The Effect of ETF's Benchmark's Market Structure and Geographic Region on Tracking Errors

By: Connor Haines

5/7/2017

Abstract: This paper looks at a sample of 20 iShare ETFs over the period from 2009-2015 that are listed on US-exchanges and benchmarked to foreign equity indexes. This paper finds that market structure (emerging or developed) and geographic region (Europe, Asia or Americas) have additional explanatory power for ETF's tracking errors when controlling for the variables: expense ratio, ETF trade volume, bid-ask spread, dividend yield, and changes in the exchange rate.

1. Introduction:

While exchange-traded funds (ETFs) are a relatively new financial innovation, they have experienced a rapid increase in use in recent years. Some of the primary benefits of ETFs are diversification, low cost, and tax efficiency. Additionally, ETFs are attractive to investors because they are traded on exchanges and thus provide high levels of liquidity and can be traded intraday (compared to mutual fund positions which can only be adjusted after market close). The first ETF ever created is commonly recognized to have been introduced in 1990, and was designed to track the TSE-35 stock index in Toronto, Canada. Since then the number of ETFs have grown rapidly. As of 2013, there were close to 4,000 different ETFs listed on over 50 exchanges around the world. From 2001 to 2013, ETF use grew over 1,400 percent to US $1.52 trillion (Charupat and Miu, 2013).

ETFs attempt to replicate the performance and risk characteristics of an underlying benchmark. Thus, an ETF's ability to replicate a benchmark accurately is an important characteristic of any ETF. Deviations between an ETF's returns and its benchmark's returns are called tracking error. My goal in this paper is to examine if the tracking error of US-listed ETFs benchmarking foreign equity indexes are explained by market structure (emerging or developed) and geography (Europe, Asia, or Americas) when controlling for determinants of tracking error found in previous financial research. I use a panel analysis of 20 ETFs over the period of 2009-2015 with either a market structure or geographic region dummy variable and interaction terms with five control variables: expense ratio, bid-ask spread, volume, dividends, and changes in the exchange rate. I find that market structure geographic region have additional explanatory power for tracking error. Several of the interaction terms are also significant.

The rest of this paper is organized as follows. Section 2 reviews the literature. Section 3 explains my motivation for this paper and how my research contributes to the existing research on ETF tracking errors. Section 4 describes how I obtained the data and provides summary statistics. Section 5 describes the econometric methodology. Section 6 reports the estimation results. Section 7 is my conclusion.

2. Literature Review

Charupat and Miu (2013) identify three main subsections of literature on ETFs. The first is research that focuses on ETF's pricing efficiency, that is, how close an ETF's market price multiplied by shares outstanding matches its assets under management (NAV), and how quickly creation/redemption arbitrage causes any discrepancies to disappear. The second topic of research concerns the effect ETF's have on related securities (such as underlying stocks and related derivatives). The third, which is summarized more fully below and is most closely related to my research, regards the performance of ETF's, ie. how well ETFs track their benchmarks. This is usually examined in terms of tracking error.

The most consistently established determinant of tracking error in previous literature is management fees. Previous research has found that the higher an ETFs expense ratio, the greater it's tracking error. This effect was found by Elton et. al. (2002), looking at Standard and Poor's Depository Receipts (SDPRs) and has since been found across many other ETFs and in many other time periods (see for example: Rompotis (2011) and Blitz et al. (2012)).

Previous literature has also identified other factors that may explain tracking error. Frino et al. (2004) and Gastineau (2002) find that index revisions can result in increased tracking errors. Dividend payments are identified as a factor affecting ETF tracking errors in Elton et al. (2002), Blitz et al. (2012) and Chu (2013). Osterhoff and Kaserer (2016) find that the underlying liquidity of the stocks comprising an ETF's benchmark affects tracking error. Rompotis (2012) finds a positive relationship between bid-ask spread and tracking error. Rompotis (2012) also finds a link between ETF replication strategy and tracking error.

While research focused on US-listed ETFs tracking foreign equity market indexes is somewhat rare, there are a couple of precedents in the literature. Johnson (2009) looks at 20 foreign country ETFs and how their tracking errors are affected by the trading hours of the country benchmarked. Shin and Soydemir (2010) look at sample of 20 iShare MSCI country ETFs from 2004-2007 and find that changes in the exchange rate are a significant source of tracking errors.

