Erasmus Universiteit Rotterdam

Erasmus School of Economics

Effect of Online Public Attention on The Returns of Cryptocurrencies: A Panel Analysis Using Google Trends

Author: Sjoek Oosting

Student number: 582299

Thesis supervisor: Ruben de Bliek

Second reader: Amy Li

Finish date: 01/07/2024

ABSTRACT

In this study the effect of online public attention on the performance of six cryptocurrencies is analyzed. To study this effect, Google Trends data is used to approximate public attention and the returns of the cryptocurrencies are considered as a measure of their performance. Furthermore, in past literature this effect has been extensively reviewed for other asset classes and established cryptocurrencies, such as Bitcoin. Therefore, we focus on less well-known cryptocurrencies, namely OKB, Bitcoin Cash, Cardano, Monero, Ripple and VeChain. Additionally, since the effects of public attention are not constricted to one specific day or point in time, we divided the analysis into two parts. The first part concerns the analysis of the direct effects of Google Trends on returns by applying data collected for the same day. In the second part, the presence of a potential delayed effect is examined by applying first-order lagged Google Trends data to estimate the returns of the following day. Due to the panel nature of our dataset, we considered using a fixed or random effects model. Therefore, a Hausman test was conducted to determine the most suitable model. The results of the test indicated that a fixed effects model should be applied in the analysis of both the delayed and current effect. The estimated coefficients of both models indicated a significantly positive effect of the variable related to Google Trends on the returns of the cryptocurrencies in the panel. Therefore, we conclude that there is a significantly positive effect of online public attention on the returns of the cryptocurrencies in our panel. Furthermore, these findings correspond to those of past research performed on other asset classes and cryptocurrencies. Thus, we consider Google Trends to be a valuable tool in the estimation of asset returns and recommend its application in future research to approximate online public attention.

Keywords: Cryptocurrency, Asset Returns, Panel Analysis, Fixed Effects Model, Google Trends

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

In the rapidly evolving landscape of financial assets, one cannot deny the strong role the internet and social media have on investors. The effect might not be apparent in average market returns but plays a central role in some of the most exceptional returns of this century. For example, the Robin Hood-like story of GameStop, where retail investors turned against institutions. These investors mainly communicated through social media platforms, such as the Reddit community WallStreetBets (NOS, 2021). Similarly, in the asset class of cryptocurrency, there is an upsurge of so-called meme coins, with Doge Coin as the most prominent example. The coin created large excessive returns after Elon Musk shared several tweets on it (Fox, 2024). We are strongly interested in this phenomenon and especially in the cryptocurrency market, due to high volatility and sensitivity to social media trends. Therefore, we will explore the relationship between returns on cryptocurrencies and online public attention in this study.

In past literature the relationship between online public attention, social media activity and returns on different asset classes has been extensively researched. These papers vary on different levels, such as the data used, their definition of online or social media activity and the asset class in question. Most define online or social media activity as social media sentiment, while others look at frequency. Ganesh and Iyer (2021) explore the relationship between the frequency of firm-initiated tweets on stock return and trading volume for thirty companies listed on the Dow Jones. They find a positive effect of firm- initiated tweets on both trading volume and stock returns. However, other papers deviate from the previous on their definition of social media activity, opting for a sentiment-based approach. For example, Bollen et al. (2016) investigate whether the collective mood, derived from Twitter feeds, is correlated with the Dow Jones Industrial Average (DJIA). They argue that including a measure for the public mood is beneficiary in determining DJIA predictions. On the other hand, Tan and Taş (2020) opt for a more specific definition of sentiment and analyse firm-specific Twitter sentiment. Again, they find a positive association between Twitter sentiment, returns and trading volume. Lastly, Oikonomopoulos et al. (2022) decided to perform a similar analysis on cryptocurrency. They examined the power of social media sentiment in predicting returns for seven of the biggest cryptocurrencies and reported an accurate forecast for two of those.

Most of the studies mentioned above, delve into social media sentiment and its effect on returns for several asset classes. In the context of cryptocurrencies, the works mainly concentrate on the most established cryptocurrencies, such as Bitcoin and Ethereum. However, we would like to shift the focus towards the less well-known alternative coins (altcoins) on the market. Since these coins seem to ‘blow up’ overnight, in some instances. Therefore, it would be interesting to estimate the effect online public attention has on these, and the cryptocurrencies selected for the analysis are: OKB, Bitcoin Cash, Cardano, Monero, Ripple and VeChain. These have a substantially smaller market capitalization than for example Bitcoin or Ethereum, with a $1.25 trillion and $367 billion market cap respectively.

Moreover, when defining social media activity, sentiment is not of interest to us, in line with the philosophy that ‘there is no such thing as bad publicity’. Investors are drawn to the cryptocurrency market for reasons that are not reflected in overall sentiment. For example, Özdemir (2022) discusses that potential investors are motivated by the prospect of a high return and ignore the risks involved. Furthermore, Haykır and Yağlı (2022) find that investors are susceptible to a fear of missing out regarding cryptocurrencies, resulting in herding behaviour and the formation of financial bubbles. For these reasons, the measure of online public attention will be shifted towards Google Trends to approximate public attention, rather than social media sentiment. Therefore, the central research question we intend to answer in this thesis, is:

“What is the effect of online public attention, approximated with Google Trends, on the performance of six different cryptocurrencies?”

With the intention to provide a substantial answer to our research question, we will perform two different analyses in this study. Firstly, we will examine the direct effect of Google Trends on returns by using data for the same day. Secondly, the delayed effect will be estimated by applying one day lagged Google Trends to estimate the returns of the following day. To perform these analyses, both Google Trends and historic price data must be sourced. Historic price data for the cryptocurrencies is collected from Tokeninsight.com, which allows users to freely download data on various cryptocurrency. Afterwards daily returns can be easily derived from the price data. The Google Trends data will be sourced directly from their website. Since Google is the most used search engine worldwide with an 86.99% market share according to Statcounter (2024), Google Trends serve as a good proxy for overall search engine requests. The tool requires a user to present a search query, period and geographical area, to base the Trends on. In the context of this analysis Google Trends are collected from 2019 throughout 2022, based on a worldwide sample and search queries corresponding to the names of the cryptocurrencies.

The data collection results in a panel dataset with observations across six cryptocurrencies over a period of four years. To analyse the panel, a fixed and random effects-model will first be estimated. Afterwards, the optimal model can be determined by applying a Hausman-test. The results of the Hausman test indicated that a fixed effects model fits our dataset best in both parts of the analysis. Furthermore, the estimations of the fixed effects models display a significantly positive current and delayed effect of Google Trends on the returns in our sample.

The rest of the paper is organized as follows. In Chapter 2 past literature on different aspect of our analysis will be reviewed, with an emphasis on Google Trends, cryptocurrencies and the relationship between them. Chapter 3 will discuss the data collection process and specify the variables included in the analysis. Subsequently, chapter 4 will delve into the methods applied to the analysis. The results derived from the analysis will be discussed in chapter 5 along with a comparison to the findings of past literature. Finally, chapter 6 and 7 contain discussions on the conclusions drawn from our results and limitations of our analysis, respectively.

CHAPTER 2 EMPIRICAL FRAMEWORK

In this chapter the findings of past literature on various aspect of our analysis will be extensively reviewed. Firstly, we will provide an overview of the online trading environment and corresponding investor behaviour. Secondly, a review will be provided on the ability of Google Trends data to approximate online public attention. Thirdly, we will examine the general characteristics of cryptocurrencies. Finally, past literature on the effect of online public attention, measured through Google Trends, on asset performance will be reviewed. Besides this, we will formulate several hypotheses based on the findings of past literature.

2.1 ONLINE INVESTOR TENDENCIES AND THE CRYPTOCURRENCY TRADE

To understand the effect online public attention, measured through Google Trends, has on crypto returns it is important to understand the online trading landscape and characteristics of crypto traders. Over the past decades the rise of the digital world has coincided with the rise of the digital economy. Consequently, both institutional and retail investors have changed their means of investing from traditional brokerage firms to online trading platforms. The online trading platform market was valued at $ 10.86 billion in 2023 and is expected to grow to $ 13.3 billion by 2026 (Statista, 2023). Among institutional and high net worth investors traditional brokerage firms still have a large market share, but especially retail investors of lesser means are shifting towards their online counterparties (Fortune Business Insights, 2024). Online trading has some clear advantages compared to traditional ways of investment. By cutting out the middleman, online trading platforms offer better information transparency, reduced trading costs, improved transaction speed and a broader range of services. However, according to Barber and Odean (2001) there are also potential risks. According to the authors improved accessibility to information may lead to “an exaggerated sense of control over the outcome”. Furthermore, traditional brokerages offer advice to less experienced investors and these services are included in the cost of trading. Online trading platforms do not offer the same personalized experience, leaving retail investors to trade on their own responsibility.

A more theoretical way to assess the motivation to trade online is by using a Technology Acceptance Model. A general TAM is based on two factors, namely Perceived Usefulness and Perceived Ease of Use (Davis, 1989). The former measures the perceived benefit a user might enjoy when using a specific technology. The latter defines to what extent a user would effortlessly adopt that technology. Roca et al. (2009) augmented this traditional model in their paper to include other factors applicable to online trading, such as privacy, perceived trust and perceived security. From the results it can be derived that perceived usefulness and trust are the strongest incentives to trade online, a result shared with other studies (Lee, 2009; Raut & Kumar, 2023). Furthermore, perceived risk also plays a large role, and investors perceive less risk in an online environment, illustrated by an increase in trading activity for both risk-seeking and risk-averse traders (Pan et al, 2023).

These findings make us wonder about the characteristics of cryptocurrency investors. Oksanen et al. (2022) find that crypto traders are more likely to be male and have an immigrant background. Moreover, they are also more likely to have taken instant loans and report higher excessive gambling, gaming and internet use than others. Corresponding to these characteristics, Talwar et al. (2021) performed an interesting investigation into different behavioural biases in the trading activity of 351 millennial men throughout the corona crises. The herding bias was the most influential, implying that investors tend to mimic each other's trades. Furthermore, hindsight bias and overconfidence also positively affected trading volume. On the contrary, there was no significant effect measured for loss aversion and mental accounting biases. Thus, the authors concluded that the individuals value the funds they allocate to investments and living expenses equally but are less affected by losing the former.

2.2 Google Trends as an Approximation of Online Public Attention

The following subsection will be dedicated to the characteristics of Google Trends and how it is derived, followed by a description of the data's application in past research. Since Google launched the feature in 2006, Google Trends has been widely used and acknowledged as an easily accessible data source for tracking online activity. Before we assess the use of Google Trends in past research, it is important to understanding how the data is derived. As mentioned before Google Trends allows the user to select an interval and geographical area base the values on. Afterwards, a sample is taken of all Google search queries that comply with the specifications and the number of times the requested search query appears in the sample is analysed (Google Trends, 2024). This number is used to scale the Google Trends to a value between 0 and 100 corresponding to the relative popularity compared to all other search queries. A sample of all Google search queries is sufficient here, since there are billions of queries for any requested period. An important note is that, unlike polling data, the scale does not reflect a reason for the queries. Therefore, it is not possible to identify whether a keyword was requested for positive or negative reasons.

The characteristics and processes mentioned above have popularised Google Trends among researchers in various field of research, due to the accessibility and broad reflection of public interest. The topics include healthcare, social sciences and economics and it has been used for purposes such as tracking viruses, measuring public interest and predicting consumer Trends respectively. In the field of healthcare research, Nuti et al. (2014) performed a systematic review of seventy studies utilizing Google Trends data. In their results, the authors highlight the versatility of the data, but also mention some potential biases. The data is especially beneficial in the field of healthcare since the real-time nature of the data enables the analysis of current events. Secondly, the authors praise the broad range of applications, such as the surveillance of outbreaks, behavioural analysis of public reactions, and testing the effectiveness of health campaigns. Furthermore, the results of studies using Google Trends data and studies using traditional data sources, such as hospital records, are often correlated. Therefore, Google Trends can be regarded as a valuable way of tracking current events. On the other hand, the review also informs us of possible biases, namely the publication and selection bias. The former is mentioned since 93% of the studies reviewed present positive results, feeding the belief that negative or inconclusive results are not published. Furthermore, the latter could be the result of poorly specified keywords in requesting the data from Google Trends.

