Financial Analysts Journal | A Publication of **CFA Institute**
https://doi.org/10.1080/0015198X.2020.1830660

Research

Levered and Inverse Exchange-Traded Products: Blessing or Curse?

Colby J. Pessina, and Robert E. Whaley

Colby J. Pessina is an investment banking analyst at Deutsche Bank, Jacksonville, Florida. Robert E. Whaley is the Valere Blair Potter Professor of Finance and director of the Financial Markets Research Center, Owen Graduate School of Management, Vanderbilt University, Nashville, Tennessee.

Levered and inverse exchange-trade products (ETPs) are designed to provide geared long and short exposures to the daily returns of various benchmark indexes. The benchmarks may be any reference index, but the popular ones are indexes of stocks, bonds, commodities, and volatility. The problem with these products is that they are not generally well understood, particularly those with futures-based benchmarks. Levered and inverse ETPs are neither suitable buy-and-hold investments nor effective hedging tools. They are unstable and exist only as mechanisms for placing short-term directional bets. Levered and inverse products are not, and cannot be, effective investment management tools.

Disclosure: The authors report no conflicts of interest.

PL Credits: 2.0

Levered exchange-trade products (ETPs) were, again, in the news in the spring of 2020. Citigroup's VelocityShares 3x Long Crude Oil exchange-traded note (UWT) nearly tanked on 9 March 2020, when its benchmark index dropped 73.5%, just shy of the 75% acceleration clause that would have triggered termination. On 15 March 2020, ProShares announced the liquidation of its UltraPro 3x Crude Oil (OILU) and UltraPro –3x Crude Oil (OILD) exchange-traded funds (ETFs). Although the 91.1% year-to-date decline of OILU was certainly basis for liquidation, OILD had risen 252.1%. Presumably, the liquidation of OILD was based not on an acceleration event or shareholder considerations but, rather, on the natural hedging that occurs in the futures replication of OILU and OILD. Invoking acceleration clauses, however, is nothing new. Two years earlier, Credit Suisse closed its Daily Inverse VIX Short Term ETN (XIV) after it lost 96.3% of its value in a single day.

Naturally, these types of events are confusing to most investors and undermine investor confidence in US financial markets. The fact of the matter is that the value-destroying characteristics of these products are not well understood. In November 2019, the US SEC proposed new regulations that would require brokers and advisers to ask their clients a series of questions before they could sell them leveraged ETPs.[1] Client responses would, presumably, help advisers determine whether investors understood all the risks inherent in the products (see Baer and Kierman 2020). The problem is inherently much deeper than investor education. Why do the products exist?

The main attraction of geared (levered) and inverse funds is that they offer an inexpensive, convenient, highly levered, and limited-liability means for profiting from a directional price view. If an investor has a strong view that the stock market will rise over next few weeks, an S&P 500 ETF, such as the ProShares UltraPro S&P500 (UPRO) or the Direxion Daily S&P 500 Bull 3X Shares (SPXL), might be suitable.

We would like to thank Stephen J. Brown, Tom Smith, Steven Thorley, CFA, and Joshua T. White.

© 2021 CFA Institute. All rights reserved.

Trading costs and barriers to entry are low. Bid–ask spreads on these products are trivial for active ETPs, and commission-free brokerage, with no account minimums, is now the norm.[2] A similar price risk exposure can be obtained in the CME Group's S&P 500 futures market, but trading costs are higher there and entering the market may not be feasible. Retail investors, for example, may not be able to open futures accounts because of account and personal wealth minimums. By buying shares of UPRO or SPXL on margin, however, the retail customer can quickly and cheaply achieve six times leverage on the S&P 500 Index. Similarly, some institutional investors are blocked from the derivatives market. Many US states, for example, limit the use of derivatives by public pension funds. Leveraged and inverse ETPs circumvent the rules. The Tennessee Consolidated Retirement System (TCRS) allows the use of stock index futures, but these vehicles may not be used for purposes of "speculative leveraging" and may not exceed 10% of the market value of the fund's total assets.[3] New York's Common Retirement Fund (CRF) limits the use of index futures to "rebalancing objectives."[4] Because leveraged and inverse ETPs are classified as equity securities, rather than derivatives, the risk management rules do not apply.

The most important problem with geared (levered) and inverse funds is that most of them are expected to collapse. The longevity of a levered or inverse ETP depends critically on the expected return and volatility of the fund's benchmark index. Cheng and Madhavan (2009) showed analytically that (1) if the expected return of the benchmark index is positive, inverse and levered inverse funds will eventually fail (i.e., their expected values are zero) and (2) if the expected benchmark return is negative, long and levered long funds will fail. We have extended these results in three important ways.

First, we demonstrate how the expected life of an ETP can be estimated by using Monte Carlo simulation, and we provide actual estimates for many of the most actively traded funds. Our simulation of a 3x natural gas ETF, for example, shows its life can be as few as 63 days and has a median of slightly more than a year (398 days).

Second, Cheng and Madhavan (2009) dismissed the negative expected index returns they found by saying that it "seems unlikely in a long-run equilibrium, but it is included for completeness" (p. 56). This conclusion is understandable. At the time, levered and inverse funds on commodity and volatility futures indexes were in their infancy. Now, they are among the most popular and the most controversial of such funds. Indeed, the recent fiascos mentioned in the opening paragraph all pertained to levered and inverse ETPs benchmarked to futures indexes. Futures-based indexes assume the idiosyncrasies of the futures contract prices on which they are based. In markets such as crude oil, natural gas, and volatility, long hedging demand generally exceeds short hedging demand. When it does, the futures price must rise to a level high enough for speculators to step in, sell futures to absorb the hedging demand imbalance, and earn a "fair" risk premium. Because the speculator sells the futures, the equilibrium expected futures return must be negative. We show, through simulation, that both highly levered long *and* short funds linked to such futures indexes are expected to fail and that ETPs on such futures indexes are expected to collapse *even if they have no leverage*![5]

Finally, the realized daily ETP return is different from its levered benchmark return for a variety of reasons, including management fees, licensing fees, operating costs, contract indivisibilities, basis risk or slippage, and front running. The difference between the daily return of the ETP and the daily return of the benchmark is called "tracking error" and is useful in assessing the degree to which issuers have achieved their stated investment objectives.

With all this information by way of backdrop, we now focus on the economics, mechanics, and empirics of levered and inverse funds—not because they provide opportunities to invest in new asset classes or manage the risk of investment portfolios in an effective way but, rather, because their inherent instability in volatile markets has led to financial market disarray and rapid-fire fund liquidations. We provide an appraisal of what these funds are intended to do—and why and when they can be expected to fail.

History of Levered and Inverse Funds

The evolution of ETPs in the United States has been fast paced. The first generation of products is characterized by the fact that they hold the securities that constitute the benchmark index. The benchmark may be stocks, bonds, or commodities, such as gold. They are classified as ETFs because their specific holdings are fully transparent on a daily basis. The first ETF launched in the United States was SPY, the SPDR S&P 500 ETF Trust, launched in January 1993.

The chief innovation of second-generation ETPs is that they are fully collateralized futures positions. In place of holding portfolios of securities, the fund holds T-bills (i.e., the collateral) and an *equal* notional amount of futures. Many are classified as ETFs because they report their holdings, the T-bills and the specific futures contract positions, each day. USCF's (United States Commodity Funds') US Oil Fund, USO, was the first to appear, in April 2006. It is long a nearby monthly crude oil futures contract. The nearby contract is rolled into the second nearby during the second week of the delivery month. ProShares' VIX (Cboe [Chicago Board Option Exchange] Volatility Index) Short-Term Futures ETF, VIXY, was launched in January 2011 as the first VIX ETF. Barclays was actually the first to launch a VIX ETP. It did so two years earlier, in January 2009. VXX, the iPath Series A S&P 500 VIX Short-Term Futures ETN (exchange-traded note) has a different structure from an ETF. As an ETN, it promises to deliver the return of the S&P 500 VIX Short-Term Futures Index, which is, itself, a fully collateralized VIX futures position. Because of the ETN structure, Barclays does not have to disclose exactly how it replicates the benchmark index. The firm may do the replication in any manner, including managing its risk with other types of volatility positions on its books. The only position that eliminates basis risk, however, is the one that replicates the index holdings—in this case, a dynamically rebalanced portfolio of VIX futures.

The third generation of ETPs are levered and inverse funds. To understand these funds, a single descriptor, *L*x, is needed, where *L* is the leverage ratio (or gear). When $L = 1$, the promised fund return is the return of the benchmark itself. Thus, the USO, VIXY, and VXX funds described previously are 1x funds. If the fund is $L = -1$, it is an inverse or short fund and promises a return of –100% that of the benchmark index. Note that neither the 1x nor the –1x funds are levered. Moreover, a –1x fund is fully collateralized. The notional amount of the futures matches the dollar amount of the T-bills; the difference is that instead of buying (going long) the futures, the fund sells (goes short). Leverage applies where the absolute value of *L* is greater than 1. The 2x designation refers to a two-times long fund, and –2x refers to a two-times inverse or short fund, and so on. Note that these funds are undercollateralized. That is, the amount of the T-bills remains the same. The futures positions in a 2x or –2x fund, however, are twice as large as they were before. Technically, 2x products can be referred to as "levered long funds" and –2x products as "levered short funds." Funds with leverage ratios between –1 and +1 (e.g., 0.5x) are overcollateralized funds; that is, the notional amount of the futures position is less than the dollar amount of the T-bills.

