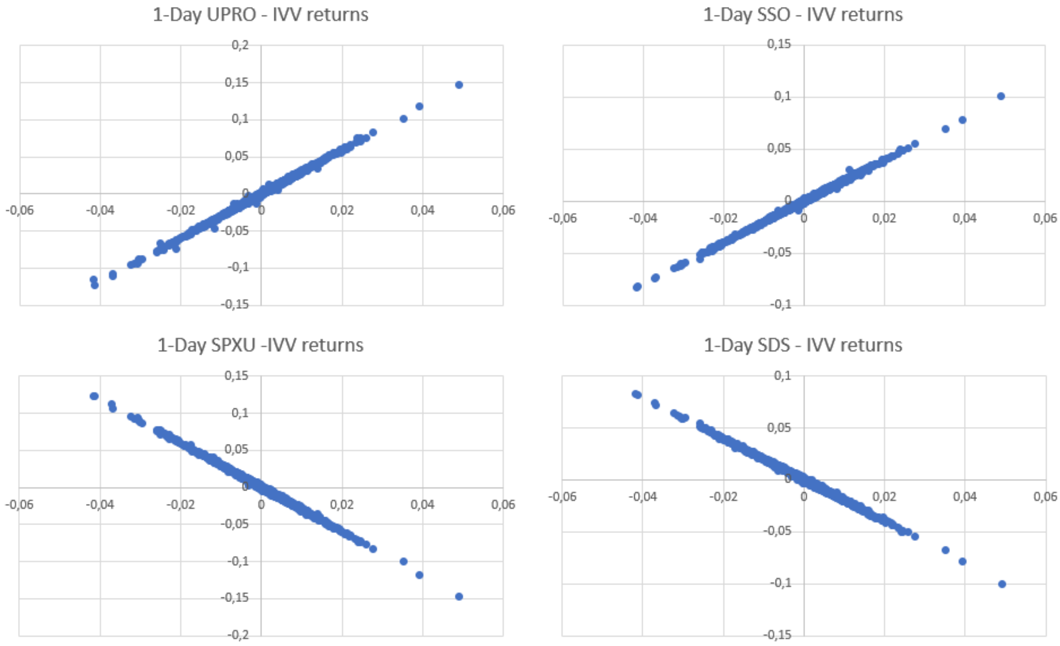

Figure 11: 1-day tracking errors of S&P500 ETFs

As we can see from the graph, the lines are straight, and we can assume a linear relationship between these points. Also, we can see that the lines pass both axis at around point 0. Considering this from the point of regression analysis presented in the methodology chapter, the intercept α would be 0, and the β would represent the slope of the graphs. For example, let's regress the "UPRO" against "IVV" as presented in the methodology. Following results occur:

Table 6: Regression results of 1-day tracking error of "UPRO" against its benchmark

R Square	0,997645403					
	Coefficient	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-0,000129439	2,82926E-05	-4,575	5,1E-06	-0,000184929	-7,3948E-05
Variable X 1	2,980975398	0,003454958	862,811	0	2,974199137	2,987751658

On table 6, the intercept α is very close to zero and "Variable X 1", which in this case is the beta, is 2.981, which is very close to the expected 3. Based on this regression, we can say that “UPRO” is able to track 3-times the non-leveraged ETF on daily basis, as it is supposed to.

This regression analysis shows us, how the leveraged ETFs should perform also on longer timeframes, if we were to rely on the naïve expectation of k-multiplied returns. On any given holding period, the beta would match the leverage-multiplier of that particular instrument. In addition to that, we would also assume the intercept, constant term α being zero. Any deviation from zero would suggest, that when the return of benchmark ETF on observed period is 0, the return of corresponding leveraged ETF would be something else than 0. This is why not only β, but also α is interesting in this concept, and there seems to be no reason to leave it out of the analysis at this point, unlike is done in some previous studies. To ease comparison on longer investment periods in the following subchapter, below is presented 1-day tracking errors on LETFs tied to Barclays U.S. 20+ Year Treasury Bond Index:

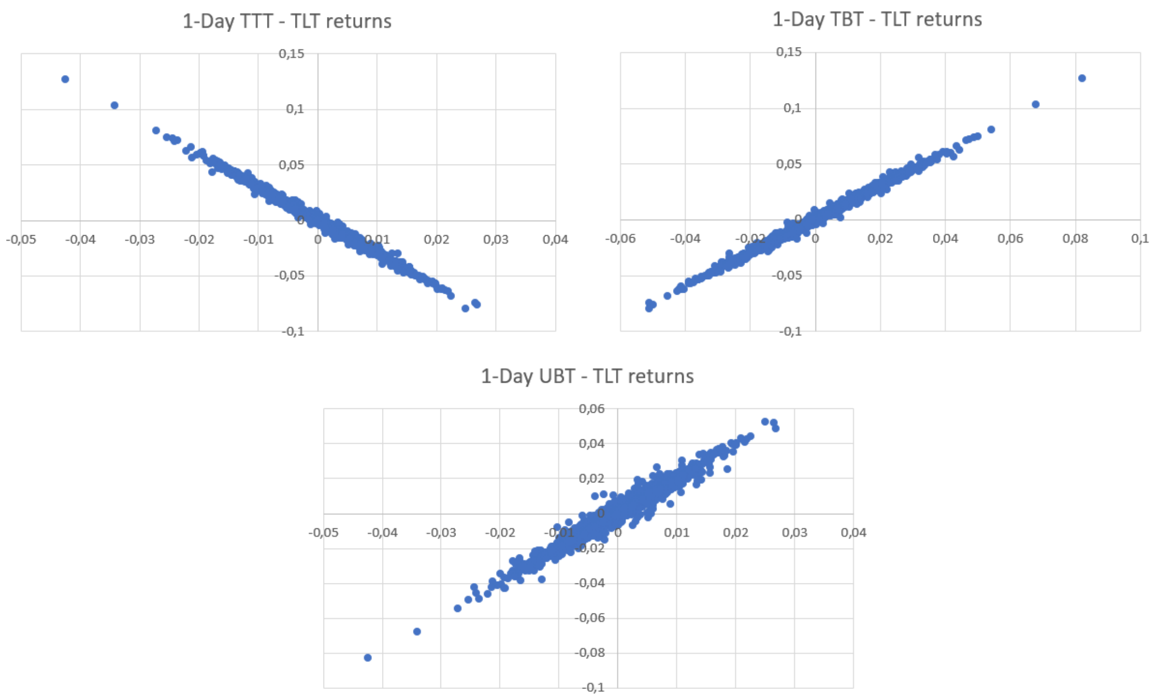

Figure 12: 1-day tracking errors of Barclays U.S. 20+ Year Treasury Bond ETFs

Again, tracking error seems to be reasonable. On the 2x-leveraged ETF “UBT” there seems to be more deviance on daily basis. However, considering a regression line through the observation, we can assume rather small tracking error, where alpha and beta would be 0 and 2, respectively.

5.2. Long-term performance – Graphical analysis

As presented in 5.1., tracking errors within LETFs on 1-day basis are expected to be minimal. In the following subchapters, the same analysis is applied to longer holding periods. From this on, tracking error is considered as the LETF’s inability to track its underlying conventional ETF on a given investment period. This way, the ability to gain k-times returns can be analyzed.

5.2.1. 3-day performance

Below is presented graphs of the returns of leveraged ETFs against their benchmark-returns on 3-day investment period. Overlapping data is used throughout these observations. To conserve space and for easier reading, only graphs from one set of equity-ETFs (S&P500) and fixed-income ETFs (Barclays U.S. 20+ Year Treasury Bond Index) are presented. After graphical observation, assumptions about performance can be made, and then further tested with regression analysis.

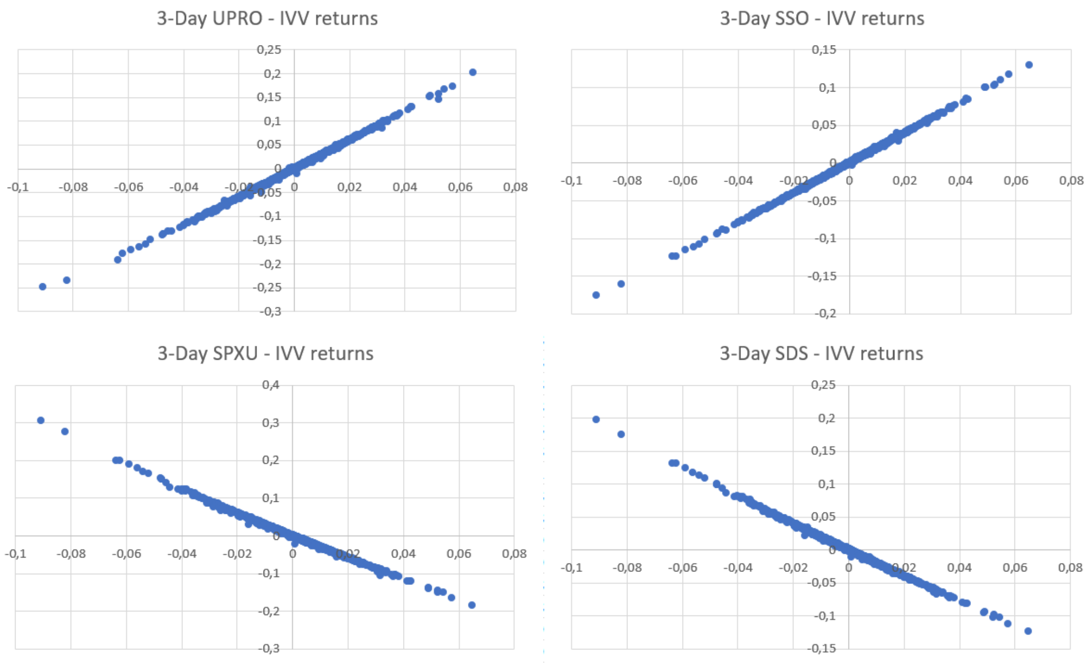

Figure 13: 3-day returns of S&P500 ETFs

The graph above presents the 3-day performance of benchmark index, S&P500-tracking "IVV", against the 4 chosen LETFs, (3x, 2x, -3x, -2x). Comparing to the 1-day returns, graphics look very similar. From the graph, we can assume very little tracking error, which would mean that on 3-day period, k-times returns would be achieved. This is because of the line crossing the axis close to 0 (intercept alpha) and the slope being close to 2, 3, -2 or -3 (coefficient of beta), depending on the LETF considered. Next graph is an example from the side of fixed-income LETFs:

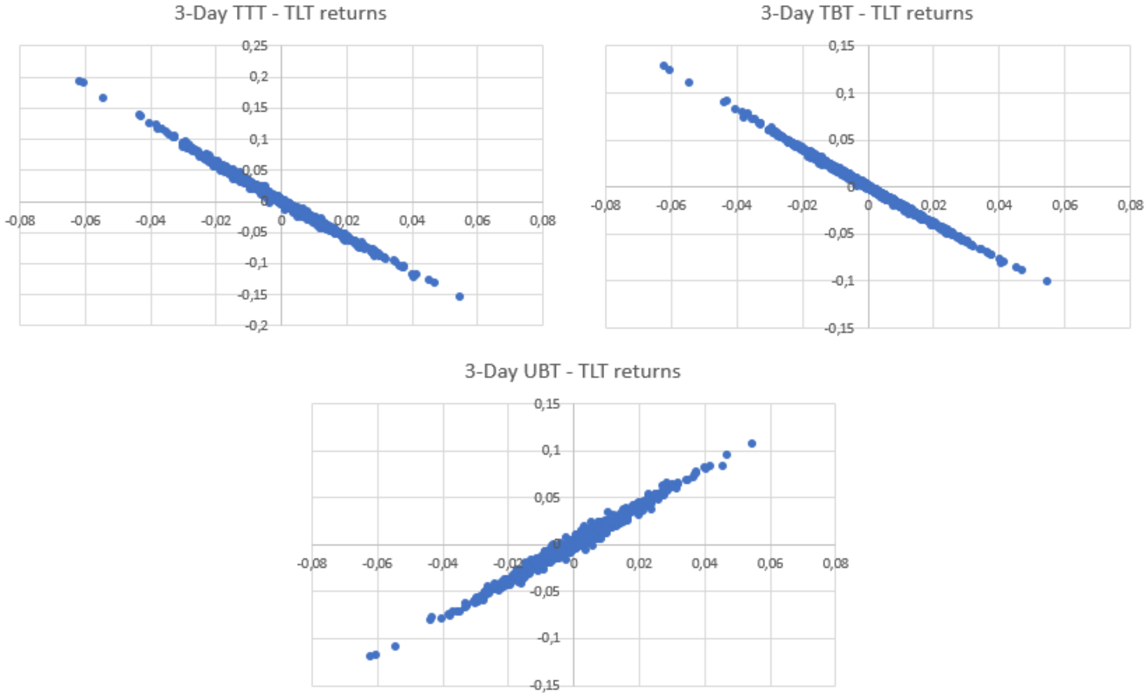

Figure 14: 3-day returns of Barclays U.S. 20+ Year Treasury Bond ETFs

The graph represents the three selected ETFs following the Barclays U.S. 20+ Year Treasury Bond Index. To recap, “TTT” is the 3-times inverse (-3x) ETF, “UBT” and “TBT” being the 2x and -2x leveraged ETFs, respectively. Results are very similar to ones within S&P500 index. Interestingly, compared to the 1-day performance, the deviation seems to get smaller on all of the three instruments.

