ERASMUS UNIVERSITEIT ROTTERDAM

On the non-compounding tracking error of leveraged exchange-traded funds

Eduard Th.J. Jansen

Erasmus Universiteit Rotterdam, Erasmus School of Economics

Master Thesis Financial Economics

Student: Eduard Th.J. Jansen

Student number: 375670

Supervisor: Ricardo Barahona

Second assessor: Dr. Judy Chalabi

Date final version: 1 May 2021

Abstract

A leveraged exchange-traded fund is an investment instrument that applies a constant leverage multiple to index-tracking portfolios. Issuers warn that the instrument is unsuitable for holding periods longer than a day due to the compounding effects of constant leverage. This paper researches the single-day performance of leveraged exchange-traded funds. After first discussing the reasons for the existence of the funds, potential determinants of tracking error are found in prior research. These are lagged tracking error, LIBOR, benchmark index volatility, fund liquidity, and leverage multiple. This paper expands on previous leveraged exchange-traded funds research by including a sample of 37 funds, with data since the launch of the first fund in 2006 until 2020. A dynamic panel regression is performed to assess the impact of the explanatory variables on tracking error and facilitates straightforward replication of the research. The regression confirms significant effects of the 1-day lagged tracking error, LIBOR, and benchmark index volatility on tracking error of leveraged exchange-traded funds. The findings provide insights for both academics and investors.

Keywords: exchange-traded funds, leverage, tracking error

Preface and acknowledgements

Before you lies my master thesis "On the non-compounding tracking error of leveraged exchange-traded funds", a research on the performance of a peculiar financial instrument that tracks diversified stock indices while applying leverage to boost returns. It was written in early 2021, during the COVID-19 lockdown. It is the conclusion to my studying at Erasmus University Rotterdam and my years of academic, professional, and personal education as a student in the Netherlands.

The inspiration for this subject stems from a personal interest in investment instruments for retail investors with a long-term outlook. My research question was formulated together with my thesis supervisor Ricardo Barahona. I thank my supervisor for his guidance and I wish him well in his academic career. I also wish to thank the researchers whose work I refer to in my thesis for their efforts in uncovering the workings of leveraged funds.

Everybody who supported me during my years as a student deserves gratitude for their support. This includes all my friends, the universities of Rotterdam and Delft, the companies that have provided me with internship opportunities, and the overall environment of the Netherlands that allows individuals to blossom. Thank you to my lovely girlfriend Fenja who encouraged me and pushed me, and still does, to continually challenge myself. Big thanks to my housemates and fellow finance students Sjoerd and Hippolyte. You have been important motivators to me the last two years, academically and professionally. I wish you well in your careers in finance. Thank you to all players and staff at the Rotterdam Student Rugby Club for providing me with lifelong memories, friendships, and a sense of passion that can never be surpassed. I also want to specifically thank my friend James for his valuable feedback, suggestions, and econometric insights for this thesis. Finally and most importantly, I wish to thank my family for giving me the opportunities that made me to be the person that I am today and I hope to have made you proud thus far.

I cherish the expectation that you enjoy your reading of this thesis and may you receive high returns on your investments.

Eduard Th.J. Jansen
Rotterdam, April 2021

Table of Contents

List of tables

1 Introduction

For modern-day retail investors, it is a common strategy to invest in mutual funds or exchange-traded funds (ETFs) that follow stock indices. In fact, the popularity of so-called passive investing instruments has exploded to pass $10 trillion of global assets under management (Wigglesworth & Janiaud, 2020), and BlackRock (2017) estimated that index investors held 18% of global equity in 2016. While index investing provides stable returns, figures show that for many retail investors, the average annual 8% of the S&P 500 is not enough. At the end of 1999, the total U.S. margin debt, money borrowed by individual investors from brokerage firms to leverage investment returns, was $229bn. This amounted to paying for 10% of total assets for some online U.S. stockbrokers (Simon, 2000). In 2021, the figure reached a record high of $814bn (Financial Industry Regulatory Authority, 2021).

Index returns can also be amplified using leveraged exchange-traded funds (LETFs). These funds promise to return a multiple, typically two or three times, of the daily return of an underlying index. The first LETF, the two-times leveraged ProShares Ultra S&P500, was introduced in 2006. From its inception until the end of 2020, it has returned over 397%. In the same period, the S&P 500 has returned over 190%. Despite these attractive returns, financial advisors and LETF issuers themselves advise against holding leveraged funds for longer than one day, while regular index investing is seen as a sound long-term strategy. This is because the constant leverage multiple of LETFs gives unpredictable returns through different compounding effects and may bring losses to investors even when the underlying index moves favorably. This constant leverage trap has been extensively researched. However, even if investors follow the fund issuers' advice and hold LETFs for one day only, the funds underperform their desired multiples.

This research aims to find the non-compounding factors that determine the daily tracking error of LETFs. The objective is to find the model with the most explanatory power, as opposed to the best predictive performance. This is done by applying dynamic panel regression to a sample of 37 funds with data since the launch of the first LETF in June 2006 until December 2020. This research adds to the existing literature, firstly because research on the tracking error of LETFs is scarce overall. Secondly, most literature focuses on the compounding tracking error of LETFs, and not the non-compounding part. Thirdly, most research uses a small sample. The research that also uses regression analysis and therefore has the most similar methodology, only uses a sample of 12 funds for the period July 2006–December 2010 (Tang & Xu, 2013). The tracking-error framework created by Bansal and Marshall (2015b) is only demonstrated on one LETF-Index pair during the period 2011–2014. Following suggestions from previous literature, the effects of several factors on daily tracking error are researched. These factors include return characteristics such as volatility, fund characteristics such as leverage multiple, and external factors such as LIBOR. Ultimately, tracking error lag, LIBOR and benchmark volatility are found to have significant explanatory value.

Following this introduction, section 2 describes a short history of the origin of LETFs. This is followed by a review of the literature available on the determinants of tracking error in index-replicating funds. Section 3 is the data section. It describes the data used to create the variables that are used in the dynamic panel regression described in the following two sections. Furthermore, it shows the characteristics of the sample, and the magnitude of tracking error on an annualized basis and on a daily basis. Section 4 describes several statistical methods and shows the considerations in determining the optimal methods for this sample and research question. Section 5 shows the results of the statistical analysis using the selected method and discusses whether these are in line with expectations raised in the literature review. Section 6 draws conclusions from the findings of the results, touches on the limitations of this research, and makes recommendations on future research on the tracking error of LETFs.

2 Literature review

This section explains the mechanics of leveraged exchange-traded funds (LETFs) and discusses the existing literature on tracking error. First the reason for the existence of LETFs is given by illustrating the history of index investing and applying leverage. Next, the workings and price movements of LETFs are shown and literature on their long-term performance is discussed. The last subsection reviews the literature on non-compounding tracking error, which is the main subject of research in this paper.

2.1 Origin of leveraged exchange-traded funds

In 1774, Amsterdam-based merchant Abraham van Ketwich founded the first investment fund enabling individuals to invest in a diversified portfolio to reduce risk. One-hundred years later, the first stock index was born when U.S. publishing firm Dow Jones & Company printed the closing prices of 11 stocks. Since then, many mutual funds and many stock indices have been founded. In 1976, the first index-tracking fund open to individual investors was launched: a mutual fund following the performance of the S&P 500 Index. In 1993, index investing was made even more accessible with the debut of exchange-traded funds (ETFs). Where mutual funds often have minimum investment requirements and can only be invested in once a day at closing price, ETFs enable investors to open and close positions in stock index funds like they can in stocks (Bansal & Marshall, 2015a; Mosselaar, 2019). For investors who do not wish to settle for the ETF returns of the underlying indices, investors may employ leverage. The simple method of using leverage entails increasing capital by borrowing a fixed amount from a stockbroker and investing this amount. By keeping the profits, investors can earn returns on a large amount of capital while using less of their own equity. Even though this method is simple, it poses risks. A sharp decline in asset value may trigger margin calls, where the collateral does not exceed the minimum value required by the broker. The borrower can then be forced to sell assets at a low price to deposit cash into the margin account, leaving less possibilities for recouping the losses they have suffered. In 2006, LETFs were introduced. These publicly traded instruments promise to return a constant multiple of the return of the underlying index. Meanwhile, the risk of margin calls is almost completely removed because of the constant leverage. The next subsection explains how LETFs work and how these instruments perform over longer holding periods.

