The fundamental analysis of the capital investment in exchange-traded fund

Racu Alina-Cristina[1*]

The Bucharest University of Economic Studies, Romania

Abstract

The popularity increase of ETFs requires the deepening of some specific aspects, in order to make a good investment. In this paper I present the concept of ETF's performance and the most important factors that influence it. I have built three studies in Eviews and analyzed their results to see what aspects should be taken into consideration when an investor decides to buy ETF's shares.

Keywords: Tracking error, passive strategy, liquidity, market capitalization, volume, risk.

JEL Classification: G11, G23

Introduction

The purpose of this paper is to investigate the relevant and useful indicators in making an investment decision in an exchange-traded fund, but also the factors that may impact the performance of this financial product.

Providing help for investors can be relevant to any capital market, because capital infusions could help the economy of a country, and in times of crisis it can even recover it. Unfortunately, the Romanian capital market does not have a developed level of investments compared to other markets, but as a result of new listings, such as Romgaz on the London Stock Exchange or the DIGI listing, the situation has begun to recover. Thus, Romania has begun to be in the institutional investor's analyzes.

Since there are currently many studies about investments in an exchange-traded fund, I find it relevant and important to study this aspect on the Romanian market as well. In addition, it is interesting to analyze whether the issues mentioned by the literature, that have an impact on the performance of an ETF, will be reflected in a market that is still under-developed, such as ours. The first ETF was listed on the Bucharest Stock Exchange in 2012, aiming to replicate the BET index, which is composed of the most liquid companies.

[1] Autor de contact, **Racu Alina-Cristina** – racu_alinacristina@yahoo.com

The paper has three chapters in its structure. The first chapter is the basis of the performance analysis theories of an ETF. The second chapter describes the ETF market in Romania, and the third chapter highlights the replication strategy that the ETF adopts on the Romanian market. The last chapter presents the performance of the ETF on the BVB and its determinants.

1. Review of the literature on the performance of ETFs

Since the main purpose of exchange-traded passive funds is to replicate the performance of an index they are pursuing, the difference between the ETF's return and that of the index at its base, called in the tracking error or error literature replication, can be used to evaluate its performance. Tracking error occurs in the performance of ETFs, as managing passive portfolios will naturally encounter difficulties. Theoretically, this type of fund requires managers to hold the same securities, in the same proportion as the index they replicate, a strategy known as full replication. In fact, fund management will encounter considerable difficulties in replicating the target index, as it has a mathematical structure that does not take into account market friction. Thus, funds will have to trade the securities that make up the index for replication, which entails transaction costs and price pressure. In addition, the literature offers various studies aimed at illustrating the various determinants of replication error.

ETF market specialist literature has increased in recent years, reflecting the importance and interest of investors for this type of financial product. Before making the investment decision in an ETF, any investor is curious about its performance or the values that the replication error records. Although there are many articles on this subject, there is still no consensus on a definition of this variable. In this subchapter I will present the main ideas that specialty studies show on this indicator, but also the factors that influence the replication error, abbreviated TE.

The paper should be the result of a research in the area, corresponding to the specific topic of the journals' issue. The structure of the paper should be clear and well emphasized by titles and subtitles placed in a logical sequence, according to the writing methodology of the scientific papers.

For any ETF that attempts to replicate the performance of a stock index, a decisive quality is its ability to track the returns of that index. The article published by Osterhoff (2016) aims to highlight the relationship between market liquidity and ETF industry performance, analyzing the impact of liquidity on the tracking error (TE) indicator. Among the main factors that influence the performance of an ETF we have the expenditure ratio, changes in the composition of the tracked index, or changes in the replication strategy, dividend payments.

Regarding the impact of market liquidity on TE, previous studies have focused on variables such as bid-ask volume and spread. It was found that for an ETF that follows the DAX index, between 2001-2006, the traded volume of the ETF has a significant negative relationship with the tracking error of the ETF. This means that, on an average, TE tends to decrease with the increase in the volume traded. The results of many studies suggest a positive relationship between bid-ask spread and TE. Milonas and Rompotis (2005) attest

this direct relationship, and Rompotis (2012) comes to similar conclusions for the German ETF market.