3. Motivation

While previous literature has found several factors that contribute to an ETF's tacking error, most of this research focuses on ETF's tracking domestic indexes. My motivation is to see if US-Listed ETFs benchmarked to foreign equity indexes have additional determinants of tracking error not accounted for by these known determinants. In particular, I'm motivated to analyze if tracking error can be attributed to the market structure of the underlying benchmark (emerging market or developed market) and the geographic region of the underlying benchmark (Europe, Asia, or Americas).

My research builds primarily off a study by Shin and Soydemier (2010). They analyzed if the tracking error of US-listed ETFs benchmarked to foreign equity indexes are explained by expense ratio, bid-ask spread, ETF trading volume, dividends, and changes in the exchange rate. I extend their analysis by seeing if, when controlling for these factors, there is additional tracking error attributable to an ETF's benchmark's market structure or geographic region. Additionally, I apply the more rigorous statistical panel analysis used by Osterhoff and Kaserer (2016).

4. Data

With the exception of annual expense ratios, all data was collected on a daily basis for the time period from the first trading day of 2009 until the last trading day of 2015. I collected data for 20 iShares Morgan Stanley Capital International (MSCI) country funds. Net asset values (NAVs), expense ratios, dividends and their benchmark index values are obtained from iShares of BlackRock Institutional Trust Company. ETF trading volume, highest bid and lowest ask price, and closing price were obtained through Wharton Research Data Services (WRDS) partnership with the Center for Research in Security Prices (CRSP). Exchange rates of each foreign currency to US dollars were acquired through WRDS partnership with the Federal Reserve's H10 Foreign Exchange Rate report.

Table 1 reports the profiles of the ETFs considered – including name of fund, ticker, geographic region, market structure, and NAV.

Table 1: Fund Overviews as of 12/31/2015

	Name of Fund	**Ticker**	**Market Structure**	**NAV ($ millions)**
Europe				
	MSCI Austria Investable Market Index Fund	EWO	DM	56.9
	MSCI Belgium Investable Market Index Fund	EWK	DM	237.9
	MSCI UK Index Fund	EWU	DM	2371.9
	MSCI France Index Fund	EWQ	DM	383.5
	MSCI Switzerland Index Fund	EWL	DM	1194.8
	MSCI Sweden Index Fund	EWD	DM	368.3
	MSCI Spain Index Fund	EWP	DM	1329.2
	MSCI Germany Index Fund	EWG	DM	6019.8
	MSCI Netherlands Investable Market Index Fund	EWN	DM	3232.3
	MSCI Italy Index Fund	EWI	DM	1089.8
Asia				
	MSCI Singapore Index Fund	EWS	DM	546.5

	MSCI Taiwan Index Fund	EWT	EM	2771.8
	MSCI South Korea Index Fund	EWY	EM	3232.3
	MSCI Malaysia Index Fund	EWM	EM	162.8
	MSCI Japan Index Fund	EWJ	DM	19899.6
	MSCI Hong Kong Index Fund	EWH	DM	2439.6
	MSCI Australia Index Fund	EWA	DM	1290.7
Americas				
	MSCI Brazil Index Fund	EWZ	EM	1947.7
	MSCI Mexico Investable Market Index Fund	EWW	EM	1276.2
	MSCI Canada Index Fund	EWC	DM	1726.4

Note: This table displays fund names, geographic region, market structure, ticker, and NAV (in $ millions) as of 12/31/2015 for all funds examined in this study.

Table 2 shows sample moments for the tracking error measurement each year of the sample. In all years, tracking error has a left-tailed skew, suggesting fatter tails than a normal distribution. I performed t-tests on the mean tracking error each year, and each result is significant at the 1% level. Additionally, the tracking errors tend to be negative. Negative tracking errors occur for 22,698 of 35,240 observations, or 64.4% of the time.

