Stock market performance of alternative energy companies and changes in oil prices

Bulliard Loris
Bruguet Marie

December 2020

1 Introduction

The energy sector is a contemporary concern because of the fast pace increasing demand and the rising worry about the environment and climate change. Oil is a fundamental raw material for our modern economies but there are several limits we are facing for this source of energy. Indeed, the extraction, refinement, transport, and consumption of oil contribute among other things to deteriorate the environment and to increase greenhouse emissions. Moreover, a large part of the reserves belongs to a few countries which are in geopolitical hotspots (Iran, Iraq, Kuwait). These countries have agreements or belong to international organizations such as Opep to control the market. This leads to instability for European economies which rely on oil importation. These global threats have triggered the interest to develop cleaner energy sources in an energy mix. Indeed, diversifying sources of energy can help reducing uncertainty.

The growth of the clean energy market depends on oil and clear energy production technologies. **Sadorsky and Henriques (2008)** and other papers bring substitution to explain the positive relationship between oil and clean energies. While there is partial substitution possibility between oil and cleaner energies for several types of industries (**Kumar et al. 2015**), some inertia still remains in the short term because of the need to amortize existing equipment and to address technical constraints. High substitution requires a long term vision and high investment in clean energy technology.

Concerning the investment in renewable energy capacity, European union invested 698 billion dollars (28% of the global) during the 2010-2019 decade while China and USA invested 758 and 356 billion dollars respectively (Global Trends in Renewable Energy Investment 2019). According to the Global Status Report published by REN21 (2020), less than 20% of global final energy consumption was from renewable energy supply in 2018. Moreover, the International Energy Agency (IEA) in their Global Energy and CO2 Status Report (2019) confirms that renewable energies met around 25% of the growth in total primary energy demand. However, the largest source of primary energy remains crude oil with a 30% share in 2019 (BP Statistical Review of world energy 2020). The major development in renewable energies in recent years is to be contrasted with the predominant place of fossil fuels.

For these reasons, studying links between oil, clean energies and technology over time is important for companies and governments wanting to invest and implement policies in the clean energy sector. It is accepted that the rise of oil price positively impacts the stock market performance of alternative energy companies. We are going to study the relationship of the stock market performance of alternative energy companies relative to oil prices. To properly identify this effect, we must control for other explanatory factors.

2 Literature review

Interest in the field began with the emergence of alternative energies in the late 2000's. **Henriques and Sadorsky (2008)** were the first to study the relationship between alternative energy stock prices and oil prices. Since, this subject of research is rapidly growing and branches out. Indeed, we have noticed several dimensions in this literature. There is a time dimension of the observations which is weekly in most of the papers but there are also daily observations for some of them. Moreover, while the majority of the papers are considering the USA, few choose to study Europe or Oecd countries or even smaller economies such as single countries. These differences in geographical area and in time can be interesting to consider for portfolio management and for implementation of policies.

As said before, one of the most important research conducted in this area was by **Sadorsky and Henriques (2008)**. This paper aims to perform an empirical analysis on the relationship between oil prices and the stock prices of renewable energy companies in the United States of America. They consider weekly data over the period 2001-2007. They used a vector autoregression (VAR) model. The paper shows that the linkages between technology and clean energy equities are stronger than with oil prices. However, at the time when the paper was published, the prospects of renewable energy being adopted as a main energy source were perceived to be lesser, and the situation has changed in the recent years.
A more recent study by **Managi and Okimoto (2013)** extended the paper of **Sadorsky and Henriques (2008)** and they showed once again that oil and clean energy are always positively correlated after structural breaks in the oil market. They added to the data considered by **Sadorsky (2008)** the US treasury bill interest rate and extended the period studied to 2010. The results indicate that there was a structural change in late 2007, a period in which there was a significant increase in the price of oil. They identified a positive relation between clean energies prices and oil prices after structural breaks. Moreover they found out a similarity in terms of the market response to both clean energy stock prices and technology stock prices. These first two papers suggested the presence of a substitution effect to explain the positive link between oil prices and clean energy stock prices.
Another similar study by **Schmitz (2009)** investigates the relationship between oil prices and alternative energy indices. The results indicate that alternative energy index returns are sensitive to changes in broad market returns and oil price returns. Specifically, an increase or decrease in the broad market or oil prices will lead to a corresponding increase or decrease in the alternative energy stocks.

A study by **Bondia et al.,(2016)** goes against previous studies and lingers on the influence of oil and technology equity on clean energy in the long run. They also consider the USA and the same period as **Managi and Okimoto(2013)**. They implemented threshold cointegration tests which allow to take into account possible endogenous regime shift of variables in the long run. Their work confirms that ignoring the presence of structural breaks in a long run series data can produce fallacious results. They showed that the influence of oil and technology equity on clean energy disappears in the long run and the growth and development of the clean energy market is independent of the oil price or technology equity.
But **Reboredo et al. (2017)** studied causality between oil prices and renewable energy stock prices using wavelet at multiple horizons in the 2006-2015 period. They found that dependence between the two variables strengthened towards the long run, i.e the opposite result shown in the study by Bondia.
The study of **Abdoh et al. (2019)** covers the 2001-2008 period in the USA and considers daily data. They combined wavelets and multivariate GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) and also found significant return transfer from oil prices and technology market to alternative energy index.

Finally, while all papers presented in this review focused on USA, **Liu and Hamori (2020)** studied differences in spillovers to renewable energy stocks both in USA and Europe using General Forecast Error Variance Decomposition (GFEVD) over the 2003-2019 period. They found that spillovers of return from all variables (crude oil, financial market index, 10-year government bond and gas futures) to renewable energy stock markets are higher in US than in Europe.
Overall, this literature review shows that there is no consensus among economists about the relationship between clean energy prices and oil prices. However, most of them conclude that there is strong correlation between stock return of renewable energy prices and high technology firms. From past studies we expect to find a positive and significant relationship between renewable energies and oil prices but also with technology stock prices. However, based on the **Bondia et al.(2016)** study we may not find any relationship between renewable energies stock prices and oil price or technology equity.