2.2 Compounding and long-term tracking error

LETFs return a constant multiple of the daily return of the underlying index. Most long or bull LETFs provide daily multiples of 2 or 3, while short, bear, or inverse LETFs aim to multiply by factors -1, -2, or -3. LETFs greatly reduce the risk of margin calls by rebalancing the leverage every day[1]. Instead of borrowing a fixed amount, investors borrow a fixed proportion of the capital employed[2]. For example, a 2x LETF promises double the daily returns by matching with borrowed funds exactly the equity invested, so that the equity invested is ½ of capital employed. Likewise, a 3x LETF returns triple that of the underlying by ensuring that equity invested is ⅓ of capital employed and ⅔ is borrowed. Daily rebalancing ensures that margin calls can only happen in unlikely events of the underlying index dropping more than 50% or 33% in a single day for 2x or 3x LETFs respectively. This daily rebalancing poses other threats in the form of compounding deviation however. LETF issuers warn that while LETFs double the returns of the underlying daily, they do not double the returns for any period longer than that and state that LETFs are not a suitable long-term investment because of their tracking error.

[1] LETFs exist with different rebalancing periods, such as monthly or quarterly, but in this paper only daily rebalancing is considered.

[2] In reality, LETFs are constructed using derivatives such as equity swaps to exchange return cash flows for interest rates (Tang & Xu, 2013). These swaps allow LETF issuers to receive the returns of equity assets without owning the assets. This technicality is important as the interest rate for these equity swaps will be suggested as a potential determinant of tracking error.

The most-documented reason for this long-term unsuitability is volatility decay, also known as the constant leverage trap (Trainor Jr. & Baryla Jr., 2008). The rebalancing of constant leverage, a main characteristic of LETFs, forces investors to buy after a rise in price of the underlying, and to sell after a decline in price. In a volatile market without a clear upward trend, investors lose money. Take the following example and compare investment A, a $100 investment in an unleveraged index, and investment B, $100 in the same index leveraged two times. If on day one the index drops 10%, investment A is worth $90 and investment B is worth $80 (a decline of 20%). On day two, the index rises 11.11% so that investment A is worth $100 again. Investment B rises 22.22% and is worth only $97.78. The order of price increases and drops does not matter: a rise of 10% followed by a decline of 9.09% brings investment A back to $100 while B has dropped to $98.18. Lack of understanding of this volatility decay has caused disgruntled investors to take legal action against LETF issuers and is a main reason why LETF issuers warn for LETFs' unsuitability for holding periods longer than one day (Tang & Xu, 2013).

The daily rebalancing can also be beneficial in the form of compound interest and can outperform the promised multiple. Bansal and Marshall (2015a) have described this as the trending effect. Let us consider again the example from the previous paragraph. If the index rises an impressive 10% for two consecutive days, investment A will increase from $100 to $110 to $121, a total return of 21%. Meanwhile, investment B rises from $100 to $120 to $144, a total return of 44% and more than twice the return of unleveraged investment A.

LETF long-term performance and the opposing forces of volatility decay and compound interest (the effect of both will be named compounding deviation from now on) have been researched for different holding periods. Avellaneda and Zhang (2010) show that a 2x leveraged long and an unleveraged short position underperform its theoretical equivalent of an unleveraged long position over any 60-day period since the inception of LETFs. Trainor, Jr. and Baryla, Jr. (2008) show that 2x LETFs also do not provide their multiple of annualized returns over 1-, 3-, 5- or 10-year periods. They do show that the LETFs outperform the unleveraged underlying, as well as the underlying leveraged with a margin account (a leveraging strategy which does not or less often rebalances the leverage, which would logically make the strategy less susceptible to volatility decay). The research shows a median annualized return of about 1.4 times the index return for 2x LETFs. A simple calculation shows that it would take little over two years for the trending effect to do its work and deliver on the 2x promise. This shows that for the long-term investor, a one-year time horizon may be too short. The effects of compounding deviation seem to turn out positive in the longer term. For holding periods longer than one year, LETFs tracking US indices do not underperform their stated goals (Loviscek, Tang, & Xu, 2014).

2.3 Non-compounding tracking error

The long-term performance of LETFs has been researched extensively. However, long-term performance of LETFs is not only dependent on compounding deviation, but also on the one-day performance of the LETF compared to its benchmark. A certain non-compounding tracking error, as is present in any exchange-traded fund, is to be expected. In the broadest sense, tracking error is expressed as the difference between portfolio return and the benchmark return and describes the divergence of price movements. Vardharaj, Fabozzi, and Jones (2004) have shown several determinants of tracking error for asset managers whose benchmark is an equity index. The asset managers in their research do not manage ETFs specifically, but the determinants they find can still be of interest. They find that increased benchmark volatility or market volatility increases tracking error. Next to this, they name market capitalization of portfolio holdings (these should be identical to the benchmark when it comes to index-replicating portfolios), the number of stocks in the portfolio (idem), investment style compared to the benchmark (idem), and sector deviation from the benchmark (again, there is no deviation in the portfolios that I will consider in this paper) as possible determinants of tracking error.

Bansal and Marshall (2015a, 2015b) created a framework suitable for decomposing tracking error for LETFs. Components of tracking error they find that are specific to daily tracking error are volatility of the underlying index and the degree of leverage employed. Later they add to this an interest rate paid component and a management component. The interest paid component is by definition negative for returns of LETFs with a positive leverage multiple, because a positive leverage multiple means a certain amount is borrowed. The management component surprisingly adds to the relative return of LETFs for the holding period and single fund they research.

Tang and Xu (2013) also investigate tracking errors for LETFs specifically. They name swap-related floating rate payout/receipt in the form of LIBOR, volatility, liquidity, and lagged tracking errors as determinants. The direction of these determinants are as follows: LIBOR increases (decreases) tracking error for bull (bear) funds as floating rates are the costs (receipt) for swap agreements. Higher volatility makes it more costly for funds to hedge and increase tracking error. Liquidity on the other hand makes it less costly for funds to hedge and thus decreases tracking error. Lagged tracking error is named as an explanatory variable, as mean reversion is expected. This is expected because of the creation/redemption feature of ETFs. This feature allows authorized participants, such as large banks, to exchange ETF shares for the underlying stocks, or vice versa. By doing this whenever there is a price difference between the ETF and the underlying, price efficiency of ETFs increases. Mean reversion implies a negative coefficient of the lag.

3 Data

This section starts with a description of the sources and variable transformations used to construct the sample. Next, several descriptive summaries are given. Firstly, the non-return characteristics of the funds will be shown. Secondly, the returns are shown on an annualized basis. Finally, the returns and tracking errors are shown on a daily basis as these are the object of the dynamic panel regression that follows.

3.1 Sources and transformation

A list of 139 LETFs has been downloaded from the ETF Database website. While by no means an academic source, this database is useful for basic operations such as compiling a list of funds. The LETFs on the list are currently active, so unfortunately there is a chance of survivorship bias. All LETFs in the list have leverage of -3, -2, -1, 2, or 3. Only funds that have at least three years of data available until 31 December 2020 have been considered, so any funds started after 31 December 2017 were removed from the sample. The total returns (including dividends) since inception for these LETFs and for the underlying indices have been downloaded from Datastream, together with the daily number of shares outstanding for each LETF (this will be used as a proxy for liquidity). All fund and index data downloaded are U.S. dollar-denominated. For several LETFs and indices, not all information could be found. These are shown in table A2 in the appendix and were eliminated from the sample. Datastream tracks the returns from the first value of the index and adjusts for stock splits. Normally, when LETF values approach zero, the LETF issuer executes a reverse-stock split. Datastream corrects for these splits so that returns are not distorted. This however means that for some LETFs, the values approach zero. Since Datastream only provides data with two decimal points, accurate returns cannot be calculated. At some point, the extent to which the data is poorly estimated is greater than the extent to which the dependent variable varies. For this reason, the decision was taken to remove any LETF with total return index values in the sample below 10.00. This threshold is chosen because below this threshold absolute returns will be either 0.0%, or 0.1% or higher. Furthermore, to check for validity of the LETF data, the intercept of the volatility of the LETF compared to its leveraged benchmark was calculated. Of LETFs with a beta over the whole sample period below 0.75, the data are deemed to be invalid. A full overview of the 37 remaining funds, as well as a list of removed LETFs can be found in appendix A.

The net expense ratio (ER) was taken from the latest fact sheet of each LETF. Although these rates are assumed to be mostly static over time, this could not be verified since only the most recently updated fact sheets are readily available. Datastream does not have the number of stocks in indices, so this figure was taken from the latest fact sheet of each index. Like the expense ratio, only the most recent values are available. Unlike the expense ratios, the number of companies may differ over time, for example for indices that include all equities within a geographic area that meet the index criteria.

Next, hypothetical perfectly leveraged indices are created by multiplying the returns with their leverage multiples. Inverse indices are considered using leverage of -1. These indices will be used as benchmarks for the LETFs to follow.

