ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis – Finance

The predictive ability of investors attention on the Cryptocurrency market activity.

Abstract

This paper studies the contemporary and predictive relationship between investors attention, as measured by Google's searches, and three cryptocurrency market activities. The findings suggest that there are a contemporary relationship and predictive ability of investors attention on cryptocurrency trading volume and volatility. Furthermore, the finding that there is no contemporary nor predictive relationship between investors attention and market returns suggests that the cryptocurrency market is efficient.

Keywords: *Cryptocurrency, market efficiency, investor attention, Google searches*

Michel L.E. Vrutaal
365900

July 19th, 2018

Supervisor: ***Dr. Ran Xing***
Second assessor: –

Table of Contents

1. Introduction

"You cannot go anywhere in the financial world or financial press without seeing a mention of Bitcoin", says Jim Paulsen (Lim, 2018). Bitcoin, the first of its kind, is the cryptocurrency with the largest market capitalization, known by its creator as a peer-to-peer version of electronic cash (Nakamoto, 2009). The importance of understanding the cryptocurrency market grows more significant with the growing market capitalization and investor's interest.

Previous research shows that Google's search volume index (SVI) is an objective way to reveal and quantify the aggregate interest of investors (Da, Engelberg, & Gao, 2011). Multiple studies expand on this and show that SVI is significantly positively related to historical and implied volatility, and trading volume (2012). This is also evident in the FX markets for the contemporary and predictive relationship (2015). Furthermore, Vozlyublennaia (2014) demonstrates that attention does influence the performance of indexes of stocks, bonds, and commodities. However, as cryptocurrencies are a relatively new commodity, this field is still relatively unexplored.

This study is determined to see if the relationship between investors attention and market activity is also existent in the cryptocurrency market. This will help fill the gap and give insights into the (informational) efficiency of the cryptocurrency market. Hence the following research question arises: ***What is the relationship between an investor's attention and the Cryptocurrency market activity, specifically Bitcoin & Ethereum, during the period 2013 – 2018?***

First, the correlation between the market activity variables, return, volatility, and trading volume are analyzed. Then, the contemporary and predictive relationship between SVI and the three variables of cryptocurrency market activity is studied.

The findings suggest that (1) investors attention is positively associated with the contemporaneous cryptocurrency market volatility and trading volume as previous studies have shown in other markets. Furthermore, (2) investors attention can be used to predict market volatility and trading volume to make investment decisions. However, (3) there is no contemporary or predictive relationship between investors attention and market returns. These findings add upon previous studies and show a similarity in behavior of cryptocurrency market and other markets.

The rest of the paper is organized in the following sections. Section 2, the theoretical framework, introduces the relevant literature for financial market efficiency, Google Trends SVI, and cryptocurrency. In section 3, the data used for this research paper is presented, and the methodology is described. Section 4 presents the results and the implications of our findings. Section 5 summarizes and concludes.

2. Theoretical Framework

Whether or not it is possible to use available information to predict market movements, and what the predictability implies for our understanding of financial markets is a widely disputed financial area. The efficient market hypothesis (EMH) states that a capital market is efficient when prices "fully reflect" all available, and relevant information (Fama, 1970). This implies that prices follow a random path as the occurrence of new information is also random (Fama, 1965). However, recent research has started to use data from information that can capture investors sentiment and attention. These are data from Google Trend SVI which is a proposed direct measure to reveal investors attention (Da, Engelberg, & Gao, 2011), and online news interaction which shows the effect of investor's sentiment on market activity (Tetlock, 2007).

Baker & Wurgler (2006) show that investor's sentiment has a significant effect on securities whose valuations are highly subjective. This shows that investors sentiment will drive a commodity that is easily traded due to no restrictions or low transaction cost. The total attention of investors, independent from their sentiment, is shown to have predictive power (Da, Engelberg, & Gao, 2011).

Lastly, Moat et al. present evidence that suggests the viewed frequency of financially related Wikipedia pages may have contained early signs of stock market movements. The majority of Wikipedia page views arrives through Google searches; this implies that Wikipedia shows only a part of the total investors attention and interest that can be captured through Google Trend.

2.1 Google Trend, investor's attention, and market activity

Google Trend (GT) is an online database where you can access different type of search volume (SVI) data that has been accumulated by Google's search engine over the years. Based on Statista, Google has held and continues to hold the largest online search engine market share of more than 85% for the past years (Statista, 2018). Hence, one can assume that this majority of users can be used as a representative data sample.

In recent years Google has been improving their algorithm to show the most relevant results based on a user's searches. This is possible through the use of multiple factors surrounding the use of keywords. The search volume of a keyword shows how much interest a specific topic is generating, and this is a representation of the aggregate user's attention towards that topic. This aligns with the claim of Da et al. (2011) that GT SVI is a direct measure of investors attention. Multiple studies build upon this claim by showing the relation between investor's attention and a variety of (capital) market activity.

2.1.1 Relationship between investors attention and market activity

The relationship between investors attention (SVI) and (capital) market activity has begun to become more evident as more studies are conducted. There seems to be an association between investors attention and each specific market activity measures, such as an asset's price, return, volatility, and trading volume. Moreover, Preis et al. (2013) find patterns that may be interpreted as "early warning signs" of stock market moves. Hence, it can also be said that there is a predictive relationship between investors attention and the market measures.

Firstly, the relationship between investors attention and asset price and return is considered. Da et al. (2011) show that an increase in investors attention predicts higher stock prices in the following two weeks but price reversal within a year. This suggests that one can use current information to construct a trading strategy that yields profitable returns. This is confirmed by Bijl et al. (2016), who shows that Google searches can be utilized in profitable strategies before transactions costs.

Second, the relationship between investors attention and market volatility is considered. Goddard et al. (2015) conclude that investors attention commoves with contemporaneous FX market volatility and predict subsequent FX market volatility. Additionally, in the US stock market, a heightened number of current Google searches is followed by an increase in market volatility the next day (Dimpfl & Jank, 2016). Evidence from the Norwegian stock market also indicates that an increase in Google searches predicts increased volatility, and increased trading volume. The predictive power of google searches is also suggested in Preis et al. (2010) by recurring patterns.

Lastly, the relationship between investors attention and trading volume is considered. As mentioned above by Dimpfl & Jank (2016), the strong relation between online searches and trading volume is also documented by Takeda and Wakao (2014) in the Japanese stock market. Aouadi et al. (2013) also shows that investors attention is correlated to trading volume, and determines stock market illiquidity and market volatility in the French market. Bank et al. (2011) reinforce the notion that an increase in investors attention is associated with a rise in trading activity, stock liquidity, and a temporarily higher future returns.

To sum it all up, multiple studies conducted in a variety of markets, ranging from the stock and FX market focusing on U.S, Norwegian, European, Asian markets, and more, show similar evidence of the relationship between investors attention and market activity. The positive relation between investors attention and the asset's return, volatility, and trading activity are shown. Furthermore, in some cases, investors attention seems to have predictive power over the market activity. While there are similarities between the stock market, the FX market, and cryptocurrency market, it is possible for them to have different associations with investors attention. Hence, the effect of investors attention on an upcoming market, such as the cryptocurrency market, is of importance to better understand its market activity.

2.1.2 Understanding the cryptocurrency market

While the effect and predictive power of Google searches are most commonly studied in the stock market, Google searches have also been used for other commodities and in other industries. In the Indian market, a presence of bidirectional causality between investors attention and the gold spot price is found, along with effects on the equity and exchange rate markets (Jain & Biswal, 2018). Similarly, Vozlyublennaia (2014) demonstrates that attention does influence performance of indexes of stocks, bonds, and commodities. Furthermore, Smith (2012) concludes that investor's attention has predictive power beyond the GARCH for foreign currencies, suggesting future studies to test this predictive power in the stock, bond, and commodities markets. Hence, the assumption is made that investor attention also influences the market indexes for cryptocurrencies.