ProShares was the pioneer in the levered and inverse product space. In June 2006, it launched collateralized futures ETFs with 2x leverage and –1x inverse ratios on several major US stock indexes, including the S&P 500, the Nasdaq 100 Index, the Dow 30 Index, and the S&P 400 Index. In July 2006, ProShares followed up with –2x ETFs on the S&P 500 and Dow 30, and two days later, with –2x ETFs on the QQQ and S&P 400. The first challenger to ProShares' dominance in the levered equity ETF space was Direxion. In November 2008, it launched 3x and –3x ETFs on the S&P 500, the Russell 2000 Index, and certain sector indexes.

ProShares was also the first mover in the levered and inverse bond space. In April 2008, it launched –2x ETFs on Barclays' US 20+ Year Treasury Bond Index and Barclays' US 7–10 Year Treasury Bond Index. Direxion followed suit in April 2009 with 3x and –3x ETFs on the same bond indexes. Following a series of business acquisitions, those benchmark indexes are now named the ICE US Treasury 20+ Year and 7–10 Year Bond Indexes.

In the commodity space, the major players are ProShares (with ETFs) and Credit Suisse (with ETNs). Here, most of the trading volume is in the levered and inverse products on crude oil and natural gas. For crude oil products, ProShares launched 2x and –2x ETFs in November 2008 and Credit Suisse launched 3x and –3x ETNs in February 2012. Curiously, the Credit Suisse ETNs were delisted in December 2016, and Citibank filled the void almost immediately with identical product structures. For natural gas products, ProShares launched 2x and –2x ETFs in November 2011, and Credit Suisse launched 3x and –3x ETNs in February 2012.

Finally, in the volatility space, the two major players are, in descending order of assets under management, Credit Suisse and Barclays. Credit Suisse was the first mover, with the controversial TVIX 2x ETN in November 2010. The firm introduced the equally controversial XIV (–1x) ETN at the same time.[6] ProShares launched the ETF counterparts, UVXY and SVXY, in October 2011.

Current Products

To provide an understanding of the scope of the ETP market in the United States, we turned to ETFdb.com to identify a complete listing of active funds as of 13 March 2020. **Table 1** provides a summary by issuer. Many things stand out. First, currently 2,330 ETPs are listed on US stock exchanges, with total assets under management (AUM) of $4,231.6 billion. Second, BlackRock (iShares), Vanguard, and State Street account for 25.2% of the number of issues and 80.9% of the AUM—dominant market positions to say the least. Third, none of the top three issuers has issued levered and inverse funds. Neither has Charles Schwab, First Trust, VanEck, nor WisdomTree. Presumably, these ETP industry leaders made early business decisions to avoid levered and inverse products because of their complexity and propensity to provide surprising investor outcomes.

Table 1. ETPs by Issuer

		All Funds				Levered and Inverse Funds			
				Percent of Total				Percent of Total	
Fund Position	Issuer	No. of Issues	AUM ($)	Issues	$ AUM	No. of Issues	AUM ($)	Issues	$ AUM
1	iShares	368	1,640,157,916,000	15.8%	38.8%				
2	Vanguard	80	1,121,965,492,600	3.4	26.5				
3	State Street SPDR	140	661,112,371,900	6.0	15.6				
4	Invesco	219	214,569,456,300	9.4	5.1	1	37,027,100	0.4%	0.1%
5	Charles Schwab	25	155,923,124,100	1.1	3.7				
6	First Trust	150	83,892,827,500	6.4	2.0				
7	VanEck	55	41,983,292,900	2.4	1.0				
8	WisdomTree	77	35,742,604,900	3.3	0.8				
9	ProShares	140	33,611,387,700	6.0	0.8	108	24,107,533,200	39.0	47.6
10	J.P. Morgan	34	32,138,565,600	1.5	0.8				
11	PIMCO	16	23,961,158,400	0.7	0.6				
12	Goldman Sachs	22	17,362,803,700	0.9	0.4	1	1,291,261,200	0.4	2.6
13	Fidelity	29	16,466,728,900	1.2	0.4				
14	DWS	37	15,156,061,000	1.6	0.4	4	167,794,600	1.4	0.3
15	FlexShares	29	14,311,807,600	1.2	0.3				
16	Direxion	97	13,010,648,400	4.2	0.3	80	12,467,279,500	28.9	24.6
17	Global X	69	10,042,180,900	3.0	0.2				
18	ALPS	16	8,425,033,600	0.7	0.2				
19	Credit Suisse	20	5,765,620,500	0.9	0.1	14	5,635,376,000	5.1	11.1
20	Pacer Financial	22	5,755,511,400	0.9	0.1				
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
140	Armor Index	1	1,802,700	0.0	0.0				
Total		2,330	4,231,618,595,600	100.0%	100.0%	277	50,630,294,000	100.0%	100.0%

Source: ETFdb.com.

Fourth, of the 2,330 ETPs, 277 (or 11.9% of total ETPs) are levered and inverse funds. They account for only 1.2% of total AUM. However, although the percentage of total AUM may seem small, the ability of this segment to generate revenue is formidable. With the total AUM for levered and inverse funds at $50.6 billion and a typical expense ratio of about 100 basis points (bps), total revenue is on the order of $506 million a year.

Table 2 slices the levered and inverse ETP data another way—by asset class. By far, the largest single category, with AUM of $42 billion, is equities; it is about 83% of the total. Apparently, investors relish the opportunity to lever long and short stock market risk exposures. The bets are clearly short term because Table 2 shows the turnover rate is 16.9%, meaning the shares turn over every six days or so. Commodities and volatility products are next largest by AUM, with market shares of 6.1% and 5.0%, respectively. Here, the turnover rates are even higher—51.3% and 114.1%, respectively. Indeed, the volatility products are being day-traded. Their holding period (i.e., the inverse of the turnover rate) is 0.88—less than a day! The top three asset classes in Table 2 account for nearly 95% of AUM, and short-term betting is the primary trading motive.

Expected Levered and Inverse Fund Performance

The performance of levered and inverse funds can be addressed in two ways—as expected performance and as actual performance. We turn first to expected fund performance, and we do so in an environment of no management fees, operating expenses, and the like. We are examining only an index and the levered and inverse products benchmarked to that index. The two key elements in this discussion are the expected benchmark return and the compounding mechanics associated with leveraging.

Expected Benchmark Returns. Cheng and Madhavan (2009) derived two important results regarding levered and inverse funds that are relevant to this study. Under standard assumptions regarding asset price movements, they showed that (1) if the expected return on the benchmark index is positive, the long-run value of an inverse or levered inverse fund is zero and (2) if the expected return on the benchmark index is negative, the long-run value of leveraged long funds is zero. Cheng and Madhavan downplayed the importance of the second result because the levered and inverse funds that existed in 2009 were almost exclusively based largely on security indexes. The only futures-based products were ProShares 2x and –2x crude oil ETFs, and they were not launched until November 2008. With the focus on the Cheng and Madhavan security indexes, long-run equilibrium index returns are commonly expected to be positive. Indeed, the Sharpe (1964)–Lintner (1965) capital asset pricing model (CAPM) states that the expected return of a risky security should exceed the risk-free rate. Security indexes, such as the S&P 500, Russell 2000, and ICE US Treasury Bond 20+ Year Index, satisfy this criterion. We call these markets "carry markets" because active arbitrage activity occurs between the security market and the futures market. In the absence of costless arbitrage opportunities, the futures price will

Table 2. Levered and Inverse ETPs by Asset Class

Asset Class	No. of ETPs	Total AUM	Percent of AUM	Average Dollar Volume	Turnover
Equity	191	42,115,083,000	83.2%	7,133,238,098	16.9%
Commodity	26	3,106,084,200	6.1	1,591,937,806	51.3
Volatility	5	2,513,256,300	5.0	2,866,578,888	114.1
Bond	22	1,524,422,600	3.0	153,728,289	10.1
Real estate	9	866,891,600	1.7	13,355,573	1.5
Multiasset	4	285,840,600	0.6	3,679,366	1.3
Currency	19	214,679,300	0.4	4,419,779	2.1
Alternatives	1	4,036,400	0.0	48,395	1.2
Total	277	50,630,294,000	100.0%		

be at full carry and reflect the difference between the risk-free interest rate and the income rate on the underlying security. The futures market is critical to providing the necessary gearing of the levered and inverse funds.[7]

The reality is, of course, that levered and inverse commodity and volatility funds have gained a strong toehold. For the benchmark indexes, long-run equilibrium returns are determined not by the CAPM relationship but by an equilibrium established between hedgers and speculators in the relevant futures market. The key economic framework is that of Keynes (1930). He argued that in commodity futures markets, the price relationship between the futures and the underlying security/commodity is not actively arbitraged as it is for stocks and bonds. These "noncarry" markets exist because the cost of delivering the underlying asset is high and effective short selling is difficult if not impossible. Without active arbitrage, the futures price curve becomes a series of market-clearing futures prices at the various contract expirations. Keynes's model involves three sets of traders: Short hedgers are those who sell futures to hedge price risk. For example, consider a farmer who, as a matter of routine, hedges the crop he seeds in the spring by selling fall wheat futures to lock in the price at which he can sell his crop at harvest. On the other side of the trade may be a long hedger who buys futures to lock in the price at which it can buy wheat—a breakfast cereal producer, perhaps, which has locked in short-term sales contracts to grocers. In this example, short-hedging demand is larger than long-hedging demand. What happens? Speculators step in to pick up the imbalance—but they wait until the futures price falls below the expected spot price at the futures expiration date. In this example, the expected return of the futures contract over its life is positive. Keynes calls this futures market equilibrium "normal backwardation," a condition common to agricultural futures markets because short-hedging demand exceeds long-hedging demand.