3 Data

3.1 Data presentation

Our dependent variable (Y) is **the European Renewable Energy Index**[1] called ERIX index which was launched in 2005 by Societe Generale. It tracks the performance of the largest stocks in the European renewable energy sector: biomass, solar energy, geothermal energy such as Siemens Gamesa and Vestas for instance. The weight is allocated based on market capitalization.

Since we aim at testing how sensitive the stock market performance of alternative energy companies are to changes in oil prices, the first explanatory variable that we consider is the **Brent index** (BRENT) [2] : Brent is a type of crude oil used as a standard in oil pricing for Europe, Africa and middle east. The price of Brent is based on a barrel which capacity is 159 liters. The evolution of Brent reflects the evolution of oil prices in Europe. And considering the substitution effect it is expected that when oil price increases then it should have a positive impact on the stock price of alternative energy companies. Because an increase in oil encourages agents to find substitution energies.

We also consider other explanatory variables that may affect the ERIX index :

1. **Technology sector index** (STK) (SPDR® MSCI Europe Technology UCITS ETF). This index is representative of all European equities belonging mainly to the information technology sector. Indeed, investors may see alternative energy companies as similar to other technology companies because in both cases the investment in research and development is massive, these are two main emerging sectors where there is a lot to discover and many people are interested to invest in. Then an higher price of technological stocks should have a positive impact on the alternative energy stocks prices.

2. **Interest rate** (EURIBOR): To represent the interest rate in our model we would like to use a treasury bond that seems more adapted. The European intervention rate doesn't seem to be adapted because it is 0 since 2016 so it will have no impact on our variable because it has been constant for 4 years. That is why we choose the Euribor at 3 months.

[1] We warmly thank Prof. Ugolini, from the Sate University of Rio de Janeiro for having provided us with the data about the ERIX index.

[2] The ERIX, BRENT, and STK indices are total return indices. It means that these are indices that consider that profits, dividends, and financial surpluses will be reinvested over time, and thus measures the capital gains realised by the shares. A return index is considered more accurate because it does not exclude redistributions, it provides a better representation of the financial performance of the underlying companies.

A lot of research has highlighted that interest rates are explaining a large part of stock price movements. Investors require a higher return for taking on extra risk by investing in stocks instead of Treasury bonds, which are guaranteed to pay a certain return. The extra return that investors can theoretically expect from stocks is referred to as the "risk premium". Following this idea, higher interest rates should have a negative impact on our alternative energy stock prices as with every stock.

3. **Exchange rate between euro and us dollar** (EURUSD): European Union has imported 87 % of it's oil in 2017. Crude oil is mainly quoted in dollars since 85% of the amount imported in Europe is paid in dollars (Eurostat). If the rate increases, oil becomes more expensive to import so it should affect positively the ERIX index, by substitution.

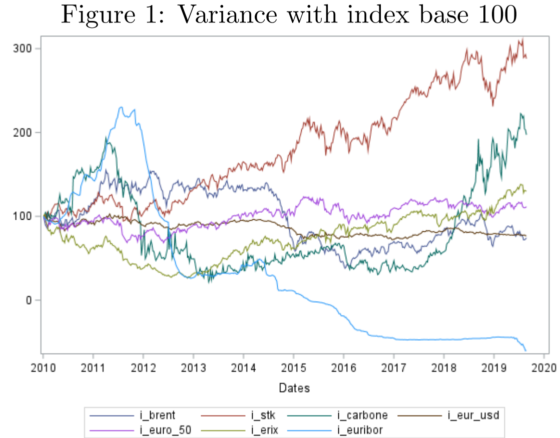
4. **Carbon emission futures by ton** (CARBONE) : We choose to use carbon price by ton in our model because an increase of the price of the ton of carbon gives incentives to invest in decarbonated energy. So a higher price of the ton should induces a positive impact on ERIX index. Here we have worked with futures which are financial instruments used to anticipate future changes in an underlying asset. You buy or sell a given quantity of the underlying asset, at a maturity date and price known in advance.

5. **Euro Stoxx 50** (EURO_50) [3] : The Euro Stoxx 50 is an index introduced in 1998 and made up of the 50 largest European stocks. It is an ideal benchmark to be used with the ERIX index. It helps us to explain variation of ERIX index independently of the evolution of the European market. Thus an increase of this index should have a positive impact on the dependant variable.

3.2 Data description

We describe the evolution of every variable for our panel going from 2010 to 2019. This description and first analysis are based on graphics that show simple correlations which may be quite different from the impact of a variable on ERIX in an econometric model, as it will be explained in a following section.
We would like to compare each variable with another, but since we have indexes, stock prices and an exchange rate, we cannot compare these time series in absolute terms. Then, a way to normalize these series is to index them by giving the same starting point of 100 in 2010. Such manipulation allows to compare the series and keeps the same percentage changes from an observation to the following one as in the non-indexed series.

Figure 1: Variance with index base 100



[3] We have checked that there are no common companies between those of the high technology index and those of the renewable index and European largest stocks. Then we avoid multicollinearity in our data.

1. First we analyse and compare the ERIX evolution with the EURO STOXX 50[4]. There has been a downfall of the ERIX index following the 2008 crisis until 2012. Since, the index only goes up, it finally reaches the pre-crisis level in 2017 and keeps rising. Compared with the most capitalised firms in Europe, the ERIX index does not have any major breakdowns since 2012 while there are two breakdowns on the EURO STOXX 50, this index seems to have cycle evolution every 5 years.
2. We analyse and compare the technology index evolution with the ERIX. The STK index has not stopped to increase since 2010 even if there are some downs but which do not seem to be significant on global development. Moreover we notice that since 2014 both indexes on renewable energies and technology have a very similar growth.
3. We analyse and compare the oil price index evolution with the ERIX. There is a huge decrease in oil price from 2014 to 2016. It is due to an overabundant offer linked to the American production which has taken off thanks to non-conventional hydrocarbons, and a falling demand linked among other things to the slowdown in world growth. This creates an imbalance likely to cause a collapse in prices. It graphically seems to have a negative correlation between the two indexes from 2010 to 2016, then the correlation seems to become positive.
4. We analyse and compare the European interest rate evolution with the ERIX. Since 2012 the trend is downward, and the rate is negative since 2016 and reaches its minimum in 2020. It's due to the sovereign debt crisis in Europe and since 2012 the European central bank uses more unconventional instruments to regulate the currency in circulation. The correlation is clearly negative between the stock price renewable energies and the interest rate. In other words, investors prefer to invest in the stock market than in treasury bonds.
5. We analyse and compare the exchange rate between the Euro and US dollar evolution with the ERIX. There is a sharp decrease of the rate around 2015. It is due to the Greek crisis in Europe and the loss of trust of investors in Europe. Moreover the dollar has more value since the United States of America were in a good economic movement with a high growth and a low unemployment. The fact that USD was above the value of Euro made the rate decrease. The exchange rate evolution can be linked with the oil price evolution, they have the same very sharp decrease in 2014. It is due to the fact that the dollar becomes much stronger than the Euro because of the gain in fossil energies from the United States of America.
6. Finally, we analyse and compare the price of carbon per ton evolution with the ERIX. It has two main situations on the carbon market. A sharp decrease around 2012 once again is linked to the sovereign debt crisis. After this event, the price was capped to 10€/ton by 2018. At this time, Europe established a stability reserve. This mechanism makes it possible to automatically modulate the quantity of allowances auctioned according to the quantity of allowances in circulation, and to withdraw a significant number of allowances. The purpose is to lead firms to speed up the energy transition from non-renewable to renewable. Since the adoption of the stability reserve, the carbon price per ton and the stock price of renewable energies seem to be positively correlated.