The main determinants mentioned in many studies are:

The annual expenditure rate that is determined as the ratio between the annual operating expenditure and the average value of the assets managed by the ETF. In many studies related to the performance of an ETF, this variable has been put in the foreground, being one of the most accepted factors in the existence of the return differences between the ETF's profitability and the index.

The size of ETF's market value may be another factor that influences the performance of the ETF. Thus, between the magnitude and the quantified performance in the form of the TE indicator, there should be an opposite relationship. In other words, a higher ETF, will be able to replicate its index as closely as possible. The negative relationship can also be explained by the fact that a higher amount of money invested can put more pressure on managers in order to perform better. In addition, large funds can also be a good magnet for professional managers who are able to perform at the target level.

The risk, normally calculated as the standard deviation of ETF's returns, may be another factor that literature has analyzed in order to explain the performance of this kind of funds. A higher risk also involves a higher volatility, with the effect of increasing tracing error. For a fund manager, it will be harder to replicate the index in the situation when volatility has increased.

Liquidity is generally quantified according to the volume traded. If ETF's market liquidity increases, cash inflows will increase, trading costs will decrease, and finally TE will decrease.

Bull (bear) periods corresponding to general growth periods (decreases) in market prices may directly affect TE, according to Wong and Shum (2010). They have analyzed 15 ETFs from 7 countries, using the daily closing prices of both funds and indices between 1999-2007, aiming to achieve a risk-return comparison between these two types of markets. For the bull market, the standard deviation was generally higher compared with the bear market, demonstrating a greater volatility in the bull markets. In terms of profitability, the results were the same, with higher values for bull market. For this reason, bull periods may be characterized by higher tracking errors than bear periods.

There are different methods to calculate the tracking error indicator. I have identified four different methods, all of them based on return differences between the ETF and the underlying index.

The simple difference between the returns (Wong and Shum, 2010):

$$TE_1 = R_{ETF,t} - R_{I,t} \quad (1)$$

Where $R_{ETF,t}$ represents the daily rentability of ETF, and $R_{I,t}$ is the daily rentability of the index followed be he ETF.

The average absolute difference in returns (Frino and Gallagher, 2001):

$$TE_2 = \frac{\sum_{t=1}^{n} |R_{ETF,t} - R_{Indice,t}|}{n} \quad (2)$$

Where *n* represents the number of observations of the analyzed period.

The standard deviation of the returns difference (Frino and Gallagher, 2001):

$$TE_3 = \sqrt{\frac{1}{N-1} * \sum_{t=1}^{n}(R_{ETF,t} - R_{Indice,t})^2} \quad (3)$$

The standard error of the regression equation, where the returns of the ETF are regressed on the returns on the benchmark index (Rompotis, 2006):

$$R_{ETF} = \alpha_i + \beta_i R_{BET} + \varepsilon \quad (4)$$

Where: R_{ETF}, R_{BET} is the daily return expressed as a percentage of the ETF or the BET index of day i.

2. ETF market in Romania

ETF BET Tradeville, under the ticker TVBETETF, is the only exchange-traded fund listed on Bucharest Stock Exchange, since 2012. It is a passive ETF and for this reason its purpose is to track and replicate the return BET. This means that the portfolio managers of the TVBETETF must invest in the entire BET index.

The share holdings of the Romanian ETF compared to BET are as follows:

Table no. 1. Composition of TVBETETF versus index BET

Ticker	*Company*	*Percentage (%) in BET*	*Percentage (%) inTVBETETF*
FP	FONDUL PROPRIETATEA	21.63%	22.65%
TLV	BANCA TRANSILVANIA S.A.	20.08%	18.24%
SNP	OMV PETROM S.A.	15.43%	14.96%
SNG	S.N.G.N. ROMGAZ S.A.	11.65%	11.24%
BRD	BRD - GROUPE SOCIETE GENERALE S.A.	10.51%	10.23%
TGN	S.N.T.G.N. TRANSGAZ S.A.	6.71%	6.54%
EL	SOCIETATEA ENERGETICA ELECTRICA S.A.	4.16%	4.05%
DIGI	Digi Communications N.V.	3.59%	3.53%
TEL	C.N.T.E.E. TRANSELECTRICA	1.73%	2.01%
SNN	S.N. NUCLEARELECTRICA S.A.	1.46%	1.43%
COTE	CONPET SA	1.27%	1.29%
M	MedLife S.A.	1.22%	0.93%
BVB	BURSA DE VALORI BUCURESTI SA	0.56%	0.55%