The two papers discussed in the literature review section that name benchmark volatility as a possible determinant for tracking error, define volatility in different manners. Vardharaj et al. (2004) take the annualized standard deviation of monthly returns of the benchmark. Tang and Xu (2013) use the Cboe Volatility Index (VIX) as their proxy for volatility. The Chicago Board Options Exchange (2019) describes the VIX as the 30-day forward volatility as implied by S&P 500 options prices. The VIX is not suitable as a volatility proxy for this paper as it is based on the S&P 500 and thus represents volatility of the U.S. market, while the sample contains indices from other geographic areas as well. To align methodology with previous literature but adapted to fit the sample, volatility in this research is calculated daily as the standard deviation of the returns of the unleveraged benchmark index for the previous 30-day period. When there is no observation at $t-30$, the nearest available data point before $t-30$ is chosen as the starting observation.

LIBOR is the annualized 3-month London Interbank Offered Rate. Daily annualized three-month rates are downloaded from the Federal Reserve Bank of St. Louis. This rate can be used in the regression in multiple ways. The rate can simply be entered as a variable. The expected influence of the rate demands a more sophisticated variable is created, however. Therefore a variable is created in which the rate is combined with the leverage multiple. As explained by Tang and Xu (2013), LETFs use equity swaps for exposure to the underlying index. Equity swaps are agreements where one party receives the returns on an asset, without being required to hold the asset. The price paid for this is a floating interest rate such as LIBOR. A new variable LIBOR Paid is created using

$$\text{LIBOR Paid} = (m - 1) \cdot \text{LIBOR}^{\frac{1}{360}} \tag{1}$$

where m is the leverage multiple. The rate is multiplied by $m - 1$ because swap agreements are made of m times assets, however the assets can be subtracted once, as the interest received on the equity can offset the interest paid on this amount.

Daily risk-free rates are downloaded from the CRSP Risk-Free Rate Series and are based on the 4-week U.S. Treasury bill rates.

3.2 Descriptive statistics

Table 1 shows the characteristics of the sample funds. The full list of funds can be found in the appendix. The geographical distribution of funds shows that has the sample has a bias towards U.S.-based indices. Of 37 funds, 28 follow a U.S.-based index. The other 11 follow an index based on either a different country or a greater geographic area (such as global indices). Within issuers, the makeup of leverage varies. Most Direxion funds are 3x leveraged, while ProShares have more -1x and 2x funds. Finally, the inception dates of the sample funds are described. The first LETFs were established in 2006 and are included in the sample. Most of the funds in the sample were launched between 2006 and 2010, while none were launched between 2011 and 2013. No LETFs launched after 2017 were considered in the sample to ensure at least three years of return data for each fund.

Table 1

Descriptive statistics of LETF sample and underlying indices

	Total (N=37)	-1x funds (N=10)	2x funds (N=18)	3x funds (N=9)
		Area		
United States	28	7	14	7
Rest of World	9	3	4	2
		Issuer		
Direxion	10	2	2	6
ProShares	27	8	16	3
		Stocks in benchmark		
Mean	366	430	299	428
Min.	20	30	30	20
Max.	1,381	1,381	1,381	1,381
		Inception year		
2006–2010	30	8	16	6
2011–2013	0	0	0	0
2014–2017	7	2	2	3

Note: This table shows summary statistics of the leveraged exchange-traded funds (LETF) sample consisting of 37 funds. Area means the count of underlying indices in a certain geographic area. Issuer is the provider company of the LETF. Stocks in benchmark is the number of companies in the underlying index. Inception year is the year in which the fund was launched and is determined by the first available data point in Datastream.

When choosing a LETF to invest in, investors should care about its performance relative to the underlying index. More specifically, they want the LETF to mimic the underlying as closely as possible, and to be able to predict the accuracy with which the

LETF can deliver this. Different levels of sophistication of investors ask for different ways of measuring this performance. The simplest manner to evaluate relative performance is to measure the absolute or relative difference between LETF and underlying index returns. A more sophisticated method is to look at these returns in a risk-adjusted way (Elton, Gruber, & Busse, 2004). This is done by using

$$R_L - R_f = \alpha_L + \beta_L\left(R_B - R_f\right) + \varepsilon_i \qquad (2)$$

where R_L is the daily return on the LETF, R_f is the daily risk-free rate and R_B is the daily return on the leveraged benchmark index. β_L is the LETF volatility in relation to the benchmark excess returns and α_L represents the risk-adjusted return that the LETF generates. An instrument that perfectly follows the underlying has alpha 0 and beta 1.

Table 2 shows the annualized returns for the sample funds and their underlying indices. Moreover, it shows differential returns and the variables from (2) to get a clear view of the simple and risk-adjusted performance of the LETFs in the sample. Panel A shows the annualized returns for the sample funds. As expected, returns strongly differ according to the leverage employed. Inverse LETFs have average annualized return of −11.6% with a 3.1% standard deviation. Two-times LETFs return 11.4% on average with standard deviation 9.4%. Triple LETFs have returns of 16.8% with a standard deviation of 15.2%. For each of the funds, the differential return has been calculated as the difference between the LETF return and the leveraged benchmark return, displayed in panel C. The mean differential return for −1x funds is −0.6%. For larger leverage multiples this increases. Two-times funds and 3x funds have −3.2% and −7.0% underperformance, respectively. Alpha, calculated using (2) and shown in panel D, denotes the risk-adjusted return. As expected, the average alpha is lower than the differential return as the differential return also consists of the beta portion of the equation.

Table 2

Annualized return statistics per leverage multiple

	−1x funds	2x funds	3x funds	−1x funds	2x funds	3x funds
	Panel A: LETF annualized return %			Panel B: Index annualized total return %		
Mean	-11.62	11.42	16.79	-11.03	14.65	23.77
Std. dev.	3.07	9.37	15.24	4.06	9.42	12.10
Min	-15.60	-13.50	-5.25	-16.81	-11.16	8.34
Max	-5.55	26.27	35.44	-5.80	29.49	39.69
	Panel C: Differential return %			Panel D: Alpha %		
Mean	-0.59	-3.23	-6.97	-0.32	-1.30	-3.15
Std. dev.	2.46	1.20	5.04	1.79	1.87	1.72
Min	-5.07	-7.23	-17.57	-4.18	-3.49	-5.62
Max	1.89	-1.64	-3.57	2.06	4.56	0.23
	Panel E: Annual expense ratio %			Panel F: Beta		
Mean	0.89	0.95	1.03	0.94	0.95	0.95
Std. dev.	-	-	-	0.06	0.04	0.06
Min	0.50	0.64	0.93	0.82	0.87	0.81
Max	0.95	1.21	1.33	0.99	0.99	0.99
	Panel G: Differential return + expenses %			Panel H: Alpha + expenses %		
Mean	0.30	-2.28	-5.94	0.57	-0.35	-2.12
Std. dev.	2.43	1.22	4.93	1.75	1.92	1.66
Min	-4.22	-6.28	-16.24	-3.33	-2.54	-4.55
Max	2.39	-0.43	-2.62	2.56	5.77	1.22

Note: This table shows summary statistics of the leveraged exchange-traded fund (LETF) sample. The annualized returns of the LETFs are calculated over the course of its existence, from the launch date in the inception year until the end of the sample period, 31 December 2020. The index annualized returns are calculated for the same period. Differential return is the difference between the return on the LETF and the annualized return on the underlying leveraged index. Alpha is calculated by taking the intercept in the regression of the daily excess LETF return (LETF return minus the risk-free rate) against the daily excess return on its underlying (leveraged) index, annualized based on 250 trading days. Beta is the volatility of the LETF in relation to the volatility of the underlying leveraged index. Expenses are the annual net expense ratio as specified in the fact sheets of the funds.

Panel E of table 2 shows that mean net expense ratios are higher for LETFs with higher leverage and are lowest for inverse leverage. The mean rates for -1x, 2x, and 3x funds are 0.89%, 0.95%, and 1.03%, respectively. Panel G tells that the economically significant underperformance of the sample funds can partially be explained by the expense ratios which range from 0.50% to 1.33%. Surprisingly, the mean tracking error before expenses is positive for inverse ETFs. However, even after correcting for expenses, large underperformance can still be observed for 2x and 3x funds.

Table 3 shows daily statistics. For each leverage multiple and for all funds, the table shows distribution data. LETF returns, Index returns, and Tracking error are shown, followed by the tracking error without the effect of expenses. The final data shown is the response variable of the regression, log of tracking error before expenses.