[4]All individual plots are page 24

As our geographical area is Europe, the sovereign debt crisis which happened in the beginning of the 2010's is an important shock for our financial data. Indeed, it seems to be a structural break (i.e a significant change of the parameters over time) in the data at the end of 2012. We operate a Chow test (Table 9) on the two sub periods 2010-2013 (150 weekly observations) and 2013-2019. So the Chow test consists in operating a regression for each of the two sub-periods, and test if the parameters are significantly different. We may face endogeneity in our model meaning that this test is maybe biased but we will ignore it to our present study

The hypothesis of Chow test, a model such as $y_t = a + bx_{1t} + cx_{2t} + \epsilon$. We suspect a structural break so we transform it into two model such as

$$y_t = a_1 + b_1 x_{1t} + c_1 x_{2t} + \epsilon \tag{1}$$

$$y_t = a_2 + b_2 x_{1t} + c_2 x_{2t} + \epsilon \tag{2}$$

The null hypothesis induces that $a_1 = a_2$, $b_1 = b_2$, and $c_1 = c_2$ i.e. there is no break and if the null hypothesis is rejected it induces a structural break. The test follows a Fisher law.

The Fisher Statistic is significant at the 1% threshold (Annexe: Table 14). Then as we want to understand the impact of oil on the ERIX index, we choose to discard the 2010-2013 sub-periods because of its high volatility thus we will discuss our work for the periods going from 2013 to 2019.

Finally, since we have financial data, we study descriptive statistics over weekly return.
The mean represents the weekly average return in percentages for each variable and the standard deviation represents the risk associated to each variable over the period considered. Investing in carbon is the most profitable in average over the period as the mean return is 0.8% but also the riskiest asset considering the standard deviation.
Looking at Skewness and Kurtosis gives information about the distribution of each financial variable. The more the Skewness is close to zero, the more the distribution is symmetric and the higher is the Kurtosis the more the distribution is centered around the mean. Our sample has quite symmetric distributions, except the STK which is less symmetric since its skewness is closer from 1 than 0, their distributions have a "tail" on the right. We also notice that ERIX and BRENT have opposite returns in average, the BRENT has a negative return since the ERIX has a positive return over the period.

Table 1: Weekly returns statistics

Variable	Min	Mean	Median	Max	Standard error	Skewness	Kurtosis
ERIX	-13.4	0.5	0.7	11.5	0.2	-0.465	1.643
BRENT	-12.7	-0.1	0.0	15.5	0.2	0.046	1.515
STK	-10.4	0.2	0.6	6.7	0.1	-0.681	1.244
CARBONE	-39.3	0.8	0.7	44.5	0.5	0.449	3.397
EUR_USD	-3.5	-0.0	0.0	4.7	0.1	0.225	1.484
EURO_50	-7.2	0.1	0.3	6.0	0.1	-0.316	0.221

4 Estimation

4.1 Linear model

4.1.1 Specification of the linear model

We denote our model as a log-log linear regression. As we have indexes and stocks, taking the logarithm of variables allows us to conveniently interpret the coefficients in terms of elasticities.
Our initial model is the following :

$$log(ERIX_t) = \alpha+\beta_1 log(BRENT_t)+\beta_2 log(EUR\_USD_t)+\beta_3 log(CARBONE_t)+\beta_4 log(STK_t)+\beta_5(EURIBOR_t)$$
$$+\beta_6 log(EURO\_50_t) + \epsilon_t$$

Where ERIX is the closure price of the European renewable energy index. BRENT the closure price of oil, EUR_USD the exchange rate between the Euro and US dollar. Then CARBON is the carbon ton price and STK the closure price for the technological index. Finally, we have the interest rate (EURIBOR) and EURO_50 the Euro Stoxx 50 that represents the European market evolution.
We make the following assumptions :

- H1 : The model is correctly specified
- H2 : $E(U|X) = 0$
- H3 : $V(U|X) = \sigma^2 I$
- H4 : There is no strict multicollinearity between the explanatory variables.
  The matrix $X'X$ is of full rank and invertible
- H5: $cov(X, U) = 0$
- H6: $p\lim_{t\to\infty} \frac{X'X}{T} = Q_{xx}$ a positive definite matrix
- H7: $p\lim_{t\to\infty} \frac{X'U}{T} = 0$
- H8: $p\lim_{t\to\infty} \frac{Z'U}{T} = 0$
- H9: $p\lim_{t\to\infty} \frac{Z'X}{T} = Q_{zx}$ a matrix of full rank
- H10: $p\lim_{t\to\infty} \frac{Z'Z}{T} = Q_{zz}$ a positive definite matrix

Where X refers to the matrix of explanatory variables. Z refers to the matrix of instrumented variables and U refers to the vector of the disturbance term.

4.1.2 Ordinary least square estimator

We assume the hypotheses we made are satisfied, we neglect endogeneity and heteroskedasticity or autocorrelation. We perform OLS to have a benchmark and watch if we don't have strange results.

