Instituto Superior de Economia e Gestão

UNIVERSIDADE TÉCNICA DE LISBOA

DESDE 1911

MESTRADO EM FINANÇAS

TRABALHO FINAL DE MESTRADO

DISSERTAÇÃO

WEAK-FORM EFFICIENCY OF EQUITY ENERGY EXCHANGE TRADED FUNDS

DIOGO OOM DE SOUSA TOVAR JALLES

ORIENTAÇÃO:

RAQUEL M. GASPAR

JULHO - 2012

Weak-Form Efficiency of Equity Energy Exchange Traded Funds

Diogo Oom de Sousa Tovar Jalles*

Technical University of Lisbon

Abstract

The main purpose of this final master dissertation is to assess the weak-form efficiency of Equity Energy Exchange Traded Funds (ETF). For the period of 2008 – 2012 we selected all equity energy ETFs traded in the U.S. stock market with inception date before 2008. The sample selected, is composed by 26 ETFs and we make use of full daily historical data and apply various tests: autocorrelation tests, runs test, unit roots structural breaks tests, panel unit roots analysis and variance ratio tests. These tests allow us to conclude that equity energy ETFs price changes follow a random walk, and so the weak-form efficiency hypothesis is not rejected.

Keywords: ETFs; Equity Energy Funds; Weak-Form Efficiency; EMH; Unit Roots; Structural Breaks; Panel Analysis; Variance Ratio

*Finance Master Student, Technical University of Lisbon, Economic School of Business.

I am grateful to Raquel M. Gaspar for advice and suggestions and to my family and friends for helpful comments.

This work is dedicated to my parents who have always inspired a lifelong pursuit of learning and knowledge.

Contents

List of Tables

List of Figures

1 Introduction

The efficiency concept is a general term and it has several distinct meanings. In capital markets one can distinguish between three types of efficiency: operational efficiency, informational efficiency and allocation efficiency. The first refers to the level of cost of carrying out transactions in capital markets. Informational efficiency, or also commonly referred as the efficient market hypothesis, now on referred as EMH, is one in which prices always fully reflect available information. So, theoretically, there is no place to technical analysis [1] and it is impossible to obtain abnormal returns by trading on the basis of information the market already knows. Lastly, allocation efficiency relates to the optimal distribution of capital among individuals in the economy.

Three informational efficiency categories are proposed by Fama (1970) depending on the nature of the information subset of interest, namely weak-form efficiency, semi-strong efficiency and strong efficiency. In the former, all information contained in historical prices is fully reflected at the current security prices. This means that successive price changes are assumed to be independent and follow a random walk. Semi-strong form efficiency, in which the concern is whether prices efficiency adjust to other information that is obviously publicly available and is relevant to forming expectations of future prices (e.g., announcements of annual earnings, stock splits, etc.). Finally, strong-form efficiency is concerned with whether a given set of investors or groups have monopolistic access to any information relevant for price formation. In addition, if the market is not efficient at the weak-form, it will never be at the semi-strong form and therefore at the strong form efficiency.

Fama (1991) redefine the market efficiency literature by changing the name of the three alternatives efficiency categories. The weak-form tests now cover the area

[1] Technical analysis, or chartism, is the practice of identifying recurring patterns in historical prices in order to forecast future price trends. The technique relies on the idea that prices "move in trends which are determined by the changing attitudes of investors in reaction to a variety of economic, monetary, political and psychological forces" (Pring 1991, pp.2) and that these trends are, therefore, predictable to some extent.

of tests for return predictability, which are not only concerned with the forecast of historical returns but also with variables like dividend yields, interest rates, size effects and seasonal effects. For the semi-strong and strong form the author only proposed changes in titles. Event studies, is now the name for semi-strong form tests of the price adjustment to public announcements. Instead of strong-form tests the author suggested the title tests for private information.

The Exchange-Traded Fund (henceforth, ETF) industry is a relatively recent innovation in financial markets that has attracted larger attention as it strongly grew in size in a relatively short period of time. The investors were attracted by these products due their simplicity, low-cost diversification benefits and the intraday trading ability. ETFs were introduced on U.S. and Canadian exchanges in the early $90s$. Deville (2006) states that the first ETF, the Toronto Index Participation Fund (TIPs), was created in Canada in 1989 and started trading in 1990 on the Toronto Stock Exchange (TSX). Often referred as the first ETF, S&P Depository Receipts (SPDRs, ticker SPY) was launched on the American Stock Exchange (AMEX) in 1993. A few years later, in 2000, ETFs come to Europe.

The industry development allowed investors to access a wide range of ETFs and improved accessibility to several asset classes. ETFs are listed on a stock exchange and they aim to replicate the performance of their benchmark indexes as closely as possible. They are a class of mutual funds but differ fundamentally from traditional ones, which do not trade midday. Traditional mutual funds take orders during trading hours, but the transactions only occur at the close of the market. The price results from the sum of the daily closing prices of all the stocks included in the fund. The same is not true for ETFs, which are traded instantaneously all day long and allow an investor to lock in a price for the underlying stocks immediately. Many authors, including Kosev and Williams (2011), assert that annual management fees and brokerage costs are typically lower compared to the average mutual fund fees. Tax effects are also not to be ignored, and ETFs perform well after-tax. Therefore, ETFs are an economical purchase, making them especially attractive for those who want to take an advantage for short-term strategy.

In sum, the ETFs present a simple, low-cost means of gaining a diversified portfolio and the capacity for intraday trading. They also offer investors the ability to invest in a range of asset classes which may otherwise be inaccessible or prohibitively expensive, including emerging market equities and commodities. However, investment in these securities is not without risk, generally concerns liquidity risk, and the industry's rapid growth has attracted increased attention from regulators.

ETF assets are growing faster, see Figure 1, than traditional mutual funds and are also becoming favorites of hedge funds and day traders. Kosev and Williams

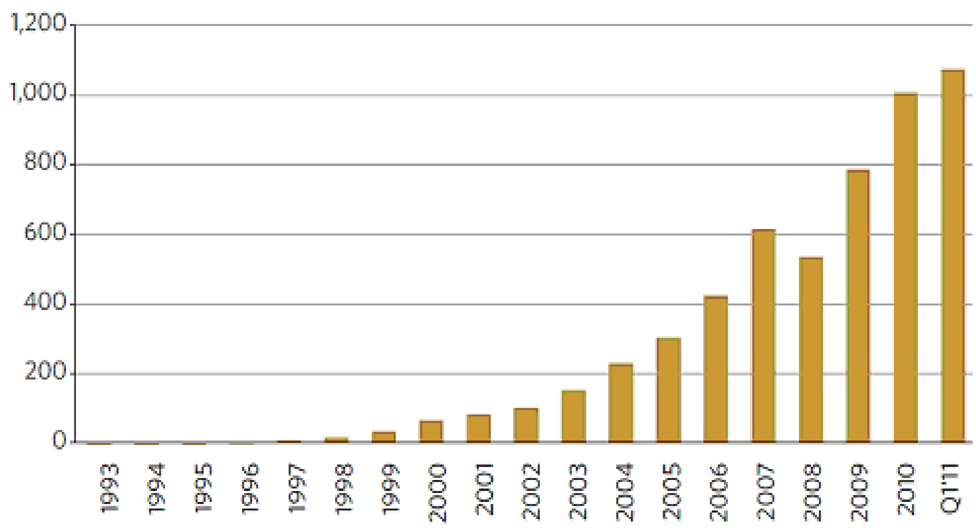


Figure 1: U.S. ETF Assets 1993 - Q1 2011 (billion dollars).
Source: Strategic Insight Simfund MF

(2011) found around 2700 ETFs globally in which 70 per cent of total ETF assets are domiciled in the United States. The majority of ETFs track equity indices from a specific country or region. Of these, Kosev and Williams (2011) says that the number and size of ETFs that invest in emerging market equities have grown strongly and in some cases, these ETF are the only way that allow investors to invest in these markets. The growth of equity ETFs is also marked, namely "in specific market sectors, such as financial, technology or energy indices and stylespecific investments, such as growth or small-cap stocks" as mentioned by Kosev and Williams (2011).

We resort to various different tests to assess the weak-form efficiency of Equity Energy ETFs. The tests performed encompass autocorrelation tests, runs test, unit roots structural breaks tests, panel unit roots analysis and variance ratio tests. These tests are performed on a sample of 26 equity energy ETFs traded in U.S. stock market for the period of $2008 - 2012$ and it covers the full historical daily net asset values (NAV) returns of the selected ETFs.

The remaining of the text is organized as follows. The next section gives a review of literature which is followed by the data. In the two final sections, the discussion of the methodology and empirical results are presented and conclusions are discussed.

2 Literature Review

The EMH has been the major issue of finance since the early 1970s and is one of the most contested and well-studied subject in all the social sciences.

The work of Samuelson (1965) provide the first economic argument for efficient markets developing the theoretical Random Walk Model (RWM) in which he refutes the idea of any price systematic pattern. The author argues that there is no way of making expected profit by extrapolating past changes into futures price. In statistical terms the theory says that successive price changes are independent and identically distributed random variables which imply that no profitable investment trading strategy can be derived based on past price. Moreover, Fama (1965) also show some empirical evidences of independence and randomness of prices, as well the non normality of price changes. After, Fama (1970), as described in the previous section, distinguish three informational efficiency categories - weak, semi-strong and strong - based on the given information set.

After these two firsts' studies, a wide range of works regarding the EMH emerged.

Nascimento (2007), states that until the late 70s and the early 80s most of empirical works are consistent with the EMH. However, from the 80s emerged several studies in which, making use of advanced econometric techniques, questioned the EMH.

Works by Sharpe (1966), Hawawini and Michel (1984), Hudson et al. (1996) and, more recently, by Evans (2006) examined the U.S. and European stock market price behaviors and found support to the fact that historical prices do not indicate fluctuations on future prices, they follow a random walk and, therefore, markets are efficient at the weak level.

As previously mentioned, the EMH is widely questioned in financial economy, therefore, many chartist theories assume that the past behavior of a security's price is rich in information concerning its future behavior. History repeats itself in that "patterns" of past price behavior will tend to recur in the future, so this can be used as a way to increase expected gains and outperform the market. Malkiel (2003) believes that stock prices are partially predictable. He also says that markets can be efficient even if they sometimes make errors in valuation, i.e. the market pricing is not always perfect. Cootner (1964), Lo and MacKinlay (1999) and Lo et al.(2000), supported Malkiel (2003) findings that short-run serial correlation are not zero so there exists short-run momentum in stock prices and recognizing some patterns in stock pricing series. Malkiel (2003) also mentions Shiller (2000) and his research on individuals psychological behavior patterns relating it with the short-term momentum, in which investors tend to under or overreact to important news announcement. Some market anomalies concerning seasonal patterns are also pointed by Malkiel (2003), such as the January effect[2] and the Mondays effect[3].

One additional aspect that questions the EMH is the professional money manager ability to outperform the market. Despite of some authors like Carhart (1997) finds

[2]The author alludes to Keim's (1983) and Haugen and Lakonishok's (1988) studies; January effect is an anomaly where financial security prices increase in the month of January creating outperform return opportunity for investors.

