Abstract

We compare tracking abilities between exchange traded funds focused on emerging and developed markets. Because the ETF is a relatively new financial instrument (first inception 1993), there is limited literature on the tracking abilities for ETFs focused on emerging and developed markets. Previous literature have shown a positive correlation between ETFs tracking errors and exchange rates, indicating that ETFs tracking foreign markets underperform ETFs tracking domestic markets. Our conclusion is that ETFs tracking developed markets exhibit lower tracking error than ETFs tracking emerging markets, with trading volume being the only factor affecting tracking error.

Acknowledgements

We would like to express our gratitude and a special thanks to Professor Måns Söderbom, Mohamed Reda Moursli and Laura Villalobos-Fiatt for sharing time, insight and comments on our econometric issues.

Contents

Introduction

The purpose is to analyze and investigate if there are differences in performance and index tracking abilities between emerging and developed markets ETFs. We will examine the hypothetical disparities in tracking performances between ETFs focused on developed and emerging markets.

Exchange Traded Funds (ETFs) are funds that trade at the stock market; they look like mutual funds but trade like stocks. ETFs are sometimes seen as alternatives to traditional mutual funds, the main differences being that ETFs, as opposed to mutual funds, can be traded at any time during the opening of the stock markets and be sold short or purchased on margin (Bodie et al. 2009:132 ff). They have a so called in-kind creation and redemption mechanism, which carries several advantages over other financial instruments. One advantage is the transparency they exhibit, which enables investors to efficiently observe and react to arbitrage opportunities, thus preserving the ETFs market price close to its net asset value (NAV).

The growing popularity of the ETF as a financial instrument is extraordinary and since the first inception in 1993 of the SPDR, or "spider" which stands for Standard and Poor's Depository Receipt, the market has seen an exponential increase in the number of ETFs and Table 1 depicts this growth over the last decade. When first introduced in 1993, the "spider" was the only ETF on the US market, and its value was approximately US$500 million. Since then, as mentioned above, the market for ETFs has grown substantially and as of 2012 more than 1,194 ETFs were traded, at a total worth of approximately US$1.34 trillion.

Table 1: Total Net Assets and Number of ETFs

Year	Total net assets, all ETFs	Number of ETFs
2001	83	102
2002	102	113
2003	151	119
2004	228	152
2005	301	204
2006	423	359
2007	608	629
2008	531	728
2009	777	797
2010	992	923
2011	1 048	1 134
2012	1 337	1 194

Billions of dollars, year-end, 2001–2012

There is a large pool of literature on conventional mutual funds which covers performance, management, fund structure and costs. Because the ETF is relatively new as a financial instrument, there is limited literature on this topic, and what little there is focuses on ETFs tracking error.

Gastineau (2001) basically provided an overview of exchange-traded funds and explains the key characteristics of the market, including the variety of ETFs. Frino and Gallagher (2001) highlighted the importance of constraining a funds tracking error to deliver identical results as the underlying index. They found that tracking error is unavoidable due to market frictions and that on average, index funds outperform actively managed funds after expenses. Shin and Soydemir (2010) discovered that exchange rates are significantly related to the ETFs tracking error. They could also show that the Asian markets display large persistence in tracking error indicating it being less efficient than the US market. Furthermore, Frino et al. (2004) found that the tracking error is significantly related to the liquidity, volatility, index replication strategy and fund size.

Elton et al. (2002) presented the characteristics and performance of the SPDR. They concluded that the SPDR underperform the S&P500 and conventional mutual funds, but on the other hand, Spiders are the most actively traded instrument on the stock exchange. Due to the creation/redemption process, the deviations between the ETFs price and its NAV should disappear within a day, indicating that the market is efficient.

There has also been much research on the differences between ETFs and mutual funds. Kostovetsky (2003) concluded that the key differences between ETFs and mutual funds are management fees, shareholder transaction costs and taxation efficiency for the ETF. According to Rompotis (2009), ETFs track their underlying indices more accurately than mutual funds, but they underperform the return on the underlying index return significantly, much due to expense ratios. Gastineau (2004) documented that the pre-tax performance of ETFs have generally underperformed large conventional mutual funds tracking the same underlying index. This is attributable to the time the funds adjust to changes in the composition of assets in the underlying index. ETFs change their composition of assets on the same day as the underlying index, whereas mutual funds adjust their composition as soon as the announcement is made. Aber et al. (2009) compared the tracking ability between ETFs and mutual funds pointing out that ETFs are more likely to trade at premium rather than at discount, with large daily price fluctuations. On average, the ETFs have larger tracking error than their competitive mutual funds.

To our knowledge there has been scarce research on the tracking ability of ETFs between developed markets and emerging markets. In this paper we contribute to the ETF literature by investigating on this matter.

ETFs Characteristics

The market for ETFs is diversified, where ETFs tracking equity indices dominate the market. These are classified as either broad-based ETFs which covers a whole market, or sector ETFs, covering specific sectors. Other types are leveraged, commodity and developed/emerging markets ETFs. The leveraged ETFs' main purpose is to use financial derivatives and debt to increase the return of the underlying index, whereas commodity ETFs invest primarily in physical commodities, e.g. precious metals or natural resources. The developed/emerging markets instruments invest in underlying indices focused on either developed or emerging markets, and they are the main focus of this thesis.