Table 2 – Tracking Error Summary Statistics by year, aggregate of all ETFs sampled

	Mean TE	Median TE	Std	Min	Max	Skew	Kurt
2009	0.088	0.016	0.269	0.000	4.589	7.209	70.919
2010	0.057	0.009	0.194	0.000	3.399	8.678	104.707
2011	0.058	0.010	0.222	0.000	6.862	12.331	249.937
2012	0.056	0.008	0.209	0.000	3.571	9.768	123.065
2013	0.020	0.003	0.144	0.000	3.361	14.079	224.121
2014	0.018	0.003	0.157	0.000	4.842	16.574	340.425
2015	0.016	0.003	0.138	0.000	3.456	15.728	280.399

Note: This table displays tracking error summary statistics by year, aggregative across all ETFs sampled. Tracking error is quoted in percent terms, ie. by using multiplying result of equation (1) by 100. Note that regression results later in this paper use raw TE from equation (1), not the percent transformation used here.

Table 3 shows sample moments for the ETFs categorized by market structure into ‘Developed Market’ and ‘Emerging Market’. 15 ETFs are categorized as developed market, while 5 are categorized as emerging market. I performed a two-sample t-test; the mean TE’s are different at the 1% level. The mean daily tracking error is .052 percent larger for Emerging Market ETFs than Developed Market ETFs.

Table 3 – Sample statistics for ETFs categorized by benchmark’s market structure

	Mean TE	Median TE	Std	Min	Max	Skew	Kurt
Developed Market	0.032	0.005	0.165	0.000	6.862	15.120	312.054
Emerging Market	0.084	0.007	0.268	0.000	4.589	6.505	62.034

Note: This table displays tracking error summary statistics by for ETFs categorized by the benchmark’s market structure. Tracking error is quoted in percent terms, ie. by using multiplying result of equation (1) by 100. Note that regression results later in this paper use raw TE from equation (1), not the percent transformation used here.

Table 4 shows sample moments for the ETFs categorized by geography into if the benchmark indexes are in Europe, Asia, or the Americas. 10 ETFs have European benchmarks, 7 have Asian benchmarks, and 3 have Americas benchmarks. I performed two-sample t-tests to see if the sample mean TE’s are different; all are different at the 1% level. The ETFs with Americas’ benchmarks have the highest mean TEs, followed by Europe and then Asia.

Table 4 – Sample Statistics for ETFs with benchmarks in Europe, Asia and the Americas

	Mean TE	Median TE	Std	Min	Max	Skew	Kurt
Europe	0.038	0.006	0.173	0.000	6.862	15.155	323.047
Asia	0.022	0.005	0.159	0.000	4.589	15.881	301.840
Americas	0.120	0.006	0.306	0.000	3.399	4.597	30.998

Note: This table displays tracking error summary statistics by for ETFs categorized by geographic region. Tracking error is quoted in percent terms, ie. by using multiplying result of equation (1) by 100. Note that regression results later in this paper use raw TE from equation (1), not the percent transformation used here.

5. Methodology

5.5 ETF and Index Returns

First, I calculate daily returns for all 20 ETFs and their corresponding benchmark indexes. In keeping with most of the literature, I use NAV return because NAV return is not affected by premiums or discounts that can cause discrepancies between an ETF's market price and actual returns on assets under management. The daily NAV return of ETF *i* and its corresponding benchmark are expressed by equations (1) and (2) respectively.

$$NR_{it} = (NAV_{it} - NAV_{it-1}) \,/\, (NAV_{it-1}) \tag{1}$$

$$IR_{it} = (Index_{it} - Index_{it-1}) \,/\, (Index_{it-1}) \tag{2}$$

5.6 Tracking Error

I calculate tracking error in the same way as Osterhoff and Kaserer (2016). They note that some previous studies use weekly or monthly data to analyze tracking error, but as ETFs are increasingly being used for short-term investment horizons and hedging strategies, it is relevant to consider daily tracking error. I calculate daily tracking error for ETF *i* at day *t* as the absolute daily deviation of *i*'s NAV return from its corresponding benchmark index's return.

$$TE_{it} = |NR_{it} - IR_{it}| \tag{3}$$

5.7 Variables: Market Structure and Geography Dummy Variables

I'm evaluating if tracking error can be attributed to market structure and geography individually. To do so, I construct market structure and geography dummy variables.

Market Structure Dummy Variable

I am testing to see if tracking error can be attributed to the benchmark's categorization as either 'Emerging Market' (EM) or 'Developed Market' (DM). This variable is '1' if the benchmark of the corresponding ETF's benchmark is classified by MSCI as EM or '0' if classified as DM. Additionally, I include a full set of interaction terms to test the additional explanatory power of each of my control variables for EM benchmarks.