In comparison to the stock market, the cryptocurrency market is relatively new and unexplored. The cryptocurrency market has broken through its all-time high and grew past an $800 billion market capitalization in January 2018 (Marshall, 2018). It is still widely debated, whether cryptocurrency is security, commodity, currency, or just a speculative digital asset. However, one can not deny the fact that a significant amount of market capitalization attracts investors for its potential for higher returns compared to the stock market. The most recent example being EOS, a cryptocurrency, that had a 193% price increase in April 2018 (Godbole, 2018). Cryptocurrency price is entirely determined by (fixed) supply that can not be altered once its made, and demand for the cryptocurrency (Nakamoto, 2009). Suggesting that any information regarding the underlying company's technology or service that reflects the added value of the currency should have a direct influence on the price through its demand. This study aims to find any association and predictive power of investors attention on cryptocurrency market activity.

2.2 What is cryptocurrency?

Considering, that capital markets have considerable importance for economic growth, it is essential that the cryptocurrency market is also studied. Especially, since cryptocurrency is not restricted to one country, but accessible worldwide (granted you have internet access). The innovative technology behind cryptocurrency has the potential to disrupt the current payment system, monetary system, and is already affecting a variety of country's policy.

The SEC chairman made a statement saying that cryptocurrencies like Bitcoin are not securities (CNBC, 2018), this was expanded on by SEC officials also confirming Ethereum as a non-security (Pisani, 2018). This study does not aim to discuss the categorization of cryptocurrency. However, in this study cryptocurrency will be assumed as the name suggests, a currency. The label is not of importance. However, it is essential that a distinction is made between the difference between the most popular cryptocurrencies and their underlying company technology/service. This distinction will make the results of this study easier to interpret.

2.2.1 Cryptocurrencies, Bitcoin, and Ethereum

As of June 2018, there are over 1600 cryptocurrencies (CoinMarketCap, 2018). Each cryptocurrency coin is required (most of the time) to use its respective underlying company's service. Hence, the more service is used, the higher the demand for that specific currency. Since most of the cryptocurrency has a fixed supply amount, an increase in demand will result in a price (denoted in US$) increase. This price increase can make a popular cryptocurrency attractive to hoard now, and sell at a later stage at a higher price. It is a popular strategy to make a profit in the cryptocurrency market. However, a price increase is not guaranteed. Hence, Selgin (2014), and Baeck and Elbeck (2014) argue that Bitcoin should be seen as a speculative commodity. Most studies of cryptocurrency omit over 1600 other cryptocurrencies, also known as Altcoins, focusing mainly on Bitcoin, and other cryptocurrencies with large market capitalization.

Bitcoin is the first cryptocurrency, which also holds the most significant market capitalization making it a good representative for the cryptocurrency market activity (ignoring the usage difference between cryptocurrencies). Bitcoin is a peer-to-peer version of electronic cash that allows direct online transactions without the need for a third party (Nakamoto, 2009). However, it seems that Bitcoin is mostly used, and considered as a digital version of gold for other cryptocurrencies, allegedly due to its name popularity. Halaburda and Gandal (2014) find that when Bitcoin becomes more valuable to the U.S. dollar, it also becomes more valuable to other cryptocurrencies. Additionally, the average monthly volatility of Bitcoin is higher than that for gold or a set of foreign currencies (Dwyer, 2015). This volatility brings forth an opportunity to make a profit if one can find a pattern or use historical information to predict market movement. Briere et al. (2013) show that Bitcoin highly distinctive features, including a high average return and volatility, offers significant diversification benefits.

Further studies on Bitcoin has focused mainly on the price discovery and efficiency under the EMH. A recent study by Urquhart (2016) concluded that Bitcoin is an inefficient market but may be in the process of moving towards an efficient market. A follow-up study does show that a power transformation of Bitcoin returns can be weakly efficient throughout the entire period used by Urquhart (Nadarajah & Chu, 2017). However, Bariviera also shows that daily return time series become more efficient across time, but is currently inefficient. From a variety of different tests, contradictory results are found. This can be partially blamed on the lack of data, as this is still a relatively new capital market as proposed by Urquhart.

While Bitcoin is a decentralized online payment service, Ethereum, the second largest cryptocurrency, has a different end goal. Ethereum is a decentralized platform for applications that run exactly as programmed without any chance of fraud, censorship or third-party interference (Ethereum Foundation, 2018). In other words, it is an infrastructure where other cryptocurrencies are built upon. There is a final competition between cryptocurrency. However, a cryptocurrency with multiple other cryptocurrencies built on top of it is sure to last longer. This distinction between cryptocurrency is of importance because infrastructure cryptocurrency might be a better representation of the cryptocurrency market.

Lastly, it is not possible to directly buy all cryptocurrency. It is easier to buy popular cryptocurrencies like Bitcoin and Ethereum and use these to trade for Altcoins.

2.3 Relevance of the cryptocurrency market

As this new innovative market emerges grabbing attention for mass adoption, one can only speculate, however, it remains a fact that the services provided by this new technology have the potential to disrupt the existing payment system and monetary system. It is of importance that such a market is efficient, and the market activity is better understood. Fama (1965) states that prices follow a random path as the occurrence of new information is also random. This implies that, for an efficient market, historical information should have no predictive power. Furthermore, studies so far show that Bitcoin is moving towards a more efficient market (Urquhart, 2016). This aligns with the EMH that there should be no predictive power from historical information.

However, it is believed that the mass adoption of cryptocurrencies will happen as it cryptocurrency becomes easier to use, and people become more aware. The awareness of people, or their attention, can be measured by Google searches. As discussed in section 2.1.1 multiple studies have shown that using the historical information of Google searches seems to have some predictive power. This contradicts the EMH, which is one of the critical cornerstones of finance. To the extent of my knowledge, the only similar study was conducted by Kristoufek, where he shows that investors attention and bitcoin prices are connected. This study builds upon this by researching the predictive power of investors attention on Bitcoin and Ethereum with more recent data. Hence, the following research question arises: ***What is the relationship between investors attention and the Cryptocurrency market activity, specifically Bitcoin & Ethereum, during the period 2013 – 2018?***

As mentioned before, this study is interested in finding evidence that shows that the cryptocurrency market is informationally efficient. This implies that Google searches data does not have any predictive power over the cryptocurrency market activity. The cryptocurrency market is divided into three activities, namely return, trading volume, and volatility. Hence, the following main hypotheses are formulated and tested:

1. **There is no statistically significant contemporary relationship between investors attention and cryptocurrency market activity.**
2. **There is no statistically significant predictive relationship between investors attention and cryptocurrency market activity.**

Lastly, to clear up these hypotheses, Google searches will be used as a measure for investors attention. Furthermore, this hypothesis is divided into three different sub-hypotheses. Focusing mainly on the trading volume, market return, and market volatility of Bitcoin and Ethereum. The sub-hypotheses are as followed:

1. ***There is no statistically contemporary/predictive relationship between Google searches and the returns of cryptocurrencies.***
2. ***There is no statistically contemporary/predictive relationship between Google searches and the market volatility of cryptocurrencies.***
3. ***There is no statistically predictive relationship between Google searches and the trading volume of cryptocurrencies.***

Rejecting these hypotheses suggests that historical information can be used to beat the cryptocurrency market systematically. Hence, this will indicate that the cryptocurrency market is inefficient in its weak-form.

