The Performance of Exchange Traded Funds and Mutual Funds

Joseph Farinella
University of North Carolina Wilmington

Raphael Kubicki
University of North Carolina Wilmington

In this study, we examine the performance of 61 exchange-traded funds (ETFs) and 61 mutual funds from January 1st, 2005 to December 31st, 2016. The sample is divided into large-cap equity, mid-cap equity, small-cap equity and fixed-income funds. We find that ETFs have a lower tracking error relative to their bench mark compared to mutual funds. We find the average bid-ask spread on ETFs is 1.73%. ETFs that invest in mid-cap stocks and small-cap stocks have higher bid-ask spreads. We find that ETFs have a statistically higher annul return and lower annual fees.

INTRODUCTION

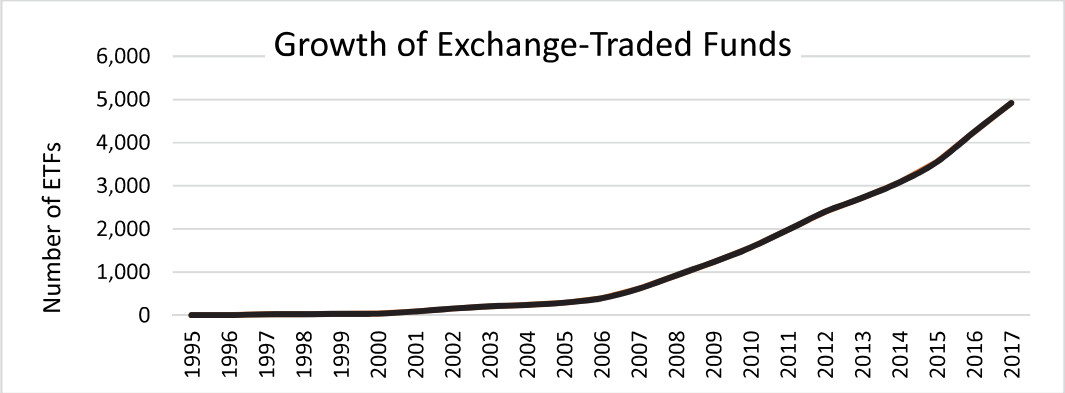
ETFs are relatively new financial vehicles and have experienced more attention over the last two decades, especially in the United States. In January 1993, the American Stock Exchange introduced the Standard & Poor's Depositary Receipts (SPDRs). Although in the beginning these products experienced less attention and slow growth mainly due to their complex and unfamiliar structure, in the mid-1990s, investors realized the benefits of indexing and passive investing. Currently, the SPDR is the largest ETF and is the second most actively traded fund behind the Vanguard 500 Mutual Fund. Sector SPDR allowed investors to focus their placement of funds on one particular segment of stocks within the S&P 500.

The flexibility of being able to short or buy ETFs on margin and the relatively low expense ratio are primary reasons assets invested in ETFs has grown. Because of the advantages that ETFs have compared to mutual funds, the demand for ETFs has increased over the last decades. At the end of 2016, exchange-traded products (ETPs) reached a new high of $3.5 trillion in total assets worldwide (Mercado, Lan, & Rajendra, 2017). This milestone underlines the tremendous growth that this relatively new investment instrument has experienced. The United States, which accounts for roughly $2.5 trillion, is by far the most active country in this segment. The total assets invested in ETFs and the number of new ETFs being created have grown.

While the first ETFs focused on the replication of the equity indices (S&P 500, Dow Jones, Russel 2000), the trend shifted in the following years more towards the replication of certain sectors (financial, healthcare, technology) or other niches. Additionally, indices from other asset classes like fixed income, commodities, or real estate, as well as minor indices in the mid-cap and small-cap sectors were incorporated. This broad supply further increased the investment opportunities and followed the demand of ETFs by meeting more specific investment goals. In 2016, there were 632 new ETFs created. This

brings the total number of registered ETFs up to 4,779 as of the end of 2016. Figure 1 provides a summary of the growth of exchange traded funds from 1995 – 2017.

FIGURE 1
GROWTH OF EXCHANGE-TRADED FUNDS



Source: Bloomberg L.P. (2017)

The goal of this research is to examine differences in performance and fees between ETFs and mutual funds that have the same or similar indices as their benchmark. While most of the prior research that has been conducted in this field either examines a shorter period of time or a small sample of funds, this study evaluates and compares 61 ETFs with 61 mutual funds and a total of 70 benchmark indices over a period of 12 years. The reason for the higher number of indices is that not all ETFs and mutual funds could be paired with the same benchmark index.

