ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business
Specialization Financial Economics

Investor Attention in the Cryptocurrency Market

Author: S.J. Konijn

Student number: 456579

Thesis supervisor: Dr. J.J.G. Lemmen

Second reader: R. Spigt MSc

Finish date: 28 February 2023

PREFACE AND ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor, Dr. Jan Lemmen, for his assistance and for providing adequate feedback during this period of writing my thesis.

ABSTRACT

This research examines the effect of investor attention on returns, realized price volatility, and the trading volume for 154 cryptocurrencies in the period between April 1st 2019 until March 31st 2022. Google search volumes, produced by Google Trends, were used as a proxy for investor attention. The paper investigates the effect using a Vector Autoregression (VAR) model, which is supported by a Granger causality test. The VAR-model results show a significant positive effect from investor attention on the returns of cryptocurrencies, realized price volatility, and trading volume. These findings are supported by the Granger causality test, which found a bi-directional relationship between investor attention and return and trading volume, and a unidirectional relationship between investor attention and realized price volatility. The findings still hold after controlling for the S&P500 price index, oil price, gold price, volatility index of the S&P500 and the cryptocurrency market capitalization.

Keywords: Cryptocurrency, Investor attention, Google Search Volume

JEL Classification: G12, G14, G19, G40

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

November 8th 2021, marked the end of approximately 11 months in which the cryptocurrency market continued to grow rapidly. At the beginning of January 2021, the total market capitalization of the cryptocurrency market rose from $750 billion to approximately $3 trillion in the following eleven months (*Mondiale cryptocurrencygrafieken*, 2022). A market growth of this size could be considered an 'investors-paradise' when looking at the returns some of them made. However, the cryptocurrency market, and more specifically Bitcoin, is characterized by high price volatility (Kuo Chuen et al., 2017), which makes investments in cryptocurrencies risky. Prior studies show that Bitcoin – the largest cryptocurrency which possesses approximately 40% of the total market capitalization – is mainly used as a speculative asset due to its high price volatility (Baur et al., 2018). In addition, Kyriazis et al. (2020) found evidence that Bitcoin and other cryptocurrencies' prices are prone to speculative price bubbles and have shown multiple bubble phases in the past. A price bubble means that prices are more driven by market sentiment and momentum than by their intrinsic value. Cheah and Fry (2015) also find evidence that Bitcoin has a fundamental value equal to zero. This is because a cryptocurrency has no underlying business model, assets or rights which will generate cash flows. The price is only driven by the probability that the network will increase in users (Lubbersen & Wierts, 2022).

All this taken into account suggests that behavioural finance theories, rather than classical finance theories, could provide solutions to understanding the cryptocurrency market. One of these behavioural theories is investor sentiment. This is examined by Baker and Wurgler (2007) who showed that investor sentiment is very present in the financial market. Investor sentiment reflects an investor's beliefs about the future performance of an asset. Investor sentiment is influenced by investor attention. Investor attention is a theory that assumes that investors buy assets that have caught their attention. The more attention an asset has gained, the more the investor's sentiment is affected. Investor attention has been researched broadly in the light of cryptocurrencies. For example, Zhu et al. (2021) found evidence that higher levels of attention cause higher volatility in cryptocurrency prices and higher returns. They also find that investor attention has 'power' in forecasting cryptocurrency prices.

There is a growing body of research on the relationship between the cryptocurrency market and investor attention. However, the results are based on research which only focuses on Bitcoin or the top 10 cryptocurrencies based on their market capitalization – at most. The cryptocurrency market has grown rapidly over time and therefore it has asked for research with a far wider range of cryptocurrencies included. Therefore, this paper will contribute to the literature by examining the relationship between investor attention and the cryptocurrency market by looking at the top 200 cryptocurrencies based on the market capitalization by Coingecko.nl. The main question that will be examined in this paper is as follows:

What is the effect of investor attention on the returns, volatility, and trading volume in the cryptocurrency market, from the period April 1st 2019 until March 31st 2022?

The following three hypotheses will be tested to support the main question:

H1: Investor attention has a significant effect on cryptocurrencies' returns.

H2: Investor attention has a significant effect on cryptocurrencies' realized price volatility.

H3: Investor attention has a significant effect on cryptocurrencies' trading volume.

All of the hypotheses will be elaborated on in chapter 2.

To answer the main research question and the hypotheses, the research will use a dataset of 154 cryptocurrencies. Data on the cryptocurrencies are obtained manually from the website Coingecko (https://www.coingecko.com/). Google search volumes, produced by Google Trends, are collected as data of investor attention. Google search volumes were introduced by Da et al. (2011) and have proven to be a very endogenous proxy for investor attention. The relationship between investor attention and the three variables of interest the returns, realized price volatility, and trading volumes of the cryptocurrencies a Vector Autoregression (VAR)-model will be conducted, which is supported by a Granger causality test. The results show a significant positive effect from investor attention on returns, realized price volatility, and trading volumes of the cryptocurrencies. The Granger causality test supported these findings, which found a bi-directional relationship between investor attention and returns and trading volume. And a unidirectional relationship between investor attention to realized price volatility. Additionally, the same results have been found after adding several control variables. The control variables are the S&P500 price index, oil price, gold price, volatility index of the S&P500 and the cryptocurrency market capitalization. The control variables will help to overcome the endogeneity problem that might arise due to the connection between the cryptocurrency market and other financial markets.

The remainder of this paper is organized as follows. In the next section, we briefly review the existing literature on cryptocurrency and investor attention. Chapters 3 and 4 describe the data and methodology, respectively. Chapter 5 shows the empirical results, and chapter 6 concludes.

2. LITERATURE REVIEW

In this chapter, classical theories about investor attention will be discussed first. Second, the world of cryptocurrencies will be explained. Third, the results of studies on investor attention in the cryptocurrency market will be discussed. Finally, these results will be synthesized into hypotheses that will help answer the main research question.

2.1 Investor Attention

The classical theory of finance is based on the Efficient Market Hypothesis (EMH). This theory suggests that markets are efficient and new information will be immediately reflected in the prices of the assets (Fama, 1970). The EMH is based on the assumption that investors are rational, and therefore assets are rationally valued. Another assumption is that investors analyse all the available information in the market, before making an investment decision. Fama (1970) introduced three forms of market efficiency: the weak, semi-strong and strong forms. The weak form of the EMH states that all past public market information is incorporated into the prices of assets. The semi-strong is a bit stricter than the weak form and states that prices of assets and securities adjust quickly to the new publicly available information. The strong form is even more strict than the semi-strong form and says that all the prices reflect all the available information (private and public) in the market (Fama, 1970). Following this theory investors are assumed fully rational, so the attention of investors should result in a more efficient way of information processing without resulting in changes in prices or trading volumes. However, in reality, attention is a scarce cognitive resource (Kahneman, 1973), and investors have limited attention. Therefore, investor attention is a subject that is broadly researched by researchers. Unfortunately, examining investor attention faces some challenges as there is no direct way of measuring the level of attention. For this reason, an endogenous proxy has to be found so it can lead to unambiguous results.

2.1.1 Advertising Expenses

Grullon, Kanatas & Weston (2003) believe investor attention is more of a 'familiarity' investors have with the firm and therefore advertising expenses can serve as a proxy for investor attention. Even though when these advertising expenses presumably are intended to gain market share, at the very least they should increase customer and investor awareness of the company's name and products. This theory builds on the words of Merton (1987), who stated that a potential investor needs at the very least be aware of a company before determining whether to buy its stock or obtain additional information. Grullon, Kanatas & Weston (2003) found that greater advertising expenses by firms are associated with more individual and institutional investors. Also, greater advertising expenses are significantly correlated with more liquidity in the firms' common stocks. Something similar has been found by Chemmanur & Yan (2009). They found that product advertising expenses are highest in the year of a firm's IPO. The theory behind their findings is that firms want to create awareness among consumers and investors before the IPO, so they could raise more new equity.

Lou (2014) also uses advertising expenses as a proxy for investor attention. He not only finds that an increase in advertising expenses is followed by an increase in retail stock purchases, but he also provides evidence that it leads to higher abnormal stock returns (Lou, 2014). However, looking further into the future, it appears that those higher returns level off with lower returns. Lou (2014) argues that managers are aware of this phenomenon in advance and use these higher advertising costs to influence short-term returns to increase their profits from insider trading.

2.1.2 Stock Characteristics

Gervais, Kaniel, and Mingelgrin (2001) believe that unusually high trading volumes should create awareness among investors which leads them into trading particular stocks. Therefore, they use trading volume as a proxy for investor attention. Their findings do confirm their hypotheses. They find that stocks whose trading volume is unusually high (small) over a certain period, have large (small) returns over the following month. In addition, Hou, Xiong, and Peng (2009) found that the price momentum effect is stronger for stocks with high trading volume. They also found that the earnings momentum effect is higher for stocks with low trading volumes. In both cases, the authors use trading volumes as an attention indicator. This is in line with the findings of Gervais, Kaniel, and Mingelgrin (2001).

Barber and Odean (2008) disagree with most papers related to investor attention. They believe individual investors' buying behaviour is impacted more by attention than their selling behaviour. While in most papers about investor attention buying and selling decisions are treated as the same side of the coin. Barber and Odean (2008) stated that investors who want to buy stocks are faced with a huge search problem because there are thousands of common stocks to choose from. While most individual investors only can sell stocks they already own, due to short-sell constraints, this makes the choice far easier. Barber and Odean (2008) also believe that attention is not as scarce a resource for institutional investors as for individual investors, as institutions have more sophisticated computer programs and more time to spend on searching than retail investors do. The authors test their beliefs with three different proxies: the abnormal trading daily volume, extreme daily returns, and whether a firm was in the news that day. In line with their predictions, they find that individual investors are prone to attention-driven buying behaviour, and institutional investors do not display that behaviour (Barber & Odean, 2008). The attention-driven buying patterns that are observed do not show any exceptional returns - which slightly contradicts the findings of Gervais, Kaniel, and Mingelgrin (2001).

Seasholes and Wu's (2007) research is an extension of that of Barber and Odean (2008). They examined the Shanghai Stock Exchange and the characterised stock price limits as a proxy for investor attention. The authors reasoned that a stock that hits the upper price limit on the stock market, shows the same characteristics as the proxies Barber and Odean (2008) used in their study. The upper price limit event of a stock is associated with three attention-grabbing events. First, stocks that hit the upper price limit are likely to generate high returns for investors. Second, the trading volume of stocks that hit the upper price limit is high. And third, the event generates news. They found that attention-grabbing stocks have positive net buy-sell differences and that the attention-grabbing events induce investors to buy new stocks they previously did not own (Seasholes & Wu, 2007). This is in line with the findings of Barber and Odean (2008). Seasholes and Wu (2007) also show that so-called 'smart traders' can generate a daily return of 1.06% by buying attention-grabbing stocks and selling them the next day. However, investors, who just buy the stocks that hit the upper price limit without selling them the day after, suffer losses due to price reversal the 5 days after the event.

2.1.3 Media Coverage

It has long been known that media has a great influence on almost everything in the world, including financial markets. But how big the influence is, differs from one case to another. To test the influence of the media on the financial markets Blankespoor et al. (2017) introduced robo-journalism. They created automated articles with an algorithm about companies with little previous media coverage. The articles that were produced lacked any private information, were exogenous in their interpretation, and spoke only about the firm's profits and decisions. Blankespoor et al. (2017) find an increase of approximately 11% in trading volume around the time the automated robo-articles have been published. Also, the authors find an increase in the depth of the market which suggests an improvement in liquidity. However, they find no evidence that automated news articles have effects on price discovery (Blankespoor et al., 2017).

Fang and Peress (2009) examined the relationship between mass media coverage and stock returns. They found that stocks that were not covered by the media significantly outperform stocks that were covered by the media. Moreover, a portfolio of stocks that were not covered by the media earned 3% more return per year than a portfolio of uncovered stocks (Fang and Peress, 2009). This even holds after controlling for the market, firm size, book-to-market ratio, and price momentum. Fang and Peress (2009) conclude that uncovered stocks have a ‘no-media premium’. Their results are neither in line with Barber and Odean (2008), nor Gervais, Kaniel, and Mingelgrin (2001). Fang and Peress (2009) give two main explanations for their findings. The first reason is that a liquidity problem is preventing arbitragers to correct the mispricing. Second, the no-media premium may be compensation for imperfect diversification (Fang and Peress, 2009). Since uncovered stocks have higher idiosyncratic risk, they should earn higher returns.