[3]The author refers French (1980); Monday effect is a anomaly in which stock prices have higher returns on Mondays.

no significance evidence of fund managers skills to obtain high returns, others as Jensen (1969) and Elton et al. (1993) analyses managers selection strategies and conclude that they can reach abnormal profitable returns.

There are several aspects in which this work contributes to the EMH literature:

- First data focus only on the U.S. exchange traded fund energy sector and covers very recent years which have not been studied yet. Rompotis (2011) use data until 2010 but for a sample of U.S. ETFs invested either in local broad market, in international capital market indexes and in sector indexes in which there is only one sector energy fund. The focus for this class of funds relies on its exponential growth, as mentioned on the previous section. Regarding the choice of energy sector is due not only to its growth in recent years but also the relevance it has on the global economy, particularly when it comes to investors' interest, as well as on a perusal of the literature which lacks such analysis.

- Second, the majority of studies on weak-form efficiency are about a comparison between countries stock markets, so there is a lack in terms of exchange traded funds literature in this respect.

- Third, the classical autocorrelation test is not always credible in the financial time series analysis once one assume that all the series are normally distributed. Therefore there are more robust tests like Runs Test to detect statistical dependencies of random variables in a series.

- Finally, this work includes some of the most recent and more powerful statistical and econometric techniques, which are rarely used in previous EMH studies. These include structural break unit root tests (as Zivot-Andrews 1992; Vogelsang & Perron 1998 and Perron & Yabu 2009), panel unit root tests (Im et al., 2003; Maddala & Wu, 1999; Pesaran, 2007) and Chow and Denning (1993) joint variance ratio test.

3 Data

The empirical analysis uses daily U.S. market return of a sample composed by 26 equity energy ETFs, see Table $A.1$ in the appendix, for the period of $1st$ January 2008 to 20 January 2012. Therefore all the inception dates are prior to 2008, so that all the ETFs have the same number of observations, adding up to 1022 daily observations for each fund. Our data was selected as to optimize the $N \times T$ panel since the alternatives of having picked either a longer T or a higher N would compromise to the total number of observations available. For example, expanding the time horizon one year would result in only 17 funds, whereas reducing it would increase the total number of funds to 44. Conducting similar analyses for different data sets accounting for such differences in T and N would go beyond the scope of this work, nonetheless one can acknowledge they may be valid complementary exercises.

The data of daily net asset values (NAV), see Figure 2, was collected from each ETF websites and then the NAV daily returns are computed.

The NAV is estimated by subtracting liabilities, on a daily basis, to the market values of the securities held, all divided by the number of shares outstanding. The price changes, expressed in U.S. dollars, were calculated using following expression:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{1}$$

where P_t and P_{t-1} are the daily prices for the periods t and $t-1$, respectively.

The choice for daily price changes is due to the fact that as we increase the variations calculation range it reduces statistical accuracy of the estimation caused by a smaller number of available observations.

Table $A.1$ in the appendix also reflects the cost advantage of ETFs, previously mentioned, as compared to other mutual funds presenting a low average expense ratio of 0.57%. Expense ratio is the annual fee that all funds or ETFs charge their shareholders. It represents the percentage of fund assets paid for operating expenses

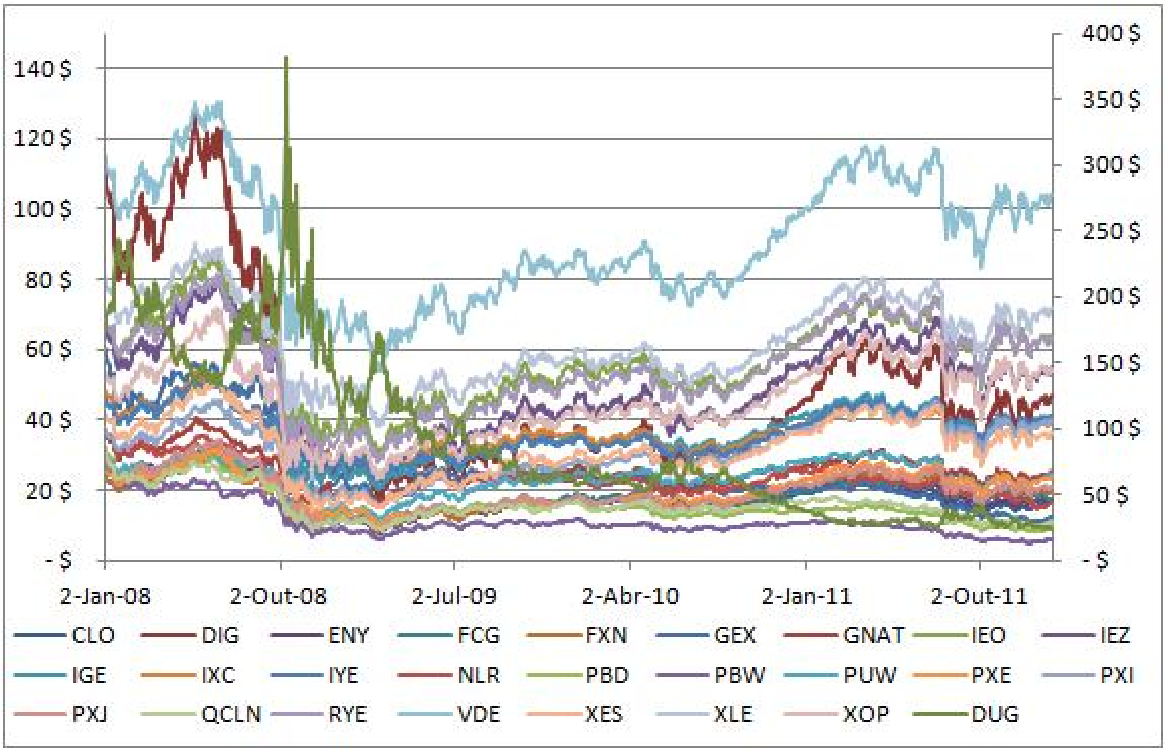


Figure 2: NAV daily evolution 2008-2012.
Notes: Author calculations.Note that DUG is drawn with a different scale and it corresponds to the right hand side axis, whereas the remaining funds should be read using the left hand side axis.

and management fees. The expense ratio does not reflect the fund's brokerage costs, or transaction fees, as well any investor sales charges. Looking at the expense ratio can help to make comparisons among funds.

Our sample includes ETFs from the most important history family funds, see Table $A.2$ in the appendix. More specifically, the well-known iShares, PowerShares, ProShares and SPDRS that track indexes like Dow Jones, NASDAQ and S&P Energy Sector Indexes among other.

4 Methodology and Empirical Results

The main purpose of this study is to shed light and contribute with new empirical evidence on whether the U.S. Equity Energy ETFs price changes follow a random walk model and, therefore, is efficient in weak form. We expect our findings to follow the existing literature on the topic previously discussed in Section 2[4].

The hypotheses are expressed as follows:

H_0: The U.S. Equity Energy ETF price changes follow a random walk.
H_1: The U.S. Equity Energy ETF price changes do not follow a random walk

4.1 Descriptive Statistics

The descriptive statistics computed are the average and the standard deviation of annualized daily returns, the minimum and maximum daily return, kurtosis and skewness coefficients and the Jarque-Bera statistics.

Table 1 exhibits the descriptive statistics of the log of the series returns.

Table 1: **Descriptive Statistics (%)**

Ticker	Average^a	Std. Dev.^{a}	Min.	Max.	Kurtosis	Skewness	Jarque-Bera
CLO	-9.776	41.171	-16.787	14.567	6.107	-0.622	1633,76**
DIG	-21.221	78.293	-37.680	31.927	8.768	-0.719	3322,20**
DUG	-49.110	78.927	-47.339	27.418	12.074	-0.916	6278,04**
ENY	-10.021	44.559	-17.303	15.392	5.543	-0.587	1350,17**
FCG	-6.701	47.165	-16.303	17.708	4.255	-0.418	790,41**
FXN	-4.288	48.928	-19.153	21.184	6.603	-0.531	1880,99**
GEX	-39.282	46.887	-14.285	17.404	5.690	-0.371	1384,62**
GNAT	-7.764	33.033	-10.334	12.005	4.317	-0.302	798,53**
IEO	-1.681	48.629	-20.027	20.023	6.780	-0.509	1977,10**
IEZ	-4.364	49.534	-18.578	18.039	5.165	-0.629	1188,60**
IGE	-3.406	41.012	-16.811	15.656	6.241	-0.483	1677,50**
IXC	-4.622	34.258	-13.828	13.682	6.544	-0.459	1836,32**
IYE	-2.767	39.227	-17.183	17.241	8.095	-0.389	2782,49**
NLR	-18.492	33.829	-11.254	10.790	4.592	-0.513	930,87**
PBD	-30.378	37.310	-11.647	13.892	5.003	-0.427	1083,07**
PBW	-38.015	45.706	-14.382	14.500	3.073	-0.263	407,92**
PUW	-4.383	39.597	-13.841	13.990	3.955	-0.470	694,51**
PXE	-1.727	46.329	-18.076	18.542	6.167	-0.459	1634,74**
PXI	2.159	42.648	-16.265	17.714	6.313	-0.523	1722,28**
PXJ	-7.426	49.383	-17.962	19.092	4.921	-0.544	1067,88**
QCLN	-25.514	44.957	-14.342	15.157	3.167	-0.343	440,82**
RYE	-2.291	46.904	-18.512	19.543	6.589	-0.527	1872,74**
VDE	-2.360	40.592	-17.288	17.546	7.456	-0.415	2367,31**
XES	-3.395	49.742	-18.674	19.570	5.618	-0.624	1392,80**
XLE	-2.744	40.324	-17.279	17.513	7.591	-0.389	2449,57**
XOP	0.409	50.139	-20.417	20.816	6.695	-0.535	1933,66**
Average	-11.507%	46.118%	-18.291%	17.728%	-	-	-

Notes: ^{a} Annualized values. ** indicates significance at 1% level. The Jarque-Bera null hypothesis is that both the skewness and excess of kurtosis are 0.

[4] Software used: E-views, SPSS, Stata and Gauss.

The average annual return is about -11.507% and the average annual standard deviation, which is a measure of risk, presents a high value equal to 46.118%. The average daily minimum and maximum return are about -18% and 18% respectively. These results show how volatile was the U.S. market during the period examined as we can observe in Figure 3.