3. Data & Methodology

To conduct this research, a set of data is required. To measure the investors attention for the cryptocurrency, the search volume data (SVI) for the keywords “Bitcoin”, and “Ethereum are used. The cryptocurrency prices, trading volume, and market capitalization are obtained from CoinMarketCap. Moreover, the returns and volatility of these cryptocurrencies are calculated. Next, a variety of regressions are run to explain the relationship between investors attention and the cryptocurrency market activity. A descriptive and predictive regression will be run. The following sub-sections discuss in details the acquiring and processing of the data, and the methodology used to run the tests.

3.1 Google Trend Data

The search volume of specific keywords are obtained from Google Trends at https://trends.google.com/trends/. The keyword “Bitcoin” is used to show the user’s attention towards the cryptocurrency Bitcoin, and the keyword “Ethereum” for the cryptocurrency Ethereum. It has been a conscious choice to omit other search terms like BTC for bitcoin, and ETH for Ethereum, since this can have a double effect.

First, it is important to note that the used SVI data will be raw. Hence, the variety of Google Trend filter options will have an impact on the relative normalized search volume. The data is normalized so that the highest value is 100, which represents the timestamp with the highest interests based on the selected options. So, the highest value might be on a different timestamp if a different time range, or option is set.

Secondly, from the variety filter options provided by Google Trend the following were chosen to acquire this research’s required data. The geographical filter is put on “worldwide” to match the global accessible nature of cryptocurrencies. Next, category is set on “all categories” to capture all interests in the keyword and topic. Additionally, GT differentiate the search words in different terms. For simplicity, this study only focuses on “search term” for all keywords. Then, the search type is set on *web searches*, which is the most popular method used to find information. However, this does not include the searches on Youtube, which seems to also be quite a popular platform. Hence, the data for the search type *youtube searches* is also exported, separately.

Lastly, the most important filter option, the time range, has to be set. This has been set based on the available data of the cryptocurrency market. For Bitcoin the sample period from May 31st, 2013 through May 31st, 2018 is studied, due to lack of data before the year 2013. This sample period might also be better, as Urquhart (2016) states that around this period is where Bitcoin starts becoming more efficient.

Next, for Ethereum the sample period is set on August 8th, 2015 through May 31st, 2018. This is since Ethereum's launch date. In Table 1 the variables obtained by Google Trend is shown. These are all independent variables.

Table 1: Summarized table of independent variables acquired from Google Trend. These variables are a measure of investor's attention.

Independent Variables	***Definitions***
BTCweb	Web search volume for Bitcoin
BTCyt	Youtube search volume for Bitcoin
ETHweb	Web search volume for Ethereum
ETHyt	Youtube search volume for Ethereum

3.2 Cryptocurrency Data

The cryptocurrency market information is obtained through CoinMarketCap (https://coinmarketcap.com/). The data consists of the daily market capitalization, trading volume, and prices corresponding their respective dates for Bitcoin and Ethereum. While the cryptocurrency market is open for trade 24/7, CoinMarketCap still provides the closing time price. This is the price used to conduct this study.

Furthermore, this daily data has to be converted to weekly data, to match the Google Trend search volume which is given on a weekly basis. The process is a simple average of these daily data. To match the specific weeks of Google Trend, the sum is taken off the daily cryptocurrency data for that respective week and then divided by 7. This provides the weekly data of prices, market capitalization, and trading volume for Ethereum and Bitcoin.

Next, two important dependent variables have yet to be calculated. Market returns and volatility. The returns for Bitcoin and Ethereum are computed directly from the weekly closed prices with the following general formula:

$$R_t = Log\left(\frac{P_t}{P_{t-1}}\right) X\ 100 \qquad (1)$$

Where R_t is the return of Bitcoin *($btcR_t$)*/Ethereum *($ethR_t$)*, P_t and P_{t-1} are the natural Bitcoin/Ethereum prices at time *t* and *t -1*.

The last financial variable that has to be calculated is the volatility. This study includes volatility as a control variable in the regression model explaining returns and volume, and also as a measure for market activity. The volatility is measured utilizing the Garman and Klass (1980) volatility estimator, however we omit the adjusted close price. Hence, the formula is as follow:

$$Variance_t = \frac{1}{2}(h_t - l_t)^2 - (2\,Log\,2 - 1)c_t^2 \qquad (2)$$

where

$$c_t = log(close_t) - log(open_t),$$
$$l_t = log(low_t) - log(open_t),$$
$$h_t = log(high_t) - log(open_t).$$

Then, with $Variance_t$ calculated, the weekly volatility can be calculated as a square root of average daily variance.

$$Volatility_t = \sqrt{\frac{1}{|S_t|}\sum_{i \in S_t} Variance_t} \qquad (3)$$

Lastly, there are a total of 150 observations for Ethereum and 264 observations for bitcoin. However, for Bitcoin, due to a lack of data for trading volume these observations are omitted. This results in a total of 232 observations for Bitcoin (trading volume).

Table 2: Summarized table of dependent financial variables.

Dependent Variable	***Definition***
btcR	Returns of Bitcoin
btcTV	Trading Volume of Bitcoin
btcVol	Volatility of Bitcoin
ethR	Returns of Ethereum
ethTV	Trading Volume of Ethereum
ethVol	Volatility of Ethereum

The dependent financial variables that will be used in the regression are shown above in Table 2. A predictive regression is run in order to test the proposed sub-hypotheses in section 2.3.

3.3 Methodology

This study investigates whether investors attention towards Bitcoin *(BTCweb, BTCyt)* and Ethereum *(ETHweb, ETHyt)*, measured by Google searches, can explain or predict Bitcoin/Ethereum return *(btcR / ethR*), trading volume *(btcTV / ethVol),* and market volatility *(btcVol / ethVol)* utilizing descriptive and predictive regressions. Additionally, the market capitalization for Bitcoin and Ethereum *(btcMC / ethMC)* is analyzed. In descriptive regressions, the independent variable and the dependent variables occur at the same time. Whereas, in the predictive regression, a lagged version of the independent variable is used to investigate whether historical information can be used to predict the three dependent variables (future returns, volatility, trading volume). In both regression models, the respective lagged dependent variable, and control variables are added. All lagged variables are set at a one week lag with t being one week. Additionally, volatility is used as a control variable, since previous studies have shown that volatility has a positive relationship with the return in the stock market (French, Schwert, & Stambaugh, 1987). This study assumes this also to be the case for the cryptocurrency market.

3.3.1 Initial analysis of variables

First, the relationship between variables is analyzed. In search of a significant correlation between independent variables and dependent variables, to add these variables to a regression. Additionally, the relationship between two independent variables is also analyzed to find if there is any significant multicollinearity. This would make one of the independent variables redundant in a multiple regression. The correlation between variables and their p-values are based on the Pearson correlation. The descriptive and predictive regression models are built based on these correlations for Bitcoin and Ethereum.