Source: Vanguard and BVB sites

Regarding the structure of the sectors of activity, the banking sector has the largest weight, almost 30%, followed by the energy sector. The financial sector weights 22%, and in the end the lowest weight is 20% of the utilities sector.

The public information provided by the site that manages the ETF, Vanguard Asset Management, assures investors that TVBETETF invests only in BET index shares, not in other derivatives, bonds or other funds. It is noted that the BET index is not adjusted for dividends, which forces the fund to reinvest the dividends received from the companies in the portfolio without the composition error being above 15%.

The fund is accessible to both institutional investors in the primary market and to the individual investors, on the secondary market. In the primary market, institutional investors can create or redeem block of ETF shares. A block of titles has 10,000 shares.

3. The replication Strategy of the ETF listed on the BVB

I will continue to analyze how well the returns of the Romanian ETF follow the returns of the target index, BET.

The first part of the case study is based on articles that have analyzed the ETF's full replication strategy, such as Purohit and Malhotra (2005) or Kayali and Unal (2009), using a simple linear regression between the returns of the ETF (dependent variable) and the returns of the target index (independent variable).

I have analyzed the daily returns of the BET index and the returns of ETF for TVBETETF. Returns were calculated using the daily closing prices for the two indices obtained from Bloomberg. The period underlying the analysis of the two returns is represented by the last 3 trading years, namely January 2015-December 2017. For the calculation of these returns, the formula used is:

$$R_i = \frac{PT_i - PT_{i-1}}{PT_{i-1}} * 100 \quad (5)$$

Where R_i is the daily return of the ETF or the BET index of day I expressed as a percentage and PT_i represents the closing price for day i or i-1 of the ETF or the BET index .

In order to analyse the replication strategy I used the linear simple regression:

$$R_{ETF} = \alpha_i + \beta_i R_{BET} + \varepsilon \quad (6)$$

Where: R_{ETF} represents ETF's return, R_{BET} indicates the returns of the BET index and ε is the error term. In the equation, α (alpha) is the free term, the return that TVBETETF could obtain constantly, independent of the index's return. Of course, in this situation, this is impossible because the ETF adopts a passive strategy. β (beta) reflects the systematic risk of the ETF and, at the same time, is used to measure the degree of aggressiveness of the management strategy. In other words, if the equation returns a β = 1, we could talk about a full replication strategy of the BET index and an identical structure of ETF portfolio as the BET index. On the other hand, if β is less than 1, this could reflect a selective replication strategy, fund managers deviating from the passive strategy, according to Rompotis (2005). Of course, there might be other factors that can influence this aspect.

The results of the linear simple regression are:

Table no. 2. Simple regression between the returns of the BET index and those of the ETF

TVBETETF			
	Constanta	**Beta**	**R-square**
Coeficient	0,0002	0,8398	0,5200
T-statistic	0,8310	28,5204	
P-value	0,4062	0,0000	

Source: own processing in E-views

Analyzing the value of R-square, the model is explained in proportion of 52% by the BET return, and it can be said that ETF's returns succeeds in pursuing profitability that corresponds to the stock index, but with a slight deviation. On the other hand, considering that TVBETETF is 5 years old and in its composition has the most liquid shares on the Romanian capital market, the value of R-square is relatively small compared to other studies from the emerging markets. For example, Kayali and Unal (2009) analyzed two ETFs listed on the Istanbul capital market, their capitalization being of two types, large and small. For the ETF with the large-capitalization, R-square = 90% vs. 75% (small capitalization). Sethi (2016) analyzes the performance of 10 listed ETFs on the Indian stock market and just one of them has a R-square bigger than 50%, concluding that ETFs fail to implement a full replication strategy as it can be seen in the case of TVBETETF. Purohit and Malhotra (2005) reach the same conclusion for passive ETFs as a result of analyzing R-squared statistics. A value different from 1 means a deviation from the full replication strategy.

