International Research Journal of Applied Finance ISSN 2229 – 6891
Vol. VII Issue – 6 June, 2016 www.irjaf.com
DOI: 10.0706/article-3

Analysis of Tracking Errors for Oil Sector Leveraged Exchange Traded Funds

Joseph Cheng* **Cameron Miller**

Abstract

The objective behind the use of leveraged and inverse ETFs is to obtain a multiple effect of the underlying index return for investors. In an efficient market, the leveraged ETF should move in multiples relative to its reference index. For example, a 3x leveraged ETF should move three times the amount of the underlier. Based on the autocorrelation regressions, we found that the market for leveraged ETFs is not perfectly efficient in that the return of the leveraged ETF is not always identical to the listed multiple times the return of the underlying index. The size of such tracking error is found to be correlated with the low return of the underlying index and low volume of the leveraged ETF. Furthermore, the regression shows negative serial correlation between past errors and current errors, which indicates that errors tend to reverse themselves over time. This suggests that while there is a tendency for the return for the leveraged ETF to converge to the listed multiple of the return for the underlying index, the convergence is not always instantaneous. Instead, there is normally a time lag of several days for the discrepancy between the leveraged ETF and the underlying index to gradually diminish.

Keywords: ETF, Leverage ETF, Tracking Error, Serial Correlation

I. Introduction

According to the article *Understanding Returns of Leveraged and Inverse Funds*, leveraged and inverse products have "attracted more than $33 billion of assets" and cover a "broad range of equity, sector, international, fixed-income, commodity and currency markets" (Hill and Foster (2009)). In other words, it is highly likely a leveraged or inverse product exists for a specific underlier. Such leveraged ETFs seek to replicate the performance before fees, for example, of the Energy Select Sector Index on a triple magnitude basis. Some investors may use these products to hedge out certain risk, or possibly for the benefit of adding leverage to a position.

In this paper, we use oil ETFs for our study. XLE is an ETF comprised of equities in the energy sector, with approximately 81% operating within the oil and industries sectors. As a percentage of the fund, the top holdings are Exxon Mobil, Chevron, and Schlumberger, making up 17.40%, 13.52%, and 7.74%, respectively.

XLE is referred to as the underlier here because it is an ETF used as a reference index or benchmark for several leveraged ETFs such as ERX and ERY. ERX is a 3x bull ETF whereas ERY is a 3x bear ETF of XLE. This means that ERX should move three times as much as XLE whereas ERY should move three times the inverse of XLE. The discrepancy between the actual return of the leveraged ETF and the return of the underlier times the multiple is the tracking error. If the triple leveraged fund rises more than 3x the underlying index, then the leveraged fund may be considered to have overshot its target of 3 times, which results in a positive error or positive tracking error. On the other hand, if the triple leverage fund rises less than 3x the underlying index, then the leveraged fund has undershot its listed multiple, which results in a negative tracking error.

II. Possible Reasons for Tracking Error

To begin analysing the tracking errors or discrepancies in returns, recognizing the differences in compounding returns is critical. Most investors believe they receive static leverage when purchasing these products, however, they in fact have exposure to dynamically rebalanced leverage. In other words, unless the return of the ETF is zero, the dynamically rebalanced fund will underperform or outperform the statically leveraged fund (Little (2010)). One notable example is the massive underperformance of leveraged ETFs, both long and short, during the credit crisis. Thus, this might be able to help explain why the coefficients on the underliers are not equal to the leveraged multiple.

One possible reason behind such discrepancy might be how compounding affects returns of leveraged and inverse ETFs for periods greater than one day. According to the journal *Understanding Returns of Leveraged and Inverse Funds*, the leveraged products are "designed to achieve a multiple of index return only on a daily basis. Over periods greater than one day, returns can be greater or less than the one-day target multiple times the index return" (Hill and Foster (2009)). For example, if an index returns 10% for two days in a row, the compounded return should be +21% ((1+10%) x (1+10%) = 1.21 or 21%). In relation to the underlier, a 2x bull leveraged ETF should return 20% per day for those two days, or the equivalent compounded return of 44% ((1+20%) x (1+20%) = 1.44 or 44%). You will notice that this compounded return is not double the underlier, but instead nearly 5% higher (44%/21% = 2.09).