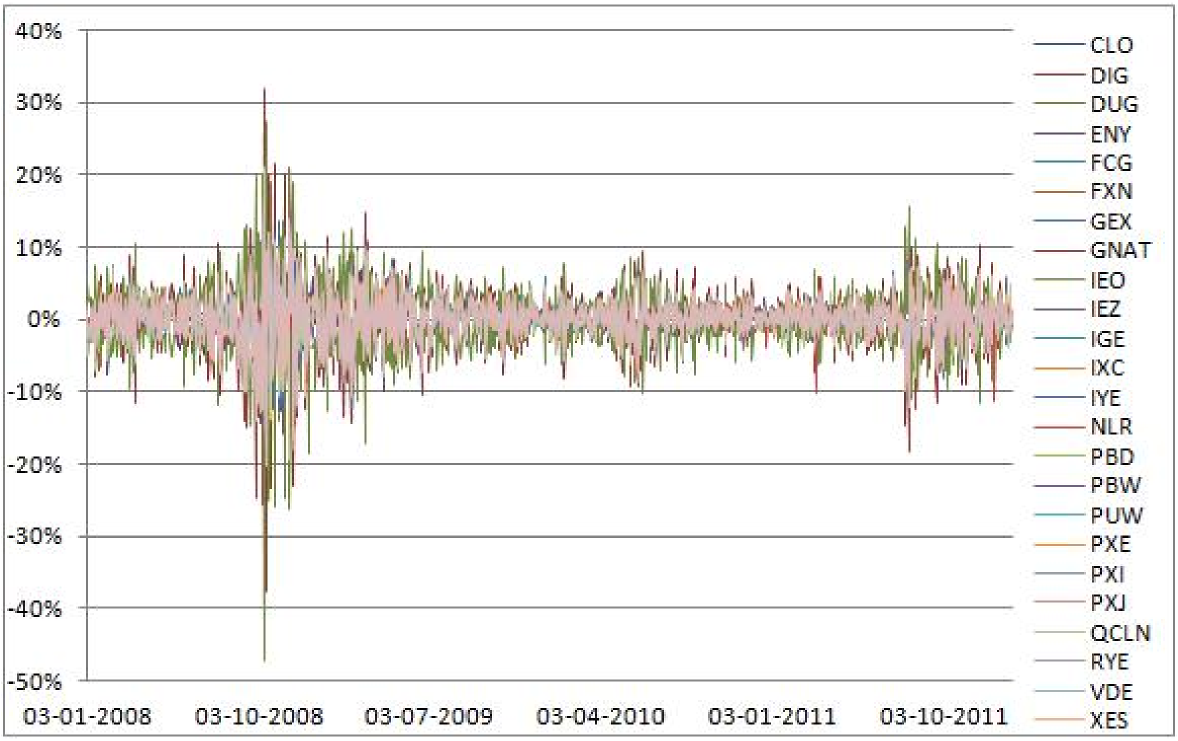


Figure 3: NAV price changes evolution 2008-2012.
Note: Author calculations

The reasons can be related to the big stroke that the world economy suffered in 2008 with the U.S. financial crisis that spread throughout the world. All this made the financial and stock markets to panic, froze credit markets and decreased the economic agents confidence.

The Jarque-Bera statistics, in Table 1, which is a goodness-of-fit test of normality based on skewness and kurtosis indicators, also rejects the hypothesis of a normal distribution of daily returns in all ETFs, at a significance level of 1%. With respect to the skewness coefficient, there is a general negative skewness, which means the distribution is left-skewed. The evidence of the generally high kurtosis indicates that the distribution is leptokurtic. That is, the series returns are not gaussian, see Figure $A.1$ in the appendix. Generally, if the observed distribution is perfectly normally distributed the values for skewness and kurtosis are respectively 0 and 3.

However, if we cut out the 2008 year, the average annual return would be positive, with $6,910\%$, the series would be less volatile with an annual risk of 36.427%, a daily average minimum return of -10.622% and a maximum of $8,731\%$. Finally, the distribution would be nearest to a normal distribution with a kurtosis and skewness coefficients about 1.9 and -0.3 respectively.

4.2 Autocorrelation Tests

An autocorrelation test is a procedure for testing independence of random variables in a series. The serial correlation coefficient measures the relationship between the values of a random variable with itself over successive time intervals. If these coefficients are equal to zero it implies that there is no correlation between variables and, thus, the series price changes follow a random walk. Otherwise, if the coefficients are significantly different from zero, the daily returns of ETFs are not independent of their lagged values and the market is inefficient. We estimate autocorrelation for different lagged time periods ($\mathrm{k} = 1, 2, 3$)[5].

According to Table 2, the majority of the variables are autocorrelated and, therefore, the prices of ETFs are affected by their pricing history, so the price changes series do not follow a random walk model.

The results of 1^{st} order autocorrelation show that there are 12 out of 26 ETFs whose first order autocorrelation is significative: 5 are significant at the 5% level and 7 are significant at 1% level. The average autocorrelation coefficient is -0.036 while the respective P-value is equal to 0.237.

With respect to 2^{nd} and 3^{rd} order autocorrelation there are 13 ETFs, in each case,

[5] The lag choice was selected from cross-checking, as a robustness exercise, different expanding lag length (up to 30). All in all, qualitatively our main conclusions do not change and remain valid, with the exception of the significance behavior of just 5 funds in line with the remaining sample.

Table 2: **Autocorrelation**

	1 Lag		2 Lag		3 Lag	
Ticker	Coef.	Q-Stat	Coef.	Q.Stat	Coef.	Q-Stat
CLO	0.003	0.012	-0.068	4.716	0.032	5.973
DIG	-0.123	15.565**	-0.110	27.897**	0.089	36.100**
DUG	-0.139	19.926**	-0.117	33.954**	0.083	41.033**
ENY	0.070	4.977*	-0.076	10.913**	0.071	16.065**
FCG	-0.020	0.408	-0.057	3.694	0.035	4.951
FXN	-0.058	3.394	-0.092	12.038**	0.072	17.286**
GEX	0.011	0.132	-0.035	1.408	0.001	1.409
GNAT	0.095	9.146**	-0.042	10.960**	-0.055	14.111**
IEO	-0.057	3.342	-0.106	14.816**	0.053	17.671**
IEZ	-0.058	3.464	-0.060	7.112*	0.074	12.760**
IGE	-0.075	5.782*	-0.088	13.753**	0.068	18.442**
IXC	0.014	0.187	-0.133	18.412**	0.047	20.683**
IYE	-0.123	15.607**	-0.109	27.829**	0.090	36.062**
NLR	0.016	0.279	0.022	0.767	0.021	1.206
PBD	0.082	6.824**	0.001	6.825*	0.022	7.332
PBW	0.007	0.054	0.001	0.055	-0.024	0.648
PUW	-0.023	0.523	-0.046	2.677	0.005	2.708
PXE	-0.071	5.231*	-0.078	11.536**	0.051	14.169**
PXI	-0.067	4.538*	-0.059	8.055*	0.067	12.653**
PXJ	-0.037	1.401	-0.046	3.554	0.070	8.530*
QCLN	0.012	0.138	-0.029	0.974	-0.002	0.979
RYE	-0.077	6.108*	-0.097	15.768**	0.074	21.339**
VDE	-0.108	11.971**	-0.101	22.524**	0.077	28.596**
XES	-0.051	2.636	-0.065	7.027*	0.077	13.096**
XLE	-0.115	13.621**	-0.109	25.784**	0.083	32.902**
XOP	-0.052	2.730	-0.100	13.075**	0.045	15.191**
Average	-0.036	-	-0.069	-	0.047	-

Notes:*,** indicates significance at 5% and 1% level respectively. The statistics reported are the Ljung-Box Q-Statistics. Under the null hypothesis that there is no autocorrelation up to order k, where k is equal to 1,2,3.

having significant autocorrelation, indicating that there are only 8 ETFs with no autocorrelation between variables. For the 2^{nd} order autocorrelation the coefficient and the P-value are about -0.069 and 0.137 respectively, and for the 3^{rd} order autocorrelation they are about -0.047 and 0.154.

In sum the ETFs price changes are dependent to their lagged values and the past returns affect future returns either in a positive or a negative way, which are not in line with the random walk hypothesis and so with the efficient market hypothesis. However it is necessary to take into consideration that the autocorrelation test assume the series are normally distributed, which is not a valid assumption, as previously shown, for the ETFs daily price changes distribution. Therefore, we applied a more robust test, runs test, to verify serial independence in the returns.

4.3 Runs Test

The runs test takes a different approach to test and detect statistical dependencies which may not be detected by the autocorrelation test. A run is defined in Siegel's (1956) as a succession of identical symbols which are followed or preceded by different symbols or no symbol at all. The runs test is a non-parametric[6] statistical procedure which examines whether a string of data is occurring randomly ignoring the properties of the distribution. The runs test analyzes the occurrence of similar events, i.e., if there are any patterns or trends in the series and, hence, if the series follow the random walk model. In other words this test is defined as the series of consecutive price changes with identical sign, such as $++, --, 00$, which means positive returns, negative returns or null returns respectively. Thus the effect of outliers is removed.

The null hypothesis is that the succeeding prices changes are independent and move randomly. When the expected number of runs (μ_R) is near from the observed number of runs (R), the null is not rejected. On the other hand, when the observed number of runs are significantly different from the expected number of runs, it indicates that the market overreacted to the information and the prices have a tendency.

As Borges (2010), we performed the runs test assuming as a positive return ($+$) any return greater than the mean return of each ETF, and a negative return ($-$) if it is below the mean return. Let n_+ and n_- reflect the totality of positive returns ($+$) and negative returns ($-$) regarding to a sample of n observations, where $n = n_+ + n_-$. For large sample size the test statistic is just about normally distributed and constructed as:

[6]Non-parametric test does not assume any assumption about the series distribution

$$Z = \frac{R - \mu_R}{\sigma_R} \approx N(0,1) \qquad (2)$$

where $\mu_R = \frac{2n_+n_-}{n} + 1$ and $\sigma_R = \sqrt{\frac{2n_+n_-(2n_+n_- - n)}{n^2(n-1)}}$

Table 3 displays the runs test result, which allows us to test the series randomness according to the return position with respect to the mean return.

Table 3: **Runs Test**

Ticker	Cases < T.V[a]	Cases ≥ T.V[a]	Number of Runs (R)	Expected Runs (μ_R)	Z	p-value
CLO	476	545	478	509,2	-1.961*	0.050
DIG	478	543	503	509,4	-0.404	0.686
DUG	506	515	522	511,5	0.660	0.509
ENY	478	543	486	509,4	-1.473	0.141
FCG	495	526	486	511,0	-1.569	0.117
FXN	478	543	490	509,4	-1.222	0.222
GEX	488	533	493	510,5	-1.099	0.272
GNAT	477	544	480	509,3	-1.843	0.065
IEO	489	532	490	510,6	-1.292	0.196
IEZ	474	547	502	508,9	-0.434	0.665
IGE	484	537	508	510,1	-0.133	0.894
IXC	481	540	486	509,8	-1.495	0.135
IYE	484	537	518	510,1	0.495	0.621
NLR	487	534	501	510,4	-0.591	0.554
PBD	471	550	441	508,4	-4.249**	0.000
PBW	467	554	491	507,8	-1.059	0.289
PUW	475	546	490	509,0	-1.198	0.231
PXE	483	538	494	510,0	-1.006	0.314
PXI	486	535	495	510,3	-0.962	0.336
PXJ	484	537	504	510,1	-0.385	0.701
QCLN	478	543	499	509,4	-0.656	0.512
RYE	500	521	523	511,3	0.734	0.463
VDE	488	533	518	510,5	0.470	0.638
XES	477	544	494	509,3	-0.962	0.336
XLE	489	532	514	510,6	0.214	0.831
XOP	479	542	488	509,6	-1.355	0.175

Notes: [a]Test Value (T.V) = Mean; Total Cases = 1021.* and ** denotes significance at 5% and 1% level respectively. The runs test tests for a statistically significant difference between the expected number of runs (μ_R) and the actual number of runs (R). Null hypothesis is that the succeeding prices changes are independent and moves randomly.