The markets for crude oil, heating oil, and VIX futures typically display an opposite condition called "contango." In these markets, short-term futures prices are usually higher than expected spot prices because long-hedging demand exceeds short-hedging demand. Speculators step in only when the futures price rises high enough that they can earn a fair premium by selling the futures. In the crude oil market, the buying demand comes largely from airlines, which need to hedge the cost of fuel because their ticket sales have locked in revenue. In the natural gas market, the largest buyers are power companies, which need to hedge input costs in the production and sale of electricity. In the volatility market, the largest buyers are stock portfolio managers, who want to hedge tail risk. In all cases, the typical futures price curve is steeply upward sloping in the short term. Speculators sell—but at prices above the expected future spot price. The difference between the futures price and the expected spot price is the speculator's *risk premium*. In equilibrium, futures returns are expected to be negative to compensate the speculator for bearing risk.

Return Compounding Mechanics.

The *compounding effect* arises because the holding-period return actually delivered or the compounded daily levered return (i.e., *CLR* in Equation 1) is different from what is expected or the levered compounded return (i.e., *LCR* in Equation 1); that is,

$$\begin{aligned} CLR &= \prod_{t=1}^{T}(1+LR_t)-1 \\ &\neq L\left[\prod_{t=1}^{T}(1+R_t)-1\right] \\ &= LCR, \end{aligned} \tag{1}$$

where L is the leverage ratio, R_t is the daily benchmark index return on day t, and T is the holding period expressed in number of days. This relationship is confusing to retail investors, and justifiably so. Consider the following scenario. An investor buys a −2x fund and plans to hold it for two days (i.e., $L = 2$ and $T = 2$). The benchmark return ends up being 5% on Day 1 and −5% on Day 2. Thus, the two-day benchmark return is −0.25%. The investor, however, expected the return of the −2x fund to be 0.50%. The reality is, of course, that the two-day fund return is −1.00%. The investor not only receives 1.50 percentage points less than expected but also finds the two-day return to be of the opposite sign! The *only* instances in which the two sides of the equation are equal is when the holding period is a single day, $T = 1$, or the leverage ratio is one, $L = 1$.

Equation 1 is also the basis for concluding that levered and inverse ETPs are *not* effective hedging instruments for any holding period greater than one day. The reason is obvious: To be an effective hedge over a *T*-day hedging interval, the hedge return needs to be equal to *LCR*. If a university endowment holds US stock index funds as its equity allocation and wants to hedge stock market risk temporarily—perhaps, as the ramifications of the Covid-19 virus become more fully understood—selling S&P 500 futures offers an effective one-to-one passive hedge. Ideally, an inverse S&P 500 ETF should do the same thing, but it does not. The *CLR* is unpredictably different from the needed hedge return, and the difference widens the longer the holding period and the greater the stock market volatility. Of course, the inverse ETF hedge could be rebalanced dynamically in such a way that it would be effective, but that step defeats the purpose. Rebalancing costs would be prohibitive. The only (uninteresting) case in which the levered ETP is an effective passive hedge over more than one day is when the investor is hedging a position that is equal and opposite to the proposed hedge in a separate trading account.

To develop a clear understanding of the practical implications of the performance of the levered and inverse products, consider modeling the problem numerically in a Monte Carlo simulation. Monte Carlo simulation is an underused analysis tool that is ideal for application in contexts such as this one. It allows one to consider thousands of possible paths that the benchmark index might take, rather than just *one* path (i.e., the one that has been realized historically) over the life of a fund. So, we present the Monte Carlo simulation we carried out. We assumed that log returns are normally distributed with mean μ and standard deviation σ. We used daily data from the period 20 December 2005 through 13 March 2020 to estimate these parameters for six benchmark indexes. We used a common sample period for all indexes so that realized returns/volatilities could be compared on an apples-to-apples basis. The newest index we used is the S&P 500 VIX Short-Term Futures Index, which began reporting on 20 December 2005. The index-level data were downloaded from Bloomberg.

Table 3 provides the results of this analysis for the first three indexes we considered—total-return cash indexes in carry markets. SPX and RTY are equity indexes and correspond to the total returns of the S&P 500 and Russell 2000, respectively. IDCOT20T is a bond index and corresponds to the total return of the ICE US Treasury 20+ Year Bond Index. In equilibrium, the carry indexes should have positive expected returns; that is, over long histories, they should have positive realized returns. The daily returns for our 14½-year sample period confirm this expectation. As Table 3 indicates, SPX (Panel A) grew at a compound annual growth rate (CAGR) of 7.87% with an annualized volatility of 19.59%. The CAGR for RTY (Panel B) is lower with a higher return volatility, and the CAGR and volatility for IDCOT20T (Panel C) are similar to the data for SPX.

The second three indexes we considered are excess-return futures indexes in noncarry markets. The crude oil index, SPGSCLP, is created from the returns of the nearby monthly crude oil futures contract.[8] As time passes, the contract approaches its delivery date. On the fifth day of the delivery month, the index begins a five-day roll period. On the first day of the roll period, 20% of the nearby contract is replaced with the second nearby contract, 20% on the second day, and so on, over the remaining three days.

The natural gas index, SPGSNGP, is created in the same manner from the returns of the nearby monthly natural gas futures contract.

The VIX futures index, the S&P 500 VIX Short-Term Futures Index (SPVXSP), is based on the returns of the two nearby VIX futures and is rebalanced each day to maintain a constant one month to expiration. Each day, a proportion of the nearby futures contract is sold and replaced with the second nearby futures.[9]

All three index constructions represent long short-term futures contracts and are, therefore, vulnerable to the effects of contango. In equilibrium, such noncarry markets have negative expected returns. Over long histories, we would expect to see negative realized returns—which is exactly what we do see over our 14½-year sample period. **Table 4**, Panel A, reports that the crude oil index, SPGSCLP, produced a CAGR of –15.51% with an annualized volatility of 36.58%, the natural gas index, SPGSNGP, produced an even worse CAGR with greater volatility, and the volatility index, SPVXSP, produced the worst growth rate and highest volatility.

Each panel in Table 3 and Table 4 contains the simulation results of a single benchmark security index. A simulation run was 5,040 trading days (20 years) in length. We generated daily index levels from the historical CAGR and volatility in the top row of each panel. Then, based on these index levels, we computed daily holding-period returns. The daily holding-period returns, in turn, were used to compute the daily returns of the various levered products. We considered six levered products with integer values ranging from –3 to +3. In each run, the product's life ended on the day the fund's value fell below 5% of its initial level.[10] We used 10,000 simulation runs.

Table 3. Simulation of Life in Days for Levered and Inverse ETPs on Three Popular Security Indexes

Measure		Leverage Ratio −3	−2	−1	1	2	3
A. SPX: CAGR = 7.87%, volatility = 19.59%							
Life in days	Min.	439	915	2,444	5,040	2,597	853
	Med.	1,644	2,795	5,040	5,040	5,040	5,040
	Prob. < 20 years (%)	99.7	93.8	24.6	0.0	0.5	7.6
CLR (%)	Min.	−43.0	−30.9	−16.7	−19.1	−34.8	−47.8
	Med.	−4.2	−2.5	−1.1	0.8	1.3	1.4
	Max.	83.7	50.6	23.0	19.4	41.9	68.1
LCR – *CLR* (%)	Min.	−26.5	−12.5	−3.9	0.0	−3.3	−10.1
	Med.	0.8	0.4	0.1	0.0	0.1	0.4
	Max.	3.7	1.9	0.6	0.0	0.7	2.1
Prob(*CLR* × *LCR* < 0) (%)		4.4	3.2	1.5	0.0	1.0	1.9
B. RTY: CAGR = 5.70%, volatility = 24.53%							
Life in days	Min.	285	521	2,160	5,040	2,056	659
	Median	1,460	2,633	5,040	5,040	5,040	5,040
	Prob. < 20 years (%)	99.7	94.6	32.3	0.0	8.0	39.1
CLR (%)	Min.	−47.7	−34.7	−18.9	−25.1	−44.5	−59.2
	Med.	−4.1	−2.3	−0.9	0.4	0.3	−0.3
	Max.	126.3	73.8	32.4	22.7	49.6	81.6
LCR – *CLR* (%)	Min.	−50.8	−23.5	−7.3	0.0	−5.8	−16.2
	Med.	1.3	0.6	0.2	0.0	0.2	0.6
	Max.	6.8	3.4	1.1	0.0	1.1	3.2
Prob(*CLR* × *LCR* < 0) (%)		6.3	4.9	2.9	0.0	0.9	3.4
C. IDCOT20T: CAGR: = 7.69%, volatility = 14.11%							
Life in days	Min.	657	1,275	3,604	5,040	5,040	4,995
	Med.	2,218	3,652	5,040	5,040	5,040	5,040
	Prob. < 20 years (%)	98.7	85.8	5.8	0.0	0.0	0.3
CLR (%)	Min.	−32.1	−22.6	−11.9	−12.2	−23.0	−32.8
	Med.	−2.6	−1.6	−0.7	0.5	0.9	1.2
	Max.	45.3	28.6	13.5	13.3	28.0	44.4
LCR – *CLR* (%)	Min.	−8.8	−4.3	−1.4	0.0	−1.5	−4.6
	Med.	0.4	0.2	0.1	0.0	0.1	0.2
	Max.	1.8	0.9	0.3	0.0	0.3	0.9
Prob(*CLR* × *LCR* < 0) (%)		3.2	2.4	1.7	0.0	0.4	1.4