A first estimation using Ordinary Least Square (OLS) is presented (Table 2), first looking at the t-Student test, we notice that all coefficients are significant i.e. statistically different from zero. However, the exchange rate between Euro and US dollar has a very high estimated coefficient and the oil price coefficient is negative while it was expected to be positive.
Putting aside the oil price coefficient, we get signs that we were expecting for the rest of the variables. Indeed, the STK, the Euro stoxx 50, the Euro-dollar exchange rate and the price of carbon impact positively the ERIX while Euribor negatively affects the ERIX.

We introduced a crossed variable between the Brent and the Euro dollar exchange rate (log(BRENT) * log(EUR_USD)) to try to isolate the indirect effect of the exchange rate by the Brent on clean energy index. Then, a second estimation is proposed (Annexe: Table 15), with the cross variable between Brent and the Euro dollar exchange rate and without the exchange rate, because we suppose that the cross variable should better integrate how the exchange rate impacts the estimation of the ERIX.
However, it appears that the coefficient values stay around the same estimation, variables are still significant and keep the same sign, even the oil prices variable that we expected to be positive. Since the introduction of the cross variable does not have the expected impact in our model, we will stay with our first specification, with the exchange rate and without the crossed variable.

If we put aside the exchange rate, from this first estimation with this model (Table 2) we can notice that the variable on the Euro stoxx 50 has the most important influence on the ERIX. An increase of one percent on the index of the fifty most capitalised European firms has an impact of 0.74 percent increase on the ERIX price. Since this variable represents the market condition in Europe, it is an expected result. In accordance with literature, we found that the technology index (STK) has also a large impact on renewable energy stocks index (ERIX). Indeed, an increase of 1 percent of the STK has an impact of 0.6 percent on ERIX.

Table 2: First OLS estimation

Variables	Estimate value	t	p-value
Intercept	-0.58	-0.62	0.53
l_BRENT	-0.36	-7.80	< 0.01
l_EUR_USD	1.09	5.46	< 0.01
EURIBOR	-0.31	-2.48	< 0.01
l_STK	0.60	3.65	< 0.01
l_CARBONE	0.11	4.66	< 0.01
l_EURO_50	0.74	4.39	< 0.01

We can already test the hypothesis H3 of our model, i.e test if we have heteroskedasticity and/or serial correlation. Since we work with time series data, we suspect our variables to be correlated with a lagged version of itself so that residuals in period t are correlated with residuals from previous periods $t - j$.
We operate a **Breush-Godfrey test** on our model estimate by OLS (Table 3), the test is such as we take H0 the rho coefficient is equal to zero and H1 the coefficient is different from zero for the following estimation: $\epsilon_t = \rho\epsilon_{t-1} + u_t$

The result of this test gives us a probability inferior to 0.001 for the Lagrange multiplier which means that we reject H0, then the estimate rho coefficient is statistically different from zero and there is serial correlation, as expected.

Moreover, we can test if the error term has heteroskedasticity. We implement a **Breusch-Pagan** test, it tests the H0 hypothesis that the error variances are all equal versus the alternative that the error variances are a multiplicative function of one or more variables. We reject H0 so our estimation by OLS gives heteroskedastic estimation which induces biased tests as the matrices of variance covariance is not properly computed.

Table 3: OLS test (1)

Test	LM	p-value
Breusch-Godfrey	301.1	< 0.01
Breusch-Pagan	84.12	< 0.01
White	188.4	< 0.01

At the end of this first part with estimating the model with ordinary least squares we have first estimations that appear to be biased because we have presence of heteroskedasticity and serial correlation which biased our variance then our test of significance. But the main problem is the suspicion of endogeneity in our regressors, we suspect simultaneity issue between the BRENT as the ERIX because they are both energy energy index. We will change the estimation method to try to improve our estimations.

4.1.3 Dealing with endogeneity: Instrumental variables and Generalized method of moments

To use IV and GMM estimator, we first have to define exogenous and endogenous variables of our model. The dependent variable i.e. the price stock of renewable energies is an endogenous variable to the model since this is the variable we try to explain.
Since the ERIX and BRENT are both energy stock prices we suspect that there is endogeneity because it might have a relationship in both directions (the evolution of these variables due to the fact that they are two subparts of energy prices in general). Thus in this model we suspect the oil price to be endogenous. We assume other variables to be exogenous to the model.

We have to look for instruments that allows to reduce endogeneity. First we took a lag variable on the ERIX and BRENT i.e. the endogenous variables. We test the IV estimation and it gives heteroskedastic and serial correlated estimations.
We produced an Hansen test (because of the heteroskedasticity) and it rejects the H0 hypothesis that is : the error term is uncorrelated with the exogenous variables if instruments are valid. In other words, instruments were endogenous so can't be used. (Annex: Table 16)

Then we decided to take single lag values from exogenous variables and to test it again. From Breusch-Pagan and Godfrey test the estimations are still heteroskedastics and serial correlated but this time the Hansen test allows to accept H0 so that instruments are well exogenous. (Annex: Table 17)

The next step is to test instrument relevance. We want to test that the endogenous variables and the instruments are sufficiently correlated (ie, test H9). If not, instruments are weak and they induce biased estimation by Instrumental Variables. And then test on this estimation are also biased since they do not have the correct size.
Our residuals are not independent and identically distributed (iid) but for simplicity of testing weak instruments, we will assume that they are. We regress endogenous variables over exogenous and instruments variables and then we use the decision rule of **Staiger-Stock (1997)** over the Fisher statistics.
We test:

H0 : instruments are not sufficiently correlated, weak instruments
H1 : instruments are sufficiently correlated, strong instruments

If Q_f is over 10 we reject H0 and instruments are not weak and if Q_f is under 10 we accept H0 and instruments are weak.
Testing our instruments give us a Q_f statistics under 10 (Annex: Table 18). Then our instruments are weak and we have to find other ones.

Finally, from **Henriques Sadorsky (2008)** one of the lag that allows for better estimation is a lag of 15 weeks. They use the **Schwarz information criterion**[5] to select the lag length, this criterion is an information criterion derived from the **Akaike information criterion** proposed by **Gideon Schwarz in 1978.** The penalty depends on the sample size and not only on the number of parameters.
Henriques and Sadorsky considered weekly data over the period 2001-2007 in the USA. Despite the fact we study the European case in a different period, we have a similar number of observations on weekly data too, we choose to apply their findings to our cases.