Bitcoin

Table 3: Summarized Pearson correlations between Bitcoin variables

	BTCweb	*BTCyt*	*btcR*	*btcVol*	*btcTV*	*btcMC*
BTCweb	1	.905**	.041	.457**	.881**	.880**
BTCyt	.905**	1	.079	.377**	.784**	.818**
btcR	.041	.079	1	-.035	-.026	-.012
btcVol	.457**	.377**	-.035	1	.339**	.319**
btcTV	.881**	.784**	-.026	.339**	1	.970**
btcMC	.880*	.818**	-.012	.319**	.970**	1

**. Correlation is significant at the 0.01 level (2-tailed).

As shown above, in Table 3, the two independent variables that measure investors attention, BTCweb and BTCyt, are strongly correlated. Hence, it is opted only to use the variable that better explains the dependent variables to avoid multicollinearity. The variable will be decided in section 4.1. Furthermore, it is opted to omit the regression with Bitcoin returns as dependent variables, due to no significant correlations with the independent variables. This leads to the following generalized descriptive (4) and predictive (5) regression models for Bitcoin:

$$btc\gamma_t = \alpha_i + \beta_1 BTC\chi_t + \beta_2 btc\gamma_{t-1} + \beta btcControls_t + \epsilon_t \qquad (4)$$
$$btc\gamma_t = \alpha_i + \beta_1 BTC\chi_{t-1} + \beta_2 btc\gamma_{t-1} + \beta btcControls_{t-1} + \epsilon_t \qquad (5)$$

where $btc\gamma_t$ is the dependent variable *btcR, btcTV,* and *btcVol* at time *t*, $BTC\chi_t$ is the independent variable *BTCweb*, and *BTCy* at time *t,* and $btcControls_t$ are *btcVol* at time *t.* ϵ_t are external factors at time *t.* For descriptive models, βs are the regression coefficients for the respective lagged dependent variable, the respective independent variable, and Bitcoin Controls. For predictive models, βs are the regression coefficients for the respective lagged dependent variable, the respective lagged independent variable, and Bitcoin lagged controls.

Ethereum

Table 4: Summarized Pearson correlations between Ethereum variables

	ETHweb	*ETHyt*	*ethTV*	*ethMC*	*ethR*	*ethVol*
ETweb	1	.812**	.816**	.780**	.116**	.442**
ETHyt	.812**	1	.502**	.483**	.135*	.443**
ethTV	.816**	.502**	1	.950**	.040	.358**
ethMC	.780**	.483**	.950**	1	-.021	.324**
ethR	.116	.135*	.040	-.021	1	.309**
ethVol	.442**	.443**	.358**	.324**	.309**	1

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

As shown above, in Table 4, the two independent variables that measure investors attention, BTCweb and BTCyt, are strongly correlated. Hence, it is opted only to use the variable that better explains the dependent variables to avoid multicollinearity. The variable will be decided in section 4.2. This leads to the following generalized descriptive (6) and predictive (7) regression models for Ethereum:

$$eth\gamma_t = \alpha_i + \beta_1 ETH\chi_t + \beta_2 eth\gamma_{t-1} + \beta ethControls_t + \epsilon_t \qquad (6)$$
$$eth\gamma_t = \alpha_i + \beta_1 ETH\chi_{t-1} + \beta_2 eth\gamma_{t-1} + \beta ethControls_{t-1} + \epsilon_t \qquad (7)$$

where $eth\gamma_t$ is the dependent variable *ethR, ethTV,* and *ethVol* at time *t*, $ETH\chi_t$ is the independent variable *ETHweb*, and *ETHyt* at time *t,* and $ethControls_t$ are *ethVol* at time *t.* ϵ_t are external factors at time *t.* For descriptive models, βs are the regression coefficients for the respective lagged dependent variable, the respective independent variable, and Ethereum controls. For predictive models, βs are the regression coefficients for the respective lagged dependent variable, the respective lagged independent variable, and Ethereum lagged controls.

3.3.2 Linear regression models

The initial regression model is run only with the independent variables. Based on the significance level, R Square (adjusted), and standard error of estimation, one of the independent variables are chosen to be used in the further regression models.

Bitcoin

Table 5: Representation of the best regression model for Bitcoin

DV	F	p-value	R-sq.adj	Std. Err	BTCweb	BTCyt
btcVol	68.349	< .001	.206	.01312302	Yes	No
btcTV	895.87	< .001	.775	1,411milj	Yes	No
btcR	1.630	> .10	.002	.003	No	Yes

As shown above, in Table 5, *BTCweb* seems to better explain the two dependent variables, *btcVol* and *btcTV,* in comparison to BTCyt. In both cases it also better than if you include BTCweb and BTCyt in the model (ignoring multicollinearity). The complete results can be found in paragraph 3.1 of the Appendix.

Ethereum

Table 6: Representation of the best regression model for Ethereum

DV	F	p-value	R-sq.adj	Std. Err	ETHweb	ETHyt
ethVol	63.378	< .001	.193	.02149	No	Yes
ethTV	516.89	< .001	.665	750 mln	Yes	No
ethR	4.793	.029	.014	.89813	No	Yes

As shown above, in Table 6, *ETHweb* seems to better explain *ethTV*, whilst *ETHyt* seems to better explain *ethVol.* In both cases, we omit the other variable as to avoid multicollinearity. Furthermore, the model for *ethR* does not seem to be significant. The complete results can be found in paragraph 3.1 of the Appendix.

4. Empirical Findings and Results

In this section, the null hypothesis is considered. First, the sub-hypotheses are tested through the use of a descriptive and predictive regression model. These models study the relationship between investors attention, measured by Google's web searches or youtube searches, on market return, volatility, and trading volume. Then, the results for each model are shown and interpreted.

4.1 Bitcoin Results

First, the results for models (equation 4 and 5) that aim to explain the Bitcoin returns (btcR) are discussed. Table 3 in section 3.3.1 shows that there are no significant correlations between bitcoin returns and investors attention. This is further confirmed when a lagged variable for investors attention and a control variable is added (volatility) to the model.

Table 7: Summarized Regression btcR result based on appendix 7.2.1: Columns (1)-(3) report results from a descriptive regression model. Each column adds a variable to the regression model. The coefficient of the respective model is shown and its explanatory power in the model. Columns (4) and (5) follow the same principle but are predictive models.

Dependent Variable: Bitcoin Returns						
	Descriptive Models			Predictive models		
	(1)	(2)	(3)		(4)	(5)
btcR_lag	.008	-.004	.063	btcR_lag	.005	.016
BTCyt		.004	.003	BTCyt_lag	.002	-.001
btcVol			-3.862	btcVol_lag		8.598
R^2	.000	.006	.01		.001	.024
Adjusted R^2	-.004	-.002	-.001		-.006	.012

**. Significant at the 0.01 level (2-tailed).

Furthermore, the predictive model is also not significant. Hence, sub-hypothesis 1 is accepted. There is no statistically significant predictive relationship between Google searches and Bitcoin returns.

Second, the results for volatility as the dependent variable are shown below in Table 8. Column 1 confirms that Bitcoin volatility is also correlated to the Bitcoin volatility of the previous period. When the model (column 3) is run as given in equation 4, it seems to be the most reliable model. This result shows a stronger positive relationship between Bitcoin volatility and Google searches when bitcoin trading volume is added as the control variable. This can be seen by the coefficients of column 3 being greater than column 2 for BTCweb. As for the predictive models the control variable, btcTV_lag is omitted due to not being significant. However, there is not much difference in the coefficients after having introduced the control variable. Based on the results in column 4, it can be concluded that there is a significant positive relationship between Google searches at time t-1 and Bitcoin volatility at time t. This suggests that there is a statistically predictive relationship. Hence, we reject sub-hypothesis 2.

Table 8: Summarized Regression btcVol result based on appendix 7.2.1: Columns (1)-(3) report results from a descriptive regression model. Each column adds a variable to the regression model. The coefficient of the respective model is shown and its explanatory power in the model. Columns (4) and (5) follow the same principle but are predictive models.