Table 3

Daily tracking error statistics per leverage multiple

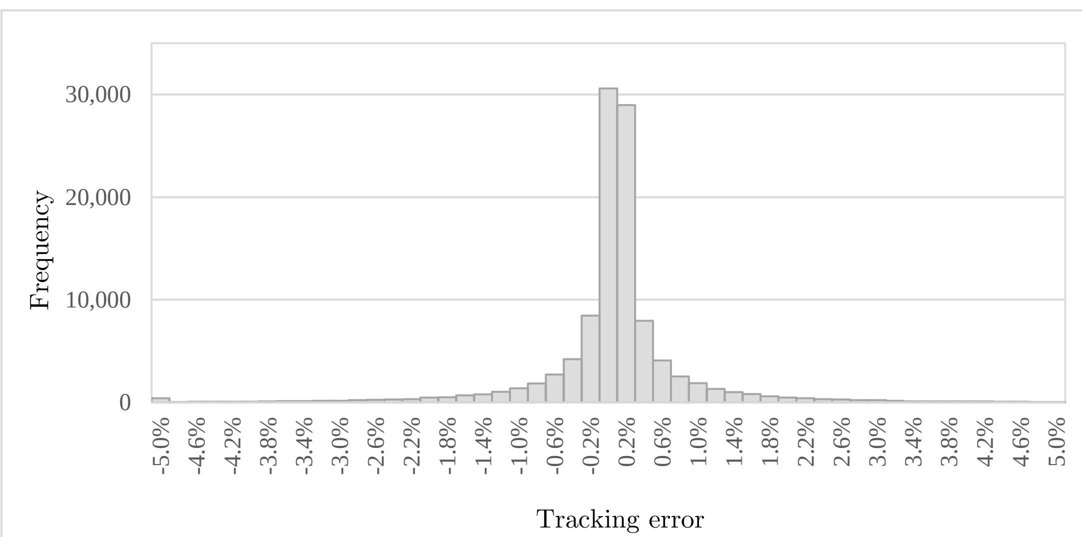
				Quantiles				
	n	μ	*SD*	Min	0.25	Mdn.	0.75	Max
				Panel A: All funds				
LETF returns %	107,449	0.06	2.84	-52.77	-0.95	0.03	1.21	43.54
Tracking error + expenses %	107,449	-0.01	1.08	-49.20	-0.17	0.00	0.16	41.70
Log (Tracking error + exp.) %	107,449	-0.01	1.08	-67.72	-0.17	0.00	0.16	34.86
				Panel B: -1x funds				
LETF returns %	28,909	-0.03	1.61	-21.39	-0.68	-0.03	0.54	20.46
Tracking error + expenses %	28,909	0.00	0.80	-16.56	-0.12	0.01	0.14	12.98
Log (Tracking error + exp.) %	28,909	0.00	0.80	-18.10	-0.12	0.01	0.13	12.21
				Panel C: 2x funds				
LETF returns %	57,231	0.08	2.82	-37.94	-1.05	0.12	1.36	40.01
Tracking error + expenses	57,231	-0.01	0.96	-15.42	-0.19	-0.01	0.17	27.18
Log (Tracking error + exp.) %	57,231	-0.01	0.96	-16.75	-0.19	-0.01	0.17	24.04
				Panel D: 3x funds				
LETF returns %	21,309	0.15	3.96	-52.77	-1.42	0.22	1.94	43.54
Tracking error + expenses %	21,309	-0.02	1.58	-49.20	-0.17	-0.01	0.15	41.70
Log (Tracking error + exp.) %	21,309	-0.03	1.60	-67.72	-0.17	-0.01	0.15	34.86

Note: Leveraged exchange-traded fund (LETF) returns and Index returns are simple arithmetic returns. Tracking error is LETF returns – Index returns. Expenses are annual net expense ratios converted to daily based on a 360-day year. Log values calculated using $r = log(R + 1)$.

In panel A of table 3, the mean tracking error, defined as the difference between the LETF return and Index return, is −0.01%. Although the mean is negative, many tracking errors are positive as can be seen from the standard deviation and the distribution in figure 1. Panel B shows the tracking error is smallest and slightly positive for -1x funds, as was also observed in table 2, panel C. The 2x funds in panel B show larger returns for both the LETF and Index, but also a larger tracking error. Returns and tracking error are largest for 3x funds as shown in panel D. The funds display large outlier returns and tracking errors. Observation of the sample leads to show that these large tracking errors often occur during periods of high volatility such as the March 2020 stock market crash. Section 5 will describe the influence of leverage multiple and volatility on tracking error, among other variables. All panels show that the expense ratio is negligible on a daily basis. Taking the logarithm of the tracking error plus expenses changes outlier returns but has a minor effect on the daily returns in the interquartile range.

Figure 1

Frequency of daily tracking errors in sample



Note: Tracking error (LETF returns – Index returns) frequencies are shown for 107,449 observations of all 37 sample funds. For visibility, only tracking errors between −5.0% and 5.0% are shown, covering over 99.6% of observations. The distribution is symmetrical but deviates from normality: it shows concentration around the mean (−0.01%) and displays fat tails.

4 Methodology

This section first motivates that a dynamic panel regression model is appropriate for the sample data. A regression model will be presented and its coefficients and their expected direction will be explained. Next, the choice between the fixed effects model and the random effects model variants of panel data analysis will be motivated. The Hausman test to further motivate the decision will be proposed. Last, this section will describe how to evaluate the models after they have been run.

4.1 Dynamic panel regression

Since the sample has multiple observations across time for different funds, some sort of panel analysis is appropriate. Stata is used to perform the statistical analysis. The panel is structured in the long format and as not all funds have observations every period, it is denoted as unbalanced. A simple version of the regression to be performed is

$$\begin{aligned} R_{\text{Diff},it} + \text{Expense}_i \\ &= \beta_0 + \beta_1\left(R_{\text{Diff},it-1} + \text{Expense}_i\right) + \beta_2\text{Leverage}_i \\ &+ \beta_3\text{Area}_i + \beta_4\text{Issuer}_i + \beta_5\text{Volatility}_i + \beta_6\text{NOSB}_i + \beta_7\text{NOSH}_{it} + \beta_8\text{LIBOR}_t \\ &+ u_{it} + \varepsilon_{it} \end{aligned} \tag{3}$$

where i is fund and t is time. $R_{\text{Diff},it}$ is the tracking error, defined as the daily differential return between the LETF and the benchmark. Expense is the net expense ratio. $R_{\text{Diff},it-1} + \text{Expense}_i$ is the previous-day value $(t-1)$ of the dependent variable. This lagged variable makes the panel regression to be called dynamic. It is believed to have a negative coefficient due to the mean reversion described in the literature. Leverage is the leverage multiple of the fund. Since a higher multiple leads to larger price swings and requires more actions by fund managers (larger daily leverage rebalancing is required) this is expected to have a negative coefficient. Area is a dummy variable for the geographic area of the underlying index being the United States or rest of world. All sample funds are maintained by U.S. issuers which could lead to better performance for funds following U.S. indices, although geographic area is not expected to have a significant effect. Since all funds and indices in the sample are dollar-denominated, currency differences would not be a logical cause of an effect. Issuer is a dummy variable for the LETF issuer being ProShares or Direxion. Again, this is not expected to have a significant effect. Volatility is the volatility of the benchmark as described in the data section and is expected to increase tracking error. NOSB is the number of stocks in the benchmark index and is expected to decrease tracking error as found by Vardharaj et al. (2004). NOSH is the number of shares outstanding for the LETF. LIBOR is the interest rate paid on equity swaps and is expected to directly increase tracking error for bull funds and decrease tracking error for inverse funds. The error term consists of between-entity error u_{it} and within-entity error ε_{it}. The sample data does not suffer from multicollinearity as demonstrated in the correlation matrix in appendix B, except for variables that are derived from each other.

Equation (3) is not the only model that will be run, although no completely new variables will be introduced. Two models will use variants of the variables in (3), such as the LIBOR Paid variable in (1). Other models will have variables removed by means of the general-to-specific approach, a strategy that starts with a general model and improves it through reduction of variables. To evaluate some of the variables, models will also be run on subsamples of the data.

4.2 Fixed effects model vs. random effects model

There are multiple methods to perform panel data regression. Two main variants are the fixed effects and random effects models. A fixed effect model controls for and omits time-invariant differences between entities because these are collinear with the entity. It instead absorbs the time-invariant models in the intercept. Leverage multiple, Issuer, Area, and Stocks in benchmark are all independent variables in (3) that are constant within each fund and thus time-invariant. Since these variables need to be researched separately, a fixed effects model is not suited for (3). A random effects model on the other hand, allows time-invariant variables to be included. For this reason, a random effects model is preferred, although this preference can also be substantiated statistically using the Hausman test.

The Hausman test can be performed as a statistical argument to use fixed effects or random effects. The Hausman test tests whether the errors of each fund are correlated with the regressors. The null hypothesis is that they are not, and that a random effects model is preferred (Torres-Reyna, 2007).

If the Hausman test confirms that a random effects model is appropriate, the regression in (3) will be performed. The significance, direction, and magnitude of its coefficients will be evaluated, and other variants of the model will be presented. Next, the model will be separately run for the different leverage multiples as a robustness check. Using these findings, a final model will be constructed at the end of the results section.

4.3 Model evaluation & selection

The results of several models that are variations of (3) will be presented in the next section. Besides the individual evaluation of each model, a comparison between models is appropriate. To select the model with the most explanatory power, several evaluation and selection criteria must be considered.