Table 4. Simulation of Life in Days for Levered and Inverse ETPs on Three Popular Futures Indexes

		Leverage Ratio					
Measure		−3	−2	−1	1	2	3
A. SPGSCLP (crude oil): CAGR = −15.51%, volatility = 36.58%							
Life in days	Min.	168	459	2,848	1,038	280	123
	Med.	2,698	5,040	5,040	4,411	1,635	859
	Prob. < 20 years (%)	87.7	49.7	2.9	66.8	98.7	100.0
CLR (%)	Min.	−67.9	−52.4	−30.6	−30.8	−52.9	−68.5
	Med.	−1.7	−0.1	0.4	−1.6	−4.1	−7.4
	Max.	178.8	100.6	42.6	42.2	99.9	177.8
LCR − *CLR* (%)	Min.	−86.3	−38.9	−11.7	0.0	−15.4	−51.1
	Med.	3.0	1.5	0.5	0.0	0.5	1.5
	Max.	12.7	6.2	2.0	0.0	2.1	6.4
Prob(*CLR* × *LCR* < 0) (%)		8.4	6.1	4.1	0.0	1.8	3.4
B. SPGSNGP (natural gas): CAGR = −35.94%, volatility = 44.90%							
Life in days	Min.	326	1,155	5,040	389	157	63
	Med.	5,040	5,040	5,040	1,688	693	398
	Prob. < 20 years (%)	48.6	10.6	0.0	99.7	100.0	100.0
CLR	Min.	−75.0	−59.4	−35.6	−39.8	−65.1	−80.6
	Med.	3.9	4.4	2.9	−4.4	−10.0	−16.8
	Max.	280.0	150.6	60.7	52.2	127.0	232.4
LCR − *CLR* (%)	Min.	−160.6	−71.0	−20.9	0.0	−22.7	−75.9
	Med.	3.6	1.8	0.6	0.0	0.6	1.5
	Max.	20.7	10.3	3.4	0.0	3.1	9.5
Prob(*CLR* × *LCR* < 0) (%)		8.5	6.6	4.6	0.0	1.4	4.2
C. SPVXSP (volatility): CAGR: −40.82%, volatility = 67.07%							
Life in days	Min.	44	107	663	258	67	41
	Med.	715	2,617	5,040	1,451	506	253
	Prob. < 20 years (%)	99.9	88.2	18.6	99.8	100.0	100.0
CLR	Min.	−87.0	−71.6	−45.2	−47.3	−74.1	−88.2
	Med.	−6.6	−1.0	1.3	−5.2	−13.0	−23.2
	Max.	404.5	206.4	79.4	74.3	191.7	370.6
LCR − *CLR* (%)	Min.	−266.8	−112.1	−32.1	0.0	−43.1	−147.6
	Med.	10.3	5.1	1.6	0.0	1.5	4.2
	Max.	40.8	21.5	7.3	0.0	7.0	25.4
Prob(*CLR* × *LCR* < 0) (%)		13.8	10.3	6.1	0.0	3.3	6.7

Before we examine the results, we repeat that the simulation results were conducted in a perfect world with no fees, costs, or other trading impediments. The finite lives of these products in these simulations were driven solely by the expected return/volatility characteristics of the benchmark and the compounding mechanics.

Turning to the results of Table 3, we focus first on Panel A, which covers SPX, the stock market poster child. The section labeled "Life in days" reports the minimum and median days of the life of the fund under different leverage assumptions. Note that when L = 1, the minimum number of days is 5,040. Recall that the two sides of Equation 1 are equal when L = 1. As we look down that column, we see that none of the simulation runs ended before 20 years had elapsed; the probability that the life is less than 20 years is zero. This result makes sense. The stock market has never fully collapsed. Now, consider the levered inverse fund –3x. The median life of the fund is 1,644 trading days, or about 6½ years, and its life could be as little as 439 days. These SPX results are consistent with the Cheng and Madhavan (2009) prediction that negative-leverage funds on benchmarks with positive expected returns will eventually collapse. Even at the other extreme, high positive leverage, fund collapse is possible. Consider L = 3 for Panel A. The median fund life is 5,040, reflecting the fact that most of the simulation runs did not lead to fund collapse within the 20-year period. But 7.6% of the runs had early termination, and one run lasted only 853 days.

The rest of the rows in each panel of Table 3 quantify the range of differences between the compounded levered return, *CLR* (i.e., what is actually earned on the levered fund), and the levered compounded return, *LCR* (i.e., the return the retail investor expects). The "*LCR* – *CLR*" rows show the difference between the levered compounded and the compounded levered returns. The last row in each panel reports the frequency with which *CLR* had a different sign from that of *LCR*, a particularly counterintuitive result. For the SPX results in Panel A of Table 3, as noted previously, the return differential is zero when L = 1 (because the two sides of Equation 1 are equal when the fund is not levered). For the L = –3 fund, the investor earned as much as 26.5 percentage points less than expected and had returns with opposite signs 4.4% of the time. At L = 3, the investor earned as much as 10.1 percentage points less than expected and had returns with opposite signs 1.9% of the time.

The results for the Russell 2000 total-return index, RTY, and the ICE US Treasury 20+ Year Bond Index, IDCOT20T, are shown in, respectively, Panels B and C of Table 3. The results are qualitatively similar to those for SPX. An additional insight brought about by the results, however, is that the range of results varies directly with the volatility of the benchmark. The median life of the various levered funds is longest for IDCOT20T, for which volatility is 14.11%, and shortest for RTY, for which volatility is 24.53%. RTY's volatility is higher than that for SPX because the Russell 2000 is a midcap rather than large-cap benchmark.

The noncarry market simulation results reported in Table 4 are much more extreme than the results in Table 3. Recall that these noncarry markets have negative expected returns because the futures price curves in the natural gas, crude oil, and volatility markets are typically in contango. In Table 4, the leverage and compounding effects behave in some unexpected ways. Whereas Cheng and Madhavan (2009) predicted that the positive-leverage funds will collapse, the simulation results indicate that highly levered long *and* highly levered inverse funds will collapse. The crude oil results, reported in Panel A of Table 4, show that the 3x fund has a median life of 859 days, might have a life as short as 123 days, and is certain to collapse within 20 years. Even at L = 1, the chance of collapse is 66.8%. Now, consider the levered inverse funds. The situation is most stable at L = –1, with only a 2.9% chance that the fund will fail. This result makes sense because L = –1 means the investor is taking an unlevered short position in a benchmark index that has a negative expected return. At L = –3, however, the chance of collapse is 87.7%. The median life is less than 11 years, and the life could be as short as 168 days.

The results for natural gas in Panel B of Table 4 and volatility in Panel C are even more dramatic. Both have expected returns lower than and volatility rates higher than crude oil. The 3x SPGSNGP fund, for example, has an expected life of about a year and a minimum of 63 days. The 3x SPVXSP fund has an expected life of 253 days, and the life could be as short as 41 days. The probability of a volatility 3x fund not having a life of 20 years is 100%. With the large negative expected returns of the natural gas and volatility funds, the –3x funds shown in Panels B and C are expected to last modestly longer. For volatility, however, the chances are 99.9% of failure within the 20-year horizon.

In summary, the Monte Carlo simulations indicate that levered and inverse funds on commodities and volatility will have a value below 5% of their initial value within a few years. What should be most disconcerting to investment professionals is that this type of analysis could have been performed before any commodity and volatility products were launched. Monte Carlo simulation is an invaluable investment analysis tool.