5.2.2. 21-day performance

Now the investment period is increased to a 1-month, or 21-day long period. As discussed in previous chapters, 1-month period is considered to be a reliable period to achieve k-times returns. To keep the comparison between timeframes reliable, the observations are proceeded using the same two sets of leveraged ETFs as previously. Below is presented the S&P500-related ETFs on 1-month period.

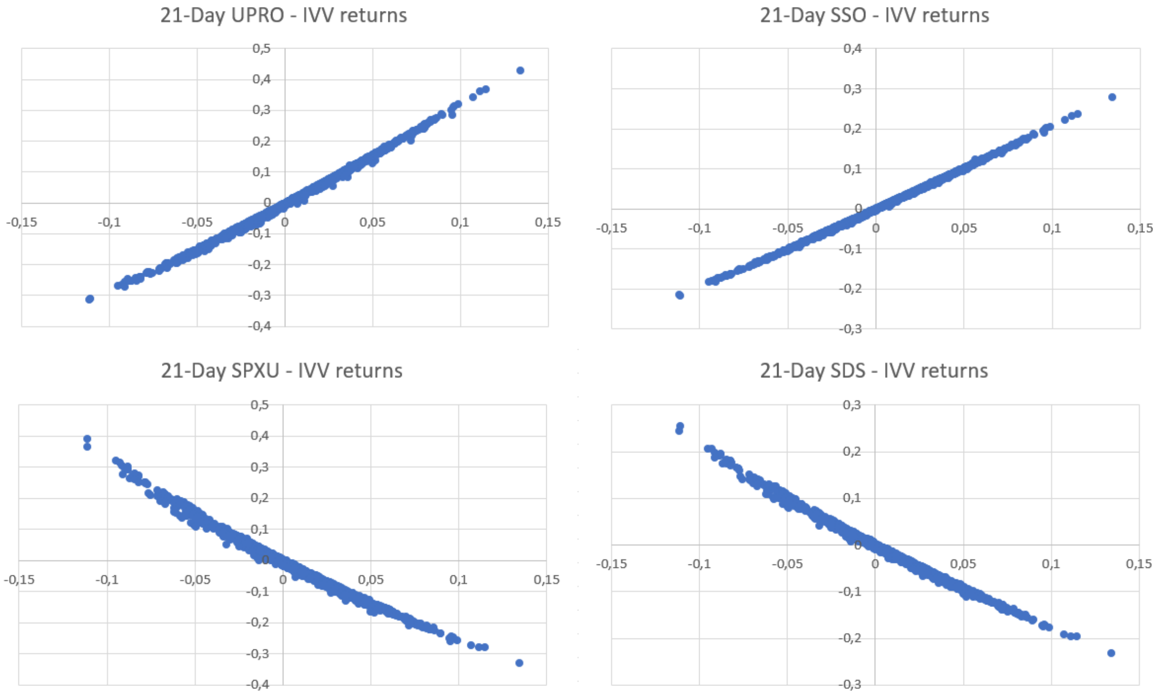

Figure 15: 21-day returns of S&P500 ETFs

As we can see from the graphs, the returns are beginning to form into a rather convex shape, especially in the case of inverse "SPXU" and "SDS" ETFs. Same trend applies to the inverse ETFs on fixed-income side:

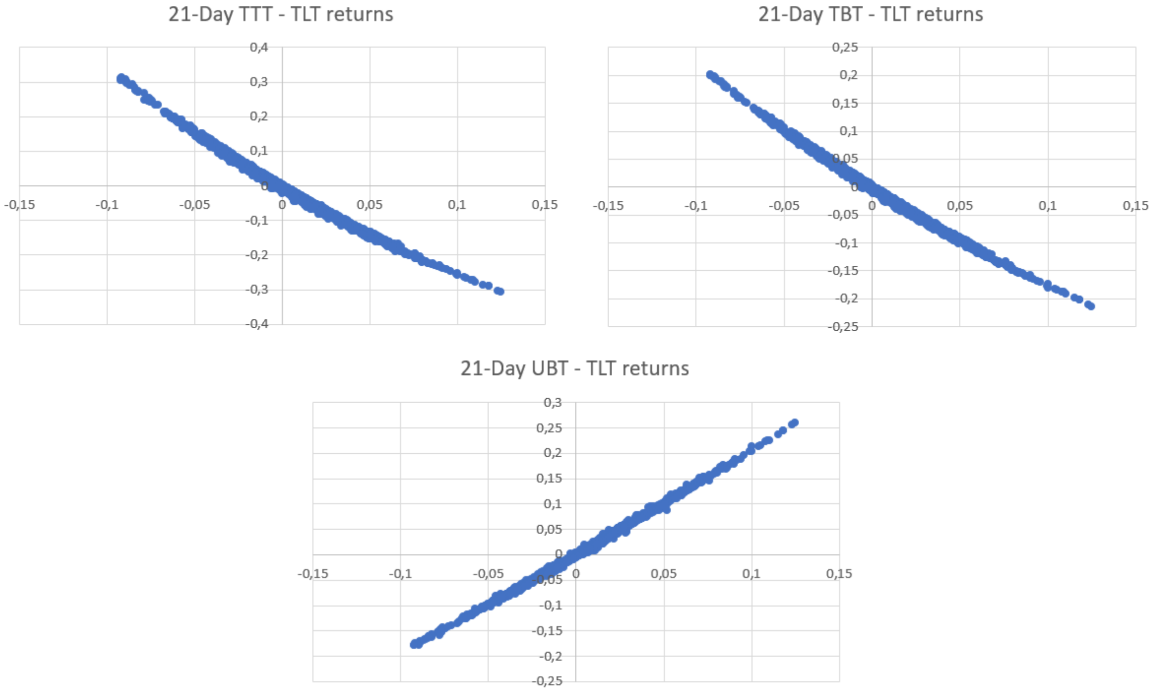

Figure 16: 21-day returns of Barclays U.S. 20+ Year Treasury Bond ETFs

Now, let's consider what the observed convexity would imply from the point of returns. Considering a regression line with expected returns (alpha=0, beta=*k*), points above the line would imply abnormally high returns, while points under the regression line would imply abnormally low returns. For example, the shape of 21-Day TTT – TLT returns seems to turn into a convex curve, which would imply abnormally high profits on the left tail of the curve, while also providing lower than expected losses on the right tail of the curve.

However, a major thing to consider here is the points between these tails. It is possible, that although tails implying abnormally good performance, the majority of observation points may actually be under the regression line, based on the convex shape of the curve.

5.2.3. 1-year and longer performance

Referring to previously presented literature of LETF performance, 1-year holding period for leveraged ETFs is often considered long. This is now the first of 3 long investment periods observed in this thesis. Below is presented 1-year, or 252-day performance of S&P500 LETFs.

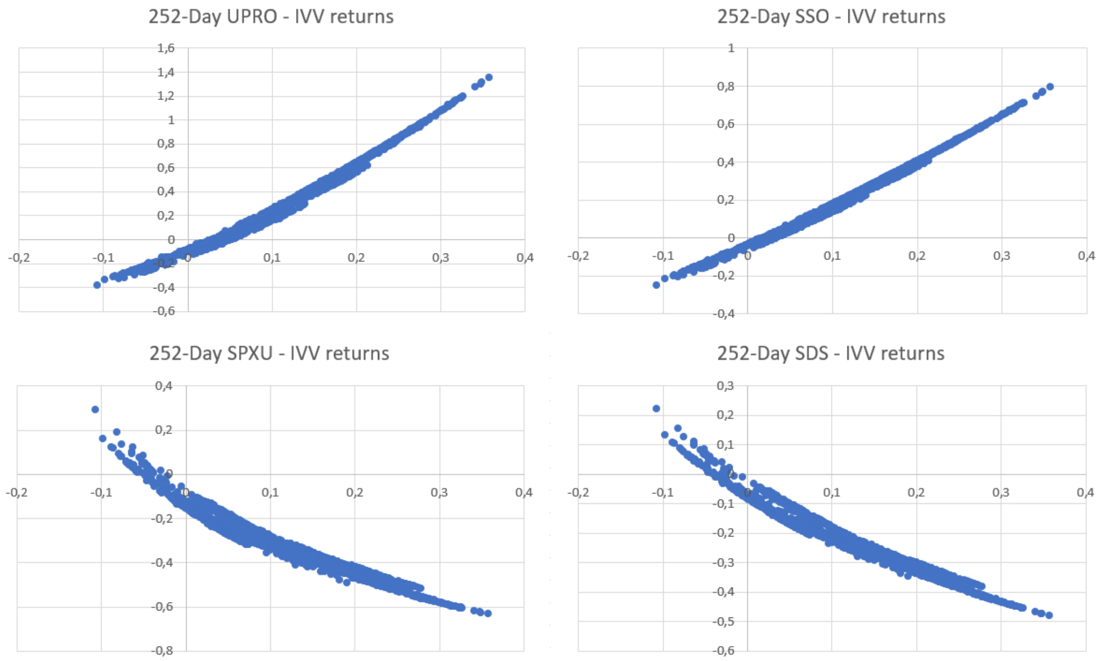

Figure 17: 252-day returns of S&P500 ETFs

The convexity seems to amplify, as the investment period gets longer. The same applies to fixed-income ETFs, as seen below.

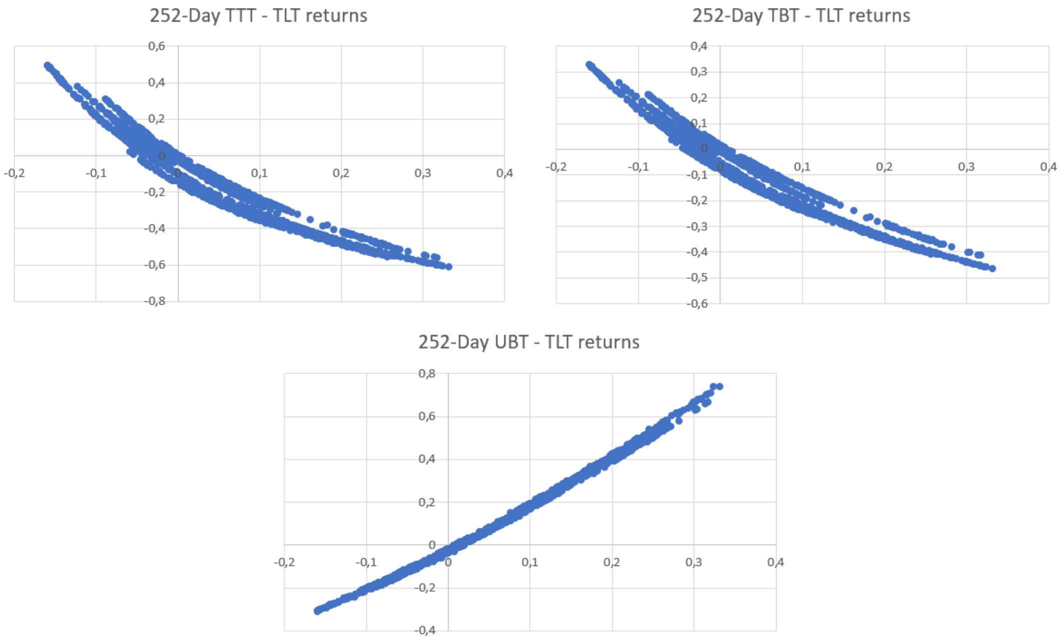

Figure 18: 252-day returns of Barclays U.S. 20+ Year Treasury Bond ETFs

Especially the returns of inverse ETFs in both categories seem to start deviating a lot from the expected regression line. Next, let’s observe 5-year, 1260-day returns of these instruments.

