MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Can Tweets by Elon Musk affect Bitcoin volatility?

verfasst von / submitted by

Benjamin Jaquemar, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 974
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Banking and Finance
Betreut von / Supervisor:	Univ.-Prof. Dr. Nikolaus Hautsch

Table of contents

1.) Introduction

Bitcoin, and cryptocurrencies in general, are receiving more and more attention in recent years due to their rise in market capitalization and both their remarkably high returns and capacity for substantial losses. Furthermore, influential, and famous people such as Tesla CEO Elon Musk, investor, and entrepreneur Mark Cuban; or even the president of El Salvador, Nayib Bukele, who made Bitcoin a legal tender in 2021, helped not only to draw more attention to Bitcoin but also to other cryptocurrencies. Launched in 2009 by its creator, who assumes the alias "Satoshi Nakamoto", Bitcoin was initially only known and used by a few people but has since experienced a stark rise in its interest.

Since then, many other cryptocurrencies have been developed and the cryptocurrency market capitalization has grown rapidly, seeing a rise from around 19 billion euros in the beginning of 2017 to over two trillion euros at the end of 2021.[1] The fact that in September 2021 the two biggest cryptocurrencies in market capitalization, Bitcoin and Ethereum were among the top 20 traded assets worldwide further evidences the growing importance of this asset class (Iyer, 2022). Cryptocurrencies have also become increasingly interesting for retail investors, fuelled by the rise in popularity of easy-to-use trading platforms like Robinhood and others. In addition, big asset management firms, such as BlackRock, have also started offering Bitcoin investment products to their clients.[2]

While Bitcoin can be defined as a digital or virtual currency within a decentralized payment network, Baur et al. (2017) find that by analysing Bitcoin transaction data, the currency can also be identified as having a predominantly speculative purpose, as opposed to being an alternative medium of exchange. In comparison with more traditional assets, a characteristic of Bitcoin and many other cryptocurrencies is their

---

[1] Cf. CoinMarketCap. Available at https://coinmarketcap.com/ (retrieved March 1st, 2022)
[2] Cf. Reuters. Available at https://www.reuters.com/business/finance/blackrock-close-filing-bitcoin-etf-coindesk-2023-06-15/ (retrieved July 2nd, 2023)

high level of volatility, where considerable price fluctuations that occur within a day, or even shorter periods of time, are commonplace (Lyócsa, et al., 2020).

One person who some claim can cause such substantial changes in volatility and the price of Bitcoin itself is Elon Musk – at least according to certain newspapers and online-media platforms. For instance, Rani Molla states in an article for Vox: “When Elon Musk tweets, crypto prices move”.[3] Being one of the wealthiest people on earth and the owner of multiple large enterprises, Musk’s stances on Bitcoin and other cryptocurrencies are considerably influential, especially when it comes “real-world” adaption and usage. For instance, when Musk added “#bitcoin” to his Twitter (now: X) bio on 29th January 2021, the price of Bitcoin spiked around 20 percent that day.[4]

By observing both the realized volatility (see chapter 3 for further details about the estimation of realized volatility) and the value of Bitcoin in response to posts on Twitter by Elon Musk, it could be hypothesised that there is a direct correlation between Elon Musk’s posts and announcements on Twitter and the price of Bitcoin.[5] To demonstrate this theory, Figure 1 and Figure 2 plot the logarithm of realized volatility and the price of Bitcoin in US-Dollar around two posts of Musk on Twitter. The value of Bitcoin and its volatility following the tweet by Elon Musk “You can now buy a Tesla with Bitcoin” on March 24th, 2021, are shown in Figure 1. This came as a major announcement for many as there were only a few large corporations who accepted Bitcoin as a means of payment. Following the CEO of Tesla’s certification of the validity of the cryptocurrency within its own commercial affairs in March 2021, there was a hope that this would be a significant step towards a more widely spread adaption of this cryptocurrency. It is possible to observe in Figure 1 the spike in volatility on the day of the tweet itself and on the day after. On the other hand, the price of Bitcoin does not appear to react positively on the day and the day after this tweet.

---

[3] Cf. Rani Molla on VOX. Available at https://www.vox.com/recode/2021/5/18/22441831/elon-musk-bitcoin-dogecoin-crypto-prices-tes (retrieved Octobe 22nd, 2021)
[4] A short text field on a person's Twitter profile that can be filled with information you want to share about yourself.
[5] Twitter was renamed to X in July 2023. However, I will continue to refer to this social media platform as Twitter in this thesis.

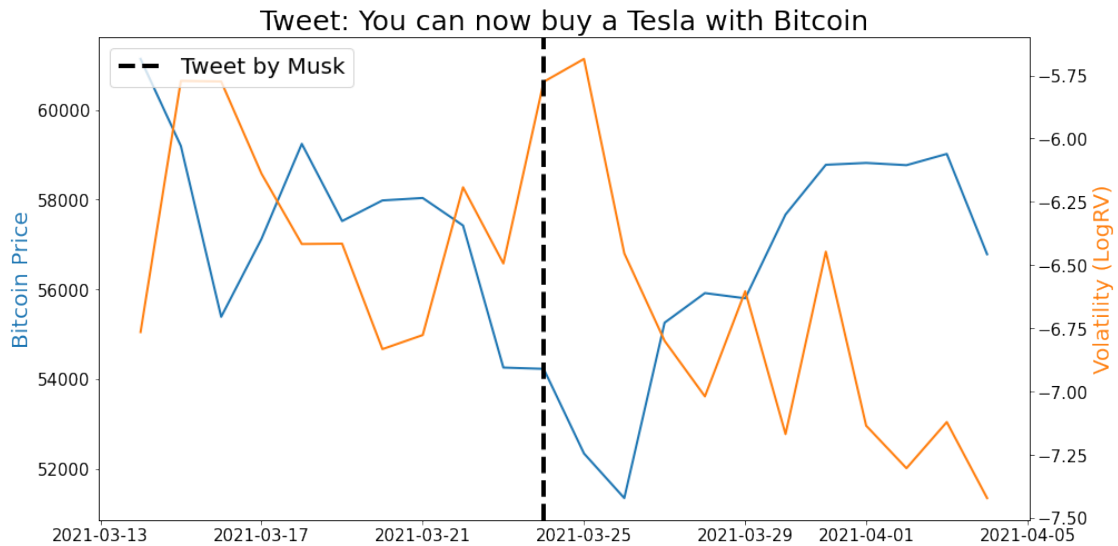

Figure 1, Bitcoin Price in USD & Log Realized Variance around Tweet by Musk on 24th March 2021