Tetlock (2007) examined the role of media pessimism in the stock market. He finds that high levels of pessimism in the media create downward pressure on the stock prices but return to their original values after a while. Secondly, extraordinary levels of pessimism predict significantly higher market trading volume. Thirdly, he finds that low market returns lead to high levels of media pessimism. Lastly, he finds that media content does not contain new fundamental information about the asset market, and that media content is not more than a sideshow to the asset market. This suggests that investor sentiment is very present in the financial sector.

2.1.4 Search Volumes

The previous chapters demonstrate that many different proxies have been used for investor attention. However, these proxies make the crucial premise that if a stock's return or turnover was unusually high, or if its name was reported in the press, investors should have taken notice. Returns and turnover, on the other hand, can be influenced by factors unrelated to investor attention, and a news piece in the Wall Street Journal does not guarantee that investors will read it (Da et al., 2011).

Therefore, Da et al. (2011) introduced Google aggregate search volumes as a new direct measure for investor attention. The authors argue that if you search for a stock on Google, you are paying attention to it. Which makes the Google aggregate search volume an unambiguous measure of investor attention. They find evidence that the search volume captures mostly the attention of (less sophisticated) retail and individual investors. This is in line with the findings of Barber and Odean (2008) who stated that individual investors are more prone to attention-buying behaviour. Da et al. (2011) also find that higher search volume stocks are associated with higher returns of 30 basis points, but these returns disappeared again at the end of the year. This also corresponds with the results of Barber and Odean (2008).

Bank et al. (2011) agree with Da et al. (2011) and find that Google search volume is a strong measure of investor attention. The authors examined the influence of Google search volumes on the German stock market. They find that an increase in a company's search volume is associated with an increase in the firm's stock liquidity. This effect is explicitly present for firms with low market capitalization, which suggest the increase in liquidity is a result of a decrease in asymmetric information costs (Bank et al, 2011). Just like Barber and Odean (2008), and Da et al. (2011), they find temporary positive returns on stocks with higher search volumes. In contrast, Bijl et al. (2016) find that high Google search volumes of stocks lead to negative returns using a more recent dataset. The authors argue that information in the search volumes is processed more quickly in the prices, which could explain the different results. Taking in mind the transaction costs, these negative returns are not large enough to create a profitable trading strategy (Bijl et al., 2016).

2.2 The World of Cryptocurrencies

2.2.1 What are Cryptocurrencies?

A cryptocurrency is a digital asset designed as a medium of exchange that uses cryptography to make sure the transactions are not tampered with (Härdle et al., 2020). Cryptocurrencies run on a Blockchain network which is a list of transactions, called blocks, that are linked using cryptography. This results in a chain of transactions. The blockchain can be compared with a distributed ledger where every user, referred to as a node in blockchain terminology, has his online copy of the ledger. Blockchain is often managed using a peer-to-peer network, allowing participants to communicate and collectively validate new blocks into the chain. Through the mechanism of blockchain the 'middle-man', like banks or other financial institutions, can be eliminated (Härdle et al., 2020). This is possible since all the data on the blockchain has been validated by other users and therefore can be trusted.

Many different cryptocurrencies can exist on the same network. Each cryptocurrency has its consensus mechanism, the number of coins outstanding, hashing algorithm, and the use of the cryptocurrency (Härdle et al., 2020; Lansky, 2018). The most well-known cryptocurrency is Bitcoin. Since the launch of Bitcoin in 2008 by Satoshi Nakatomo, the price of bitcoin increased by more than 500% in July 2016 (Urquhart, 2016). As of 1 June 2022, Bitcoin has a total market capitalisation of more than $600 billion, which represents approximately half of the total market capitalisation of the cryptocurrencies, based on data from Coingecko. Bitcoin originally was introduced as a currency, but nowadays there is a debate as to whether bitcoin is better to be labelled as an asset. Like in every debate, there are always arguments in favour and against, but most researchers agree that because of the high price, volatility, and zero correlation with other currencies, it is hard to serve as a currency (Glaser et al., 2014; Yermack, 2015; White, 2020)

2.2.2 Pricing of Cryptocurrencies

Bitcoin is not the only cryptocurrency that has high price volatility. Investors can get lucky and receive extreme returns that they would not have thought possible beforehand, and others can have bad luck and lose a fortune from one day to the next. Therefore, the pricing of cryptocurrencies is a field of interest for many researchers.

Where cryptocurrencies differ from other financial assets, is that no dividends or any other future cash flows are generated by having a cryptocurrency in the portfolio. Therefore, unbacked cryptocurrencies lack intrinsic value. For example, the fundamental value of Bitcoin is equal to zero (Cheah & Fry, 2015). This is because unbacked cryptocurrency has no underlying business model, assets or rights which will generate cash flows. The price is only driven by the probability that the network will increase in users (Lubbersen & Wierts, 2022). However, Bitcoin believers think that Bitcoin is the only digital asset that will be needed in the future. Because of its decentralized structure, its anonymity of the users, and limited supply, its price will increase considerably even from where it is now after the recent Bitcoin crash (23000 US Dollars in July 2022 compared to 67000 US Dollars in November 2021). On the other hand, some investors believe because of the lack of intrinsic value, Bitcoin and other cryptocurrencies are in a financial bubble which will burst (Cheun et al., 2015).

The supply of most unbacked cryptocurrencies is limitless, therefore it creates scarcity among these cryptocurrencies when there is a bull market. For example, Bitcoin handles a maximum Bitcoin supply of 21 million Bitcoins. The problem that arises from a fixed 'money supply' is that it cannot react to certain market developments and economic conditions. This makes Bitcoin and other unbacked limitless cryptocurrencies very prone to high price volatility (Lubbersen & Wierts, 2022). Besides unbacked cryptocurrencies, there are also stablecoins. Stablecoins are cryptocurrencies that are backed with assets. Most stablecoins are backed by a single fiat currency, but some stablecoins are backed by other assets or multiple fiat currencies. Stablecoins are designed to serve as money or investment in the world of cryptocurrencies because the stablecoins intend to be stable without price fluctuations. However, if the stablecoins are backed by other cryptocurrencies or unstable assets this will not be the case and the value will fluctuate (Lubbersen & Wierts, 2022).

García-Monleón et al. (2021) have tried to develop a theoretical framework that points out that the level of utility, number of layers, and value of the nodes of the cryptocurrencies has an impact on the prices. Since this is still very theoretical, and other research on intrinsic value is very limited, investors struggle for determining the intrinsic value, which serves as a threshold in valuing other financial assets. Therefore, the prices of cryptocurrencies are speculative (Yermack, 2015). This ensures that there are high risks associated with trading in cryptocurrencies, but it can also create high returns. Thus, trading in cryptocurrencies is suitable for risk-seeking investors. This claim is supported by many types of research that point out that cryptocurrencies are characterised by high volatility (Kuo Chuen et al., 2017; Abu Bakar & Rosbi, 2017; Liu & Tsyvinski, 2021).

Liu and Tsyvinski (2021) establish that network growth is significantly and positively correlated with the prices of cryptocurrencies. Also, they find that time-series momentum and investor attention have a significant effect on the prices, but more on this subject in the next chapter. Furthermore, several papers have shown that the prices of cryptocurrencies can exhibit behaviour like financial bubbles (Kyriazis et al., 2020; Cheah & Fry, 2015). In addition, Caferra et al. (2021) stated that where firms, and thus stock prices, are related to the state of the economy, cryptocurrency prices are connected to the behaviour of traders and investors. Therefore, behavioural finance could play a key role in predicting cryptocurrency prices.

2.3 Investor Attention in the Cryptocurrency Market

Previous research has shown that cryptocurrencies do not behave like other assets on the financial market and that classical financial theories do not offer a solution. Therefore, theories from behavioural finance are tested to see if they can explain the unorthodox behaviour of cryptocurrencies in contrast to other assets.

Smales (2022) uses search engines as a proxy for investor attention. He finds that an increase in search volume is associated with higher returns for the largest cryptocurrencies, greater volatility, and higher illiquidity (Smales, 2022). The author explains that irrational investors are afraid of missing out on peaks in the market. Liu and Tsyvinski (2021) find similar results while examining the predictability of cryptocurrency returns. They find that high levels of investor attention predict high future returns over the one-to-six-week horizons. An increase of one standard deviation in the investor attention measure leads to an increase of 3.0% in the returns of cryptocurrency by using Google search volume as a proxy of investor attention. These results were similar when the number of Twitter posts was used as a proxy for investor attention. In addition, Kraaijeveld & De Smedt (2020) found in their research on investor sentiment on Twitter, that Twitter sentiment causes trading volumes and returns on the cryptocurrency market rather than Twitter activities following the events on the cryptocurrency market.

Subramaniam and Chakraborty (2019) examine the causal relationship between investor attention and the prices of cryptocurrencies. They find that spikes in investor attention cause positive returns among Bitcoin, Ripple, and Ethereum. They also find that for newer cryptocurrencies, like Ripple, investor attention only influences the returns on an exceptional performance. Therefore, there are differences across the cryptocurrencies, and the results of some coins cannot be applied to other cryptocurrencies (Subramaniam and Chakraborty, 2019).

Smuts (2019) is one of the first to show that the relationship between Google search volume and cryptocurrency prices is no longer consistently positive. He finds especially strong negative correlations for the two largest cryptocurrencies: Bitcoin and Ethereum. Lin (2021) finds contradictory results to those of Smales (2022) and Liu and Tsyvinski (2021). He finds hardly any evidence that investor attention has any influence on the returns of cryptocurrencies. Lin (2021) does find evidence that past returns have a positive significant impact on future investor attention.

Zhang and Wang (2020) show the bi-directional Granger causality between investor attention and cryptocurrency return and return volatility. This means that investor attention not only affects returns and return volatility but also the other way around. Zhang and Wang (2020) also find that high levels of attention always indicate positive returns for the top twenty cryptocurrencies. However, low levels of investor attention lead to inconsistent results (Zhang & Wang, 2020). Urquhart (2018) uses Google search volumes to find out what causes the attention of Bitcoin. He finds that realized volatility and trading volume are important determinants of next-day investor attention, but investor attention has no significant capacity in predicting and forecasting realized volatility, trading volume, and Bitcoin returns. Zhu et al. (2021) examined the relationship between investor attention and Bitcoin in different ways. First, the authors show that investor attention Granger causes Bitcoin returns and realized volatility, which is in line with the findings of Zhang and Wang (2020). Second, they find non-linear connections between investor attention and the bitcoin market. Third, several out-of-sample predictions were implemented and showed that the predictive model outperformed the benchmark model in the first period. This means that investor attention has some predictive power in Bitcoin's return (Zhu et al., 2021). However, predictive models did not outperform the benchmark model for realized volatility.

Shen et al. (2019) used the number of tweets while examining the relationship between investor attention and Bitcoin. They argue that Twitter is a better proxy for investor attention than Google search volumes because it captures the attention of more retail investors. They find that the volume of tweets is a substantial determinant of realized volatility and trading volume. Al Guindy (2021) also uses the number of tweets as a proxy and found that the proxy corresponds with greater cryptocurrency price volatility. Also, the author showed with a Vector Autoregression that investor attention could predict future price volatility. Moreover, attention-grabbing events (during which investors are 'distracted') complies with lower price volatility (Al Guindy, 2021).

Table 1 summarises the articles covered in section 2.3. It shows the sample period of the research, what was examined, the methods and proxies used for the study, and the results the study obtained.