In the field of economics Google Trends have been applied in a broad range of subjects in the field of both macro- and microeconomics. A good example of the former is a paper published by Wolozko (2020). In his paper Wolozko utilizes Google Trends to develop an economic activity tracker for the Organisation of Economic Co-operation and Development (OECD). To develop the tracker, he collected data on search topics and categories related to economic activity, namely "economic crises", "mortgage", "bankruptcies", "Luggage", "Mortgage". The tracker utilizes a non-linear neural network model to accurately predict the complex effect of search engine data on economic activity. In the end, the topics mentioned before were found to be strong predictors of different economic developments. Especially during the COVID-19 pandemic, the tracker provided valuable insights in economic movements. Therefore, the study highly valued the use of Google Trends data in economic research, although certain alterations may be necessary to prevent biases. In this study, Wolozko (2020) applied additional normalization to the data besides the normalization performed by Google to prevent long-term biases. He discusses that a bias could arise from non-economic factors influencing search results, such as shifts in Google user behaviour or changes in the Google algorithm.

Besides macroeconomic topics, there are also numerous studies that delve into subjects in the fields of microeconomics. Most of these papers regard revolve around the prediction of market trends or consumer behaviour for different products. For example, Silva et al. (2022) detected an explanatory power of Google web searches in the fashion industry. In the paper a case study is performed on the luxury brand Burberry and the Google Trends were collected related to the brand's name. The main research goal is to evaluate different forecasting models. In the results they report that a Denoised Neural Network Autoregression Model fits the data best. Moreover, they presented a positive correlation coefficient of 0.72 between the web search data and shopping trends. The authors dedicate the success of the search engine data to the broad scope of general consumer interest it provides.

A different paper centred around the prediction of product sales, is the paper "Predicting consumer behaviour with web search" by Goel et al. (2010). The authors use Google Trends data to predict the sales of different products before their release, focussing on events such as the box-office success of films, first-month sales of videogames and weekly ranking on the billboard top 100 charts. The nature of these events varies strongly from the research topic in this paper, since the independent variables are all upcoming events and not a continuous phenomenon such as the returns on cryptocurrencies. However, the paper does emphasize that models based on search counts outperform models based on other publicly available data. Additionally, some studies shift the time perspective of their research, focusing on the detection of current trends rather than the prediction of future trends. For example, Carriere-Swallow & Labbé (2011) performed a case study on consumer trends in the Chilean automobile market. The authors emphasize that due to a lag in many economic variables, it is difficult to make an accurate assessment of current conditions. In their studies, they aggregated the search request for all car brands to form a Google Trends Automotive Index. After the addition of the index, they found that the model fit improved for several nowcasting models.

All in all, we can conclude that Google Trends data provides valuable insights in various fields, such as healthcare, measuring overall economic activity, measuring consumer Trends and behaviour. Furthermore, it can be applied in both predictive and detective analyses.

2.3 Defining Cryptocurrency Characteristics

The following subsection is dedicated to a comprehensive review of the characteristics of cryptocurrencies. First, information will be provided on the underlying technologies that enable cryptocurrencies to be traded without the interference of financial institutions. Afterwards, we will compare the fundamental characteristics of cryptocurrencies to those of other asset classes. Finally, the ability of returns to measure asset performance will be assessed.

The main underlying technology that cryptocurrencies are built upon is blockchain technology. Blockchain technology stores transactions between parties in a public digital ledger (Joo et al., 2019). These transactions are known as blocks and contain general information on the transaction, such as the parties involved, the time of the transaction and previous transactions. All users can access this ledger, ensuring that transactions are transparent and secure. Since all transactions are built upon the record of other transactions, it is nearly impossible to alter a transaction. These measures of security allow cryptocurrencies to be traded without a central authority. When a transaction is requested, there are two main methods of validation, namely the Proof of Work (PoW) and the Delegated Proof of Stake (DPoS) algorithms (Bach et al., 2018). The former utilizes so-called miners, who are required to solve a complex mathematical problem to ensure that the requesting party has sufficient funds. As a reward the miners obtain new coins. This algorithm is used by cryptocurrencies such as Bitcoin, Litecoin and Bitcoin Cash. On the other hand, the DPoS algorithm works slightly different. In the DPoS blockchain, validators are assigned based on the number of coins they stake. Staking entails that the validator temporarily deposits some of his coins as collateral. The validator then runs the software that generates new blocks. In return the validator receives a transaction fee. Cryptocurrencies that work on a DPoS algorithm include Ethereum and Binance Coin.

New cryptocurrencies are launched through the process of an Initial Coin Offering (ICO), a process similar to a crowdfunding campaign. In the ICO startups raise money by selling and distributing their new token to a large pool of investors. In a study performed by Benedetti and Kostovetsky (2021) this process was evaluated for 2390 ICOs raising $12 billion in total. In their findings they report an average return of 179% between the preset ICO price and opening market price. Furthermore, tokens average a 48% abnormal return over the first 30 days after the ICO benchmarked against the prices of Bitcoin. Furthermore, there are a few important differences between an ICO and an initial public offering (IPO) of stocks. Firstly, the regulation on ICOs is considerably less than on IPOs. The latter are regulated by financial authorities and require admission of financial information, while the former are subject to varying standards of regulation, but less vigorous than on IPOs (Ofir & Sadeh, 2019). This lack of regulation often causes a high level of information asymmetry, preventing investors from making well informed decisions. Furthermore, the two offerings attract a different type of investor. ICOs are mostly dominated by retail investors. Their involvement can lead to higher volatility, due to more speculative trading behaviour, especially in the presence of information asymmetry and a lack of trader experience. However, when it comes to IPOs, institutional investors tend to stabilize the market and reduce volatility by taking part in the initial pricing of the stock (Ofir & Sadeh, 2019). Thus, it can be concluded that there is a large difference between the initial offerings of common stocks and cryptocurrencies. However, extensive research indicates that the overall cryptocurrency market has evolved from a peripheral market to one resembling an intermediate sized stock exchange (Wątorek et al., 2021). This resemblance is based upon several market characteristics. For example, return distributions in the cryptocurrency market show ‘fat tails’, a trait often seen in mature financial markets. Furthermore, the market shows signs of volatility clustering, a phenomenon where high volatility periods tend to be followed by low volatility periods.

This transition towards a more mature market can be partially attributed to external events, particularly the COVID-19 pandemic. Caferra and Vidal-Tomás (2021) evaluate co-movement between the cryptocurrency and stock market during the pandemic. They found a long-term correlation between the two markets, although they seem to move independently in the short-run. These findings can be attributed to the quick recovery the cryptocurrency market made compared to the stock market after both markets crashed in March 2020. Therefore, the authors highlight potential hedging abilities of cryptocurrencies in times of crises. The market’s reaction to the pandemic prompts us to investigate the influence the pandemic had on the returns of our cryptocurrencies of interest. Therefore, we will include dummy variables related to the stage of the pandemic to approximate the effect this period had. Hence, we formulate the following hypothesis on this effect:

Hypothesis 1: The returns on the six cryptocurrencies of interest in our analysis were positively affected by the Covid-19 pandemic.

These findings also shed light on a key characteristic of cryptocurrencies, namely that cryptocurrency portray higher levels of volatility than other asset classes (Brini & Lenz, 2024). Contrary to traditional markets, a reversed leverage effect is present in the cryptocurrency market, implying that market volatility increases more due to positive returns than negative returns. This phenomenon could be explained by behavioural factors, for example because investors experience a Fear of Missing Out (FOMO) or display speculative trading behaviour.

For the final part of this subsection, we will examine the use of returns as a measure of asset performance, rather than price. In their study Mallikarjuna and Rao (2019) mention several reasons to utilize returns rather than price. Firstly, returns are stationary, which implies that the mean, covariance, and other statistical properties are more consistent. This property is crucial for the application of variables in forecasting models and time series analysis. Secondly, returns enable comparison across different assets and asset classes, because returns normalize price changes. Furthermore, returns are an essential element in analysing volatility as current positive or negative returns can imply future volatility. However, the last property is not of interest in the analysis performed in this study.

2.4 Google Trends and Cryptocurrency Returns

In this subsection we will combine the topics that have been previously discussed and review past literature on the relationship between Google Trends and asset returns. First, an overview will be provided for this relationship in the field of stocks. Consequently, the topic of interest will be shifted towards the cryptocurrency market. Furthermore, throughout this section we will formulate two hypotheses about different aspects of our analysis, based on the findings presented in past studies.

The first study we will consider is a paper by Hu et al. (2018). In their paper the authors propose a model to predict the direction of returns in the Dow Jones Industrial Average (DJIA) and S&P 500 and study the effects of adding a Google Trends variable. They do so by comparing the prediction abilities of two different sets of variables. The first includes variables reflecting traditional stock data, such as price and trading volume. The second also includes Google Trends data. The data collected from Google Trends regards search queries for a wide range of financial terms, such as ‘cash’, ‘returns’ and ‘dividend’. Afterwards, the authors evaluated the two models based on the hit-ratio, a percentage resembling the number of correct predictions. Their findings indicate that the hit-ratio, for the prediction of both the DJIA and the S&P 500, increased with the inclusion of the Google Trends variable. A different study performed by Huang et al. (2019) also assesses the strength of utilizing Google Trends in a forecasting model. In their study the authors apply a Granger Causality model on the returns of the S&P 500 index. The results indicate that ten search terms are Granger causal with the S&P 500 index. These terms are not only related to finance, such as ‘cash’, ‘return’ and ‘dividend’, but also to more general topics, such as ‘war’ and 'society’. Furthermore, Huang et al. (2019) expand their field of research compared to Hu et al. (2018), by applying their findings in an investment strategy. The strategy was based on a different model that included Google Trends data to predicts the movement of index and release a buy or sell signal. The model resulted in an 83% return over a period of two years, whereas a buy-and-hold strategy would have achieved a 40,87% return.

Besides the stock market, extensive research has also been conducted in the relationship between Google Trends and different characteristics of cryptocurrencies. For example, a study performed by Nasir et al. (2019), who analyse the effect weekly Google Trends have on weekly returns on Bitcoin using a Vector Autoregressive model (VAR). In their model the authors utilized three week lagged Google Trends as the independent variables. This analysis resulted in the conclusion that Google Trends data has a significant positive effect on the returns of Bitcoin. Anastasiadis and Papadamou (2022) altered and elaborated a similar research design, opting for a one week lagged VAR-model. Furthermore, the research is performed on the returns of four cryptocurrencies, namely Bitcoin, Ethereum, Stellar and Monero. The results of the regressions indicate that Bitcoin stands out as a market leader, since Bitcoin’s lagged returns and Google Trends influence both the Google Trends and returns of the other three cryptocurrencies. These findings motivate us to include the Google Trends of Bitcoin as a control variable in our research. Although the effect is stronger for the returns on Bitcoin, we believe that including Bitcoin’s Google Trends fits the research design better, since it reflects a level of public attention towards cryptocurrencies. These insights allow is to formulate our second hypothesis:

Hypothesis 2: Bitcoin’s Google Trends will have a positive effect on the returns of the six cryptocurrencies.

Focusing on the dynamics of public attention, Aslanidis et al. (2022) studied the information flow between various economic factors, the cryptocurrencies themselves and some of their market statistics using information entropy. In their research they utilized the Google Trends for two sets of search queries. Firstly, the authors constructed a General Uncertainty Index (GUI) based on search queries related to economic uncertainty, such as ‘recession’ and ‘inflation’. Secondly, a Google Trends Cryptocurrency Attention Index (GTC) was constructed by summing the Google Trends of five cryptocurrencies, namely Bitcoin, Ethereum, Ripple, Litecoin and Bitcoin Cash. From the results it can be concluded that cryptocurrency market behaviours are not related to the GUI. However, significant information flows were detected in both directions between the GTC and market returns. This indicates that increased returns and online public attention have a positive relationship in both directions. Furthermore, bidirectional information flows were also detected between market volatility and the GTC, where the flow from volatility to Google Trends attention was stronger.