Geography Dummy Variables

I am testing to see if tracking error can be attributed to the benchmark's categorization as 'Europe', 'Asia,' or 'America' based on MSCI categorization. To do so, I construct two dummy variables. One takes on the value of '1' if the benchmark is in Asia and '0' otherwise, while the other takes on a value of '1' if the benchmark is in the Americas and '0' otherwise. I also include a full set of interaction terms to test the additional explanatory power of each of my control variables for different geographies.

5.8 Variables: Control Variables

I construct five factors that I will use in my analysis as control variables. These are the factors used by Shin and Soydemir (2010) with some minor adjustments to prepare the factors for panel analysis. Below I highlight any relevant calculations in constructing the factors, and explain theoretically why each factor should affect tracking error.

Daily Expense Ratio

Expense Ratios serve as a general proxy for management fees of ETFs. Management fees affect tracking errors because they are assessed to an ETF's NAV and thus reduce NAV return relative to the ETFs benchmark index. In general, the higher the management fees the more an ETF is expected to underperform its underlying index (Charupat and Miu 2013). Expense ratios are quoted at an annual rate, but assessed daily and so I transform annual expense ratios into a daily fee for my analysis. As a result, the expense ratio coefficients in my regression results are large, however this is due how small the daily expense ratio tends to be (see control summary statistics following this section).

Scaled Bid-Ask Spread

I construct this factor in the same way as Shin and Soydemier (2010), adjusted so that it is a measure for each ETF each day rather than a daily average across ETFs. To do so, I calculate the average rate of daily price changes by taking the highest trading price (CRSP's ask hi data) of the day and the lowest trading price of the day (CRSP's bid low data) and dividing it by the closing price.

$$Spread_{it} = (AskHi_{it} - BidLow_{it}) / (Close_{it}) \tag{4}$$

This factor serves as a proxy for the underlying liquidity and risk of the benchmark index. The reason this works as a proxy is related to the unique supply and demand dynamics of the ETF market. Unlike stocks, ETFs do not have a fixed supply. As more money is added to an ETF, more shares of the ETF are created. Large institutional investors, called Authorized Participants (AP's), are allowed to place an order with an ETF provider for new shares of an ETF. The AP then sells these shares on the secondary market. APs act as market makers and liquidity providers. They tend to be agnostic about which ETFs they purchase or sell, instead hedging their positions and attempting to make a profit through the bid-ask spread.

However, the less liquid and more risky an ETF's underlying benchmark's market is, the more difficult it will be for an AP to hedge their position. In order to account for this additional risk, APs increase the bid-ask spread. Thus this factor becomes a proxy for underlying market liquidity and risk, and thus a higher coefficient for this factor is expected to have a positive relationship with tracking error.

Log of ETF Trading Volume

I construct the volume factor in the same way as Shin and Soydemir (2010) adjusted to be the log transformed daily trading volume of each ETF rather than an average across ETFs. To do so, I take the log of each ETF's trading volume each day. Volume serves as a proxy for other transaction costs involved in managing an ETF, such as creation and redemption of shares, and may also proxy other ETF liquidity constraints that could affect tracking error. However, the precedent for using volume is relatively less conclusive in the literature with some studies finding a positive correlation between volume and tracking error and other studies finding a negative correlation (Charaput and Miu, 2013).

Dividend Yield

Most ETFs obtain dividends from their constituent stocks, but do no reinvest the dividends. There tends to be a delay between when ETFs obtain dividends and when they aggregate these dividends and pay them to ETF shareholders. All ETFs in my sample declare dividends twice a year. Since ETFs do not reinvest dividends this results in a portion of NAV that is not earning returns, and thus it is expected that dividend yield has a positive relationship with tracking error. To create this as a daily factor I assign daily dividend yield as the dividend yield that the corresponding ETF paid out last time it paid (or declared it would not pay) dividends. As a result, the dividend yield coefficient in my regression results appear small; however, this is due to assigning the semi-annual dividend yield to every date between payments.