Creation and redemption

The creation and redemption mechanism is one of the unique features of the ETFs. When the ETF provider are about to launch a new product or create new shares of an existing ETF, it turns to an authorized participant (AP), usually a market maker or a financial institution with buying power.

To create ETF shares, the AP buys a basket of the underlying ETF stocks from the market and delivers them to the ETF provider. In return, the AP gets a number of shares of equal value, called a creation unit. In most cases, the ETF provider imposes a fee to cover the administrative costs of handling the creation or redemption process. A minimum of 50,000 ETF shares has to be created and placed before the market close. The AP can choose to hold these shares or sell them at the stock market, the latter is usually the case. The redemption process is just the reverse of the creation process. The AP delivers a minimum of 50,000 ETF shares to the ETF provider, and the ETF company provides the AP with the underlying stocks of equal value.

The purpose of this mechanism is to keep the ETFs' price in line with its NAV. Because the ETF trades like a stock, its price will fluctuate during the day. The price will at some point be different from its NAV which creates an arbitrage opportunity. The AP will then create or redeem shares to make use of this arbitrage opportunity.

To make the creation/redemption mechanism efficient, the ETF must be transparent so that APs can replicate the underlying stocks perfectly. Thus, the ETF market is one of the most transparent markets in the world. During any time of the day you can see which the underlying stocks are and the exact weights of the portfolio. A mutual fund on the other hand, discloses their underlying assets quarterly or even less frequently, implying that the price can be far away from its NAV.

Liquidity is another important feature for the efficiency of the creation/redemption process. In a highly liquid market, the ETFs' price will be closer to its NAV, and with a low liquid market, the ETFs' price will be further away from its NAV. There is also a liquidity perspective from the investors on the secondary stock markets point of view. With a high turnover of the ETF, it makes it easier to quickly sell it to other investors. On the contrary, a low turnover could make an investor wait for days to be able to sell the ETF (ETF Education Center 2010).

Tax advantages

ETFs provide significant tax efficiencies compared to conventional mutual funds, much due to ETFs in-kind creation and redemption process. One might ask why, and the simple answer is that there are no taxes that transmits to existing investors since the only time taxes would impose a cost is when the entire holding of ETFs are sold by the investor. By contrast, all mutual fund holders incur capital gains when a mutual fund is redeemed, which bring tax disadvantages to existing fund holders. For the tax deferred investor though, there is no advantage in holding ETFs over mutual funds (Dellva 2001). This has also been shown by Gastineau (2004) who concluded that conventional mutual funds outperform ETFs when it comes to the pre-tax performance.

Expense Ratios

The expense ratio is a measure of operating costs for the provider of an ETF or a mutual fund. It is an annual expense taken directly from the ETFs' assets and therefore it lowers the investors' return. Compared to mutual funds, ETFs often have as low as or lower expense ratios. One interesting fact is that there is a general difference in expense ratios among ETFs tracking developed and emerging markets.

Bid/Ask Spread

The bid/ask spread is the gap between the price buyers are willing to pay, and the price sellers are willing to sell ETFs for. When trading an ETF, one buys it at the higher asking price, and then sells it at the lower bid price. The wider the spread is, the more it will eat up of the total return. The spread on an exchange-traded fund depends on the liquidity and volume of the ETF. For an ETF with high liquidity, equivalently large volume, the spread will be low, often just pennies. When an ETF has low liquidity, the spread is higher. This spread is often a bigger concern for frequent intra-day traders rather than the buy-and-hold investors (ETF Education Center 2010).

Commissions

ETFs trade at the stock market, as pointed out earlier, and therefore they generate commission fees. This fee is generally determined by factors such as which type of broker is used, the amount invested and the timing of the trade. Commissions are not a concern for institutional investors, since all the transactions are made directly via the fund managers, thus there are no brokerage fees to take into account.

Dellva (2001) points out that for small dollar investors making monthly deposits, mutual funds typically generate lower costs than ETFs. This is due to ETFs frequent inquiry of commissions and bid/ask spread costs over time. On the other hand, if the investment is of a lump sum type, the lower expense ratios for ETFs have an advantage over the mutual funds.

Research Methodology

The conducted research is a comparative quantitative study. We have chosen five different ETFs to cover the world markets, all from the leading provider of ETFs, State Street Global Advisors. We use exchange-traded funds from the same provider to avoid any bias between management styles, rules and other factors that might differ between providers (Johnson 2009). Two ETFs that largely reflect the developed markets are SPY, tracking the S&P500® index, and FEZ, tracking the Euro STOXX 50® index. In addition, we have three emerging market ETFs covering three different markets; Middle East & Africa, Latin America and Asia excluding Japan. These three ETFs are chosen because, combined, they reflect a large part of the emerging markets in the world. Individually, they represent many emerging countries in the same region. We will compare these three emerging market ETFs against the two developed market ETFs. Furthermore, we use their main benchmark indices as a comparative value measure for the ETFs performances. We decided not to include an ETF for the BRICS, since we are also interested in comparing the individual difference in performance between the emerging markets. Table 2 below depicts the chosen ETFs along with their corresponding benchmark index and their tickers.