In another study the causality between Google Trends data is examined closer, by employing a nonparametric causality in quantiles analysis. This method examines the causality between two variables within predetermined quantiles of their distribution. In the paper written by Raza et al. (2022) the causality between Google Trends data and the return on six cryptocurrencies -Bitcoin, Ethereum, Litecoin, Dash, Nem, and Litecoin- was assessed in quantiles of size 0.1. The results were varying for different cryptocurrencies. Bitcoin for example showed causality with Google Trends data up to the 0.7 quantile, indicating that the 60% lowest values of Google Trends can predict the 60% lowest returns. On the other hand, Ethereum showed significant causality in specific middle quantiles. Overall, a significant relationship was found in some quantiles for all cryptocurrencies except for Dash. The study emphasizes the application possibilities in trading strategies, because the results provide a detailed insight for which values in the distribution the relationship occurs.

Through the extensive review of past literature, this subsection provides us with a clear understanding of the relationship between Google Trends and cryptocurrencies. Although most of the past research vary in research design from the analysis performed in this paper, we believe the information reviewed enables us to formulate a general hypothesis for our research question:

Hypothesis 3: online public attention, approximated by Google Trends, has a positive effect on the returns on the six cryptocurrencies of interest.

Furthermore, many of the previously discussed papers have included lagged values of the Google Trends variables in their analysis to observe potential delayed effects. This motivates us to perform a second analysis utilizing the lagged values of the Google Trends data, which will be further discussed in the methodology section. In line with past research, we formulate the last hypothesis:

Hypothesis 4: besides a current effect, public attention, approximated by Google Trends, also displays a delayed positive effect on the returns on the six cryptocurrencies of interest.

CHAPTER 3 DATA

In this chapter we will provide an overview of the data and methodology used in the analysis of the relationship between Google Trends and cryptocurrency returns. First, we will explain our motivation behind the selection of the cryptocurrencies of interest and sample period. Afterwards the data collection process and alterations made to the data will be discussed. Thirdly we will define the variables included in the analysis and provide summary statistics and Pearson's correlation matrices to gain an overview of the data. Lastly, a comprehensive review of the statistical models that were applied to perform the analysis.

3.1 SAMPLE DESCRIPTION

In this subchapter a detailed description will be provided on the data utilized in our analysis. To begin we will define our sample period and cryptocurrencies of interest. All data was sampled for the period of December 1, 2019, through January 31, 2022. This sample period was selected to incorporate data throughout the Covid-19 pandemic. As mentioned before in the empirical framework the pandemic had strong implications for the cryptocurrency market. To control for these implications, dummies will be incorporated to distinguish between the periods before, during and after the pandemic.

Before collecting the data, suitable cryptocurrencies had to be selected to perform the analysis on. The cryptocurrencies were selected in line with two conditions. Firstly, the coins should have a maximum market capitalization of $20 billion on December 31, 2022. Secondly, the coins should have a distinguishable name, to ensure that the search request data collected from Google Trends can only regard the coin in question. In the end the six coins that were selected are: OKB (OKB), Cardano (ADA), Bitcoin Cash (BCH), Monero (XMR), Ripple (XRP) and VeChain (VET), with their ticker presented in parentheses. In the following sections, an explanation will be provided on the data collection of all variables along with a specification of the variables.

The result is a panel data set with 1461 observations for each of the six cryptocurrencies, resulting in 8766 observations in total for the independent and dependent variables. A panel dataset is defined by a combination of cross sectional and timeseries characteristics. This type of data not only enables us to track the development of the cryptocurrencies over time, but also creates a better understanding of the individual characteristics of the cryptocurrencies.

3.2 Data Collection and Variable Specification

In this section we will discuss the data collection and characteristics of the different variables of interest. To begin with the dependent variable and proceeding to our independent and control variables.

3.2.1 Cryptocurrency Returns

As mentioned before, the dependent variables in the analysis are the daily returns of the cryptocurrencies mentioned above, used as a measure of performance. To derive these returns end-of-day prices for the six coins were collected from Tokeninsight.com in US dollars. After the data was collected the daily returns were calculated in excel using the following formula:

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1$$

In this formula $R_{i,t}$ represents the return and $P_{i,t}$ the end of day price of cryptocurrency *i* at time *t.* To describe the development of the coins over the sample period, table 3.1 below provides some general information on the end-of-day price and the market capitalization (MCAP) at the beginning and ending of our sample period.

Table 3.1

Price and MCAP development throughout the sample period and corresponding cumulative returns.

	Price (USD)		MCAP (Billion USD)		Cumulative
	01/01/2019	31/12/2022	01/01/2019	31/12/2022	Return
OKB	0.6586	26.0480	0.1976	6.3732	36.5459
BCH	148.8404	97.0354	2.6121	1.8595	-0.3962
ADA	0.0405	0.2456	1.2612	8.6080	4.7463
XMR	46.5457	147.2207	0.7766	2.6534	2.0588
XRP	0.3492	0.3389	14.2497	17.3834	-0.0694
VET	0.0039	0.0158	0.2176	1.1440	2.7927

Note: General information on the six cryptocurrencies presented by their ticker. Prices are presented in US dollars and MCAP in terms of billion US dollars. Returns are displayed in decimal forms, where an increase of 0.01 corresponds to an increase of 1%.

As shown in table 3.1 the cryptocurrencies have a wide range in price and MCAP. Besides this, not all prices have increased over the period, resulting in negative cumulative returns for Bitcoin Cash and Ripple. In contrast to the prices, it can be observed that all market capitalizations increase over the sample period, which might seem contradictory since some prices decrease. However, this is due to the constant release of new tokens through mining. To visualize the price changes over time, the cumulative returns of OKB are plotted in figure 3.1 below:

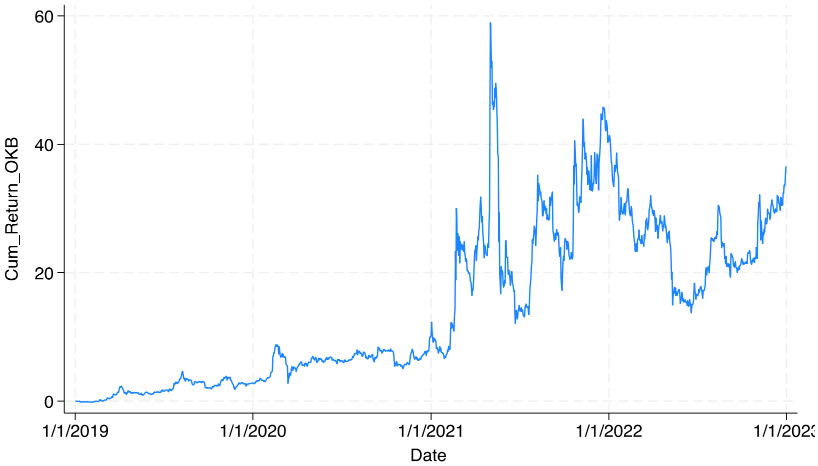

Figure 1: Cumulative returns on OKB from January 1, 2019, through December 31, 2022.

The figure illustrates that the cumulative returns vary strongly over time. Overall, a positive cumulative return can be observed corresponding the value in table 3.1. Similar graphs for the other cryptocurrency are displayed in figures A1-A5 of appendix A. In the following we will further discuss the data collected from Google Trends and provide summary statistics on all data discussed in this section

3.2.2 Google Trends Data

In this subchapter, the collection and alteration of the Google Trends data will be discussed. For each of the six cryptocurrencies Google Trends data was collected on the search queries directly related to their names, namely ‘OKB’, ‘Cardano’, ‘Bitcoin Cash’, ‘Monero’, ‘Ripple’ and ‘VeChain’. Above a 9-month threshold Google Trends provides weekly data, while we are interested in daily data. Therefore, daily data was collected for six-month periods, adding up to the full length of the sample period. However, a problem occurs when applying this method. As we mentioned before, Google Trends takes a sample of all search queries for the requested period and scales the requested search query based on its relative popularity within the sample. As a result, the daily data collected is based on different scales and cannot be compared. To counteract this, weekly data was collected for the entire four-year period and each daily data point was multiplied with the value of the corresponding week. Thus, the data is normalized against the overall scale throughout the four-year sample period. This method was previously discussed in posts on Towardsdatascience.com (Tseng, 2021), and is provided in an Application Programming Interface (API) in the statistical software programs R and Python. Unfortunately, a similar API is not provided by Stata and the modifications were made in excel by hand. To visualize the change, the weekly, daily and normalized data for OKB are plotted in figures A7-A8 of appendix A.

3.2.3 Control Variables

Several control variables are added in our analysis, not only to isolate the effect of public attention, but also to observe the effects of these control variables. Firstly, the Google Trends of Bitcoin are added as a control variable. In the empirical framework we provide an explanation on the decision to include these. The data on Bitcoin was collected and normalized in the same way as the other cryptocurrencies.

Secondly, dummy variables will be included to reference the periods before and during the Covid-19 pandemic. As discussed before this period spiked public interest in the cryptocurrency market and unobservable effects related to this period may otherwise interfere with the results of our analysis. The classification of the periods before, during and after the pandemic, was based on both political implementations and the performance of financial markets. Lock-down measures started to be implicated worldwide in late February and Early March of 2020. Around the same time the S&P 500 experienced a historic crash (Mazur et al., 2021). Therefore, we consider the pandemic to have started on March first, 2020. Although, the stock market recovered by July 2020 (Sunder, 2020), many of the measures taken to counteract the pandemic still held (U.S. Centre for Disease Control and Prevention, 2023). Since these implications also affected the market, we consider the pandemic to have ended later than the market's recovery, namely December 31, 2021. After establishing these dates, the dummy variables were constructed in Stata with an if-function.

3.3 SUMMARY STATISTICS

In table 3.2 below summary statistics are provided for the variables included in our analysis. Besides these, the summary statistics of the grouped returns and adjusted Google Trends data are also included to give an overview of the entire sample.

Before we asses these statistics, it is worth to mention that we tested the returns and normalized Google Trends data for skewness and kurtosis. Upon testing the returns showed a high level of kurtosis, at 15.679. Kurtosis is a statistical phenomenon that implies a low probability of outliers in the distribution of a variable. Therefore, the variation of a variable is most likely caused by a few extreme outliers (Ruppert, 1987). Consequently, the variance does not reflect the distribution of the variable, resulting in an increased possibility of bias (Shete et al., 2004). To prevent this, the data was winsorized in Stata. This process adjusts the 1% highest and lowest values to the border values of the 99% interval. As a results, a minimum return of -0.14 and a maximum of 0.174 can be observed for all return variables in table 3.2. Furthermore, due to the winsorization of the data, the variances for the returns are relatively low, which can be derived from the standard deviations. Furthermore, all cryptocurrencies show a positive average return ranging from 0.1% to 0.4%.

With regards to the Google Trends data, it showed moderate skewness. Therefore, we contemplated to apply a logarithmic scale. However, the model used in this analysis does not assume normality (Gardiner et al., 2008) and we decided not to apply the scale to observe the direct effects. When looking at table 3.2, OKB shows the highest mean rating at 30.276. The average ratings of the other currencies range from 3.99, for VeChain, to 21.716, for Monero. Furthermore, it is interesting that Bitcoin does not display the highest mean rating, even though it is widely considered as the most established cryptocurrency. As mentioned before, a part of our analysis will be dedicated to the one-day lagged Google Trends values. The summary statistics for the one-day lagged values correspond to those in the table below, since only the date of the value that is applied changes, not the value itself. Furthermore, this also applies to the correlations discussed in the following subsection.