In the simulations, we used a 5% stopping criterion to represent the maximum loss an investor is willing to tolerate before liquidating her investment in the fund. As a practical matter, the reason levered and inverse ETP prices never actually reach zero is that issuers engage in periodic reverse splits (ostensibly to avoid liquidation and preserve the management fee revenue stream). A reverse stock split is an action taken by a corporation (or, in this case, an ETP issuer) to inflate the share price. It is done by consolidating the number of shares by a factor such as 2 or 5 in the case of a 1-for-2 or 1-for-5 reverse split. Because the aggregate market value of shares does not change, share prices increase by the same factor. An example occurred on 21 April 2020, when ProShares issued a 1-for-25 reverse stock split for its Ultra Bloomberg Crude Oil (UCO) fund, a 3x crude oil ETF. As a result of a dramatic oil oversupply and a lack of storage capacity, UCO's share price had dropped to $1.35. The 1-for-25 reverse stock split occurred overnight, and by the next morning, UCO was trading at nearly $34 a share, momentarily alleviating ProShares' concern about potential fund collapse and liquidation (Dierking 2020).

To provide a flavor of this amusing practice, we collected the reverse split histories of four VIX ETPs—VXX, VIXY, TVIX, and SVXY. We then computed the number of days between adjacent splits, the ETP returns since the last splits, and the implied stopping criteria. The results are reported in **Table 5**. For an interpretation of the results, consider, first, Panel A

Table 5. Reverse Split Histories of 1x and 2x VIX ETPs since Each Fund's Inception

		Since Last Change					Since Last Change		
Date	Reverse Ratio	Days	Return	Stop	Date	Reverse Ratio	Days	Return	Stop
A. VXX (1x): Inception 2009/01/30					*B. VIXY (1x): Inception 2011/01/04*				
2010/11/09	1:04	648	−89.1%	10.9%	2013/06/10	1:05	888	−89.1%	10.9%
2012/10/05	1:04	696	−81.2	18.8	2016/07/25	1:05	1,141	−81.2	18.8
2013/11/08	1:04	399	−64.4	35.6	2017/07/17	1:04	357	−64.4	35.6
2016/08/09	1:04	1,005	−81.2	18.8					
2017/08/23	1:04	379	−67.4	32.6					
Mean		625.4	−76.7	23.3	Mean		795.3	−78.2	21.8
Median		648	−81.2%	18.8%	Median		888	−81.2%	18.8%
C. TVIX (2x): Inception 2010/11/30					*D. UVXY (2x): Inception 2011/10/04*				
2012/12/21	1:10	752	−99.1%	0.9%	2012/03/08	1:06	156	−85.2%	14.8%
2013/08/30	1:10	252	−79.7	20.3	2012/09/07	1:10	183	−88.6	11.4
2015/06/23	1:10	662	−96.6	3.4	2013/06/10	1:10	276	−82.4	17.6
2016/08/09	1:25	413	−87.3	12.7	2014/01/24	1:04	228	−70.2	29.8
2017/03/16	1:10	219	−83.7	16.3	2015/05/20	1:05	481	−88.7	11.3
2018/06/08	1:10	449	−88.3	11.7	2016/07/25	1:05	432	−85.4	14.6
2019/12/02	1:10	542	−83.8	16.2	2017/01/12	1:05	171	−79.8	20.2
					2017/07/17	1:04	186	−73.4	26.6
Mean		469.9	−88.4	11.6	Mean		264.1	−81.7	18.3
Median		449	−87.3%	12.7%	Median		207	−83.8%	16.2%

for VXX (Barclays iPath Series A S&P 500 VIX Short-Term Futures ETN), a 1x ETN whose first day of trading was 30 January 2009. Just 648 days later, its price had fallen –89.1%, whereupon Barclays performed a 1-for-4 reverse split. In other words, when the share price reached 10.9% of its original amount, it was stopped and recalibrated.[11] After another 696 days, the share price dropped by –81.2% and the issuer performed another 1-for-4 reverse split. The median number of days between reverse splits for VXX is 648.

VIXY (ProShares VIX Short-Term Futures ETF) is the ETF counterpart to VXX. It has a shorter history and a median number of days between reverse splits, as shown in Panel B of Table 5, of 888. TVIX (Credit Suisse's VelocityShares Daily 2x VIX Short-Term ETN), depicted in Panel C, is a 2x ETN on the same benchmark, SPVXSP. Because its leverage ratio is higher than that of VXX, we would expect the frequency of reverse splits to be higher. The results, shown in Panel C, confirm this expectation. TVIX had more splits than VXX, and the splits occurred more frequently. The median number of days between splits is 449. The most interesting comparison is TVIX and UVXY (ProShares Ultra VIX Short-Term Futures ETF), which is depicted in Panel D. It is also a 2x fund and has a median number of days between reverse splits of only 207. The difference is, of course, that the issuers have different risk management policies. Credit Suisse allowed TVIX to fall to a much lower price before engaging in a reverse split than ProShares did with UVXY. A possible reason for this difference is that TVIX had no options listed on it whereas UVXY did. Trading in options is encumbered when share prices get to exceedingly low levels.

Actual Fund Performance

Consider now the actual performance of ETPs. Aside from the fact that levered and inverse funds are theoretically flawed from a long-term investment/risk management perspective, they may also have problems with implementation. The question is how well issuers perform at providing the levered benchmark return. To analyze this issue, we performed tracking-error analyses of 35 popular funds. *Tracking error* here is the difference between the daily returns of the ETP and the daily returns of the benchmark:

$$TE_t = R_{F,t} - LR_{B,t}, \quad (2)$$

where $R_{F,t}$ and $R_{B,t}$ are the daily returns of, respectively, the fund and its benchmark index and L is the leverage ratio. Note that this measurement is not risk adjusted; it does not have to be. Whatever market price risk factors affect the ETP returns also affect the benchmark returns. The mean of the tracking error is called the "tracking difference." The standard deviation of the tracking error shows the effectiveness of the issuer at replication. The *t*-ratio for the null hypothesis is that the tracking difference is zero; it was computed by dividing the tracking difference by its standard error (i.e., the standard deviation divided by $\sqrt{T}$). The meaning of this ratio, however, extends beyond statistical significance. It is a risk-adjusted tracking difference (RATD). For example, if two funds have the same tracking difference, the fund with the highest ratio is preferable. Because of its simple, intuitive appeal, we dub it the "RATD ratio." With competing funds benchmarked to the same index, the fund with the highest RATD ratio over the evaluation period, even if it is negative, has provided the best performance. The practical reasons why RATD ratios vary fall into two broad categories—(1) differences between the market price of the ETP and its net asset value (NAV) and (2) differences between the NAV and the level of the underlying benchmark.

Market Price vs. NAV per Share.

Differences between market price and NAV per share are driven largely by the authorized participant (AP) and his or her interactions with the ETP issuer. Working together, the AP and the issuer ensure that the market value of the ETP's shares outstanding equals the sum of the market values of the securities/derivatives that compose the benchmark index (i.e., the aggregate NAV of the fund).

The essential elements of this arbitrage are best illustrated by narrowing the focus for the moment. Assume the fund is created with the use of a collateralized futures position. Recall that the second-generation funds were unit leverage (1x and –1x) funds on commodities or volatility. Hence, they were created on the basis of a fully collateralized futures position. The NAV of the example fund equals the size of its cash (T-bills). The fund also buys a 1x or sells a –1x notional amount of futures. ProShares' VIXY is an example of such a 1x product. Suppose that on a given day, excess buying pressure is occurring on VIXY in the secondary market. On the other side of the customer trades is the AP. As he sells the shares of VIXY, he hedges by buying the nearby and second nearby VIX futures in the weights prescribed by the fund's benchmark, the S&P 500 VIX Short-Term Futures Index (SPVXSP).[12] Just before the close of trading, the AP tallies up the notional amount of his futures positions and notifies ProShares of his intention to do a *creation*. The AP delivers the notional amount in cash, receives shares (or units) of VIXY, and liquidates the VIX futures position. This "in-kind" transaction ensures that the market price of the ETF is exactly equal to the benchmark's indicative value.

A *redemption* works exactly in reverse. If VIXY is experiencing excessive selling pressure on a given day, the AP delivers shares of VIXY and receives cash, which he then uses to cover his long positions. As a practical matter, the issuer creates or redeems in blocks of shares. For VIXY, a block or creation unit is 50,000 shares.

The difference between the market price and the NAV is the *premium/discount* on the fund. For such products as VIXY, the premium/discount is small.[13] The VIX futures market is deep and liquid, resulting in low trading costs. The less deep and liquid the underlying futures market, however, the greater the premium/discount. To the extent that the level of the premium/discount persists over time, it will be reflected in the tracking difference (TD). Also, APs pay a fee for each creation/redemption. The effects of such fees are usually small. In the case of VIXY, for example, the maximum fixed fee is $250 per *creation unit*.[14]

The third-generation funds were the L ≠ 1 products. They remain collateralized futures positions, but they are either undercollateralized or overcollateralized. Again, the NAV of the fund equals the amount of cash (T-bills) it holds. In this case, however, the fund also buys or sells a notional amount of futures equal to *L* times the NAV. ProShares has two such VIX ETFs: The UVXY fund is a 2x ETF based on the SPVXSP, and the SVXY fund is –1x.[15]

A final caveat regarding the market price/NAV relationship is in order. If anything impedes the creation/redemption arbitrage, the ETP is, in effect, transformed from an open-end to a closed-end mutual fund, and the premiums/discounts can be substantial. On 21 February 2012, for example, Credit Suisse announced that it was suspending creations and redemptions of its popular 2x TVIX ETN. Excess demand to buy TVIX drove the share price up over the ensuing days to a premium of nearly 90% more than NAV per share on 21 March 2012. When the creation/redemption process resumed on 22 March, the share price dropped 30%, followed by another 30% drop on 23 March.