Table 1 Meta Table

PAPER	PERIOD	WHAT?	PROXY FOR ATTENTION	METHOD (CONTROL VARIABLES)	RESULTS
Urquhart (2018)	August 2010 – July 2017	Volatility, Trading volume, Return	Google Search Volume	VAR-model	Previous day volatility and trading volume are drivers of attention of Bitcoin
Shen et al. (2019)	September 2014-August 2018	Volatility, Return, Trading volume	Twitter	VAR-Model	Number of tweets is a significant driver of volatility and trading volume for Bitcoin
Subramaniam and Chakraborty (2019)	January 2013-March 2018	Causality return	Google Search Volume	Quantile Causality Approach	Search volume ↑ → returns Bitcoin, Ethereum, and Ripple ↑
Smuts (2019)	December 2017 – June 2018	Predicting Prices	Telegram, Google Search Volume	LSTM Models	Investor attention can predict prices: best 1-week horizon
Zhang and Wang (2020)	April 2013 – April 2018	Causality Return and volatility	Google Search Volume	(non) Linear-Granger Causality Test,	Bi-directional Granger causality / High levels of attention predict high returns

				Quantile regression	
Zhu et al. (2021)	July 2013 – May 2020	Return, Volatility	Google Search Volume	VAR-Model, Granger Causality Test	Investor attention causes volatility and returns / Investor attention has predictive power for returns
Al Guindy (2021)	November 2017 - November 2018	Volatility	Twitter	OLS-Regression (daily twitter sentiment, coin-specific effects)	Number of tweets ↑ → price volatility ↑
Lin (2021)	April 2017- February 2020	Return	Google Search Volume	Granger Causality test, VAR-model	Higher past returns predict higher future attention
Liu and Tsyvinski (2021)	January 2014- July 2020	Predicting returns	Google Search Volume, Twitter	OLS-regression	Search volume ↑ → +3.0% return
Smales (2022)	January 2014- June 2021	Volatility, Return, Liquidity	Google Search Volume	OLS-regression (macro, uncertainty)	Search volume ↑ → greater volatility, return, and illiquidity

2.5 Hypotheses

Previous literature by Subramaniam and Chakraborty (2019) has shown that there is a causal relationship between investor attention and the return on cryptocurrencies. Moreover, Zhang and Wang (2020) even showed that there is a bi-directional relationship between the two. Additionally, previous literature has proven that high levels of investor attention cause higher returns for cryptocurrencies (Zhang and Wang, 2020; Subramaniam and Chakraborty, 2019; Liu and Tsyvinski,2021; Zhu et al.,2021). Therefore, the first hypothesis is as follows:

H1: Investor attention has a significant effect on cryptocurrencies' returns.

The relation between investor attention and price volatility of cryptocurrencies has also been thoroughly researched. Urquhart (2018) found that previous-day volatility is one of the drivers of investor attention to Bitcoin. Similar results have been found by Shen et al. (2019) that show that the number of tweets, as a measure of investor attention, is a significant driver of price volatility of Bitcoin. Additionally, Zhu et al. (2021) show that higher levels of investor attention cause an increase in price volatility of cryptocurrencies. Moreover, Smales (2022) found similar results and shows that an increase in investor attention leads to higher price volatility. Therefore, the second hypothesis is as follows:

H2: Investor attention has a significant effect on cryptocurrencies' realized price volatility.

The relationship between investor attention and trading volume is examined often among researchers. Shen et al. (2019) found that the number of tweets is a significant driver of the trading volumes of Bitcoin. And Urquhart (2018) shows that previous-day trading volume is one of the drivers behind investor attention on Bitcoin. Additionally, Smales (2022) found that higher levels of investor attention led to more illiquidity among cryptocurrencies. These returns seem counterintuitive compared to traditional financial markets, however when the price of cryptocurrencies is rising, there tends to be an increase in volatility, price jumps and liquidity decreases. This could be explained by retail investors who have a 'fear of missing out' on potentially high returns and therefore take liquidity from the market. Therefore, the third hypothesis is as follows:

H3: Investor attention has a significant effect on cryptocurrencies' trading volume.

3. DATA

In this chapter the data collection process will be elaborated. First, it will be discussed how the data on investor attention is obtained and modified. Second, the same will be done with the data on the cryptocurrency market. Finally, the descriptive statistics of the dataset will be covered.

3.1 Investor Attention Data

Google Trends provides insights on the frequencies of search terms in Google, dating back to January 2004 (Da et al., 2011). These frequencies are expressed in the *Search Volume Index* (SVI). Additionally, Google Trends provides data per Country and category on these search volumes. To ensure completeness of the data collection: country-filter is set to global since cryptocurrencies can be traded all over the world. Also, for full completeness, all categories are selected so no search volumes are ignored for this analysis.

The value of the SVI is measured in values between 0 and 100, where 0 refers to a period where the term does not comply with a certain threshold value and 100 refers to a period where the term was searched on its highest relative volume (Bank, Larch & Peter, 2011). The SVI, therefore, indicates how much relative attention a certain topic is receiving. This paper will use the weekly *Search Volume Index* (SVI) provided by Google Trends as a direct measure of investor attention for the period between April 1st 2019 until March 31st 2022. Weekly SVI for a search term is the number of searches for that term scaled by its time-series average (Da et al., 2011). Each cryptocurrency has its cryptocurrency ticker, a short code or text that identifies cryptocurrencies. This ticker will be used as a search term in Google Trends because it is less ambiguous than the cryptocurrency's name. The ticker for each cryptocurrency that was used in Google Trends can be found in Appendix A.

3.1.2 Investor Attention Data Modifications

Following the paper of Da et al. (2011), the log of SVI during the week minus the log of median SVI during the previous 8 weeks has been taken to calculate the abnormal search volume index (ASVI). The formula is as follows:

$$ASVI = \log(SVI_t) - \log\left[Med(SVI_{t-1}, \dots, SVI_{t-8})\right] \quad (1)$$

Where log(SVI_t) is the logarithm of the Search Volume Index (SVI_t) during week t. The logarithm of the SVI_t in the previous 8 weeks is been given by $log[Med(SVI_t{-}1, \dots, GSVI_t{-}8)]$. The 'normal' level of attention is captured by the median that is taken from the previous weeks. In this way, the ASVI is resistant to recent jumps or low-frequencies seasonalities (Da et al., 2011). A large positive ASVI represents a large spike in investor attention to that cryptocurrency.

3.2 Cryptocurrency Data

3.2.1 Cryptocurrency Data Collection

The data of each cryptocurrency is manually downloaded from the website Coingecko (https://www.coingecko.com/). Coingecko provides a fundamental analysis of the crypto market and tracks more than 10.000 cryptocurrencies. For this research, the daily opening and closing prices, daily trading volume, and daily market capitalization are collected from the top 200 cryptocurrencies on this website, as of August 10th 2022. Coingecko combines data from many different cryptocurrency exchanges, therefore the data will be more accurate than for some single exchanges, which could suffer from miscalculations.

Daily data is collected over the sample period of April 1st 2019 until March 31st 2022. To match the weekly data provided by Google Trends, weekly averages are calculated for the variables. This will be more elaborated in section 3.1.2 Cryptocurrencies that were released after September 2021 will not be included in the sample due to insufficient data points. Also, it will exclude all the cryptocurrencies, which coin ticker is a normal word. For example, the cryptocurrency 'Chainlink' has a coin ticker 'Link'. Using the coin ticker 'Link' as a search term for Google Trends would lead to very ambiguous results since it would not show me the direct measure of investor attention. It would also display the results of everybody who searched 'link' on Google. Finally, some coins were not included because the data from either Google Trends or Coingecko about the coin was missing. At the end of the selection process, 154 cryptocurrencies were included in the dataset. A list including all the cryptocurrencies can be found in Appendix A.

3.2.2 Cryptocurrency Data Modifications

The weekly average closing price is computed out of the daily data available on Coingecko.com to represent the weekly cryptocurrency prices. The returns of the cryptocurrencies are calculated using the following formula:

$$R_t = Ln\,(P_t) - LN(P_{t-1}) \qquad (2)$$

where Rt is the return of the cryptocurrency and Pt and $Pt{-}1$ are the average weekly closing prices at respectively weeks t and t – 1.

The second variable that is calculated is realized volatility. Andersen and Bollerslev (2003) introduced calculating the weekly realized volatility by using the daily returns, which are calculated in equation 2. This paper follows the paper by Urquhart (2018) and therefore, the weekly realized volatility will be calculated as the root of the squared sum of daily returns. Equation 3 is therefore stated as:

$$RV_t = \sqrt{\sum_{i=1}^{N} R_t^2} \tag{3}$$

Where R_t^2 is the squared daily return of a cryptocurrency.

The third variable that is calculated is the weekly trading volume. Also, for this variable, the weekly average is calculated from the daily data on Coingecko.com. Many papers have established a strong relationship between trading volume and returns. Therefore, a detrended volume is used. Based on previous research by Llorente et al. (2002), a volume shock is defined as the log deviation of trading volume from its trend over a rolling period for a certain period. Based on this theory the weekly volume will be calculated as follows in equation 4:

$$VLM_t = \log(Volume_{\ t}) - \log\left(\frac{\sum_{i=t-11}^{t} Volume_i}{12}\right) \tag{4}$$

where $\log(Volume_t)$ is the logarithm of the weekly trading volume.

3.3 Figures and Descriptive Statistics

To show that google search volumes and the cryptocurrency market are highly correlated, Figures 1 to 4 have been created. The figures show that the amount of times 'BTC' is googled follows roughly the same trend as Bitcoin's price, price volatility, and trading volume. Bitcoin has been used in this example because it is the largest cryptocurrency.

Figure 1 Price development and Google search volume of Bitcoin

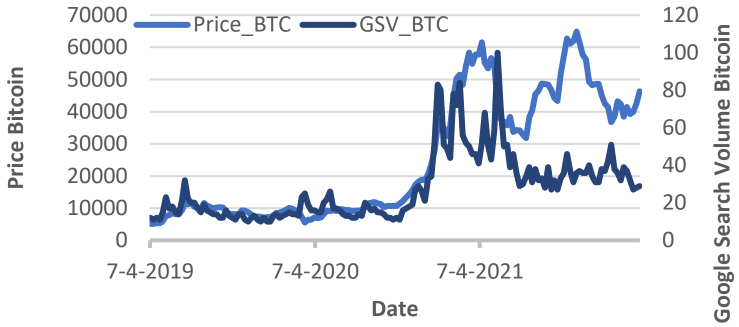

Figure 2 Price volatility development and Google search volume of Bitcoin

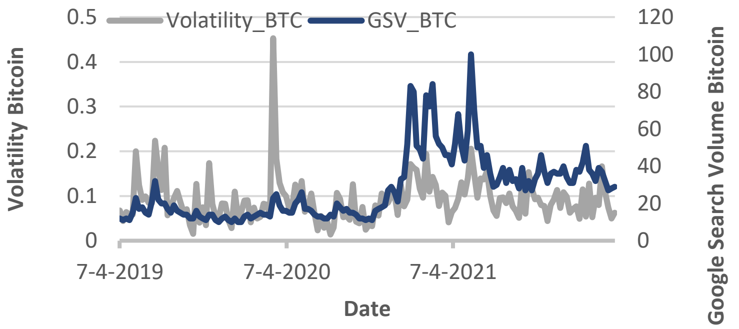

Figure 3 Trading volume development and Google search volume of Bitcoin

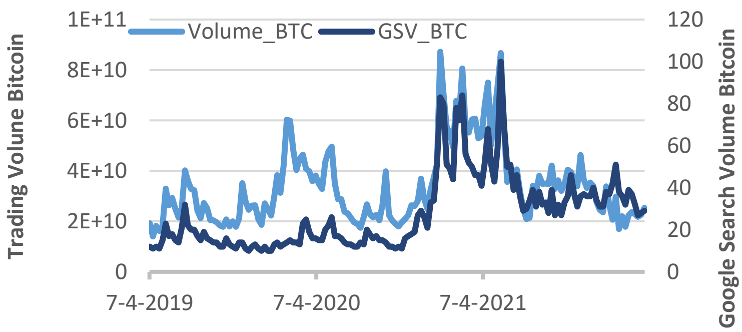

Table 2 shows the descriptive statistics for the data obtained from Coingecko.com and Google Trends. It can be seen that the 154 different cryptocurrencies on average generated a weekly return of 2%. However, from the minimum and maximum weekly returns, respectively -21.7% and 35.5%, can be seen that cryptocurrencies are risky investments. This is extra underlined by the high positive skewness and high kurtosis. The realized volatility is calculated as the root of the squared sum of daily returns, therefore there are no negative values for realized volatility. This also explains the extreme value of the kurtosis, since the observations are now spread into a smaller distribution.