LITERATURE REVIEW

Compared to mutual funds, which can only be bought from and sold to the fund company, ETFs trade like stocks, and can be purchased and resold in the secondary market. Another similarity to stocks is that ETFs trade during the trading day and not just at the end of the day like mutual funds. This allows investors to participate in intraday price movements. (Gastineau, 2001).

In order to offer new ETF shares on an exchange, the creation process has to be undertaken. The creation of ETFs can only be executed through an authorized participant (AP). Authorized participants are usually large dealers or brokerage firms. AP deposits an entire basket of securities in the exact weight as the underlying index. Thus the ETF should always be able to closely mirror its benchmark. In exchange for the deposited basket of securities, the AP receives shares of the ETF from the ETF manager for an equal value to the deposit. This process usually takes place in so-called "creation units." Each of these units contains a certain number of shares which is usually in the order of 25,000, 50,000 or 100,000. After receiving the shares, the AP is able to offer them on the exchange to make them available to investors, which automatically gives the AP a central market-to-market function (Hill, Nadig, & Hougan, 2015).

The return of shares is called redemption and works in the exact reverse way. The AP can redeem the received shares with the ETF provider in exchange for the securities (Foucher & Gray, 2014). The securities can be sold afterwards in the open market at the current market price. Since there are many APs in each ETF, several creation and redemption baskets are produced every day. This back-and-forth ensures that the net asset value (NAV) of the securities is always within a very close range to the price of the ETF.

To keep the entire process transparent, each ETF has to post its creation and redemption baskets with the National Securities Clearing Corporation (NSCC) for publication. Even though the creation and

redemption baskets are identical most of the time, sometimes minor deviations between the holdings of the ETF and those in the underlying index can occur, (Kosev & Williams 2011).

The pricing of ETFs follows the rule of supply and demand and is identical to stocks. This mechanism can interfere with the ETFs' "fair" value, which is based on the net asset value of the securities in the underlying index. Nevertheless, in cases where the demand is greater than the current supply, the price of ETFs can increase above the fair value. These under- or overpricing situations are called discounts and premiums. As a consequence of such discounts or premiums, an arbitrage opportunity arises which can be implemented. By buying the underlying securities, exchanging them for new ETF shares (creation process), and selling the new shares to the market, the AP is able to make a riskless profit. The increase in demand of the underlying securities and the increase in supply of the ETF bring the prices back to unity.

However, only when the underlying securities are liquid can APs execute their arbitrage strategy immediately. For this reason, ETFs that track smaller and less liquid indices usually have higher premiums and discounts. According to Vanguard Financial Advisor Services (2015), premiums and discounts are defined as the positive or negative difference between the ETF's market price and its NAV. The main reason that causes such differences can be explained by two factors. First, there is the variation in trading hours of the exchange where the ETF is traded and the one where the underlying securities are traded. Second, a significant decrease in the liquidity of the underlying securities can generate pricing differences. Consequently, it can be stated that ETFs with a frequently traded domestic underlying index (e.g. U.S., large-cap) should have smaller premiums and discounts than those who trade infrequently and trade internationally.

Evidence that ETFs have lower returns than their underlying indices can be found in research by Svetina and Wahal (2008), Shin and Soydemir (2009), Rompotis (2005), Harper et al. (2004), and Gastineau (2004). These papers show that ETFs slightly underperform their benchmarks; primarily due to expenses. The relative performance of mutual funds compared to their benchmarks is less clear. Poterba and Shoven (2002) find an underperformance of index funds while Gastineau (2004) states that index mutual funds outperform their benchmarks.

Generally, fees that are connected to ETFs and mutual funds can be distinguished between nonrecurring fees and annual expenses. The most common nonrecurring fees are front-end and back-end loads, as well as early withdrawal fees. According to the U.S. Securities and Exchange Commission, sales loads are charges that an investor has to pay to the broker during the process of the purchase (front-end load) or redemption (back-end load). Early withdrawal charges occur when the investor sells shares before the required holding period ends. These nonrecurring fees often apply to open-end funds and can appear in all kinds of combinations. Except for some actively managed variations, ETFs, usually do not have nonrecurring charges.