Change in the Exchange Rate

Shin and Soydemir (2010) were the first to consider changes in the exchange rate as a possible factor contributing to tracking error, and I construct this factor similar to their method using the following equation:

$$FX_{it} = (ExchRate_{it} - ExchRate_{it-1}) / (ExchRate_{it-1}) \tag{5}$$

ETFs that are listed on US exchanges but composed of stocks traded on foreign exchanges are subject to exchange rate risk. Since NAV returns and Index returns are USD denominated, any exchange rate fluctuations can either mitigate or magnify any existing tracking errors. Tracking error will be mitigated when USD strengthens and magnified when USD weakens. In my calculations, a positive exchange rate factor value signifies USD has strengthened and a negative value that USD has weakened, thus I would expect the coefficient for this factor to be negative (leading to an increase in tracking error when USD weakens and a decrease in tracking error when USD strengthens).

5.9 Summary Statistics of Control Variables

In Table 5, I display summary statistics for the control variable. Each metric was recorded daily for each of the 20 ETFs in the sample over the period 2009 -2015. The mean daily expense ratio is .00002 percent. For comparison, the mean daily tracking error is .045 percent. The mean dividend yield is .39 percent, but is only paid semi-annually. A positive mean FX measure indicates that, on average, USD appreciated over this period.

Table 5 – Sample Statistic of Control Variables

	Mean	Median	Std	Min	Max
Expense	2.115E-05	2.098E-05	2.374E-06	1.900E-05	3.634E-05
Spread	1.409E-02	1.135E-02	1.013E-02	1.685E-03	3.212E-01
Volume	1.393E+01	1.423E+01	1.825E+00	7.090E+00	1.980E+01
Dividend	3.970E-01	3.519E-01	3.562E-01	0.000E+00	2.333E+00
FX	1.053E-02	0.000E+00	6.428E-01	-1.221E+01	9.298E+00

Note: This table displays sample statistics for daily expense ratio (Expense), daily scaled bid-ask spread (Spread), daily log of ETF trading volume (Volume), semi-annual dividend yield (Dividend), and daily change in the exchange rate (FX).

5.10 Panel Regression

I perform two panel analyses. The first model tests to see if tracking error can be attributed to the benchmark's market structure, while the second model tests to see if tracking error can be attributed to the benchmark's geographic region.

The first model can be expressed as follows:

$$TE_{it} = a + b_1{*}EMdummy_{it} + b_2{*}Expense_{it} + b_3{*}Spread_{it} + b_{3*}Volume_{it} + b_4{*}Dividend_{it} + b_5{*}FX_{it} + b_6{*}(EM_{it}{*}Exp_{it}) + b_7{*}(EM{*}Spr_{it}) + b_8{*}(EM{*}Vol_{it}) + b_9(EM{*}Div_{it}) + b_{10}{*}(EM{*}FX_{it}) + error \quad (6)$$

Where *TE* is daily tracking error, *EMdummy* is the emerging market dummy variable, *Expense* is daily expense ratio, *Spread* is the scaled daily bid-ask spread, *Volume* is log of daily volume, *Dividend* is the previous quarter's dividend yield, *FX* is daily change in exchange rate, and *EM*Exp, EM*Spr, EM*Vol, EM*Div,* and *EM*FX* are the respective emerging market interaction terms, all for fund *i* at time *t.*

Additionally, I regress daily tracking error on each independent variable individually to better compare how the effect changes when I add the dummy variable and interaction terms. I also run the regression with and without controlling for interaction terms.

The second model testing geographic regions can be expressed as follows:

$$TE_{it} = a + b_1{*}AsiaDummy_{it} + b_2{*}AmerDummy_{it} + b_3{*}Expense_{it} + b_4{*}Spread_{it} + b_{5*}Volume_{it} + b_6{*}Dividend_{it} + b_7{*}FX_{it} + b_8{*}(Asia{*}Exp_{it}) + b_9{*}(Amer{*}Exp_{it}) + b_{10}{*}(Asia{*}Spr_{it}) + b_{11}(Amer{*}Spr_{it}) + b_{12}{*}(Asia{*}Vol_{it}) + b_{13}{*}(Amer{*}Vol_{it}) + b_{14}{*}(Asia{*}Div_{it}) + b_{15}{*}(Amer{*}Div_{it}) + b_{16}(Asia{*}FX_{it}) + b_{17}{*}(Amer{*}FX_{it}) + error \quad (7)$$