Table 3.2

Summary statistics on the variables of interest

	N	Mean	Std. Dev	Min	Median	Max
$\text{Return}_{\text{Sample}}$	8766	.002	0.050	-.14	.001	.174
$\text{Return}_{\text{OKB}}$	1461	.003	0.050	-.14	.001	.174
$\text{Return}_{\text{BCH}}$	1461	.001	0.049	-.14	0	.174
$\text{Return}_{\text{ADA}}$	1461	.003	0.052	-.14	.001	.174
$\text{Return}_{\text{XMR}}$	1461	.002	0.045	-.14	.003	.174
$\text{Return}_{\text{XRP}}$	1461	.001	0.047	-.14	0	.174
$\text{Return}_{\text{VET}}$	1461	.003	0.056	-.14	.001	.174
Adj. $\text{Trends}_{\text{Sample}}$	8766	13.855	13.296	0	9.75	100
Adj. $\text{Trends}_{\text{OKB}}$	1461	30.276	10.166	0	29.15	84.15
Adj. $\text{Trends}_{\text{BCH}}$	1461	10.051	8.108	2.47	8.16	95
Adj. $\text{Trends}_{\text{ADA}}$	1461	5.557	7.966	.32	2.76	100
Adj. $\text{Trends}_{\text{XMR}}$	1461	21.716	14.252	2.55	18.9	94
Adj. $\text{Trends}_{\text{XRP}}$	1461	11.541	7.104	1.95	10.54	90
Adj. $\text{Trends}_{\text{VET}}$	1461	3.99	7.572	0	1.56	100
Adj. $\text{Trends}_{\text{BTC}}$	1461	11.984	10.464	1.26	9.3	100
Pre-Pandemic	1461	.291	0.454	0	0	1
Pandemic	1461	.459	0.499	0	0	1

Note: summary statistics on all variables. Returns are presented in decimal points where a 0.01 average return corresponds to 1%. Furthermore, the returns and adjusted Google Trends of the cryptocurrencies are distinguishable by the ticker in subscript. The grouped variable represents the entire sample.

3.4 Correlation Matrices

To create a better understanding of the relationships between the variables, Pearson's correlations were determined between all variables in our analysis. The Pearson's correlation takes the form of a value between -1 and 1 that represents the strength and direction of a linear relationship (Sedgwick, 2012). Table 3.3 below displays the correlations between the winsorized returns of each cryptocurrency. Furthermore, table 3.4 displays those between the normalized Google Trends variables of each cryptocurrency.

Table 3.3

Pearson's correlations between the returns of each individual cryptocurrency

	$Return_{OKB}$	$Return_{BCH}$	$Return_{ADA}$	$Return_{XMR}$	$Return_{XRP}$	$Return_{VET}$
$Return_{OKB}$	1.000					
$Return_{BCH}$	0.505***	1.000				
$Return_{ADA}$	0.457***	0.695***	1.000			
$Return_{XMR}$	0.486***	0.685***	0.630***	1.000		
$Return_{XRP}$	0.465***	0.683***	0.643***	0.586***	1.000	
$Return_{VET}$	0.472***	0.664***	0.687***	0.600***	0.601***	1.000

Table 3.3: Correlation matrix for the returns of all cryptocurrencies, distinguishable by their ticker in subscript. With significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 3.4

Pearson's correlations between the adjusted Google Trends of each individual cryptocurrency

	Adj. $Trends_{OKB}$	Adj. $Trends_{BCH}$	Adj. $Trends_{ADA}$	Adj. $Trends_{XMR}$	Adj. $Trends_{XRP}$	Adj. $Trends_{VET}$	Adj. $Trends_{BTC}$
Adj. $Trends_{OKB}$	1.000						
Adj. $Trends_{BCH}$	-0.031	1.000					
Adj. $Trends_{ADA}$	0.005	0.349***	1.000				
Adj. $Trends_{XMR}$	-0.034	0.588***	0.537***	1.000			
Adj. $Trends_{XRP}$	-0.032	0.521***	0.206***	0.339***	1.000		
Adj. $Trends_{VET}$	0.082***	0.596***	0.499***	0.620***	0.452***	1.000	
Adj. $Trends_{BTC}$	-0.092***	0.485***	0.434***	0.643***	0.266***	0.330***	1.000

Note: Correlation matrix for the adjusted Google Trends variables of all cryptocurrencies, distinguishable by their ticker in subscript. With significance levels: *** p<0.01, ** p<0.05, * p<0.1

The values presented in the table illustrate that there is a strong positive correlation between the returns of all cryptocurrencies, significant at the 1% level. This is in line with past research and could be accredited to general market movement. Besides this, the table displays a significant positive correlation between most adjusted Google Trends variables. This can be interpreted as a result of general public attention towards the cryptocurrency market. However, only OKB exhibits divergent correlations. Not only are the correlations between OKB and the other variables smaller, but most are insignificant. Furthermore, some of the coefficients have negative values, including the correlation between OKB and Bitcoin that is significant at the 1% level.

Having discussed the correlation within the dependent and independent variable, we will now devote our attention to the correlations between the dependent and independent variables. Table 3.5 displays these correlations for the variables reflecting the entire sample. The matrices related to each individual cryptocurrency can be found in table A1-A6 of appendix A.

Table 3.5

Correlation matrix for the sample returns and independent variables

	$Return_{Sample}$	Adj. $Trends_{Sample}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{Sample}$	1.000				
Adj. $Trends_{Sample}$	0.061***	1.000			
Adj. $Trends_{BTC}$	-0.049***	0.251***	1.000		
Pre-Pandemic	-0.003	-0.133***	-0.362***	1.000	
Pandemic	0.042***	0.128***	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between sample returns and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

From the table it can be concluded that there is a significant positive correlation between the dependent return variable and the independent Google Trends variable at the 1% significance level. This result is not shared by all individual cryptocurrencies, since OKB and Monero display insignificant positive correlations. Contrary to the findings of past research, the correlation between the Google Trends ratings of Bitcoin and the returns of the sample is negative. Then again, the correlation between Bitcoin's Google trends and the sample Google Trends is positive, which corresponds to the findings of past literature. The individual cryptocurrencies again exhibit similar relationships. Proceeding to the dummy variables concerning the status of pandemic. The dummy related to the period before the pandemic exhibits negative correlations with both the sample and Bitcoin's adjusted Google Trends. However, the variable expressing the period during the pandemic has a positive correlation with both the sample returns and Google Trends variable, significant at the 1% level.

Many of the correlations displayed in table 3.5 are relatively high. As a result, there is an increased risk of multicollinearity. Multicollinearity entails that there is a strong correlation between the independent variables in an analysis. As a result, standard errors increase, and variables could be wrongly specified as significant (Daoud, 2017). A common way to detect multicollinearity is by calculating the Variance Inflation Factor (VIF) for the independent variables. A VIF between 1 and 5 can be interpreted as a moderate correlation between independent variables. As can be seen in table A7 of appendix A, there is an acceptable amount of multicollinearity between our independent variables. This removes the need for any corrections.

CHAPTER 4 METHODOLOGY

In the following chapter, the analytical methods that are employed in this paper will be extensively reviewed. As mentioned before in the data section, the dataset utilized in our analysis is a panel dataset. When it comes to panel data analysis, there are several potential statical models, of which the fixed and random effects models are among the most used (Yaffee, 2003). Therefore, one of these models will be applied for the analysis conducted in this study. However, the suitability of each model strongly depends on the characteristics of the data. The following subsections will describe the characteristics of both models, followed by an explanation of the Hausman statistical test used to determine the most suitable model.

4.1 FIXED EFFECTS MODEL

The fixed effects model is built upon the believe that in a panel dataset there are time-invariant effects that are specific to each group, or cryptocurrencies in the context of our analysis . By controlling for those variations, the fixed effects model isolates the effect of the independent variable on the dependent variable. The model could be interpreted as a dummy variable for each separate group. Then the coefficient of the dummy would reflect the fixed effects for each group (De Haan, 2020). Furthermore, an important assumption of the model is that it allows for correlation between these fixed effects and the independent variables.

A favourable property of the fixed effects model is that it eliminates the possibility of omitted variable bias. The omitted variable bias occurs when a variable is excluded that influences both the independent and dependent variable. For example, in the context of studying the effect of online public attention on cryptocurrency returns, such a variable could be the launch date. The model eliminates this bias by de-meaning the variables. In this process, the mean of each variable is determined per group and then subtracted from the observed variables during estimation of the model (deHaan, 2020). By de-meaning the data the time-invariant characteristics within groups are eliminated. This allows the model to estimate the relationship between the variables from the deviations from the mean. However, this process has a potential downside, as a considerable amount of the within group variation is lost, thereby decreasing the variance to base estimations on. As a result, the standard errors of the estimated coefficients can increase, complicating the identification of the true effects of the independent variables (Clark & Linzer, 2015). To create a better understanding of the model, the general formula is presented below, modified to include the variables in our analysis:

$$R_{i,t} = \alpha + \beta_1 GT_{i,t} + \beta_2 GT_{BTC,t} + \beta_3 PP + \beta_4 P + u_{i,t} \text{ where: } u_{i,t} = \mu_i + v_{i,t}$$

With:	$R_{i,t}$	=	The returns on Cryptocurrency *i* at time *t*
	α	=	Intercept
	β	=	Coefficients corresponding to the variables
	$GT_{i,t}$	=	The adjusted Google Trends of cryptocurrency *i* at time *t*
	$GT_{BTC,t}$	=	The adjusted Google Trends of Bitcoin at time *t*
	PP	=	Dummy for the pre-pandemic period
	P	=	Dummy for the period during the pandemic
	μ_i	=	Fixed effect for cryptocurrency *i*
	$v_{i,t}$	=	Error term for cryptocurrency *i* at time *t*

In the model above the Google Trends values at time *t-1* are applied to analyse the relationship between the one-day lagged Google Trends and cryptocurrency returns. Furthermore, the inclusion of time-invariant control variables is made redundant, since these effects are captured in the fixed effects. However, the control variables included in the model above vary over time and are not subject to this redundancy.

4.2 Random Effects Model

The random effects model offers an alternative approach to panel data analysis. Although the models exhibit large similarities, the random effects model varies from the fixed effects model on some key characteristics and assumptions. Contrary to the fixed effects model, the random effects model assumes that time-invariant group-specific effects are random and normally distributed (Clark & Linzer, 2015). The random effects allow the model to make prediction based on both the within and between group variations. As mentioned before, in the fixed effects model the between group variation is lost by the demeaning process, which eliminates the time-invariant effects captured in the mean. As a result, the random effects model can estimate the coefficients of the independent variables more efficiently than the fixed effects model. However, the model makes a critical assumption, namely that the random effects are uncorrelated with the independent variables (Brooks, 2019). To provide a visualization of the model, the general formula applied to our research is presented below:

$$R_{i,t} = \alpha + \beta_1 GT_{i,t} + \beta_2 GT_{BTC,t} + \beta_3 PP + \beta_4 P + \omega_{i,t} \text{ where: } \omega_{i,t} = \varepsilon_i + v_{i,t}$$

With:	$R_{i,t}$	=	The returns on Cryptocurrency *i* at time *t*
	α	=	Intercept
	β	=	Coefficients corresponding to the variables
	$GT_{i,t}$	=	The adjusted Google Trends of cryptocurrency *i* at time *t*
	$GT_{BTC,t}$	=	The adjusted Google Trends of Bitcoin at time *t*
	PP	=	Dummy for the pre-pandemic period
	P	=	Dummy for the period during the pandemic
	ε_i	=	Random effect for cryptocurrency *i*
	$v_{i,t}$	=	Error term for cryptocurrency *i* at time *t*

When comparing the general formulas of the random and fixed effects model, a strong similarity stands out. In the formula the sole difference seems to be the inclusion of random effects (ε_i) rather than fixed effects (μ_i). However, due to the definition of these effects and the assumptions made, the models differ. In the following subsection, the implications of these assumptions and the specification of a suitable model will be further discussed. Besides this, the general formula for the analysis of the one-day lagged effects is derived in a similar way as for the fixed effects model.