During the 5-year observed time period from November 2010 to October 2015, the holding period returns of the reference index (XLE), 3x bull (ERX), and 3x bear (ERY) were -0.50%, -9.63%, and -89.42%, respectively. Based on such spread in returns, we can firmly state that while a relationship does exist between leveraged and inverse products and their respective underliers for any given day or week, the closeness of that relationship is diminished over an extended period of time.

There are several additional problems besides the compounding return calculations that plague leveraged ETFs. For example, it is important for any leveraged ETF that the leverage ratio be kept constant, thus the fund must add to its winning positions during a bull market, and sell its losing positions during a bear market. To illustrate this point, if an investor has an account worth $1,000 and chooses to invest in a 2x leveraged ETF, they are effectively buying $2,000 worth of the index. If that index rises by 10% the following day, the investor's account is worth $1,200, where the assets of the index are worth $2,200. Through the simple calculation of dividing the worth of the assets of the index over the value of the investor's account, the leverage ratio can be calculated, in this case result in a factor of 1.83. The investor must now purchase more of the index to keep the leverage ratio constant. If the underlier fails to continue moving in the favourable direction, the investor can quickly lose large sums of money. In addition, according to *Understanding the Tracking Errors of Commodity Leveraged ETFs,* "should the leveraged ETFs exhibit high volatility but no significant movement in price over a period of time, the constant daily re-balancing would cause the fund to decline in value. Therefore, leveraged ETFs can be viewed as long momentum but short volatility, and the value erosion due to realized variance of the reference is called volatility decay" (Guo and Leung (2014)).

III. The Size of Tracking Errors

To illustrate how the discrepancy or tracking error vary with the returns of the index, we have two graphs showing the relative returns of the leverage ETFs and the underlyer with respect to the return of the underlyer. Note the graph (Figure 1) illustrates the relative return ratio, which is the daily return of the leveraged fund (ERX) divided by the daily return of the underlier (XLE). In this case, the leverage fund is the 3x bull so the relative return ratio is expected to be about three. As can be seen, most of the days the relative returns do fall along the horizontal line correspoinding to a multiple of 3. The vertical distance between the actual point and that horizontal line represents the deviation from the listed multiple; this deviation is the tracking error. Interestingly, while the average ratio is close to three, there are times when the relative return ratio deviate substantially above or below 3. Such occurrences are most common along the vertical line where return of XLE is zero. The return or changes in XLE, which is the denominator of the relative return ratio, can cause the relative return ratio to be abnormally large. This is because when the denominator, which is the return for XLE, is very small, even a small discrepancy between the leverage fund return and the listed multiple can lead to abnormally high deviation from the norm, as observed in many instances in the graph. Nevertheless, it can be seen that the leverage ETF, in general, acts as a multiplier as described. Thus, the product can be said to generally deliver a level of risk that is consistent to what has been stated.

Note the graph (Figure 2) illustrates the relative return ratio, which is the daily return of the inverse leveraged fund (ERY) divided by the daily return of the underlier (XLE). In this case, the leverage fund is the 3x bear so the relative return ratio is expected to be about -3. Interestingly, while the average ratio is close to -3, there are times when the relative return ratio deviates substantially above or below -3, especially when the return of XLE is close to zero, yielding a small denominator. Also, notice that the points for the inverse leveraged fund in Figure 2 is a bit more scattered than that in Figure 1. This is likely due to the fact that the inverse leveraged fund has a much lower volume, resulting in lower market efficiency and higher tracking errors. In summary, the size of the tracking error is correlated with low volume and low return for the underlying index.