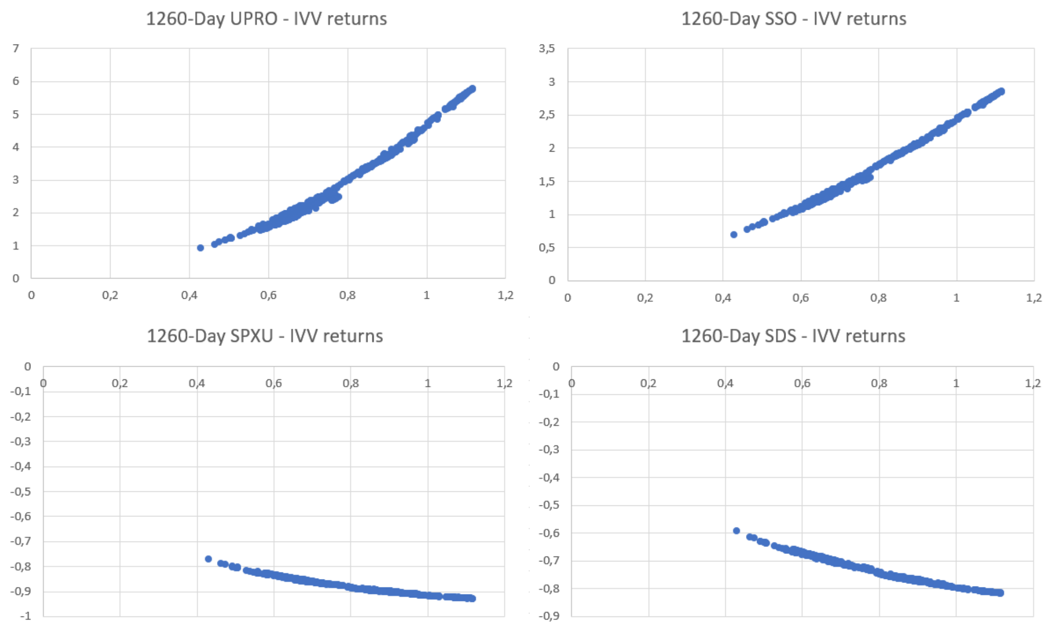

Figure 19: 1260-day returns of S&P500 ETFs

From the positively leveraged ETFs “UPRO” and “SSO”, we can see that there are many periods of performing better than what the naïve expectation would assume. As seen in the data-chapter of this thesis, the overall performance of S&P500 index, based on the benchmark ETF “IVV”, has been good during the last 7 years. This reflects strongly in the inverse ETFs of “SPXU” and “SDS”, as neither of those could have provided any profit on any of the overlapping 5-year investment periods. Next, the performance graphs of 2 other sets of leveraged ETFs are brought in. Following graph represents the performance of ETFs tracking NASDAQ-100 index.

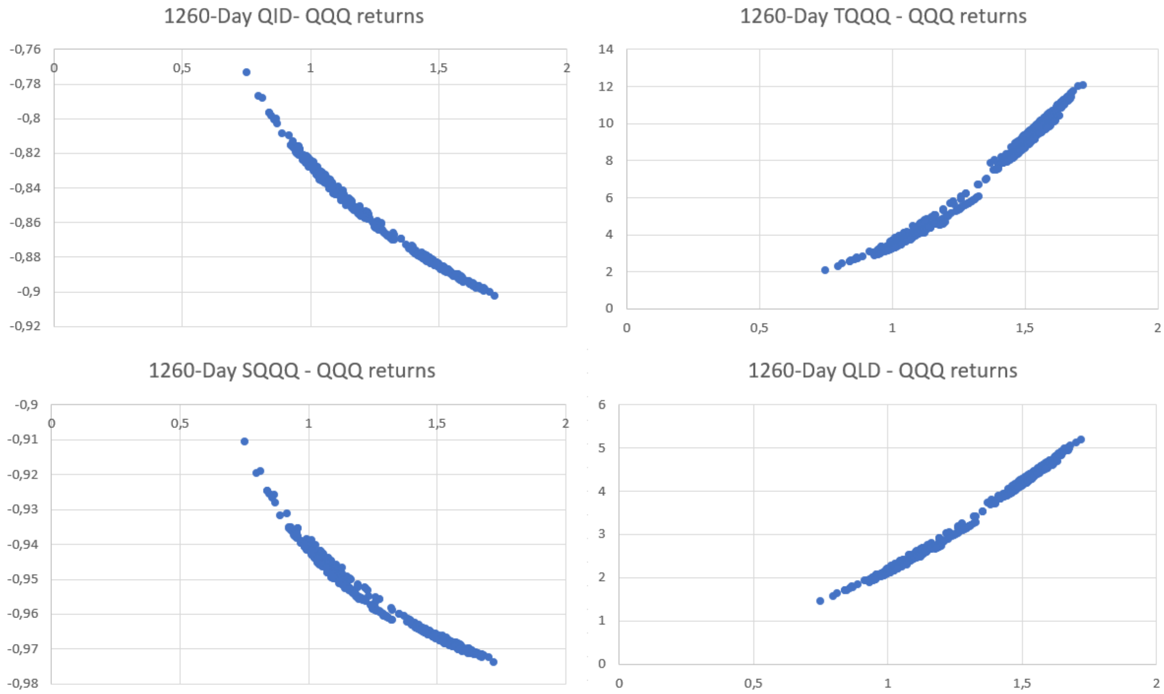

Figure 20: 1260-day returns of NASDAQ-100 ETFs

Especially, performance of “TQQQ” is interesting here. Convex shape shows us that the relationship between returns is not linear. When the returns from investing in the benchmark ETF during this period increase, the returns of 3x-leveraged ETF increase more than they are supposed to, based on the naïve expectation of k-multiplied returns.

Looking at the graphs of inverse ETFs on both set of graphs, an interesting conclusion can be made. Although devastating performance during the observed 5-year periods, there are no k-multiplied losses, compared to the index as we can see that the value of the investments diminishes towards zero. This backs the idea of using inverse ETFs as shorting instruments, as investor’s losses are limited to the amount of invested capital, unlike when practicing actual short-selling or using certain derivatives.

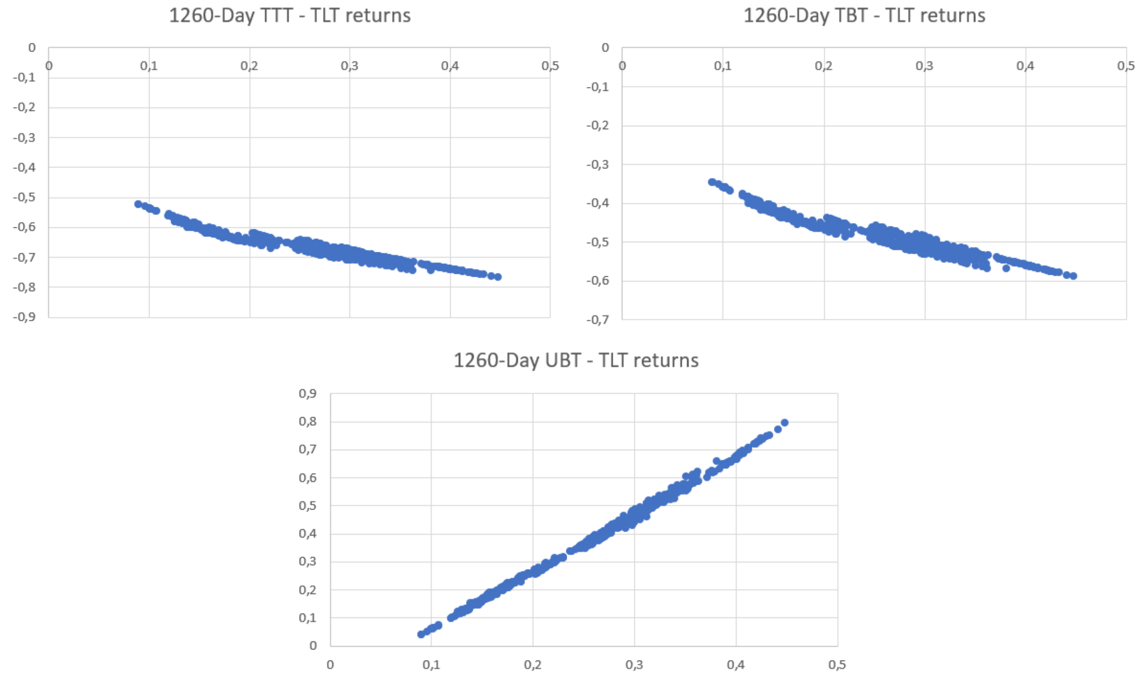

Figure 21: 1260-day returns of Barclays U.S. 20+ Year Treasury Bond ETFs

Above is presented the corresponding graph of the fixed-income ETFs. Performance of the inverse ETFs are rather similar to ones within equity ETFs. Interestingly, the shape of UBT – TLT returns is very linear. At this point, I am assuming that this is very much the shape we would expect in more volatile times. The overall performance of underlying “TLT” during the 7-year period has been positive. However, as seen in the data chapter, it has been more volatile and not as straightforward, as observed equity benchmark ETFs.

In summary, observations of these graphs provide us a strong expectation, that the linear assumption made in the following regression analysis will break at least around 1-year investment period, if not earlier. The convexity seen in the most of the long-term return figures is likely to root from the compounding effect as suggested by Zhang (2010) and Avellaneda et al. (2010b).

5.3. K-times returns

This subchapter presents the results from tests conducted as presented in the methodology chapter. To not only rely on vague estimation by looking at the plotted data from different investment periods, I provide results from regression analyses conducted between every combination of LETF-ETF pairs, for every investment period defined in methodology. Relevant numbers of results from these analyses are provided in tables, one investment period at time. After presenting the results in tables, graphs of the results are presented for easier interpretation and conclusions.

Every single of the LETFs is analyzed in their own regression analysis. A table combines the results of these analyses from 1 investment period at a time. Each table presents 7 numbers per regression. First is presented R-squared, which tells us how much of the variance of the sample is predicted by the model. Often considered as goodness of fit, the closer the value of R-squared is to 1, the better the data should fit the model in terms of variance.

Alpha and beta are the estimates from the regression analysis. All of the presented estimates are statistically significant with a p-value of <0.05. Standard errors of estimates are presented below them in parenthesis. Because of the studied data being overlapping, standard errors are calculated as Newey-West standard errors. “Expected” column presents the excepted beta, which is based on the leverage-multiplier of corresponding LETF. Last column, “Error margin”, represents the coefficient beta’s proportional deviation from its excepted value. I have defined an acceptable error margin, where we can reliably assume k-multiplied returns from that period as 5 percentage. Where error margins are larger than 0.05, the particular

LETF can not reliably provide k-times returns on that period. The error margin E is calculated as absolute value of percentual difference between the realized (β_r) and expected beta (β_e):

$$E = \left|\frac{\beta_r}{\beta_e} - 1\right| \tag{23}$$

Below is presented the table of results from 3-day overlapping investment periods from the observed timeframe:

Table 7: Regression analysis results: 3-day investment periods

TICKER	R-Squared	Alpha	Beta	Expected	Error margin
SSO	0,9991	-0,0003	1,9921	2,000	0,004
		(0,0000)	*(-0,0034)*		
UPRO	0,9985	-0,0004	2,9856	3,000	0,005
		(0,0000)	*(-0,0109)*		
SDS	0,9977	0,0002	-1,9987	-2,000	0,001
		(0,0000)	*(-0,0087)*		
SPXU	0,9965	0,0002	-2,9974	-3,000	0,001
		(-0,0001)	*(-0,0186)*		
QLD	0,9992	-0,0002	1,9922	2,000	0,004
		(0,0000)	*(-0,0039)*		
TQQQ	0,9984	-0,0004	2,9810	3,000	0,006
		(-0,0001)	*(-0,0107)*		
QID	0,9973	0,0001	-1,9929	-2,000	0,004
		(-0,0001)	*(-0,0092)*		
SQQQ	0,9957	0,0001	-2,9901	-3,000	0,003
		(-0,0001)	*(-0,0195)*		
UYG	0,9983	-0,0002	1,9876	2,000	0,006
		(0,0000)	*(-0,0027)*		
FINU	0,9016	-0,0006	3,1212	3,000	0,040
		(-0,0005)	*(-0,0848)*		
SKF	0,9953	0,0000	-1,9877	-2,000	0,006
		(-0,0001)	*(-0,0087)*		
FINZ	0,7559	-0,0003	-2,8728	-3,000	0,042
		(-0,0007)	*(-0,0732)*		

UBT	0,9889	-0,0002	1,9847	2,000	0,008
		(0,0000)	*(-0,0053)*		
TBT	0,9977	0,0000	-1,9841	-2,000	0,008
		(0,0000)	*(-0,0048)*		
TTT	0,9962	0,0000	-2,9778	-3,000	0,007
		(-0,0001)	*(-0,0087)*		
UST	0,9915	-0,0002	1,9838	2,000	0,008
		(0,0000)	*(-0,0043)*		
PST	0,9901	0,0000	-1,9842	-2,000	0,008
		(0,0000)	*(-0,0055)*		

We can see from table 7, that all of the alphas are very close to 0, as well as betas are close to their expected values of *k*. Two ETFs to stand out in these results are “FINU” and “FINZ”, which are 3x and -3x leveraged ETFs of Dow Jones U.S. Financials Index. Although still staying under the critical 0.05 limit of the error margin, they seem to have trouble to match k-times returns. However, this phenomenon has an explanation out of these results: Both of these LETFs were introduced at the beginning of the observed timeframe (7/2012). Very small trading volume or days of no trading at all shortly after the introduction has caused an actual tracking error of 1-day performance, which now reflects to the 3-day performance also.

Generally, we can say that on 3-day investment periods, the leveraged ETFs can reliably provide k-multiplied returns. In the next table is presented results of regression analyses for 21-day period.