Dependent Variable: Bitcoin Volatility						
	Descriptive Models			Predictive models		
	(1)	(2)	(3)		(4)	(5)
btcVol_lag	.555**	.444**	.437**	btcVol_lag	.462**	.454**
BTCweb		.282**	.497**	BTCweb_lag	.203**	.305**
btcTV		-	-.241**	btcTV_lag		-.111
R^2	.308	.375	.388		.341	.343
Adjusted R^2	.306	.371	.381		.336	.336

**. Significant at the 0.01 level (2-tailed).

Finally, the drivers of the trading volume movements are explored. Table 9 shows the coefficients of each model. Column 1 confirms that Bitcoin trading volume is also correlated to the Bitcoin trading volume of the previous period. When the model is run as given in equation 4, the control variable, volatility, is not significant. Hence, we omit this variable from the equation and come to a more reliable model as shown in column (2) of the descriptive models. This model shows a positive relationship between Google's searches and Bitcoin trading volume. As for the predictive models (equation 5), column (4) shows a more promising result. The results indicate a significant positive relationship between that Google searches at time t-1 and Bitcoin trading volume at time t. This suggests that there is a statistically predictive relationship. Hence, we reject sub-hypothesis 3.

Table 9: Summarized Regression btcTV results based on appendix 7.2.1: Columns (1)-(3) report results from a descriptive regression model. Each column adds a variable to the regression model. The coefficient of the respective model is shown and its explanatory power in the model. Columns (4) and (5) follow the same principle but are predictive models.

Dependent Variable: Bitcoin Trading Volume					
	Descriptive Models		Predictive models		
	(1)	(2)		(3)	(4)
btcTV_lag	.968**	.748**	btcTV_lag	.805**	.792**
BTCweb		.271**	BTCweb_lag	.185**	.217**
btcVol		-	btcVol_lag		-.045**
R^2	.938	.962		.945	.947
Adjusted R^2	.938	.962		.945	.946

**. Significant at the 0.01 level (2-tailed).

The Bitcoin results show that there is a positive relationship between Google searches, as a proxy for investors attention, and bitcoin trading volume and volatility. This seems to be the case in both contemporaneous and lagged models. However, there is no correlation between Google searches and bitcoin returns. Furthermore, the evidence suggests that there is a statistically significant predictive relationship between Google searches and Bitcoin trading volume and volatility. This predictive power seems to be greater as a control variable is added for Bitcoin trading volume.

4.2 Ethereum Results

First, the results for models (equation 6 and 7) that aim to explain the Ethereum returns *(ethR)* are discussed. In section 3.3.2 it is confirmed that there is no correlation between the measures for investors attention and Ethereum returns. This is further confirmed after adding a lagged Ethereum returns variable. While the model is significant, the coefficient for the google searches is not significant.

Table 10: Summarized Regression ethR results based on appendix 7.2.2: Columns (1)-(3) report results from a descriptive regression model. Each column adds a variable to the regression model. The coefficient of the respective model is shown and its explanatory power in the model. Columns (4) and (5) follow the same principle but are predictive models.

Dependent Variable: Ethereum Returns						
	Descriptive Models			Predictive models		
	(1)	(2)	(3)		(4)	(5)
ethR_lag	.125	.102	.05	ethR_lag	.122	.098
ETHyt		.007	-.001	ETHyt_lag	.002	-.001
ethVol			11.361	ethVol_lag		3.569
R^2	.016	.028	.098		.016	.023
Adjusted R^2	.012	.021	.087		.009	.011

**. Significant at the 0.01 level (2-tailed).

Furthermore, the predictive models are not significant. Hence, sub-hypothesis 1 is accepted. There is no statistically significant predictive relationship between Google searches and Ethereum returns.

Second, Ethereum volatility is considered as a dependent variable. The results are shown below in Table 11. Column 1 confirms that Ethereum volatility is also correlated to the Ethereum volatility of the previous period. Ethereum trading volume is not significant when added as the control variable in the descriptive and predictive model. Hence, column 2 shows the stronger model. The result of column 2 shows a stronger positive relationship between Google searches and Ethereum volatility. This can be seen by the coefficient in column 2 is more significant than in column 3 for ETHyt. Additionally, it seems that when the lagged dependent variable is added, ETHweb also has a similar influence on Ethreum volatility as ETHyt. This suggests that Google searches through web searches and youtube searches have a positive relationship with Ethereum volatility. As for the predictive models, column 4 is the stronger model. Based on the results in column 4, it can be concluded that there is a significant positive relationship between Google searches at time t-1 and Ethereum volatility at time t. This suggests that there is a statistically predictive relationship. Hence, we reject sub-hypothesis 2.

Table 11: Summarized Regression ethVol results based on appendix 7.2.2: Columns (1)-(3) report results from a descriptive regression model. Each column adds a variable to the regression model. The coefficient of the respective model is shown and its explanatory power in the model. Columns (4) and (5) follow the same principle but are predictive models.

Dependent Variable: Ethereum Volatility						
	Descriptive Models			Predictive models		
	(1)	(2)	(3)		(4)	(5)
ethVol_lag	.722**	.651**	.639**	ethVol_lag	.677**	.671**
ETHyt		.172**	.132**	ETHyt_lag	.103**	.088
ethTV		-	.090	ethTV_lag		.035
R^2	.521	.546	.552		.530	.531
Adjusted R^2	.520	.542	.547		.526	.525

**. Significant at the 0.01 level (2-tailed).

Finally, the drivers of the Ethereum trading volume movements are explored. Table 12 shows the coefficients of each model. Column 1 confirms that Bitcoin trading volume is also correlated to the Bitcoin trading volume of the previous period. The models in column 2 and 3 are quite similar. However, column 3 is omitted due to the control variable, volatility, not being significant. The model in column 3 shows a positive relationship between Google searches and Ethereum trading volume. As for the predictive model, there also seems to be a positive predictive relationship between Google searches and Ethereum trading volume.

Moreover, when the control variable volatility is added, evidence shows that the relationship is stronger. The results indicate a significant positive relationship between that Google searches at time t-1 and Bitcoin trading volume at time t. This suggests that there is a statistically predictive relationship. Hence, we reject sub-hypothesis 3.

Table 12: Summarized Regression ethTV results based on appendix 7.2.2: Columns (1)-(3) report results from a descriptive regression model. Each column adds a variable to the regression model. The coefficient of the respective model is shown and its explanatory power in the model. Columns (4) and (5) follow the same principle but are predictive models.

Dependent Variable: Ethereum Trading Volume						
	Descriptive Models			Predictive models		
	(1)	(2)	(3)		(4)	(5)
ethTV_lag	.937**	.733**	.733**	ethTV_lag	.821**	.820**
ETHweb		.278**	.277**	ETHweb_lag	.141**	.164**
ethVol		-	.001	ethVol_lag		-.051*

R^2	.878	.913	.913		.884	.886
Adjusted R^2	.877	.913	.912		.884	.885

**. Significant at the 0.01 level (2-tailed). *. Significant at the 0.05 level (2-tailed).

The Ethereum results show that there is a positive relationship between Google searches, as a proxy for investors attention, and Ethereum trading volume and volatility. This seems to be the case in both contemporaneous and lagged models. However, there is no correlation between Google searches and Ethereum returns. Furthermore, the evidence suggests that there is a statistically significant predictive relationship between Google searches and Bitcoin trading volume and volatility. This predictive power seems to be greater as a control variable, Ethereum volatility, is added in the Ethereum trading volume model.