IV. Regression Results

In this study, daily and weekly returns were generated from data for XLE, ERX, and ERY and gathered from DataStream from November 2010 to October 2015 to evaluate the relationships between the leveraged products and their respective underliers. For example, since the return for a 3x Leverage ETF is expected to generate three times the return of the underlying ETF against which the leveraged ETF is benchmarked, the coefficient of the return for the underlying ETF should be about 3 when the leveraged ETF is regressed against the underlying ETF. In order to estimate both the concurrent correlation between these products and possible serial correlation of tracking errors, a tenth-order serial correlation regression was applied to the returns of ERX and returns of XLE, which allows us to go back ten days for a thorough detection of a possible lagged correlation. E-Views was used for these regressions, as shown in the following tables:

Table 1 shows the regression result of Return of ERX against the Return of XLE (the underlying index or benchmark) on a daily basis. The coefficient is 2.93, which is, as expected, very close to 3. What is interesting however, is that the coefficient for the one-day-lag error, two-day-lag error, and the three-day-lag error are all negative and statistically significant at the 95% level. The negative coefficients for these lagged errors suggest that there is a tendency for discrepancies between the actual return of the leveraged fund and the listed multiple of the underlying index to be reversed by errors with opposite signs over the course of several days. This is referred to as the compensating effect where the discrepancy is neutralized gradually over time. Also, note that the t-statistic for the lagged errors tend to decline over time, indicating that the tendency for reversal is the strongest in the first day, and then tends to become weaker on the second day and so forth. This suggests that while the market is not perfectly efficient, the market plays catch-up quite quickly.

Table 2 shows the regression result of Return of ERX against the Return of XLE (the underlying index or benchmark) on a weekly basis. The coefficient is also 2.93, which is very close to the result we got in Table 1 where daily return was used. The difference here is that the coefficient for the one-week-lag error is no longer statistically significant. This is not surprising since the compensating effect takes only a few days, as shown in Table 1. Thus, the negative coefficient for the one-week-lag error should no longer be statistically significant.

Table 3 shows the regression result of Return of ERY against the Return of XLE (the underlying index or benchmark) on a daily basis. Since ERY is 3x bear, the coefficient should be -3. The actual estimated coefficient turned out to be -2.94, which is, as expected, very close to -3. What is interesting however, is that the coefficients for the one-day-lag error, two-day-lag error, and all the way back to the five-day-lag error are all negative and statistical significant at the 95% level. This suggests that there is a tendency for the discrepancy between the actual return of the leveraged fund and the listed multiple of the underlying index to be reversed over the course of several days. Note that the t-statistics for the coefficients for the lagged errors decline over time. Such phenomenon is similar to what we found for ERX. The compensating effect for ERY seems to last longer (5 days instead of 3 days) than for ERX. That is, it takes longer for the discrepancy to vanish. Perhaps this is due to the possibility that ERY is a bear ETF that has lower volume than the bull ETF. The lower volume might contribute to more market inefficiency and thus a longer time is needed for the compensating effect to run its course.

Table 4 shows the regression results of Return of ERY against the Return of XLE (the underlying index or benchmark) on a weekly basis. Since ERY is 3x bear, the coefficient should be -3. The actual estimated coefficient turned out to be -2.97, which is similar to the result obtained using daily return as shown in Table 3. The negative coefficients for the one-week-lag error and the two-week-lag error are statistically significant, which suggests that the compensating effect lasts as long as two weeks. This might indicate that market efficiency might be lagging for such lower volume products.

V. Conclusion

Based on our analysis, it appears that the leveraged funds generally have small tracking errors. Larger errors have been observed in ETFs with low volume. Although the errors are generally small, there seems to be a tendency for these errors to correct themselves over time. Based on the autocorrelation regressions, we noted that the coefficients for the lagged errors are negative, which suggests that most errors or discrepancies between the actual return of the leveraged fund and the listed multiple of the index tend to be reversed by an error of the opposite sign in the following days. This phenomenon that we have observed is referred to as the compensating effect and seems to last only two days or so. On a weekly return basis, the compensating effect is not readily observable because the compensating effect takes place rather quickly, usually in a matter of days rather than a matter of weeks. Thus, the compensating effect is not readily detected when weekly return data is used for serial correlation regressions. However, the compensating effect is quite apparent when daily return data are used. In conclusion, while the leveraged ETF might not be perfectly efficient in terms of tracking its corresponding reference index in any given day, they are generally efficient over time.