Table 2: List of ETFs

ETF	ETF Ticker	Benchmark Index	Index Ticker
SPDR® S&P500®	SPY	S&P500® Index	SPTR
SPDR® EURO STOXX 50®	FEZ	EURO STOXX 50® Index	SX5U
SPDR® S&P® Emerging Middle East & Africa	GAF	S&P® Mid-East and Africa BMI Index	STBMMEU
SPDR® S&P® Emerging Latin America	GML	S&P® Latin America BMI Index	STBMLAU
SPDR® S&P® Emerging Asia Pacific	GMF	S&P® Asia Pacific Emerging BMI Index	STBMAEU

List of ETFs name, its ticker, the benchmark index and its ticker

For all of our chosen ETFs, we have collected weekly data from November 7, 2008 until November 1, 2013, which gives us a sufficient volume of data to work with, 260 observations per ETF. Since we suspect finding autocorrelation in the error terms, Newey-West estimations will be used to correct for it. We have chosen weekly data to avoid some of the time effect on our results due to the different opening and closing hours of some of the ETFs and their underlying indices (from now on called *time effect*). There will still be a time effect, but using weekly data will hopefully reduce this effect. We presume no time effects on the SPY and GML since their underlying indices are based on the American continents, thus there cannot be any substantial time effect. However, FEZ, GMF and GAF all track indices based in other continents and thus trades at different opening and closing hours. By using weekly data, we expect the impact of the time effect to be lower on the performance and tracking error results, which in turn provides less biased regression results. When it comes to the premium/discount, we are interested in the daily effect, and as a result we will use daily data in our premium/discount analysis.

The ETF data we have collected include the historical NAV, the premium/discount to NAV and the bid/ask spread, all prices are given in US dollars. The NAV is calculated once each trading day, at 1600 hours Eastern Time and both the premium/discount to NAV and bid/ask spread are expressed in percentage form. We collected weekly last prices, where dividends are included, for all of our ETFs and their benchmark indices. Finally, all of our data are collected from Bloomberg.

Performance and risk

To begin with, we convert the weekly prices into weekly returns in percentage form using the following equation,

$$r_t = \frac{(\mathrm{Price}_t - \mathrm{Price}_{t-1})}{\mathrm{Price}_{t-1}} \times 100\% \qquad (1)$$

where r_t is the weekly return, Price_t is this week's price and Price_{t-1} is last week's price.

Thereafter, we will calculate the average weekly returns ($\bar{r}$), which is the sum of the weekly returns dividing by the number of weeks. As a measure of risk we use standard deviation, σ.

Finally, to make comparisons between the ETFs and indices we use a risk/return ratio,

$$\frac{\sigma}{\bar{r}} \qquad (2)$$

Premium/discount

The creation and redemption process, which is a function of the trading between the ETF provider and an AP, should cancel out any premiums when the AP trades in ETFs to the provider in return for stocks which sell back to the market. Any discounts should be canceled out when the AP trades in stocks against ETFs and then sells them back to the market. Yet, it is common that ETFs trade below or above the NAV. Following Aber et al. (2009), the percentage by which the ETFs' closing or midpoint price differs from its NAV is viewed as the premium or discount of an ETF. The premium/discount equation is:

$$\alpha_t = \frac{\text{Price}_t - \text{NAV}_t}{\text{NAV}_t} \qquad (3)$$

α_t is the measure unit of the premium/discount from NAV. $\alpha_t > 0$ represents a premium and $\alpha_t < 0$ represents a discount.

Regression analysis

We make use of time-series to evaluate the tracking ability on each of the ETFs,

$$R_{Et} = \alpha_i + \beta_i R_{Bt} + \varepsilon_{Et} \qquad (4)$$

where R_{Et} is the return of the ETFs, R_{Bt} is the return of the benchmark index and ε_{Et} presents the residual error of the ETF. α_i (alpha) indicates whether the return of the ETFs is above or below the return of the benchmark index, thus a positive alpha indicates a return above the benchmark index return and vice versa. We assume the alphas to be negative and statistically insignificant but yet economically significant since they all inherit expense ratios.

The coefficient β_i (beta) will be interpreted as a measure of the ETFs' aggressiveness towards the benchmark index, i.e. how well the ETF replicates the benchmark index. Beta higher than one indicates that the ETF moves more aggressive than its benchmark index, so if the ETF exhibits a beta of 1.05 the intuition will be that for every 1 unit increase in the benchmark index, the ETF will increase by 1.05 units. By contrast, a beta below unity infers that the ETF follows the benchmark index more passively. Furthermore, with a beta equal to unity, the ETF replicates the benchmark index perfectly. The beta can be considered as a measure of systematic risk.

Tracking error

The returns on ETFs are supposed to replicate the returns of the underlying indices, but Elton et al. (2001) showed that ETFs tend to underperform the returns of the underlying indices due to e.g. fees and other service expenses. To measure how large this underperformance is, we compare the weekly returns of the ETFs to the underlying indices by applying three different methods suggested by Frino and Gallagher (2001) and Rompotis (2009). We then calculate the average tracking error from these three methods, since we believe them to exhibit different tracking errors. The first method uses the root MSE of regression (4).

Second, we measure the tracking error (TE) by taking the average absolute difference between the return of the ETF and its underlying index,

$$TE_1 = \frac{\sum_{t=1}^{n} |r_{it} - b_{it}|}{n} \qquad (6)$$

where r_{it} and b_{it} is the weekly return of the ETF and its underlying index, respectively. We use the absolute difference so the negative and positive value does not cancel each other out.