The results are suggest to admit that the series returns are not dependent and moves randomly. The number of expected runs is less than the number of observed runs in 22 ETFs and larger in 4 ETFs. However this difference is only significant in two cases on the sample, namely the CLO and the PBD fund which are significant at 5% and 1% level respectively. That is, these two funds are the only ones who reject the null hypothesis that the series follow a random walk. It means that the observed number of runs are close to the expected number of runs in 24 out of 26 ETFs and hence these returns series are independent and follow a random walk. Furthermore, this gives support that the classical autocorrelation test may be unreliable in the analysis of financial time series.

4.4 Unit Roots and Structural Breaks

We performed some different unit roots tests. Firstly we resort to the standard Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests and then we will also conduct three unit roots tests allowing for structural break, particularly the Zivot-Andrews (1992) (ZA), Vogelsang and Perron (1998) (VP) and Perron and Yabu (2009) (PY).

The unit root test is applied to check the stationarity as a necessary condition for random walk. According to the random walk hypothesis the log price series must have a unit root whereas the returns series must be stationary. A series is said to be stationary if the mean and autocovariances of the series do not depend on time. A nonstationary time series is said to be integrated to order one, or $I(1)$, if the series of its first differences, $\Delta R_t = R_t - R_{t-1}$, is $I(0)$. More generally, a series is integrated to order d, or $I(d)$, if it must be differenced d times before a stationary series, $I(0)$, result. A series is non-stationary, $I(1)$, if it contains a unit root.

A structural break is an econometrics concept that appears when we see an unexpected shift in a time series. This can conduct to large forecasting errors and unreliability of the model. Those shifts can be caused by outliers, wars, strikes, economic crisis, changes in fiscal policies, among others factors. Whenever a series is $I(1)$ one or more breaks may be falsely suggested by the data even if the series are stable over time.

The Augmented Dickey-Fuller Unit Root Test (ADF) constructs a parametric correction for higher order serial correlation by adding lagged differenced terms on the right-hand side of the test regression. In the ADF test, we have to decide whether or not to include a constant and/or time trend or neither in the regression. So we will apply three different alternative estimations of ADF tests: including a constant; including both constant and time trend; and without any constant and time trend in the regression. This test is compared to MacKinnon's (1996) critical values for the rejection of the hypothesis of a unit root. If the ADF coefficient is lower than MacKinnon's critical values, at a 95% confidence level, then the hypothesis of a unit root will be rejected and, consequently, the ETF sample may follow a random walk.

The Phillips-Perron (PP) test offers an alternative non-parametric method for correcting for serial correlation in unit root testing. Basically, they use the standard Dickey-Fuller or ADF test, but modify the t-test statistic so that the serial correlation does not affect the asymptotic distribution of the test statistic. The test is robust with respect to unspecified autocorrelation and heteroscedasticity in the disturbance process of the test equation. As in the case of the ADF test, in the PP test we also have to decide whether or not to include a constant and/or time trend or neither and we also have to compare the PP coefficient to the MacKinnon's (1996) critical values.

Perron (1989) showed that the traditional unit root test had little power to differentiate between a path of a stationary unit root when there was a structural change and that can lead to a mistrustfully rejection of a unit root null hypothesis. Thus, beyond the standard unit root tests, we proceed to three unit root test allowing for breaks: Zivot-Andrews (1992) (ZA); Vogelsang and Perron (1998) (VP); and Perron and Yabu (2009) (PY). As stated by Glynn et al. (2007), "this procedure can identify when the possible presence of a structural break occurred, so it would provide valuable information for analyzing whether a structural break on a certain variable is associated with a particular factor" as the ones that were mentioned upon.

The ZA is a sequential test that allows the endogenous treatment of one identified break in time series in either the intercept, the linear trend or in both. This approach estimates, across all possible breaks, the breakpoint to be where the ADF t-test statistic is minimized, i.e. the most negative, and identifies the corresponding date.[7]

We complement this with the modified ADF test proposed by Vogelsang and Perron (1988) (VP) also allowing for one endogenously determined break. VP enables a class of test statistics that allow for two different forms of structural break, additive outlier (AO) and innovational outlier (IO). The AO model allows for an abrupt shift in level, whereas the IO results in more gradual changes from the initial level to a new one. For the unit root test that allow for one endogenously determined breaks it is assumed that the shift can be modeled by a dummy variable $DU_t = 1$ for t $\leq TB$ and 0 otherwise, where TB is the shift date (time break).

Finally, Perron and Yabu (2009) (PY) proposed a new test for structural changes in the trend function of the time series without any prior knowledge as to whether the noise component is stationary or contains an autoregressive unit root. This test has emerged after has been shown that, in some unit root tests as VP, the critical values are substantially smaller in the $I(0)$ case than in the $I(1)$ case, suggesting that the test is conservative in the stationary case.

Table 4 presents the three different ADF and PP tests respectively: including a constant; including both constant and time trend; and without any constant and time trend in the regression.

The evidence in the three alternatives in both tests suggests the rejection of the null of a unit root in all ETFs and so the conclusion that the series returns are stationary which may support that the equity energy ETFs follow a random walk.

Observing all funds in the three alternatives we found much lower coefficients than the MacKinnon's (1996) critical values. For the first, the second and the third alternative the average coefficients are respectively -29.276, -29.284 and -27.149

[7]The ZA (1992) critical values differs from the Perron's (1989), owing to the fact that the selection of the time of the break is treated as the outcome of an estimation procedure, instead of predetermined exogenously.

Table 4: **Standard Unit Root Tests**

	Augmented Dickey Fuller			Phillips Perron		
Ticker	Constant	Trend and Constant	None	Constant	Trend and constant	None
CLO	-31.801**	-31.802**	-31.809**	-31.908**	-31.909**	-31.911**
DIG	-27.144**	-27.196**	-27.142**	-36.402**	-36.190**	-36.405**
DUG	-27.686**	-27.674**	27.610**	-37.345**	-37.328**	-37.210**
ENY	-29.757**	-29.759**	-29.765**	-29.683**	-29.685**	-29.692**
FCG	-32.550**	-32.539**	-32.563**	-32.550**	-32.539**	-32.563**
FXN	-25.512**	-25.519**	-25.524**	-33.886**	-33.889**	-33.902**
GEX	-31.546**	-31.537**	-31.474**	-31.552**	-31.543**	-31.470**
GNAT	-29.024**	-29.028**	-29.032**	-28.893**	-28.898**	-28.902**
IEO	-25.872**	-25.873**	-25.884**	-33.916**	-33.913**	-33.934**
IEZ	-33.822**	-33.833**	-33.837**	-33.890**	-33.906**	-33.906**
IGE	-25.696**	-25.706**	-25.708**	-34.522**	-34.530**	-34.538**
IXC	-25.608**	-25.626**	-25.618**	-31.602**	-31.649**	-31.616**
IYE	-27.134**	-27.159**	-27.147**	-36.410**	-36.432**	-36.427**
NLR	-31.392**	-31.382**	-31.369**	-31.409**	-31.400**	-31.389**
PBD	-29.399**	-29.395**	-29.342**	-29.395**	-29.390**	-29.342**
PBW	-31.674**	-31.677**	-31.602**	-31.674**	-31.676**	-31.603**
PUW	-32.635**	-32.636**	-32.650**	-32.670**	-32.671**	-32.685**
PXE	-25.391**	-25.411**	-25.404**	-34.380**	-34.396**	-34.398**
PXI	-34.107**	-34.122**	-34.123**	-34.159**	-34.178**	-34.176**
PXJ	-33.114**	-33.128**	-33.127**	-33.157**	-33.195**	-33.170**
QCLN	-31.537**	-31.537**	-31.512**	-31.556**	-31.557**	-31.520**
RYE	-25.972**	-25.984**	-25.985**	-34.683**	-34.695**	-34.701**
VDE	-26.626**	-26.646**	-26.639**	-35.809**	-35.829**	-35.827**
XES	-33.569**	-33.577**	-33.585**	-33.651**	-33.691**	-33.667**
XLE	-26.970**	-26.994**	-26.983**	-36.078**	-36.115**	-36.095**
XOP	-25.648**	-25.651**	-25.661**	-33.743**	-33.789**	-33.760**

Notes:** indicates significance at 1% level. Both under the null hypothesis that the time series has a unit root. Values presented are ADF and PP test statistics.

for the ADF test and -33.266, -33.269 and -33.269 for the PP test. The corresponding critical values are -3.437, -3.967 and -2.567 for the ADF test and for the PP test at the 1% level.

As the previous tests do not take into account the structural breaks in the series, we resort unit root tests that allow for endogenously determined breaks, and that can be observed in Table 5.According to ZA outcome we reject the unit root at 1% significance level for all the 26 funds. The ZA 1% critical value is -5.57. Also for the VP tests, there is a lack of unit roots for all the ETFs, which leads us to reject the null hypothesis once again, but in this case at 5% significance level. As for the VP test, for 5% level, the critical values are -3.56 and -4.27 respectively for the additional outlier model and for the innovational outlier model.

Table 5: **Structural Breaks Unit Root Tests**

	Allowing for endogenously determined breaks			
	ZA	VP(AO)	VP(IO)	PY2009
Ticker	(1)	(2)	(3)	(4)
CLO	21/11/08**	18/11/08*	19/11/08*	20/11/08*
DIG	28/11/08**	07/10/08*	08/10/08*	15/10/08
DUG	13/10/08**	16/10/08*	17/10/08*	10/10/08
ENY	13/10/08**	18/11/08*	19/11/08*	10/10/08
FCG	21/11/08**	18/11/08*	19/11/08*	20/11/08
FXN	21/11/08**	18/11/08*	19/11/08*	20/11/08*
GEX	21/11/08**	18/11/08*	19/11/08*	20/11/08
GNAT	28/10/08**	08/10/08*	09/10/08*	27/10/08
IEO	13/10/08**	18/11/08*	19/11/08*	10/10/08
IEZ	21/11/08**	18/11/08*	19/11/08*	20/11/08
IGE	28/10/08**	18/11/08*	19/11/08*	20/11/08
IXC	28/10/08**	08/10/08*	09/10/08*	10/10/08
IYE	13/10/08**	18/11/08*	17/10/08*	10/10/08
NLR	28/10/08**	02/10/08*	03/10/08*	27/10/08*
PBD	21/11/08**	18/11/08*	19/11/08*	20/11/08**
PBW	21/11/08**	18/11/08*	19/11/08*	20/11/08
PUW	21/11/08**	18/11/08*	19/11/08*	20/11/08*
PXE	21/11/08**	18/11/08*	19/11/08*	20/11/08
PXI	21/11/08**	18/11/08*	19/11/08*	20/11/08
PXJ	21/11/08**	18/11/08*	19/11/08*	20/11/08
QCLN	21/11/08**	18/11/08*	19/11/08*	20/11/08
RYE	21/11/08**	18/11/08*	19/11/08*	20/11/08
VDE	28/10/08**	18/11/08*	19/11/08*	10/10/08
XES	21/11/08**	18/11/08*	19/11/08*	20/11/08*
XLE	13/10/08**	16/10/08*	17/10/08*	10/10/08
XOP	13/10/08**	18/11/08*	19/11/08*	10/10/08

Notes: For the structural-break type tests only dates are presented and when applicable, a statistically significant symbol is added.