Figure 2 demonstrates the price and log realized volatility of Bitcoin after an announcement made by Musk on Twitter in May 2021, which disclosed that the possibility of using Bitcoin as a means of payment for Tesla cars was to be suspended (Tweet: "Tesla & Bitcoin" together with a screenshot of text with a more detailed explanation). This Tweet seems to be consistent with a temporary change in the realized volatility and price of the cryptocurrency, demonstrating a spike in the volatility of Bitcoin because of Musk's Tweet on the day of the Tweet and the following day, and a significant decline in price following the post, which declared that Tesla would no longer accept Bitcoin as a medium of exchange.

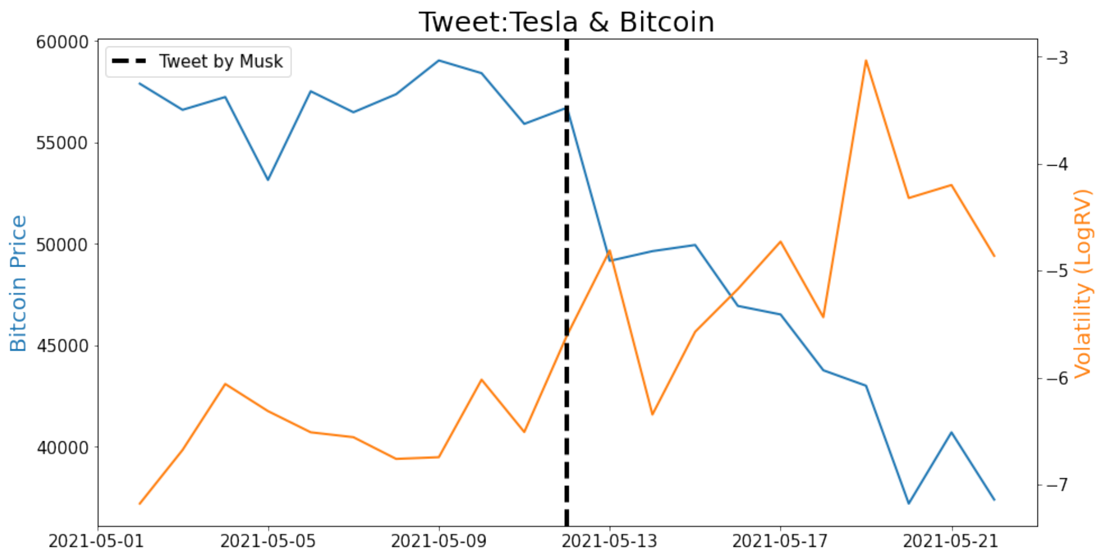

Figure 2, Bitcoin Price in USD & Log Realized Volatility around Tweet by Musk on 12th May 2021

Further to the comments on Twitter regarding Bitcoin and Tesla, Musk also used the social media platform, which he now owns, to talk more broadly about Bitcoin and other cryptocurrencies. And, as previously illustrated, these posts on Twitter seem to have caused changes in the price and volatility of Bitcoin. But is this just a coincidence or can Tweets by Elon Musk indeed cause significant changes in Bitcoin volatility?

In this master's thesis, I will examine this question by modelling and forecasting the volatility of Bitcoin returns, applying a heterogeneous autoregressive model of realized volatility (HAR-RV) and utilising high-frequency intraday data of Bitcoin prices. The aim of this thesis is to find evidence that Tweets made by Musk on cryptocurrencies do, in fact, influence the volatility of Bitcoin prices, and thus volatility models can also benefit from this information.

By augmenting HAR-RV models with data regarding Musk's cryptocurrency-related posts on Twitter, I find that these posts shared on the social networking platform are associated with an increase in realized volatility, and that augmenting HAR-RV models with such data can help to strengthen modelling and forecasting realized volatility of Bitcoin.

This is the first thesis to investigate the effect of Tweets made by Musk on volatility using a HAR-RV model. The examination contributes to the research on volatility modelling and the forecasting of Bitcoin prices and indicates that considering posts by influential people on Twitter, or possibly also other social media platforms, can improve volatility models. Additionally, this study also approaches the field of information economics through its analysis of how public information shared on Twitter can influence economic decisions made by traders and how it affects the volatility of important assets like Bitcoin.

2.) Literature Review

There is a growing body of scientific literature on modelling and forecasting the volatility of Bitcoin, concurrent with its increase in market capitalization, economical adaption, and the growing interest in Bitcoin and cryptocurrencies in general.

There are several papers on modelling and forecasting Bitcoin volatility, with one of the most recent by Bergsli et al. (2021), who investigate the volatility of Bitcoin considering GARCH (generalized autoregressive conditional heteroskedasticity and HAR (heterogenous autoregressive) models for forecasting Bitcoin volatility using realized variance from high frequency data. They find, among other things, that RV-based HAR models are better in forecasting volatility than GARCH models, especially for short-period volatility forecasts.

Lyócsa et al. (2020) examine the impacts of macroeconomic news, regulation, and hacking exchange markets on the volatility of bitcoin using an HAR-RV-model, augmented with data on these shocks. They find Bitcoin's volatility to be much higher when compared to other financial assets, and that especially Bitcoin-specific factors like regulation or hacking drives Bitcoin volatility.