Mutual funds and ETFs charge annual expenses which is measured by the expense ratio. This ratio includes all annual fees such as management fees, 12b-1 distribution fees (only applicable for mutual funds), and other expenses. This ratio can vary between 0.1% and almost 2% and is charged on an annual basis.

ETFs also have a hidden expense which is the bid-ask spread. Investors must purchase the ETF at the (higher) ask price while they sell at the (lower) bid price. The difference between these two prices is the bid-ask spread and is an implicit cost that the market maker receives for their services.

Kostovetsky (2003) examines the cost differences between index mutual funds and exchange-traded funds from the perspective of private investors. The author notes three important non-tracking-error differences between mutual funds and ETFs, which are management fees, shareholder transaction costs, and taxation costs. By using a one-period model (one year), he measures the cost differences to determine whether investors earn higher returns in mutual funds or ETFs. The results show that investors who invest more than $60,000 find ETFs more cost-efficient. Using a longer time investment horizon, lowers the threshold needed to make the ETFs more cost efficient.

If a fund manager is not able to match the returns of the benchmark, a tracking error occurs. Thus the tracking error is defined as the difference between the return of the investment instrument (ETF or mutual

fund) and its benchmark. The tracking error can partially explain the relative performance, since the difference over time between the fund and the benchmark is shown. However, the tracking error only measures the volatility and does not express whether the return differential is negative or positive (Brandon, Rottschafer, & Zvingelis, 2013). The measure of volatility also measures the risk of the investment by showing how much it deviates from its benchmark.

In this study we apply three established methods to measure the tracking error. Prior studies are mixed depending on the period and funds examined. Shin and Soydemir, (2009), Frino & Gallagher, (2001) and Svetina & Wahal (2008) find that ETFs and index mutual funds have small but significant tracking error. Harper et al., (2004) finds almost no tracking error. Chang et al. (2015) finds that closed end funds have a higher tracking error compared to ETFs. Rompotis (2005) finds similar tracking error differences between ETFs and index mutual funds.

The last part of this paper compares the cost structure and its impact on the overall performance of ETFs and mutual funds. Each of these types of funds can be related to different kinds of expenses that the investors have to consider. For the ETFs, the bid-ask spread, premiums or discounts, and expense ratios are considered. Mutual funds have an expense ratio, and often front-end and back-end loads.

Hilliard (2014) finds an average premium of almost zero for domestic-equity ETFs. Chang et al. (2015) focus on the expense ratio. Their findings in the fixed-income segment are that ETFs, compared to closed end funds, have a lower average expense ratio of 0.42%. Rompotis (2005) confirms these findings in his comparison of expense ratios between equity ETFs and index mutual funds. A priori, we expect ETFs to have a lower expense ratio compared to mutual funds.

Most of the research that has been done in this field has investigated either a short observation period, a relatively small group of funds, or only certain segments of funds. Poterba and Shoven (2002) compare one ETF with one index fund over the period of six years, and Gastineau (2004) examine one ETF and two closed end funds over five years. Rompotis (2005) evaluates 16 pairs of ETFs and index mutual funds over one and a half years. Harper et al. (2004) studies 22 pairs of ETFs and closed end funds over five years using monthly data. Shin and Soydemir (2009) examine 26 ETFs over three years. The most comprehensive is by Svetina and Wahal (2008) who examine 584 ETFs for the period from each fund's inception to 2007.

DATA

The data that is used in this study consists of ETFs, mutual funds, and their underlying benchmark indices. We examine a period of 12 years from January 1st, 2005 - December 31st, 2016. We select funds that have the highest total assets over this period. The sample consists of 61 ETFs and 61 mutual funds that have comparable benchmarks. Further, the sample is grouped into equity and fixed-income, and the equity funds is split into large-cap, mid-cap, and small-cap. We also collect data for 70 underlying or benchmark indices that Bloomberg indicates are the performance benchmarks for a given ETF and mutual fund. The higher number of indices is due to the fact there are a total of 122 investments being examined (61 ETFs and 61 mutual funds) and each has a benchmark for their fund. The summary of the selected funds and indices, as well as the number of observations and the total assets, are listed in Table 1.