Another important thing to note from Table 2 is some cryptocurrencies were founded around the start of the sample period. That is why in some cases there was no data available for the first few weeks regarding the trading volume. This explains the differences in the number of observations between return, volatility and trading volume. Finally, the number of observations from investor attention (ASVI) largely exceeds the number of observations from the cryptocurrencies. The reason for this is that Google Trends is unable to recognize typos from the search terms and thus will these typos be included in the received output. Also, it can be the case that people search for the ticker of the cryptocurrency just to learn more about the Coin before it is introduced to the market.

Table 2 Descriptive statistics

Variable	Obs	Mean	Std.Dev.	Min	Max	Skewness	Kurtosis
Return	17218	.002	.027	-.217	.355	1.264246	14.95808
Volatility	17218	.152	.132	.001	2.91	4.062598	42.97186
Volume	16180	.046	.385	-4.821	3.763	.4239005	14.14462
ASVI	22592	.012	.141	-1.154	1.699	1.942566	20.88256

4 METHODOLOGY

In this chapter, the methodology that will be used in this research will be elaborated. Important to note is that the three hypotheses will be answered in the same way using the same methodology. The following methodology will be applied to the data. First, the data will be tested for stationary. Second, a VAR-model will be elaborated and the variables of interest will be highlighted. Lastly, the Granger-causality test will be conducted to test the direction of the observed effect.

4.1 Stationary

In this research a vector autoregressive (VAR)-model will be used. However, for a VAR-model to work, the data has to be stationary. A stationary process is one whose statistical properties do not change over time (Nason, 2006). In other words, the series is flat looking with no visible trend or autocorrelation. If data is non-stationary it can lead to spurious and misleading results (Brooks & Chris, 2019).

The Dickey-Fuller or Augmented Dickey-Fuller tests are both fit to test if the data is stationary (Brooks & Chris, 2019). However, the Augmented Dickey-Fuller test is used when the error term is likely to be white noise (Mushtaq, 2011). White noise means that, in this case, the error terms have no mean, finite variance, and the covariance between the dependent variable and the first lag of the dependent variable is equal to zero. Unfortunately, the Augmented Dickey-Fuller cannot be used on panel data in Stata. Therefore, a Fischer-type unit root test is used to test if the data is stationary, which is based on the Augmented Dickey-Fuller test. The Fisher-type unit root test uses the p-values from the panel-specific Augmented Dickey-Fuller test and uses the four methods proposed by Choi (2001). The Augmented Dickey-Fuller test has as the null-hypothesis that α = 1 in equation 9, where α is the coefficient of the first lag on the dependent variable y. This means that the data is non-stationary. The alternative hypothesis from the test is that α < 1, when this is the case the data is stationary.

$$y_t = c + \beta_t + \alpha y_{t-1} + \emptyset \Delta y_{t-1} + \varepsilon_t \tag{5}$$

$$H_0: \alpha = 1$$

$$H_1: \alpha < 1$$

The results of the Fisher-type unit root test are presented in Table 3.

Table 3 Fisher-type unit root test

Variable	**Test**	**Statistic**
Return	Inversed chi-squared *(308)*	6785.7117***
	Inverse normal	6785.7117***
	Inverse logit *(774)*	-150.6994***

	Modified inv. chi-squared	260.9944***
Realized	Inversed chi-squared *(308)*	4157.2333***
Volatility	Inverse normal	-55.6260***
	Inverse logit *(774)*	-92.3568***
	Modified inv. chi-squared	155.0900***
Volume	Inversed chi-squared *(306)*	1574.9885***
	Inverse normal	-28.1763***
	Inverse logit *(769)*	-34.1758***
	Modified inv. chi-squared	51.2958***
ASVI	Inversed chi-squared *(308)*	4238.4359***
	Inverse normal	-57.4736***
	Inverse logit *(764)*	-94.8227***
	Modified inv. chi-squared	158.9593***

*, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively

4.2 VAR-Model

Because the data has been found stationary, a VAR-model will be used to examine the relationship between investor attention and the three variables of interest (return, realized volatility, and trading volume). VAR-models are used for multivariate time series models. Multivariate time series models are datasets with more than 1 dependent variable which varies over time, also called time-dependent (Brooks & Chris, 2019). VAR-models are very useful for this type of analysis, because lagged values (past values) for both the dependent and independent variables are allowed in the regression, and will help to predict future values. Following Urquhart (2018), a VAR-model with only one variable can be described as:

$$X_t = c + \sum_{i=1}^{P} \beta_i X_{t-i} + \varepsilon_t \quad (6)$$

Where X_t is the variable of interest (return, volatility, trading volume, and the SVI), c is a constant, and ε_t is the error term. β_i is the coefficient expressing the effect of the lagged component, and ρ captures the models lag duration.

The next step is to determine how many lags from each variable need to be added to the VAR-model. A VAR-model that uses too many lags, leads to a model that is overfitting. While using a VAR-model with too few lags can lead to autocorrelated errors (Brooks & Chris, 2019). Akaike information criterion (AIC) and Bayesian information criterion (BIC) are the 2 most commonly used information criteria for determining the number of lags. However, there is no clear consensus in the literature as to which of the two should be used. In this paper, the Bayesian information criterion will be used. The procedure for selecting the appropriate leg length can be found in Appendix B.

4.3 Granger Causality test

In addition to the VAR-model, a Granger causality test will be conducted to better understand all of three hypotheses. A Granger causality test helps to understand which variable Granger causes the other variable. In other words, the Granger causality test will not test the real causal relationship, but the test will check the direction of the relationship between two variables. A Granger causality test can have four different outcomes: a bi-directional causality, unidirectional causality (in both directions), and no Granger causality at all (Brooks & Chris, 2019). The models used to test for a Granger causality are the following:

$$R_t = \mathrm{a}_{01} + \mathrm{a}_{11}R_{t-1} + \cdots + \mathrm{a}_{n1}R_{t-n} + \beta_{11}IA_{t-1} + \cdots + \beta_{n1}IA_{t-n} + \varepsilon_t \tag{7}$$

$$IA_t = \mathrm{a}_{02} + \mathrm{a}_{12}R_{t-1} + \cdots + \mathrm{a}_{n2}R_{t-n} + \beta_{12}IA_{t-1} + \cdots + \beta_{n2}IA_{t-n} + \epsilon_t \tag{8}$$

$$RV_t = \mathrm{a}_{03} + \mathrm{a}_{13}RV_{t-1} + \cdots + \mathrm{a}_{n3}RV_{t-n} + \beta_{13}IA_{t-1} + \cdots + \beta_{n3}IA_{t-n} + \mu_t \tag{9}$$

$$IA_t = \mathrm{a}_{04} + \mathrm{a}_{14}RV_{t-1} + \cdots + \mathrm{a}_{n4}RV_{t-n} + \beta_{14}IA_{t-1} + \cdots + \beta_{n4}IA_{t-n} + \upsilon_t \tag{10}$$

$$VLM_t = \mathrm{a}_{05} + \mathrm{a}_{15}VLM_{t-1} + \cdots + \mathrm{a}_{n5}VLM_{t-n} + \beta_{15}IA_{t-1} + \cdots + \beta_{n5}IA_{t-n} + \nu_t \tag{11}$$

$$IA_t = \mathrm{a}_{05} + \mathrm{a}_{16}VLM_{t-1} + \cdots + \mathrm{a}_{n6}VLM_{t-n} + \beta_{16}IA_{t-1} + \cdots + \beta_{n6}IA_{t-n} + \psi_t \tag{12}$$

R_t, RV_t, and VLM_t respectively represent the returns, realized volatility, and trading volume of the cryptocurrencies at week t. IA_t is the investor attention at week t and a_{01}, a_{02}, a_{03}, a_{04}, a_{05} and a_{06} are constants in the equations. Additionally, ε_t, ϵ_t, μ_t, υ_t, ν_t, and ψ_t represent the error terms in the equations.

The Granger causality tests whether the relevant coefficients are equal to zero, which means that there is no Granger causality present. For example, in equation (7) the null hypothesis says that investor attention does not Granger cause returns of the cryptocurrencies. The alternative hypothesis states that investor attention does Granger cause returns of cryptocurrencies.

$$H_0\colon \beta_{11} = \cdots = \beta_{\mathrm{n}1} = 0$$

To test the Granger causality, an F-test will be conducted (Lin, 2021).

5 RESULTS

In this chapter, the results will be discussed based on a Vector Autoregression model which is supported by a Granger causality test. First, the relationship between investor attention and the returns of cryptocurrencies will be handled. Second, the effect of investor attention on realized price volatility will be covered. Third, the relationship between investor attention and trading volumes of cryptocurrencies will be discussed. And lastly, several control variables will be added to serve as a robustness check.

5.1 Return and Investor Attention

Table 4 shows the results of the Vector Autoregression model for return and investor attention variables. The results suggest that investor attention from the four previous weeks does significantly influence the future returns of the cryptocurrencies. However, investor attention in weeks t-2 and t-4 has a negative influence on the return of the 154 cryptocurrencies. Meaning, an increase in weeks t-2 and t-4 leads to lower future returns. An explanation for this is that the market is correcting the initial price increases. Additionally, it can be concluded from Table 4 that the returns of cryptocurrencies in the previous four weeks have a significantly positive effect on investor attention. Thinking about this makes sense because if a cryptocurrency had positive returns in the last weeks, (retail) investors might expect the same returns of the cryptocurrency in the next week. Therefore, (retail) investors will pay more attention to cryptocurrency, which will result in higher search volumes.

Looking at the weeks t-1 and t-3 the results are in line with Smales (2022), Liu & Tsyvinski (2021), Subramaniam & Chakraborty (2019), and Zhang & Wang (2020). They all find that investor attention has a significant positive influence on the return of cryptocurrencies. Nevertheless, the findings at weeks t-2 and t-4 are in line with Smuts (2019). He finds that the relationship between investor attention and cryptocurrency returns is no longer consistently positive. Finally, the results correspond to the findings of Lin (2021), who finds that past returns have a positive significant effect on future investor attention.

In addition to the vector autoregression model, a Granger causality test will be conducted. As is mentioned in section 4.3, a Granger causality test helps to understand which variable Granger causes the other variable. The results from the Granger causality test in Table 5 show that there is a bi-directional Granger causal relation between investor attention and the return of cryptocurrencies. This means that investor attention not only affects returns but also the other way around. These results match with the findings of Zhang & Wang (2020) who found a bi-directional relation between investor attention and return in 11 out of a total sample of 20 cryptocurrencies. Also, the results are partially in line with Lin (2021), and Zhu et al. (2021) who found that investor attention Granger causes returns of cryptocurrencies.

Table 4 Vector autoregression results for return and investor attention

	R_t	ASVI
R_{t-1}	0.041*** (0.014)	0.438*** (0.037)
R_{t-2}	0.054*** (0.011)	0.158*** (0.034)
R_{t-3}	0.030*** (0.011)	0.198*** (0.034)
R_{t-4}	0.085*** (0.012)	0.139*** (0.032)
$ASVI_{t-1}$	0.005*** (0.002)	0.571*** (0.018)
$ASVI_{t-2}$	-0.009*** (0.003)	0.071*** (0.019)
$ASVI_{t-3}$	0.004** (0.002)	-0.022 (0.016)
$ASVI_{t-4}$	-0.006*** (0.002)	-0.067*** (0.014)

This table reports the Vector autoregression results of the analysis between returns (Rt) and investor attention ($ASVIt$). The robust standard errors are presented in the brackets. *, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively.