The third and most commonly used measure of tracking error is the standard deviation of the return difference between the ETF and its benchmark index,

$$TE_2 = \sqrt{\frac{\sum_{t=1}^{n} (e_{it} - \bar{e}_i)^2}{n-1}} \qquad (7)$$

where e_{it} is the difference of returns at week t between the ETF and its benchmark index and $\bar{e}_i$ is the average weekly return difference. Pope and Yadav (1994) note that TE_2 and root MSE will show the same tracking error of the beta of regression (4) is equal to unity. If the beta is not exactly equal to one, root MSE will overestimate the tracking error.

Factors that affects tracking error

Now that we have determined the tracking error we would like to examine what factors might affect the tracking error. Rompotis (2009) tested for expense ratio and risk as factors driving tracking error. He concluded that the expense ratio has a significant negative effect; a higher expense ratio results in lower tracking error. The risk on the other hand showed a positive effect on tracking error, however this result was statistically insignificant. As a result, we choose not to test for this factor.

Since the liquidity differs largely between the ETFs, we would want to see how it affects the tracking error. We use trading volume as a measure of liquidity inferring that a larger trading volume should result in a lower tracking error. We also include the premium/discount factor, where a value over zero indicates a premium and a value below zero indicates a discount. The coefficient of premium/discount should be positive because a higher premium/discount should result in a higher tracking error. Finally, the third factor we believe affect the tracking error is the bid/ask spread. This coefficient ought to be positive as well, a higher bid/ask spread should indicate a higher tracking error. We will make use of cross-sectional regression,

$$\text{logTE} = \beta_0 + \beta_1 \text{logVol} + \beta_2 \text{PremDisc} + \beta_3 \text{BidAsk} + \varepsilon \tag{8}$$

where logTE is the log of tracking error and logVol is the log of trading volume. PremDisc and BidAsk is the premium/discount and the bid/ask spread, respectively. Both PremDisc and BidAsk are expressed in percentage form. We use logarithmic form on the dependent variable *tracking error* and the independent variable *volume* since they are positive and show large spreads their values.

Cumulative Returns

We calculate the cumulative returns (R) to construct graphs displaying the performance of the ETFs and their benchmark index,

$$R = \sum_{t=1}^{T} r_t \qquad (5)$$

where r_t is the weekly returns.

Results

Initially we will present the performance results of the ETFs compared to their underlying indices. We will also display the premium/discount and the regression results for the ETFs. Furthermore, the different measures of tracking error and the factors affecting tracking error will be presented here.

Performance and risk

The performance of the ETFs and their underlying indices are presented in table 3. The results are presented as the average weekly return and risk. The difference in return compared to the underlying indices is higher for the emerging markets.

Table 3: Returns and risk

ETF Ticker	Return		Risk		Risk/Return	
	ETF	Index	ETF	Index	ETF	Index
SPY	0.322	0.325	2.751	2.728	8.553	8.399
FEZ	0.242	0.241	4.268	4.316	17.636	17.909
GML	0.304	0.375	4.116	4.153	13.539	11.075
GMF	0.342	0.394	3.330	2.977	9.744	7.566
GAF	0.331	0.339	3.413	3.185	10.311	9.395

This table presents the average weekly return and risk, both measured in percentage, of the ETFs and their benchmark indices. It also shows the risk/return ratio, indicating the risk per unit of returns.

GMF (Asia, excluding Japan) yield the highest average return and FEZ (Europe) the lowest, with approximately 34 and 24 b.p., respectively. FEZ and its benchmark index, the EURO STOXX 50, suffered huge losses due to the Europe crisis starting in 2011 and this might be the reason FEZ exhibits low average returns. GML (Latin America), GMF and GAF (Africa & Middle East) displays differences in returns compared to their benchmark index of approximately 7, 5 and 1 b.p., respectively. We were expecting GML to display lower discrepancy than its benchmark index since it is not affected by different trading hours, compared to GMF and GAF which are both affected by different trading hours. Furthermore, it is surprising that GAF only has a difference of 1 b.p. On the other hand, FEZ and SPY shows the lowest discrepancy in return compared to their benchmark index with 0.1 and 0.3 b.p., respectively, well in line with our expectations. FEZ has 0.1 b.p. higher average return in comparison to its benchmark index, which is very surprising since FEZ should underperform EURO STOXX 50 by the expense ratio, corresponding to 29 b.p.

SPY has the lowest average weekly standard deviation with 2.751 percent and FEZ the highest with 4.268 percent. FEZ along with GML even had lower standard deviation than their benchmark index. The difference in risk between the ETF and its benchmark index is lowest for SPY, GML and FEZ with 2.3, 3.7 and 4.8 b.p., respectively. GMF and GAF exhibit highest differences with 33.3 and 22.8 b.p., respectively.

All ETFs have higher risk to return ratio than their benchmark indices, except FEZ, indicating that the ETFs need to take on more risk than their benchmark indices to generate higher returns. FEZ has 27.3 b.p. lower risk/return ratio towards its benchmark index, which is due to a higher return and lower risk than its underlying index. It's surprising that FEZ has the highest risk to return ratio. This might be, as earlier explained, due to the effect of the Europe crisis during the chosen sample period. One possible explanation to the inconsistencies in return and risk between the ETFs and their corresponding benchmark index is that the ETFs might have betas (β) not equal to 1, meaning the ETFs does not perfectly replicate their benchmark index. The betas will be shown in the subsequent section where we provide the betas together with the regression results.