For all 122 funds and 70 indices, we use daily data thus most funds have a total of 3,020 observations and the entire sample is approximately 537,800 observations. The data is primarily collected from the Bloomberg terminal and if necessary supplemented with data from finance.yahoo.com. Data for the indices is retrieved from Wharton Research Data Service (WRDS). Since ETFs are traded with a spread, additionally, all bid and ask prices for the observation period are obtained. Moreover, the daily closing prices for all funds and indices are collected. We collect the daily net asset value for all ETFs in order to calculate the premiums and discounts. The expense ratios are obtained for all funds as well as other expenses such as front or back loads, early withdrawal charges, and other fees that are not included in the expense ratio.

TABLE 1
SELECTION OF FUNDS AND INDICES

Category	Number Funds/Indices	Number of Observations	Total Assets (M)
ETF-Large	20	59,129	$647,620
ETF-Mid	16	37,595	$82,834
ETF-Small	10	26,036	$2,467
ETF-Fixed-Inc.	15	40,196	$262,518
MF-Large	20	59,396	$1,284,352
MF-Mid	16	46,922	$49,954
MF-Small	10	30,179	$10,882
MF-Fixed-Inc.	15	39,741	$136,243
Index-Large	22	59,193	
Index-Mid	15	44,602	
Index-Small	13	37,075	
Index-Fixed-Inc.	20	57,774	
Sum	**192**	**537,838**	**$2,476,870**

METHODOLOGY

First, we analyze the performance of ETFs, mutual funds and their benchmark indices. Therefore, the daily percentage returns are calculated. We calculate the daily returns for the ETFs, mutual funds and indices using equation (1):

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100\% \quad (1)$$

where R_t is the percentage return on day t and P_t is the closing price of the fund or index on day t. Since ETFs can be traded during the day and have a spread, we want to apply a second approach. This time the mid-point between the bid and ask price used as shown in equation (2):

$$R_t = \frac{\frac{B_t + A_t}{2} - \frac{B_{t-1} + A_{t-1}}{2}}{\frac{B_{t-1} + A_{t-1}}{2}} \times 100\% \quad (2)$$

where R_t is the percentage return on day t, B_t is the bid price and A_t the ask price of the ETF. With the second approach we try to find confirmation for the results for the ETF returns that were calculated with the last price in equation (1). However, Aber et al. (2009) highlight in their study that most of the trading during the day happens around the mid-price. Nevertheless, for further calculations regarding the ETFs the returns resulting from equation (1) are used, based on the assumption that the last price is equivalent to the mid-price over a period of 12 years.

The tracking errors measure the fund's ability to track or mirror its benchmark. In the previous literature, different approaches for the calculation of tracking errors are used. Aber et al. (2009) and Svetina and Wahal (2008) use the standard deviation of differences between the fund and the benchmark index. Harper et al. (2004) use the difference of returns between the ETF and the index for their study.

Gallagher and Segara (2006) use both of the approaches. The most common way of measuring the tracking error, which is established by Frino and Gallagher (2001) and later adapted by Rompotis (2005) and Shin and Soydemir (2009), is to use all three approaches. In our study, we use the same set of three different methods, which are the average of absolute return differences, the standard deviation of return differences, and the standard error from the regression analysis.

First, the tracking error is estimated with the average of absolute return differences between the ETF and its underlying index. For this purpose, we use the following equation:

$$TE_1 = \frac{\sum_{t=1}^{n} |FR_t - IR_t|}{n} \tag{3}$$

where TE_1 stands for the first tracking error, FR_t is the fund's returns in day t, while IR_t represents the index return in day t. Both returns are based on the results from equation (1).

The second method calculates the tracking error by using the standard deviation of differences between the returns. This model is:

$$TE_2 = \sqrt{\frac{\sum_{t=1}^{n} (D_t - \overline{D})^2}{n-1}} \tag{4}$$

Again, ETF as well as index returns are calculated with equation (1). D_t is the difference of ETF returns and the returns of the underlying index in day t and $\overline{D}$ the difference between average returns over the period of n days.

The last method to estimate the tracking error also uses the daily returns from equation (1). The standard error from the regression analysis can be found with this model:

$$FR_t = \alpha_t + \beta_t \cdot IR_t + \varepsilon_t \tag{5}$$

where FR_t is the return on the fund in day t and IR_t represents the return of the tracking index portfolio at the same day t. The alpha coefficient α_t expresses the excess return. This shows the fund's ability to outperform its benchmark. Since the ETFs used in this study are only replicating, the alpha is expected to be negative or zero.