Based on this we try new instruments that are lag variables from exogenous variables but with 15 lags against only one in the previous estimation.
Moreover, it seems better to keep a single lag on the EURO 50 variables since this is the market conditions representation, the financial world integer very quickly changes in these variables, it is not necessary to have a large lag over it.
So, as instruments, we choose STK, Carbon and Euro-dollar exchange rate with 15 lags and only one lag for the Euro 50 variable.

We recompute the IV estimator and test it again. We find heteroskedasticity and serial correlated estimation but this time the Hansen test allows to accept H0 and the weak test induces a Q_f statistic over 10 so we reject H0 and instruments are not weak. .

Table 4: Tests on IV estimator (3)

Test	Statistic	p-value
Breusch-Pagan	49.41	< 0.01
Breusch-Godfrey	275.9	< 0.01
Hansen	4.74	0.19

Table 5: Tests weak instruments (2)

Variable	Q_f	Statistic
l_BRENT	13.83	$Q_f > 10$

As we face an (unknown) form of heteroskestacity[6], we will use **Generalized method of moments (GMM)**. Indeed, as we have a large sample, GMM brings an advantage of consistency in the presence of heteroskedasticity. Moreover we use the **Bartlett Kernel** to produce a covariance matrix robust to heteroskedasticity and autocorrelation. This allows us to construct asymptotically valid tests if the sample is large, which is the case in this model that has 330 observations.

We produce a GMM model with the same instruments as with IV. We test the exogeneity of our initial endogenous variables to check if the use of GMM model allows to reduce the endogeneity of the model. (Table 6)

[5] Also name the Bayesian information criterion
[6] In other words, disturbances are not independent and identically distributed

Table 6: GMM estimator

Variables	Estimate value	t	p-value
Intercept	3.14	2.34	0.02
l_BRENT	-0.64	-4.17	< 0.01
l_EUR_USD	1.97	3.72	< 0.01
l_STK	0.75	3.29	< 0.01
l_CARBONE	0.12	4.86	< 0.01
L_EURO_50	0.34	1.56	0.12
L_EURIBOR	-0.14	-0.81	0.42

As we take into account endogeneity, heteroskedasticity and serial correlation to estimate our model by GMM, it is the best estimation we can provide. As we have a large sample and endogenous variables, we can say that IV and GMM are consistent but GMM is more efficient.

From this new estimation the first important information is that the interest rate variable becomes non significant, it is more logical than with precedent estimation since the interest rate is under the threshold of 0 since 2016 and therefore has limited impact on the financial market.

The variables that have an impact on the stock prices of renewable energy are the oil prices (BRENT), the technological index (STK), the exchange rate (EUR_USD) and the carbon ton prices (CARBONE). Then, we notice that the market condition coefficient (EUR_50) keeps a positive sign but becomes not significant. However, the technology market index (STK) also keeps a positive sign but is significant, it seems that the dependant variable (ERIX) has a stronger relation with the technological market than with the European market's biggest firms.

First an increase of 1% of the oil prices induces a decrease of 0.64% of the ERIX index. It means that at the opposite of what we expected, an increase in oil price does not influence investors to invest in renewable area but at the opposite it pushes them to concentrate their investment in oil. We are aware that the state of the art is to use models which allow us to study shocks on variables and the dynamics of time series, however we use a linear model which reflects long term behavior and that is maybe why we find that for our period the substitutability hypothesis is not verified.
Then an increase of 1% in the technological index induces an increase of 0.75% of the dependent variable. As expected and shown in **Sadorsky and Henriques (2008)**, the linkages between technology and clean energy equities are stronger than with oil prices. Investors assimilated clean renewable energies with research in technology because these are two main emerging sectors. A possible interpretation is that the renewable energy sector is not yet perceived as a mature sector so the progress in technology has a major impact on development and stock prices of alternative energy companies.

Finally, an increase of 1% in carbon ton prices induces an increase of 0.12% of the Erix index. It confirms the idea that an increase in the price of the ton of carbon gives incentives to invest in decarbonated energy.

We test the exogeneity of our regressors and find out that the vector of variable that we suspect to be endogenous is still endogenous and the other vector is exogenous. In this situation we can identify that estimations by OLS are inconsistent and so are estimation by FGLS. But IV and GMM estimates are consistent and since we have non independant and identical disturbances, GMM estimator is the best.

4.2 Autoregressive model

4.2.1 Specifiaction of the autoregressive model

To try to treat the serial correlation in another way, we made up a new model which is autoregressive. To do so we add a regressor which is a lag value of the dependent variable, the ERIX index.

$$log(ERIX_t) = \alpha + \beta_1 log(BRENT_t) + \beta_2 log(EUR\_USD_t) + \beta_3 log(CARBONE_t) + \beta_4 log(STK_t) + \beta_5 (EURIBOR_t)$$
$$+ \beta_6 log(EURO\_50_t) + \beta_7 log(ERIX_{t-1}) + \epsilon_t$$

This model should allow to take into account the link with previous observation so that it will no more appear in the disturbance term.
We make the following assumptions :

- H1 : The model is correctly specified
- H2 : $E(U|X) \neq 0$
- H3 : $V(U|X) = \sigma^2 I$
- H4 : There is no strict multicollinearity between the explanatory variables.
  The matrix $X'X$ is of full rank and invertible
- H5: $cov(X,U) = 0$
- H6: $p \lim\limits_{t\to\infty} \frac{X'X}{T} = Q_{xx}$ a positive definite matrix
- H7: $p \lim\limits_{t\to\infty} \frac{X'U}{T} = 0$
- H8: $p \lim\limits_{t\to\infty} \frac{Z'U}{T} = 0$
- H9:$p \lim\limits_{t\to\infty} \frac{Z'X}{T} = Q_{zx}$ a matrix of full rank
- H10: $p \lim\limits_{t\to\infty} \frac{Z'Z}{T} = Q_{zz}$ a positive definite matrix

4.2.2 Ordinary least square estimator

We assume the hypotheses we made are satisfied, we neglect endogeneity and heteroskedasticity. We perform OLS to have a benchmark and watch if we don't have strange results.