When it comes to variable selection, simplicity states that when a choice between models is presented, the simplest model is often best. Studies show that complexity in models frequently reduces precision. Complexity and overfitting makes models less generizable (Zellner, Keuzenkamp, & McAleer, 2002). This is in line with the general-to-specific approach, which starts with a general model and aims to improve it by removing unneeded variables. On the other hand, trimming variables is advised against unless one is certain of the unimportance of the variable (Jaccard, 2001). Excluding variables contradicts validation of the explanatory powers of the model in some cases. Even the insignificance of variables in the theoretical model is of importance, depending on the theoretical justification of their presence. To find the model with most explanatory power, all significant variables must be included. This is opposed to searching for a model with the most predictive power, where variables can be removed if their coefficients are too small to influence the prediction, even if they are significant (Shmueli, 2010). In this paper, the advantages and disadvantages of adding and removing variables will be balanced by removing insignificant variables that have no theoretical justification. To show the effect of exclusion, all models will be shown in parallel.

For simple regressions, R-squared (R^2) is a popular assessment of goodness of fit. It represents the proportion of the variance of the dependent variable that is explained by the variance of the independent variables. Panel regression presents not one but three values. Overall R^2 is the R^2 value seen in usual regressions. Between R^2 is the R^2 if the time component were removed from the regression and only measures the variance between entities. Within R^2 concerns the variance within an entity over time and disregards the variance between entities. Overall R^2 will be used as the measure of choice to compare different models in the results section. It must be noted that R^2 values are usually lower for panel data than for cross-sectional or time series regressions. However, R^2 will not strictly be used to evaluate the absolute effectiveness of models, but rather the quality of the models relative to each other. Also, non-adjusted R^2 increases when new variables are added and not only when the explanatory power of the model increases. This implies that when R^2 values are the same, the model with the least variables is best.

5 Results

In this section, the tests described in the methodology section are performed and their results described. First, the results of the Hausman test will confirm the choice for the random effects model. Next, several dynamic panel regression models are run starting with (3). The results of these will be discussed in relation to the expectations raised in the literature section. Finally, the quality of the models is evaluated side by side.

5.1 Model choice

The previous section claimed that a random effects model is preferred over a fixed effects model for (3) because of the presence of time-invariant variables that need to be researched. To statistically substantiate this preference, a Hausman test will be performed. For the Hausman test, both the fixed effects model and the random effects model are modelled in table 4. The coefficients and their significance will be discussed in the next section. The fixed effects model omits Leverage multiple, Issuer, Area and Stocks in benchmark because these variables are time-invariant. The null hypothesis of the Hausman test cannot be rejected, $\chi^2(4) = 5.01$, $p = .29$, which leads to believe that the random effects model is indeed preferred, as was expected.

Table 4

Fixed effects model vs. random effects model for Hausman test

	Fixed effects	Random effects
Intercept	0.0053	-0.0064
Lag log tracking error + expenses	-0.4381***	-0.4380***
Volatility	-0.0189***	-0.0189***
Simple LIBOR	-0.0017	-0.0018
LETF shares outstanding	0.0000	0.0000
Leverage multiple	Omitted	-0.0087***
Issuer	Omitted	0.0153*
Area	Omitted	0.0193
Stocks in benchmark	Omitted	-0.0000
Hausman $\chi^2(4)$	5.01	
Probability > χ^2	0.2863	

Note: Table note on next page.

Table 4 note: Fixed effects dynamic panel regression and random effects dynamic panel regression of log tracking error on multiple independent variables. Independent variable is log tracking error plus expenses. Leverage multiple, Area, Issuer and Stocks in benchmark are time-invariant variables and are omitted in the fixed effects model due to collinearity with the funds. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

5.2 Determinants of tracking error

As was determined in the previous subsection, the random effects dynamic panel regression model is used to determine the coefficients and significance of the effect of independent variables on tracking error as described in (3). In this section, the seven models in table 5 are described. Since each model is an alteration or improvement of the previous model, for the first few models only the significance and the direction of the variables will be described. In the latter models the coefficients and their interpretations will be described in more detail. In all models, the regressand is log daily tracking error plus expenses: the logarithm of the difference between daily returns of LETF and returns of underlying leveraged benchmark before subtraction of daily expenses. Since tracking errors are usually described as negative numbers and since a positive coefficient leads to an increase in LETF return compared to the benchmark, a negative (positive) coefficient will be described as having an increasing (decreasing) effect on tracking error.

Table 5

LETF daily tracking error before expenses

	All funds				-1x fund	2x fund	3x fund
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Intercept	-0.0064	0.0128	-0.0030	0.0044	0.0020	0.0025	0.0211
	(0.0161)	(0.0283)	(0.0139)	(0.0052)	(0.0080)	(0.0070)	(0.0209)
Lag log tracking error + expenses[a]	-0.4380***	-0.4380***	-0.4380***	-0.4379***	-0.4250***	-0.4504***	-0.4304***
	(0.0027)	(0.0027)	(0.0027)	(0.0027)	(0.0053)	(0.0037)	(0.0062)
Leverage multiple	-0.0087***						
	(0.0021)						
Absolute leverage		-0.0167*					
		(0.0100)					
Inverse dummy[b]		0.0063					
		(0.0139)					
Issuer dummy[c]	0.0153*	0.0109					
	(0.0085)	(0.0100)					
Simple LIBOR[a]	-0.0018	-0.0020					
	(0.0024)	(0.0024)					
LIBOR Paid[a]			-2.5071***	-2.3216***	-1.8741***	-2.0175**	-2.1078
			(0.4446)	(0.4424)	(0.6325)	(0.9652)	(2.2078)
Volatility[a]	-0.0189***	-0.0189***	-0.0189***	-0.0172***	-0.0116**	-0.0122***	-0.0488***
	(0.0037)	(0.0037)	(0.0037)	(0.0036)	(0.0047)	(0.0045)	(0.0136)
Area[d]	0.0193	0.0220*	0.0189				
	(0.0118)	(0.0123)	(0.0117)				
Stocks in benchmark	-0.0000	-0.0000	-0.0000				
	(0.0000)	(0.0000)	(0.0000)				
LETF shares outstanding [e]	0.0000	0.0000	0.0000				
Overall R^2	0.1920	0.1921	0.1921	0.1919	0.1808	0.2030	0.1855
Number of funds	37	37	37	37	10	18	9
Observations	107,441	107,441	107,441	107,441	28,905	57,227	21,309

Note: Table note on next page.

Table 5 note: Random effects dynamic panel regression of log tracking error on multiple independent variables. Independent variable is log tracking error is the logarithmic difference between returns of leveraged exchange-traded funds (LETFs) and returns of underlying leveraged benchmark, without expenses. Leverage is the leverage level ranging from -1.0x to 3.0x. Simple LIBOR is the annualized rate, and LIBOR Paid is the daily rate multiplied by leverage multiple – 1. Volatility is the trailing 30-day volatility of the underlying index. Stocks in benchmark is the number of companies in the underlying index. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

[a] Percentage.

[b] 0 = Positive leverage, 1 = Negative leverage.

[c] 0 = Direxion, 1 = ProShares.

[d] 0 = rest of world, 1 = United States

[e] Expressed in thousands.

Model I in table 5 is the most basic model and improvements are added to it in the following models. In model I, all independent variables from (3) are used as regressors. The lagged tracking error is most significant and has a highly negative coefficient, implying that a log tracking error of one percent leads to a log tracking error of -0.44% the next day. This indicates the mean reversion that has been described in section 2. An arbitrage opportunity does not necessarily present itself here, because the mean reversion could in fact be explained by arbitrage due to the creation/redemption feature of exchange-traded funds. Recall that this feature allows authorized participants to exchange ETF shares for the underlying stocks. The next significant regressor is Leverage multiple. Its direction implies that higher leverage leads to a higher negative tracking error, which was expected because higher leverage is more costly to apply. The significance of the Issuer variable is somewhat surprising. It can be explained by ProShares hiring better managers than Direxion, but a different explanation is found by transforming the Leverage variable in model II. This will be addressed in the next paragraph. The final significant regressor in model I is benchmark volatility, of which the coefficient implies that an increase in volatility leads to a larger tracking error, ceteris paribus. This is in line with expectations from the literature. Simple LIBOR is not significant in model I. A more sophisticated version of the variable is used in model II. Area is insignificant. There was no reason to believe it was and no literature supporting it, so this is line with expectations. Stocks in benchmark and LETF shares outstanding are also not significant. Stocks in benchmark was expected to be significant with a positive coefficient, although the literature that provided this expectation applied the variable to a different type of tracking error research. Therefore its insignificance comes as no surprise. LETF shares outstanding was used as a proxy for liquidity and was thus expected to have a significantly positive coefficient too. Perhaps number of shares outstanding is not properly representative of liquidity and a different variable can be used in future research.