Regarding the coefficients, for a significance level of 5%, it can be seen that beta has a Probability= 0.000< 5%, which means it is statistically significant. On the other hand, the p-value for alpha= 0.4062 > 5%, which means that it is not statistically significant. The interpretation of coefficients β and α is as follows: when the BET index increases by 1 percentage point, the ETF increases by 0.83976 percentage points. When the market is inactive, the return of the ETF will be 0.000194, and the return of the BET would take the unlikely value of 0.

When comparing with the studies mentioned above, the results for the α and β coefficients are in accordance with them. β is different and less than 1 in all studies, which is why we cannot talk about a full replication strategy, but we can see an evolution of returns of ETFs very close to the evolution of the followed index. α is also statistically insignificant in all studies because the ETF is trying to adopt a passive strategy and the manager of an exchange-traded fund cannot obtain a constant and independent profitability from that of the target index.

The conclusion of this chapter is similar to that of studies that analyzed poorly developed capital markets, like Romania, and this is that the managers of an exchange-traded fund fail to perfectly replicate the returns of the index they are pursuing. Thus, there is also a first characteristic of passive ETFs in this type of market, namely that the investment objectives of these financial products are not achieved, which gives rise to the tracking error.

As a result of these results, the second part of the case study requires an analysis of the ETF tracking error indicator and the factors that have an impact on it.

3. Determinants of tracking error

Following the quantification of the replication error indicator using the formulas mentioned in Chapter 1, I decided to run two multiple regressions having this statistic as a dependent variable, the purpose being to determine the variables that influence it. I will only consider the last two methods of estimating the tracking error indicator as relevant, as they have proven to give similar results.

The independent variables I will use in the regression were also mentioned in the knowledge stage. They are the size of the ETF, the risk and the liquidity.

The equation of the multiple regression is:

$$TE_i = \propto_0 + \beta_0 LnMărime\_ETF + \beta_1 LnVolum + \beta_2 Risc + \beta_3 AbsPremium + \varepsilon_i \quad (7)$$

I have built a model for each dependent variable:

- TE2= the average absolute difference in returns

- TE3= the standard deviation of the returns difference

The independent variables were:

LnMărime_ETF represents log of the market value of ETF, the data source being represented by the Vanguard site and the frequency is a daily one.

AbsPremium is the absolute premium value expressed as a percentage, the formula is as it follows:

$$PR_i = \frac{PT_i - NAV_i}{NAV_i} * 100$$

Where: PR_i = premium rate on day i

PT_i = the trading price of day i

NAV_i= the net asset value of day i

The variables have a daily frequency and are taken from Bloomerg, except the NAV, which is from Vanguard site.

LnVolum was calculated as the natural logarithm of the volume, being extracted from the Bloomberg site and representing the number of transactions made in a day.

Risc variable is calculated using the ETF's standard deviation for the past 22 days, which is based on a trading month.

For all four independent variables with daily frequency, the period is 3 years, January 2015-December 2017.

In order to determine whether the factors analysed affect the performance of the ETF in a relevant way, observed by a smaller tracking error indicator, I have established the following hypotheses according to the literature:

Hypothesis 1: The size of the ETF is in a negative relationship with the tracking error indicator, according to Chu (2011).

Hypothesis 2: The trading volume would influence indirectly the tracking error indicator as described by Chu (2011).

Hypothesis 3: The absolute premium values are expressed in percentages which are in a positive relationship with the tracking error variable as a result of some liquidity issues highlighted by Rompotis (2012).

Hypothesis 4: Between risk and the tracking error indicator there should be a direct relationship, Rompotis (2012).

The first step in running multiple regressions for the period 2015-2017 was to test if the daily data is stationary for the four independent variables. In this regard, the Augmented Dickey-Fuller test, the Akaike criterion, was applied in E-views, based on the assumptions:

H_0: there is a unit root (the series is non-stationary)

H_1: there is no single root

According to Table no. 3, the Risc, LnVolum, AbsPremium variables result in a significant probability for a 5% level, which is why I can conclude that the series are stationary, and another test is not necessary. The exception is the variable LnMariemeETF, for which the null hypothesis is accepted, Prob (LnMeasureFET)> 5%. In this case, it was necessary to apply the first difference on the data series, D_LnMărimeETF, to have a stationary series.