Lastly, from the last column of Table 5, we present the PY test results. In this case, the interpretation of the null is not as homogenous as before. Overall, we reject the $I(1)$ hypothesis for only 6 times out of 26. However, the rejections happen with different levels of significance: five funds are statistically significant at 5% level and one at 1% level. PY critical values are 3.12 and 4.47 for 5% and 1% levels respectively.

In this context, all structural breaks took place in October and November of 2008. This evidence can be related to the U.S. financial crisis, that made the financial and stock markets to panic.

In sum, the previous unit root tests set forth that the ETFs price changes are stationary and this could mean that they move randomly.

4.5 Panel Unit Roots

Panel unit root tests are a generalization of the ADF individual unit root tests to a common panel unit root test. As unit root tests, generally, have lower power to differentiate stationary from nonstationary series in small sample sizes, we can increase the number of observations by using panel analysis. The idea is that this panel analysis will be more powerful than performing individual unit root tests for each fund. The null hypothesis is that each individual time series contains a unit root against the alternative that each time series is stationary.

We resort three different types of panel unit root tests: two first generations tests and one second generation test. The former ones concern Im, Pesaran and Shin (2003) test (IPS) and Maddala and Wu (1999) test (MW). The latter one comprise the Pesaran (2007) CIPS test[8].

The first generation of panel unit root tests is based on the cross-sectional independency hypothesis. The second generation tests differ from the first ones by relaxing the cross-sectional independence assumption and accounting for cross-sectional dependencies of the error terms.

The results of first generation tests are displayed in Table 6.This generation is based on the cross-sectional independency hypothesis. The Im, Pesaran and Shin (2003) test (IPS) clearly reject the null hypothesis. With a critical value equal to -1.73 for 5% significance level, a test statistical so negative as -120 is a sufficient evidence against the hypothesis that all return series contain a nonstationary process, $I(1)$.

Also in Table 6 we have the Maddala and Wu (1999) (MW) panel unit root tests. One stated without trend and the other one with trend. The results are not very different: both tests do not reject the null hypothesis for short lags. The alternative without trend reject the presence of unit roots in the second and third level, whereas the test modeled with trend do not reject the $I(0)$ only at third level.

[8]For more detailed discussion of these tests, the reader should refer to the original references.

Table 6: **First Generation Unit Root Tests**

Im, Pesaran and Shin (2003) Panel Unit Root Test (IPS)[a]

Full	returns
in levels	
lags	[t-bar]
1.42	-120**

Maddala and Wu (1999) Panel Unit Root Test (MW)[b]

Full	returns	without trend	returns	with trend
lags	p_λ	(p)	p_λ	(p)
in levels				
0	0.000	1.000	0.000	1.000
1	0.000	1.000	0.000	1.000
2	1771.723**	0.000	0.000	1.000
3	3395.319**	0.000	905.428**	0.000

Notes: [a] We report the average of the return-specific "ideal" lag-augmentation (via AIC). Distribution is approximately t. ** indicates the cases where the null is rejected at 1% level. [b] The MW statistic for different lag-augmentations and the p-values for each of the MW tests are reported. The interested reader can find details on the test statistics in the original references.

The second generation test is an approach allowing cross-sectional dependence and its results are presented in Table 7. Before performing the Pesaran (2007) CIPS test, and to justify the use of the second generation test, we run a simple regression of returns on a constant to test whether there is cross-sectional dependence in sample. Cross-sectional dependence in the error term of the estimated model results then in inconsistent coefficient estimates if the independent variables are correlated with the unspecified common variables or shocks.

Table 7: **Second Generation Unit Root Tests**

Pesaran (2007) Panel Unit Root Test (CIPS)

Full	returns	without trend	returns	with trend
lags	p_λ	(p)	p_λ	(p)
in levels				
0	-24.655	0.000	-24.942	0.000
1	-24.655	0.000	-24.942	0.000
2	-24.655	0.000	-24.942	0.000
3	-24.655	0.000	-24.942	0.000

Notes: Null hypothesis of non-stationarity. P-values for each CIPS tests are reported.

With this in mind, we test the presence of cross-sectional dependence with Pesaran's (2004) CD[9] test statistic. We find a statistic of 417.895, corresponding to a

[9] CD means cross-sectional dependence

P-value of zero (the null hypothesis is cross-sectional independence). We confirmed our result with other versions of this test, such as Friedman's (1937) statistic or the Frees' (1995), obtaining, respectively, a statistic of 18722.036 and 18.291 both matching a zero P-value. Therefore, after confirmed that there is cross-sectional dependence in the series returns we applied the CIPS test.

The CIPS test reject, for all lags, the null hypothesis of nonstationarity, which implies there is no unit roots in our sample.

In this context, our panel unit roots results suggest that equity energy ETF returns may follow a random walk.

4.6 Variance Ratio Tests

The last test, the variance ratio test, VR now on, is another test that has an important role when testing the random walk hypothesis. This method was developed by Lo and Mackinlay (1988) arguing that this test is thought to have more attractive features than others like Dickey-Fuller unit root test. VR examines the predictability of time series data by comparing the returns variances calculated over different intervals. The ratio of the variance of the q-period difference scaled by q to the variance of the first difference must equal to one to assume data follows a random walk,

$$VR(q) = \frac{\sigma^2(q)}{\sigma^2(1)} \tag{3}$$

where $\sigma^2(q)$ is the variance of returns of the period $q > 1$ and $\sigma^2(1)$ is the variance of returns in the first period and are defined as,

$$\sigma^2(q) = \frac{1}{Tq}\sum_{t=1}^{T}(y_t - y_{t-q} - q\hat{\mu})^2 \tag{4}$$

$$\sigma^2(1) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{\mu})^2 \tag{5}$$

where $\hat{\mu} = (1/t)\sum_{t=1}^{T}(y_t - y_{t-1})$

Firstly, Lo and Mackinlay (1988) assume that the errors, ϵ_t, are homescedastic, i.e., ϵ_t are independent and identically distributed, $\epsilon_t \sim iid\ N(0, \sigma_0^2)$. Then, they argue that there are evidences in which the financial series volatility changes with time, and so there is heteroscedasticity, which can lead to reject the random walk null hypothesis. To face this, Lo and Mackinlay (1988) recommend two alternative test statistics: one assuming homoscedasticity, $Z(q)$, and other one allowing heteroscedasticity, $Z^*(q)$.

The statistic under the homoscedasticity follow the standard normal distribution asymptotically $N(0,1)$,

$$Z(q) = \frac{VR(q) - 1}{\varphi(q)^{1/2}} \tag{6}$$

where

$$\varphi(q) = \frac{(2(2q-1)(q-1))}{3qT} \tag{7}$$

The second test statistic under the assumption of heteroscedasticity

$$Z^*(q) = \frac{VR(q) - 1}{\varphi^*(q)^{1/2}} \tag{8}$$

is also asymptotically distributed as $N(0,1)$, where

$$\varphi^*(q) = \sum_{j=1}^{q-1} [\frac{2(2q-j)}{q}]^2 \times \delta(j) \tag{9}$$

and

$$\delta(j) = \frac{\sum_{t=j+1}^{k-1}(y_t - \hat{\mu})^2 (y_{t-j} - \hat{\mu})^2}{[\sum_{t=1}^{T}(y_t - \hat{\mu})^2]^2} \tag{10}$$

The process suggested by Lo and MacKinlay (1988) is created to test single VR for a particular q-difference. However, under the random walk hypothesis there must have $VR(q)$ equal to 1 for all q. If returns are positively (negatively) autocorrelated, $VR(q)$ should be greater (less) than 1. Still, a series has mean-reversion if $VR(q)$ is significantly less than unity at long horizons, and has mean-aversion, or persistence, if $VR(q)$ is significantly greater than unity for long horizons. The rejection of the null hypothesis for a single q value, implies the rejection of the random walk hypothesis. To face this, Chow and Denning (1993), developed a multiple VR test that consider the maximum absolute value of a set of multiple VR statistics, and is classified as a joint variance ratio test[10] since it test the joint null hypothesis for all periods.

In order to make this study consistent with other recent research[11], the VR test were conducted with different lag lengths, i.e. $q = 2$, 5, 10 and 30, for daily data.

Table 8 exhibit the results of the variance ratio test for equity energy ETFs under the assumption of homoscedasticity.There are 3 funds (ENY, NLR and PBD)

Table 8: **Variance Ratio Tests under the assumption of Homoscedasticity**

	Lag 2		Lag 5		Lag 10		Lag 30		C-D
Ticker	VR(2)	Z(2)	VR(5)	Z(5)	VR(10)	Z(10)	VR(30)	Z(30)	Z
CLO	1.005	0.169	0.976	-0.347	0.896	-0.986	0.938	-0.321	0.986
DIG	0.878	-3.886**	0.748	-3.670**	0.712	-2.728**	0.644	-1.844	3.886**
DUG	0.862	-4.404**	0.698	-4.408**	0.611	-3.679**	0.429	-2.960**	4.408**
ENY	1.072	2.290*	1.098	1.435	1.021	0.194	1.031	0.161	2.290
FCG	0.982	-0.579	0.939	-0.889	0.893	-1.015	0.817	-0.951	1.015
FXN	0.944	-1.782	0.852	-2.161*	0.804	-1.855	0.796	-1.055	2.161
GEX	1.013	0.425	0.984	-0.239	0.905	-0.896	0.951	-0.253	0.896
GNAT	1.097	3.084**	1.090	1.313	1.031	0.290	0.911	-0.460	3.084**
IEO	0.945	-1.767	0.823	-2.580**	0.761	-2.261*	0.638	-1.876	2.580*
IEZ	0.944	-1.802	0.885	-1.670	0.826	-1.645	0.861	-0.720	1.802
IGE	0.927	-2.345*	0.830	-2.474*	0.767	-2.207*	0.676	-1.679	2.474
IXC	1.015	0.491	0.920	-1.172	0.858	-1.340	0.722	-1.441	1.441
IYE	0.878	-3.891**	0.743	-3.749**	0.683	-3.001**	0.562	-2.272*	3.891**
NLR	1.018	0.583	1.082	1.199	1.042	0.393	1.113	0.583	1.199
PBD	1.084	2.675**	1.155	2.263*	1.128	1.211	1.317	1.641	2.675*
PBW	1.009	0.294	1.005	0.072	0.967	-0.314	0.951	-0.253	0.314
PUW	0.979	-0.661	0.915	-1.236	0.881	-1.129	0.921	-0.411	1.236
PXE	0.930	-2.226*	0.829	-2.488*	0.769	-2.185*	0.661	-1.759	2.488
PXI	0.935	-2.070*	0.870	-1.893	0.813	-1.767	0.800	-1.035	2.070
PXJ	0.965	-1.127	0.929	-1.035	0.858	-1.342	0.898	-0.531	1.342
QCLN	1.013	0.431	0.981	-0.283	0.911	-0.843	0.925	-0.387	0.843
RYE	0.925	-2.410*	0.814	-2.717**	0.743	-2.427*	0.670	-1.712	2.717*
VDE	0.894	-3.400**	0.767	-3.397**	0.700	-2.840**	0.598	-2.085*	3.400**
XES	0.951	-1.564	0.893	-1.558	0.825	-1.652	0.856	-0.747	1.652
XLE	0.886	-3.630**	0.754	-3.591**	0.694	-2.893**	0.581	-2.171*	3.630**
XOP	0.950	-1.591	0.832	-2.443*	0.766	-2.215**	0.652	-1.806	2.443