Hu et al. (2019) investigate the risk dynamics of Bitcoin from the perspective of realized volatility utilizing HAR-type models. Their analysis demonstrates that the one-day lagged RV estimator drives the future realized variance significantly, and that for short-term one-day ahead forecasting, price jumps do not add extra forecasting accuracy.

Aalborg, Molnár and de Vries (2019) study the drivers of volatility as well as the return and trading volume of Bitcoin. For their volatility models they employ HAR-RV models and show that these models work well to model Bitcoin volatility. Moreover, when including additional variables to their volatility models, Aalborg et al. (2019) find Bitcoin trading volume and changes in the number Bitcoin addresses to be highly significant drivers and Google Trends to be a weakly significant driver of Bitcoin volatility.

In his investigation of the effect of Elon Musk's Twitter activity on cryptocurrency markets, undertaken using an event study approach and a sample of 47 posts on Twitter, Ante (2023) finds that Musk's tweets cause significant abnormal returns, as well as a highly significant increase in the trading volume of the respective cryptocurrency mentioned by Musk, drawing the conclusion that these Tweets do have an influence on cryptocurrency markets. For Ante, this raises questions on how to deal with the contradiction between investors protection and the right of freedom of expression by people of public interest or "influencers", who can move markets with their comments.

Shen et al. (2018) consider the link between the number of tweets from Twitter and the volatility and the trading volume of Bitcoin in their paper. Using VAR (vector autoregressive) models and applying causality tests, it is discovered that the number of tweets notably influence the following day's Bitcoin realized volatility and volume, but not the cryptocurrency's returns.

3.) Data

The data used for the calculation in this thesis I obtained from cryptodownload.com, Google as well as from the social network service Twitter. The sample period of the data ranges from January 1st, 2017, until October 31st, 2021. This amounts to 1765 daily observations for Bitcoin prices, over a time span of more than three years. I consider the period between 2017 and 2021 to be sufficient for my estimations. As data sets had timestamps from different time zones, I use UTC, the coordinated universal time, as the time standard.

In contrast to volatility models of other more traditional assets, Bitcoin is traded 24 hours a day, without breaks on weekends or bank holidays. Therefore, there are data entries for each of these days within the sample period. In the following sections, I would like to outline a detailed description of the data used in my thesis.

3.1) Bitcoin Volatility

As volatility is latent, an adequate measure to estimate the volatility of Bitcoin is required. Therefore, the concept of realized variance (RV) as a proxy of the true variance is used. The high frequency data, I use was downloaded from the website cryptodatadownload.com which provides historical cryptocurrency data from various crypto currency exchanges for hundreds of cryptocurrency pairs.[6] The various data sets on this website are retrieved by the CryptoDataDownload from different cryptocurrency exchanges, and differ in trading pair, period and the intraday frequency of the data entries provided. The data I employ for the estimations in this thesis stem from the exchanges Bitstamp and Bitfinex, providing Bitcoin prices in US Dollar on a minutely basis. The decision to use the data of these two exchanges was based on data quality, as both are well-established business platforms and therefore had sufficient liquidity on their platforms to provide the high-frequency data requisite for the entire sample period. The primary data set that I utilized was Bitstamp, but due to missing data entries for around thirty days in November and December 2020, the application of data from Bitfinex was necessary to replace the absent information. The merging of these two data sets was accomplishable due to there being no variation in the prices of their respective Bitcoin/USD quotes. In Figure 3, the entire sample of Bitcoin prices in US-Dollars is shown. It can be observed that the price fluctuated a substantial amount during this period, with the most visible spikes occurring around the end of 2017, the beginning of 2018 and towards the start of 2021.

---

[6] Cf. CryptoDataDownload. Available at https://www.cryptodatadownload.com/ (retrieved October 2021).

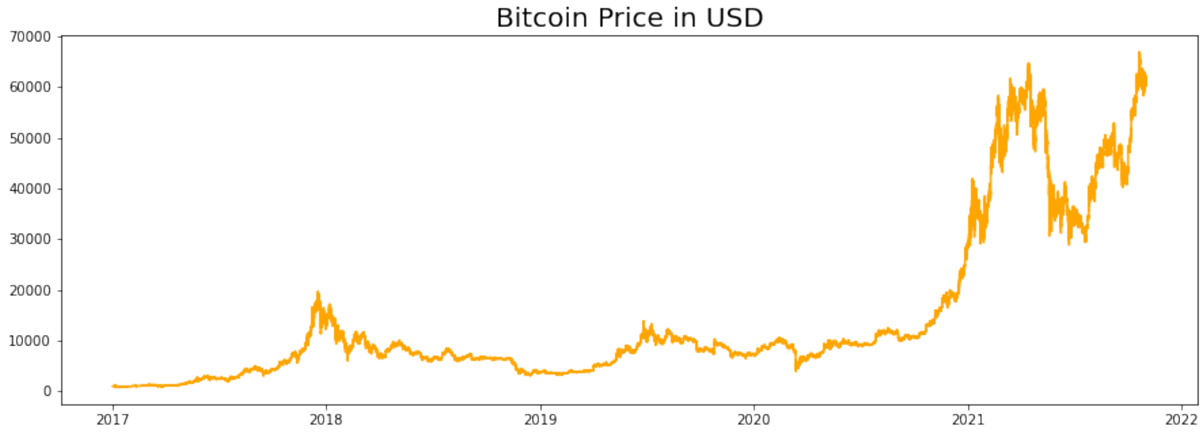

Figure 3, Bitcoin Price in US-Dollar

3.1.1) Realized Variance

This high-frequency data of Bitcoin prices is used to estimate the realized variance (RV). To estimate the RV, I followed the steps of Christoffersen (2012). Firstly, I calculated intraday logarithmic returns as shown in (1), where *m* is the total number of observations on a given day, $R_{t+j/m}$ denoting the *j*th return and $S_{t+j/m}$ denoting the *j*th closing price on day *t+1*. For my analysis I use five-minute returns, where the value of *m* is equal to *288*. In general, a lower time grid is desirable when constructing realized measures, but this can also cause difficulties, as problems with microstructural effects may be more likely to arise. The decision to choose a five-minute time grid is based on the paper by Liu et al. (2015) who find that after a thorough investigation of several assets in different asset classes, that it is "difficult to significantly beat 5-minute RV". The time series of these five-minute logarithmic returns is shown in Figure 4.