4.3 Results implications on cryptocurrency market

This study aims to explain the relationship between investors attention and the cryptocurrency market activity. Hence, a set of 3 sub-hypotheses are tested for Bitcoin and Ethereum. These are formulated to conclude the main hypothesis: ***There is no statistically significant contemporary/predictive relationship between investors attention and cryptocurrency market activity.***

This study focuses on Bitcoin and Ethereum, the two cryptocurrencies with the largest market capitalization, as it is assumed these can be used as representatives for the rest of cryptocurrencies. Especially, Ethereum, as this is a platform where multiple cryptocurrencies are built on. These cryptocurrencies are based on the ERC20 token of Ethereum and are likely to have a positive relationship with the Ethereum market activity. Similarly, most cryptocurrencies are associated with Bitcoin as it is the means to buy them.

To sum up our findings, based on the results, Google searches can tell us more about future trading activity and volatility of Bitcoin and Ethereum. With web searches being more prominent for Bitcoin market activities and Ethereum trading volume, and youtube searches for Ethereum volatility. Therefore, the same can be concluded for the cryptocurrency market activity as a whole since a majority of cryptocurrencies follow these two major cryptocurrencies. In other words, we reject the main hypothesis that there is no statistically significant predictive relationship between investors attention and cryptocurrency market activity. An investor can use Google information to predict trading volume and volatility. However, it is not possible to predict returns. Hence, unless a different trading strategy is used that can consistently produce returns, through predicting trading volume and volatility by using investors attention it is concluded that the market is efficient. This aligns with the Efficient Market Hypothesis as one can not predict returns.

5. Conclusion

This study aims to investigate the relationship between investors attention, as measured by Google searches, and cryptocurrency market activity. The focus lies mainly on three market activities, return, trading volume, and volatility. Bitcoin and Ethereum are used as representatives for the cryptocurrency market, since these two combined accounts for the majority of total cryptocurrency market capitalization. The data for this research is obtained from CoinMarketCap and Google Trend. Next, the data has also been converted to a weekly basis. Based on this data a linear regression model is built to show the effect of Google searches on return, trading volume, and volatility.

Regarding the returns, there is neither a contemporary relationship nor a predictive relationship between investors attention and Bitcoin or Ethereum returns. However, this result does not necessarily mean that such a relationship does not exist in the cryptocurrency market. This could be caused by the definition of returns in this paper – different trading strategies could define returns differently. On the other hand, investors attention can both explain and predict trading volume and volatility. This indicates that investors can use information from Google Trend in making investment decisions. These results are in accordance with market efficiency. In addition to the main findings, it also concluded that web searches are more related to cryptocurrency market than YouTube searches. However, the difference between the two is small.

To conduct this study, a specific set of limitations were faced. Due to time constraint, the focus lies only on one keyword per cryptocurrency, however, this does not represent all the possible keywords used to search for this topic. Moreover, this keyword does not differentiate between positive and negative user's interests and does not show the user's intent behind the search. A user can only be searching for this topic for entertainment or education purposes instead of investing reasons.

The possibilities for further studies is vast. Market capitalization seems to be correlated to Google searches. The effects of investors attention on the market capitalization could be explored. Additionally, further essential research to conduct is an extension of this study where a distinction is made between negative and positive investors attention. This will make clear if the negative and positive information has an equal effect on market activity. This idea is mainly inspired by the efficient market hypothesis critics. These critics suggest that loss aversion could cause that negative information has a stronger effect on market activity, hence it is possible to take advantage of this to beat the market and make profitable returns.

6. Bibliography

Aouadi, A., Arouri, M., & Teulon, F. (2013, September). Investor attention and stock market activity: Evidence from France. *Economic Modelling, 35*, 674-681.

Baeck, C., & Elbeck, M. (2014, August 26). Bitcoins as an investment or speculative vehicle? A first look. 30-34.

Baker, M., & Wurgler, J. (2006, August 3). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance, 61*(4), 1645-1680.

Bank, M., Larch, M., & Peter, G. (2011). Google Search Volume and its Influence on Liquidity and Returns of German Stocks. *Financial Markets and Portfolio Management, 25*(3), 239-264.

Bariviera, A. F. (2017, December). The inefficiency of Bitcoin revisited: A dynamic approach. *Economic Letters, 161*, 1-4.

Bijl, L., Kringhaug, G., Molnar, P., & Sandvik, E. (2016, May). Google searches and stock returns. *International Review of Financial Analysis, 45*, 150-156.

Briere, M., Oosterlinck, K., & Szafarz, A. (2013, Sep 14). Virtual Currency, Tangible Return: Portfolio Diversification with Bitcoin. *Journal of Asset Management, 16*(6), 365-373.

CNBC. (2018, June 6). *SEC chairman: Cryptocurrencies like bitcoin are not securities*. Retrieved June 14, 2018, from CNBC: https://www.cnbc.com/video/2018/06/06/sec-chairman-cryptocurrencies-like-bitcoin--not-securities.html

CoinMarketCap. (2018, June 14). *Cryptocurrencies By Market Capitalization*. Retrieved June 14, 2018, from CoinMArketCap: https://coinmarketcap.com/

Da, Z., Engelberg, J., & Gao, P. (2011, September 21). In Search of Attention. *The Journal of Finance, 66*(5), 1461-1499.

Dimpfl, T., & Jank, S. (2016, March). Can Internet Search Queries Help to Predict Stock Market Volatility? *European Financial Management, 22*(2), 171-192.

Dwyer, G. P. (2015, April). The Economics of Bitcoin and Similar Private Digital Currencies. *Journal of Financial Stability, 17*, 81-91.

Ethereum Foundation. (2018). *Ethereum*. Retrieved June 15, 2018, from Ethereum: https://www.ethereum.org/

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

Fama, E. F. (1970, May). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25*(2), 383-417.

Fink, C., & Johann, T. (2012, September 1). May I Have Your Attention, Please: The Market Microstructure of Investor Attention. 1-59.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987, September). Expected stock returns and volatility. *Journal of Financial Economics, 19*(1), 3-29.

Garman, M. B., & Klass, M. J. (1980, Jan). On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business, 53*(1), 67-78.

Godbole, O. (2018, May 9). *EOS Looks Poised for a Move Toward Record Highs*. Retrieved from Coindesk: https://www.coindesk.com/eos-looks-poised-for-a-move-toward-record-highs/

Goddard, J., Kita, A., & Wang, Q. (2015, September). Investor attention and FX market volatility. *Journa of International Financial Markets, Institutions and Money, 38*, 79-96.

Halaburda, H., & Gandal, N. (2014, October 6). Competition in the Cryptocurrency Market. *NET Institute Working Paper No. 14-17*, 1-32.

Jain, A., & Biswal, P. C. (2018, May 7). Does internet search interest for gold move the gold spot, stock and exchange rate markets? A study from India. *Resources Policy*.

Kim, N., Lucivjanska, K., Molnar, P., & Villa, R. (2017, November 29). Google searches and stock market activity: evidence from Norway. *Finance Research Letters*.

Kristoufek, L. (2013, December 4). Bitcoin meets Google Trends and Wikipedia: Quantifying the relationshio between phenomena of the internet era. *Scientific Reports 3*.

Lim, P. J. (2018, January 18). *Here's How a Bitcoin Crash Could Bring Down the Entire Stock Market*. Retrieved from Time: http://time.com/money/5105912/bitcoin-bubble-stock-market/

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

Marshall, A. (2018, January 7). *Combined Crypto Market Capitalization Races Past $800 Bln*. Retrieved from Cointelegraph: https://cointelegraph.com/news/combined-crypto-market-capitalization-races-past-800-bln

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013, May 8). Quantifying Wikipedia Usage Patterns Before Stock Market Moves. *Scientific Reports , 3*.