Table 8: Regression analysis results: 21-day investment periods

TICKER	R-Squared	Alpha	Beta	Expected	Error margin
SSO	0,9991	-0,0023	2,0210	2,000	0,011
		(0,0001)	*(0,0055)*		
UPRO	0,9972	-0,0045	3,0594	3,000	0,020
		(0,0004)	*(0,0164)*		
SDS	0,9933	-0,0004	-1,9237	-2,000	0,038
		(0,0004)	*(0,0177)*		
SPXU	0,9894	-0,0017	-2,8392	-3,000	0,054
		(0,0008)	*(0,0331)*		
QLD	0,9988	-0,0026	2,0304	2,000	0,015
		(0,0002)	*(0,0067)*		
TQQQ	0,9959	-0,0054	3,0851	3,000	0,028
		(0,0006)	*(0,0194)*		
QID	0,9914	-0,0025	-1,8891	-2,000	0,055
		(0,0006)	*(0,0180)*		
SQQQ	0,9850	-0,0055	-2,7710	-3,000	0,076
		(0,0011)	*(0,0355)*		
UYG	0,9987	-0,0017	2,0155	2,000	0,008
		(0,0002)	*(0,0084)*		
FINU	0,9373	-0,0074	3,4116	3,000	0,137
		(0,0021)	*(0,0939)*		
SKF	0,9908	-0,0012	-1,9404	-2,000	0,030
		(0,0006)	*(0,0267)*		
FINZ	0,8998	0,0002	-3,1226	-3,000	0,041
		(0,0027)	*(0,0837)*		
UBT	0,9977	-0,0013	2,0067	2,000	0,003
		(0,0001)	*(0,0073)*		
TBT	0,9932	0,0001	-1,9586	-2,000	0,021
		(0,0004)	*(0,0215)*		
TTT	0,9895	-0,0001	-2,9197	-3,000	0,027
		(-0,0007)	*(0,0415)*		
UST	0,9979	-0,0012	1,9959	2,000	0,002
		(0,0001)	*(0,0046)*		
PST	0,9914	0,0002	-1,9614	-2,000	0,019
		(0,0002)	*(0,0141)*		

On a 1-month investment period, betas of 4 LETFs deviate critically from their expected values. Interestingly, all of these are either inverse, or 3x-leveraged. LETFs with a positive leverage of 2, still seem to hold on to their expected returns. Next, let’s observe results of 3-

month investment period. Previous studies have often considered 3-month period to be too long to reliably provide k-times returns.

Table 9: Regression analysis results: 63-day investment periods

TICKER	R-Squared	Alpha	Beta	Expected	Error margin
SSO	0,9975	-0,0082	2,0474	2,000	0,024
		(0,0007)	*(0,0129)*		
UPRO	0,9919	-0,0170	3,1357	3,000	0,045
		(0,0019)	*(0,0403)*		
SDS	0,9781	-0,0045	-1,8327	-2,000	0,084
		(0,0022)	*(0,0480)*		
SPXU	0,9659	-0,0116	-2,6451	-3,000	0,118
		(0,0041)	*(0,0892)*		
QLD	0,9967	-0,0097	2,0687	2,000	0,034
		(0,0010)	*(0,0153)*		
TQQQ	0,9887	-0,0218	3,1924	3,000	0,064
		(0,0029)	*(0,0473)*		
QID	0,9753	-0,1311	-1,7463	-2,000	0,127
		(0,0028)	*(0,0464)*		
SQQQ	0,9582	-0,0281	-2,4692	-3,000	0,177
		(0,0051)	*(0,0846)*		
UYG	0,9973	-0,0070	2,0563	2,000	0,028
		(0,0007)	*(0,0148)*		
FINU	0,9072	-0,0304	3,5813	3,000	0,194
		(0,0048)	*(0,1277)*		
SKF	0,9775	-0,0084	-1,8301	-2,000	0,085
		(0,0022)	*(0,0436)*		
FINZ	0,9027	-0,0094	-2,9156	-3,000	0,028
		(0,0061)	*(0,1153)*		
UBT	0,9978	-0,0037	1,9927	2,000	0,004
		(0,0004)	*(0,0102)*		
TBT	0,9835	0,0006	-1,9673	-2,000	0,016
		(0,0012)	*(0,0311)*		
TTT	0,9750	0,0008	-2,9439	-3,000	0,019
		(-0,0022)	*(0,0611)*		
UST	0,9977	-0,0035	1,9694	2,000	0,015
		(0,0002)	*(0,0063)*		
PST	0,9769	0,0004	-1,9306	-2,000	0,035
		(0,0006)	*(0,0256)*		

These results of table 9 are interesting. If considering a deviation of under 5 percentage from the expected reasonable, some of the observed LETFs are still able to track the underlying ETF. Smallest error margins of 0.004, 0.015, 0.016 and 0,019 are all from fixed-income LETFs. Largest errors seem to be within inverse equity ETFs.

At this point is good to note, that the error margin is not necessarily a bad thing when considering the overall returns from that particular investment period. Considering the graphical presentation in the data chapter and 5.2., the equity LETFs have actually performed really well, and it is likely that these error margins represent rather positively abnormal returns. For example, looking at the graphs of overall performance over the whole 7-year period, equity LETFs seem to have performed better than the fixed-income LETFs. However, based on these regression analyses, the fixed-income LETFs follow a more linear relationship against their benchmark ETFs, and can so provide k-times returns more reliably.

Next is presented a similar table of results from regression analyses from 1 year overlapping investment periods. As discussed, 3-months already being a “long” investment period among leveraged ETFs, 1-year period could be considered “unsuitable” for this type of instruments.

Table 10: Regression analysis results: 252-day investment periods

TICKER	R-Squared	Alpha	Beta	Expected	Error margin
SSO	0,9934	-0,0551	2,2612	2,000	0,131
		(0,0031)	*(0,0186)*		
UPRO	0,9831	-0,1343	3,7814	3,000	0,260
		(0,0085)	*(0,0542)*		
SDS	0,9638	-0,0627	-1,2969	-2,000	0,352
		(0,0046)	*(0,0268)*		
SPXU	0,9529	-0,1290	-1,6098	-3,000	0,463
		(0,0069)	*(0,0416)*		
QLD	0,9935	-0,0796	2,3676	2,000	0,184
		(0,0040)	*(0,0169)*		
TQQQ	0,9816	-0,2116	4,1127	3,000	0,371
		(0,0114)	*(0,0501)*		
QID	0,9800	-0,1236	-1,0573	-2,000	0,471
		(0,0032)	*(0,0142)*		
SQQQ	0,9655	-0,2325	-1,2017	-3,000	0,599
		(0,0052)	*(0,0223)*		
UYG	0,9949	-0,0474	2,2568	2,000	0,128
		(0,0029)	*(0,0184)*		
FINU	0,9353	-0,2005	4,5384	3,000	0,513
		(0,0185)	*(0,1271)*		
SKF	0,9555	-0,0735	-1,3094	-2,000	0,345
		(0,0056)	*(0,0328)*		
FINZ	0,8682	-0,1077	-1,9626	-3,000	0,346
		(0,0139)	*(0,0809)*		
UBT	0,9947	-0,0227	2,1036	2,000	0,052
		(0,0014)	*(0,0195)*		
TBT	0,9448	-0,0208	-1,6192	-2,000	0,190
		(0,0036)	*(0,0370)*		
TTT	0,9232	-0,0426	-2,2345	-3,000	0,255
		(0,0062)	*(0,0648)*		
UST	0,9940	-0,0149	1,9940	2,000	0,003
		(0,0006)	*(0,0139)*		
PST	0,9090	-0,0033	-1,7796	-2,000	0,110
		(0,0022)	*(0,0582)*		

These results are well in line with previous studies. Almost none of the observed LETF betas are under the defined error margin. In addition, the alphas presenting intercepts are getting further from 0. This was expected, based on the graphical analysis of chapter 5.2.

However, as in table of 3-month results, the fixed-income LETFs still stand out from the 1-year results. 2x-leveraged “UBT” and “UST” have notably lower error margins, than other products. Especially the beta of 1.9940 of “UST” is very close to its expected value of 2. Although there begins to be concerns of reliability and suitability of the model in periods this long, based on the increasing standard errors and alpha deviating further from zero, it is reasonable to analyze these two instruments on even longer, 2-year investment period. Following table presents the results from regression analyses of 2-year investment periods, on these two fixed-income LETFs.

Table 11: Regression analysis results: 504-day investment periods

TICKER	R-Squared	Alpha	Beta	Expected	Error margin
UBT	0,9953	-0,0540	2,1054	2,000	0,053
		(0,0019)	*(0,0132)*		
UST	0,9870	-0,0326	2,0603	2,000	0,030
		(0,0009)	*(0,0165)*		

Surprisingly, even on 2-year investment periods, the error margins are still relatively low. However, the estimate of alpha gets even further from 0 as we lengthen the investment period. Although being more volatile during the whole 7-year period, fixed-income LETFs seem to provide k-times returns more reliably than equity LETFs. However, as seen in the data chapter, the overall performance of equity LETFs has been better during the observation period.

To get a summary and better overall view of the results presented in the tables above, a graphical presentation of the error margins calculated from the regression results is provided. The graphical representation helps to conclude and better understand the amount of error margins between different types of LETFs studied in these analyses.

Below is presented error margins of all studied LETFs and investment periods.

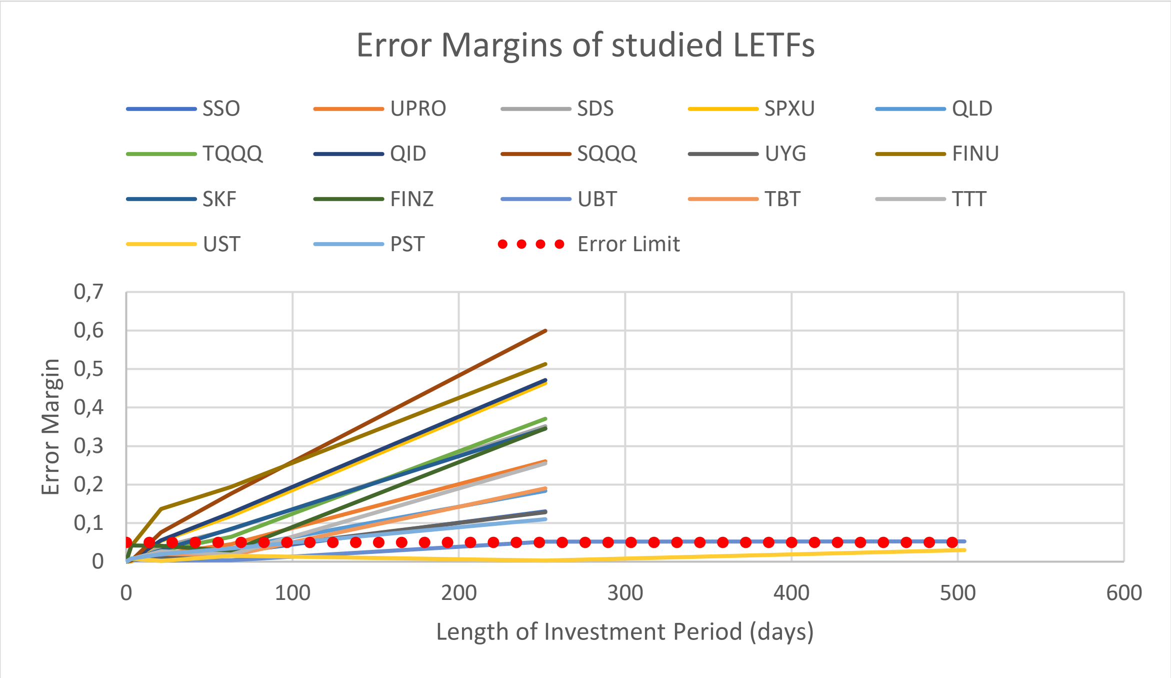

Figure 22: Error margins of studied LETFs

The overall graph of all studied ETFs won't tell us much. From this, we can see that the fixed-income "UST" and "UBT" stay under the dotted line of error margin 0.05 while the others pass it on rather early stage. Let's drill down to observe the error margins by groups. Following graphs represent the error margins of 3x and -3x-leveraged ETFs. The investment period is limited to the 3-month period, as it is enough to provide the information needed and helps the interpretation of the graphs.