Next we investigate expenses that are associated with ETFs and mutual funds. We measure the bid-ask spread on ETFs by applying two different models. The first method is the most common one and can be calculated with the following equation:

$$S_1 = \frac{PA_t - PB_t}{PA_t} \tag{6}$$

where PA_t is the ask price at day t and PB_t the bid price at day t. To confirm these results we apply a second method. Roll (1984) provided this alternative that has also been used by Rompotis (2005). The equation is:

$$S_2 = \frac{s_t}{\sqrt{PA_t \cdot PB_t}} \tag{7}$$

where s_t is the difference between ask and bid prices at day t. Next we take a closer look at the discount or premium, which can result in a benefit or cost for the investor. Even though it is assumed that every price that differs from the equilibrium price is causing an arbitrage situation and should be shortly offset, prior research has shown that this statement is not always true.

The most established method to calculate the ETF's premium or discount is by using the following equation:

$$PD_1 = \frac{P_t - NAV_t}{NAV_t} \times 100\% \quad (8)$$

where P_t is the closing price at day t and NAV_t is the net asset value at day t. The results are displayed in percentages. Positive results reflect a premium while negative numbers show that the fund was trading at a discount on this day.

EMPIRICAL RESULTS

Table 2 presents the average daily returns of ETFs, mutual funds, and their underlying indices. The daily return is highest for mid-cap followed by small-cap, large cap and fixed income. The small cap has the highest standard deviation followed by mid-cap, large-cap and fixed income.

TABLE 2
SUMMARY OF DAILY RETURM

Category	Mean	Std. Dev.	Median	Min	Max	N
ETF-Large	0.027%	1.275%	0.064%	-9.712%	12.756%	54,810
ETF-Mid	0.043%	1.433%	0.075%	-9.734%	10.143%	31,300
ETF-Small	0.035%	1.976%	0.064%	-13.939%	15.487%	21,435
ETF-Fixed	0.003%	0.489%	0.016%	-4.567%	4.712%	37,417
ETF – Avg.	**0.027%**	**1.293%**	**0.055%**	**-9.488%**	**10.775%**	
MF-Large	0.027%	1.243%	0.071%	-9.257%	11.595%	59,396
MF-Mid	0.023%	1.474%	0.096%	-13.986%	10.181%	46,922
MF-Small	0.012%	1.445%	0.088%	-15.847%	9.071%	30,179
MF-Fixed	0.002%	0.290%	0.008%	-3.330%	2.588%	39,741
MF – Avg.	**0.016%**	**1.113%**	**0.066%**	**-10.605%**	**8.359%**	
Index Large	0.042%	1.194%	0.072%	-8.553%	10.055%	59,193
Index-Mid	0.045%	1.465%	0.100%	-11.152%	10.288%	44,602
Index-Small	0.042%	1.604%	0.099%	-11.371%	9.129%	37,075
Index-Fixed	0.017%	0.331%	0.023%	-2.563%	3.014%	57,774
Index-Avg.	**0.037%**	**1.148%**	**0.073%**	**-8.410%**	**8.121%**	
Total Avg.	**0.027%**	**1.185%**	**0.065%**	**-9.501%**	**9.085%**	

Table 2 shows the descriptive statistics on an average daily basis for the entire sample from January 1st, 2005 to December 31st, 2016. The ETF statistics are calculated with an average of equation (1) and equation (2). Mutual fund and index calculations are based on equation (1).

In Table 3, we examine the tracking errors which shows how each ETF and mutual fund performed relative to their benchmark. To estimate the tracking errors, we apply three different approaches as explained in the methodology section. TE_1 = Average of absolute return differences between funds and benchmarks. TE_2 = Standard deviation of differences between returns. TE_3 = Root mean square error from regression analysis. P-values are in parenthesis.

For all the categories of funds, the three tracking errors are lower for ETFs compared to mutual funds. The results from the t-tests show that the mean absolute return (TE_1) is statistically different at the 5% level. The last two rows show the average of the three tracking errors. The tracking error is lower for the equity funds and also for all funds. The average tracking error for the equity ETFs is 0.301% compared to

0.371% for equity mutual funds, the difference is statistically significant. The average tracking error for the all ETFs is 0.2641% compared to 0.3127% for mutual funds and is statistically significant.