Table 5 Granger causality test between return and investor attention

Equation	Excluded	P-value
Returns	Investor attention	0.000***
Investor attention	Returns	0.000***

*, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively

Taking into mind the results from Tables 4 and 5, it can be concluded that investor attention and the returns of cryptocurrencies have a significant impact on each other. The results show that if returns are increasing, the attention of investors is also increasing. Additionally, a change in levels of attention by investors has a very significant effect on the returns of cryptocurrencies. These findings are strengthened by the Granger causality test, which found a bi-directional relationship between investor attention and the returns of all cryptocurrencies. Lastly, Fama's (1970) theory suggests that markets are efficient and new information will be immediately reflected in the prices of the assets. In this case, since investor attention from 2, 3 or 4 weeks ago still has an impact on the prices today, it can be concluded that the market is suffering from high inefficiencies.

5.2 Volatility and Investor Attention

The results from the vector autoregression model that used the volatility and investor attention measurement of 154 cryptocurrencies, are shown in Table 6. The results in week t-1 suggest that investor attention has a significant effect on price volatility. This means that higher investor attention in week t-1 results in higher price volatility in the future. Intuitively this makes sense because when investors pay little attention to the market, information is only partially processed into the prices since learning is slow. Therefore, low attention leads to low realized volatility (Andrei & Hasler, 2014). Contrarily, highly attentive investors directly incorporate new information into the prices and therefore cause high price volatility. In addition to this, because higher attention generates higher price volatility, investors need a higher risk premium to cover this extra risk caused by attention (Andrei & Hasler, 2014). The results correspond to those of Al Guindy (2021), Shen et al. (2019), Zhang and Wang (2020), and Zhu et al. (2021) who show that higher levels of investor attention cause an increase in price volatility.

Surprisingly, the result for week t-3 suggests that higher investor attention leads to lower price volatility in the future, which are not in line with the results found by any other of the authors. Additionally, from the same table can be concluded that previous price volatility in weeks t-2 and t-4 has a significant negative impact on investor attention in the future. This means that if price volatility increases, the attention of investors in the cryptocurrencies decreases. Intuitively, this could make sense, because if cryptocurrencies are very unstable, investors might think the cryptocurrencies are too risky to invest in and therefore the attention the cryptocurrencies receive is decreasing. Urquhart (2018) also finds evidence that realized volatility affects investor attention. However, his results suggest a positive relationship between the two variables. In that way, the results do not correspond to the result found by Urquhart (2018).

Table 6 Vector autoregression results for volatility and investor attention

	RV_t	ASVI
RV_{t-1}	0.327*** (0.029)	-0.005 (0.011)
RV_{t-2}	0.104*** (0.022)	-0.023** (0.010)
RV_{t-3}	0.081*** (0.023)	0.005 (0.011)
RV_{t-4}	0.118*** (0.022)	-0.019* (0.010)
$ASVI_{t-1}$	0.022** (0.011)	0.609*** (0.018)
$ASVI_{t-2}$	0.012 (0.012)	0.074*** (0.019)
$ASVI_{t-3}$	-0.020* (0.011)	-0.022 (0.017)
$ASVI_{t-4}$	0.000 (0.008)	-0.063*** (0.014)

This table reports the Vector autoregression results of the analysis between realized volatility (RVt) and investor attention ($ASVIt$). The robust standard errors are presented in the brackets. *, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively.

Similar to section 5.1, a Granger causality test will be conducted. The results from the Granger causality test in Table 7 show that there is a unidirectional relationship between price volatility and investor attention. According to the results in Table 7, investor attention Granger causes price volatility and not the other way around. These findings are in line with earlier literature of Zhu et al. (2021) who also found a unidirectional Granger causal relationship. The results do not completely correspond with the results of Shen et al (2019), and Zhang and Wang (2020). They both find bi-directional relationships between price volatility and investor attention.

Table 7 Granger causality test between volatility and investor attention

Equation	Excluded	P-value
Volatility	Investor attention	0.026**
Investor attention	Volatility	0.102

*, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively

To conclude the results based on Tables 6 and 7, investor attention has a significant positive influence on realized volatility. This means that when the levels of attention increase among investors, it causes more price volatility in the cryptocurrency market. This could be explained by high-attentive investors, who incorporated information faster into prices, which cause an increase in realized price volatility.

Therefore, taking in that higher investor attention leads to higher returns, it can be concluded that higher (lower) returns is associated with higher (lower) price volatility.

On the other hand, it can be seen that realized volatility has a negative significant effect on investor attention. As explained, this could make sense if investors are deterred by higher risk in the market, which causes investors' attention to decrease. However, this result is not supported by the Granger causality test that is conducted, which only found a unidirectional relationship.

5.3 Volume and Investor Attention

The VAR results for trading volume and investor attention are displayed in Table 8. Matching the expectations based on previous research, the results show a strong positive significant relationship between trading volume and investor attention in the first previous week. Meaning that an increase in Google search volume on cryptocurrencies in the previous week results in higher trading volume of the cryptocurrencies in the future. Logically, if attention of investors increases it also increases the volume of trading in the cryptocurrency market. However, from the same Table it can be concluded that the effect of an increase in trading volume will be partially neutralized in the next week, due to the negative coefficient in the second lag of investor attention. Taking in mind the results of section 5.1 this could mean that investors are realizing the market is correcting the price increases and therefore trading volume decreases. These results correspond with Nasir et al. (2019) who find that an increase in Google search volumes has a positive impact on the trading volume of cryptocurrencies. Also, the results record a diminishing effect in the following weeks. Additionally, the results show a significant positive relationship between investor attention and trading volume in the first previous week. This implies that a positive shock to trading volumes is associated with higher levels of attention by investors. This can be reasonable if you think that abnormally high trading volumes alert investors, which can cause Google search volumes to increase. The results are in line with Urquhart (2018) who finds that trading volume is an important determinant of next-day investor attention. However, his paper does not find that investor attention has any predicting power over trading volume. Which contradicts some of the results this paper founds.

Table 8 Vector autoregression results for volume and investor attention

	VLM_t	**ASVI**
VLM_{t-1}	0.714*** (0.019)	0.026*** (0.004)
VLM_{t-2}	0.072*** (0.017)	-0.017*** (0.004)
$ASVI_{t-1}$	0.135*** (0.024)	0.597*** (0.019)
$ASVI_{t-2}$	-0.055** (0.022)	0.028 (0.019)

This table reports the Vector autoregression results of the analysis between trading volume ($VLMt$) and investor attention ($ASVIt$). The robust standard errors are presented in the brackets. *, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively.

The results from the Granger causality in Table 9 show that there is a bi-directional Granger causality between trading volume and investor attention. This means that trading volume not only has an impact on investor attention but also the other way around. Urquhart (2018) found in his research that trading volume Granger causes investor attention. An explanation for the contradicting results can be that Urquhart's research is specified on only Bitcoin, and this paper uses a cross-sectional research method. Also, the results contradict those of Nasir et al. (2019) who found no Granger causality between trading volume and investor attention. But again, they only examined Bitcoin which can be an explanation for those different results.

Table 9 Granger causality test between volume and investor attention

Equation	**Excluded**	**P-value**
Volume	Investor attention	0.000***
Investor attention	Volume	0.000***

*, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively

Not contrary to expectations, it can be concluded that there is a very strong and significant relationship between investor attention and trading volume. The results show, logically, that higher levels of attention by investors cause higher volumes of trading in the cryptocurrency market. Nevertheless, these higher levels of trading volume diminish in the next week. Additionally, an increase in trading volume causes higher levels of attention from investors. Both results are supported by the Granger causality test, which shows a bi-directional relationship between investor attention and trading volume.

5.4 Control Variables

Five control variables are added to an Autoregressive Distributed Lag (ARDL)-model to serve as a robustness check. An ARDL-model is fit to use as a VAR-model including multiple (control) variables with lags that vary over time (Stock & Watson, 2019). These control variables could each influence the returns, trading volume, price volatility and investor attention. The first added control variable is the weekly return of the S&P500, which is one of the most commonly used indicators for the performance of the American economy. For example, Georgoula et al. (2015) who, due to a vector error correction model, found a negative correlation between Bitcoin's price and the S&P500 market index.

The second and third control variables are the weekly price change of (WTI) oil price and the weekly price change of gold. Both assets have been commonly investments for investors who want to hedge against economic instability. Cryptocurrencies could provide investors, who want to hedge against this instability, with a new hedging alternative. For example, Okorie and Lin (2020) show that there is a bi-directional relationship between volatility spillovers from the oil market to the cryptocurrency market and the other way around. Also, Zeng et al. (2020) find that there is a weak connection between the price of gold and the price of Bitcoin. However, there is still a debate going on because Kjærland et al. (2018) and Erdas & Caglar (2018) found that there is no significant relationship between oil, gold and the cryptocurrency market.

The fourth control variable that is added is the VIX, which is broadly used by researchers as a proxy for the volatility of the S&P500. Ghorbel & Jeribi (2021) found that current conditional stock indices (S&P500 and VIX) not only depend on their previous volatility but also on the past volatility of cryptocurrencies. This shows that the cryptocurrency market and the traditional financial market are connected.

The last control variable that is used is the total market capitalization for each cryptocurrency. This is done to minimize the endogeneity because it could be the case that there could be more attention for a single cryptocurrency among investors. After all, the cryptocurrency is getting bigger. Therefore, the control variable market capitalization is used to control for the size of the cryptocurrency.

The ARDL-model that will be used can be found below in equation 13:

$$X_t = c + \sum_{i=1}^{P} \beta_i X_{t-i} + \sum_{i=1}^{P} \delta_i SP500_{t-i} + \sum_{i=1}^{P} \gamma_i OIL_{t-i} + \sum_{i=1}^{P} \theta_i GOLD_{t-i} + \sum_{i=1}^{P} \vartheta_i VIX_{t-i} + \sum_{i=1}^{P} \rho_i MARKET_{t-i} + \varepsilon_t \quad (13)$$

Where X_t is the variable of interest (return, volatility, trading volume, and the SVI), c is a constant, and ε_t is the error term. $\beta_i, \delta_i, \gamma_i, \theta_i, \vartheta_i,$ and ρ_i are the coefficient expressing the effect of the lagged component, and ρ captures the models lag duration.

5.4.1 Data Modifications and Method

Weekly data of the control variables (S&P500 index, WTI oil price, gold price and VIX) is derived from finance.yahoo.com. The descriptive statistics about the control variables can be found in Appendix C. The returns of the control variables have the same calculation as the return of cryptocurrencies calculated in equation 2 in section 3.2.2. Data about the market capitalization of each cryptocurrency is calculated from the daily data available on Coingecko.com. A weekly average is taken to match the same timespan as other variables. Thereafter, the logarithm is taken so it could correct certain outliers.

The same methods will be applied as in chapter 4, but now including the control variables. The first step is to check if the control variables are stationary with the Fisher-type unit root test. The results of this test are presented in Table 10. From Table 10 it can be seen that all the variables are stationary. Noteworthy, the inverse normal and inverse logit statistics for the variable LogMarketCap are not significant. However, when the number of panels is finite, the inverse chi-squared is applicable Choi (2001). Appendix D shows the result of the lag length determination for the VAR-models including the control variables. Again, the Bayesian information criterion will be handled so each model will use four lags.

Table 10 Fisher-type unit root test for control variables

Variable	Test	Statistic
S&P500	Inversed chi-squared *(308)*	9068.0415***
	Inverse normal	-90.3765***
	Inverse logit *(774)*	-201.5649***
	Modified inv. chi-squared	352.9521***
WTI_Oil	Inversed chi-squared *(308)*	8446.7878***
	Inverse normal	-86.9378***
	Inverse logit *(774)*	-187.7557***
	Modified inv. chi-squared	327.9211***
Gold	Inversed chi-squared *(308)*	1.06e+04***

	Inverse normal	-98.5173***
	Inverse logit *(774)*	-236.3177***
	Modified inv. chi-squared	415.9458***
VIX	Inversed chi-squared *(308)*	1.11e+04***
	Inverse normal	-100.8397***
	Inverse logit *(774)*	-246.7636***
	Modified inv. chi-squared	434.8803***
LogMarketCap	Inversed chi-squared *(308)*	374.6530**
	Inverse normal	0.5780
	Inverse logit *(774)*	-0.7153
	Modified inv. chi-squared	2.7751**

*, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively

5.4.2 Return

Comparing the results in Table 11 of the VAR model including the control variables with the VAR model without the control variables, multiple things can be noticed. Looking at the effects on investor attention, the first thing that can be seen is that returns of the cryptocurrencies still has a significant positive effect. It is also noticeable that after adding the control variables, the coefficient for the weeks t-1, t-2 and t-3 is slightly higher compared to the results without control variables for returns. This means that the effect of positive returns has a stronger effect on the attention of investors, meaning that the attention will increase more. Another thing to notice is that some lags of the control variables (S&P500, WTI Oil price, gold price, VIX and market capitalization) have a significant effect on the level of attention from investors. This contradicts the earlier findings of Kjærland et al. (2018) and Erdas & Caglar (2018) who found that there is no significant relationship between the prices of oil and gold and the cryptocurrency market.