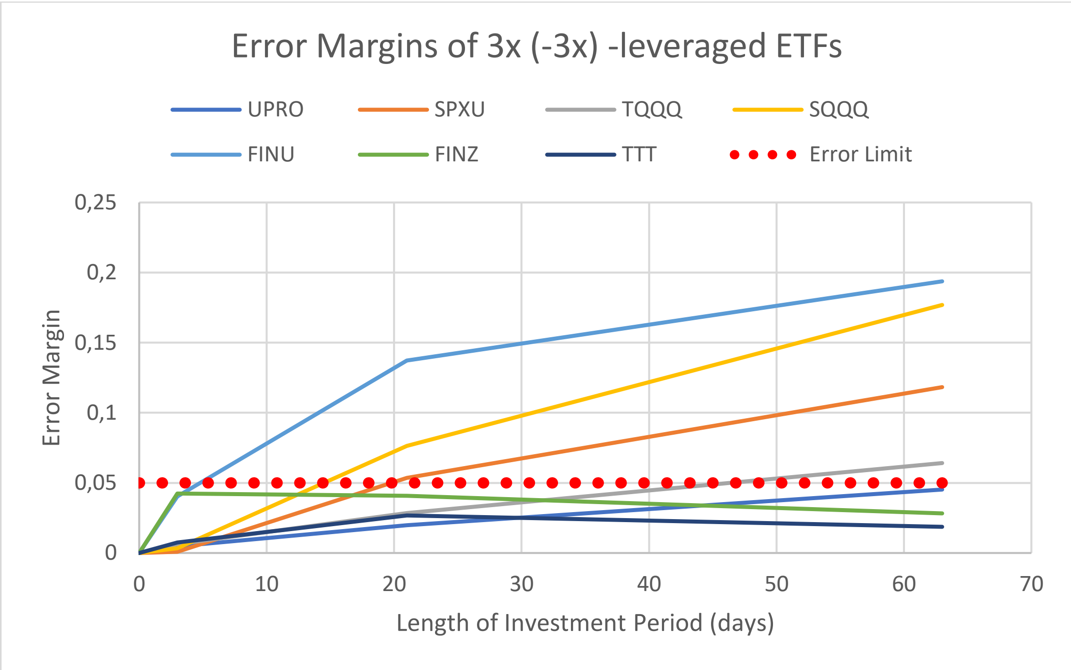

Figure 23: Error margins of 3x (-3x) -leveraged ETFs

Error margins of 2x and -2x-leveraged ETFs for comparison:

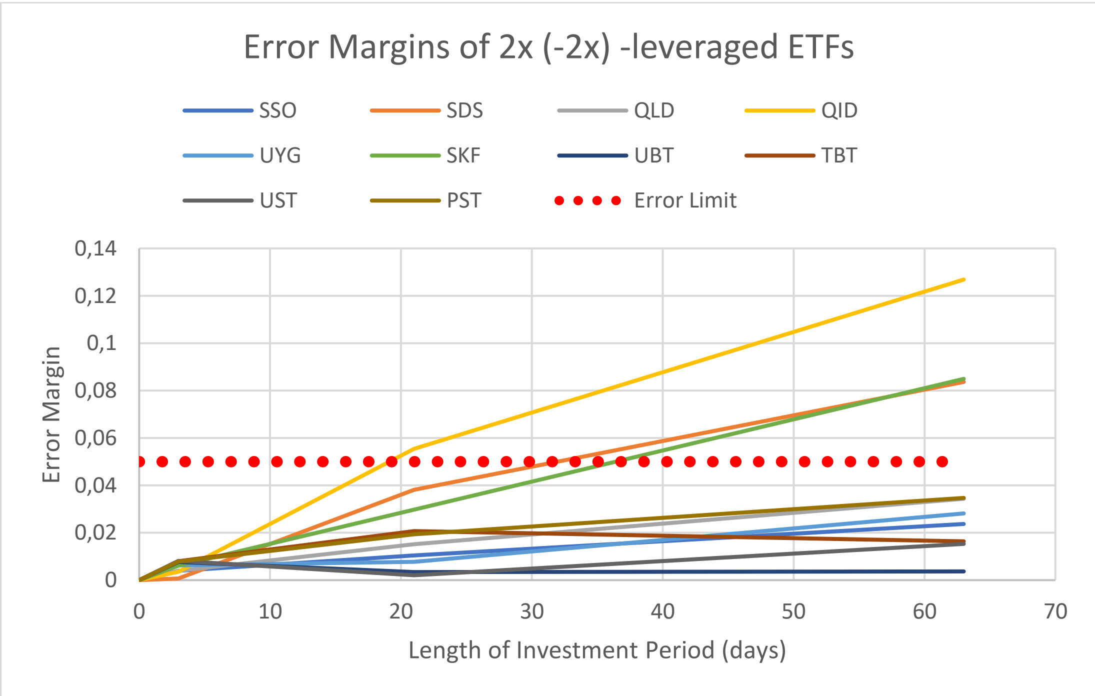

Figure 24: Error margins of 2x (-2x) -leveraged ETFs

Looking at graph of 2x-leveraged ETFs, we can see that most of them stay under the dotted line of error limit, meaning that they would probably match their k-multiplied returns on investment periods less than 63 days (3 months). Three lines representing “QID”, “SDS” and “SKF” are all inverse, -2x-leveraged ETFs. We can conclude that the positively leveraged hold better on k-times returns than negatively leveraged. It seems to be, that the 2x (-2x)-leveraged ETFs are also more likely to hold on k-times returns than 3x (-3x)-leveraged ETFs.

Another pair to compare is the differences between error margins on positively leveraged ETFs and inverse LETFs. Below is presented a graph of error margins on positively leveraged ETFs, again up to 3-month investment periods.

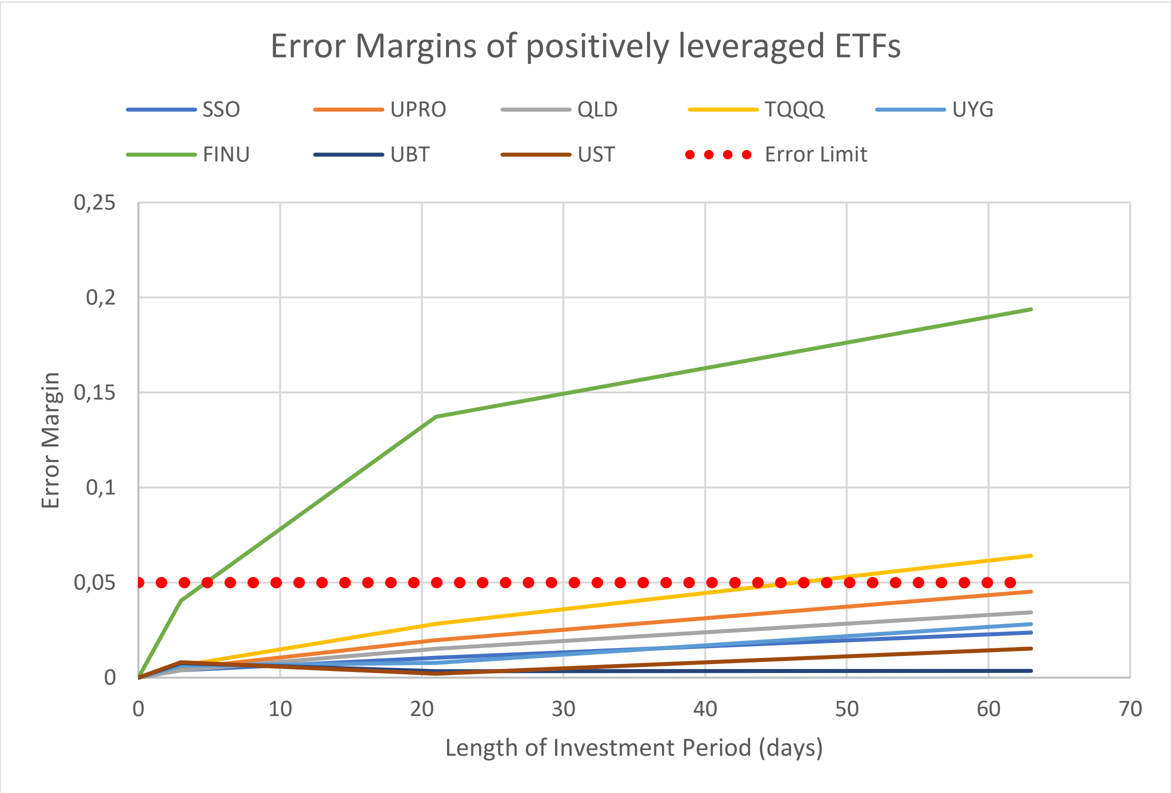

Figure 25: Error margins of positively leveraged ETFs

Inverse, "negatively" leveraged ETFs for comparison:

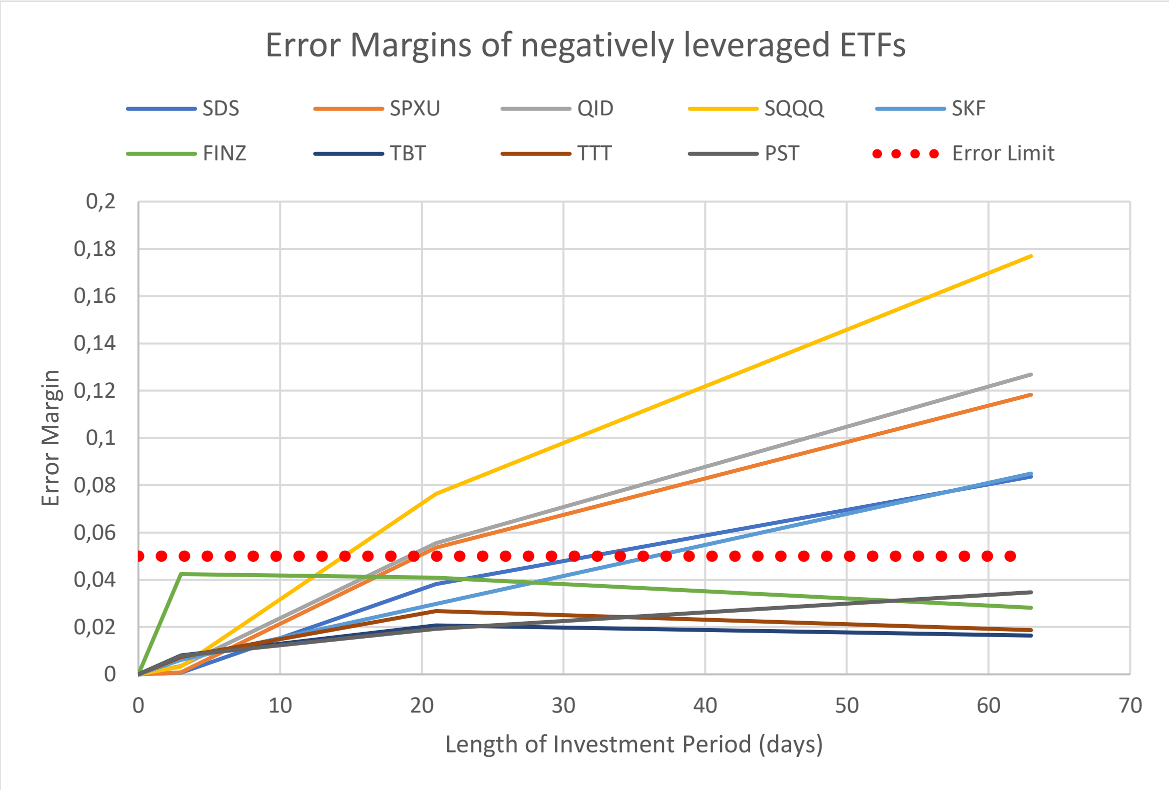

Figure 26: Error margins of negatively leveraged ETFs

Difference between these two groups is remarkable. Looking at graph of ETFs with positive leverage, we can see good matching of expected returns, except for the 3x-leveraged "FINU" which already suffered from actual tracking error on 1-day basis. It is also observable, that at 63-day period, the 3x-leveraged "FINU", "TQQQ" and "UPRO" have higher error margins than other, 2x-leveraged products.

Looking at the graph of ETFs with negative leverage, it is remarkable that except for "FINZ", all the other products staying under the critical error limit are fixed-income ETFs. At this point, it is good to make a comparison of error margins between equity and fixed-income ETFs.