Net Asset Value vs. the Benchmark.

Differences between the fund's NAV and the underlying benchmark's value come in a variety of forms. The only explicit cost is the *total expense ratio*. It is expressed as a simple annual rate and covers the fund's operating costs and management fee. The amount of cash taken from the assets of the fund each day is usually computed as

$$\text{Cash} = ER \times \left(\frac{n}{365} \right) \times NAV_{t-1}, \tag{3}$$

where *ER* is the expense ratio, *n* is the number of calendar days between business days $t-1$ and t, and NAV_{t-1} is the net value of the fund at the close on business day $t-1$. For the ETPs considered in our sample, the expense ratio ranged from 0.89% to 1.65%. On average, the ratio amounts to a half of 1 bp per day.

The remaining costs are implicit. *Indivisibility costs* arise from the fact that futures contracts have a standardized size and fractional contracts cannot be traded. The easiest way to see this situation in practice is by examining ETF holdings. In **Table 6**, we show the holdings of VIXY, UVXY, and SVXY on 13 March 2020.[16] The highlighted fields are directly from the funds' files, which are available to investors each day. The other entries are from various sources (as provided in Table 6) and were used to verify ProShares' computations.

Table 6 is informative in a number of ways. Panel A contains the holdings of the 1x VIXY. As of the close on 13 March 2020, the fund held 663 18 March 2020 VIX futures with an exposure value of $35,420,775. Because VIX futures have a contract denomination of 1,000, the *implied* price of the 18 March 2020 futures contract is $35,420,775/663,000, or $53.425. From the Cboe website, we found that the *actual* settlement price was $53.425. The implied and actual values match exactly. We performed the same analysis for the 15 April 2020 VIX futures and produced the same result. Beneath the futures contracts in Panel A is a row labeled "Net other assets/Cash." Presumably, this label refers to cash or T-bills, and the amount should equal the NAV of the fund. We confirmed this equality by downloading a second data set

Table 6. Holdings and NAVs of ProShares VIX ETFs at the Close on 13 March 2020

Futures	Settlement Price	SPVXSP Weight	Source
VIX futures 2020/03/18	53.425	10.000%	Cboe
VIX futures 2020/04/15	43.900	90.000%	Cboe

ETN	Closing Price	Source
VXX	43.200	Bloomberg

A. VIXY: ProShares VIX Short-Term Futures ETF (1x)

Holdings	Shares/Contracts	Exposure Value	Market Value	Implied Price	Index Weights
VIX futures 2020/03/18	663	35,420,775		53.425	10.009%
VIX futures 2020/04/15	5,961	261,687,900		43.900	89.991%
Net other assets/Cash	297,283,619		297,283,619		
Total exposure value		297,108,675			
Cash/NAV/Difference		297,283,619	297,279,064	4,555	
Leverage ratio		0.999412			

B. UVXY: ProShares Ultra VIX Short-Term Futures ETF (1.5x)

Holdings	Shares/Contracts	Exposure Value	Market Value	Implied Price	Index Weights
VIX futures 2020/03/18	2,318	123,839,150		53.425	10.018%
VIX futures 2020/04/15	20,820	913,998,000		43.900	89.982%
VXX*	1,542,184	67,014,447		43.454	
Net other assets/Cash	736,727,443		736,727,443		
Total exposure value		1,104,851,597			
Cash/NAV/Difference		736,727,443	736,734,249	–6,805	
Leverage ratio		1.499675			

C. SVXY: ProShares Short VIX Short-Term Futures ETF (–0.5x)

Holdings	Shares/Contracts	Exposure Value	Market Value	Implied Price	Index Weights
VIX futures 2020/03/18	–609	–32,535,825		53.425	9.935%
VIX futures 2020/04/15	–5,521	–242,371,900		43.900	90.065%
Net other assets/Cash	549,485,607		549,485,607		
Total exposure value		–274,907,725			
Cash/NAV/Difference		549,485,607	549,508,907	–23,300	
Leverage ratio		–0.500300			

Notes: Shaded values in panels are directly from 13 March 2020 ProShares' daily holdings file. All other panel values are computations based on ProShares' data. VXX* is Barclays' iPath Series B S&P 500 VIX Short-Term Futures Swap – GS.

Source: ETF holdings at www.proshares.com/resources/data_downloads.html.

from the ProShares website. It contains the entire daily NAV histories for all ProShares ETFs.[17] The 13 March 2020 NAV for VIXY from the NAV file was $297,279,064, as indicated in a shaded cell in Table 6. The reported cash amount from the holding file was $297,283,619. The small difference is probably the result of rounding or interest income considerations.

At the top of Table 6 are the weights applied to the nearby and second nearby VIX futures settlement prices on 13 March 2020 when we computed the SPVXSP settlement level. The weights are 10% on the 18 March 2020 contract and 90% on the 15 April 2020 contract. The ProShares positions reflect these weights as best they can. Because fractional numbers of contracts cannot be bought or sold, the actual leverage ratio will generally not be equal to the promised leverage ratio. For VIXY, in Panel A, the actual leverage ratio is 0.999412x, not the promised 1x. As a result, NAV returns will not, other considerations being held constant, exactly match the SPVXSP returns. Although the effect of this discrepancy should be randomly distributed around zero and not affect the level of TD, it causes the standard deviation of the tracking error to increase and the RATD ratio to fall.

The two remaining panels of Table 6 reveal similar contract indivisibility effects. The ProShares replication holdings, in theory, should reflect the benchmark index weights. In the case of SVXY, in Panel C, the actual leverage ratio, –0.500300, is only slightly less than promised. The same is true for UVXY in Panel B. The replication holdings include a swap position. Judging by the note, "iPath Series B S&P 500 VIX Short-Term Futures Swap – GS," the swap has Goldman Sachs as a counterparty and VXX as a reference asset. Although ProShares is on the receive side of the VXX return, what they pay is not clear. The closing price of VXX on 13 March 2020 was $43.20, according to Bloomberg. The implied price from the information in the holdings file is $67,014,447/1,542,184, or $43.454. The cause of the discrepancy is not clear, but the 13 March 2020 date lies in a period of unusually high volatility because of the Covid-19 virus. Similar analyses of earlier dates showed little difference between the implied and actual VXX prices. Such differences contribute to the *basis risk* between NAV and the benchmark.

End-of-day rebalancing costs may contribute to the basis risk in at least two ways. First, a levered and inverse fund must rebalance its derivatives position at the end of day in order to deliver its promised levered return on the following day. For security indexes, the rebalancing is accomplished by using total-return swaps, as pointed out by Cheng and Madhavan (2009). Because these trades are OTC trades, basis risk can be minimized since the timing of the swap trade is unencumbered by exchange hours and can be negotiated when settlement prices are known. But for futures-based indexes, the primary hedge instruments are exchange-traded futures contracts. The incremental number of futures contracts required at the end-of-day settlement price is

$$n_{F,t} - n_{F,t-1} = \left(\frac{NAV_{t-1}}{F_t}\right)(L^2 - L)R_{F,t}, \qquad (4)$$

where $n_{F,t}$ is the number of futures held at the close of day t, NAV_{t-1} is the net asset value of assets under management on day $t-1$, F_t is the futures price at the close on day t, and $R_{F,t}$ is the benchmark return on day t.[18] These futures demands are directly proportional to the term $L^2 - L$. For both 3x and –2x funds, for example, $L^2 - L = 6$. More importantly, the incremental hedging demand is in the same direction as the return of the futures index benchmark for the day, which is not fully resolved until end-of-day settlement. Because the transactions must be executed before the trading period ends, the trade prices, by definition, will have to be different from settlement prices unless they are the same by chance; hence, basis risk will be incurred.

Front running is the second concern with these types of trades. Levered and inverse ETPs must execute their replication trades as close as possible to settlement to minimize basis risk. The end-of-day futures demand expressed by Equation 4, however, although not fully resolved before the close, is almost fully known by the market well before the close. This gap leaves open the possibility that certain traders, knowing that the rebalancing demand at the close will be large and positive (negative), will step in ahead of issuers, buy (sell) futures, and then unwind at the close. Although this type of activity exacerbates the end-of-day price movement (which should reverse in the next trading interval), it does not necessarily contribute to basis risk or, hence, tracking error.

Tracking-Error Analysis. Now that we have identified the sources of tracking error, we turn to analyzing tracking error for 35 of the largest and most active ETPs in our four asset categories—stocks, bonds, commodities, and volatility.

We proceed in two steps. First, we examine the tracking error associated with stock and bond ETFs. Recall that these funds are characterized as carry markets because of the active arbitrage between the derivatives markets and the underlying security markets. **Table 7** contains the results. Second, we show the same analyses of the noncarry markets—commodities and volatility. The benchmark indexes are affected by the persistent contango in the futures market and have negative expected rates of return. The results are in **Table 8**.