Looking at the effects on returns of the cryptocurrencies, it can be seen that the effects of investor attention are relatively less significant for the first three previous weeks than the effects before adding the control variables. However, the sign and magnitude of the coefficient are almost identical. But the most notable observation from Table 11 is that most of the alternative investment options, like WTI oil, Gold and VIX, have a significant negative effect on the return of cryptocurrency. This means that if those markets are performing badly, investors are probably switching preferences and are more interested in investing in cryptocurrencies which results in a positive price shock in the cryptocurrency market. However, an opposing argument for this theory is the significant positive effect of the S&P500.

This finding is also not in line with Georgoula et al. (2015) who found a negative significant relationship between the S&P500 index and the price of Bitcoin.

Table 11 Vector autoregression with the returns, investor attention and control variables

	R_t	ASVI	R_t	ASVI
R_{t-1}	0.021 (0.016)	0.515*** (0.054)	0.041*** (0.014)	0.438*** (0.037)
R_{t-2}	0.070*** (0.019)	0.251*** (0.063)	0.054*** (0.011)	0.158*** (0.034)
R_{t-3}	0.083*** (0.017)	0.207*** (0.058)	0.030*** (0.011)	0.198*** (0.034)
R_{t-4}	0.118*** (0.017)	0.089** (0.041)	0.085*** (0.012)	0.139*** (0.032)
$ASVI_{t-1}$	0.005* (0.003)	0.578*** (0.019)	0.005*** (0.002)	0.571*** (0.018)
$ASVI_{t-2}$	-0.007** (0.003)	0.068*** (0.020)	-0.009*** (0.003)	0.071*** (0.019)
$ASVI_{t-3}$	0.004 (0.003)	-0.015 (0.018)	0.004** (0.002)	-0.022 (0.016)
$ASVI_{t-4}$	-0.006*** (0.002)	-0.071*** (0.015)	-0.006*** (0.002)	-0.067*** (0.014)
$S\&P500_{t-1}$	0.061*** (0.013)	0.185*** (0.054)		
$S\&P500_{t-2}$	0.139*** (0.015)	0.298*** (0.054)		
$S\&P500_{t-3}$	0.074*** (0.014)	0.122** (0.057)		
$S\&P500_{t-4}$	-0.025** (0.013)	-0.066 (0.055)		
WTI_Oil_{t-1}	-0.020*** (0.002)	0.016* (0.010)		
WTI_Oil_{t-2}	-0.005** (0.002)	-0.027*** (0.009)		
WTI_Oil_{t-3}	0.003* (0.002)	-0.014* (0.008)		
WTI_Oil_{t-4}	-0.020*** (0.002)	0.002 (0.009)		
$Gold_{t-1}$	-0.098*** (0.011)	0.018 (0.043)		
$Gold_{t-2}$	0.031*** (0.011)	0.087** (0.043)		
$Gold_{t-3}$	-0.116*** (0.012)	0.030 (0.048)		
$Gold_{t-4}$	-0.077*** (0.011)	-0.035 (0.047)		
VIX_{t-1}	-0.019*** (0.003)	0.011 (0.011)		
VIX_{t-2}	0.021*** (0.003)	0.052*** (0.012)		

VIX_{t-3}	0.020*** (0.003)	0.043*** (0.012)
VIX_{t-4}	-0.008*** (0.003)	-0.023** (0.011)
$LogMarketCap_{t-1}$	0.008 (0.007)	-0.048** (0.021)
$LogMarketCap_{t-2}$	-0.029*** (0.009)	0.036 (0.026)
$LogMarketCap_{t-3}$	0.016* (0.009)	0.025 (0.026)
$LogMarketCap_{t-4}$	0.010* (0.006)	-0.008 (0.017)

This table reports the Vector autoregression results of the analysis between returns (Rt), investor attention ($ASVIt$) and control variables in panel 1 and 2. Panel 3 and 4 are the results obtained from previous sections, but are displayed again so it is easier to compare the results. The robust standard errors are presented in the brackets. *, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively.

5.4.3 Volatility

The results of the added control variables to the VAR model with realized volatility and investor attention are shown in Table 12. Looking at the variables of interest investor attention and realized volatility the differences in adding the control variables are almost negligible. The results for investor attention on realized volatility and vice versa, almost have the same sign, magnitude of coefficient, and have the same significance. A notable observation is that S&P500 again is strongly connected to the cryptocurrency market, with this time significant effects on both realized volatility and investor attention. Again, the findings are not in line with the work of Kjærland et al. (2018) and Erdas & Caglar (2018), they did not find any correlation between the prices of oil and gold and the cryptocurrency market. Nevertheless, the results are in line with Okorie and Lin (2020) who show that there is a bi-directional relationship between volatility spillovers from the oil market to the cryptocurrency.

Table 12 Vector autoregression with realized volatility, investor attention and control variables

	RV_t	**ASVI**	**RV_t**	**ASVI**
RV_{t-1}	0.273*** (0.037)	-0.019 (0.012)	0.327*** (0.029)	-0.005 (0.011)
RV_{t-2}	0.059** (0.028)	-0.027** (0.011)	0.104*** (0.022)	-0.023** (0.010)
RV_{t-3}	0.049* (0.028)	-0.004 (0.012)	0.081*** (0.023)	0.005 (0.011)
RV_{t-4}	0.125*** (0.026)	-0.019* (0.010)	0.118*** (0.022)	-0.019* (0.010)
$ASVI_{t-1}$	0.021* (0.011)	0.595*** (0.019)	0.022** (0.011)	0.609*** (0.018)

$ASVI_{t-2}$	-0.005 (0.011)	0.066*** (0.020)	0.012 (0.012)	0.074*** (0.019)
$ASVI_{t-3}$	-0.019* (0.010)	-0.015 (0.018)	-0.020* (0.011)	-0.022 (0.017)
$ASVI_{t-4}$	0.009 (0.008)	-0.067*** (0.015)	0.000 (0.008)	-0.063*** (0.014)
$S\&P500_{t-1}$	0.261*** (0.065)	0.155*** (0.056)		
$S\&P500_{t-2}$	-0.376*** (0.069)	0.282*** (0.058)		
$S\&P500_{t-3}$	0.216*** (0.072)	0.107* (0.061)		
$S\&P500_{t-4}$	0.280*** (0.061)	-0.101* (0.058)		
WTI_Oil_{t-1}	0.063*** (0.010)	0.016* (0.010)		
WTI_Oil_{t-2}	0.013 (0.011)	-0.037*** (0.009)		
WTI_Oil_{t-3}	-0.017** (0.007)	-0.012 (0.008)		
WTI_Oil_{t-4}	0.051*** (0.007)	0.003 (0.009)		
$Gold_{t-1}$	0.204*** (0.050)	0.003 (0.043)		
$Gold_{t-2}$	0.030 (0.044)	0.053 (0.043)		
$Gold_{t-3}$	0.385*** (0.051)	0.046 (0.048)		
$Gold_{t-4}$	0.166*** (0.044)	-0.000 (0.047)		
VIX_{t-1}	0.141*** (0.014)	-0.004 (0.011)		
VIX_{t-2}	0.002 (0.012)	0.048*** (0.012)		
VIX_{t-3}	0.073*** (0.012)	0.040*** (0.012)		
VIX_{t-4}	0.043*** (0.012)	-0.025** (0.012)		
$LogMarketCap_{t-1}$	-0.002 (0.025)	0.095*** (0.015)		
$LogMarketCap_{t-2}$	0.090** (0.037)	-0.055** (0.022)		
$LogMarketCap_{t-3}$	-0.111*** (0.032)	0.018 (0.020)		
$LogMarketCap_{t-4}$	0.005 (0.019)	-0.054*** (0.013)		

This table reports the Vector autoregression results of the analysis between realized volatility (RVt), investor attention ($ASVIt$) and control variables in panel 1 and 2. Panel 3 and 4 are the results obtained from previous sections, but are displayed again so it is easier to compare the results. The robust standard errors are presented in the brackets. *, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively.

5.4.4 Volume

Lastly, the results of the added control variables to the VAR model with realized volatility and investor attention are shown in Table 13. Looking at the effects on investor attention, the results do somewhat differ from the results without control variables. The results show a positive significant effect of trading volume in week t-1 at a significance level of 5%. Without adding the control variables this yields a significance level of 1%. Additionally, after adding the control variables the coefficient slightly decreased for the first previous week and slightly increased for the second previous week. Also, it can be concluded that trading volumes three and four weeks ago from now have no significant effect on the level of attention from investors.

Looking at the effects on trading volume, the results do not differ much from the results obtained without control variables. Again, only the first two previous weeks of investor attention show significant results. The results imply that higher investor attention in week t-1 leads to higher trading volumes, this effect diminishes in the next week, just like the results found in section 5.3. Again, the S&P500 index is very strongly connected to the trading volume of cryptocurrencies. Suggesting that higher returns in the S&P500 would result in higher trading volumes on the cryptocurrency market. Additionally, the positive and significant effect of market capitalization in week t-1 suggests that trading volume is higher for the biggest cryptocurrencies. Intuitively, this makes sense since coins with higher market capitalization are more known among investors and therefore the trading volume is higher. Moreover, higher market capitalization also means higher liquidity. as liquidity increases, it becomes easier for investors to trade which has a positive effect on trading volume. However, the results for t-2 and t-4 suggest otherwise where a negative and significant effect is established.

Table 13 Vector autoregression with trading volume, investor attention and control variables

	VLM_t	**ASVI**	**VLM_t**	**ASVI**
VLM_{t-1}	0.653*** (0.025)	0.011** (0.005)	0.714*** (0.019)	0.026*** (0.004)
VLM_{t-2}	0.066*** (0.025)	-0.011* (0.006)	0.072*** (0.017)	-0.017*** (0.004)
VLM_{t-3}	0.039** (0.017)	0.009 (0.006)		
VLM_{t-4}	0.014 (0.015)	-0.005 (0.005)		
$ASVI_{t-1}$	0.102*** (0.026)	0.595*** (0.020)	0.135*** (0.024)	0.597*** (0.019)
$ASVI_{t-2}$	-0.065** (0.029)	0.063*** (0.021)	-0.055** (0.022)	0.028 (0.019)

$ASVI_{t-3}$	-0.022 (0.028)	-0.014 (0.019)
$ASVI_{t-4}$	0.021 (0.022)	-0.073*** (0.016)
$S\&P500_{t-1}$	0.804*** (0.120)	0.192*** (0.056)
$S\&P500_{t-2}$	0.812*** (0.129)	0.312*** (0.056)
$S\&P500_{t-3}$	0.829** (0.129)	0.146** (0.058)
$S\&P500_{t-4}$	0.163 (0.125)	-0.066 (0.056)
WTI_Oil_{t-1}	0.015 (0.024)	0.017* (0.010)
WTI_Oil_{t-2}	-0.093*** (0.021)	-0.038*** (0.010)
WTI_Oil_{t-3}	-0.084*** (0.022)	-0.011 (0.008)
WTI_Oil_{t-4}	-0.042** (0.020)	0.003 (0.009)
$Gold_{t-1}$	0.160* (0.095)	0.026 (0.044)
$Gold_{t-2}$	-0.039 (0.099)	0.053 (0.043)
$Gold_{t-3}$	0.161 (0.110)	0.040 (0.049)
$Gold_{t-4}$	-0.205* (0.108)	-0.015 (0.048)
VIX_{t-1}	0.108*** (0.025)	-0.001 (0.011)
VIX_{t-2}	0.091*** (0.026)	0.047*** (0.012)
VIX_{t-3}	0.152*** (0.027)	0.045*** (0.012)
VIX_{t-4}	0.009 (0.015)	-0.026** (0.012)
$LogMarketCap_{t-1}$	0.373*** (0.052)	0.079*** (0.017)
$LogMarketCap_{t-2}$	-0.287*** (0.075)	-0.048** (0.024)
$LogMarketCap_{t-3}$	0.035 (0.062)	0.011 (0.021)
$LogMarketCap_{t-4}$	-0.104** (0.041)	-0.037* (0.015)

This table reports the Vector autoregression results of the analysis between trading volume ($VLMt$), investor attention ($ASVIt$) and control variables in panel 1 and 2. Panel 3 and 4 are the results obtained from previous sections, but are displayed again so it is easier to compare the results. The robust standard errors are presented in the brackets. *, ** and *** represent the significance of the t-statistics at 10%, 5% and 1%, respectively.