References

Carver, A. B. (2009, October). Do Leveraged and Inverse ETFs Converge to Zero? *ETFs and Indexing, 2009*, 144-149. doi:10.3905/ETF.2009.2009.1.144

Guo, K., & Leung, T. (2014, October). Understanding the Tracking Errors of Commodity Leveraged ETFs. SSRN Electronic Journal SSRN Journal. doi:10.2139/ssrn.2389411

Hill, J., & Foster, G. (2009, November). Understanding Returns of Leveraged and Inverse Funds. Journal of Indexes.

Little, P. K. (2010). Inverse and Leveraged ETFs: Not Your Father's ETF. Index Investing The Journal of Index Investing, 1(1), 83-89. doi:10.3905/jii.2010.1.1.083

Authors

Dr. Joseph Cheng
Department of Finance and International Business
Ithaca College, Ithaca, NY
cheng@ithaca.edu

Cameron Miller
Ithaca College, Ithaca, NY
cmiller5@ithaca.edu

Figure 1

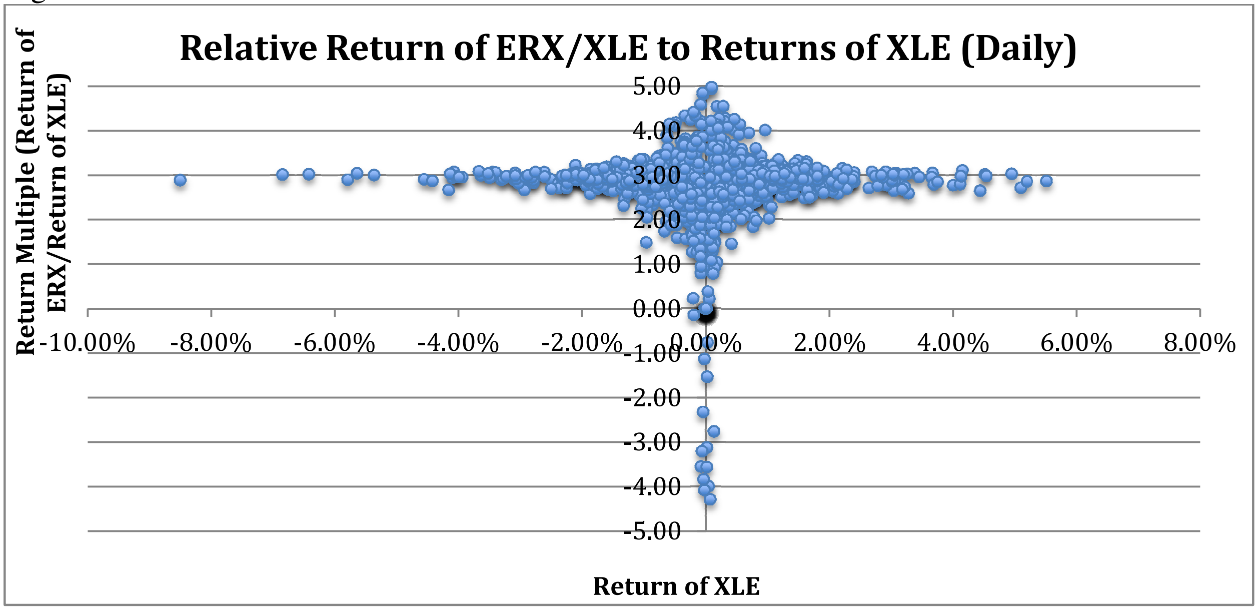

Figure 2

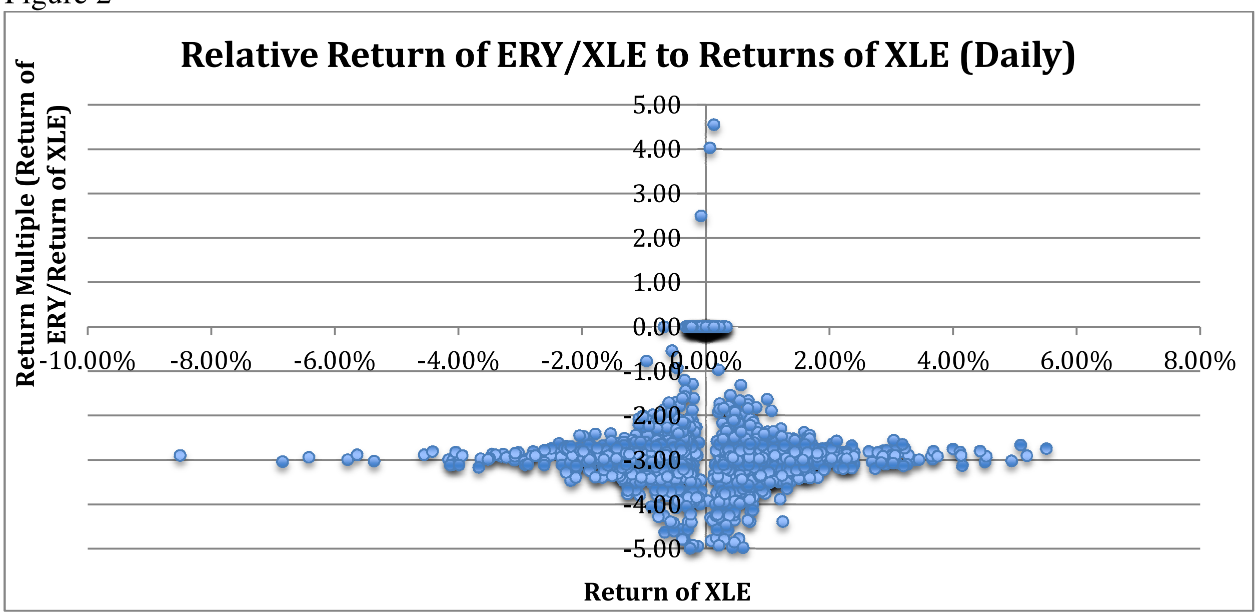

Table I. Autocorrelation Regression Results*: ERX Returns vs. XLE Returns Daily*
Dependent Variable: ERX_RETURNS
Method: Least Squares
Date: 04/21/16 Time: 13:59
Sample (adjusted): 11/08/2010 10/23/2015
Included observations: 1295 after adjustments
Convergence achieved after 6 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.0001	0.0000	-1.7144	0.0867
XLE_RETURNS	2.9320	0.0048	607.5847	0.0000
AR(1)	-0.2188	0.0280	-7.8274	0.0000
AR(2)	-0.1147	0.0286	-4.0121	0.0001