4.3 MODEL SPECIFICATION & HAUSMAN TEST

The decision whether a fixed or random effects model should be applied, could be considered as a trade-off between biasedness and efficiency. The central factor in this trade-off is the assumption made about the correlation between the group specific effects and the independent variables. The fixed effects model allows for correlation between the group specific fixed effects and the independent variables. Besides this, the estimated coefficients are generally unbiased, because the model controls for all omitted variables. However, since the model only uses within-group variation to estimate the coefficients, these can be less precise (Clark & Linzer, 2015). On the other hand, the random effects model assumes there is no correlation between the random effects and independent variables. If this assumption holds, the random effects model produces more efficient estimates, since it incorporates both within- and between-group variation in the estimation. Then again, if the assumption is violated, the estimated coefficients will be biased (Clark & Linzer, 2015).

In conclusion, the model specification depends on whether the assumption of no correlation is met. This can be tested by using the Durban-Wu-Hausman test, or simply Hausman test. This test compares the consistency of two estimated coefficients against each other (Hausman, 1978). A coefficient is considered consistent if it converges to the true population coefficient when the sample size approaches infinity. However, a coefficient cannot be consistent if it is biased. When the test is applied, under the null hypothesis both coefficients are consistent. Thus, in the context of a fixed and random effects model, the test compares the estimated coefficients of both models based on their consistency, if the null can be rejected, we can conclude that the coefficient of the random effects model is inconsistent and therefore subject to bias due to correlation with the random effects. The hypotheses of the test are as follows:

H_0: *Both coefficients are consistent,*

H_a: *only the fixed effects coefficient is consistent*

The test statistic of the Hausman test follows a chi-squared distribution. If the null is rejected, the fixed effects model will be applied, since the Random Effects model has violated the assumption of no correlation. However, if the null cannot be rejected, the random effects model will be applied, since it is more efficient.

CHAPTER 5 RESULTS

In this chapter we will discuss the findings of the analyses discussed above. We will start by discussing the results of the Hausman-test for both models, where one uses the same-day Google Trends values and the other the one-day lagged values. Subsequently, the tests conducted for heterogeneity and autocorrelation will be evaluated. After the tests are conducted and appropriate adjustments are made, we will provide an extensive assessment of both models and their estimations. Lastly, we will compare the findings of analysis to those of past literature.

5.1 HAUSMAN TEST

As mentioned before in the methodology section, the Hausman test is conducted to determine whether a fixed or random effects model suits the dataset best. Under the null hypothesis, the estimated coefficients of both the random and fixed effects model are consistent, implying that the random effects model is more suitable, due to a higher level of efficiency. The test statistic follows a chi-squared distribution, where the degrees of freedom are determined by the number of parameters. The test is conducted in Stata by running both a fixed and random effects model, saving the estimations and applying the test on these estimates. The results of the Hausman test conducted on the estimates of both models are provided in table 5.1 below:

Table 5.1

Results of the Hausman test for same-day and lagged Google Trends data

	same-day	*Lagged*
X^2	41.210	19.790
P-value	0.000	0.001

Note: results of the Hausman-tests conducted to determine the most suitable models to analyse both the same-day effect of Google Trends on returns and the delayed effect. The values included in the table are the chi-squared test statistic and the corresponding p-value for the two separate tests.

Considering the p-values reported in the table, we can conclude that there is sufficient evidence to reject the null hypothesis for both tests. Therefore, the fixed effects model will be applied in both parts of the analysis, since the coefficients of the random effects model are considered to be inconsistent.

5.2 Testing for Autocorrelation and Heteroskedasticity

To ensure the fixed effects models provide optimal estimations, we first conducted several tests to detect possible autocorrelation and heteroscedasticity in the models.

5.2.1 Heteroskedasticity

In the presence of heteroskedasticity the variance of the error terms is not constant over time. As a result, the predicted values may be inefficient, and the estimated standard errors biased. To test for heteroskedasticity in the fixed effects models, the xttest3 command was applied in Stata. This command performs a modified Wald-test on the variance of the error terms. Under the null hypothesis the variances of the error terms are constant across all cryptocurrencies. The test statistic of the Wald test follows a chi-squared distribution with the number of groups in the panel as the degrees of freedom. For the models incorporating same-day and one-day lagged Google Trends, the results of the test are presented in table 5.2 below:

Table 5.2

Results of the modified Wald-tests for the same-day and lagged fixed effects model

	same-day FE-model	*Lagged FE-model*
X^2	52.270	52.260
P-value	0.000	0.000

Note: the results of the modified Wald test conducted to detect heteroskedasticity in the error terms of both fixed effects models. The values included in the table are the chi-squared test statistic and the corresponding p-value.

The p-values reported in the table imply that we can reject the null hypothesis of no heteroskedasticity at the 1% significance level. This might seem contradictory, since fixed effects generally control for heteroskedasticity within groups. However, heteroskedasticity can still occur in the error terms across groups. Furthermore, the model only controls for the effects of unobserved time-invariant variables. Therefore, the remaining heteroskedasticity most likely occurs because of variation in error-terms between groups or the effects of omitted variables that vary over time. To eliminate the remaining heteroskedasticity in the model, we will apply clustered standard errors.

5.2.2 Autocorrelation

Autocorrelation is a statistical phenomenon entailing that variables are correlated with their lagged values. The presence of autocorrelation does not directly imply biasedness of coefficient estimates but can result in the underestimation of standard errors. Consequently, t-statistics increase, and coefficients are more likely to be misspecified as significant. To test for autocorrelation a Wooldridge-test was conducted in Stata, with the command xtserial. Under the null-hypotheses of the test there is no first order autocorrelation. Furthermore, the test-statistic follows an F-distribution with 1 as the standard degrees of freedom for the numerator, and the number of groups minus 1 as de degrees of freedom for the denominator. The results of the Wooldridge test are presented in the table below:

Table 5.3

Results of the Wooldridge test for autocorrelation

	Coefficient
F-statistic	26.814
P-value	0.004

Note: Results of the Wooldridge test conducted to test for first-order autocorrelation in the dependent and independent variables in our model. The values included in the table are the F-test statistic and the corresponding p-value.

The values presented in the table indicate that we can reject the null hypothesis of no autocorrelation, since the p-value of 0.004 is smaller than the critical value at a 5% significance level, namely 0.05. Therefore, we must control for autocorrelation in both models. However, the clustering process discussed before also controls for autocorrelation in the variables within groups.

5.3 Model Estimation Results

In this section we will discuss the final estimations of the models applied in this study. Firstly, we will review the estimations of the fixed effects model applied to examine the relationship between the returns and Google Trends of the six cryptocurrencies on the same day. Afterwards, we will shed light on the fixed effects model applied to estimate the delayed effect of Google Trends on returns. Thirdly we will examine the fixed effects of the cryptocurrencies in both models. Finally, we will compare the fixed effects models based on their prediction ability.

5.3.1 Same-Day Fixed Effects Model

In table 5.4 below the estimated coefficients are presented for the first fixed effects model. In this model the winsorized returns of the six cryptocurrencies served as the dependent variable and the adjusted Google Trends data as the independent variable of interest. Furthermore, the coefficients of the dummy variables are also presented, which represent the effects of the adjusted Google Trends data of Bitcoin and the status of the Covid-19 pandemic.

Table 5.4

Estimations of the fixed effects model for the same day

	Coefficient	Robust std. err.	t-value	p-value
Adj. $\text{Trends}_{\text{grouped}}$	0.0006**	0.0002	2.62	0.047
Adj. $\text{Trends}_{\text{BTC}}$	-0.0004***	0.0001	-4.11	0.009
Pre-Pandemic	0.0016	0.0015	1.10	0.322
Pandemic	0.0054***	0.0013	4.33	0.008
Constant	-0.0031	0.0026	-1.18	0.291

Note: the table exhibits the estimated coefficients of the variables included in the fixed effects model that was applied to analyse the relationship between the sample returns and adjusted Google Trends on the same day. Furthermore, the standard deviations, p- and t-values corresponding to the coefficients are also presented. All coefficients are presented in decimal points, where a one unit increase in a variable with a coefficient of 0.01 implies an increase in estimated return of 1%. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The coefficient of the adjusted trends variable illustrates that there is a positive relationship between Google Trends and returns in our sample, which is significant at the 5% level. The coefficient can be interpreted as an average increase of six basis point if the Google Trends value experiences a one-unit increase. These findings are in line with past research on other cryptocurrencies and asset classes. Furthermore, the results deliver sufficient evidence to accept hypothesis 3 formulated in the empirical framework. To provide a better understanding of the scale of this effect within the groups of the panel, the results of six OLS-regressions performed on the returns and Google Trends of each individual cryptocurrency are presented in table B1 of appendix B. The contents of this table indicate that the direction and significance of the individual OLS-coefficients correspond to those of the fixed effects model for all cryptocurrencies, except OKB. Furthermore, the individual OLS-coefficients reveal that the effect of Google Trends on returns is strongest for Bitcoin Cash and weakest for Monero, with coefficients of 0.0015 and 0.0003 respectively.

Contrary to the findings of past research, Bitcoin's Google Trends exhibit a negative coefficient of -0.005, significant at the 1% level. Besides this, the coefficient opposes the previously formulated hypothesis 2. To further analyse this relationship, we performed an OLS-regression with the Google Trends variable of our sample, as the dependent variable, and the adjusted Google Trends of Bitcoin, as the independent variable. The results of this regression are presented in table B2 of appendix B and indicate a significant positive relationship between the two variables. Therefore, we conclude that the spillover effect mentioned by Anastasiadis and Papadamou (2022) only applies to the online public attention towards the cryptocurrencies in our sample, and not for the returns. Furthermore, the OLS-coefficients of the individual cryptocurrencies again show the same significance and direction for all currencies except OKB.

Regarding the dummy variables that resemble the stage of the pandemic, it can be observed that the pandemic had a significant positive effect on the returns in our sample. On average the returns of the coins were 5.4 basis points higher during the pandemic than in the period in our sample after the pandemic, which serves as the base value. These findings correspond to those of past research that the pandemic increased overall interest and returns in the cryptocurrency market and supports hypothesis 1. However, the coefficient of the period before the pandemic is insignificant and cannot be interpreted.

5.3.2 First-Order Lag Fixed Effects Model

In table 5.5 below the estimated coefficients of the fixed effects model applied to the second part of our analysis are displayed. Compared to the model above, one alteration is made, namely the application of first-order lagged Google Trends values as the independent variable, rather than the values from the same day. The application of the lagged values should enable us to detect a delayed effect, if present.

Table 5.5

Estimations of the fixed effects model for delayed effects

	Coefficient	Robust std. err.	t-value	p-value
Adj. $Trends_{grouped}$	0.0004**	0.0001	3.11	0.027
Adj. $Trends_{BTC}$	-0.0003***	0.0001	-4.01	0.010
Pre-Pandemic	0.0023	0.0014	1.415	0.152
Pandemic	0.0057***	0.0012	4.565	0.006
Constant	0.0009	0.0023	0.357	0.309

Note: the table exhibits the estimated coefficients of the variables included in the fixed effects model applied to analyse the relationship between the sample returns and first-order lagged adjusted Google Trends. Furthermore, the standard deviations, p- and t-values corresponding to the coefficients are also presented. All coefficients are presented in decimal points, where a one unit increase in a variable with a coefficient of 0.01 implies an increase in estimated return of 1%. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The table illustrates that the coefficients of the two models are similar in direction and significance but vary in size. For example, the coefficient of the lagged sample Google Trends variable is 0.0004, compared to 0.0006 for the same-day variable in the other model. Although the effect is smaller, we can still conclude that Google Trends have a delayed effect on returns, which allows us to accept hypothesis 4. In line with the analysis of the same-day effects, we again performed six OLS-regressions to observe the effect for each individual cryptocurrency. The regression results are displayed in table B3 of appendix B. The coefficients in the table range from 0.0004 to 0.0010, indicating that the effect of Google Trends also diminishes in the individual analyses. Contrary to the findings of the same-day analysis, where OKB displayed an insignificant effect, now only Monero displays insignificant effects.