$$R_{t+j/m} = \ln(S_{t+j/m}) - \ln(S_{t+(j-1)/m}) \tag{1}$$

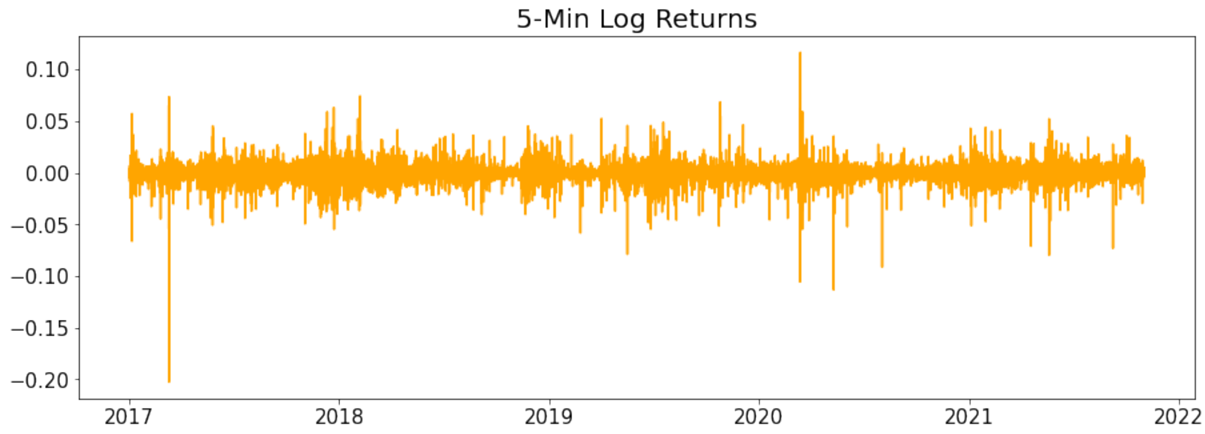

Figure 4, Bitcoin 5-min Log Returns

In a second step, the estimate of the daily variance is calculated following the formula shown in (2), again following the steps and the notation of Christoffersen (2012). The realized variance is the sum of the intraday squared returns, calculated by the formula in (1):

$$\mathrm{RV}_{t+1}^{\mathrm{m}} = \sum_{j=1}^{m} R_{t+j/m}^{2} \tag{2}$$

Equation (2) gives the daily RV for the whole 24-hour period. The time series of the daily RV is plotted in Figure 5. In Figure 6, the graph for the logarithmic form of the realized variance of Bitcoin returns is plotted.