In model II in table 5, the Leverage variable from model I is split up into the absolute leverage multiple and a dummy variable for inverse funds. Model II shows that only the absolute leverage level has a significant coefficient, and that the Issuer and the Inverse dummy have insignificant effects. An increase in leverage from one to two, or from two to three, leads to a 0.02% decrease in log daily LETF return compared to its benchmark, ceteris paribus. Again, higher leverage leading to higher tracking error is in line with expectations because higher leverage is more costly for LETF issuers to apply. Like model I, model II shows similar significant coefficients for the tracking error lag and benchmark volatility. Simple LIBOR remains insignificant. Surprisingly, the Area coefficient becomes significant at the ten percent level. Although theoretical explanations can be found, the sudden significance after transforming other variables implies that its significance is a result of correlation with another variable. Regardless, from a statistical point of view the coefficient implies that funds tracking U.S. indices perform better compared to their benchmark, all else being equal. Stocks in benchmark and LETF shares outstanding are again not significant.

Model III in table 5 introduces another modification to the variables. Simple LIBOR and Leverage are combined into LIBOR Paid by converting LIBOR into a daily rate and multiplying by $m - 1$ as described in (1) in the data section. As expected, the coefficient is significantly negative. The magnitude of the coefficient, however, is surprising. One would expect the LIBOR Paid to have a literal one-to-one relationship with the tracking error and the coefficient to be -1. Models III and IV show that the coefficient is larger than expected by factor 2.5. The size of the coefficient being greater than one can be caused by a practical error in the statistical analysis or data collection, or it could have a theoretical explanation. To eliminate the suspicion that the coefficient is incorrect due to its calculation in (1) by involving the leverage multiple, models V, VI, and VII are created. These three models are modelled separately per leverage multiple. Although the LIBOR Paid coefficients change, they do not approach 1 enough to give reason to think that the error lies in the calculation of the variable. Surprisingly, the coefficient becomes insignificant for the 3x leveraged funds. Apart from this observation, the separate models provide a robustness check to see if the model performs similarly for funds with different leverage multiples.

Model IV is created using the improvements from models I and II and the robustness check in models V, VI, and VII, and by removing the insignificant variables from model III. This final model contains three variables that are significant at the 1% level. The first significant variable is lag tracking error. The coefficient implies that a 1% increased log tracking error leads to a -0.43% log tracking error the following day, ceteris paribus. This mean reversion is in line with expectations from the literature due to the creation/redemption feature of exchange-traded funds. The second variable is LIBOR Paid. This variable is calculated using (1). The coefficient can be interpreted as a 1% increase in daily LIBOR per $m-1$, with m the leverage multiple, leading to a −2.3% change in tracking error, all else being equal. Finally, a 1% increase in the volatility of the benchmark index leads to a −0.02% change in log tracking error, ceteris paribus. If one were to create a prediction model, this can be combined in

$$\begin{aligned} log\left(\mathrm{R}_{\mathrm{Diff},it} + \mathrm{Expense}_i\right) \\ = -0.4379 \cdot log\left(\mathrm{R}_{\mathrm{Diff},it-1} + \mathrm{Expense}_i\right) - 2.3216 \cdot (m_i - 1) \cdot LIBOR_t^{\frac{1}{360}} \\ - 0.0172 \cdot \mathrm{Volatility}_{it} + \varepsilon_{it}. \end{aligned} \tag{4}$$

As described in the methodology section, overall R^2 is used to compare the quality of the models. For all models, overall R^2 remains constant around 19%. This indicates that the models have the same explanatory power. However, R^2 not only increases when the explanatory power of the model increases, but also simply when new variables are added (R^2 here is not adjusted for this). This knowledge implies that when R^2 values are the same, the model with the least variables is best. Combined with the principle of simplicity, this gives added comfort to the conclusion that model IV is the best model for the sample funds.

6 Conclusion

LETFs are an increasingly popular way of applying leverage to index-tracking portfolios. While historical returns have been attractive, issuers warn that the instrument is unsuitable for holding periods longer than a day. This paper researches the performance of LETFs on a single day, specifically in comparison to their benchmark index, and adds to the available literature with the use of sample data until 2020 and a methodology that is easy to replicate. Using dynamic panel regression, three variables have been found to have significant effects on tracking error of LETFs in the sample, in line with the research described in the literature section. First, the 1-day lag has a significantly negative effect. This supports the existing literature and can be explained by mean reversion due to the creation/redemption feature of exchange-traded funds. Second, LIBOR has a significantly negative effect. This is explained by the derivative construction of LETFs, which uses total return swaps to exchange cash flows for floating interest rates. Last, volatility of the underlying benchmark has a significantly negative effect. Previous literature has found the same: more volatile indices are more difficult and therefore more costly to track.

Limitations to the research mostly occurred in the data collection. Datastream does not have data available for each fund or for each underlying index. It also does not have time-varying data for the number of stocks in benchmark indices or for expense ratios, so only the latest figures have been used. Next to this, Datastream corrected for stock splits, which occur often with inverse funds, in a way that made the data unsuitable for this research. All -2x and -3x funds have had to be eliminated, even though these could have given interesting insights in the more extreme variants of the LETF. This also has the unfortunate consequence that conclusions drawn from the sample in this paper cannot necessarily be projected onto the full LETF population. Judging from past literature, Bloomberg seems to provide higher quality data for LETFs and indices than Datastream does. Next to the data collection, some variables could have been constructed in a different way. As a variable representing liquidity, the number of LETF shares outstanding was used. It was not found to have a significant effect, but perhaps another variable such as fund flows are better indicators of liquidity. Finally, no information was available on fund managers, even though these are expected to influence fund performance.

To continue the quest for explanation of tracking errors of LETFs, I suggest future researchers start with the suggestions from the limitations mentioned in the previous paragraph. Using a larger sample with more detailed variables, while applying the same methodology, may provide researchers and investors with an even more detailed image of what to expect from investing with constant leverage ranging from -3 or lower to 3 or higher.

References

Avellaneda, M., & Zhang, S. (2010). Path-Dependence of LETF Returns. *SIAM Journal on Financial Mathematics, 1*(1), 586-603.

Bansal, V. K., & Marshall, J. F. (2015a). A tracking error approach to LETFs: Are they really that bad? *Global Finance Journal, 26*(1), 47-63.

Bansal, V. K., & Marshall, J. F. (2015b). Tracking error decomposition and return attribution for leveraged exchange traded funds. *Global Finance Journal, 28*(1), 84-94.

BlackRock. (2017). *Viewpoint: Index Investing Supports Vibrant Capital Markets* [White paper]. https://www.blackrock.com/corporate/literature/whitepaper/viewpoint-index-investing-supports-vibrant-capital-markets-oct-2017.pdf

Chicago Board Options Exchange. (2019). *Cboe Volatility Index®* [White paper]. https://cdn.cboe.com/resources/vix/vixwhite.pdf

Elton, E. J., Gruber, M. J., & Busse, J. A. (2004). Are Investors Rational? Choices among Index Funds. *The Journal of Finance, 59*(1), 261-288.

Financial Industry Regulatory Authority. (2021, February). *Margin Statistics*. https://www.finra.org/investors/learn-to-invest/advanced-investing/margin-statistics

Jaccard, J. (2001). *Interaction Effects in Logistic Regression.* SAGE Publications.

Loviscek, A., Tang, H., & Xu, X. E. (2014). Do leveraged exchange-traded products deliver their stated multiples? *Journal of Banking & Finance, 43*(1), 29-47.

Mosselaar, J. S. (2019). *A brief history of Robeco and the mutual fund industry* [White paper]. Robeco: https://www.robeco.com/media/0/8/a/08ad3a249b40c7c316ce00c0349c6874_a-brief-history-of-robeco-and-the-mutual-fund-industry-hk-sg_tcm19-22051.pdf

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science, 25*(3), 289-310.

Simon, R. (2000, April 27). Margin Investors Learn the Hard Way That Brokers Can Get Tough on Loans. *The Wall Street Journal*, C1.

Tang, H., & Xu, X. E. (2013). Solving the return deviation conundrum of leveraged exchange-traded funds. *Journal of Financial and Quantitative Analysis, 48*(1), 309-342.

Torres-Reyna, O. (2007). *Panel Data Analysis Fixed and Random Effects using Stata (v. 4.2).* Princeton University: https://dss.princeton.edu/training/Panel101.pdf

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

Vardharaj, R., Fabozzi, F. J., & Jones, F. J. (2004). Determinants of Tracking Error for Equity. *Journal of Investing, 13*(2), 37-47.