Nadarajah, S., & Chu, J. (2017, January). On the inefficiency of Bitcoin. *Economic Letters, 150*, 6-9.

Nakamoto, S. (2009). *Bitcoin.* Retrieved May 13, 2018, from bitcoin: https://bitcoin.org/bitcoin.pdf

Nambiampurath, R. (2018, March 29). *Analysts are trying to figure out a relationship between cryptocurrency prices and the stock market*. Retrieved from btcmanager: https://btcmanager.com/analysts-are-trying-to-figure-out-a-relationship-between-cryptocurrency-prices-and-the-stock-market/

Peng, L., Xiong, W., & Bollerslev, T. (2007, June). Investor Attention and Time-varying Comovements. *European Financial Management, 13*(3), 394-422.

Pisani, B. (2018, June 14). *Bitcoin and ethereum are not securities, but some cryptocurrencies may be, SEC official says*. Retrieved June 14, 2018, from CNBC: https://www.cnbc.com/2018/06/14/bitcoin-and-ethereum-are-not-securities-but-some-cryptocurrencies-may-be-sec-official-says.html

Preis, T., Moat, H. S., & Stanley, H. E. (2013, May 4). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports, 3*, 1684.

Preis, T., Reith, D., & Stanley, H. E. (2010). Complex Dynamics of Our Economic Life on Different Scales: Insights from Search Engine Query Data. *Philosophical Transactions of the Royal Society A, 368*, 5707-5719.

Selgin, G. (2014, November 12). Bitcoin: Problems and Prospects.

Smith, G. P. (2012, June). Google Internet search activity and volatility prediction in the market for foreign currency. *Finance Research Letters, 9*(2), 103-110.

Statista. (2018). *Worldwide market share of search engines*. Retrieved June 13, 2018, from Statista: https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Takeda, F., & Wakao, T. (2014, April). Google Search Intensity and Its Relationship with Returns and Trading Volume of Japanese Stocks. *Pacific-Basin Finance Journal*.

Tantaopas, P., Padungsaksawasi, C., & Treepongkaruna, S. (2016, June). Attention effect via internet search intensity in Asia-Pacific stock markets. *Pacific-Basin Finance Journal, 38*, 107-124.

Tetlock, P. C. (2007, June). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance, 62*(3), 1139-1168.

Urquhart, A. (2016, November). The inefficiency of Bitcoin. *Economic Letters, 148*, 80-82.

van der Walt, E. (2018, March 27). *Bitcoin's a Pretty Good Guide to Animal Spirits in Stocks*. Retrieved from Bloomberg: https://www.bloomberg.com/crypto

Vlastakis, N., & Markellos, R. N. (2012, June). Information demand and stock market volatility. *Journal of Banking & Finance, 36*(6), 1808-1821.

Vozlyublennaia, N. (2014, April). Investor attention, index performance, and return predictability. *Journal of Banking & Finance, 41*, 17-35.

7. Appendix

7.1 Statistics Output – Correlation

I. Bitcoin variables correlations

Correlations

		BTCweb	BTCyt	btcR	btcVol	btcTV	btcMC
BTCweb	Pearson Correlation	1	.905**	.041	.457**	.881**	.880**
	Sig. (2-tailed)		.000	.505	.000	.000	.000
	N	261	261	261	261	261	261
BTCyt	Pearson Correlation	.905**	1	.079	.377**	.784**	.818**
	Sig. (2-tailed)	.000		.203	.000	.000	.000
	N	261	261	261	261	261	261
btcR	Pearson Correlation	.041	.079	1	-.035	-.026	-.012
	Sig. (2-tailed)	.505	.203		.574	.678	.844
	N	261	261	261	261	261	261
btcVol	Pearson Correlation	.457**	.377**	-.035	1	.339**	.319**
	Sig. (2-tailed)	.000	.000	.574		.000	.000
	N	261	261	261	261	261	261
btcTV	Pearson Correlation	.881**	.784**	-.026	.339**	1	.970**
	Sig. (2-tailed)	.000	.000	.678	.000		.000
	N	261	261	261	261	261	261
btcMC	Pearson Correlation	.880**	.818**	-.012	.319**	.970**	1
	Sig. (2-tailed)	.000	.000	.844	.000	.000	
	N	261	261	261	261	261	261

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

II. Ethereum variables correlations

Correlations

		ETHweb	ETHyt	ethTV	ethMC	ethR	ethVol
ETHweb	Pearson Correlation	1	.812**	.816**	.780**	.116	.442**
	Sig. (2-tailed)		.000	.000	.000	.061	.000
	N	261	261	261	261	261	261
ETHyt	Pearson Correlation	.812**	1	.502**	.483**	.135*	.443**
	Sig. (2-tailed)	.000		.000	.000	.029	.000
	N	261	261	261	261	261	261
ethTV	Pearson Correlation	.816**	.502**	1	.950**	.040	.358**
	Sig. (2-tailed)	.000	.000		.000	.522	.000
	N	261	261	261	261	261	261
ethMC	Pearson Correlation	.780**	.483**	.950**	1	-.021	.324**
	Sig. (2-tailed)	.000	.000	.000		.737	.000
	N	261	261	261	261	261	261
ethR	Pearson Correlation	.116	.135*	.040	-.021	1	.309**
	Sig. (2-tailed)	.061	.029	.522	.737		.000
	N	261	261	261	261	261	261
ethVol	Pearson Correlation	.442**	.443**	.358**	.324**	.309**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	261	261	261	261	261	261

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

7.2 Statistics Output – Regression

7.2.1 Bitcoin Regression Models

I. btcR = constant + btcR_lag1 + BTCyt + btcVol + ε

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.101[a]	.010	-.001	.786750772678

a. Predictors: (Constant), btcVol, LAGS(btcR,1), BTCyt

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.619	3	.540	.872	.456[b]
	Residual	158.458	256	.619		
	Total	160.077	259			

a. Dependent Variable: btcR

b. Predictors: (Constant), btcVol, LAGS(btcR,1), BTCyt

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.108	.086		1.248	.213
	BTCyt	.005	.003	.102	1.511	.132
	LAGS(btcR,1)	.000	.063	.000	.004	.997
	btcVol	-3.862	3.585	-.072	-1.077	.282

a. Dependent Variable: btcR

II. $btcR = constant + btcR\_lag1 + BTCyt\_lag1 + btcVol\_lag1 + \varepsilon$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.154[a]	.024	.012	.781379772659

a. Predictors: (Constant), LAGS(btcVol,1), LAGS(btcR,1), LAGS(BTCyt,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.775	3	1.258	2.061	.106[b]
	Residual	156.302	256	.611		
	Total	160.077	259			

a. Dependent Variable: btcR

b. Predictors: (Constant), LAGS(btcVol,1), LAGS(btcR,1), LAGS(BTCyt,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.059	.086		-.683	.495
	LAGS(btcR,1)	.016	.062	.016	.255	.799
	LAGS(BTCyt,1)	-.001	.003	-.025	-.376	.707
	LAGS(btcVol,1)	8.598	3.564	.161	2.412	.017

a. Dependent Variable: btcR

III. $btcVol = constant + btcVol\ \_lag1 + BTCweb + btcTV + \varepsilon$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.623[a]	.388	.381	.011604580163

a. Predictors: (Constant), btcTV, LAGS(btcVol,1), BTCweb

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.022	3	.007	54.177	.000[b]
	Residual	.034	256	.000		
	Total	.056	259			