TABLE 3
TRACKING ERROR ANALYSIS (ALL FIGURES ARE PERCENTAGES)

Category	TE1		TE2		TE3		Avg.	
	ETF	MF	ETF	MF	ETF	MF	ETF	MF
Large	0.011	0.006	0.206	0.389	0.205	0.348	0.141	0.248
Mid	0.010	0.020	0.709	0.898	0.499	0.512	0.406	0.477
Small	0.005	0.024	0.842	0.753	0.630	0.731	0.492	0.503
Fixed	0.013	0.015	0.254	0.240	0.191	0.149	0.153	0.134
Avg. Equity	0.010	0.014	0.487	0.605	0.406	0.494	0.301	0.371
T-stat.		-1.89		-1.37		-1.58		-1.58
P-val.		0.03		0.09		0.06		(0.06)
Avg. All	0.011	0.014	0.430	0.515	0.353	0.410	0.265	0.313
T-stat.		-2.14		-1.21		-1.13		-1.58
P-val.		0.02			0.11		0.13	0.06

Next we evaluate the expenses associated with ETFs and mutual funds. In Table 4, we show the bid-ask spreads for ETFs using equation (6) for Spread 1, and equation (7) for Spread 2. Table 4 reports the descriptive statistics for the spread analysis. The percentile rows display the average of the bottom 80% of the sample.

TABLE 4
SPREAD ANALYSIS (ALL FIGURES ARE PERCENTAGES)

Category	Spread	Mean	Std. Dev.	Med.	Min	Max	Count
ETF- Large	Spread 1	1.335	1.656	0.633	-0.123	8.699	49,512
	Spread 2	1.308	1.607	0.630	-0.124	8.306	49,512
	Percentile	(0.652)	(0.732)	(0.313)	-(0.123)	(2.570)	(39,599)
ETF-Mid	Spread 1	2.657	3.559	0.704	0.000	16.599	25,998
	Spread 2	2.557	3.372	0.700	0.000	15.304	25,998
	Percentile	(1.164)	(1.570)	(0.336)	(0.000)	(5.357)	(20,793)
ETF-Small	Spread 1	2.569	3.188	1.000	-0.033	11.937	14,730
	Spread 2	2.487	3.045	0.992	-0.033	11.260	14,730
	Percentile	(1.203)	(1.389)	(0.48)	-(0.033)	(5.056)	(11,780)
ETF Fixed-Inc.	Spread 1	0.458	1.178	0.110	-0.002	12.004	35,134
	Spread 2	0.450	1.140	0.110	-0.002	11.321	35,134
	Percentile	(0.101)	(0.081)	(0.087)	-(0.002)	(0.361)	(28,098)
	Avg.	**1.728**	**2.343**	**0.610**	**-0.039**	**11.929**	**250,748**
		(0.780)	(0.943)	(0.306)	-(0.039)	(3.336)	(100,270)

The average spread ranges between 0.450% for fixed-income ETFs and 2.657% for the equity mid-cap ETFs. The median is significantly lower and ranges from 0.110% to 1.00%. The lower median indicates that the mean is driven by outliers. This statement finds confirmation in the maximum spread that reaches up to 16.69%. By measuring the average spread for the bottom 80% of our sample (these numbers are displayed in parentheses) we find a lower average spread across all categories. Such impact on the average spread by the top 20% can be caused by a very low trading volume of the ETF or a higher degree of illiquidity in the underlying asset on some days which leads to higher spreads.

The area of ETF spread analysis has not been thoroughly examined in the literature making these results particularly interesting. A few comparable studies find spreads between 0.01% and 0.03% in the U.S. large-cap section (Ivanov, 2016), 0.082% to 0.794% (Engle & Sarkar, 2002) and 0.009% to 1.27% (Agrrawal & Clark, 2009). A possible explanation for higher spreads in our study is the longer observation period.

Next, we examine if ETFs trade at premiums or discounts relative to their net asset value. The premiums and discounts are calculated by using equation (8). The results are presented in Table 5, the average premium for all segments is 0.04%. The premium/discount for the equity segments is much lower and ranges from -0.029% to 0.005%. The relatively small values indicates that the arbitrage mechanism works and prices are close to net asset value. The slightly higher premium of 0.178% for the fixed-income segment may be due to the less liquid nature of bonds. The largest discounts and premiums are 5.302% and 5.036%, respectively.