The top panel of Table 7 contains the results for the ETFs benchmarked to the S&P 500 total-return index. Beginning with the −3x products, note Direxion's SPXS and ProShares' SPXU. SPXS has an expense ratio of 1.08%, whereas SPXU has one of 0.91%. The tracking difference is higher for SPXU than SPXS by about half a basis point, and the standard deviation is higher by 3.3 bps. The RATD ratio for SPXU is higher, meaning that it outperformed SPXS, and its absolute magnitude indicates that the tracking difference is nearly significantly different from zero from a statistical standpoint. Continuing across the rows, we show the results from a regression of the daily ETF return on the daily return of the S&P 500 benchmark. The estimated beta for SPXS, −2.998, is nearly exactly its promised level of −3. The *t*-ratio of 0.70 corresponds to the null hypothesis that the slope equals −3. It is positive, reflecting the fact that −2.998 is greater than −3, and its level does not nearly approach its cutoff at the 5% probability level. In other words, we cannot reject the hypothesis that Direxion delivered on its promise to deliver a −3x product. The adjusted R^2 from the regression, at 0.998, indicates that the relationship is

Table 7. Tracking Errors and Daily Returns of ETPs Benchmarked to Total Returns of Security Indexes

		Tracking-Error Summary			Return Summary			
Fund	Expense Ratio	Tracking Difference	Standard Deviation	RATD Ratio	β	H_0: $\beta = L$ (*t*-ratio)	Adj. R^2	Holding-Period Return
A. S&P 500 (sample period 26 June 2009–13 March 2020; number of obs. = 2,694; benchmark HPR = 267.87%)								
SPXS, Direxion, −3x ETF	1.08%	0.002%	0.183%	0.564	−2.998	0.70	0.998	−99.59%
SPXU, ProShares, −3x ETF	0.91	0.007	0.216	1.615	−2.972	6.96	0.997	−99.53
SDS, ProShares, −2x ETF	0.89	0.004	0.109	2.070	−1.989	5.31	0.999	−96.33
SH, ProShares, −1x ETF	0.89	0.002	0.064	1.265	−0.997	2.69	0.998	−78.44
SSO, ProShares, 2x ETF	0.90	−0.008	0.116	−3.710	1.983	−7.78	0.998	726.41
SPXL, Direxion, 3x ETF	1.02	−0.014	0.187	−3.916	2.981	−5.29	0.998	1,389.01
UPRO, ProShares, 3x ETF	0.92	−0.013	0.212	−3.242	2.970	−7.41	0.998	1,435.59
B. Russell 2000 (sample period 6 November 2008–13 March 2020; number of obs. = 2,856; benchmark HPR = 175.09%)								
TZA, Direxion, −3x TF	1.11%	−0.009%	0.367%	−1.369	−2.918	18.84	0.994	−99.91%
TWN, ProShares, −2x ETF	0.95	−0.005	0.268	−0.937	−1.948	16.26	0.992	−98.19
RWM, ProShares, −1x ETF	0.95	−0.002	0.127	−0.904	−0.976	15.49	0.993	−81.72
UWM, ProShares, 2x ETF	0.95	−0.008	0.280	−1.614	1.940	−17.97	0.992	230.86
TNA, Direxion, 3x ETF	1.14	−0.017	0.414	−2.154	2.886	−24.21	0.992	118.89
C. 20-year T-bond (sample period 22 January 2010–13 March 2020; number of obs. = 2,550; benchmark HPR = 132.29%)								
TMV, Direxion, −3x ETF	1.04%	0.000%	0.889%	0.007	−2.845	7.77	0.889	−97.54%
TBT, ProShares, −2x ETF	0.90	−0.003	0.595	−0.208	−1.868	9.94	0.886	−90.24
TBF, ProShares, −1x ETF	0.92	−0.001	0.297	−0.170	−0.940	9.05	0.887	−65.44
UBT, ProShares, 2x ETF	0.95	−0.010	0.580	−0.850	1.918	−6.27	0.894	242.25
TMF, Direxion, 3x ETF	1.05	−0.014	0.892	−0.797	2.844	−7.82	0.888	382.75

Table 8. Analysis of Tracking Errors and Daily Returns of ETPs Benchmarked to Futures Indexes

		Tracking-Error Summary			Return Summary			
Fund	Expense Ratio	Tracking Difference	Standard Deviation	RATD Ratio	β	H_0: β = *L* (*t*-ratio)	Adj. R^2	Holding-Period Return
A. Crude oil (sample period 28 March 2017–13 March 2020; number of obs. = 745; benchmark HPR = –36.15%)								
DWT, Citigroup, –3x ETN	1.50%	–0.026%	1.674%	–0.431	–2.859	5.19	0.937	–60.37%
OILD, ProShares, –3x ETF	0.49	–0.040	1.676	–0.649	–2.830	6.28	0.936	–63.26
SCO, ProShares, –2x ETF	0.95	–0.022	1.110	–0.549	–1.878	6.80	0.937	–26.21
USO, USCF, 1x ETF	0.73	0.010	0.557	0.485	0.944	–6.24	0.936	–30.68
UCO, ProShares, 2x ETF	0.95	0.009	1.110	0.230	1.849	–8.60	0.937	–69.29
UWT, Citigroup, 3x ETN	1.50	–0.012	1.767	–0.181	2.732	–9.70	0.929	–92.81
OILU, ProShares, 3x ETF	0.49	–0.006	1.759	–0.091	2.719	–10.29	0.930	–92.20
B. Natural gas (sample period 8 February 2012–13 March 2020; number of obs. = 2,037; benchmark HPR = –83.51%)								
DGAZ, Credit Suisse, –3x ETN	1.65%	–0.043%	2.190%	–0.878	–2.702	17.01	0.921	–95.46%
KOLD, ProShares, –2x ETF	0.95	0.008	1.539	0.243	–1.727	23.38	0.915	15.90
UNG, USCF, 1x ETF	1.28	0.000	0.732	–0.004	0.918	–13.61	0.920	–82.54
BOIL, ProShares, 2x ETF	0.95	–0.047	1.545	–1.358	1.722	–23.81	0.914	–99.55
UGAZ, Credit Suisse, 3x ETN	1.65	–0.041	2.157	–0.847	2.692	–17.93	0.924	–99.99
C. Volatility (sample period 5 October 2011–2 February 2018; no. of obs. = 1,593; benchmark HPR = –99.74%)								
XIV, Credit Suisse, –1x ETN	1.65%	–0.016%	1.367%	–0.468	–0.868	17.90	0.897	2016.61%
SVXY, ProShares, –1x ETF	1.38	–0.019	1.363	–0.541	–0.873	17.15	0.898	1906.65
VXX, Barclays, 1x ETN	0.89	–0.017	1.311	–0.524	0.871	–18.36	0.906	–99.76
VIXY, ProShares, 1x ETF	0.87	–0.017	1.310	–0.524	0.875	–17.72	0.906	–99.76
TVIX, Credit Suisse, 2x ETN	1.65	–0.092	3.123	–1.175	1.618	–24.25	0.869	–100.00
UVXY, ProShares, 2x ETF	1.65	–0.066	2.643	–1.002	1.737	–18.60	0.905	–100.00

practically perfect. Finally, the last column in Panel A is the holding-period return for the period 26 June 2009 through 13 March 2020. All data series from product launch to the end date were collected from Bloomberg. Because we wanted to compare the various products linked to the same benchmark, the time series begins with the ETP with the latest launch date. Again, the results are consistent with the prediction that if the benchmark return is expected to be positive, funds with negative leverage ratios will fail. An investment of $100 in SPXS on 26 June 2009 is at the end of the period worth 41 cents.

From a performance perspective, Panel A of Table 7 indicates that the other S&P 500 ETFs performed as well as the –3x funds. The tracking differences are all ±1 bp a day. According to the RATD ratio, ProShares' 3x UPRO performed better than Direxion's 3x SPXL. On the basis of holding-period return, UPRO had a 1,436% return; SPXL's return was 1,389%. Highly positively levered S&P 500 funds can provide extraordinary holding-period returns in a stock market that is slowly rising upward, such as that experienced during the 11-year bull market that ended in March 2020.

The tracking differences provided in Panel B of Table 7 for all the Russell 2000 ETFs are less than zero and are slightly lower than the funds' expense ratios. The standard deviations are all higher than their S&P 500 ETF counterparts. Because the Russell 2000 comprises 2,000 midcap stocks, which have less deep and liquid markets than large-cap stocks, we would expect more tracking error. Noise can also be seen in the estimated regression relationship. The betas for the various products are lower in an absolute value sense than they were for the comparably levered S&P 500 products. Downwardly biased slope coefficients are expected when regression variables are noisy.

As shown in Panel C of Table 7, the results for the ETFs written on the 20+ year bond index exhibit yet even more noise. This finding is understandable. Relative to stock markets, bond markets are not deep, have high spreads, and are traded infrequently. The tracking differences indicate that all the funds performed reasonably well in relation to their stated objectives. The standard deviations are much higher than for the other carry benchmarks. The adjusted R^2 levels from the return regression are uniformly below 0.90.