a. Dependent Variable: btcVol

b. Predictors: (Constant), btcTV, LAGS(btcVol,1), BTCweb

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.009	.001		6.915	.000
	LAGS(btcVol,1)	.437	.053	.437	8.205	.000
	BTCweb	.001	.000	.497	4.669	.000
	btcTV	-1.191E-012	.000	-.241	-2.328	.021

a. Dependent Variable: btcVol

IV. btcVol = constant + btcVol_lag1 + BTCweb_lag1 + ε

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.584[a]	.341	.336	.012023555595

a. Predictors: (Constant), LAGS(BTCweb,1), LAGS(btcVol,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.019	2	.010	66.436	.000[b]
	Residual	.037	257	.000		
	Total	.056	259			

a. Dependent Variable: btcVol

b. Predictors: (Constant), LAGS(BTCweb,1), LAGS(btcVol,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.009	.001		7.212	.000
	LAGS(btcVol,1)	.462	.057	.462	8.116	.000
	LAGS(BTCweb,1)	.000	.000	.203	3.566	.000

a. Dependent Variable: btcVol

V. btcTV = constant + btcTV_lag1 + BTCweb + ε

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.981[a]	.962	.962	581451020.569663

a. Predictors: (Constant), BTCweb, LAGS(btcTV,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2212538235675317400000.000	2	1106269117837658700000.000	3272.160	.000[b]
	Residual	86887919355626230000.000	257	338085289321502850.000		
	Total	2299426155030943600000.000	259			

a. Dependent Variable: btcTV

b. Predictors: (Constant), BTCweb, LAGS(btcTV,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	-164865294.370	42288118.554		-3.899	.000
1	LAGS(btcTV,1)	.750	.021	.748	35.626	.000
	BTCweb	60028737.429	4653922.333	.271	12.899	.000

a. Dependent Variable: btcTV

VI. btcTV = constant + btcTV_lag1 + BTCweb_lag1 + btcTV_lag1 + ε

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.973[a]	.947	.946	690708693.708622

a. Predictors: (Constant), LAGS(btcVol,1), LAGS(btcTV,1), LAGS(BTCweb,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
	Regression	2177294059142387900000.000	3	725764686380796000000.000	1521.269	.000[b]
1	Residual	122132095888555880000.000	256	477078499564671420.000		
	Total	2299426155030943600000.000	259			

a. Dependent Variable: btcTV

b. Predictors: (Constant), LAGS(btcVol,1), LAGS(btcTV,1), LAGS(BTCweb,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	65025946.662	73711118.904		.882	.379
	LAGS(btcTV,1)	.795	.031	.792	25.559	.000
	LAGS(BTCweb,1)	48083003.450	7262243.639	.217	6.621	.000
	LAGS(btcVol,1)	-9049490086.154	3309214512.162	-.045	-2.735	.007

a. Dependent Variable: btcTV

7.2.2 Ethereum Regression Models

I. ***ethR= constant + ethR_lag1 + ETHyt + ethVol + ε***

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.313[a]	.098	.087	.86591

a. Predictors: (Constant), ethVol, LAGS(ethR,1), ETHyt

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	20.814	3	6.938	9.253	.000[b]
	Residual	191.949	256	.750		
	Total	212.762	259			

a. Dependent Variable: ethR

b. Predictors: (Constant), ethVol, LAGS(ethR,1), ETHyt

Coefficients[a]

Model	Unstandardized Coefficients	Standardized Coefficients	t	Sig.

		B	Std. Error	Beta		
1	(Constant)	-.083	.071		-1.167	.244
	LAGS(ethR,1)	.050	.062	.050	.803	.423
	ETHyt	-.001	.004	-.008	-.123	.903
	ethVol	11.361	2.555	.300	4.446	.000

a. Dependent Variable: ethR

II. ethR= constant + ethR_lag1 + ETHyt_lag1 + ethVol_lag1 + ε

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.151[a]	.023	.011	.90113

a. Predictors: (Constant), LAGS(ETHyt,1), LAGS(ethR,1), LAGS(ethVol,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.882	3	1.627	2.004	.114[b]
	Residual	207.881	256	.812		
	Total	212.762	259			

a. Dependent Variable: ethR

b. Predictors: (Constant), LAGS(ETHyt,1), LAGS(ethR,1), LAGS(ethVol,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.069	.074		.944	.346
	LAGS(ethR,1)	.098	.065	.098	1.505	.133
	LAGS(ethVol,1)	3.569	2.718	.094	1.313	.190
	LAGS(ETHyt,1)	-.001	.004	-.013	-.183	.855

a. Dependent Variable: ethR

III. $ethVol = constant + ethVol\_lag1 + ETHyt + \varepsilon$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.739[a]	.546	.542	.01619

a. Predictors: (Constant), ETHyt, LAGS(ethVol,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.081	2	.040	154.522	.000[b]
	Residual	.067	257	.000		
	Total	.148	259			

a. Dependent Variable: ethVol

b. Predictors: (Constant), ETHyt, LAGS(ethVol,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.005	.001		3.816	.000
	LAGS(ethVol,1)	.650	.046	.651	14.077	.000
	ETHyt	.000	.000	.172	3.723	.000

a. Dependent Variable: ethVol

IV. $ethVol = constant + ethVol\_lag1 + ETHyt\_lag1 + \varepsilon$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.728[a]	.530	.526	.01647

a. Predictors: (Constant), LAGS(ETHyt,1), LAGS(ethVol,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.079	2	.039	144.873	.000[b]
	Residual	.070	257	.000		
	Total	.148	259			

a. Dependent Variable: ethVol

b. Predictors: (Constant), LAGS(ETHyt,1), LAGS(ethVol,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.005	.001		4.001	.000
	LAGS(ethVol,1)	.676	.048	.677	14.178	.000
	LAGS(ETHyt,1)	.000	.000	.103	2.151	.032

a. Dependent Variable: ethVol

V. ***ethTV = constant + ethTV_lag1 + ETHweb + ε***

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.956[a]	.913	.913	291839105.87561

a. Predictors: (Constant), ETHweb, LAGS(ethTV,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	230694493815130300000.000	2	115347246907565150000.000	1354.317	.000[b]
	Residual	21888706375596847000.000	257	85170063718275664.000		

	Total	2525832001907 27130000.000	259			

a. Dependent Variable: ethTV

b. Predictors: (Constant), ETHweb, LAGS(ethTV,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-26521565.159	20261674.533		-1.309	.192
	LAGS(ethTV,1)	.737	.027	.733	27.085	.000
	ETHweb	15719615.899	1532222.723	.278	10.259	.000

a. Dependent Variable: ethTV

VI. *ethTV = constant + ethTV_lag1 + ETHweb_lag1 + ethVol_lag1 + ε*

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.942[a]	.886	.885	334654990.924 84

a. Predictors: (Constant), LAGS(ethVol,1), LAGS(ethTV,1), LAGS(ETHweb,1)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2239127456752 95230000.000	3	7463758189176 5080000.000	666.443	.000[b]

	Residual	28670454515431907000.000	256	111993962950905888.000		
	Total	252583200190727130000.000	259			

a. Dependent Variable: ethTV

b. Predictors: (Constant), LAGS(ethVol,1), LAGS(ethTV,1), LAGS(ETHweb,1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	37365918.910	27222646.773		1.373	.171
	LAGS(ethTV,1)	.825	.037	.820	22.344	.000
	LAGS(ETHweb,1)	9307018.600	2163783.207	.164	4.301	.000
	LAGS(ethVol,1)	-2086680628.971	967971299.111	-.051	-2.156	.032

a. Dependent Variable: ethTV