Premium/discount

Table 4 shows the premium or discounts to NAV of the ETFs. It tells us how many days the ETFs trade at a premium, discount or at a price equal to the underlying indices. As expected SPY trades at equal on most days. What is surprising is that GML trades at equal more times than FEZ which might be due to the time difference in trading hours affecting FEZ (Europe), but not GML (Latin America). All ETFs, except SPY, trades mostly at a premium.

Table 4: Number of days at premium/discount

ETF	Premium	Discount	Equal	Total days	t-test
SPY	516	573	160	1249	-0.787
FEZ	767	474	8	1249	5.391
GAF	667	578	5	1250	3.558
GML	646	587	17	1250	6.606
GMF	713	529	8	1250	5.388

Number of days the ETF trades at premium, above its NAV, at discount, below its NAV, or at a price equal to its NAV. The t-values for all ETFs, beside SPY, are significant at any conventional significance level, which states that they frequently trade at prices above or below their NAV.

Regression analysis

The results of the time-series regression (4) are shown in Table 5. All ETFs, except FEZ, exhibit alphas below zero. Though they are all statistically insignificant, indicating they have the same performance as their underlying index, these alphas are economically significant, especially to investors making large dollar investments. In practice, differences in returns amounting to single basis points can be a matter of earnings or loses of millions of dollars. Therefore, it should not be hard to grasp the economic importance of alphas, whether or not they are statistically significant. The beta, which represents the systematic risk and replication aggressiveness, are close to unity for all ETFs. Generally, R^2 is high for all ETFs, indicating good replication strategies. All ETFs showed significant autocorrelation in the error terms and therefore we used the Newey-West estimations with different lags to correct for the autocorrelations.

Table 5: Regression results

ETF Ticker	Alpha (α)	t-test	Beta (β)	t-test	R^2
SPY	-0.005	-0.74	1.006	0.52	0.996
FEZ	0.012	0.42	0.956	-1.45	0.934
GML	-0.056	-1.53	0.959	-0.75	0.936
GMF	-0.056	-1.22	1.012	0.43	0.819
GAF	-0.0005	-0.01	0.977	-0.38	0.831

This table shows the results from regressing each ETF against its benchmark index using regression (4): $R_{Et} = \alpha_i + \beta_i R_{Bt} + \varepsilon_{Et}$, where R_{Et} is the return of the ETF, R_{Bt} is the return of the benchmark index and ε_{Et} presents the residual error of the ETF. The alpha indicates wheather the return of the ETFs are above or below the return of its benchmark index. The beta coefficient indicates the ETFs aggressiveness towards its benchmark index, i.e. how well the ETF replicates its benchmark index. The t-test for alpha states if it differs from zero, the t-test for beta states if it differs from one.

For SPY, the alpha coefficient of -0.005 indicates that it underperforms the benchmark index by 0.5 b.p. per week. FEZ has a positive alpha of 1.2 b.p., indicating it outperforms its benchmark index and this result is consistent with the previous return analysis from table 3. GML and GMF have the same alpha value of -5.6 b.p., both underperforming its underlying index. GAF shows an alpha of shockingly -0.0005, underperforming by only 0.05 b.p. per week, which is the lowest of all ETFs. We expected GAF and GMF to have the highest alphas, in absolute terms, mostly due to time effects. SPY should have the lowest alpha, followed by FEZ, GML, GMF and GAF. Therefore, if we were to exclude GAF from our analysis, our results would be in line with our expectations.

SPY has a beta value of 1.006, which indicates a slightly higher risk and a marginally higher aggressiveness towards its benchmark index, with the intuition being that if S&P500 increase by 1 unit, SPY increase by 1.006 units. The remaining ETFs all have betas with larger disparity than SPY. GMF, with a beta of 1.012, can be interpreted the same way as SPY, with a higher risk and aggressiveness towards its benchmark index. FEZ, GML and GAF have betas below unity, displaying more conservative replication strategies toward their benchmark indices and lower systematic risk. We expect these ETFs to exhibit greater discrepancy than SPY when compared to their benchmark index over the long run. The t-test on beta, testing if beta differs from one, shows that FEZ has the highest t-value, but still insignificant. Hence, we cannot reject the null hypothesis that the ETFs fully replicate their benchmark index.

The reported yearly expense ratio for SPY and FEZ are 11.02 and 29 b.p. respectively, whereas for GML, GMF and GAF it is 59 b.p. Converting the generated alphas to yearly rates, we find the expense ratios are much higher than what is reported by the ETF provider State Street Global Advisors. Our findings shows that the yearly expense ratios for the two developed markets ETFs SPY and FEZ are approximately 26 and 63 b.p., respectively. For the emerging markets ETFs GML and GMF the yearly expense ratio are 2.95 percent, and for GAF it is 2.6 b.p.

Tracking error

Table 6 presents the results from the three different methods used for calculating the tracking error. It also shows the average tracking error derived from these individual tracking errors. The ranking of the average tracking error is in line with our expectations that developed market ETFs exhibit lower tracking error than emerging market ETFs.

Table 6: Tracking Error

ETF ticker	Root MSE	TE_1	TE_2	Average TE
SPY	0.181	0.086	0.181	0.149
FEZ	1.098	0.783	0.857	0.913
GML	1.040	0.657	1.050	0.916
GMF	1.421	1.050	1.416	1.296
GAF	1.405	0.959	1.402	1.255

The tracking error reflects the deviations between the return of the ETFs and its benchmark index. The tracking error are measured using three different methods; Root mean square error, TE_1 and TE_2. Root MSE is the standard error of regression (4). TE_1 is the average absolute difference in return between ETFs and benchmark indices. TE_2 refers to the standard deviation in return difference between ETFs and benchmark indices. The average TE is the sum of Root MSE, TE_1 and TE_2 divided by 3.