Notes: Variance ratios, VR and the test statistics, $Z(j)$ with $j = 2, 5, 10, 30$, for homoscedasticity are reported. The null hypothesis is that the variance ratio equal unity, which means that the series follow a random walk. The Chow and Denning (1993), C-D, statistic, which tests all the $Z(q)$ together is also presented.*,** indicates significance at 5% and 1% level respectively.

[10] For more detailed discussion of these test, the reader should refer to the original references.

[11] As in Borges (2010).

that presents all VR larger than unity, which indicates that the variances grow more than proportionally with time. Besides these, there are also six ETFs (CLO, FXN, GEX,IXC, PBW and QCLN) with at least one VR larger than one. All the other ETFs display VR less than one. With respect to the individual tests we observe that the number of rejections decreases as lag length increases.

Table 9: **Variance Ratio Tests under the assumption of Heteroscedasticity**

	Lag 2		Lag 5		Lag 10		Lag 30		C-D
Ticker	VR(2)	Z*(2)	VR(5)	Z*(5)	VR(10)	Z*(10)	VR(30)	Z*(30)	Z*
CLO	1.005	0.095	0.976	-0.199	0.896	-0.561	0.938	-0.186	0.561
DIG	0.878	-2.071*	0.748	-1.785	0.712	-1.321	0.644	-0.952	2.071
DUG	0.862	-2.413*	0.698	-2.133*	0.611	-1.779	0.429	0.429	2.413
ENY	1.072	1.375	1.098	0.839	1.021	0.112	1.031	0.095	1.375
FCG	0.982	-0.385	0.939	-0.551	0.893	-0.623	0.817	-0.603	0.623
FXN	0.944	-1.040	0.852	-1.166	0.804	-0.998	0.796	-0.595	1.166
GEX	1.013	0.218	0.984	-0,123	0.905	-0.454	0.951	-0.129	0.454
GNAT	1.097	2.168*	1.090	0.854	1.031	0.183	0.911	-0.292	2.168
IEO	0.945	-1.020	0.823	-1.365	0.761	-1.195	0.638	-1.047	1.365
IEZ	0.944	-1.212	0.885	-1.020	0.826	-0.978	0.861	-0.438	1.212
IGE	0.927	-1.358	0.830	-1.336	0.767	-1.175	0.676	-0.926	1.358
IXC	1.015	0.280	0.920	-0.617	0.858	-0.698	0.722	-0.775	0.775
IYE	0.878	-2.097*	0.743	-1.843	0.683	-1.474	0.562	-1.187	2.097
NLR	1.018	0.402	1.082	0.787	1.042	0.252	1.113	0.375	0.787
PBD	1.084	1.622	1.155	1.305	1.128	0.681	1.317	0.925	1.622
PBW	1.009	0.193	1.005	0.046	0.967	-0.196	0.951	-0.161	0.196
PUW	0.979	-0.455	0.915	-0.782	0.881	-0.699	0.921	-0.259	0.782
PXE	0.930	-1.259	0.829	-1.332	0.769	-1.171	0.661	-0.991	1.332
PXI	0.935	-1.252	0.870	-1.033	0.813	-0.946	0.800	-0.578	1.252
PXJ	0.965	-0.777	0.929	-0.639	0.858	-0.804	0.898	-0.326	0.804
QCLN	1.013	0.296	0.981	-0.183	0.911	-0.535	0.925	-0.250	0.535
RYE	0.925	-1.434	0.814	-1.477	0.743	-1.306	0.670	-0.962	1.477
VDE	0.894	-1.920	0.767	-1.737	0.700	-1.444	0.598	-1.120	1.920
XES	0.951	-0.977	0.893	-0.904	0.825	-0.944	0.856	-0.441	0.977
XLE	0.886	-2.005*	0.754	-1.806	0.694	-1.451	0.581	-1.153	2.005
XOP	0.950	-0.914	0.832	-1.298	0.766	-1.181	0.652	-1.023	1.298

Notes: Variance ratios, VR and the test statistics, $Z^*(j)$ with $j = 2, 5, 10, 30$, for heteroscedasticity are reported. The null hypothesis is that the variance ratio equal unity, which means that the series follow a random walk. The Chow and Denning (1993), C-D, statistic, which tests all the $Z^*(q)$ together is also presented.*,** indicates significance at 5% and 1% level respectively.

It also evident in Table 9 that the H_0, under heteroscedasticity assumption, is practically never rejected. As regards to the joint variance ratio test, namely Chow and Denning (1993), we find under homoscedasticity assumption only 9 out of 26 ETFs that do not follow a random walk. However, under the hypothesis allowing for heteroscedasticity, which deem more credible and appropriate among financial economists, the random walk hypothesis is never rejected.

5 Conclusions

This empirical work enlarge the existing literature on the EMH by investigating the weak form efficiency of equity energy exchange traded fund returns. The purpose of the study is to investigate whether the selected ETFs follows the random walk model. As stated by Lo and Mackinlay (1988) "the rejection of the random walk hypothesis does not necessarily imply the inefficiency of stock market", because the conclusions are sample-based. In the case of this study, the sample consists of 26 equity energy ETFs listed on U.S. market.

The results reveal that equity energy ETF returns are efficient in the weak form. Table 10 summarizes the results of all the tests performed. Jarque-Bera test clearly rejects the normality of the daily NAV returns, showing that the distribution is left skewed and leptokurtic. With respect to the autocorrelation test the price changes are dependent to their lagged values, so the past returns affect future returns either in a positive or a negative way. However, the autocorrelation test assume the series are normally distributed, which is not a valid assumption in this sample. Nevertheless, the runs test, which is a non-parametric approach to test statistical dependencies, admits the series returns are not dependent and moves randomly, which consequently means that they can be weakly efficient (in line with Rompotis,2011) and showing that autocorrelation results do not necessarily imply that they are predictable.

For assess the stationarity several different methods are used. First the standard Augmented Dickey-Fuller and Phillips-Perron conclude the series price changes are stationary which suggests that the equity energy ETFs are considered efficient in the weak form. Secondly, 3 out of 4 structural break unit root tests (Zivot-Andrews, 1992; two forms of Vogelsang and Perron, 1998), set forth that ETF returns are stationary, which is a necessary conditions for random walk. Only the Perron and

Table 10: **Summary of Test Results: Random Walk Hypothesis Rejected?**

			Unit Root Tests					Panel Unit Root[a]			VR	
Ticker	AC	Runs	ADF	PP	ZA	VP	PY	IPS	MW	CIPS	Z	Z*
CLO	NO	YES	NO	NO	NO	NO	NO				NO	NO
DIG	YES	NO	NO	NO	NO	NO	YES				YES	NO
DUG	YES	NO	NO	NO	NO	NO	YES				YES	NO
ENY	YES	NO	NO	NO	NO	NO	YES				NO	NO
FCG	NO	NO	NO	NO	NO	NO	YES				NO	NO
FXN	YES	NO	NO	NO	NO	NO	NO				NO	NO
GEX	NO	NO	NO	NO	NO	NO	YES				NO	NO
GNAT	YES	NO	NO	NO	NO	NO	YES				YES	NO
IEO	YES	NO	NO	NO	NO	NO	YES				YES	NO
IEZ	YES	NO	NO	NO	NO	NO	YES				NO	NO
IGE	YES	NO	NO	NO	NO	NO	YES				NO	NO
IXC	YES	NO	NO	NO	NO	NO	YES				NO	NO
IYE	YES	NO	NO	NO	NO	NO	YES	NO	NO	NO	YES	NO
NLR	NO	NO	NO	NO	NO	NO	NO				NO	NO
PBD	NO	YES	NO	NO	NO	NO	NO				YES	NO
PBW	NO	NO	NO	NO	NO	NO	YES				NO	NO
PUW	NO	NO	NO	NO	NO	NO	NO				NO	NO
PXE	YES	NO	NO	NO	NO	NO	YES				NO	NO
PXI	YES	NO	NO	NO	NO	NO	YES				NO	NO
PXJ	NO	NO	NO	NO	NO	NO	YES				NO	NO
QCLN	NO	NO	NO	NO	NO	NO	YES				NO	NO
RYE	YES	NO	NO	NO	NO	NO	YES				YES	NO
VDE	YES	NO	NO	NO	NO	NO	YES				YES	NO
XES	YES	NO	NO	NO	NO	NO	NO				NO	NO
XLE	YES	NO	NO	NO	NO	NO	YES				YES	NO
XOP	YES	NO	NO	NO	NO	NO	YES				NO	NO

Notes: AC = Autocorrelation; ADF = Augmented Dickey-Fuller; PP = Phillips-Perron; VP = Vogelsang-Perron in both Additive and Inovational Outlier forms;[a]These tests do not reject RWH in panel. IPS = Im, Pesaran and Shin; MW = Maddala-Wu; CIPS = Pesaran (2007); VR = Variance Ratio; Z and Z* = Homoscedasticity and Hetroscedasticity assumptions respectively.

Yabu (2009) test reject the interpretation of the null is not as homogenous as before and only rejects the stationarity hypothesis for six times out of 26. Lastly, three different types of panel unit root tests are used, namely Im, Pesaran and Shin (2003) test (IPS), Maddala and Wu (1999) test and the Pesaran (2007) CIPS test, and all of them reject the null hypothesis.

Finally, the last method to test the random walk hypothesis is the Variance Ratio test which is broadly in line with the previous tests, i.e., do not reject that ETF price changes moves randomly and therefore they could be weakly-efficient.

It is worth noting that is important to replicate studies on EMH, whereas sometimes different studies find contradicting evidences for the sample and periods. As pointed by Borges (2010) "several studies suggest that market efficiency tends to develop over time, which justifies updating previous studies, using more recent data and a new set of more powerful techniques".