where *TE* is daily tracking error, *AsiaDummy* is the Asia region dummy variable, Amerdummy is the Americas region dummy variable, *Expense* is daily expense ratio, *Spread* is the scaled daily bid-ask spread, *Volume* is log of daily volume, *Dividend* is the previous quarter's dividend yield, *FX* is daily change in exchange rate, and *Asia*Exp, Asia*Spr, Asia*Vol, Asia*Div,* and *Asia*FX* are the respective Asia interaction terms, and *Amer*Exp, Amer*Spr, Amer*Vol, Amer*Div,* and *Amer*FX* are the respective Americas interaction terms, all for fund *i* at time *t.*

I run this regression individually on the variables not yet regressed, and on the geographic region dummy variables with and without controlling for interaction terms.

The panels display heteroscedasticity, time-series autocorrelation and cross-sectional dependence. I therefore use robust standard errors following the method of used by Osterhoff and Kaserer (2016). This involves using the computational method devised by Driscoll and Kraay (1998) and cited in Hoechle (2007). This method generates standard errors which are heteroscedasticity and autocorrelation consistent, as well as consistent for cross-sectional dependence.

6. Empirical Results

(See Tables 6 and 7 on pages following this section.)

First, I looked at each independent variable's effect on tracking error individually. I found every variable with the exception of the change in the exchange rate to be significant. While Shin and Sodeymir (2010) found exchange rate to be a significant factor explaining tracking error, they are the only precedent in the literature to have found this. Additionally, they use a different methodology that aggregates across ETF's rather than using a panel analysis. Finally, their analyses uses the same ETF's but over a different sample period of 2004-2007. In all of the regressions I ran, exchange rate never appears as a significant factor.

Table 6 shows the outcome of my regression focused on market structure. When I ran the regression including the control variables, I found a statistically significant effect for the EM dummy variable. This corroborates my hypothesis that even controlling for effects known to affect tracking error there is additional explanatory power by sorting the sample into ETFs benchmarked to emerging market and developed market equity indexes. The dummy variable shows this effect to be positive; ETFs benchmarked to emerging market indexes have a higher tracking error when compared to ETFs benchmarked to developed market indexes. The dummy variable has a coefficient of .0004, meaning the tracking error for ETFs with emerging market benchmarks is expected to be .04% higher than for ETFs with developed marked benchmarks, even when controlling or known factors affecting tracking error. Additionally, running the regression with interaction terms suggests that the known control variables affect tracking errors for ETFs benchmarked to emerging markets and developed markets differently. The interaction terms for spread, volume, and divided yield are all significant. My results suggest spread, volume, and dividend yield all have an additional positive relationship to tracking error for ETFs benchmarked to emerging market equity indexes. The change in coefficient as a result of including the interaction term is greatest for spread. Since spread proxies the risk and liquidity of the underlying market, it makes theoretical sense that spread would have a greater affect for ETFs benchmarked to emerging markets. To the extent volume is a proxy for transaction costs, it also makes sense that this effect may be greater for ETFs benchmarked to emerging markets, although since the effect of volume is inconclusive across ETF literature this result is not definitive. The theoretical link with dividend yield is also less clear; my findings may present an opportunity for additional research exploring why dividend yield has more of an effect on tracking error for ETFs benchmarked to emerging markets. The r-squared values for all of my regressions are relatively small, suggesting a better model might be found that better explains tracking error.