We estimate this new model always through OLS (Table 7). This new specification of the model induces changes in the estimation of coefficients. From the analysis of the results we can deduce that the oil price (BRENT), the carbon ton price (CARBONE) and the index of the fifty most capitalised European firms variables (EURO_50) becomes no more significantly different from zero. It is comprehensible that the euro stoxx 50 becomes non significant since it should represent the market evolution, but adding a lag value allows to have the past market evolution included in our dependant index.
However, the technological index (STK) remained statistically significant while crude oil was not. This is an appropriate estimation since some papers **(Bondia et al.,(2016))** also found that oil prices and renewable energies stocks are independant in the long run.

Then the lag value is significant with 1% of errors and the interest rate is significant with 5% of errors. It is common that the lag value capture a large part of the explanatory power.

Table 7: OLS estimation: Autoregressive model

Variables	Estimate value	t	p-value
Intercept	0.18	0.68	0.49
l_BRENT	0.00	0.09	0.93
l_EUR_USD	-0.09	-1.54	0.12
l_STK	0.13	2.84	< 0.01
l_CARBONE	0.00	-1.10	0.27
L_EURO_50	-0.03	-0.77	0.44
L_EURIBOR	0.7	1.94	0.05
L_LAG_ERIX	0.95	62.33	< 0.01

We proceed a Breusch-Godfrey test on this autoregressive model estimated by OLS, to study if the introduction of a lag value allows to reduce the autocorrelation.
This test give us a Lagrange multiplier of 2.16, since this follows a chi-two law, our statistics is under the threshold and we can accept the H0 hypothesis and conclude that there is no autocorrelation in this estimation of the autoregressive model.
Since we do not have autocorrelation anymore we can figure out if the estimation is homoskedastic. If the estimation is homoskedastic then the variance of all disturbances terms is constant and uniform, otherwise variance diverges and there is heteroskedasticity.
By the Breusch-Pagan and White test we denote H0 the hypothesis where variance of disturbances are identical and H1 variance of disturbances are not identical. Here we reject the null hypothesis because we reject the H0 hypothesis, i.e. there is heteroskedasticity in this autoregressive model estimate by OLS.

Table 8: OLS test (1)

Test	LM	p-value
Breusch-Godfrey	2.16	0.14
Breusch-Pagan	27.67	< 0.01
White	99.37	< 0.01

This model estimate by Ordinary Lest Square is no more serial correlated but has always heteroskedasticity, we could corrected it by computed better covariance/variance matrix. However this model has at least two endogenous variables, the BRENT as we suspect it previously and the lag value of the dependent variable is by construction endogenous. Then we changed the method of estimation to correct the biases induced by endogenous variables.

4.2.3 Dealing with endogeneity: Instrumental variables and Generalized method of moments

To use IV and GMM estimator, we first have to define exogenous and endogenous of our model. As previously, the dependent variable i.e. the price stock of renewable energies is an endogenous variable to the model and so is the price stock of oil. Then by definition, the regressor that is built with a lag on the independent variable should also be endogenous. Thus in this model we suspect two regressors to be endogenous: the BRENT and the LAG_ERIX. We assume the other variables to be exogenous to the model.

As previously with the linear model, we use a lag of 15 in our instrumental variables to estimate our model. We then choose as instruments the lag of 15 on the STK variable, the CARBONE variable, the EUR_USD and finally the LAG_ERIX. Moreover, it seems better to keep a single lag on the EURO_50 variable since this is a market conditions representation, the financial world integer very quickly changes in this variable, it is not necessary to have a large lag over it.

We produce an estimation by the instrumental variable method and we test our instruments to study if they are well exogenous and strong.
We find a statistic of test equal to 4.36 for the Hansen test and as previously we regress endogenous variables over exogenous and instruments variables and then we use the decision rule of **Staiger-Stock (1997)** over the Fisher statistics.[7] We found out that our instruments are valid and strong.

Table 9: Tests on IV estimator ()

Test	Statistic	p-value
Breusch-Pagan	27.21	< 0.01
Breusch-Godfrey	18.03	< 0.01
Hansen	4.36	0.23

Table 10: Tests weak instruments (2)

Variable	Q_f	Statistic
l_BRENT	21.59	$Q_f > 10$
l_LAG_ERIX	144.10	$Q_f > 10$

We reject H0 for both tests so we still have heteroskedasticity and serial correlation but the serial correlation is much less strong than before.
As we still face an unknown form of heteroskedasticity we estimate the model by GMM. As we have a large sample, GMM brings an advantage of consistency in the presence of heteroskedasticity. Moreover we use the Bartlett Kernel to produce a covariance matrix robust to heteroskedasticity and autocorrelation.

However, the estimation by GMM with correction of Bartlett on the autoregressive model induces that only the lag value stays significant. The regressor built on a lag of the explanatory variable i.e. the ERIX index on log, captures all the explanatory power.
In other words, the specification of our model should not be optimal.

Table 11: GMM estimation: Autoregressive model

Variables	Estimate value	t	p-value
Intercept	0.25	0.68	0.50
l_BRENT	0.03	0.58	0.56
l_EUR_USD	-0.18	-1.01	0.31
l_STK	0.10	1.25	0.22
l_CARBONE	-0.01	-1.33	0.18
L_EURO_50	-0.09	-1.54	0.12
L_EURIBOR	0.7	1.67	0.09
L_LAG_ERIX	0.99	20.96	< 0.01

[7] Our residuals are not independent and identically distributed (iid) but for simplicity of testing weak instruments, we will assume that they are.

We test the exogeneity of our regressors and find out that the vector of variable that we suspect to be endogenous is still endogenous and the other vector is exogenous. In this situation we can identify that estimations by OLS are inconsistent and so are estimation by FGLS. But IV and GMM estimates are consistent and since we have non independant and identical disturbances, GMM estimator is, one more time, the best estimatore for the autoregressive model.

4.3 Overview of the different estimators

For both model, we produced an overview of the differences estimators. IV and GMM provide relatively similar results as they both face the endogeneity issue, but GMM offers better results as it takes into account the fact that disturbances are not iid and moreover it can take into account the serial correlation with Bart option. OLS and FGLS are inconsistent FGLS should allow to have a matrix of variance-covariance robust to heteroskedasticity and perform unbiased tests. But, as we have endogeneity and serial correlation, FGLS is not a valid estimator and we see it well on the estimations it gives.