According to Pope and Yadav (1994), root MSE and TE_2 should give the same result if the beta of regression (4) is exactly equal to unity, which is not the case in our regression, where none of the ETFs has a beta exactly equal to one. Therefore, root MSE overestimates the tracking error. Because of this overestimation, we compute the average tracking error which gives more justified tracking error estimation. As a result, SPY has the lowest average tracking error at 0.149 followed by FEZ at 0.913. The relatively high average tracking error for FEZ compared to SPY, might be due to the difference in trading hours between the underlying index and the ETF. The three emerging markets GML, GMF and GAF have an average tracking error of 0.916, 1.296 and 1.255, respectively. One remarkable aspect is that we believed GML to exhibit a tracking error closer to GMF and GAF, rather than to FEZ. Here we can also draw the analogous conclusion that because the GML does not exhibit a time effect, since the underlying index trades approximately at similar hours as the ETF, it lowers the tracking error.

Factors that affects tracking error

Table 7 shows the results from the regression (8) of which factors that affect the tracking error. The tested factors are trading volume, representing the liquidity, premium/discount to NAV and bid/ask spread.

Table 7: Factors affecting tracking error

	logTE		
	β	t-test	p-value
logVol	-0.200	-12.22	0.000
Premium/Discount*	-0.065	-0.48	0.632
Bid/Ask spread*	-0.025	-0.51	0.608
		$R^2 = 0.25$	

This table presents results from the regression (8): $logTE = \beta_0 + \beta_1 logVol + \beta_2 PremDisc + \beta_3 BidAsk + \varepsilon$. logVol is the natural logarithm of volume, PremDisc is the premium and discount prices of the ETFs expressed in percentage form and BidAsk refers to the bid/ask spread in the ETF prices.

*** Variable expressed in percentage**

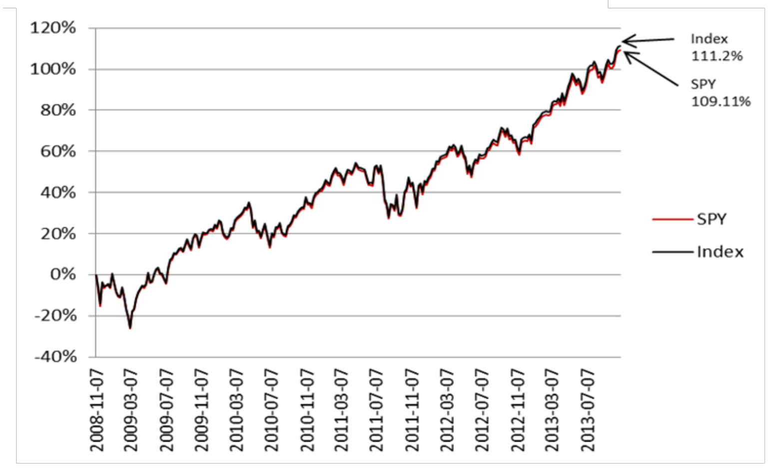
The only significant factor is the trading volume, significant at the one percent level. The coefficient value of -0.200 indicates that for every percentage point increase in volume, the tracking error decreases by 0.2 percentage points. This is in line with results from Frino et al. (2004) and Elton et al. (2002), that a high trading volume gives investors the opportunity to exploit arbitrage opportunities quickly, which will reduce the tracking error. The factors premium/discount and bid/ask spread have a negative coefficient. This is not what we expected; an increase in premium/discount or bid/ask spread should not lower the tracking error. These coefficients were insignificant, thus they do not affect tracking error. The explanatory power for this model is 25%, as can be seen by R^2 measure in the table.

Cumulative Returns

Figure 1 to 5, seen below, illustrates the cumulative performances of all ETFs (red lines) towards their benchmark index (black lines) during the five-year sample period using equation (5). These figures show that the developed market ETFs exhibit better tracking abilities than the emerging market ETFs. All ETFs fall into line with our expectations that the ETFs tracking developed markets display higher tracking accuracy than ETFs tracking emerging markets.

During the five-year sample period, SPY underperformed S&P500 by 2.09 percentage points, where 55 b.p. of the underperformance can be addressed to the expense ratio stated by the provider. This difference of 2.09 percentage points during this five-year period implies that SPY has higher total costs, including expense ratio, than stated. Furthermore, in Figure 1 we can now visually distinguish not only the effect of alpha, but also beta. The alpha of -0.005 indicate that SPY underperforms S&P500 by 0.5 b.p. per week, and the beta of 1.006 imply SPY is slightly more aggressive towards S&P500. Both alpha and beta are statistically insignificant; nevertheless, the economic significance is apparent over time. The average trading volume during the chosen period is approximately 660 million per week, and because we have shown that higher trading volume lowers tracking error we were expecting SPY to replicate the S&P500 almost perfectly.

Figure 1: Cumulative return SPY and S&P500



This figure displays the weekly cumulative return for SPY and SPTR (S&P500) closing prices.