Table 7 displays my regression results focused on the benchmark's geographic region. When I include the dummy variables and control variables, volume is no longer a significant explanatory variable, but the geographic dummy variables are significant. This collaborates my hypothesis that the benchmark's geographic location can provide additional explanatory power for ETF tracking errors, although there may be material multicollinearity with other variables. My regression shows that the ETFs in my sample benchmarked to the Americas have the largest tracking error, followed by Europe, followed by Asia. ETFs with Americas benchmarks are expected to have 0.05% higher tracking errors than ETFs with Europe benchmarks, while ETFs with Asia benchmarks are expected to have 0.01% lower tracking errors than ETFs benchmarked to Europe, even when controlling for known factors that affect tracking errors. Given my sample of ETFs, it makes intuitive sense that the Americas have the highest tracking error – two of the three benchmark to the volatile markets of Mexico and Brazil. I find it more interesting that the ETF's benchmarked to Europe tend to have a higher tracking error than those benchmarked to Asia. Including interaction terms in the regression does not show any compelling evidence for what may be driving European tracking errors to be higher than Asian tracking errors as only a few of the interaction terms are significant. This suggests that most of what is driving the additional effect on tracking errors in the geographic region control is not accounted for in known control variables. The one variable that is significant for all regions when including interaction terms is volume, suggesting there was multicollinearity between the regional dummy variables and volume when not using interaction terms. The volume effect varies highly based on geographic region: an increase in ETF trading volume has a negative relationship to tracking error for ETF's benchmarked to Europe, roughly no net effect on ETFs benchmarked to Asia, and a positive relationship to ETFs benchmarked to the Americas. Relative to ETFs benchmarked to Europe, ETFs benchmarked to Asia have a 0.01% increase in TE for a unit increase in the log transformed daily trading volume, while ETFs with Americas benchmarks have 0.05% increase in TE for an equal increase in volume. As with the market structure analysis, the theoretical reason for why volume has a the effect it does is not evident, but as it is significant here as well it further the case for additional research looking at volume's effect on tracking error, and noting that the volume effect varies greatly based on geographic region could provide a starting point for additional analysis. Overall, my results collaborate my hypothesis that an ETF's benchmark geographic region has an effect on tracking error, but with the exception of the region-specific volume variable, my results do not paint a clear picture of what is driving this geographic region difference in tracking error, and suggest there are other relevant variables that are not utilized in my model.

Table 6: Regression Results for Analysis of ETF's Benchmark's Market Structure

Variables	Dependent variable TE							
EMdummy	0.0005 (0.0000)***						0.0004 (0.0000)***	-0.0056 (0.0008)***
Expense		85.8836 (10.9741)***					24.1458 (9.2647)***	13.5149 (9.2617)
Spread			0.0318 (0.0036)***				0.0254 (0.0033)***	0.0099 (0.0030)***
Vol				0.0001 (0.0000)***			-0.00002 (0.0000)**	-0.0001 (0.0000)***
Div					0.0012 (0.0001)***		0.0011 (0.0001)***	0.0004 (0.0001)***
FX						-0.0004 (0.0037)	-0.0031 (0.0035)	-0.0021 (0.0039)
EM*Exp								-31.5671 (21.9395)
EM*Spr								0.0365 (0.0056)***
EM*Vol								0.0004 (0.0001)***
EM*Div								0.0013 (0.0002)***
EM*FX								-0.0027 (0.0056)
Constant	0.0003 (0.0000)***	-0.0014 (0.0002)***	0.0000 (0.0000)	0.0005 (0.0001)***	-0.0000 (0.0000)	0.0005 (0.0000)***	-0.0007 (0.0002)***	0.0006 (0.0002)***
R-squared	0.0130	0.0107	0.0267	0.0037	0.0487	0.0000	0.0777	0.1180

Note: This table displays regression results for Tracking Error (TE) on individual variables, as well as TE regressed on the EMdummy variable with control variables and the EM dummy variable on control variables and interaction terms. The tested independent variables are the emerging market dummy variable (EMdummy), daily expense ratio (Expense), scaled bid-ask spread (Spread), log of volume (Vol), dividend yield (Div), and change in exchange rate (FX). Standard errors are in parenthesis. *** is significant at the 1% level. **is significant at the 5% level. R-squared values are also included.

Table 7: Regression Results for Analysis of ETF's Benchmark's Geographic Region

Variable	**Dependent variable TE**		
AsiaDummy	-0.0001 (0.0000)***	-0.0001 (0.0000)***	-0.0010 (0.0004)**
AmerDummy	0.0008 (0.0000)***	0.0005 (0.0001)***	-0.0094 (0.0015)***
Expense		56.468 (10.2121)***	10.7058 (8.6219)
Spread		0.0213 (0.0031)***	0.0121 (0.003)***
Vol		0.0000 (0.0000)	-0.0001 (0.0000)***
Div		0.0010 (0.0001)***	0.0004 (.0001)***
FX		-0.0026 (0.0035)	-0.0027 (0.0044)
Asia*Exp			6.4281 (13.6606)
Amer*Exp			100.5877 (68.9243)
Asia*Spr			-0.0023 (0.0037)
Amer*Spr			0.0295 (0.0037)***
Asia*Vol			0.0001 (0.0000)***
Amer*Vol			0.0005 (0.0001)***
Asia*Div			-0.0001 (0.0001)
Amer*Div			0.0013 (0.0002)***
Asia*FX			-0.0001 (0.0045)
Amer*FX			0.0001 (.0068)
Constant	0.0004 (0.0000)***	-0.0015 (0.00027)***	0.0009 (0.0002)***
R-squared	0.0267	0.0825	0.1286