Table no. 3. The results of the ADF-AIC test

	Risc	*LnMărimeETF*	*LnVolum*	*AbsPremium*
	Prob. ADF-CRITERIUL AIC			
Intercept	0,0023	0,8932	0,0009	0,0000
Trend and intercept	0,0006	0,6947	0,0021	0,0000
	D_LnMărimeETF			
	Prob. ADF-CRITERIUL AIC			
Intercept	0,0000			
Trend and intercept	0,0000			

Source: own processing in E-views

To be sure of the bonity of the E-views models, I have verified that the average rate of errors is 0 and I applied the homoscedasticity test for the errors, the test of the error correlation and normality test.

The final results from the hypotheses analysed and the specific tests were appropriate for the following models:

Table no. 4. Eviews Model 1 Final Results

	Coef.	*Erori std.*	*T-st.*	*Prob.*	*R-squared*	*Durbin-Watson*	*Prob.(F-statistic)*
D_LnMărime_ETF	-0,0036	0,0023	-1,5331	0,096			
LnVolum	-0,0002	0,0000	-9,9513	0,000	39%	1,882	0,000
Risc	0,2453	0,0105	23,3328	0,000			
AbsPremium	0,0422	0,0105	4,0338	0,000			

Source: own processing in E-views

Table no. 5. Eviews Model 2 Final Results

	Coef.	*Erori std.*	*T-st.*	*Prob.*	*R-squared*	*Durbin-Watson*	*Prob.(F-statistic)*
LnVolum	-0,0002	0,0000	-10,8991	0,000	34%	1,911	0,000
Risc	0,2870	0,0141	20,4024	0,000			
AbsPremium	0,0524	0,0134	3,9092	0,000			

Source: own processing in E-views

For model 1, all independent variables are statistically significant. For model 2, the size of the ETF is statistically insignificant.

Regarding the size of the ETF, there is a negative relationship between this variable and the tracking error, only for model 1. This can be explained by the fact that, with the increase of a fund, its ability to track the target index increases. In addition, the more a fund attracts more investors, the responsibility gets bigger and the ability of managers to achieve their strategy will increase. Thus, the first hypothesis is respected only for the TE2 calculated as the average absolute difference in returns.

The relationship between tracking error and volume shows that the increase in liquidity on the Romanian market positively contributes to the diminishing of the variations between the ETF's and BET return. Transactions are more cost-effective, as the spread diminishes with the cash inflows in the fund, which results in a decrease in the tracking error variable. In addition, transactions will be made even faster, as a result of a liquid market. It is well known that Romania is still a frontier market, characterized by the lack of liquidity, but the statistic regressions, although reflecting a relatively small coefficient, brings statistically significant results of this aspect. For both models, with TE2 and TE3 as dependent variables, a log volume unit determines the decrease in tracking error by -0,0002 percentage points. Thus, the first conclusion is that in the Romanian ETF market, what influences liquidity, may be the decision of investors to act in terms of information about the volume traded.

In addition to the liquidity aspect, it can be explained by the absolute premium variable expressed as a percentage, which shows for the first model that, with a change of one percentage point in the premium absolute value, the ETF replication error increases by 0.0422 percentage points. For the second model, TE3 increases by 0.0524 percentage points to a change of one percentage point in the absolute premium value. Absolute premium percentages were considered because any deviation from the NAV is considered

ineffectiveness according to Rompotis. This inefficiency may result in a decrease in the market liquidity. Even if these price deviations are temporary or permanent, they will lead to inefficient and difficult transactions, and eventually the tracking error may increase.

Finally, the risk variable is significant and explains most of the tracking error evolution. A changed unit of the return's ETF standard deviation increase tracking error by 0.245 percentage points and by 0.287 percentage points for TE3. In other words, there is a positive relationship between TE and risk. The increase in risk leads to an increase in market volatility, which will negatively influence the transactions made by fund managers. These will become more difficult to achieve, and hence the impossibility of following closely the return of the target index.