TABLE 5
PREMIUM ANALYSIS

Category	Mean	Std. Dev.	Med.	Min	Max	Count
ETF-Large	0.004%	0.163%	0.011%	-2.030%	2.665%	58,739
ETF-Mid	0.005%	0.279%	0.016%	-3.751%	2.949%	37,411
ETF-Small	-0.029%	0.434%	-0.017%	-5.036%	5.302%	26,318
ETF-Fixed Inc.	0.178%	0.331%	0.108%	-3.626%	3.118%	40,315
Average	**0.040%**	**0.302%**	**0.029%**	**-3.611%**	**3.509%**	162,783

Table 6 displays the percentage of days at which the ETFs trade at a premium, on par, or at a discount relative to net asset value. The days that equity funds trade at premium or a discount is comparable. The fixed income funds trade at a premium 80% of the time.

TABLE 6
DAYS AT PREMIUM, PAR, OR DISCOUNT

Category	Premium	Par	Discount	Count
ETF-Large	53.5%	5.9%	40.5%	58,739
ETF-Mid	55.3%	2.1%	42.6%	37,411
ETF-Small	44.9%	3.1%	52.1%	26,318
ETF-Fixed-Inc.	80.1%	3.4%	16.5%	40,315
Average	**58.4%**	**3.6%**	**38.0%**	162,783

In Table 7, we examine the nonrecurring charges of mutual funds and ETFs. Mutual funds can charge loads, which range from 5.75% for front loads or as low as 0%. In our sample, only 25 out of the 61 mutual funds charge a load.

TABLE 7
NONRECURRING EXPENSES

	Mutual Funds			ETFs	
Category	Front Load	Back Load	Early Withdraw	Prem.	Spread
Large	1.088%	0.350%	0.025%	0.004%	1.321%
Mid	1.797%	0.250%	0.250%	0.005%	2.607%
Small	0.575%	0.100%	1.100%	-0.029%	2.528%
Fixed-Inc.	1.350%	0.117%	0.033%	0.178%	0.454%
Average	1.202%	0.204%	0.352%	0.040%	1.728%
Sum	**1.759%**			**1.767%**	

We examine net annual returns, annual expenses and gross annual returns in Table 8. The annual return is calculated by multiplying the daily return by 250; which is the average number of trading days in a year. The annual expenses are reported by each fund and the gross return is calculated by adding net annual returns and expenses for each fund. We report the results for each category. The net annual returns in each category are lower for mutual funds compared to ETFs the largest difference is in the small cap and mid-cap categories. The expenses in each category are higher for mutual funds compared to ETFs in each category. Once again the largest difference is in the small cap and mid cap categories. The gross annual returns are lower for mutual funds compared to ETFs. The small cap and mid cap funds have the largest difference.

TABLE 8
NET ANNUAL RETURNS, EXPENSES, AND GROSS ANNUAL RETURNS

Category	**Net Annual Return**		**Annual Expenses**		**Gross Annual Return**	
	MF	ETF	MF	ETF	MF	ETF
Large	6.045%	7.533%	0.62%	0.25%	6.67%	7.78%
Mid	5.740%	11.124%	0.89%	0.25%	6.63%	11.37%
Small	3.044%	9.410%	1.30%	0.47%	4.34%	9.88%
Fixed Inc.	0.563%	0.956%	0.53%	0.17%	1.09%	1.12%
Avg. Equity	5.2866%	9.1899%	0.862%	0.296%	6.149%	9.486%
T-stat.		7.24		-8.07		6.28
P-value		(0.0000)		(0.0000)		(0.0000)
Avg.- All	4.1251%	7.1652%	0.7798%	0.2643%	4.9049%	7.4296%
T-stat.		-8.78		-8.78		3.72
P-value		(0.0000)		(0.0000)		(0.0000)

Table 8 show the average for each category, all equity funds, and all funds. For the equity funds, the net annual returns for mutual funds is 5.2866% compared to 9.1899% for ETFs; the net returns are statistically different. The expenses for mutual funds are 0.862% compared to 0.296% for ETFs which is also statistically different. The last two columns show that the gross annual returns for mutual funds is 6.149% compared to 9.486% for ETFs; this is also statistically different. The difference in gross annual returns supports the notion that the higher return for ETFs is not solely driven by differences in the expenses but also by the higher gross returns on the ETFs.