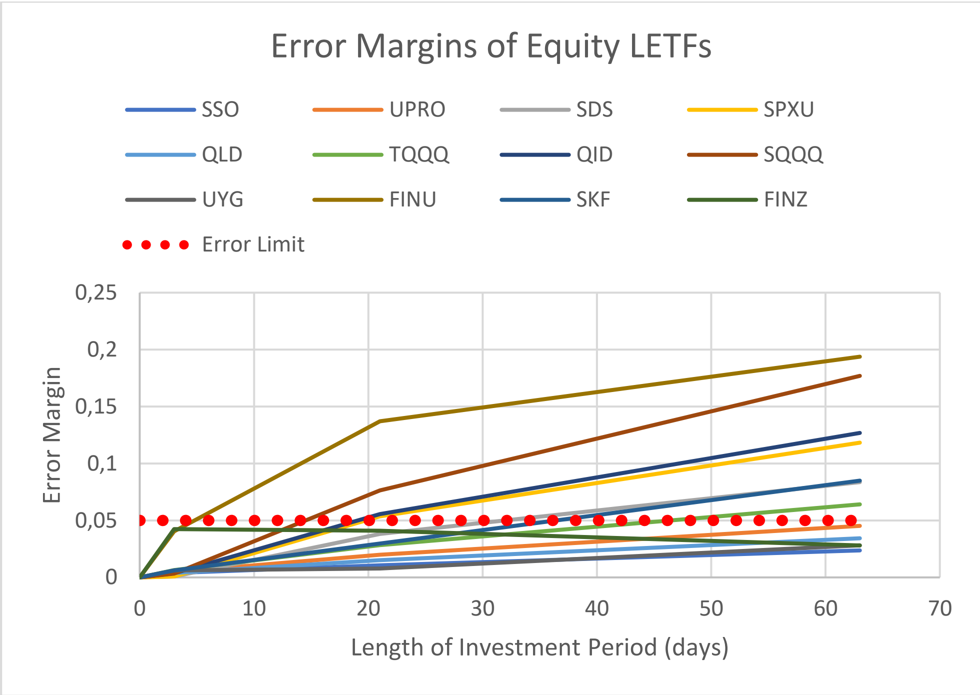

Figure 27: Error margins of equity LETFs

Error margins of fixed-income LETFs:

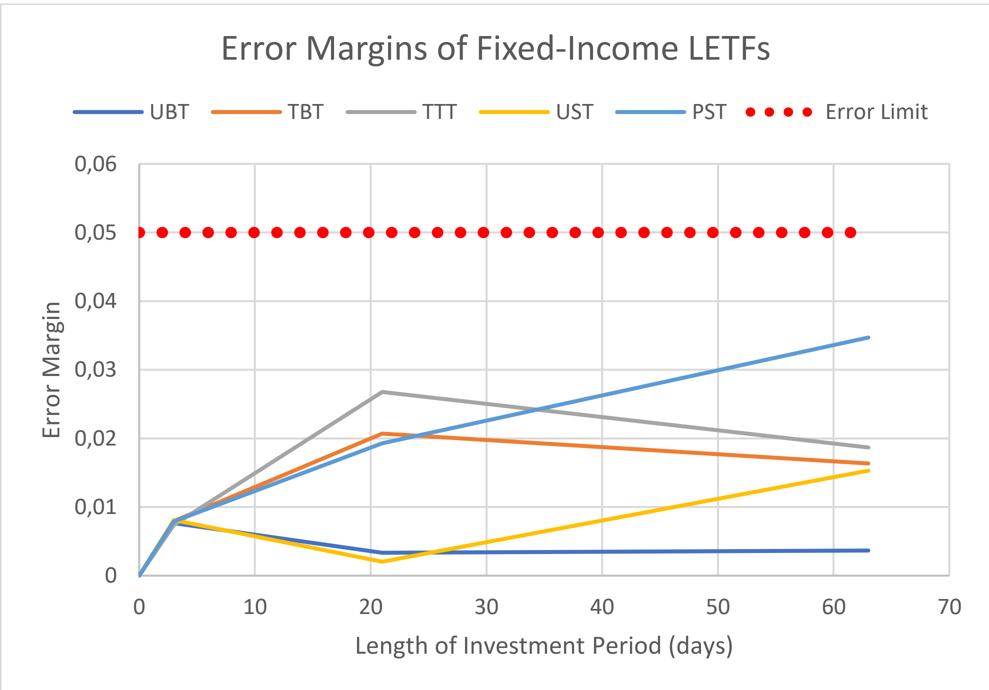

Figure 28: Error margins of fixed-income LETFs

This comparison provides largest difference between the groups on investment periods of 3-months and less. All of the observed fixed-income LETFs' error margins stay under the critical limit of 0.05 on these periods. Looking at the equity LETFs, similar conclusions can be made, as was with previous comparisons: In the case of equity LETFs, the products to hold the k-multiplied returns on less than 3-month investment periods, are the ones with positive 2x-leverage. 3x-leveraged products, either positive or inverse, seem to drift faster from the expected returns, as well as -2x-leveraged inverse ETFs. As seen in the results, this is also the case within fixed-income ETFs, as the inverse ETFs are less likely to hold on to the expected returns, although they seem to hold for longer, over 3-month periods, unlike inverse equity LETFs.

5.4. Risk-adjusted returns – Abnormal returns and effects of volatility

As we observed from the regression results, fixed-income LETFs were more likely to hold on to k-multiplied returns. However, the overall performance of equity LETFs, more notably 3x-leveraged ETFs have been better, as seen from the plot graphs on chapter 8.2. as well as seen from on the data-chapter. The following tests have been conducted in attempt to find out confirmation and reliability to the results and conclusions of regression analysis results.

By calculating risk-adjusted returns based on the volatility of LETFs as presented in chapter 4.3., these results are provided in attempt to find out, whether leveraged ETF provide return and risk in the same ratio as their non-leveraged benchmark ETFs. As well as when observing the performance based on naïve expectation in previous chapter, the observing of risk-adjusted returns is done with different investment periods. Table below presents the results from t-tests for every LETF risk-adjusted returns against their benchmark ETF risk-adjusted returns, on 3-, 21-, 63- and 252-day periods.

Table 12: T-test results of risk-adjusted returns

TICKER	t Stat 3d	t Stat 21d	t Stat 63d	t Stat 252d
SSO	-0,22466	0,884127	1,599224	1,6834405
UPRO	-0,06747	0,944849	1,840735	1,2202482
SDS	-0,20055	-0,63538	-1,1235	-5,559769
SPXU	0,965677	-0,51984	-0,99391	-7,064893
QLD	0,076009	0,529853	1,06382	-0,073201
TQQQ	-0,38977	0,539075	1,123723	-1,534004
QID	0,233465	-0,06749	-0,33023	-7,779288
SQQQ	0,175272	-0,09958	-0,44985	-10,71797
UYG	0,645738	0,484478	1,022447	1,0355575
FINU	0,696631	2,206422	3,541515	2,3534181
SKF	-1,14046	0,077676	0,496997	-2,33601
FINZ	-1,6868	-1,33039	-0,95683	-5,783742
UBT	0,441168	0,648323	1,094881	2,1323726
TBT	-0,76394	-0,14289	-0,3515	-0,245629
TTT	-0,08499	-0,09978	-0,30722	-0,07717
UST	0,328825	1,35625	2,273727	4,1073847
PST	-0,54176	-0,57627	-0,87268	-1,408028

In table 12 is presented t-stat values of the t-tests. With alpha of 0.05, we can say that there is no statistical difference between the two groups, if the presented t-stat value is between -1.96 and 1.96, or the absolute value of t-stat is under 1.96. Again, for better interpretation of the results, below is presented a graphical comparison of t-stats between positively and inverse leveraged ETFs. LETFs which provide risk and return in the same ratio, should have a t-stat value between the limits of -1.96 and 1.96.

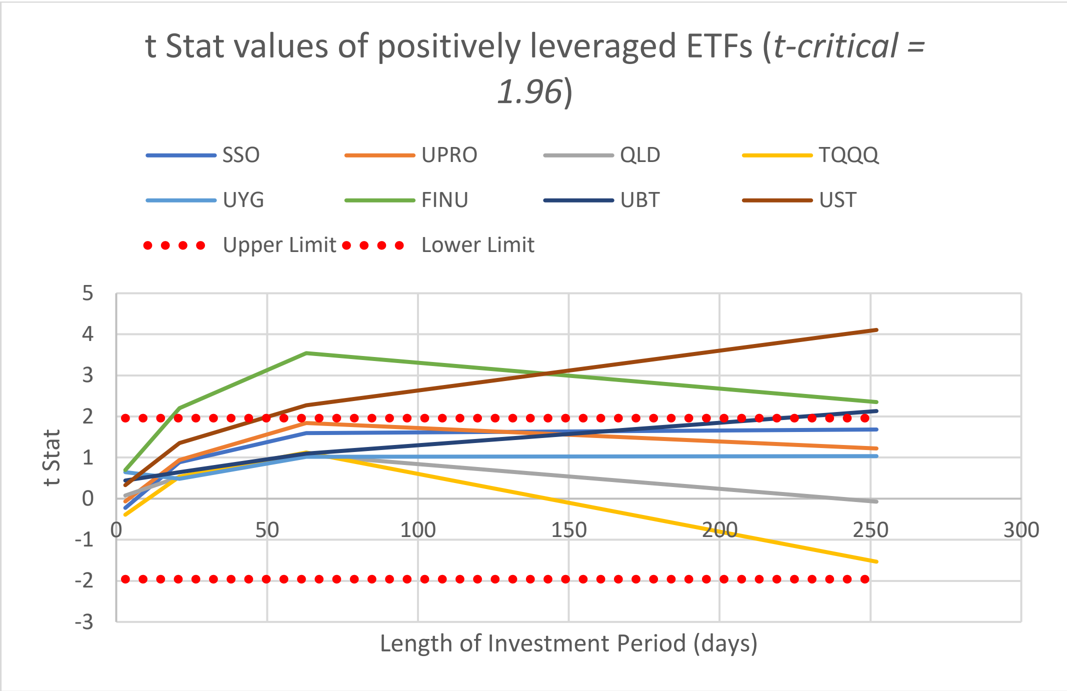

Figure 29: T-stat values of positively leveraged ETFs

Negatively leveraged for comparison:

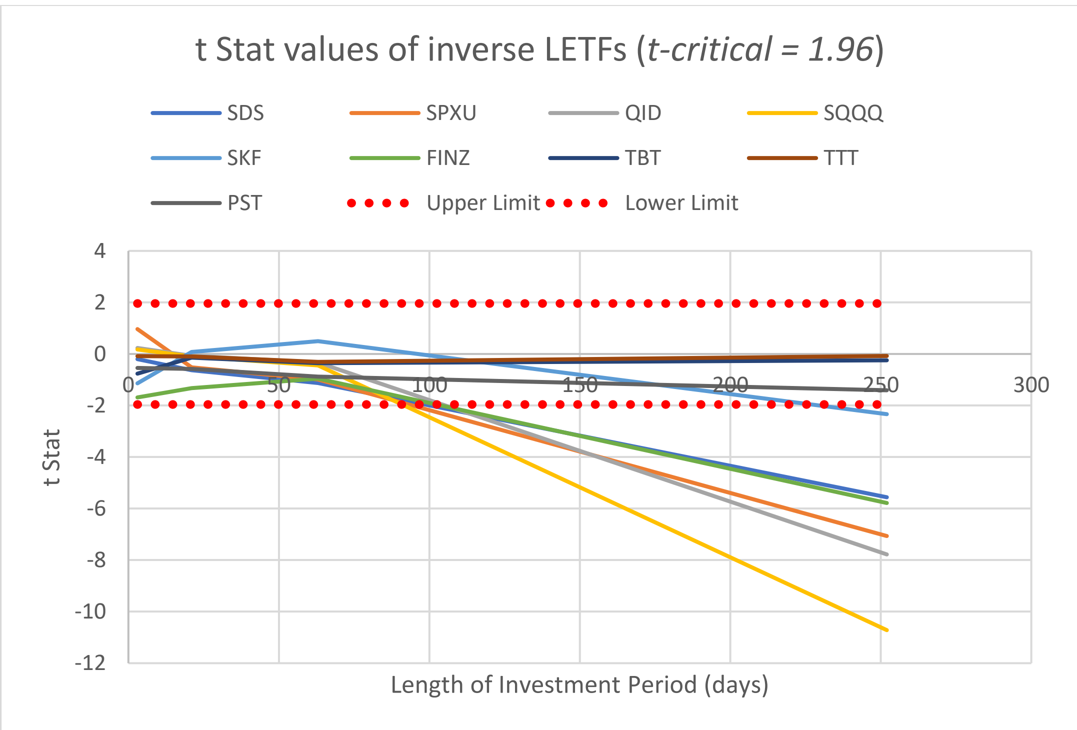

Figure 30: T-stat values of inverse LETFs

Interestingly, although 2x-leveraged fixed-income ETFs “UBT” and “UST” seemed to hold best on the k-multiplier returns when analyzed with beta, in the t-test analysis their risk-adjusted returns do not equal the risk-adjusted returns of benchmark ETFs. The compounding problem is visible in the case of inverse LETFs. These t-tests help to further interpret the results from regression analyses, as it strengthens the findings that inverse LETFs can not provide returns in same risk-ratio as their benchmarks. In addition, these tests show some weakness in the positively leveraged fixed-income ETFs, which were supposed to track their benchmark most precisely based on the regression analyses.

Lastly, the effects of volatility on the returns of studied LETFs is observed in the following. Theory behind leveraged ETFs and much discussed problems rooting from constant leverage trap and nature of compounding returns suggest, that volatility restricts LETF to perform in a manner that they would provide risk and return in the same ratio as their non-leveraged equivalents. In figure 31 is presented risk-adjusted returns of 2x-leveraged QID against its benchmark ETF QQQ, and volatility of the benchmark ETF, on overlapping 1-year investment periods.