6 Future Research

For future research it could be interesting to:

- Include more advanced econometric techniques, such as the variance ratio tests used by Borges (2010).
- Test the Exchange Traded Funds ability to replicate the performance of their underlying benchmark indexes, i.e. their tracking efficiency.
- Verify, for other markets, to what extent the classical autocorrelation tests are considered reliable in the analysis of the independence of random variables.
- Conduct similar analyses in a few years to get a larger period and split the analysis into sub-periods so that they can be interpreted separately.

References

Borges, M. R. (2010). Efficient market hypothesis in European stock markets, *The European Journal of Finance*, 16 (7), 711-726.

Carhart, M. M. (1997). On persistence in mutual fund performance, *Journal of Finance*, 52 (1), pp. 57-82.

Chow, K., and Denning, K. (1993). A simple multiple variance ratio test, *Journal of Econometrics* , 58, 385-401.

Cootner, P. (1964). *The random character of stock market prices*, M.I.T. Press, Cambridge, MA.

Deville, L. (2006). Exchange Traded Funds: History, Trading and Research. In: Zopounidis, C., Doumpos, M., Pardalos, P. M., (Eds.) *Handbook of Financial Engineering*, Springer, pp. 1-37.

Elton, E. J., Gruber, M.J., Das, S. and Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios, *Review of Financial Studies*, 6 (1), pp. 1-22.

Evans, T. (2006). Efficiency tests of the UK financial futures markets and the impact of electronic trading systems, *Applied Financial Economics,* 16 (17), pp. 1273-1283.

Fama, E. F. (1965). The behavior of stock market prices, *Journal of Business,* 38 (1), pp. 34-105.

Fama, E. F. (1970). Efficient Capital Markets: A Review of theory and empirical work, *Journal of Finance,* 25 (2), pp. 383-417.

Fama, E. F. (1991). Efficient capital markets: II, *Journal of Finance,* 46 (5), pp. 1575-1617.

Frees, E. W. (1995). Assessing cross-sectional correlations in panel data, *Journal of Econometrics,* 69, 393-414.

French, K. (1980). Stock returns and the weekend effect, *Journal of Financial Economics*, 8, pp. 55-69.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Journal of the American Statistical Association,* 32 (200), 675-701.

Glynn, J. , Perera, N. and Verma, R. (2007). Unit Root Tests and Structural Breaks : A Survey with Applications, *Revista De Metodos Cuntitativos para La Econofa La Empresa,* 3, 63-79.

Haugen, R. A. and Lakonishok, J. (1988). *The incredible January Effect*, Dow Jones-Irwin, Homewood.

Hawawini, G. and Michel, P. (1984). *European equity markets: A review of the evidence on price behaviour and efficiency in European equity markets: Risk, return and efficiency*, Garland Publishing Company, New York and London.

Hudson, R., Dempsey, M. and Keasev, K. (1996). A note on the weak-form efficiency of capital markets: The application of simple technical trading rules to UK stock prices - 1935 to 1994, *Journal of Banking and Finance,* 20, pp. 1121-1132.

Im, K. S., Pesaran, M. H. and Shin, Y. (2003), Testing for unit roots in heterogeneous panels, *Journal of Econometrics,* 115, 53-74.

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios, *Journal of Business*, 42 (2), pp. 167-247.

Keim, D. B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence, *Journal of Financial Economics*, 12, pp. 13-32.

Kosev, M. and Williams, T. (2011). Exchange-traded Funds, *Reserve Bank of Australia Quarterly Bulletin, March*, pp. 51-59.

Lo, A. W. and MacKinlay, A.C. (1988). Stock Market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies*, 1 (1), pp. 41-66.

Lo, A. W. and MacKinlay, A.C. (1999). *A non-Random walk down wall street*, Princeton University Press, Princeton.

Lo, A.W., Mamaysky, H. and Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation, *Journal of Finance,* 55 (4), pp. 1705-1765.

MacKinnon, J.G. (1996). Numerical Distribution Functions for Unit Root and Cointegration Tests, *Journal of Applied Econometrics*, 11 (6), pp. 601-618.

Maddala, G. S., and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and New Simple Test, *Oxford Bulletin of Economics and Statistics,* 61, 631-652.

Malkiel, B. G. (2003). The Efficient Market Hypothesis and its critics, *Journal of Economic Perspectives,* 17 (1), pp. 59-82.

Nascimento, V. (2007). Eficiência Informacional do Mercado de Acções: *O Caso Português*, Tese de Mestrado, Faculdade de Economia da Universidade do Porto.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Econometrica,* 57 (6), 1361-1401.

Perron, P. and Yabu, T. (2009). Testing for Shifts in Trend with an Integrated or Stationary Noise Component, *Journal of Business and Economics Statistics,* 27 (3), 369-396.

Pesaran, M. H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels, *Cambridge Working Papers in Economics, No. 0435*, Faculty of Economics, University of Cambridge.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross section dependence, *Journal of Applied Econometrics,* 22, 265-312.

Pring, M. J. (1991). *Technical Analysis Explained*, Second Edition, New York: McGraw-Hill.

Rompotis, G. G. (2011). Testing Weak-Form Efficiency of Exchange Traded Funds Market, *The IEB International Journal of Finance,* 2, pp.2-33.

Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly, *Industrial Management Review,* 6 (2), pp. 41-49.

Sharpe W. (1966). Mutual fund performance, *Journal of Business,* 39 (1), pp. 119-138.

Shiller, R.J. (2000). *Irrational exuberance,* Princeton University Press, Princeton.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, Inc., New York.

Vogelsang, T. and Perron, P. (1998). Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time, *International Economic Review,* 39 (4), 1073-1100.

Zivot, E. and Andrews, D. W. K. (1992). Further Evidence on the Great Crash, the Oil-Price Shock and the Unit Root Hypothesis, *Journal of Business and Economic Statistics,* 10 (3), 251-270.

Appendix

Table A.1 - ETFs Profiles

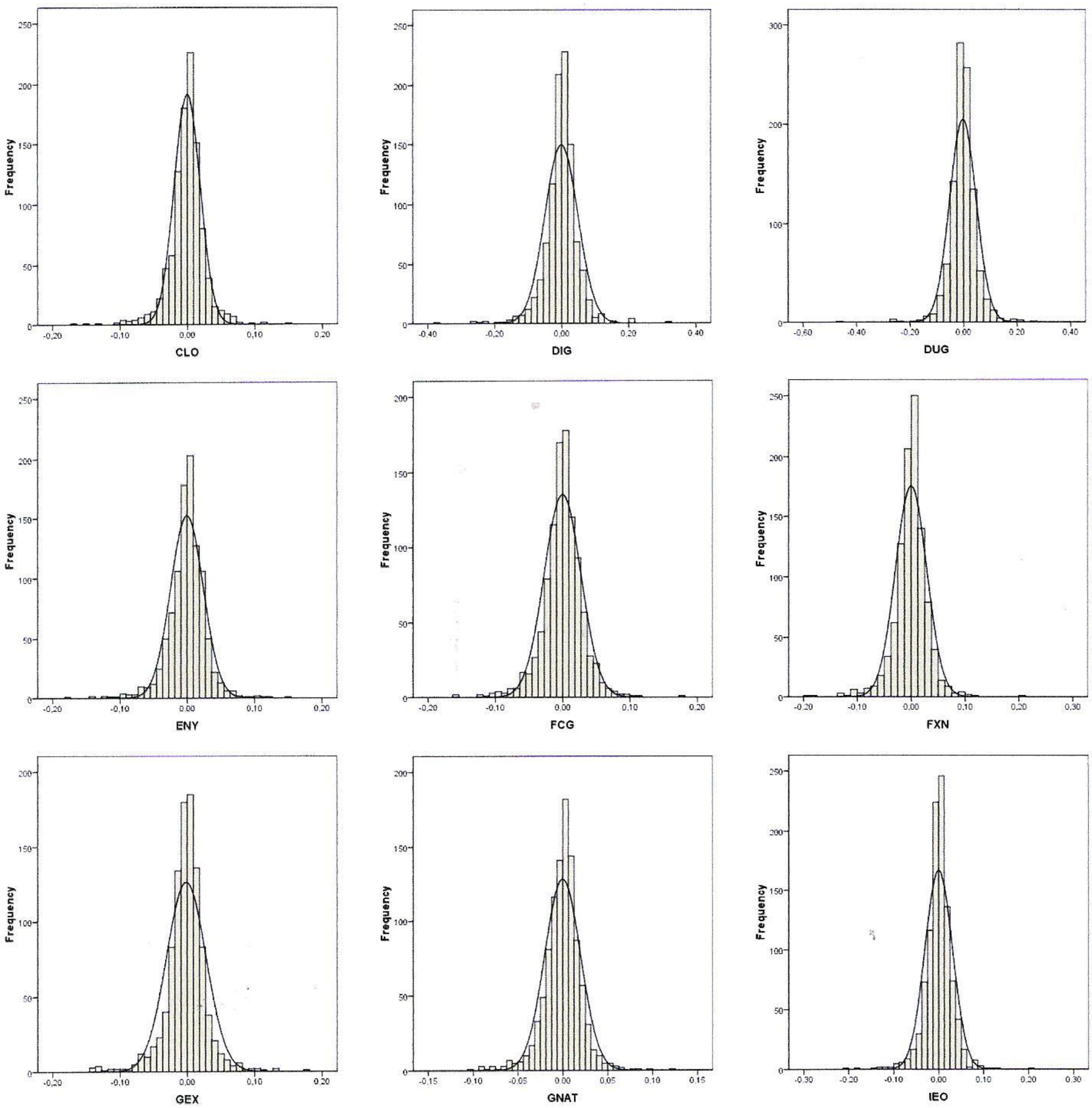
Ticker	Name	Inception Date	Expense Ratio(%)
CLO	Oil Sands Sector ETF	26/10/2006	0,67
DIG	ProShares Ultra Oil & Gas	30/01/2007	0.95
DUG	ProShares UltraShort Oil & Gas	30/01/2007	0.95
ENY	Guggenheim Canadiam Energy Income ETF	03/07/2007	0.65
FCG	First Trust ISE-Revere Natural Gas Index Fund	08/05/2007	0.60
FXN	First Trust Energy Alpha Dex Fund	08/05/2007	0.70
GEX	Market Vectors Global Alternative Energy ETF	03/05/2007	0.60
GNAT	Wisdom Tree Global Natural Resources Fund	13/10/2006	0.58
IEO	iShares Dow Jones US Oil & Gas Exploration & Production Index Fund	01/05/2006	0.47
IEZ	iShares Dow Jones US Oil Equipment & Services Index Fund	01/05/2006	0.47
IGE	iShares S&P North American Natural Resources Sector Index Fund	22/10/2001	0.48
IXC	iShares S&P Global Energy Sector Index Fund	12/11/2001	0.48
IYE	iShares Dow Jones US Energy Sector Index Fund	12/06/2000	0.47
NLR	Market Vectors Uranium + Nuclear Energy ETF	13/08/2007	0.57
PBD	PowerShares Global Clean Energy	13/06/2007	0.75
PBW	PowerShares WilderHill Clean Energy	03/03/2005	0.70
PUW	PowerShares WilderHill Progressive Energy Portfolio	24/10/2006	0.70
PXE	PowerShares Dynamic Energy Exploration & Production Portfolio	26/10/2005	0.63
PXI	PowerShares Dynamic Energy Sector Portfolio	12/10/2006	0.65
PXJ	PowerShares Dynamic Oil & Gas Services Portfolio	26/10/2005	0.63
QCLN	First Trust NASDAQ Clean Edge Green Energy Idx	08/02/2007	0.60
RYE	Guggenheim S&P 500 Equal Weight Energy ETF	01/11/2006	0.50
VDE	Vanguard Energy ETF	23/09/2004	0.19
XES	SPDR S&P Oil & Gas Equipment & Services ETF	19/06/2006	0.35
XLE	Energy Select Sector SPDR Fund	16/12/1998	0.18
XOP	SPDR S&P Oil & Gas Exploration & Production ETF	19/06/2006	0.35