Wigglesworth, R., & Janiaud, A. (2020, January 8). Index funds break through $10tn-in-assets mark amid active exodus. *Financial Times*. https://www.ft.com/content/a7e20d96-318c-11ea-9703-eea0cae3f0de

Zellner, A., Keuzenkamp, H. A., & McAleer, M. (2002). *Simplicity, Inference and Modelling: Keeping it Sophisticatedly Simple.* Cambridge University Press. doi:10.1017/CBO9780511493164

Appendices

Appendix A. List of LETFs in sample

The final sample uses 37 LETFs as shown in table 7. Of 139 funds in the initial sample, 102 were removed for various reasons. The full list of removals is shown in table 8 along with the corresponding reason for removal.

Table A1

List of LETFs in sample

Symbol	Fund name	Underlying index	Leverage
CHAD	Direxion Daily CSI 300 China A Share Bear 1X	CSI 300	-1
CHAU	Direxion Daily CSI 300 China A Share Bull 2X	CSI 300	2
UDOW	ProShares UltraPro Dow30	Dow Jones Industrial Average	3
DDM	ProShares Ultra Dow30	Dow Jones Industrial Average	2
DOG	ProShares Short Dow 30	Dow Jones Industrial Average	-1
TPOR	Direxion Daily Transportation Bull 3X	Dow Jones Transportation Average	3
SBM	ProShares Short Basic Materials	Dow Jones U.S. Basic Materials	-1
UYM	ProShares Ultra Basic Materials	Dow Jones U.S. Basic Materials	2
UGE	ProShares Ultra Consumer Goods	Dow Jones U.S. Consumer Goods	2
UCC	ProShares Ultra Consumer Services	Dow Jones U.S. Consumer Services	2
SEF	Short Financials ProShares	Dow Jones U.S. Financials	-1
RXL	ProShares Ultra Health Care	Dow Jones U.S. Health Care	2
UXI	ProShares Ultra Industrials	Dow Jones U.S. Industrials	2
DDG	Short Oil & Gas ProShares	Dow Jones U.S. Oil & Gas	-1
DIG	ProShares Ultra Oil & Gas	Dow Jones U.S. Oil & Gas	2
NAIL	Direxion Daily Homebuilders Bull 3X	Dow Jones U.S. Select Home Constr.	3
USD	ProShares Ultra Semiconductors	Dow Jones U.S. Semiconductors	2
ROM	ProShares Ultra Technology	Dow Jones U.S. Technology	2
UPW	ProShares Ultra Utilities	Dow Jones U.S. Utilities	2
EURL	Direxion Daily FTSE Europe Bull 3X	FTSE Developed Europe	3
EFO	ProShares Ultra MSCI EAFE	MSCI EAFE	2
EFZ	Short MSCI EAFE ProShares	MSCI EAFE	-1
UPV	ProShares Ultra Europe	MSCI Europe	2
EDC	Direxion Daily Emerging Markets Bull 3X	MSCI Emerging Markets	3
EET	ProShares Ultra MSCI Emerging Markets	MSCI Emerging Markets	2
EUM	Short MSCI Emerging Markets ProShares	MSCI Emerging Markets	-1

Symbol	Fund name	Underlying index	Leverage
QLD	ProShares Ultra QQQ	NASDAQ-100	2
SH	ProShares Short S&P 500	S&P 500	-1
SPDN	Direxion Daily S&P 500 Bear 1X	S&P 500	-1
SPUU	Direxion Daily S&P 500 Bull 2X	S&P 500	2
SPXL	Direxion Daily S&P 500 Bull 3X	S&P 500	3
SSO	ProShares Ultra S&P 500	S&P 500	2
UPRO	ProShares UltraPro S&P 500	S&P 500	3
MIDU	Direxion Daily Mid Cap Bull 3X	S&P MidCap 400	3
MVV	ProShares Ultra Midcap 400	S&P MidCap 400	2
MYY	ProShares Short Midcap 400	S&P MidCap 400	-1
UMDD	ProShares UltraPro MidCap400	S&P MidCap 400	3

Note: This table shows the 37 leveraged exchange-traded funds used in the sample, along with their underlying index and leverage multiple.

Table A2

List of LETFs not considered in sample

Symbol	Fund name	Underlying	Leverage	Reason for removal
HDGE	AdvisorShares Ranger Equity Bear ETF	Actively managed	-1	Actively managed
KORU	Direxion Daily South Korea Bull 3X	MSCI Korea 25-50	3	Beta below 0.75
EZJ	ProShares Ultra MSCI Japan	MSCI Japan	2	Beta below 0.75
CWEB	Direxion Daily CSI China Internet Index Bull 2X	CSI China Overseas Internet	2	Fund data unavailable
UYG	ProShares Ultra Financials	Dow Jones U.S. Financials	2	Fund data unavailable
BIB	ProShares Ultra Nasdaq Biotechnology	NASDAQ Biotechnology	2	Fund data unavailable
TQQQ	ProShares UltraPro QQQ	NASDAQ-100	3	Fund data unavailable
MZZ	ProShares UltraShort Midcap 400	S&P MidCap 400	-2	Fund data unavailable
BIS	ProShares UltraShort Nasdaq Biotechnology	NASDAQ Biotechnology	-2	Fund data unavailable
WEBS	Daily Dow Jones Internet Bear 3X	Dow Jones Internet Composite	-3	Index data unavailable
WEBL	Daily Dow Jones Internet Bull 3X	Dow Jones Internet Composite	3	Index data unavailable
DFEN	Direxion Daily Aerospace & Defense Bull 3X	Dow Jones U.S. Select Aero & Defense	3	Index data unavailable
YANG	Direxion Daily China 3x Bear Shares	FTSE China 50	-3	Index data unavailable
YINN	Direxion Daily China 3x Bull Shares	FTSE China 50	3	Index data unavailable
WANT	Direxion Daily Consumer Discretionary Bull 3X	S&P Consumer Discretionary Select	3	Index data unavailable
ERY	Direxion Daily Energy Bear 2X	Energy Select Sector	-3	Index data unavailable
ERX	Direxion Daily Energy Bull 2X	Energy Select Sector	3	Index data unavailable
FAZ	Direxion Daily Financial Bear 3X	Russell 1000 Financial Services	-3	Index data unavailable
FAS	Direxion Daily Financial Bull 3X	Russell 1000 Financial Services	3	Index data unavailable
DUST	Direxion Daily Gold Miners Bear 2X	NYSE Arca Gold Miners	-3	Index data unavailable
NUGT	Direxion Daily Gold Miners Bull 2X	NYSE Arca Gold Miners	3	Index data unavailable

Symbol	Fund name	Underlying	Leverage	Reason for removal
CURE	Direxion Daily Healthcare Bull 3X	Health Care Select Sector	3	Index data unavailable
INDL	Direxion Daily India Bull 3X	MSCI India	3	Index data unavailable
DUSL	Direxion Daily Industrials Bull 3X	Utilities Select Sector	3	Index data unavailable
PILL	Direxion Daily Pharmaceutical & Medical Bull 3X	Dynamic Pharmaceuticals Intellidex	3	Index data unavailable
DPST	Direxion Daily Regional Banks Bull 3X	Solactive US Regional Banks	3	Index data unavailable
RETL	Direxion Daily Retail Bull 3X	Russell 1000 Retail	3	Index data unavailable
RUSL	Direxion Daily Russia Bull 2X	MVIS Russia	3	Index data unavailable
LABD	Direxion Daily S&P Biotech Bear 3X	S&P Biotechnology Select Industry	-3	Index data unavailable
LABU	Direxion Daily S&P Biotech Bull 3X	S&P Biotechnology Select Industry	3	Index data unavailable
SOXS	Direxion Daily Semiconductor Bear 3X	PHLX Semiconductor	-3	Index data unavailable
SOXL	Direxion Daily Semiconductor Bull 3X	PHLX Semiconductor	3	Index data unavailable
TZA	Direxion Daily Small Cap Bear 3X	Russell 2000	-3	Index data unavailable
TNA	Direxion Daily Small Cap Bull 3X	Russell 2000	3	Index data unavailable
TECS	Direxion Daily Technology Bear 3X	Technology Select Sector	-3	Index data unavailable
TECL	Direxion Daily Technology Bull 3X	Technology Select Sector	3	Index data unavailable
UTSL	Direxion Daily Utilities Bull 3X	Industrial Select Sector	3	Index data unavailable
UBOT	Direxion Robotics, Artificial Intelligence & Automation Index Bull 3X	Indxx Global Robotics & Artificial Intelligence Thematic	3	Index data unavailable
SMHB	ETRACS 2xMonthly Pay US Small Cap High Dividend ETN Series B	Solactive US Small Cap High Dividend	2	Index data unavailable
HDLB	ETRACS Monthly Pay 2x US High Dividend Low Volatility ETN Series B	Solactive US High Dividend Low Volatility	2	Index data unavailable
MJO	Indxx MicroSectors Cannabis 2X ETN	Indxx MicroSectors North American Cannabis	2	Index data unavailable