Figure 2 shows that FEZ was succesfull in replicating the Euro STOXX 50. At the end of the five-year period FEZ overperformed its benchmark index by 1.33 percentage points. This result is striking, as we were expecting FEZ to underperform the benchmark index with 1.45 percent, which is addressed to the stated expense ratio. This implices that the total costs over this period was 12 b.p. due to the ETFs' ability to overperform its benchmark index.

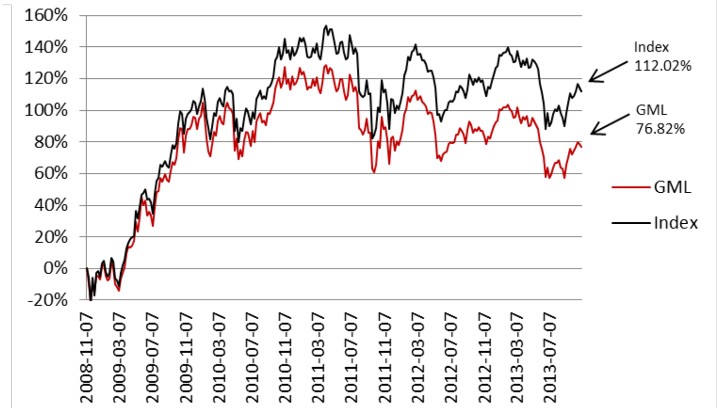
FEZ has a beta below unity, β = 0.956, indicating it should passively mimic Euro STOXX 50 over time. But since it is not statistically significant, this effect cannot be emphasised on the tracking ability. Due to the average trading volume of approximately 3,6 million per week, it results in lowering the tracking error. It is therefore not surprising that FEZ mimics the underlying benchmark well.

Figure 2: Cumulative return FEZ and EURO STOXX 50

This figure displays the weekly cumulative return for FEZ and SX5U (EURO STOXX 50) closing prices.

Figure 3 shows that GML underperformed S&P Latin America by astonishingly 35.2 percentage points during the five-year period, where only 2.95 percent of the underperformance can be explained by the expense ratio stated by the provider. This difference shows that the total costs for GML is much higher than the stated expense ratio. Although the beta of GML, β=0.959, should result in a better replication than Figure 3 shows, this beta is statistically insignificant and cannot be emphasized. The average trading volume of 60 000 per week, compared to SPY and FEZ that has over millions in trading volume every week, indicates it is not surprising that GML exhibit a high tracking error. Thus, this might be the reason for the large underperformance.

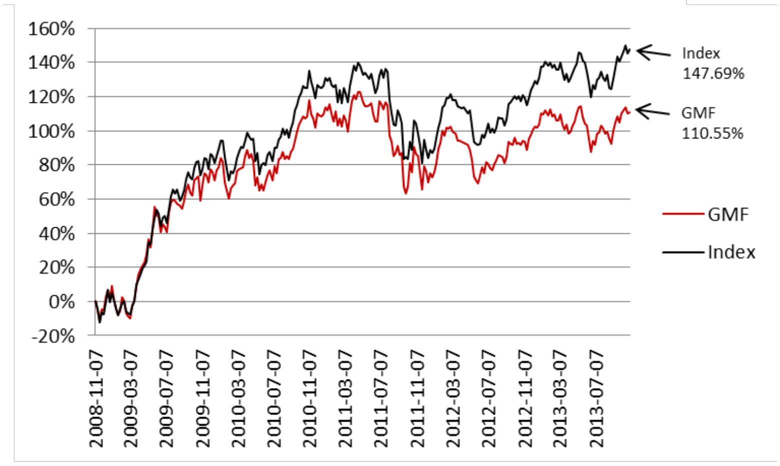
Figure 3: Cumulative return GML and S&P Latin America



This figure displays the weekly cumulative return for GML and STBMLAU (S&P Latin America) closing prices.

Figure 4, showing the cumulative returns for GMF and S&P Asia Pacific Emerging, the underperformance of GMF is even higher than GML and amounts to 37.14 percentage points, during the same time period. Just like for GML, 2.95 percent can be explained by the stated expense ratio. This is interpreted the same way as for GML, showing that GMF has much higher total costs than the stated expense ratio. The beta for GMF is 1.012, but it is statistically insignificant and should not be stressed as an explanation to the discrepancy in cumulative returns between the GMF and its benchmark index. The point that rather should be stressed is the average trading volume, amounting to approximately 260 000 per week, which is fairly low compared to the developed market ETFs.

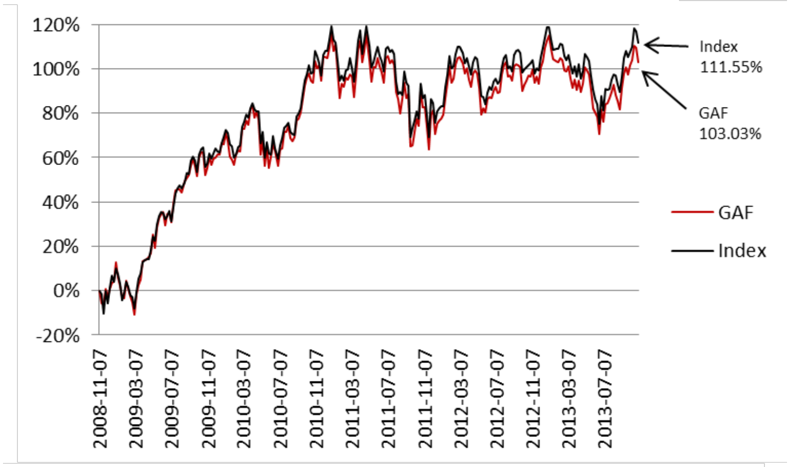
Figure 4: Cumulative return GMF and S&P Asia Pacific Emerging



This figure displays the weekly cumulative return for GMF and STBMAEU (S&P Asia Pacific Emerging) closing prices.