6. CONCLUSION

This paper studied the relationship between investor attention and the cryptocurrency market. This paper tries to answer the question: *What is the effect of investor attention on the returns, volatility, and trading volume in the cryptocurrency market, from the period April 1st 2019 until March 31st 2022?* The research question is answered with the help of three hypotheses. The conclusion will be constructed by first giving a conclusion for each hypothesis followed by an overall conclusion. In the end, the limitations will be covered and suggestions will be made for further research.

6.1 Discussion

The first hypothesis is '*Investor attention has a significant effect on cryptocurrencies' returns'.* The results show indeed a significant positive effect of investor attention on the returns on the cryptocurrency market. Meaning that an increase in the level of attention by investors increases the returns on the cryptocurrency market. Additionally, this research also found a significant positive effect of the returns of cryptocurrencies on the levels of attention by investors. The results support the idea that, if cryptocurrencies are performing better and returns are increasing, it generates more attention from investors. After adding the control variables, the same relationships between two variables were observed. The hypothesis is even more supported by the results of the Granger causality test that showed a bi-directional relationship between investor attention and the returns of cryptocurrencies. All this together makes enough to accept the first hypotheses.

The second hypothesis is '*Investor attention has a significant effect on cryptocurrencies' realized price volatility'.* From this paper, it can be concluded that the second hypothesis is supported by the results. The paper found a significant positive effect of investor attention on realized price volatility. Meaning that an increase in the attention of investors also causes an increase in the realized price volatility. This result is supported by the Granger causality test that showed that investor attention affects realized price volatility. These results show that the attention of investors is an important aspect of cryptocurrency markets becoming efficient. When there is less attention, realized price volatility is lower which suggests that prices might not be corrected as much as what would be optimal. This makes the market less efficient. On the other hand, the results show a negative significant relationship between realized volatility and the attention of investors. Meaning that if prices become more volatile, the attention of investors will decrease. However, the Granger causality test did not support this finding.

The third and last hypothesis is *'Investor attention has a significant effect on cryptocurrencies' trading volume'.* This hypothesis has also been accepted. The results indicate that higher levels of attention cause higher volumes of trading in the cryptocurrency market, as expected from theory. However, these effects dimmish in the next week. Moreover, the results show also that higher trading volumes cause investors to pay more attention because the level of attention is increasing. Both results are supported by the Granger causality test, since this yields a bi-directional relationship. Also, after adding the control variables to the model, the hypothesis still is accepted.

There is still a lot of research going on to better understand the cryptocurrency market and the valuation of cryptocurrencies. But this research shows that in order to better understand the cryptocurrency market, one must look beyond the classical financial theories and include behavioural finance theories in their research. Because the results show investor attention has a very strong impact on the cryptocurrency market. If investor attention increases, it cause higher returns, higher realized price volatility, and higher trading volumes in the cryptocurrency market. However, the cryptocurrency market still shows some inefficiencies that investor attention is not able to clarify. Therefore, other behavioural theories have to be found to answer these problems.

6.2 Limitations and Further Research

Unfortunately, this research has a couple of limitations. One of the main limitations of this research is that Google search volume cannot capture whether the attention on the cryptocurrency is from good or bad news. Therefore, it is difficult to assess the full effect of investor attention on cryptocurrencies. To better understand the attention of investors in the cryptocurrency market, it would be very helpful to get more insight into the sentiment of the cryptocurrency investors. Therefore, a suggestion for further research is to add the investor sentiment analysis to the analysis of investor attention.

Another shortcoming of using Google search volume as a proxy for investor attention is that it captures only the attention of retail investors. Most professional investors use other tools like Reuters and Bloomberg. Therefore, probably not all the investors are included in the sample, thus not 'all the attention' is used to do the analysis. Hence, a suggestion for further research is to include also the attention of professional investors. By adding this extra attention to the sample, stronger or new relationships may be found. In addition, Google search volume does also capture the attention of people that are not interested in investing. To solve this problem, the tickers of each cryptocurrency are used as a search term. However, this still might not completely capture all the attention of investors. Therefore, a small sidenote has to be made next to the results. To counter this problem, additional proxies such as number of tweets, for investor attention could be added so it captures a broader range of attention.

Some belief that (most) cryptocurrencies can be an effective hedge against inflation, because (most) cryptocurrencies have a capped supply. However, if you look at recent price decreases in the cryptocurrency market, others will argue that the hedge is not as effective as some have suggested. Therefore, it is interesting to see what effect the inflation has on the cryptocurrency market. In addition to this suggestion, in times of high inflation, it might be interesting to look at the real returns of the cryptocurrencies. That are returns that are corrected for price inflation.

Another limitation of this paper is the availability of Google Trends data. For periods longer than 3 months, only weekly data is available on Google Trends. Thus, daily data about the cryptocurrencies have been adjusted to weekly averages so it matches the same weekly data as Google Trends. However, since prices of cryptocurrencies are very volatile it is preferred to use daily data because it will probably give more complete results. Some cryptocurrencies will have huge volume, price and volatility shocks in just a few hours. Lastly, besides the added control variables, there could be other variables that are still left out that could have a strong relationship with the dependent variables and cause an endogeneity problem.

Reference List

Abu Bakar, N., & Rosbi, S. (2017). High volatility detection method using statistical process control for cryptocurrency exchange rate: A case study of Bitcoin. *The International Journal of Engineering and Science*, *6*(11), 39-48.

Al Guindy, M. (2021). Cryptocurrency price volatility and investor attention. *International Review of Economics & Finance*, *76*, 556–570. https://doi.org/10.1016/j.iref.2021.06.007

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, *71*(2), 579–625. https://doi.org/10.1111/1468-0262.00418

Andrei, D., & Hasler, M. (2014). Investor Attention and Stock Market Volatility. *Review of Financial Studies*, *28*(1), 33–72. https://doi.org/10.1093/rfs/hhu059

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, *21*(2), 129-152.

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. https://doi.org/10.1007/s11408-011-0165-y

Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, *21*(2), 785–818. https://doi.org/10.1093/rfs/hhm079

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets?. *Journal of International Financial Markets, Institutions and Money*, *54*, 177-189.

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, *45*, 150–156. https://doi.org/10.1016/j.irfa.2016.03.015

Blankespoor, E., de Haan, E., & Zhu, C. (2017). Capital market effects of media synthesis and dissemination: evidence from robo-journalism. *Review of Accounting Studies*, *23*(1), 1–36. https://doi.org/10.1007/s11142-017-9422-2

Brooks, C. & Chris, B. (2019). *Introductory Econometrics for Finance* (4th Revised edition). Cambridge University Press.

Caferra, R., & Vidal-Tomás, D. (2021). Who raised from the abyss? A comparison between cryptocurrency and stock market dynamics during the COVID-19 pandemic. *Finance Research Letters*, *43*, 101954. https://doi.org/10.1016/j.frl.2021.101954

Cheah, E. T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Chemmanur, T., & Yan, A. (2009). Product market advertising and new equity issues. *Journal of Financial Economics*, *92*(1), 40–65. https://doi.org/10.1016/j.jfineco.2008.02.009

Choi, I. 2001. Unit root tests for panel data. Journal of International Money and Finance 20: 249–272. https://doi.org/10.1016/S0261-5606(00)00048-6.

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *The Journal of Finance*, *66*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Erdas, M. L., & Caglar, A. E. (2018). Analysis of the relationships between Bitcoin and exchange rate, commodities and global indexes by asymmetric causality test. *Eastern Journal of European Studies*, 9(2), 27.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, *25*(2), 383–417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Fang, L., & Peress, J. (2009). Media Coverage and the Cross-section of Stock Returns. *The Journal of Finance*, *64*(5), 2023–2052. https://doi.org/10.1111/j.1540-6261.2009.01493.x

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C., & Siering, M. (2014). Bitcoin-asset or currency? revealing users' hidden intentions. *Revealing Users' Hidden Intentions (April 15, 2014). ECIS*.

García-Monleón, F., Danvila-del-Valle, I., & Lara, F. J. (2021). Intrinsic value in crypto currencies. *Technological Forecasting and Social Change*, *162*, 120393. https://doi.org/10.1016/j.techfore.2020.120393

Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D. N. & Giaglis, G. M. (2015). Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2607167

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The High-Volume Return Premium. *The Journal of Finance*, *56*(3), 877–919. https://doi.org/10.1111/0022-1082.00349

Ghorbel, A. & Jeribi, A. (2021). Investigating the relationship between volatilities of cryptocurrencies and other financial assets. *Decisions in Economics and Finance*, *44*(2), 817–843. https://doi.org/10.1007/s10203-020-00312-9

Grullon, G., Kanatas, G., & Weston, J. P. (2003). Advertising, Breadth of Ownership, and Liquidity. *Review of Financial Studies*, *17*(2), 439–461. https://doi.org/10.1093/rfs/hhg039

Härdle, W. K., Harvey, C. R., & Reule, R. C. G. (2020). Understanding Cryptocurrencies*. *Journal of Financial Econometrics*, *18*(2), 181–208. https://doi.org/10.1093/jjfinec/nbz033

Hou, K., Xiong, W., & Peng, L. (2009). A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.976394

Kahneman, D. (1973). *Attention and effort (Prentice-Hall series in experimental psychology)* (First Edition). Prentice-Hall.

Kjærland, F., Khazal, A., Krogstad, E. A., Nordstrøm, F. B., & Oust, A. (2018). An analysis of bitcoin's price dynamics. *Journal of Risk and Financial Management*, 11(4), 63.

Kraaijeveld, O. & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, *65*, 101188. https://doi.org/10.1016/j.intfin.2020.101188

Kyriazis, N., Papadamou, S., & Corbet, S. (2020). A systematic review of the bubble dynamics of cryptocurrency prices. *Research in International Business and Finance*, *54*, 101254. https://doi.org/10.1016/j.ribaf.2020.101254

Kuo Chuen, D. L., Guo, L., & Wang, Y. (2017). Cryptocurrency: A New Investment Opportunity? *The Journal of Alternative Investments*, *20*(3), 16–40. https://doi.org/10.3905/jai.2018.20.3.016

Lansky, J. (2018). Possible State Approaches to Cryptocurrencies. *Journal of Systems Integration*, *9*(1), 19–31. https://doi.org/10.20470/jsi.v9i1.335

Lin, Z. Y. (2021). Investor attention and cryptocurrency performance. *Finance Research Letters*, *40*. https://doi.org/10.1016/j.frl.2020.101702

Liu, Y., & Tsyvinski, A. (2021). Risks and Returns of Cryptocurrency. *The Review of Financial Studies*, *34*(6), 2689–2727. https://doi.org/10.1093/rfs/hhaa113

Llorente, G., Michaely, R., Saar, G., & Wang, J. (2002). Dynamic Volume-Return Relation of Individual Stocks. *Review of Financial Studies*, *15*(4), 1005–1047.
https://doi.org/10.1093/rfs/15.4.1005

Lou, D. (2014). Attracting Investor Attention through Advertising. *Review of Financial Studies*, *27*(6), 1797–1829. https://doi.org/10.1093/rfs/hhu019

Lubbersen, V. & Wierts, P. (2022, november). *Crypto-assets: Evolution and policy response*. De Nederlandsche Bank. https://www.dnb.nl/media/ieliqxqc/dnb-occasional-study-crypto-s-engels.pdf

Merton, R. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance, 42*, 483–510.