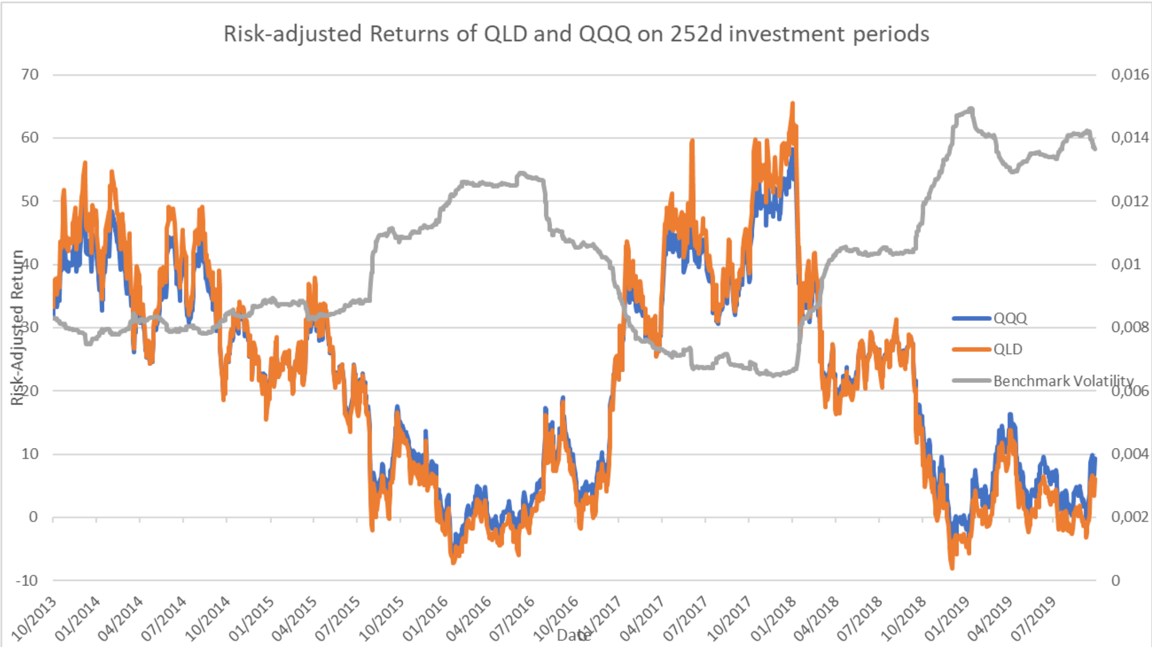

Figure 31: Risk-adjusted returns of QLD and QQQ on 252d investment periods

Notable aspect from this graph is, that abnormal returns (QLD returns over benchmark QQQ) are higher, when the risk-adjusted returns of benchmark are high, and vice versa. In addition, we can see a possible negative effect of volatility on the abnormal returns.

It is well known fact, as also seen in the literature review discussing compounding returns, that volatility has a negative effect on returns of any conventional investment vehicle. However, figure 31 suggests that there is even stronger effect on leveraged ETF returns, as was also suggested by theories presented in the literature review. It is observable from the graph, that in times of high volatility, there are negative abnormal returns, while in times of lower volatility, the leveraged ETF can provide higher returns than the benchmark ETF, in other words positively abnormal returns. From the figure 32 below, it can be seen how risk-adjusted returns of both 2x- and 3x-leveraged ETFs are higher, when volatility is low. In addition, in times of high volatility, 2x-leveraged "QLD" would perform better than 3x-leveraged "TQQQ", when measuring risk-adjusted returns.

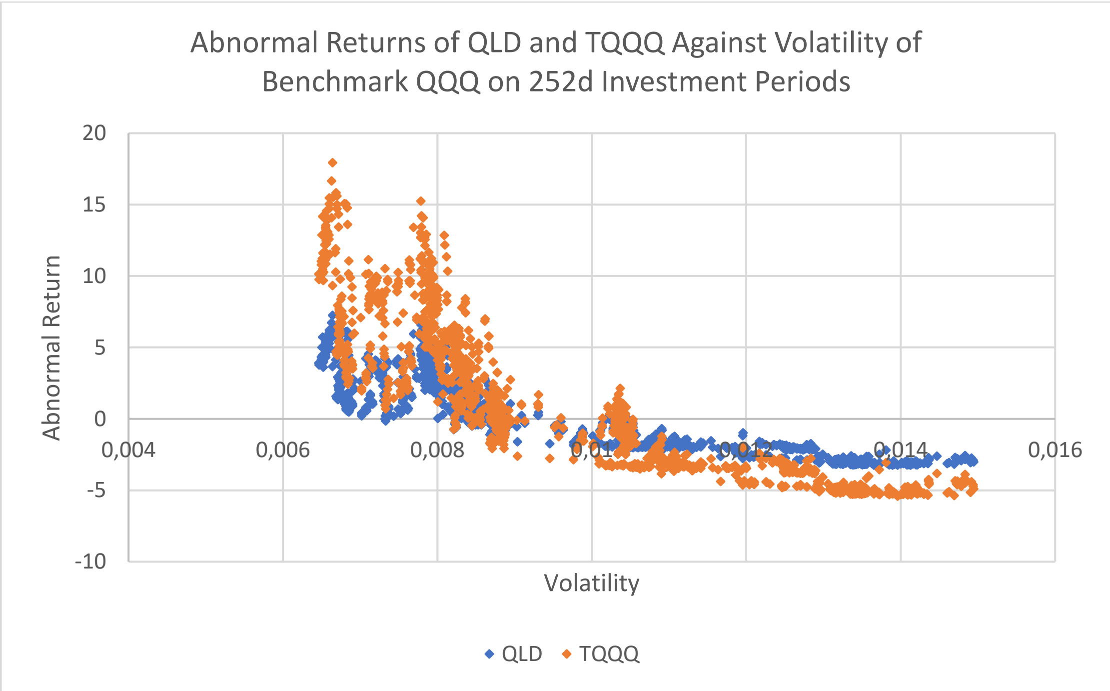

Figure 32: Abnormal returns of QLD and TQQQ against volatility of benchmark on 252d investment periods

Although within instruments tracking NASDAQ-100 index there is clear correlation visible from the graph, it was not the case with all instruments. To find out how volatility affects the abnormal returns, below is presented results of correlation tests between abnormal returns of LETFs and their benchmark ETFs' volatility. Shape of the graph on figure 32 suggests using monotonic Spearman's rank correlation, rather than Pearson correlation, which measures linear correlation between two samples. Period of 1-year (252 days) is used for the analyses as it is long enough to have provided abnormal returns between the samples. For examples, on 3-day periods, any abnormal returns are hardly present as the studied LETFs reach k-multiplier returns on that period very well, meaning that risk and return match. This leads to a fact that correlation coefficients between abnormal returns and benchmark's volatility are very close to zero.

Below is presented a table of Spearman's correlation coefficients between the abnormal returns and volatility of the benchmark indices on overlapping 252-day investment periods. All of the presented values are significant on 95% significance level.

Table 13: Spearman's correlation coefficients between abnormal returns and volatility

TICKER	Spearman's Rank Correlation Coefficient
SSO	-0,8155
UPRO	-0,8525
SDS	-0,8411
SPXU	-0,8676
QLD	-0,9435
TQQQ	-0,9476
QID	-0,8779
SQQQ	-0,8942
UYG	-0,5595
FINU	-0,4174
SKF	-0,6707
FINZ	-0,5690
UBT	0,4633
TBT	-0,6616
TTT	-0,5776
UST	0,1624
PST	-0,8246

On table 13, value of -1 would imply perfect negative correlation between the abnormal returns and volatility and 0 would mean that there is no correlation between the samples. As with studying the performance on different investment periods, also here 2x-leveraged fixed-income LETFs stand out of the rest. Their abnormal returns seem to have positive correlation with the volatility. As there is no theoretical reason for this phenomenon, it is very likely that the sample of these two instruments is not adequate enough, and the positive coefficients should be considered with extreme caution, no matter the acceptable p-values of the coefficients.

6. CONCLUSIONS

This chapter discusses the conclusion of results achieved in the empirical testing and compares them to previous studies. Summary of the study, as well as answering the research questions are in this chapter. In addition, limitations of the thesis and possible further research topics are discussed.

6.1. Answering the research questions

Summary of the results, and conclusions are made based on the defined research questions. The first research question of this thesis was defined as:

Can leveraged ETFs provide k-times returns on longer than 1-day holding periods?

Followed by the second question of:

Are there significant differences in the behavior of returns between groups of LETFs?

These questions were based on the idea that by holding on to the k-multiplied returns the risk-return relationship would stay the same through these periods, equaling the risk-return-ratio of non-leveraged ETFs.

Answering to these questions was done by comparing leveraged equity and fixed-income ETFs to their non-leveraged counterparties, which acted as benchmarks, on investment periods of 3, 21, 63, 252 and 504 days. Empirical analysis was done by calculating returns for each of the chosen LETFs and their benchmark ETFs for the defined investment periods, by using overlapping daily closing prices. Then each of the LETFs were modeled against their benchmark ETF with regression analyses. Deviation from the expected betas was calculated to observe an error margin, which was then used to approximate the LETF’s inability to provide k-multiplied returns on that given investment period.

Not only was an answer found to the first research question, but in addition, differences between types of leveraged ETFs were recognized. Shortly, the answer to the first research question is Yes – Leveraged ETFs can provide k-times returns on longer than 1-day periods. I defined an error margin of 0.05 for the difference between excepted returns and the estimated returns suggested by beta of regression analyses. It came out, that on 21-day investment periods, leveraged ETFs can still quite reliably provide k-times returns. However, there are differences between groups. Firstly, inverse LETFs seem to start drifting away faster from the expected returns, than positively leveraged. Secondly, both positive and inverse 3-times leveraged ETFs start to drift earlier from the expected returns, than 2x-leveraged ETFs. Null-hypothesis can safely be accepted on investment periods up to 21-days on 2x-leveraged, 3x-leveraged and -2x-leveraged ETFs. In case of -3x-leveraged, null-hypothesis is rejected. On 3-month, 63-day investment periods, both 2x- and 3x- positively leveraged ETFs could provide risk and return in the same ratio, accepting the null-hypothesis. In the case of inverse ETFs, the situation is opposite, as null-hypothesis is rejected on 3-month investment periods. Considering these results, the null-hypothesis of research question 2 is rejected, as there are clearly difference in the returns between the groups.

Interestingly, difference of results between fixed-income and equity ETFs is large. Positively leveraged 2x fixed-income ETFs were able to hold on to k-times returns on investment periods even longer than 1 year. This is interesting because during the observed 7 years of trading days, the overall performance of benchmark equity-ETFs was better than the corresponding performance of fixed-income benchmark ETFs. However, it cannot be concluded reliably, that fixed-income LETFs as a group would more likely provide risk and return in the same ratio as their non-leveraged equivalents, than group of leveraged equity ETFs and their benchmark ETFs.

Although t-tests conducted to test significant difference between risk-adjusted returns of LETFs and their benchmark ETFs helped to assess and further evaluate the results of the regression analyses and answer to research questions 1 and 2, they provided information to answer to research question 3, which had a null-hypothesis of:

H0: Risk-adjusted returns between leveraged and non-leveraged ETFs are the same

Based on the t-tests, risk-adjusted returns do not differ significantly at 1-month investment periods, and results are rather mixed on 3-month periods. The most important outcome of these tests, and considering this null-hypothesis, ended up being the value it gave to evaluate the results of regression analyses and validate the methodology used on them, rather than using these t-test results on their own. These t-test results also back up the whole fundamental idea of k-times returns presenting the risk-return ratio, as is done in this thesis.

In their entirety, results are in line with previous studies. On 1-month investment period, investors can expect k-multiplied returns rather reliably, which would imply that the risk-return ratio would be the same as with non-leveraged ETFs on those periods. The used methodology suggests that it could also be possible with 3-month long investment periods, but the type of LETFs used has to be considered, as inverse and 3x-leveraged ETFs are less likely to perform in that expected manner, as seen from the results. The empirical part of this thesis strengthens previous findings with similar results, but also takes account to different type of LETFs, for example finding out that inverse ETFs, at least -3x-leveraged may not be suitable for holding periods of 21-days. As suggested by previous literature, on longer periods, investor should probably go with some of the presented strategies such as shorting inverse LETF pairs. In previous studies, use of other investment strategies than straight investing in LETFs is further recommended in times of high volatility of underlying indices. This is backed up by negative correlation between volatility and LETF returns presented in this study, when additional strategy is not used.

6.2. Limitations

Although differences between k-times returns between groups can be done and reliable conclusions made, minor downside is associated with the used methodology. While reliably pointing out the point of time where the expected linear relationship breaks, as well as the magnitude of estimates, after a certain point of lengthening the investment periods the estimates themselves are not very reliable as the data will not fit the linear model anymore. As discussed in the methodology chapter, although being used in these kinds of studies, the regression model can not reliably interpret the meanings of alpha and beta after the linear relationship of the data breaks. This concern has been presented by Charupat et al. (2014), as they generated more precise and sophisticated model to measure performance of leveraged ETFs. Another limitation of this study is the relatively short data. 7-years is relatively small sample, when starting to consider longer investment periods. Because of short timeframe of availability of data and the need to use overlapping investment periods backs up the idea of using simulation-based returns, as has been done in several other studies. The model used in this study also lacks the ability to indicate the exact length of investment period where k-times returns can still be expected. Using this methodology, it would require senseless number of analyses to iterate the exact period lengths, and that is why the selection of different investment periods is based mostly on previous studies.