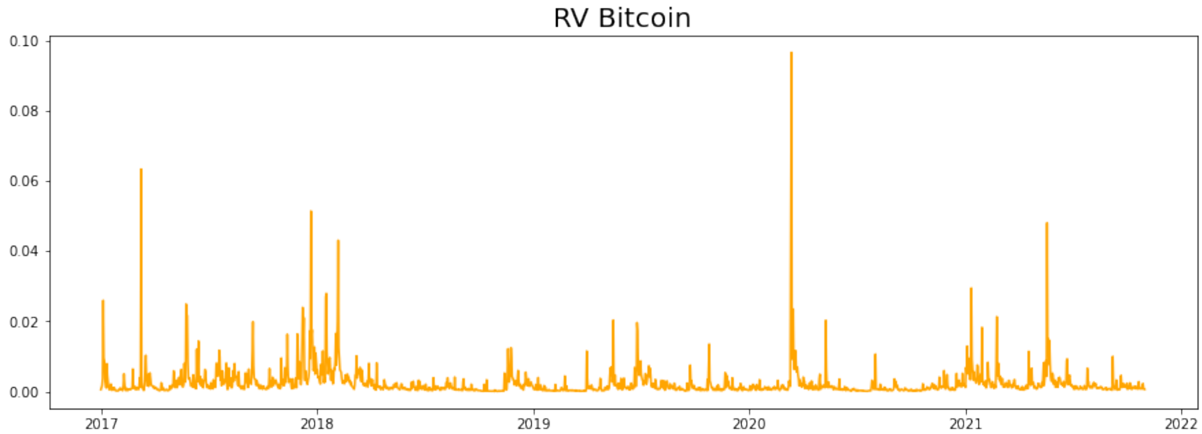

Figure 5, Time series Daily Realized Variance Bitcoin

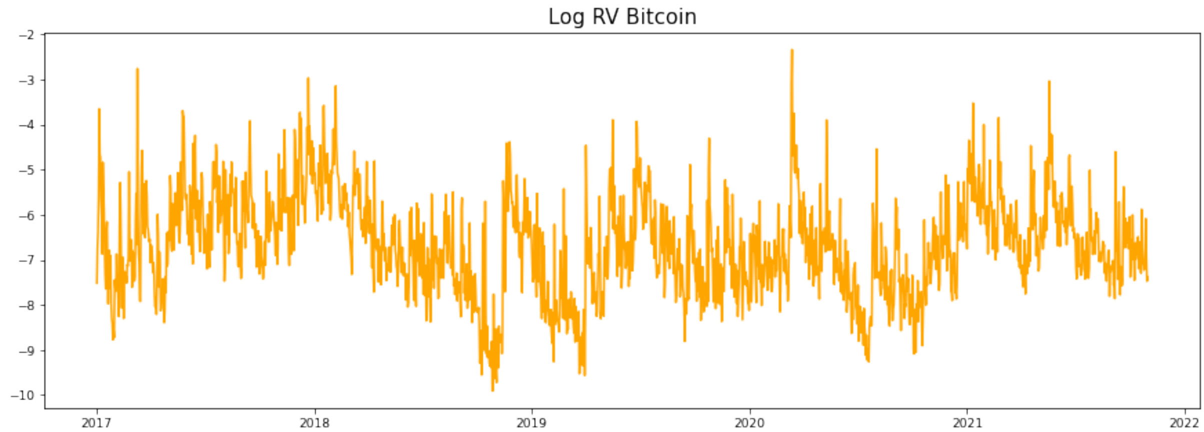

Figure 6, Time series Daily Log Realized Variance Bitcoin

3.1.2) Stylized Facts of RV

In "Elements of Financial Risk Management", Christoffersen (2012) outlines four stylized facts that should be considered when building a suitable model for forecasting realized variance: first, that the RV is superior to daily squared returns and a "more precise indicator" when estimating the daily variance. The second stylized fact of realized variance is that for a big number of lags, it will exhibit a large positive autocorrelation. The third and fourth stylized facts are that the log transformation of the realized variance demonstrates approximately a normal distribution, and that when $\mathrm{R_t}$ (daily return) is divided by the square root of the realized variance, the resulting values will approximately exhibit an independently and identically standard normal distribution. In the following section, these four stylized facts will briefly be investigated to support the grounds that a reasonable model for forecasting RV with this data can be set up.

Figure 5 shows the time series of daily realized variance, calculated from the five-minute intraday squared returns and the squared daily returns plotted for the whole sample size. Comparing these two graphically, indicates that the first of four stylized facts of RV is fulfilled, as squared returns appear to be noisier and more ridged than the RV. When looking at Figure 8, and Figure 9 we can see that the RV and especially the log RV exhibit large autocorrelation and high, therefore satisfying the second stylized fact. Investigating the distribution, it can be observed in Figure 9 that the distribution of log RV appears to be approximately normally distributed, and in Figure 10, we can see that the values of daily return divided by the square root of RV seem to be fairly i.i.d. standard normally distributed. This also shows that the third and fourth stylized facts seem to be fulfilled.