Mushtaq, R. (2011). Augmented Dickey Fuller Test. *Social Science Research Network*. https://doi.org/10.2139/ssrn.1911068

Mondiale cryptocurrencygrafieken. (2022, 10 juli). CoinGecko. Visited on:
https://www.coingecko.com/nl/global_charts

Nasir, M. A., Huynh, T. L. D., Nguyen, S. P. & Duong, D. (2019). Forecasting cryptocurrency returns and volume using search engines. *Financial Innovation*, *5*(1). https://doi.org/10.1186/s40854-018-0119-8

Nason, G. P. (2006). Stationary and non-stationary time series. *Statistics in Volcanology*, *60*. https://doi.org/10.1144/IAVCEI001.11

Okorie, D. I., & Lin, B. (2020). Crude oil price and cryptocurrencies: Evidence of volatility connectedness and hedging strategy. *Energy Economics*, 87, 104703.

Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical Finance*, *14*(5), 590–610. https://doi.org/10.1016/j.jempfin.2007.03.002

Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict Bitcoin? *Economics Letters*, *174*, 118–122. https://doi.org/10.1016/j.econlet.2018.11.007

Smales, L. (2022). Investor attention in cryptocurrency markets. *International Review of Financial Analysis*, *79*, 101972. https://doi.org/10.1016/j.irfa.2021.101972

Smuts, N. (2019). What drives cryptocurrency prices? an investigation of google trends and telegram sentiment. *ACM SIGMETRICS Performance Evaluation Review*, *46*(3), 131-134.

Stock, J. H., & Watson, M. W. (2019). *Introduction to econometrics* (Vol. 104). Boston: Addison Wesley.

Subramaniam, S., & Chakraborty, M. (2019). Investor Attention and Cryptocurrency Returns: Evidence from Quantile Causality Approach. *Journal of Behavioral Finance*, *21*(1), 103–115. https://doi.org/10.1080/15427560.2019.1629587

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, *148*, 80–82. https://doi.org/10.1016/j.econlet.2016.09.019

Urquhart, A. (2018). What causes the attention of Bitcoin? *Economics Letters*, *166*, 40–44. https://doi.org/10.1016/j.econlet.2018.02.017

White, R., Marinakis, Y., Islam, N., & Walsh, S. (2020). Is Bitcoin a currency, a technology-based product, or something else?. *Technological Forecasting and Social Change*, *151*, 119877.

Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. In *Handbook of digital currency* (pp. 31-43). Cambridge: Academic Press.

Zeng, T., Yang, M., & Shen, Y. (2020). Fancy Bitcoin and conventional financial assets: Measuring market integration based on connectedness networks. *Economic Modelling*, 90, 209-220.

Zhang, W., & Wang, P. (2020). Investor attention and the pricing of cryptocurrency market. *Evolutionary and Institutional Economics Review*, *17*(2), 445–468. https://doi.org/10.1007/s40844-020-00182-1

Zhu, P., Zhang, X., Wu, Y., Zheng, H., & Zhang, Y. (2021). Investor attention and cryptocurrency: Evidence from the Bitcoin market. *PLOS ONE*, *16*(2). https://doi.org/10.1371/journal.pone.0246331

Appendix A: List of Cryptocurrencies

Coin	Ticker	Coin	Ticker	Coin	Ticker
Bitcoin	BTC	Fantom	FTM	Pax Gold	PAXG
tenset	10SET	FTX	FTT	Playdapp	PLA
1inch	1INCH	Frax share	FXS	Polymath	POLY
Aave	AAVE	Moonbeam	GLMR	Quant	QNT
Cardano	ADA	GMX	GMX	Qtum	QTUM
Algogrand	ALGO	Gnosis	GNO	Render	RNDR
Alchemix USD	ALUSD	The graph	GRT	Rocked Pool	RPL
Amp	AMP	Gate	GT	Reserve Rights	RSR
Ankr	ANKR	Gemini Dollar	GUSD	Ravencoin	RVN
Arweave	AR	Hedera	HBAR	Safemoon [OLD]	SAFEMOON
Avalanche	AVAX	Huobi BTC	HBTC	Siacoin	SC
Axie infinity	AXS	HIVE	HIVE	Secret	SCRT
Baby Doge Coin	BABYDOGE	Helium	HNT	Skale	SKL
Bitcoin Cash	BCH	huobi	HT	Smooth love Potion	SLP
BNB	BNB	Internet Computer	ICP	Synthetix Network	SNX
Bitcoin SV	BSV	Icon	ICX	Solana	SOL
Bitcoin Gold	BTG	Immutable X	IMX	Serum	SRM
Bittorrent	BTT	IOST	IOST	Lido Staked Ether	STETH
Binance USD	BUSD	Iotex	IOTX	Stacks	STX
cDai	CDAI	Just	JST	Sxp	SXP
Ceek smart vr	CEEK	Kava	KAVA	Synapse	SYN
Celo	CELO	KuCoin	KCS	Theta Fuel	TFUEL
cETH	CETH	Kadena	KDA	Theta Network	THETA
Swissborg	CHSB	Klaytn	KLAY	Tokenize xchange	TKX
Chiliz	CHZ	Kyber network crystal	KNC	TRON	TRX
Compound	COMP	Kusama	KSM	TrueUSD	TUSD
Cronos	CRO	Bitkub Coin	KUB	Trust wallet	TWT
Curve dao	CRV	Lido dao	LDO	Uma	UMA
Casper Network	CSPR	Link	LN	USD Coin	USDC
Convex finance	CVX	Livepeer	LPT	Neutrino USD	USDN
Convex CRV	CVXCRV	Loopring	LRC	PAX Dollar	USDP
DAI	DAI	Lisk	LSK	Tether	USDT
DAO Maker	DAO	Litecoin	LTC	TerraClassicUSD	USTC
Decred	DCR	Liquidity USD	LUSD	Vvs finance	VVS
DeFiChain	DFI	Decentraland	MANA	WAX	WAXP
DIGIByte	DGB	Polygon	MATIC	Wrapped Bitcoin	WBTC
Dogecoin	DOGE	Merrit Circle	MC	Woo Network	WOO
DyDx	DYDX	Metis	METIS	Theter Gold	XAUT
Elrond	EGLD	Magic Internet Money	MIM	Chia	XCH
Escoin	ELG	Mina Protocol	MINA	Coinmetro	XCM
Enjin Coin	ENJ	Iota	MIOTA	eCash	XEC

Ethereum Name Service	ENS	maker	MKR	Nem	XEM
EOS	EOS	Marinade staked SOL	MSOL	Stellar	XLM
Ergo	ERG	mxc	MXC	Monero	XMR
Ethereum Classix	ETC	Nexo	NEXO	Radix	XRD
Ehtereum	ETH	Nucypher	NU	XRP	XRP
Euro Tether	EURT	Nexus Mutual	NXM	Tezos	XTZ
Energy Web	EWT	OKB	OKB	Yearn.Finance	YFI
Fei USD	FEI	OKC	OKT	zCash	ZEC
Filecoin	FIL	Ecomi	OMI	Zilliqa	ZIL
Flux	FLUX	Ontology	ONT	0x	ZRX
Frax	FRAX	Osmosis	OSMO		

Appendix B Lag Selection

Table 14 Optimal Lag Length Investor Attention and Return

The lowest values of MBIC determine the number of lags. The lowest values are indicated by *.

Lag	CD	J	J-Value	MBIC	MAIC	MQIC
1	0.516	221.945	0.000	67.695	189.945	149.433
2	0.519	151.829	0.000	36.142	127.829	97.446
3	0.522	101.865	0.000	24.740	85.865	65.609
4	0.427	37.016	0.000	-1.547*	29.016*	18.888*

Table 15 Optimal Lag Length Investor Attention and Volatility

The lowest values of MBIC determine the number of lags. The lowest values are indicated by *.

Lag	CD	J	J-Value	MBIC	MAIC	MQIC
1	0.655	192.655	0.000	38.406	160.655	120.143
2	0.658	150.326	0.000	34.639	126.326	95.942
3	0.661	113.380	0.000	36.255	97.380	77.124
4	0.585	51.222	0.000	12.659*	43.222*	33.094*

Table 16 Optimal Lag Length Investor Attention and Volume

The lowest values of MBIC determine the number of lags. The lowest values are indicated by *.

Lag	CD	J	J-Value	MBIC	MAIC	MQIC
1	0.782	158.993	0.000	5.668	126.993	86.674
2	0.784	107.957	0.000	-7.037*	83.957	53.718
3	0.786	93.927	0.000	17.265	77.927	57.768
4	0.748	57.321	0.000	18.990	49.321*	39.241*

Appendix C Descriptive statistics control variables

Table 17 Descriptive statistics control variables

Variable	Obs	Mean	Std.Dev.	Min	Max	Skewness	Kurtosis
SP500	23870	.003	.029	-.162	.114	-1.249801	12.03945
WTI_OIL	23870	.003	.106	-.821	.504	-2.615645	28.66202
Gold	23870	.003	.022	-.09	.08	-.3916014	5.752286
VIX	23870	.004	.129	-.372	.442	.5233089	4.005814
LogMarketCap	15701	8.709	.949	2.777	12.088	-.1668086	5.296061

Appendix D Lag Selection Control Variables

Table 17 Return and investor attention

Lag	CD	J	J-Value	MBIC	MAIC	MQIC
1	0.999	7305.061	0.000	5433.717	6913.061	6420.592
2	0.999	3958.691	0.000	2555.183	3664.691	3295.339
3	0.999	4234.98	0.000	3299.308	4038.98	3792.745
4	0.999	1153.242	2.4e-209	685.406*	1055.242 *	932.1247*

Table 18 Volume and investor attention

Lag	CD	J	J-Value	MBIC	MAIC	MQIC
1	0.998	7036.134	0.000	5164.791	6644.134	6151.666
2	0.999	4256.35	0.000	2852.842	3962.35	3592.999
3	0.999	4248.243	0.000	3312.572	4052.243	3806.009
4	0.999	1229.456	3.0e-225	761.6201*	1131.456*	1008.339*

Table 19 Volatility and investor attention

Lag	CD	J	J-Value	MBIC	MAIC	MQIC
1	0.999	6875.826	0	5013.587	6483.826	5993.269
2	0.999	4.08E+03	0	2681.551	3784.23	3416.313
3	0.999	4128.174	0	3197.054	3932.174	3686.895
4	0.999	1101.311	1.50E-198	635.7518*	1003.311*	880.6723*

Appendix E Stata do file

```
//panel data
xtset Coin Date, daily delta(7)

//test for stationary
xtunitroot fisher Return, dfuller lags(1)
xtunitroot fisher Volatility , dfuller lags(1)
xtunitroot fisher Volume, dfuller lags(1)
xtunitroot fisher ASVI, dfuller lags(1)


//optimal lag selection
asdoc pvarsoc ASVI Return, pvaro(instl(1/5))
asdoc pvarsoc ASVI Volatility, pvaro(instl(1/5))
asdoc pvarsoc ASVI Volume, pvaro(instl(1/5))

//Vector autoregression and granger causality
pvar Return ASVI, lags(4) vce(robust)
pvargranger
pvar Volatility ASVI, lags(4) vce(robust)
pvargranger
pvar Volume ASVI, lags(2) vce(robust)
pvargranger

//Vector autoregression including control variables
pvar ASVI Volume SP500 WTI_OIL Gold VIX LogMarketCap, lags(4)
pvar ASVI Return SP500 WTI_OIL Gold VIX LogMarketCap, lags(4)
pvar ASVI Volatility SP500 WTI_OIL Gold VIX LogMarketCap, lags(4)
```