Table 12: Overview of estimations on the linear model

		OLS			**FGLS**		
Variable	Coefficient	t	p-value	Variable	Coefficient	t	p-value
Intercept	-0.59	-0.62	0.53	Intercept	-26.73	-1.66	0.09
l_BRENT	-0.36	-7.80	< 0.01	l_BRENT	2.29	2.94	< 0.01
l_EUR_USD	1.08	5.46	< 0.01	l_EUR_USD	3.91	1.16	0.25
l_STK	0.60	3.65	< 0.01	l_STK	-2.93	-1.05	0.29
l_CARBONE	0.11	4.66	< 0.01	l_CARBONE	-0.22	-0.56	0.57
l_EURO_50	0.74	4.39	< 0.01	l_EURO_50	2.78	0.97	0.33
EURIBOR	-0.31	-2.48	< 0.01	EURIBOR	-4.46	-2.12	< 0.01
		IV			**GMM**		
Variable	Coefficient	t	p-value	Variable	Coefficient	t	p-value
Intercept	3.51	3.25	< 0.01	Intercept	3.13	2.34	0.02
l_BRENT	-0.71	-5.74	< 0.01	l_BRENT	-0.64	-4.17	< 0.01
l_EUR_USD	2.17	5.65	< 0.01	l_EUR_USD	1.97	3.72	< 0.01
l_STK	0.86	4.21	< 0.01	l_STK	0.75	3.29	< 0.01
l_CARBONE	0.13	5.20	< 0.01	l_CARBONE	0.12	4.86	< 0.01
l_EURO_50	0.27	1.48	0.14	l_EURO_50	0.34	1.56	0.12
EURIBOR	-0.04	-0.22	0.82	EURIBOR	-0.15	-0.81	0.42

Table 13: Overview of estimations on the autoregressive model

		OLS			FGLS		
Variable	Coefficient	t	p-value	Variable	Coefficient	t	p-value
Intercept	0.17	0.68	0.50	Intercept	0.18	0.68	0.49
l_BRENT	0.00	0.09	0.93	l_BRENT	0.00	0.09	0.93
l_EUR_USD	-0.09	-1.54	0.12	l_EUR_USD	-0.09	-1.54	0.12
l_STK	0.13	2.84	< 0.01	l_STK	0.13	2.84	< 0.01
l_CARBONE	-0.00	-1.10	0.27	l_CARBONE	-0.01	-1.10	0.27
l_EURO_50	-0.30	-0.77	0.44	l_EURO_50	-0.03	-0.77	0.44
EURIBOR	0.07	1.94	0.05	EURIBOR	0.07	1.94	0.05
l_LAG_ERIX	0.95	62.33	< 0.01	l_LAG_ERIX	0.95	62.33	< 0.01
		IV			**GMM**		
Variable	Coefficient	t	p-value	Variable	Coefficient	t	p-value
Intercept	0.26	0.70	0.48	Intercept	0.25	0.68	0.5
l_BRENT	0.02	0.41	0.68	l_BRENT	0.03	0.58	0.56
l_EUR_USD	-0.16	-0.86	0.39	l_EUR_USD	-0.18	-1.01	0.31
l_STK	0.11	1.41	0.16	l_STK	0.10	1.25	0.21
l_CARBONE	-0.01	-1.37	0.17	l_CARBONE	-0.01	-1.33	0.18
l_EURO_50	-0.08	-1.57	0.12	l_EURO_50	-0.09	-1.54	0.12
EURIBOR	0.08	1.91	0.06	EURIBOR	0.07	1.67	0.09
l_LAG_ERIX	0.98	19.75	< 0.01	l_LAG_ERIX	0.99	20.96	< 0.01

The principal interpretation for the estimations of the Autoregressive model is that the introduction of the lagged value capture all the explanatory power, which is a classical issue in econometric.

5 Conclusion

With the increasing global environmental pollution and energy crisis, investment in the renewable energy sector has become a concern for investors. Moreover, the unprecedented growth of the clean energy sector over the last decade which is expected to continue in the coming years have generated a great deal of interest in understanding the relationship between renewable energy stocks prices and crude oil price.

In this study, we investigated the relationship between the stock market performance of renewable energy and oil crude price in Europe over the 2013-2019 period using a linear model. As we detected an endogenous variable in our model, the OLS and FGLS estimates are biased and inconsistent. We also identified heteroskedasticity and serial correlation in our error term which led us to prefer the GMM estimates with Bart option over the IV's estimates. Indeed, in this case, GMM estimates are found to be unbiased and asymptotically more efficient than IV's estimates. Concerning the results, we found a negative sign for the brent coefficient, which is the opposite of what we expected to find. But the european technology index (STK) largely impacts the renewable energy stocks performance, which is in line with the literature.

Then, we introduced an autoregressive model to treat the serial correlation by adding a lagged variable of the Erix index to the previous model. We found that a large part of the explanatory power has been captured by the lagged variable which is a common issue.

Finally, we are aware that the state of the art is to use models which allow us to study shocks on variables and the dynamics of time series as GARCH model[8], however we use a linear model which reflects long term behavior.

[8] See our estimation in annexe: "Autoregressive conditional heteroskedasticity using the GARCH specification"

6 References

Baum C., Schaffer M. (2003). "Instrumental variables and GMM: estimation and testing." The Stata Journal 1, 1-31.

Broadstock, D.C., Cao, H., Zhang, D. (2012). "Oil shocks and their impact on energy related stocks in China." Energy Econ. 34, 1888–1895.

Henriques, I., Sadorsky, P. (2008). "Oil prices and the stock prices of alternative energy companies." Energy Econ. 30, 998–1010.

Huang et al. (2013). "Conditional jump dynamics in the stock prices of alternative energy companies." International journal of energy economics and policy, 3, 288-296.

Juan C. Reboredo, Ugolini A. (2018). " The impact of energy prices on clean energy stock prices. A multivariate quantile dependence approach." Energy Econ. 76, 136–152.

Kiefer N.,Vogelsang T. (2001). "Heteroskedasticity-Autocorrelation robust standard errors using the Bartlett kernel without truncation."