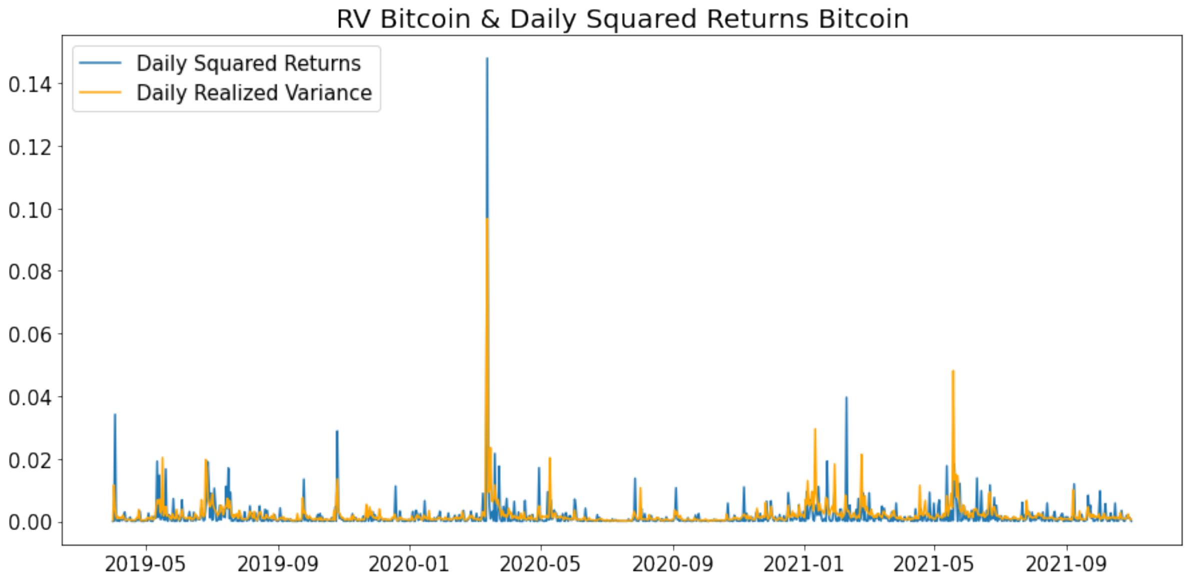

Figure 7, Realized variance Bitcoin & Daily Squared Returns of Bitcoin

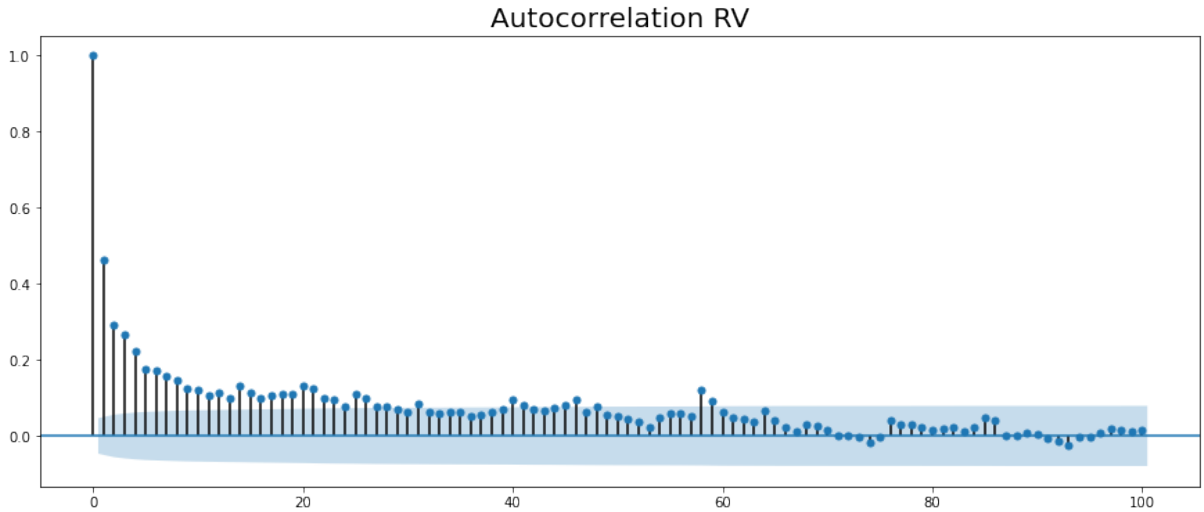

Figure 8, Autocorrelation of Realized Variance

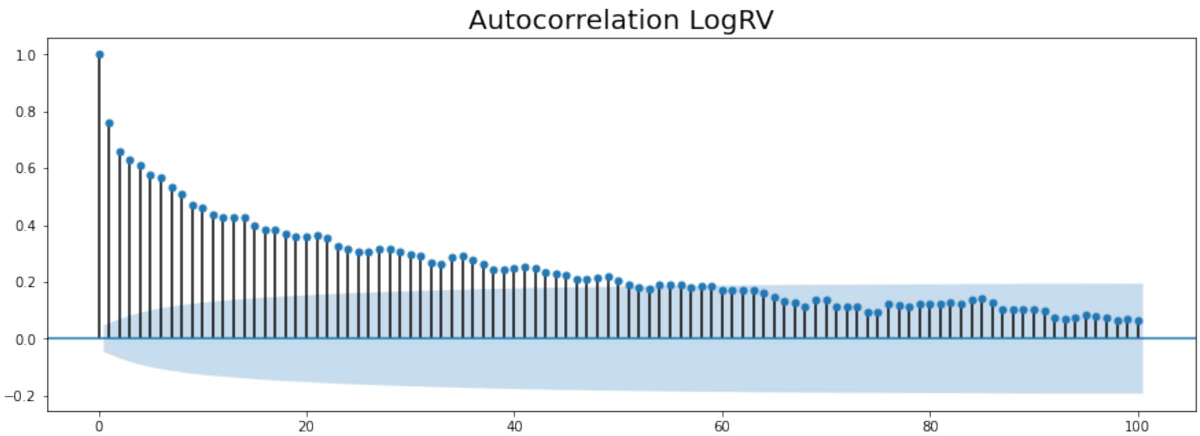

Figure 9, Autocorrelation of Log Realized Variance

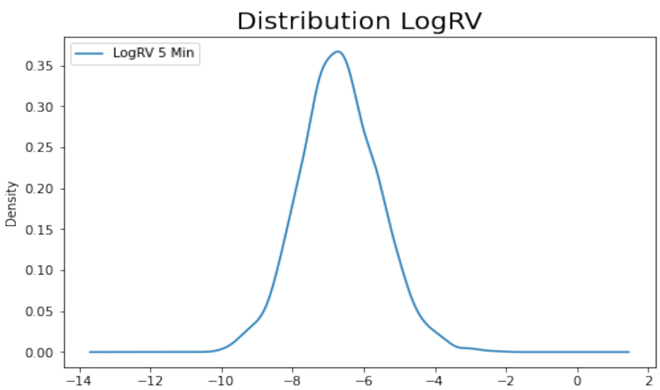

Figure 10, Distribution LogRV

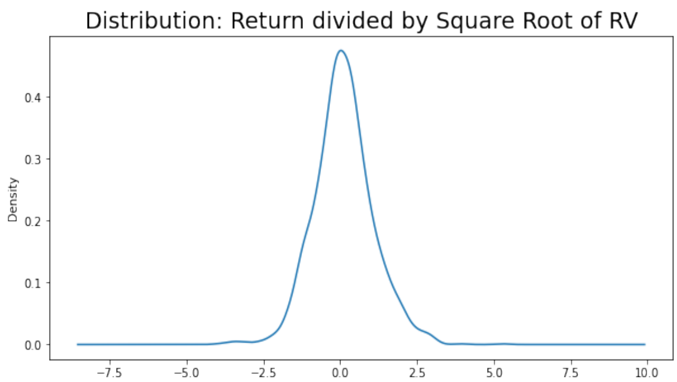

Figure 11, Return scaled by the Square Root of RV

3.2) Google Trends

Google Trends is an online tool provided by the technology company Google, which according to their website, "provides access to a largely unfiltered sample of actual search requests made to Google".[7] According to one of Google's data editors, this data helps to "explore the global reaction to major events" and is "one of the world's largest real time datasets".[8] It is possible to use Google Trends to investigate the trends of Google searches with a specific keyword or broader topic on both a regional and global level. Depending on the period chosen Google Trends supplies data on a weekly, daily, hourly or minute by minute basis. The search data is normalized by Google itself to "make comparisons between terms easier", and the numbers provided range from zero to 100, with higher numbers indicating a greater quantity of Google searches for a certain keyword or topic. [9]

The keyword of interest for this thesis was "Bitcoin", and I retrieved daily Google Trends data for the whole sample period. As there were limitations in downloading the daily Google Trends data directly from the website over such an extensive period, the API "pytrends" in Python was used, which permitted the import and daily use of data from Google trends for the whole sample period.[10]

Kim et al. (2019) suggest in their paper "Google searches and stock market activity: Evidence from Norway" that data of Google Searches provided by Google Trends needs "some standardization relatively to its past", because the values are subject to changes depending on the requested time range. For this analysis, I chose to scale the Google Trend data by its rolling four-weeks-mean, as shown in equation (3),

---

[7] Cf. Google. Available at https://support.google.com/trends/answer/4365533?hl=en (retrieved December 17th, 2022)

[8] Cf. Medium.com. Available at https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8 (retrieved March 11th, 2022)

[9] Google describes the scaling process of Google Trends as follows: "Each data point is divided by the total searches of the geography and time range it represents to compare relative popularity. Otherwise, places with the most search volume would always be ranked highest. The resulting numbers are then scaled on a range of 0 to 100 based on a topic's proportion to all searches on all topics. Different regions that show the same search interest for a term don't always have the same total search volumes."