Similar to the sample Google Trends, the effect of Bitcoin's Google Trends also diminished with the application of lagged values. However, the coefficient of -0.0003 indicates that there still is a significant negative effect at the 1% level. If we consider that both model's estimated coefficient representing the Google Trends of Bitcoin have negative signs, we can conclude that Bitcoin's Google Trends have an overall negative effect on the returns in our sample. This understanding provides us with sufficient evidence to reject hypothesis 2.

An interesting observation is that the coefficient of the variable related to the period during the pandemic has increased compared to the same-day fixed effects model, from 0.0054 to 0.0057. Although the change is relatively small, it is most likely related to the alteration of the temporal design. Since same day Google Trends capture the direct effect of online public attention on returns, they also capture fluctuations in public attention caused by the pandemic. However, by including the lagged variables these direct fluctuations are smoothed out, thereby increasing the observed effect of the pandemic. Furthermore, now we have established that the pandemic has a significant positive effect on returns in both models, we obtained sufficient evidence to accept hypothesis 4.

5.3.3 Fixed effects

Besides the coefficients, a key aspect of the fixed effects model we have not yet discussed is the inclusion of cryptocurrency specific fixed effects. The fixed effects could be interpreted as a base level of return for each individual cryptocurrency that is estimated by the model. As we have discussed before the fixed effects reflect underlying time-invariant factors specific to each cryptocurrency. The Fixed effects of the cryptocurrencies in our sample are displayed in table 5.6 for both models:

Table 5.6

Fixed effects for each cryptocurrency in the sample estimated by the same-day and one-day lagged model

	Fixed Effects	
	Same-day	Lagged
OKB	-0.0080	-0.0054
Bitcoin Cash	0.0009	0.0003
Cardano	0.0052	0.0038
Monero	-0.0044	-0.0031
Ripple	0.0000	-0.0004
Vechain	0.0063	0.0047

Note: the table displays the fixed effects of all six cryptocurrencies for both models used in the analysis. The left column represents the values for the same-day model, and the right column the values for the lagged model. The fixed effects are represented in decimal points.

The values presented in table 5.6 indicate that the fixed effects in the same-day model range from -0.0080 to 0.0063, for OKB and Vechain respectively. This implies that Vechain is estimated to have the highest return if the same conditions apply to all the cryptocurrencies, and OKB to have the lowest. Furthermore, the fixed effects of the lagged model are similar to those of the same-day model when it comes to their sign and size relative to each other. However, the values of the lagged model are consistently smaller than their counterparties of the same-day model, ranging from -0.0054 to 0.0047, for OKB and Vechain respectively. This observation might seem contradictory, since the fixed effects represent the same time-invariant characteristics in both models. However, since the lagged model represents the delayed effect Google Trends have on returns, the fixed effects in this model also reflect the delayed effect of the time-invariant characteristics on returns. Therefore, we can conclude that the direct effect of the cryptocurrency specific time-invariant characteristics on returns, reflected by the fixed effects of the same-day model, is stronger than the delayed effect, represented by the fixed effects of the lagged model.

5.3.4 Comparing models

The past sections highlight the similarities and differences between the models applying same-day and one-day lagged Google Trends. However, the results reported above do not allow us to draw a concrete conclusion about the estimation strength of the models. Therefore, we derive the Aikake Information Criterion in Stata. The AIC is determined based on the number of parameters and the log-likelihood estimate that captures the probability that a model could have predicted the observed independent value. Afterwards, the AICs of the two model are compared and the model with the lowest value is considered to be the best fit. The AICs for both models are reported in table 5.7 below:

Table 5.7

AICs for the same-day and lagged fixed effects model

	AIC
Same-day FE-model	-27,783.02
Lagged FE-model.	-27,709.55

Note: AICs for the same-day and lagged fixed effects model. The values do not have a specific unit of measurement.

From the table we can conclude that the FE-model has better predictive capabilities and fits the data best. A rule of thumb in comparing models using AIC is that a difference of more than two indicates that a model is significantly better. Therefore, we can conclude that in our analysis the FE-model is significantly better.

5.4 Discussion

In this chapter we will compare the findings of our analysis to those of past literature. The research design in this paper deviated from past literature in the methods applied and cryptocurrencies of interest. However, our findings largely correspond to those previously established in past literature.

First and foremost, the results of our analysis on the relationship between Google Trends and the returns of the six cryptocurrencies of interest. Similar to Raza et al. (2022) we established a positive relationship between these two variables. Furthermore, our results also correspond to those of similar analysis performed on different asset classes. For example, Hu et al. (2018) who found that Google Trends increased their model's ability to predict the direction of the stock market. The second model in our analysis is dedicated to the delayed effect of Google Trends. In line with the findings of Nasir et al. (2019), in their VAR analysis of the returns on Bitcoin, we also established a positive effect.

On the other hand, our findings contradict to those of past research when it comes to the role of Bitcoin as a market leader. Anastasias and Papadomou (2021) described that both the returns and Google Trends of Bitcoin positively influenced the returns and Google Trends values of several other cryptocurrencies. Our results indicate the opposite and display a negative effect of Bitcoin's Google Trends on the returns of the cryptocurrencies in our sample. Then again, a positive effect was established on the Google Trends values of the cryptocurrencies.

Regarding the Covid-19 pandemic, we used two dummy variables to examine the effect of this period, resulting in the observation of a positive effect of the pandemic on returns. Although we did not encounter research that examined the effect of this period in a similar way, we can compare our results to the general notions about this period and the cryptocurrency market. Wątorek et al. (2021) examine the development of cryptocurrency market characteristics throughout a period that coincides with the pandemic. The authors conclude that the market has developed from a peripheral state to one resembling an intermediate sized stock index. In line with these findings, Caferra and Vidal Tomás (2021) describe a long-run correlation between the stock and cryptocurrency market following the Covid-19 crash. Furthermore, the authors highlight that the cryptocurrency market recovered faster than the stock market in the short run. In conclusion, past research highlights a positive effect of the pandemic on the cryptocurrency market, which corresponds to our findings.

CHAPTER 6 CONCLUSION

In this paper we have examined the relationship between online public attention and the returns on six different cryptocurrencies. In recent years, promising stories about excessive cryptocurrency returns have been hard to elude. However, it also seems the positive stories outweigh the negative, and in a market that strongly resembles a financial bubble at times, it is hard to evaluate whether cryptocurrencies could be a rational investment. Therefore, we aimed to establish a method of quantifying the effect of online public attention on the returns on six different cryptocurrencies using Google Trends data. Although extensive research had been conducted on this relationship for different asset classes and more established cryptocurrencies, such as Bitcoin, smaller coins had been overlooked. Therefore, this study concentrated on smaller and lesser-known cryptocurrencies, namely OKB, Bitcoin Cash, Cardano, Monero, Ripple and VeChain, in the aim to answer the research question: "What is the effect of online public attention, approximated with Google Trends, on the performance of six different cryptocurrencies?"

To provide a broad understanding of the relationship between online public attention and returns, we divided our analysis into two parts, varying in temporal design. In the first part, the direct effect was studied by using Google Trends and returns data for the same day. The second part was devoted to analysing a potential delayed effect of public attention, by using the first-order lagged Google Trends data. Furthermore, several control variables were included in the analysis. Since past literature highlights that Bitcoin serves as a market leader that positively influences the returns on other cryptocurrencies, Google trends data related to Bitcoin was included as a control variable. Moreover, dummy variables representing the stage of the pandemic were also included, because past literature describes the pandemic as a period of strong development in the cryptocurrency market. Due to the panel nature of our dataset, containing observations across six cryptocurrencies over a period of four years, we contemplated using either a fixed or random effects model. The Hausman test was applied to determine the best fitting model, resulting in the application of the fixed effects model in both parts of the analysis.

The results of the analyses indicate that online public attention, approximated through Google Trends data, has a significantly positive current and delayed effect on the returns in our sample. Furthermore, contrary to the findings of past research, public attention towards Bitcoin exhibited a significantly negative effect on the returns in our sample. However, in line with past research we observed a positive effect of the pandemic on the returns in both parts of the analysis.

In conclusion, the results of this study show that online public attention, measured through Google Trends data, is a valid instrument to estimate the returns on the cryptocurrencies in our sample. These findings correspond to those of past research on other asset classes and cryptocurrencies. Google Trends data can therefore be considered as a valid instrument in the estimation of asset returns, across different asset classes.

CHAPTER 7 LIMITATIONS

In this chapter we will discuss some limitations we noticed during our analysis. Consequently, we will provide recommendation for future analysis based on these limitations. Revisiting the different aspect of our analysis, we have noticed three main limitations. Firstly, we would like to reconsider the inclusion of OKB as one of the cryptocurrencies of interest in this analysis. In hindsight we believe that the Google Trends data related to the search query ‘OKB’ also reflected public attention towards other goods or services. As a result, OKB displayed divergent results throughout the analysis in comparison to the other cryptocurrency. Not only is the reported mean of the normalized Google Trends values substantially higher for OKB, but OKB also showed little significant correlations and coefficients in comparison to the other coins. Therefore, we believe the Google Trends data on OKB is biased. Moreover, the same problems could be discussed for Monero, but we believe these problems to be less prevalent.

Secondly, we would like to devote our attention to the application ability of the results established. The main effect that was assessed in this study is the effect of Google Trends on returns, with the objective to establish a relationship that could possibly be applied in, for example, an investment strategy. However, in practice the application possibilities are limited. Although current Google Trends data is available at any point in time during a day, this can still be subject to change throughout the day. Therefore, it would be unwise to base an investment decision on the Google Trends value that is observed on the same day. However, the delayed relationship we established does offer potential applications in investment strategies. Therefore, we would recommend further analysis to be conducted on the potential application abilities of lagged Google Trends values.

Lastly, we will consider an aspect of our analysis we do not directly consider as a limitation, namely the inclusion of Bitcoin’s Google Trends. As we mentioned before, the effect of the variable representing Bitcoin’s Google Trends resulted in a negative coefficient, contrary to the findings of past research. Our decision to include this variable was based on the previously observed positive effect of both Bitcoin’s Google Trends data and returns on the Google Trends and returns of other cryptocurrencies. In hindsight we could have included both the returns and Google Trends of Bitcoin and would therefore recommend the inclusion of both in future research, to examine these effects closer.

REFERENCES

Anastasiadis, P., & Papadamou, S. (2022). The dimension of popularity in the cryptocurrency market. SN Business & Economics, 2(5). https://doi.org/10.1007/s43546-022-00206-5

Aslanidis, N., Bariviera, A. F., & López, Ó. G. (2022). The link between cryptocurrencies and Google Trends attention. Finance Research Letters, 47, 102654. https://doi.org/10.1016/j.frl.2021.102654

Barber, B. M., & Odean, T. (2001). The Internet and the Investor. The Journal of Economic Perspectives, 15(1), 41–54. http://www.jstor.org/stable/2696538

Bach, L. M., Mihaljevic, B., & Zagar, M. (2018). Comparative analysis of blockchain consensus algorithms. 2018 41st International Convention On Information And Communication Technology, Electronics And Microelectronics (MIPRO), 1545–1550. https://doi.org/10.23919/mipro.2018.8400278

Benedetti, H., & Kostovetsky, L. (2021). Digital Tulips? Returns to investors in initial coin offerings. Journal Of Corporate Finance, 66, 101786. https://doi.org/10.1016/j.jcorpfin.2020.101786

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of* Computational *Science*, *2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Brini, A., & Lenz, J. (2024). A Comparison of Cryptocurrency Volatility-benchmarking New and Mature Asset Classes. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2404.04962

Brooks, C. (2019). *Introductory Econometrics for Finance* (4th ed.). Cambridge University Press. https://doi.org/10.1017/9781108524872

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

Carrière-Swallow, Y., & Labbé, F. (2011). Nowcasting with Google Trends in an Emerging Market. Journal Of Forecasting, 32(4), 289–298. https://doi.org/10.1002/for.1252

Clark, T. S., & Linzer, D. A. (2014). Should I use fixed or random effects? *Political Science Research and Methods*, *3*(2), 399–408. https://doi.org/10.1017/psrm.2014.32