The commodity and volatility results are reported in Table 8. Focusing on the crude oil results in Panel A first, we see that many of the funds, particularly the highly levered ones, had tracking differences well in excess of their respective expense ratios. OILD has a TD of 4 bps per day. The RATD ratio shows that OILD was the poorest performing of the seven funds considered. The realized betas for all the crude oil funds are less in absolute magnitude than what was promised, both in an economic and a statistical sense. The adjusted R^2 levels are not close to 1.0. Although the specific causes of the drag are not isolated, note that OILD is not close to replicating the performance of the benchmark. The holding-period results in the last column confirm our simulation predictions. In a noncarry market in which the expected benchmark return is negative, both levered short *and* levered long funds will eventually fail. The holding-period returns of the 3x UWT and OILU funds are both less than –92%. At the other end of the spectrum are the –3x DWT and OILD funds, both with returns less than –60%.

Our prediction of failure was, sadly, accurate. On 15 March 2020, ProShares announced the liquidation of its OILU and OILD ETFs. And on 19 March 2020, Citibank announced the acceleration of DWT and UWT; they are no longer with us.

The results for natural gas shown in Panel B of Table 8 are qualitatively similar to those for crude oil. Tracking differences are mostly negative and large. The standard deviations of the tracking errors are larger for the 3x natural gas products than they are for the crude oil products. The highly levered 3x long and –3x funds have dismal holding-period returns, –99.99% and –95.46%, respectively, as does 2x BOIL, at –99.55%. The –2x KOLD fund is the single fund with a positive holding-period return. It is also the single natural gas ETF with a positive RATD ratio.

The final panel in Table 8 contains the results for the volatility, VIX, ETPs. These results are amusing in a perverse sort of way. First, note that the 1x funds, VXX and VIXY, are *unlevered*, lost 2 bps a day, and had holding-period returns of nearly –100%. This behavior was first documented by Whaley (2013) many years ago in a warning about the effects of the contango trap in the VIX futures market. That same VIX futures price behavior has persisted in the ensuing years. As expected, the 2x products, TVIX and UVXY, performed worse than the unlevered products, with tracking errors exceeding 6 bps a day and holding-period returns of –100%. The most startling aspect of Panel C of Table 8 is the returns of the inverse products—XIV at 2,017% and SVXY at 1,906.7%. But note that the sample ends on 2 February 2018, the Friday before XIV collapsed. On Monday, 5 February 2018, the underlying futures index, SPVXSP, spiked by 96.7%. This rise triggered an acceleration event and liquidation. Although SVXY was not liquidated, ProShares changed the leverage ratio from –1x to –0.5x, reducing the levered benchmark volatility by 50%. Currently, the only way to get a 100% short exposure to SPVXSP is to sell the two nearby futures and rebalance the portfolio daily to maintain the constant one-month horizon.

Conclusion

Levered and inverse funds are controversial for good reason. Unlike typical securities traded on exchanges, their expected long-run values are zero. Is this fact hidden? No. Product issuers say so in their prospectuses. The economics, mechanics, and empirics of these product structures hold the key to understanding the reasons for the litany of recent fund failures. Given the expected return–risk characteristics of the underlying benchmark indexes, fund collapses were predictable from the get-go. We demonstrated this fact by using Monte Carlo simulation and return–risk parameters estimated in a period *before* the spike in volatility that began in March 2020. Compounding mechanics and dynamic rebalancing activity play roles, to be sure, as does an equally important (but often neglected) aspect of levered and inverse fund instability—negative expected returns for futures-based benchmark indexes. Most of the levered and inverse product controversies in recent times are for funds benchmarked to futures indexes linked to crude oil, natural gas, or volatility. These indexes have issues in their own right. In the futures markets, equilibrium expected benchmark returns are negative because the demands of long hedgers often exceed those of short hedgers. Speculators step in to absorb the hedging imbalance only when the futures price is high enough to earn a satisfactory risk premium. Levering these futures-based benchmarks merely exacerbates the performance problem and accelerates a fund's demise. Levered and inverse products are not, and cannot be, effective investment management tools.

Editor's Note

This article was externally reviewed using our double-blind peer-review process. When the article was accepted for publication, the authors thanked the reviewers in their acknowledgments. David Chambers and Fanesca Young were the reviewers for this article.

Submitted 22 April 2020

Accepted 23 September 2020 by Stephen J. Brown

Notes

1. See SEC Press Release 2019-242, "SEC Proposes to Modernize Regulation of the Use of Derivatives by Registered Funds and Business Development Companies" (25 November 2019).
2. Popular brokers, such as Fidelity, E*TRADE, Vanguard, TD Ameritrade, Charles Schwab, and Merrill, allow customers to open a trading account with no account minimum. In addition, no brokers charge commissions for online equity and ETP trades (see Frankel 2020; Beilfuss and Osipovich 2019).
3. The TCRS regulations may be found at https://treasury.tn.gov/Portals/0/Documents/Retirement/Policies/TCRS_Investment_Policy_Approved_10-16-2018.pdf.
4. The New York CRF regulations may be found at www.osc.state.ny.us/pension/generalpolicies.pdf.
5. The eventual collapse of levered products is not hidden by issuers. In its prospectus for VelocityShares Daily 2x VIX Short-Term ETN (TVIX), Credit Suisse explicitly states, "At higher ranges of volatility, there is a significant chance of a complete loss of the value of ETNs even if the performance of the applicable underlying index is flat" (p. 28). At some level, the logic here seems circular. On the one hand, Credit Suisse knows that the attraction of levered funds is their ability to create increased volatility. On the other hand, if the volatility of the benchmark index on which the levered return is based becomes too high, the fund will collapse.
6. We describe the fate of these funds later.
7. Most of the carry funds are replicated through OTC total return swap agreements. The OTC market maker, in turn, uses the futures market as a key hedging instrument in managing her swap exposure.
8. See "S&P GSCI Crude Oil: Methodology" (December 2019) at www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-gsci.pdf; the roll methodology is described at www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-gsci-crude-oil-annual-roll-supplement.pdf.
9. See "S&P VIX Futures Indices: Methodology" (May 2019) at www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-vix-futures-indices.pdf.
10. The 5% stopping criterion is probably unduly conservative. We look at the actual histories of the frequency of reverse splits for VIX ETPs at the end of this section.
11. This procedure is in the same spirit as the stopping criteria used in the simulations.
12. SPVXSP has a constant one month to expiration. The weights change each day as the futures contracts approach expiration. Nearby contracts are sold and replaced with second nearby contracts. On one day each month, the index will have a single contract (with 30 days to expiration as the futures contracts are rolled).
13. ETF.com estimates that 83% of ETPs traded in the United States trade within 1% of their NAVs and 94% trade within 2% (Roy 2017).
14. APs may also pay a variable fee of up to 0.10% of the value of the creation unit.
15. These ratios were unexpectedly changed to 1.5x and –0.50x on 26 February 2018.
16. ProShares reports the holdings of all of its ETFs on a daily basis. Go to www.proshares.com/resources/data_downloads.html.
17. www.proshares.com/resources/data_downloads.html.
18. This equation first appeared in O'Neill and Whaley (2020).

References

Baer, Justin, and Paul Kierman. 2020. “Leveraged-ETF Seller Urges Investors to Fight Proposed Regulation.” *Wall Street Journal* (31 January).

Beilfuss, Lisa, and Alexander Osipovich. 2019. “The Race to Zero Commissions.” *Wall Street Journal* (5 October). www.wsj.com/articles/the-race-to-zero-commissions-11570267802?mod=article_inline.

Cheng, Minder, and Ananth Madhavan. 2009. “The Dynamics of Levered and Inverse Exchange-Traded Funds.” *Journal of Investment Management* 7 (4): 43–62.

Dierking, David. 2020. “Leveraged Oil ETF UCO Undergoing a 1:25 Reverse Split.” TheStreet (21 April). www.thestreet.com/etffocus/trade-ideas/leveraged-oil-etf-uco-undergoing-a-1-25-reverse-split.

Frankel, Matt. 2020. “Best Online Stock Brokers for October 2020.” Ascent (1 October). www.fool.com/the-ascent/buying-stocks/.

Keynes, John M. 1930. *The Applied Theory of Money*. London: MacMillan.

Lintner, John. 1965. “The Valuation of Risk Assets and the Selection of Risky Investments and Capital Budgets.” *Review of Economics and Statistics* 47 (1): 13–57.

O’Neill, Michael J., and Robert E. Whaley. 2020. “Effects of Nondiscretionary Trading on Futures Prices.” Working paper, Financial Markets Research Center, Vanderbilt University.

Roy, Sumit. 2017. “ETFs with the Largest Premiums and Discounts.” ETF.com (22 March). www.etf.com/sections/features-and-news/etfs-largest-premiums-discounts.

Sharpe, Willian F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *Journal of Finance* 19 (3): 425–42.

Whaley, Robert E. 2013. “Volatility Trading: At What Cost?” *Journal of Portfolio Management* 40 (1): 95–108.