CONCLUSION

In this study, we examine the tracking error, expenses, and returns on 61 ETFs and 61 mutual funds. In each category, we find that the mutual funds have higher tracking errors compared to the ETFs. The results from the t-tests show that the tracking errors are significantly different. The tracking error does not show if the variation from the benchmark is positive or negative; thus we examine the returns to determine if ETFs or mutual funds have better performance. The returns are partially determined by the expenses in the ETFs and mutual funds; thus we provide a detailed examination of all the expenses.

We examine annual and nonrecurring expenses in mutual funds and ETFs. We find that mutual funds have higher expense ratios than ETFs. We also examine the bid-ask spread and find relatively high spreads for ETFs, especially in the mid- and small-cap sections. Another interesting finding is that ETFs often trade at discounts or premiums relative to their net asset value, although the magnitude is relatively small.

The last part of the paper examines the net annual returns, annual expenses and gross annual returns. We find that for each category ETFs have higher net annual returns, lower expenses and higher gross annual returns. These differences are statistically significant. The results indicate that the ETFs higher returns are driven by lower annual fees and superior investing.

REFERENCES

Agrrawal, P., & Clark, J. M., (2009). A Multivariate Liquidity Score and Ranking Device for ETFs. *Academy of Financial Services*, 2009.

Aber, J. W., Li, D., & Can, L. (2009). Price volatility and tracking ability of ETFs. *Journal of Asset Management*, 10(4), 210-221.

Brandon, T., Rottschafer, D., & Zvingelis, J. (2013). *A Tracking Error Primer*.

Chang, C. E., Krueger, T. M., & Witte, H. D. (2015). Do ETFs outperform CEFs in fixed income investing? *American Journal of Business*, 30(4), 231-246.

Engle, R., & Sakar, D. (2002) *Pricing Exchange Traded Funds*.

Foucher, I., & Gray, K. (2014). Exchange-Traded Funds: Evolution of Benefits, Vulnerabilities and Risk. Bank of Canada, *Financial System Review*.

Frino, A., & Gallagher, D. R. (2001). Tracking S&P 500 Index Funds. *Journal of Portfolio Management*, 28(1), 44-55.

Gallagher, D. R., & Segara, R. (2006). The performance and trading characteristics of exchange-traded funds. *Journal of Investment Strategy*, 1(1), 47-58.

Gastineau, G. L. (2001). An Introduction to Exchange-Traded-Funds (ETFs). *Journal of Portfolio Management*, 27 (3), 88-96.

Gastineau, G. L. (2004). The Benchmark Index ETF Performance Problem. *ETFs and Indexing*, 30 (2), 96-103.

Harper, A. T., Madura, J., & Schnusenberg, O. (2004). Performance comparison between exchange-traded funds and closed-end country funds. *Journal of International Financial Markets, Institutions and Money*, 16(2), 104-122.

Hill, J. M., Nadig, D., & Hougan, M. (2015). A comprehensive guide to exchange traded funds (ETFs). *Research Foundation Publication*, 2015(3), 1-181.

Hilliard, J. (2014). Premiums and Discounts in ETFs: An analysis of the arbitrage mechanism in domestic and international funds. *Global Finance Journal*, 25(2), 90-107.

Ivanov, S. I. (2016). Analysis of ETF bid-ask spread components. *The Quarterly Review of Economics and Finance*, 61, 249-259.

Kosev, M., & Williams, T. (2011). Exchange-traded Funds. *Bulletin*, 1, 51-60.

Kostovetsky, L. (2003). Index Mutual Funds and Exchange-Traded Funds. *The Journal of Portfolio Management*. 29(4), 80-92.

Mercado, S., Lan, S., & Rajendra, A. (2017). *ETF Annual Review & Outlook*.

Poterba, J., & Shoven, J. (2002). Exchange-Traded Funds: A New Investment Option for Taxable Investors. *The American Economic Review*, 92(2), 422-427.

Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance*, 39(4), 1127-1139.

Rompotis, G. G. (2005). An Empirical Comparing Investigation on Exchange Traded Funds and Index Funds Performance.

Shin, S., & Soydemir, G. (2009). Exchange-traded funds, persistence in tracking errors and information dissemination. *Journal of Multinational Finance Management*, 20(4-5), 214-234.

Svetina, M., & Wahal, S. (2008). Exchange Traded Funds: Performance and Competition.

Vanguard Financial Advisor Services. (2015). Premium and discounts. *Investor education*.