AR(3)	-0.0659	0.0287	-2.2938	0.0220
AR(4)	-0.0265	0.0287	-0.9216	0.3569
AR(5)	0.0520	0.0286	1.8198	0.0690
AR(6)	-0.0378	0.0287	-1.3201	0.1870
AR(7)	-0.0115	0.0286	-0.4006	0.6888
AR(8)	0.0446	0.0285	1.5636	0.1182
AR(9)	-0.0060	0.0284	-0.2123	0.8319
AR(10)	-0.0072	0.0277	-0.2600	0.7949

R-squared	0.9963	Mean dependent var	0.0006
Adjusted R-squared	0.9962	S.D. dependent var	0.0397
S.E. of regression	0.0024	Akaike info criterion	-9.1826
Sum squared resid	0.0077	Schwarz criterion	-9.1348
Log likelihood	5957.7650	Hannan-Quinn criter.	-9.1647
F-statistic	31037.0400	Durbin-Watson stat	1.9980
Prob(F-statistic)	0.0000		

Table II. Autocorrelation Regression Results*: ERX Returns vs. XLE Returns Weekly*

Dependent Variable: ERX_RETURN

Method: Least Squares

Date: 04/21/16 Time: 14:10

Sample (adjusted): 1/07/2011 10/23/2015

Included observations: 251 after adjustments

Convergence achieved after 6 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.0005	0.0002	-2.2241	0.0271
XLE_RETURN	2.9367	0.0118	248.8501	0.0000
AR(1)	-0.1166	0.0639	-1.8256	0.0692
AR(2)	-0.0333	0.0642	-0.5183	0.6048
AR(3)	0.0257	0.0637	0.4031	0.6873
AR(4)	-0.0123	0.0641	-0.1920	0.8479
AR(5)	-0.0737	0.0567	-1.2986	0.1953
AR(6)	-0.4574	0.0570	-8.0277	0.0000
AR(7)	-0.1262	0.0637	-1.9809	0.0488
AR(8)	0.1274	0.0642	1.9832	0.0485
AR(9)	-0.0719	0.0644	-1.1153	0.2658
AR(10)	0.1299	0.0663	1.9601	0.0511