Daoud, J. I. (2017). Multicollinearity and Regression Analysis. *Journal Of Physics. Conference Series*, *949*, 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. Management Information Systems Quarterly, 13(3), 319. https://doi.org/10.2307/249008

deHaan, E. (2020). Practical guidance on using and interpreting fixed effects models. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3699777

FAQ about Google Trends data – Trends Help. (2024). https://support.Google.com/Trends/answer/4365533?hl=en

Fox, M. (2024, March 15). *Dit zijn de 8 grootste "memecoins" op de cryptomarkt*. Business Insider Nederland. https://www.businessinsider.nl/dit-zijn-de-8-grootste-memecoins-op-de-cryptomarkt-en-zo-hard-stijgen-ze/

Ganesh, A. U. J., & Iyer, S. S. (2021). Impact of Firm-Initiated tweets on stock return and trading volume. Journal of Behavioral Finance, 24(2), 171-182. https://doi.org/10.1080/15427560.2021.1949717

Gardiner, J. C., Luo, Z., & Roman, L. A. (2008). Fixed effects, random effects and GEE: What are the differences? *Statistics in Medicine*, *28*(2), 221–239. https://doi.org/10.1002/sim.3478

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., & Watts, D. J. (2010). Predicting consumer behavior with Web search. Proceedings Of The National Academy Of Sciences Of The United States Of America, 107(41), 17486–17490. https://doi.org/10.1073/pnas.1005962107

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*(6), 1251. https://doi.org/10.2307/1913827

Haykır, Ö., & Yağlı, İ. (2022). Speculative bubbles and herding in cryptocurrencies. Financial Innovation, 8(1). https://doi.org/10.1186/s40854-022-00383-0

Hu, H., Tang, L., Zhang, S., & Wang, H. (2018). Predicting the direction of stock markets using optimized neural networks with Google Trends. Neurocomputing, 285, 188–195. https://doi.org/10.1016/j.neucom.2018.01.038

Huang, M. Y., Rojas, R. R., & Convery, P. D. (2019). Forecasting stock market movements using Google Trend searches. Empirical Economics, 59(6), 2821–2839. https://doi.org/10.1007/s00181-019-01725-1

Joo, M. H., Nishikawa, Y., & Dandapani, K. (2019). Cryptocurrency, a successful application of blockchain technology. Managerial Finance, 46(6), 715–733. https://doi.org/10.1108/mf-09-2018-0451

Kaiser, L. (2019). Seasonality in cryptocurrencies. *Finance Research Letters*, *31*. https://doi.org/10.1016/j.frl.2018.11.007

Lee, M. (2009). Predicting and explaining the adoption of online trading: An empirical study in Taiwan. Decision Support Systems, 47(2), 133–142. https://doi.org/10.1016/j.dss.2009.02.003

Mazur, M., Dang, M., & Vega, M. (2021). COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. *Finance Research Letters*, *38*, 101690. https://doi.org/10.1016/j.frl.2020.101690

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

NOS. (2021, January 27). Beurskoersgekte losgebroken rond videogame-winkel GameStop. NOS. https://nos.nl/artikel/2366194-beurskoersgekte-losgebroken-rond-videogame-winkel-gamestop

Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., & Murugiah, K. (2014). The Use of Google Trends in Health Care Research: A Systematic Review. PloS One, 9(10), e109583. https://doi.org/10.1371/journal.pone.0109583

Ofir, M., & Sadeh, I. (2019). ICO vs IPO: Empirical Findings, Information Asymmetry and the Appropriate Regulatory Framework. Social Science Research Network. https://doi.org/10.2139/ssrn.3338067

Oikonomopoulos, S., Tzafilkou, K., Karapiperis, D., & Verykios, V. S. (n.d.). Cryptocurrency Price Prediction using Social Media Sentiment Analysis. 2022 13th International Conference on

Information, Intelligence, Systems & Applications (IISA). https://doi.org/10.1109/iisa56318.2022.9904351

Oksanen, A., Mantere, E., Vuorinen, I., & Savolainen, I. (2022). Gambling and online trading: emerging risks of real-time stock and cryptocurrency trading platforms. Public Health, 205, 72–78. https://doi.org/10.1016/j.puhe.2022.01.027

Online trading platform market Size | Global Growth [2032]. (2024). https://www.fortunebusinessinsights.com/online-trading-platform-market-104934

Özdemir, O. (2022). Cue the volatility spillover in the cryptocurrency markets during the COVID-19 pandemic: evidence from DCC-GARCH and wavelet analysis. Financial Innovation, 8(1). https://doi.org/10.1186/s40854-021-00319-0

Pan, Y., Mithas, S., Hsieh, J. P., & Liu, C. (2023). Do Risk Preferences Shape the Effect of Online Trading on Trading Frequency, Volume, and Portfolio Performance? Journal Of Management Information Systems, 40(2), 440–469. https://doi.org/10.1080/07421222.2023.2196777

Raut, R. K., & Kumar, S. (2023). An integrated approach of TAM and TPB with financial literacy and perceived risk for influence on online trading intention. Digital Policy, Regulation And Governance. https://doi.org/10.1108/dprg-07-2023-0101

Raza, S. A., Yarovaya, L., Guesmi, K., & Shah, N. (2022). Google Trends and cryptocurrencies: a nonparametric causality-in-quantiles analysis. International Journal Of Emerging Markets, 18(12), 5972–5989. https://doi.org/10.1108/ijoem-10-2021-1522

Roca, J. C., García, J. J., & De La Vega, J. J. (2009). The importance of perceived trust, security and privacy in online trading systems. Information Management & Computer Security, 17(2), 96–113. https://doi.org/10.1108/09685220910963983

Ruppert, D. (1987). What is Kurtosis? An Influence Function Approach. *˜the œAmerican Statistician*, *41*(1), 1–5. https://doi.org/10.1080/00031305.1987.10475431

Sedgwick, P. (2012). Pearson's correlation coefficient. *BMJ*, *345*(jul04 1), e4483. https://doi.org/10.1136/bmj.e4483

Shete, S., Beasley, T. M., Etzel, C. J., Fernández, J. R., Chen, J., Allison, D. B., & Amos, C. I. (2004). Effect of winsorization on power and Type 1 error of variance components and related methods of

QTL detection. *Behavior Genetics*, *34*(2), 153–159. https://doi.org/10.1023/b:bege.0000013729.26354.da

Silva, E. S., Hassani, H., Madsen, D. Ø., & Gee, L. (2019). Googling Fashion: Forecasting Fashion Consumer Behaviour Using Google Trends. Social Sciences, 8(4), 111. https://doi.org/10.3390/socsci8040111

StatCounter Global Stats. (2024). Search engine market share worldwide. StatCounter Global Stats. https://gs.statcounter.com/search-engine-market-share

Statista. (2023, November 30). Size of the global online trading market 2020, with forecasts up until 2026. https://www.statista.com/statistics/1260026/forecast-global- online-trading-platform-market/

Sunder, S. (2020). How did the U.S. stock market recover from the Covid-19 contagion? *Mind & Society*, *20*(2), 261–263. https://doi.org/10.1007/s11299-020-00238-0

Taboada, M. (2016). Sentiment Analysis: An Overview from Linguistics. Annual Review of Linguistics, 2(1), 325–347. https://doi.org/10.1146/annurev-linguistics-011415-040518

Tan, S. D., & Taş, O. (2020). Social media sentiment in international stock returns and trading activity.˜ the œJournal of Behavioral Finance, 22(2), 221–234. https://doi.org/10.1080/15427560.2020.1772261

Talwar, S., Talwar, M., Tarjanne, V., & Dhir, A. (2021). Why retail investors traded equity during the pandemic? An application of artificial neural networks to examine behavioral biases. Psychology & Marketing, 38(11), 2142–2163. https://doi.org/10.1002/mar.21550

Tseng, Q. (2021, December 12). Reconstruct Google Trends daily data for extended period. *Medium*. https://towardsdatascience.com/reconstruct-Google-Trends-daily-data-for-extended-period-75b6ca1d3420

U.S. Centre for Disease Control and Prevention. (2023, March 15). *CDC Museum COVID-19 timeline*. Centers for Disease Control and Prevention. https://www.cdc.gov/museum/timeline/covid19.html

Wątorek, M., Drożdż, S., Kwapień, J., Minati, L., Oświęcimka, P., & Stanuszek, M. (2021). Multiscale characteristics of the emerging global cryptocurrency market. *Physics Reports*, *901*, 1–82. https://doi.org/10.1016/j.physrep.2020.10.005

Woloszko, N. (2020). Tracking activity in real time with Google Trends. OECD Economics Department Working Papers. https://doi.org/10.1787/6b9c7518-en

Yaffee, R. (2003). A primer for panel data analysis. *Connect: Information Technology at NYU*, *8*(3), 1-11.

Zamfir, I., & Iordache, A. M. (2022). The influences of covid-19 pandemic on macroeconomic indexes for European countries. *Applied Economics*, *54*(39), 4519–4531. https://doi.org/10.1080/00036846.2022.2031858

APPENDIX A: ADDITIONAL TABLES AND FIGURES OF CHAPTER 3

Figure A1: Cumulative returns on Bitcoin Cash throughout the sample period from January 1, 2019, through December 31, 2022.

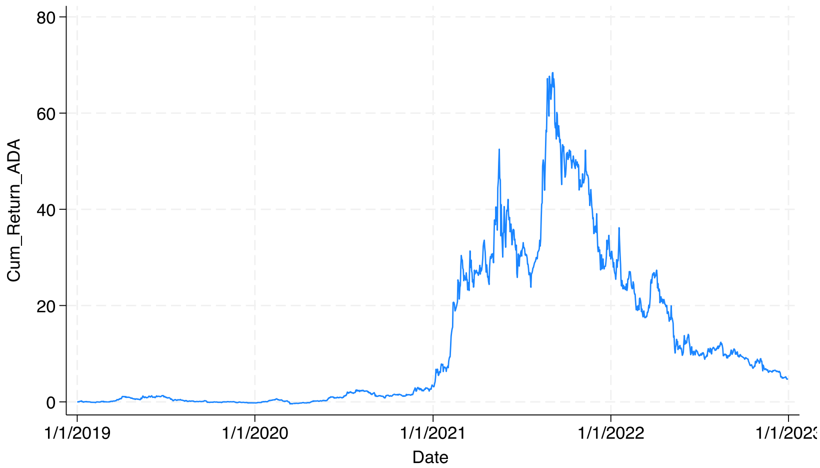

Figure A2: Cumulative returns on Cardano throughout the sample period from January 1, 2019, through December 31, 2022.

Figure A3: Cumulative returns on Monero throughout the sample period from January 1, 2019, through December 31, 2022.

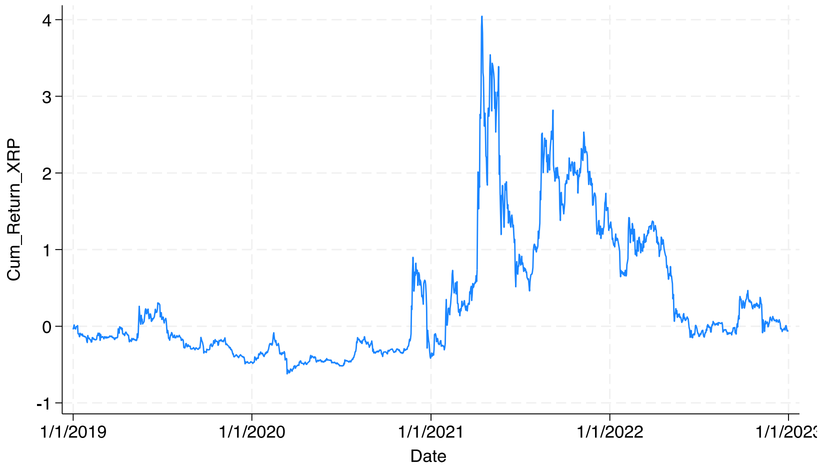

Figure A4: Cumulative returns on Ripple throughout the sample period from January 1, 2019, through December 31, 2022.