Kumar, S., Managi, S.,Matsuda, A. (2012). "Stock prices of clean energy firms, oil and carbon markets: a vector autoregressive analysis." Energy Econ. 34, 215–226.

Kumar S., Fujii H. and Managi S. (2015). "Substitute or Complement? Assessing Renewable and Nonrenewable Energy in OECD Countries" Applied economics 47, 1438-1459

Liu.T Hamori.S (2020). "Spillovers to Renewable Energy Stocks in the US and Europe: Are they different?" MDPI

Managi, S., Okimoto, T. (2013). "Does the price of oil interact with clean energy prices in the stock market?" Jpn. World Econ. 27, 1–9.

Ravichandran A., Nayak S. (2019). "Impact of crude oil prices on renewable energy stocks in india." JETIR volume 6, issue 6

Schmitz A. (2009). "Effect of Oil Prices on returns to alternative energy investments, thesis defensed in the Georgia Institute of Technology."

7 Appendices

7.1 Table of contents

Contents

7.2 Sources of data

- **ERIX :** We sent emails to several searchers that worked on our topic because we can't find historical data about the ERIX index, except on Bloomberg but we haven't a licence to log in.
  Then Andrea Ugolini, professor at the state university of Rio de Janeiro sent us a database extract from Bloomberg, we warmly thank him for allow us to work with these data.

- **BRENT :** Yahoo the 23/09/2020
  Available at, https://cutt.ly/oh1sjlm

- **EURIBOR :** Banque de France the 23/09/2020
  Available at, https://www.banque-france.fr/statistiques/taux-et-cours/taux-interbancaires

- **CARBONE :** Investing the 06/10/2020
  Available at, https://fr.investing.com/commodities/carbon-emissions-historical-data

- **EURO_50 :** Yahoo the 06/10/2020
  Available at, https://fr.investing.com/etfs/spdr-msci-europe-info-tech-historical-data

- **STK :** Investing website the 06/10/2020
  Available at, https://fr.investing.com/etfs/spdr-msci-europe-info-tech-historical-data

- **EUR_USD :** Investing the 06/10/2020
  Available at, https://www.investing.com/currencies/eur-usd-historical-data

7.3 Table annexes

Table 14: Chow test

Breaking point	DDL num	DDL den	F-value	p-value
150	7	463	50.35	< 0.01

Table 15: OLS estimation with the BRENT_RATE variable and without the EUR_USD variable

Variables	Estimate value	t	p-value
Intercept	-0.69	-0.74	0.46
l_BRENT	-0.43	-8.52	< 0.01
l_BRENT_RATE	0.29	6.37	< 0.01
l_STK	0.56	3.46	< 0.01
l_CARBONE	0.12	5.07	< 0.01
l_EURO_50	0.80	4.79	< 0.01
EURIBOR	-0.36	-2.92	< 0.01

Table 16: Tests on IV estimator (1)

Test	Statistic	p-value
Breusch-Pagan	81.24	< 0.01
Breusch-Godfrey	297.5	< 0.01
Hansen	44.7	< 0.01

Table 17: Tests on IV estimator (2)

Test	Statistic	p-value
Breusch-Pagan	37.24	< 0.01
Breusch-Godfrey	292.7	< 0.01
Hansen	2.72	0.60

Table 18: Tests weak instruments (1)

Variable	Q_f	Statistic
l_BRENT	2.89	$Q_f < 10$

7.4 Autoregressive conditional heteroskedasticity using the GARCH specification

Table 19: GARCH estimation

Variables	Estimate value	t	p-value
Intercept	2.57	4.16	< 0.01
l_BRENT	-0.00	-0.02	0.98
l_BRENT_RATE	0.06	0.39	0.70
l_STK	0.79	13.52	< 0.01
l_CARBONE	0.04	2.30	< 0.05
l_EURO_50	0.05	0.75	0.45
EURIBOR	-0.33	-1.65	0.09
AR 1	-0.83	-11.36	< 0.01
AR 2	-0.16	-2.31	< 0.05
ARCH 0	0.00	1.29	0.20
ARCH 1	0.06	0.85	0.39
ARCH 2	0.03	0.43	0.67
ARCH 3	-0.00	-0.00	0.99
ARCH 4	0.03	0.55	0.58
ARCH 5	0.04	0.63	0.53
ARCH 6	0.04	0.59	0.55
ARCH 7	-0.00	-0.00	0.99
ARCH 8	0.06	0.92	0.36
GARCH 1	0.00	0.00	0.99

We have tried to estimate a GARCH model by following SAS documentation and other resources. ARCH models are introduced by Engle (1982) and allow to take into account time-varying volatilities. As we saw it in literature review, the majority of papers use this kind of models. A better estimation could be produced by using PROC AUTOREG on SAS software, the procedure estimates and forecasts linear regression models for time series data when the errors are autocorrelated or/and heteroscedastic. The autoregressive error model is used to correct for autocorrelation, and the generalized autoregressive conditional heteroscedasticity (GARCH) model and its variants are used to model and correct for heteroscedasticity.
When time series data are used in regression analysis, often the error term is not independent through time. Instead, the errors are serially correlated as we show it with our financial data. If the error term is autocorrelated, the efficiency of ordinary least squares (OLS) parameter estimates is adversely affected and standard error estimates are biased, as proved in the previous sections.

We will not discuss the results here as we are not sure about how to make proper interpretations. But we can summarize our methodology to estimate the model which can be found in the SAS document. Firstly, we have test for autocorrelation using the Durbin-Watson test. We found to have first order autocorrelation. This test is not appropriate to find the autoregressive order as test for higher order assume to have absence of lower order autocorrelation. Instead, we can use the stepwise regression. We choose to use an initial order of 5 lags. Only lags of order 1 and 2 remain significant. So we stay with an AR(2). Then we test for heteroskedasticity using arch test. The Q statistics was significant for changes in variance across time using lag windows ranging from 1 through 12. Then the test strongly indicates heteroskedasticity. The test is also supposed to help determining the order of the ARCH model, here we just can say that the order q is high. Then we estimated a model AR(2)-GARCH(q=8, p=1). We choose $p > 0$ to allows a long memory process.

7.5 Graphic annexes