R-squared	0.9954	Mean dependent var	0.0019
Adjusted R-squared	0.9951	S.D. dependent var	0.0878
S.E. of regression	0.0061	Akaike info criterion	-7.3076
Sum squared resid	0.0090	Schwarz criterion	-7.1390
Log likelihood	929.0988	Hannan-Quinn criter.	-7.2397
F-statistic	4654.0590	Durbin-Watson stat	1.9925
Prob(F-statistic)	0.0000		

Table III. Autocorrelation Regression Results*: ERY Returns vs. XLE Returns Daily*
Dependent Variable: ERY_RETURNS
Method: Least Squares
Date: 04/21/16 Time: 14:08
Sample (adjusted): 11/08/2010 10/23/2015
Included observations: 1295 after adjustments
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.0001	0.0000	-2.3792	0.0175
XLE_RETURNS	-2.9484	0.0066	-449.9885	0.0000
AR(1)	-0.5044	0.0279	-18.0524	0.0000
AR(2)	-0.3146	0.0313	-10.0590	0.0000
AR(3)	-0.1510	0.0325	-4.6482	0.0000
AR(4)	-0.0616	0.0328	-1.8786	0.0605
AR(5)	-0.0708	0.0327	-2.1621	0.0308
AR(6)	-0.0312	0.0327	-0.9540	0.3403
AR(7)	-0.0070	0.0327	-0.2129	0.8314
AR(8)	-0.0144	0.0324	-0.4439	0.6572
AR(9)	-0.0284	0.0313	-0.9076	0.3643
AR(10)	0.0176	0.0279	0.6319	0.5275

R-squared	0.9912	Mean dependent var	-0.0008
Adjusted R-squared	0.9911	S.D. dependent var	0.0401
S.E. of regression	0.0038	Akaike info criterion	-8.3106
Sum squared resid	0.0183	Schwarz criterion	-8.2627
Log likelihood	5393.1100	Hannan-Quinn criter.	-8.2926
F-statistic	13119.0200	Durbin-Watson stat	1.9966
Prob(F-statistic)	0.0000		

International Research Journal of Applied Finance ISSN 2229 – 6891
Vol. VII Issue – 6 June, 2016 www.irjaf.com
DOI: 10.0706/article-3

Table IV. Autocorrelation Regression Results*: ERY Returns vs. XLE Returns Weekly*

Dependent Variable: ERY_RETURN
Method: Least Squares
Date: 04/21/16 Time: 14:11
Sample (adjusted): 1/07/2011 10/23/2015
Included observations: 251 after adjustments
Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.0006	0.0003	-2.0304	0.0434
XLE_RETURN	-2.9734	0.0208	-143.1382	0.0000
AR(1)	-0.2061	0.0651	-3.1649	0.0018
AR(2)	-0.1665	0.0662	-2.5140	0.0126
AR(3)	-0.0777	0.0669	-1.1616	0.2466
AR(4)	-0.0577	0.0669	-0.8634	0.3888
AR(5)	-0.0456	0.0644	-0.7087	0.4792
AR(6)	-0.2994	0.0645	-4.6397	0.0000
AR(7)	-0.0394	0.0677	-0.5823	0.5609
AR(8)	-0.0680	0.0675	-1.0078	0.3146
AR(9)	-0.0909	0.0680	-1.3365	0.1826
AR(10)	-0.0083	0.0677	-0.1222	0.9029

R-squared	0.9874	Mean dependent var	-0.0031
Adjusted R-squared	0.9868	S.D. dependent var	0.0901
S.E. of regression	0.0103	Akaike info criterion	-6.2589
Sum squared resid	0.0256	Schwarz criterion	-6.0903
Log likelihood	797.4860	Hannan-Quinn criter.	-6.1910
F-statistic	1705.0790	Durbin-Watson stat	1.9905
Prob(F-statistic)	0.0000		