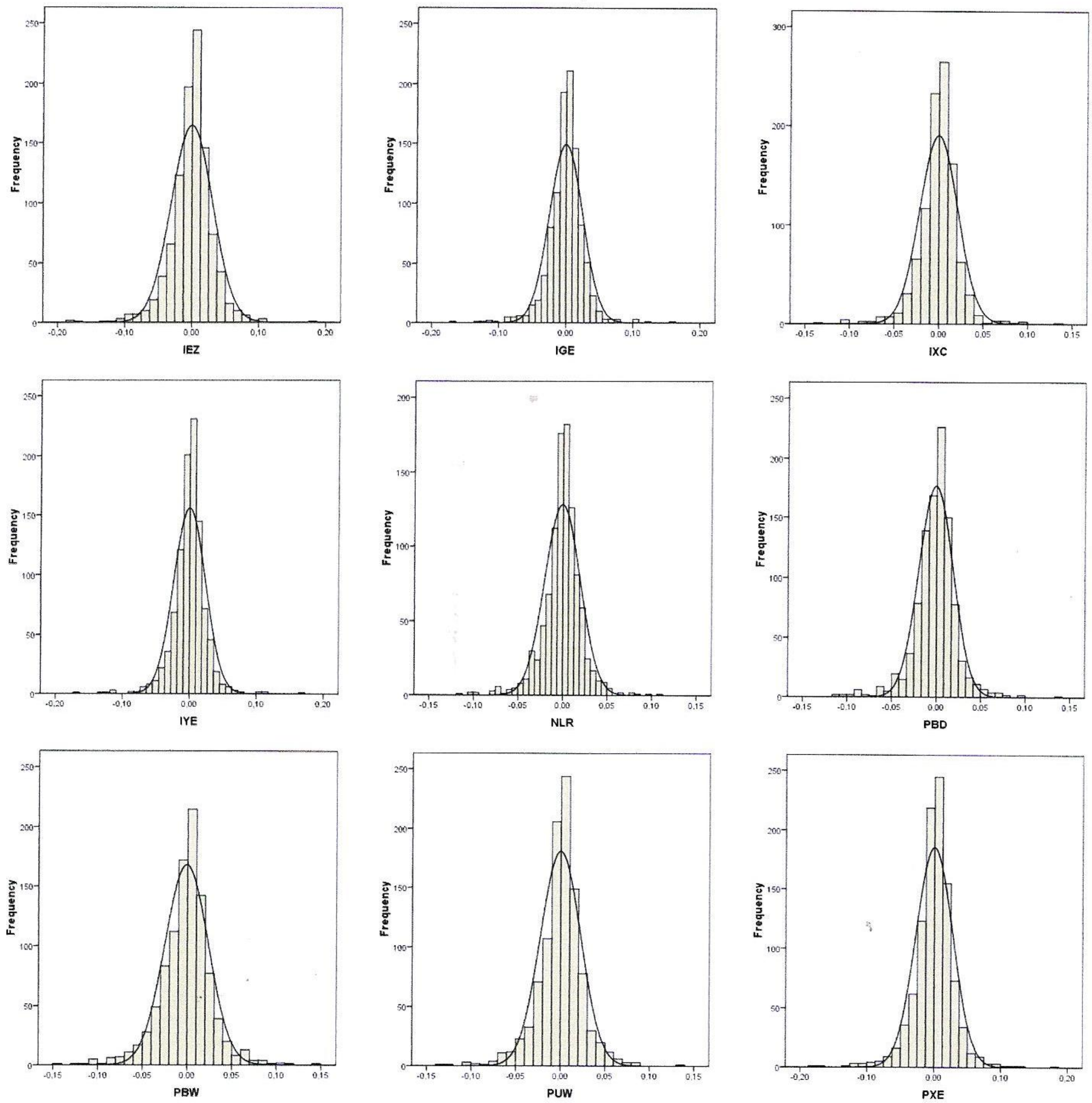
Average	0.57

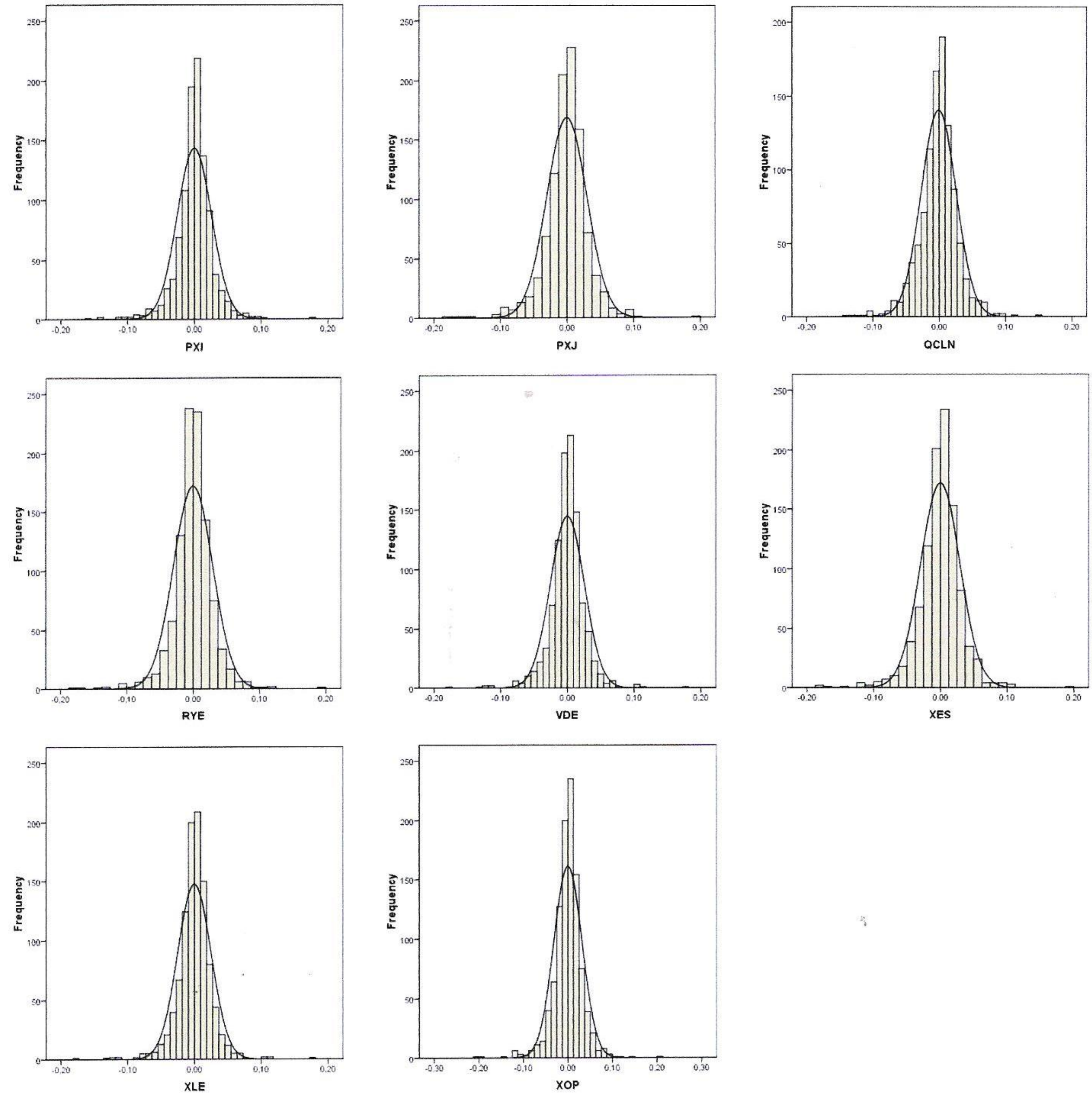
Table A.2 - ETFs Fund Family and Tracked Indexes

Ticker	Fund Family Name	Tracked Indexes
CLO	Claymore	Sustainable Wealth Oil Sands Sector Index
DIG	ProShares	Dow Jones U.S. Oil & Gas Index
DUG	ProShares	Dow Jones U.S. Oil & Gas Index
ENY	Guggenheim Investments	Sustainable Canadian Energy Income Index
FCG	First Trust	ISE-Revere Natural Gas Index
FXN	First Trust	StrataQuant Energy Index
GEX	Van Eck	Ardour Global Index
GNAT	WisdomTree Trust	WisdomTree Global Natural Resources Index
IEO	iShares	Dow Jones U.S. Select Oil Exploration & Production Index
IEZ	iShares	Dow Jones U.S. Select Oil Equipment & Services Index
IGE	iShares	S&P North American Natural Resources Sector Index
IXC	iShares	S&P Global Energy Sector Index
IYE	iShares	Dow Jones U.S. Oil & Gas Index
NLR	Van Eck	DAX global Nuclear Energy Index
PBD	PowerShares	WilderHill New Energy Global Innovation Index
PBW	PowerShares	WilderHill Clean Energy Index
PUW	PowerShares	WilderHill Progressive Energy Index
PXE	PowerShares	Dynamic Energy Exploration & Production Intellidex Index
PXI	PowerShares	Dynamic Energy Sector Intellidex Index
PXJ	PowerShares	Dynamic Oil & Gas Services Intellidex Index
QCLN	First Trust	NASDAQ Clean Edge Green Energy Index
RYE	Guggenheim Investments	S&P Equal Weight Index Energy
VDE	Vanguard	MSCI U.S. Investable Market Energy Index
XES	State Street Global Advisors	S&P Oil & Gas Equipment & Services Select Industry Index
XLE	State Street Global Advisors	Energy Select Sector Index
XOP	State Street Global Advisors	S&P Oil & Gas Exploration & Production Select Industry Index

Figure A.1: Histogram of ETFs Daily Returns vs Normal Distribution




Frequency
IEZ
IGE
IXC
IYE
NLR
PBD
PBW
PUW
PXE


Frequency
PXI
PXJ
QCLN
RYE
VDE
XES
XLE
XOP