Comparing GAF to the other emerging market ETFs, it exhibits the best performance towards its benchmark index. Figure 5 shows that it only underperformed the S&P Mid-East and Africa by 8.52 percentage points. Furthermore, comparing to the stated expense ratio of 2.95 percent for the entire five-year period, GAF has higher total costs than the stated expense ratio. As can be seen in figures 3 to 5, GAF tracks its benchmark index more closely than GML and GMF. This is surprising due to the fact that GAF has the lowest average trading volume of approximately 42 000 per week, compared to 260 000 and 60 000 for GMF and GML, respectively. According to the findings in Table 7, which shows the factors affecting tracking error, GAF should exhibit lower tracking ability due to its lower average trading volume.

Figure 5: Cumulative return GAF and S&P Mid-East and Africa



This figure displays the weekly cumulative return for GAF and STBMMEU (S&P Mid-East and Africa) closing prices.

Conclusions

This paper contributes to the literature about exchange-traded funds by focusing on the aspect of emerging versus developed markets. The focus is mainly on ETFs performance and tracking error, but it also takes on the aspect of costs. We compare two developed markets, SPY and FEZ, against three emerging markets, GML, GMF and GAF.

In our performance analysis, we did not find any consistent results that developed market ETFs exhibit lower risk than emerging market ETFs. All ETFs, beside SPY, trades mostly at premium. Our regression analysis shows that alphas are statistically insignificant for all ETFs, with the interpretation that none of the ETFs perform different from its benchmark index. Moreover, an interesting finding is that the total costs are much higher than the stated expense ratios by the provider. The beta, which is a measure of systematic risk for the ETFs, does not differ significantly from one indicating that all ETFs exhibit full replication strategies.

The rank of tracking error corresponds to our expectations that developed market ETFs should have lower tracking error than emerging market ETFs, indicating the emerging markets is less efficient. The tracking error of GML and FEZ does not differ substantially and we believe this is due to the time effect that FEZ exhibits but not GML. We tested for volume, premium/discount and bid/ask spread as factors that might affect tracking error. We have shown that the only significant factor is volume, resulting in a decreasing tracking error as trading volume increases.

The figures displaying the cumulative returns shows that the developed market ETFs track their benchmark index better than the emerging market ETFs. One corresponding factor for the discrepancy in cumulative returns between the ETFs and their underlying benchmark is the large differences in trading volumes. The developed market ETFs trade in volumes of hundred millions, whereas the emerging market ETFs trade in significantly lower volumes of no more than a few hundred thousand.

When we tested for possible factors affecting the tracking ability, we did not include the time effect. We believe it might have a large impact on the creation/redemption process because the difference in trading hours restricts authorized participants from trading the underlying stocks. Another factor not considered in this paper is the exchange rates. Since the underlying stocks of the international ETFs are traded in foreign currencies (sometimes even different currencies within the same ETF), it might also have an impact on the tracking ability. It would be interesting to study these factors and their effect on tracking ability.

References

Articles

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

Dellva, L. W. (2001), "Exchange-traded funds not for everyone", *Journal of Financial Planning*, vol. 14, pp. 110-124.

Elton, E. J., Gruber, M. J., Comer, G., Li, K. (2002), "Spiders: Where are the bugs?", *Journal of Business,* vol. 75**,** pp. 453-472.

Frino, A., Gallagher D. R. (2001), "Tracking S&P 500 index funds", *Journal of Portfolio Management*, vol. 28, pp. 44-55.

Frino, A., Gallagher D. R. (2004), "Index design and implications for index tracking: evidence from S&P 500 index funds", *Journal of Portfolio Management*, vol. 30, pp.89-95

Gastineau, L. G. (2001), "Exchange-traded funds: an introduction", *Journal of Portfolio Management*, vol. 27, pp. 88-96.

Gastineau, L. G. (2004), "The benchmark index ETF performance problem", *Journal of Portfolio Management*, vol. 30, pp. 96-103.

Johnson, F. W. (2009), "Tracking errors of exchange traded funds", *Journal of Asset Management*, vol. 10, pp. 253-262.

Kostovetsky, L. (2003), "Index mutual funds and exchange-traded funds", *Journal of Portfolio Management*, vol. 29, pp. 80-92.

Pope, F. P., Yadav, P. K. (1994), "Discovering errors in tracking error", *Journal of Portfolio Management*, vol. 20, pp. 27-32

Rompotis, G. G. (2009), "Interfamily competition on index tracking: The case of the vanguard ETFs and index funds", *Journal of Asset Management*, vol. 10, pp. 263-278.

Shin, S., Soydemir, G. (2010), "Exchange-traded funds, persistence in tracking errors and information dissemination", *Journal of Multinational Financial Management*, vol. 20, pp. 214-234.

Book

Bodie, Z., Kane, A. & Marcus, A. J. (2011), *Investments and Portfolio Management*, 9th ed, New York: McGraw Hill Higher Education.

Internet Source

ETF Education Center, 2010. Available (online): <http://www.indexuniverse.com/etf-education-ce.html>. [23 October 2013].