A more sophisticated model would be needed to reliably estimate the effects of alpha and beta deviating from their expected value. Methodology of this thesis can assess the direction and magnitude of the deviation, but as an estimate it is not very accurate, as also the alpha starts to deviate from its expected value of zero. The reason why this is problematic is behind the convexity of the returns, as the datapoints will not follow a linear relationship, after the investment period reaches a certain length. Probably another model would be needed to deal with the convex data, if one would like to end up with reliable estimates after the linear relationship breaks.

Another limitation is within the dataset of fixed-income LETFs. There is very limited number of fixed-income LETFs in market, making it difficult to draw conclusions of that group. As there was only two sets of fixed-income LETFs, it was not possible to get distinctive results for them as a group, although different behavior than with equity LETFs was observed.

6.3. Topics for further research

From a viewpoint of an investor, to reliably know whether their expected returns would equal the risk they are taking is surely in a great interest. This will however not take account on the realized returns which should surely be another aspect an investor would care about. Breaking of the linear relationship, and LETF not being able to provide k-multiplied returns is not necessary a bad thing. As seen from plotting the data, there is often a possibility for returns well over the naïve expectation. Even staying under the expected return, LETF can outperform the non-leveraged benchmark. This type of performing over and under the expectations could provide ideas for further studies of this subject. Deviation of beta from expectation could be further studied, to observe whether the deviation from naïve expectation is good or bad thing from the point of returns.

Research question remains open for very-long investment periods. Although there is consensus and theories about the volatility decay effecting negatively on leveraged ETFs on long periods, which this study also backs up, there is not yet enough data to empirically assess the risk and returns of for example 10- or 15-year investment periods. As discussed in the data chapter, the overall performance of LETFs for last 7-years have been tremendous and would possibly suggest that LETFs could be usable for ultra-long investment periods, given a low-volatility investment period.

At the moment, it would be possible to conduct a study for ultra-long research periods by creating simulated data based on the underlying conventional ETFs or indices. As seen in the literature review of this thesis, simulation-based data is used in some previous studies, but ultra-long 10- or even longer investment periods are not assessed. Considering the overall performance of these products on last 7-years, it could be possible that the positive path-dependence could outweigh the volatility decay. In my opinion, testing for ultra-long periods with simulation-based data would be in great interest.

As said, precise effects of volatility on LETF returns is also very interesting topic from the point of view of an investor. Previous literature and findings of volatility in this thesis drive for studying different investment strategies. As presented in the literature review, strategies like shorting leveraged ETF pairs and rebalancing the ETF portfolio are most efficient in times of high volatility. Part of this thesis studying volatility, as well as existing theories, show that volatility has a decaying effect on LETF returns, if no specific strategy is used.

This creates a research gap to assess optimal LETF investment strategies for a given time, based to the volatility of underlying indices, as in times of high volatility one would likely need a specific strategy to maximize returns, and vice versa.

REFERENCES

Avellaneda, M. & Zhang, S. 2010a. Leveraged ETFs: All You Wanted to Know but Were Afraid to Ask. Risk Professional, February 2010, pp.54-60.

Avellaneda, M. & Zhang, S. 2010b. Path-Dependence of Leveraged ETF Returns. SIAM Journal on Financial Mathematics 2010, Vol.1(1), pp.586-603.

Bai, Q., Bond, S., Hatch, B. 2012. The Impact of Leveraged and Inverse ETFs on Underlying Stock Returns. Working paper. University of Cincinnati. Department of Finance.

Bansal, V. K. & Marshall, J. F. 2015. A Tracking error approach to leveraged ETFs: Are they really that bad? Global Finance Journal May 2015, Vol.26, pp.47-63.

Barnhorst, B., Cocozza, C. 2011. Inverse and Leveraged ETFs: Considering the Alternatives. Journal of Financial Planning Jan 2011, Vol.24(1), pp.44-49.

Bilello, C. 2018. Do All Leveraged ETFs Go To Zero? Seeking Alpha. [Online document]. [Accessed 22 November 2019]. Available at https://seekingalpha.com/article/4203362-leveraged-etfs-go-zero

Charupat, N. & Miu, P. 2011. The pricing and performance of leveraged exchange-traded funds. Journal of Banking & Finance 35 (2011) 966-977.

Charupat, N. & Miu, P. 2014. A New Method to Measure the Performance of Leveraged Exchange-Traded Funds. The Financial Review 49 (2014) 735-763.

Chen, J. & Diaz, J. 2012. Spillover and Asymmetric-volatility Effects of Leveraged and Inverse Leveraged Exchange Traded Funds. Journal of Business and Policy Research Vol.7. No.3. September 2012 Special Issue. pp.1-10.

Cheng, J. & Miller, C. 2016. Analysis of Tracking Errors for Oil Sector Leveraged Exchange Traded Funds. International Research Journal of Applied Finance 2016, Vol.7(6), pp.89-98.

Cheng, M. & Madhavan, A. 2009. The Dynamics of Leveraged and Inverse Exchange-Traded Funds. Journal of Investment Management: JOIM, p. 7.

Co, R. 2009. Leveraged ETFs vs. Futures: Where Is the Missing Performance? Research and Product Development. CME Group. [Online document]. [Accessed 22 November 2019]. Available at https://www.cmegroup.com/trading/equity-index/files/Leveraged_ETFs.pdf

etf.com 2019. Leveraged ETF Overview. [Online document]. [Accessed 20 November 2019]. Available at https://www.etf.com/channels/leveraged-etfs

ETF Guide 2019. Investment & Pensions Europe. Liam Kennedy. London: IPE International Publishers Ltd

ETFdb.com 2019a. The Most Significant Events in ETF History. [Online document]. [Accessed 30 September 2019]. Available at https://etfdb.com/the-most-significant-events-in-etf-history/

ETFdb.com 2019b. Leveraged Equity ETF List. [Online document]. [Accessed 17 October 2019]. Available at https://etfdb.com/etfs/leveraged/equity/

ETFdb.com 2019c. Bond ETF List. [Online document]. [Accessed 17 October 2019]. Available at https://etfdb.com/etfs/asset-class/bond/

ETFdb.com 2019d. Leveraged Bond ETFs. [Online document]. [Accessed 17 October 2019]. Available at https://etfdb.com/etfdb-category/leveraged-bonds/

ETFdb.com 2019e. Barclays Capital U.S. 20+ Year Treasury Bond Index – ETF Tracker. [Online document]. [Accessed 5 November 2019]. Available at https://etfdb.com/index/barclays-capital-us-20-year-treasury-bond-index/

ETFdb.com 2019f. iBoxx $ Liquid High Yield Index – ETF Tracker. [Online document]. [Accessed 5 November 2019]. Available at https://etfdb.com/index/iboxx--liquid-high-yield-index/

Guedj, I., Li, G., Mccann, G. 2010. Leveraged and Inverse ETFs, Holding Periods and Investment Shortfalls. The Journal of Index Investing 2010 Vol.1(3), pp.45-57.

Guo, K. & Leung, T. 2015. Understanding the tracking errors of commodity leveraged ETFs. Fields Institute Communications 2015, Vol.74, pp.39-63.

Harri, A. & Brorsen, B. W. 1998. The Overlapping Data Problem. Quantitative and Qualitative Analysis in Social Sciences. Vol. 3, Iss. 3, 2009, 78-115.

Hessel, C., Nam, J., Wang, J., Xing, C., Zhang, G. 2018. Shorting Leveraged ETF Pairs. Journal of Trading Spring 2018, pp.69-79.

Hill, J. & Foster, G. 2009a. Understanding Returns of Leveraged and Inverse Funds. Journal of Indices, 12(5), pp.40-58.

Hill, J. & Teller, S. 2009b. Rebalancing Leveraged and Inverse Funds. Institutional Investor Journals. ETFs and Indexing Fall 2009 (1) pp.67-76.

Hunter, M., Lu, X., McLeod, R., Mehran, J. 2013. Bad news bears: Effects of expected market volatility on daily tracking error of leveraged bull and bear ETFs. Managerial Finance 14 October 2013, Vol.39(12), pp.1169-1187.

Jarrow, R. A. 2010. Understanding the risk of leveraged ETFs. Finance Research Letters 7 (2010) 135-139.

Jiang, X. & Peterburgsky, S. 2017 Investment performance of shorted leveraged ETF pairs. Applied Economics 20 September 2017, Vol.49(44), pp.4410-4427.

Lau, A. 2014. Leveraged ETFs not just for day traders anymore. [Online document]. [Accessed 30 September 2019]. Available at https://www.reuters.com/article/us-investing-etfs-leveraged/leveraged-etfs-not-just-for-day-traders-anymore-idUSBREA1B0WK20140212

Lauricella, T. 2009. Investing in Funds: A Quarterly Analysis --- Fund Fiend – ETF Math Lesson: Leverage Can Produce Unexpected Returns. Wall Street Journal, Eastern edition; New York, N.Y. 05 Jan 2009: R.1.

Leung, T. & Santoli, S. 2016. Leveraged Exchange-Traded Funds. Price Dynamics and Options Valuation. eBook. SpringerBriefs in Quantitative Finance.

Liebowitz, S. J. & Margolis, S., E. 1999. Path dependence. Encyclopedia of Law and Economics. 0770. p.981-998.

Loehr, R. & Lamb, R. 2013. Long-Term Investing With Leveraged Exchange Traded Funds. International Journal of Arts and Commerce- Vol.2 No.4. pp.27-34.

Lu, L., Wang, J., Zhang, G. 2012. Long term performance of leveraged ETFs. Financial Services Review 21 (2012) 63-80.

Militaru, R. & Dzekounoff, D. 2010. Trading with leveraged and inverse ETFs. Futures; Chicago Vol. 39, Iss. 3, (Mar 2010): 32-35.

Murphy, R. & Wright, C. 2010. An Empirical Investigation of the Performance of Commodity-Based Leveraged ETFs. The Journal of Index Investing Winter 2010, 1 (3) 14-23.

Newey, W. & West, K. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, Vol. 55, Iss. 3, (May 1987): 703.

Real Statistics Add-In for Excel. 2019. Add-in and documentation available at: http://www.real-statistics.com/

Rompotis, G. 2014. On leveraged and inverse leveraged exchange traded funds. Aestimatio 2014, Issue 9, pp.150-180.

Sharpe, W. 1966. Mutual Fund Performance. The Journal of Business 1 January 1966, Vol.39(1), pp. 119-138.

Shum, P. & Kang, J. 2012. Leveraged and inverse ETF performance during the financial crisis. Managerial Finance, Vol.39(5), pp.476-508.

Shum, P., Hejazi, W., Haryanto, E., Rodier, A. 2016. Intraday Share Price Volatility and Leveraged ETF Rebalancing. Review of Finance Oct 2016, Vol-20(6), p.2379.

Širůček, M., Ruml, V., Strejček, P. 2018. Measuring the performance of leveraged and non-leveraged ETF'S. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis 2018, Vol.66(5), pp.1357-1367.

The Motley Fool. What Is a Leveraged ETF? [Online document]. [Accessed 30 October 2019]. Available at https://www.fool.com/knowledge-center/what-is-a-leveraged-etf.aspx

Trainor, W. J. & Baryla, E. A. 2008. Leveraged ETFs: A Risky Double That Doesn't Multiply by Two. Journal of Financial Planning; Denver Vol. 21, Iss. 5, (May 2008); 48-50,52-55.

Trainor, W. J. 2010. Do Leveraged ETFs Increase Volatility. Technology and Investment. Aug 2010, Vol.1(3), pp.215-220.

Trainor, W. J. 2011. Daily vs. monthly rebalanced leveraged funds. Journal of Finance and Accountancy. Vol.6, pp.1-14.

Trainor, W. J., Carroll, M. 2013. Forecasting Holding Periods for Leveraged ETFs Using Decay Thresholds: Theory and Applications. Journal of Financial Studies & Research. Vol.2013 pp.1-13.

Yahoo! Finance 2019. Daily adjusted closing prices for selected ETFs from 22/10/2012 to 22/10/2019

Yates, T. 2007. Leveraged ETFs: A Value Destruction Trap? Seeking Alpha. [Online document]. [Accessed 7 October 2019]. Available at https://seekingalpha.com/article/31195-leveraged-etfs-a-value-destruction-trap

Zhang, J. 2010. Path-Dependence Propertied of Leveraged Exchange-Traded Funds: Compounding, Volatility and Option Pricing. Doctoral dissertation. New York University Department of Mathematics.

Zweig, J. 2009. The Intelligent Investor: How Managing Risk With ETFs Can Backfire. Wall Street Journal, Eastern edition; New York, N.Y. 28 Feb 2009: B.1.