[10] https://pypi.org/project/pytrends/

which means that the variable is now based on the absolute number provided by Google Trends (which is represented as $GoogleTrend_t$ in (3) on a given day t and is given in relation to the average level of the past four weeks' data). This system of scaling converts the data to positive values around one, with values above one for days with a higher number of Google searches compared to the previous four weeks, and values below one for days when Google searches drop below the four-weeks average.

$$\mathrm{GT}_\mathrm{t}^\mathrm{d} = \frac{GoogleTrend_t}{\frac{1}{28}\sum_{i=1}^{28} GoogleTrend_{t-i}} \tag{3}$$

I want to point out that this standardization is slightly different to methods used in other papers where Google search volume was also used in the analysis of volatility and return of various assets. The transformation used here gives positive values only, whereas the standardization utilized for instance by Bijl et al. (2016) or Kim et al. (2019), for example, can take values that can be both negative and positive.[11] This will make a difference when it comes to the interpretation of the effect of this variable, not only but especially for below-rolling-average values. This approach to standardize this data is chosen as I expect Google Trends to have a positive effect on the realized variance overall, which gets less for below average values, but not negative.

The Google Trends variable is used in this thesis to reflect the attention of the public and especially of retail investors to Bitcoin. For example, a higher number of Google searches for the term Bitcoin can be attributable to certain shocks or news events, affecting the price of Bitcoin. Figure 12 plots the time series of the standardized Google Trends variable together with the price of Bitcoin. It can be gathered that during periods of greater price fluctuation that the Google Trends variable

[11] The standardization in these papers is done by subtracting the rolling average from the actual value and dividing this value by the rolling standard deviation.

experiences a sharp increase, as can be seen at the end of 2018, in the middle of 2019 and during 2021.

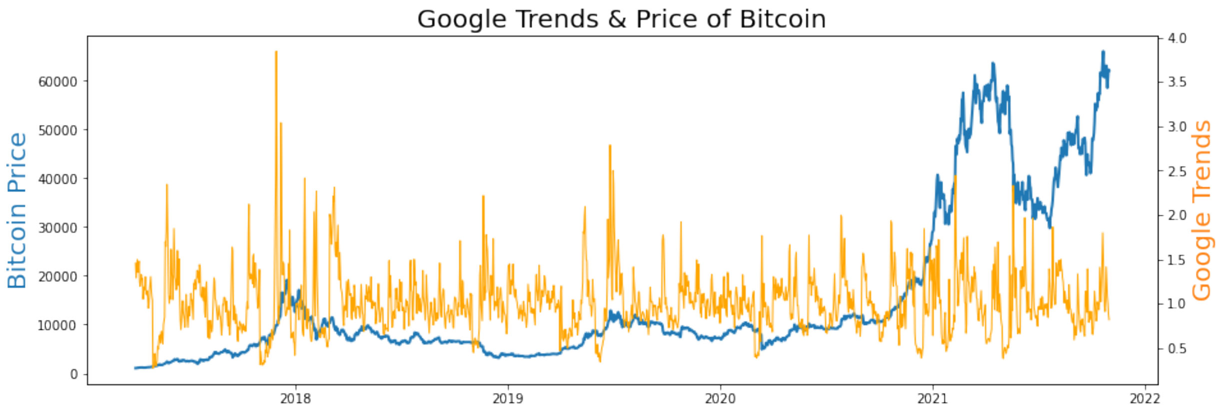

Figure 12, Standardized Google Trends & Bitcoin Prices in US-Dollar

3.3) Tweets

Tweets in this thesis are posts published on the social media network Twitter by Elon Musk, the CEO of Tesla, and various other technology companies like Neuralink or SpaceX and Twitter (now: X) itself now.[12] The Tweets exemplified in this thesis were selected by me by reviewing all of Musk's Tweets in the respective period. They include all postings published by Elon Musk on the topic of cryptocurrencies in this time range and come with an exact time stamp (year, day, hour, and minute). The Tweets also include replies to posts on other people's posts on Twitter. In Figure 3, the distribution of the Tweets over the sample period is plotted, with one line indicating one Tweet, and the blue line showing Bitcoin prices in US-Dollar. This graph demonstrates that most of Musk's posts occur in the final year of the sample and that the Tweets are not evenly distributed over the displayed period. In total, I gathered

[12] I want to note that Musk became CEO of Twitter (now: X) subsequent to the selection of this thesis topic.

128 Tweets about cryptocurrencies during the selected period, with 42 of them pertaining to Bitcoin and 86 regarding other cryptocurrencies.

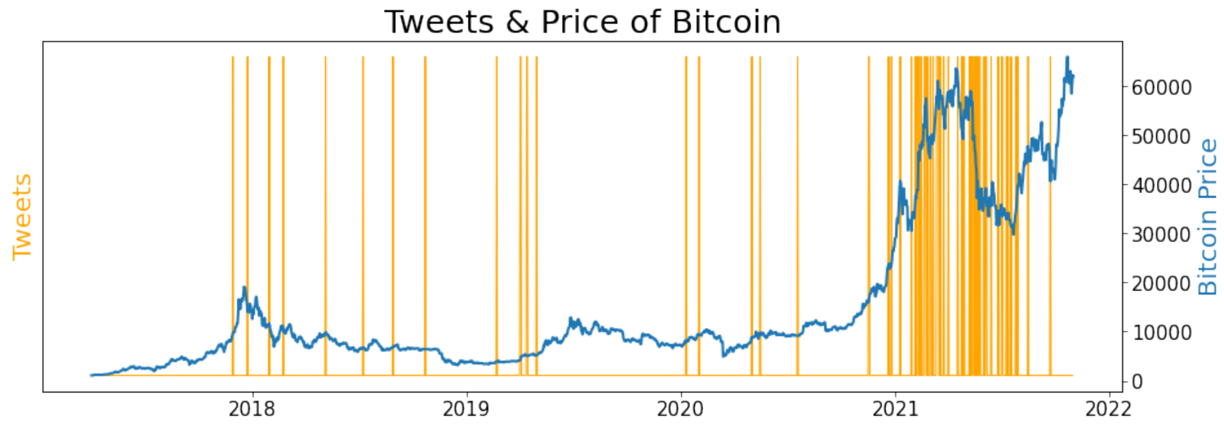

Figure 13, Distribution of Tweets over Sample Period & Bitcoin Prices in USD

It is difficult to summarize the content of the Tweets by Musk, as his posts range from memes about cryptocurrencies to more official and important announcements about the usage of cryptocurrencies for his companies, and others not appearing to make any sense at all. To give an example, in February 2021 Musk posted a “meme” saying that Chuck Norris can withdraw Bitcoins from the (insolvent) Bitcoin exchange Mt. Gox. The following month, March 2021, Musk announces, “You can now buy a Tesla with Bitcoin”, a major announcement for the car manufacturer.