Symbol	Fund name	Underlying	Leverage	Reason for removal
BNKD	MicroSectors U.S. Big Banks Index -3x Inverse	Solactive MicroSectors U.S. Big Banks	-3	Index data unavailable
BNKU	MicroSectors U.S. Big Banks Index 3x ETN	Solactive MicroSectors U.S. Big Banks	3	Index data unavailable
NRGD	MicroSectors U.S. Big Oil Index -3X Inverse ETN	Solactive MicroSectors U.S. Big Oil	-3	Index data unavailable
EMTY	ProShares Decline of the Retail Store ETF	Solactive-ProShares Bricks and Mortar Retail Store	-1	Index data unavailable
YXI	ProShares Short FTSE China 50	FTSE/Xinhua China 25	-1	Index data unavailable
RWM	ProShares Short Russell 2000	Russell 2000	-1	Index data unavailable
SBB	ProShares Short Small Cap 600	S&P SmallCap 600	-1	Index data unavailable
XPP	ProShares Ultra FTSE China 50	FTSE/Xinhua China 25	2	Index data unavailable
UWM	ProShares Ultra Russell2000	Russell 2000	2	Index data unavailable
SAA	ProShares Ultra SmallCap600	S&P SmallCap 600	2	Index data unavailable
LTL	ProShares Ultra Telecommunications	Dow Jones U.S. Select Telecomm.	2	Index data unavailable
URTY	ProShares UltraPro Russell2000	Russell 2000	3	Index data unavailable
SRTY	ProShares UltraPro Short Russell2000	Russell 2000	-3	Index data unavailable
FXP	ProShares UltraShort China 50	FTSE/Xinhua China 25	-2	Index data unavailable
TWM	ProShares UltraShort Russell 2000	Russell 2000	-2	Index data unavailable
SDD	ProShares UltraShort Small Cap 600	S&P SmallCap 600	-2	Index data unavailable
HIBS	Direxion Daily S&P 500 High Beta Bear 3X	S&P 500 High Beta	-3	Launched after 2017
HIBL	Direxion Daily S&P 500 High Beta Bull 3X	S&P 500 High Beta	3	Launched after 2017
FNGZ	MicroSectors FANG+ Index -2X Inverse ETN	NYSE FANG+	-2	No fact sheet or exp. ratio
FNGO	MicroSectors FANG+ Index 2X ETN	NYSE FANG+	2	No fact sheet or exp. ratio
GNAF	MicroSectors FANG+ Index Inverse ETN	NYSE FANG+	-1	No fact sheet or exp. ratio
FNGD	MicroSectors FANG+™ Index -3X Inverse ETN	NYSE FANG+	-3	No fact sheet or exp. ratio
FNGU	MicroSectors FANG+™ Index 3X ETN	NYSE FANG+	3	No fact sheet or exp. ratio

Symbol	Fund name	Underlying	Leverage	Reason for removal
FIEE	UBS AG FI Enhanced Europe 50 ETN	STOXX Europe 50 USD	2	No fact sheet or exp. ratio
FIHD	UBS AG FI Enhanced Global High Yield ETN	MSCI World High Dividend Yield	2	No fact sheet or exp. ratio
FBGX	UBS AG FI Enhanced Large Cap Growth ETN	Russell 1000 Growth	2	No fact sheet or exp. ratio
MLPR	ETRACS Quarterly Pay 1.5x Alerian MLP	Alerian MLP	1.5	Quarterly rebalanced
BDCX	ETRACS Quarterly Pay 1.5x Wells Fargo BDC	Wells Fargo Business Development	1.5	Quarterly rebalanced
BRZU	Direxion Daily Brazil Bull 2X	MSCI Brazil 25/50	3	Values below 10.00
EDZ	Direxion Daily Emerging Markets Bear 3X	MSCI Emerging Markets	-3	Values below 10.00
JDST	Direxion Daily Junior Gold Miners Index Bear 2X	MVIS Global Junior Gold Miners	-3	Values below 10.00
JNUG	Direxion Daily Junior Gold Miners Index Bull 2X	MVIS Global Junior Gold Miners	3	Values below 10.00
LBJ	Direxion Daily Latin America 3x Bull Shares	S&P Latin America 40	3	Values below 10.00
MEXX	Direxion Daily MSCI Mexico Bull 3X	MSCI Mexico IMI 25-50	3	Values below 10.00
SPXS	Direxion Daily S&P 500 Bear 3X	S&P 500	-3	Values below 10.00
DRIP	Direxion Daily S&P Oil & Gas Exploration & Production Bear 2X	S&P Oil & Gas Exploration & Production Select Industry	-3	Values below 10.00
GUSH	Direxion Daily S&P Oil & Gas Exploration & Production Bull 2X	S&P Oil & Gas Exploration & Production Select Industry	3	Values below 10.00
PSQ	ProShares Short QQQ	NASDAQ-100	-1	Values below 10.00
UBR	ProShares Ultra MSCI Brazil	MSCI Brazil	2	Values below 10.00
SDOW	ProShares UltraPro Short Dow30	Dow Jones Industrial Average	-3	Values below 10.00
SMDD	ProShares UltraPro Short MidCap400	S&P MidCap 400	-3	Values below 10.00
SQQQ	ProShares UltraPro Short QQQ	NASDAQ-100	-3	Values below 10.00
SPXU	ProShares UltraPro Short S&P 500	S&P 500	-3	Values below 10.00
SMN	ProShares UltraShort Basic Materials	Dow Jones U.S. Basic Materials	-2	Values below 10.00
SZK	ProShares UltraShort Consumer Goods	Dow Jones U.S. Consumer Goods	-2	Values below 10.00

Symbol	Fund name	Underlying	Leverage	Reason for removal
SCC	ProShares UltraShort Consumer Services	Dow Jones U.S. Consumer Services	-2	Values below 10.00
DXD	ProShares UltraShort Dow 30	Dow Jones Industrial Average	-2	Values below 10.00
EPV	ProShares UltraShort Europe	MSCI Europe	-2	Values below 10.00
SKF	ProShares UltraShort Financials	Dow Jones U.S. Financials	-2	Values below 10.00
RXD	ProShares UltraShort Health Care	Dow Jones U.S. Health Care	-2	Values below 10.00
SIJ	ProShares UltraShort Industrials	Dow Jones U.S. Industrials	-2	Values below 10.00
BZQ	ProShares UltraShort MSCI Brazil	MSCI Brazil 25/50	-2	Values below 10.00
EEV	ProShares UltraShort MSCI EM	MSCI Emerging Markets	-2	Values below 10.00
EWV	ProShares UltraShort MSCI Japan	MSCI Japan	-2	Values below 10.00
DUG	ProShares UltraShort Oil & Gas	Dow Jones U.S. Oil & Gas	-2	Values below 10.00
QID	ProShares UltraShort QQQ	NASDAQ-100	-2	Values below 10.00
SDS	ProShares UltraShort S&P 500	S&P 500	-2	Values below 10.00
SSG	ProShares UltraShort Semiconductors	Dow Jones U.S. Semiconductors	-2	Values below 10.00
REW	ProShares UltraShort Technology	Dow Jones U.S. Technology	-2	Values below 10.00
SDP	ProShares UltraShort Utilities	Dow Jones U.S. Utilities	-2	Values below 10.00
EFU	ProSharesUltraShort MSCI EAFE	MSCI EAFE	-2	Values below 10.00

Note: This table shows the full list of leveraged exchange-traded funds (LETFs) that were not considered in the sample, along with their underlying index, leverage multiple, and reason for removal from the sample.

Appendix B. Correlation matrix

Table B1

Correlation matrix of independent variables

	Leverage multiple	Leverage absolute	Inverse dummy	Issuer dummy	LIBOR Simple	LIBOR Paid	Benchma rkolatility	Area dummy	Stocks in benchmark	LETF shares
Leverage multiple	1.0000									
Leverage absolute	0.9438	1.0000								
Inverse dummy	-0.9673	-0.8291	1.0000							
Issuer dummy	-0.2750	-0.4269	0.1377	1.0000						
LIBOR Simple	-0.0313	-0.0600	0.0070	0.0277	1.0000					
LIBOR Paid	0.6849	0.6198	-0.6829	-0.1501	0.1271	1.0000				
Benchmark volatility	-0.0455	-0.0569	0.0333	0.0247	0.0481	-0.0237	1.0000			
Area dummy	0.0821	0.0545	-0.0971	0.2198	0.0785	0.0798	0.0672	1.0000		
Stocks in benchmark	-0.0624	0.0209	0.1215	-0.2187	-0.0841	-0.0969	-0.1073	-0.7587	1.0000	
LETF shares outstanding	-0.0110	0.0143	0.0296	-0.0248	-0.0770	-0.0215	0.0531	0.1183	0.0881	1.0000

Note: This table shows the correlations between explanatory variables.