Conclusions

Regarding the ability of the ETF to track the target index, I have applied a simple regression, using as a dependent variable the daily returns of ETF and as an independent variable the daily returns of BET. The results showed a close relationship between the two indicators, with a Beta = 0.83. The result was consistent with other studies for poorly developed financial markets. Some studies have showed that a Beta different than 1 can mean another replication strategy used by managers. In my case, this idea cannot be fully accepted because the shares in the composition of ETF are very similar to those of BET, the data being public. In addition, a stronger explanation of the specialists is that the strategy deviates as a result of the influence of several factors.

In the second part of the case study I have identified possible factors which could influence tracking error indicator. It is desirable that this indicator should be as small as possible so that the ETF performs very close to the target, and their returns to be similar. The results have showed that the size of the ETF, the market liquidity expressed as the volume traded and the absolute value premiums, as well as ETF's risk, are statistically significant and could have an impact on the performance of TVBETETF. I have noticed that an increase in the ETF's size can improve its performance, due to the fact that there is a negative relation between tracing error and size. Also, traded volume influences tracking error negatively, because an improvement in market liquidity resulted in lower tracking errors. Absolute premium percentage and risk values are directly related to the tracking error. Thus, performance decreases if the ETF's risk increases or ETF pricing inefficiencies appear in relation to the net asset value.

That being said, it would be relevant for an investor to take into consideration all these types of variables in order to identify the best ETF. Although the present study has analyzed just one ETF, the results are in line with the literature and can be applied to a sample of several exchange traded funds.

Bibliography

[1] Blake, C., R., Elton, E. J., & Gruber, M. J., 1993. The performance of bond mutual funds. *Journal of business*, 371-403.

[2] Chu, P.K.-K., 2011. *Study on the tracking errors and their determinants: evidence from Hong Kong exchange traded funds*. Applied Financial Economics, 21(5), 309-315.

[3] Delcoure, N., & Zhong, M., 2007. *On the premiums of iShares*. Journal of Empirical Finance, 14(2), 168-195.

[4] Frino, A., & Gallagher, D., 2001. *Tracking S&P 500 index funds*. Journal of portfolio Management, 28(1).

[5] Kayali, M., & Unal, S., 2009. *Tracking Accuracy of Large and Small-CAP ETFS: an Empirical Analysis Of The Istanbul Stock Exchange.* South East European Journal of Economics and Business, 4(1), 107-112.

[6] Markowitz, H., 1952. *Portfolio selection*. The journal of finance, 7(1), 77-91.

[7] Malhotra, N., & Purohit, H., 2015. *Pricing Efficiency and Performance of Exchange Traded Funds in India.* Journal of Applied Finance, 21(3), 16-35.

[8] Milonas, N. T., & Rompotis, G.G., 2006. *Investigating European ETFs: The case of the Swiss exchange traded funds.* In Conference of HFAA in Thessaloniki, Greece.

[9] Osterhoff, F., & Kaserer, C., 2016. *Determinants of tracking error in German ETFs–the role of market liquidity.* Managerial Finance, 42(5), 417-437.

[10]Pástor, Ľ., & Stambaugh, R.F., 2003. *Liquidity risk and expected stock returns.* Journal of Political economy, 111(3), 642-685.

[11] Rompotis, G.G., 2005. *An empirical comparing investigation on exchange traded funds and index funds performance.* Available through: papers.ssrn.com.

[12] Rompotis, G.G., 2010. *Active versus passive ETFs: An investigation of bid-Ask spread.* IUP Journal of Applied Finance, 16(3), 5.

[13] Rompotis, G.G., 2012. *The german exchange traded funds.* IUP Journal of Applied Finance, 18(4), 62.

[14] Sethi, N., 2016. *Examining the Performance of Indian Exchange Traded Funds (ETFs).* Eurasian Journal of Business and Economics, 9(18), 61-79.

[15] Wong, K. H., & Shum, W. C., 2010. *Exchange-traded funds in bullish and bearish markets.* Applied Economics Letters, 17(16), 1615-1624.