Note: This table displays regression results for Tracking Error (TE) on the two region dummy variables (AsiaDummy, AmerDummy), as well as TE regressed on the region dummy variables with control variables and the region dummy variables on control variables and interaction terms. *** is significant at the 1% level. **is significant at the 5% level. R-squared values are also included.

7. Conclusion

In this paper, I examine if US-listed ETFs benchmarked to foreign equity indexes have tracking errors that can be explained by underlying market structure and geographic region. I run regressions controlling for factors that have been found to affect ETF tracking errors in previous studies: expense ratio, bid-ask spread, ETF trading volume, dividend yield, and changes in the exchange rate. As postulated, I found that market structure and geographic region have additional explanatory power for tracking errors when controlling for these variables. I found that there is a positive effect on tracking error when ETF's are benchmarked to emerging market indexes. I also found that the effect of spread, volume and dividend yield are all greater when an ETF is benchmarked to an emerging market index.

When testing geographic region, I found a greater marginal positive effect on tracking error for ETF's benchmarked to Europe over Asia, and America over Europe. My sample of ETFs in the Americas was small, so this is a result that could benefit from further study in the future. The interaction terms with geographic region suggest that with the exception of volume, the marginal change in tracking error based on region is largely separate from region-specific effects of the control variables used.

The majority of variables found to affect ETF's tracking error in previous studies have been determined from research that looked primarily at US-listed ETFs benchmarked to domestic indexes. It was my goal in this paper to broaden this discussion to focus on US-listed ETFs that are benchmarked to foreign equity indexes. I found that known variables are not exhaustive for explaining the tracking error of these ETFs, as a benchmark's market structure and geographic region have additional explanatory power for tracking error.

Building off my study, additional research can focus on if there are particular factors that matter for different regions beyond those tested in this study. Furthermore, my model could be extended to account for ways in which variables may not have a linear effect on tracking error, perhaps increasing the explanatory power of my model.

References

Blitz, D. , Huij, J. and Swinkels, L. (2012), “The performance of European index funds and exchange-traded funds”, European Financial Management , Vol. 18 No. 4, pp. 649-662.

Chu, P.K.-K. (2013), “Tracking Errors and Their Determinants: Evidence from Hong Kong Exchange Traded Funds” , University of Macau, Macau.

Elton, Edwin J. and Gruber, Martin J. and Commer, George and Li, Kai, Spiders: “Where are the Bugs.” The Journal of Business, Vol. 75, No. 3, July 2002. Available at SSRN: https://ssrn.com/abstract=316323

Hoechle, D. (2007), “Robust standard errors for panel regressions with cross-sectional dependence”, *The Stata Journal* , Vol. 7 No. 3, pp. 281-312.

Johnson, W.F. (2009), “Tracking errors of exchange traded funds”, Journal of Asset Management, Vol. 10 No. 4, pp. 253-262.

Milonas, N.T. and Rompotis, G.G. (2006), “Investigating European ETFs: the case of the Swiss exchange traded funds”, working paper for Annual Conference of HFAA, University of Athens, Thessaloniki

Narat Charupat, Peter Miu, (2013) "Recent developments in exchange-traded fund literature: Pricing efficiency, tracking ability, and effects on underlying securities", Managerial Finance, Vol. 39 Issue: 5, pp.427-443, doi: 10.1108/03074351311313816

Rompotis, G.G. (2011), “Predictable patterns in ETFs’ return and tracking error”, Studies in Economics and Finance , Vol. 28 No. 1, pp. 14-35.

Rompotis, G.G. (2012), “The German exchange traded funds”, The IUP Journal of Applied Finance , Vol. 18 No. 4, pp. 62-82.li

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