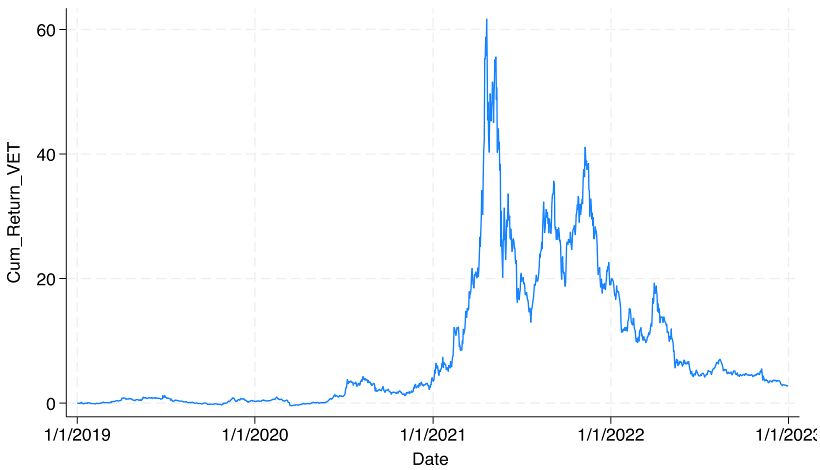

Figure A5: Cumulative returns on VeChain throughout the sample period from January 1, 2019, through December 31, 2022.

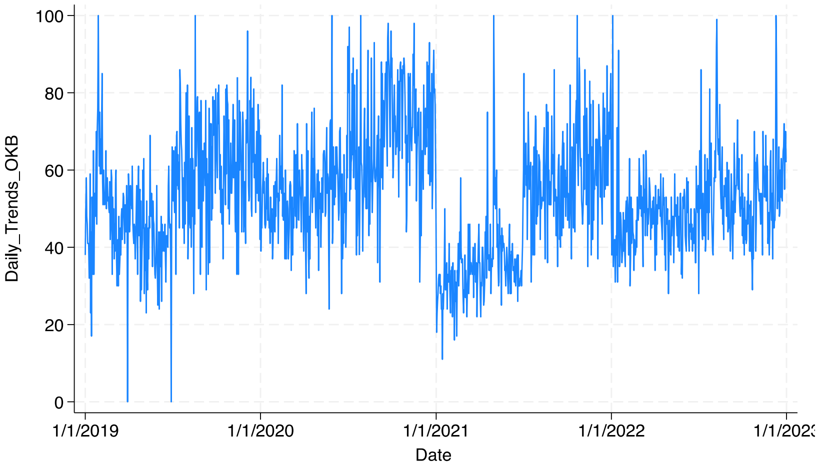

Figure A6: the plotted daily Google Trends values of OKB, which were collected in 6-month intervals, adding up to the length of the sample period.

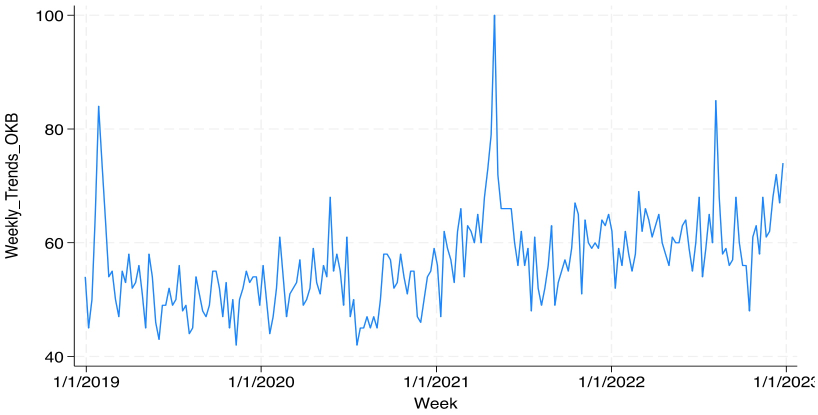

Figure A7: plot of the weekly Google Trends values of OKB, which were collected at once for the entire sample period.

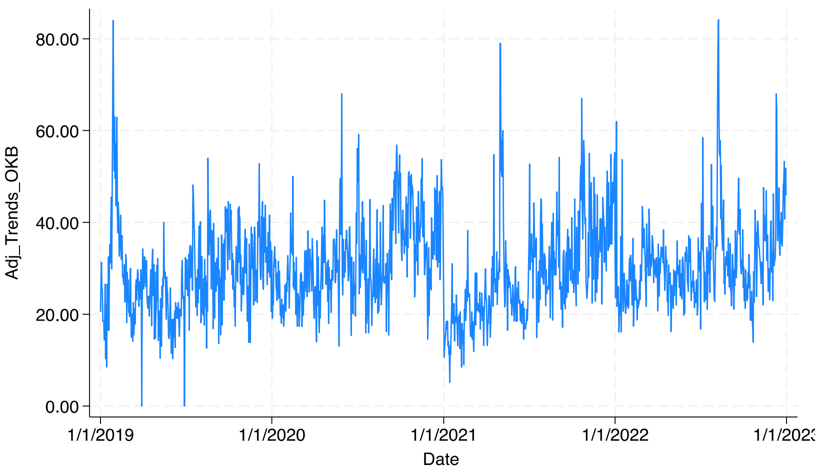

Figure A8: plot of the normalized daily Google Trends values of OKB, which were derived by multiplying the collected daily values by the value of the corresponding week.

Table A1

Pearson's correlation matrix for the returns on OKB and the independent variables

	$Return_{OKB}$	Adj. $Trends_{OKB}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{OKB}$	1.000				
Adj. $Trends_{OKB}$	0.036	1.000			
Adj. $Trends_{BTC}$	-0.050*	-0.092***	1.000		
Pre-Pandemic	0.028	-0.129***	-0.362***	1.000	
Pandemic	0.001	0.057**	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between the returns on OKB and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2

Pearson's correlation matrix for the returns on Bitcoin Cash and the independent variables

	$Return_{BCH}$	Adj. $Trends_{BCH}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{BCH}$	1.000				
Adj. $Trends_{BCH}$	0.164***	1.000			
Adj. $Trends_{BTC}$	-0.053**	0.485***	1.000		
Pre-Pandemic	0.016	-0.062**	-0.362***	1.000	
Pandemic	0.028	0.215***	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between the returns on Bitcoin Cash and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3

Pearson's correlation matrix for the returns on Cardano and the independent variables

	$Return_{ADA}$	Adj. $Trends_{ADA}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{ADA}$	1.000				
Adj. $Trends_{ADA}$	0.079***	1.000			
Adj. $Trends_{BTC}$	-0.024	0.434***	1.000		
Pre-Pandemic	-0.018	-0.324***	-0.362***	1.000	
Pandemic	0.075***	0.186***	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between the returns on Cardano and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4

Pearson's correlation matrix for the returns on Monero and the independent variables

	$Return_{XMR}$	Adj. $Trends_{XMR}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{XMR}$	1.000				
Adj. $Trends_{XMR}$	0.016	1.000			
Adj. $Trends_{BTC}$	-0.077***	0.643***	1.000		
Pre-Pandemic	-0.007	-0.370***	-0.362***	1.000	
Pandemic	0.032	0.248***	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between the returns on Monero and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A5

Pearson's correlation matrix for the returns on Ripple and the independent variables

	$Return_{XRP}$	Adj. $Trends_{XRP}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{XRP}$	1.000				
Adj. $Trends_{XRP}$	0.148***	1.000			
Adj. $Trends_{BTC}$	-0.051*	0.266***	1.000		
Pre-Pandemic	-0.020	0.133***	-0.362***	1.000	
Pandemic	0.044*	0.043*	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between the returns on ripple and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6

Pearson's correlation matrix for the returns on VeChain and the independent variables

	$Return_{VET}$	Adj. $Trends_{VET}$	Adj. $Trends_{BTC}$	Pre-Pandemic	Pandemic
$Return_{VET}$	1.000				
Adj. $Trends_{VET}$	0.125***	1.000			
Adj. $Trends_{BTC}$	-0.047*	0.330***	1.000		
Pre-Pandemic	-0.016	-0.253***	-0.362***	1.000	
Pandemic	0.068***	0.336***	0.235***	-0.590***	1.000

Note: matrix of the Pearson's correlations between the returns on VeChain and the independent variables. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7

Variance Inflation Factor (VIF) of the independent variables

	VIF
Adj. $Trends_{Grouped}$	1.07
Adj. $Trends_{BTC}$	1.21
Pre-Pandemic	1.67
Pandemic	1.54
Average	1.37

Note: VIF of all independent variables in our analysis determined to check for multicollinearity. A rule of thumb is that a VIF < 5 is acceptable.

APPENDIX B: ADDITIONAL TABLES AND FIGURES OF CHAPTER 5

Table B1

Results of the OLS-regressions of the individual cryptocurrency returns on the same-day independent variables

	Return					
	OKB	BCH	ADA	XMR	XRP	VET
Adj. Trends	0.0002	0.0015***	0.0007***	0.0003**	0.0013***	0.0011***
	(0.225)	(0.000)	(0.009)	(0.019)	(0.000)	(0.000)
Adj. $Trends_{BTC}$	-0.0002	-0.0009***	-0.0004*	-0.0007***	-0.0006***	-0.0005**
	(0.303)	(0.000)	(0.078)	(0.004)	(0.003)	(0.013)
Pre-Pandemic	0.0038	-0.0044	0.0058	0.00022	-0.0088**	0.0018
	(0.266)	(0.238)	(0.110)	(0.949)	(0.015)	(0.619)
Pandemic	0.0029	-0.0007	0.0107***	0.0040	0.00171	0.0059*
	(0.329)	(0.828)	(0.001)	(0.175)	(0.559)	(0.098)
Constant	-0.00190	-0.00224	-0.00325	0.0011	-0.00484	0.00198
	(0.739)	(0.475)	(0.368)	(0.767)	(0.195)	(0.583)

Note: estimated coefficients of the OLS-regressions of the individual returns of each cryptocurrency on the independent same-day variables. The corresponding p-values are presented in brackets underneath the coefficients. With significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1

Table B2

Results of the OLS-regression of the sample Google Trends on the Google Trends of Bitcoin

	Coefficient	Robust std. err.	t-value	p-value
Adj. $Trends_{BTC}$	0.3184***	0.0174	18.350	0.000
Constant	10.0392***	0.2196	45.710	0.000

Note: estimated coefficients of the OLS-regression of the sample Google Trends on the Google Trends of Bitcoin. The corresponding standard errors, p-value and t-value are presented in the remaining columns. With significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1

Table B3

Results of the OLS-regressions of the individual cryptocurrency returns on the lagged independent variables

	Return					
	OKB	BCH	ADA	XMR	XRP	VET
Lag. Adj. Trends	0.0004**	0.0010***	0.0005*	0.0002	0.0007*	0.0007**
	(0.015)	(0.002)	(0.075)	(0.132)	(0.086)	(0.023)
Lag. Adj. $Trends_{BTC}$	-0.0001	-0.0006**	-0.0002	-0.0005**	-0.0004*	-0.0004
	(0.813)	(0.013)	(0.395)	(0.022)	(0.052)	(0.109)
Pre-Pandemic	0.0057	-0.0013	0.0060	0.0000	-0.0049	0.0027
	(0.113)	(0.728)	(0.112)	(0.993)	(0.198)	(0.466)
Pandemic	0.0029	0.0014	0.0106***	0.0040	0.0031	0.0074**
	(0.319)	(0.637)	(0.001)	(0.169)	(0.288)	(0.040)
Constant	-0.0102	-0.0020	-0.0041	0.0021	-0.0022	0.0006
	(0.109)	(0.555)	(0.300)	(0.993)	(0.579)	(0.883)

Note: estimated coefficients of the OLS-regressions of the individual returns of each cryptocurrency on the independent one-day lagged variables. The corresponding p-values are presented in brackets underneath the coefficients. With significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$