Tweets are included in the estimations as dummy variables, by assigning the value of one if a tweet was posted, and the value zero if there was no Tweet on a given day. If there was more than one Tweet on a given day, only the Tweet with the highest number of interactions (based on retweets) is included in the calculations. During the data gathering process, I indicated whether the Tweets were about Bitcoin specifically or about cryptocurrencies in general (all other cryptocurrencies but Bitcoin) to identify possible variations depending on the content of the posts on Twitter. In addition, I included the “Like” interactions, as well as the “Retweets” in the data set.

I then divided the Tweets according to their number of interactions. As the absolute number of the reach of the Tweets is unknown, I used the number of interactions as a proxy for the popularity and reach.[13] Tweets with lower and higher interactions are divided by the mean number of Retweets of all the Tweets included in the data set. Although the number of Likes and Retweets seem to have a strong positive correlation (with a Pearson correlation coefficient of 0.92), I chose to use Retweets instead of Likes. I expected Retweets to be a slightly better indicator for the reach of a Tweet than Likes, as Tweets directly gain additional reach with each Retweet.

In my calculations, I used following Tweet-related variables: Tweets as a dummy variable, where all the Tweets about Bitcoin and cryptocurrencies in general are included; Tweets with higher and lower interactions; Tweets specifically referring to Bitcoin; and Tweets about cryptocurrencies in general (with both Bitcoin-specific Tweets and Tweets relating to cryptocurrencies in general being separated into higher and lower interactions, based on the number of Retweets). On account of there being multiple Tweets on certain days, only the Tweet with the highest quantity of interactions (based on the number Retweets) is kept.

The complete list of Tweets, including their web links, the number of Likes and Retweets, as well as the classification into Tweets about Bitcoin and Tweets about cryptocurrencies in general can be found in full detail in the appendix.

3.4) Weekend

Weekend is a simple dummy variable, which takes the value zero for weekdays, and the value one for Saturdays and Sundays, to indicate the difference in the dependent variable between weekends and weekdays.

[13] There have been changes on Twitter recently, which allows people to see the total reach of Tweets now. For the period used, this data is not available.

3.5) Stylized Facts and sample characteristics

In Table 1 the stylized facts of each of the variables are shown for the data used for the analysis, meaning that the data for days without complete entries for a given day (because of the construction of the variables using rolling averages like the weekly or monthly RV variables RV^w and RV^w or the Google Trends variable GT^d) were omitted, leaving us with 1706 observations. For the stylized facts for $\mathrm{G}T^d * Tweets_H$ and $GT^d * Tweets_L$ the mean, the standard deviation, etc. is calculated excluding days without zero-values for the Tweet variable. It can be viewed that on the days with Tweets (both for Tweets with above and below average interactions), the average value for Google Trends is higher than the overall mean. In Table 2, the count of the Tweet dummy variables is shown. As only the Tweet with the highest number of interactions is considered each day, the number of Tweets is reduced to 76. It can also be seen that the number of Tweets for the subcategories becomes quite small, especially for $\mathrm{BTCTweets}_H$, with only five observations. Therefore, the corresponding results should be interpreted with some caution. More details on the construction of the various Tweets and RV variables are given in subchapters 4.1 and 4.2.

	$\boldsymbol{RV^d}$	$\boldsymbol{RV^w}$	$\boldsymbol{RV^m}$	$\boldsymbol{GT^d}$	$\mathbf{GT}^d * \mathbf{Tweets}_H$	$\mathbf{GT}^d * \mathbf{Tweets}_L$
Count	1706	1706	1706	1706	27	49
Mean	-6.677475	-6.676433	-6.678396	1.012178	1.066961	1.106679
Std	1.105202	0.934936	0.809429	0.338870	0.454008	0.339123
Min	-9.907971	-9.401186	-9.053657	0.272854	0.438393	0.497462
25%	-7.398821	-7.242595	-7.168240	0.813574	0.813778	0.913540
50%	-6.723153	-6.737809	-6.732250	0.968029	0.985507	1.055072
75%	-5.949976	-6.054742	-6.026580	1.158870	1.248535	1.283627
max	-2.336395	-4.066262	-4.830267	3.846154	2.333333	2.423493

Table 1: Stylized Facts

	$\mathbf{Tweets}$	$\mathbf{Tweets}_H$	$\mathbf{Tweets}_L$	$\mathbf{BTCTweets}_H$	$\mathbf{BTCTweets}_L$	$\mathbf{CryptoTweets}_H$	$\mathbf{CryptoTweets}_L$
Count	76	27	49	5	19	22	30

Table 2: Count of Tweet Dummy Variables

4.) Volatility Model

For my analysis, a heterogeneous autoregressive model of realized volatility (HAR-RV) is used. To assess the effect of Tweets on volatility, this model is augmented by other regressors. HAR-RV models are, as already mentioned, a powerful yet simple approach to forecast volatility, which can be easily expanded through the inclusion of other variables. This chapter begins by discussing and explaining the HAR-RV model developed by Corsi (2009), followed by a closer observation of the models and their individual expansions.

4.1) HAR-RV Model

The HAR-RV model is a simple linear regression model that can incorporate the persistence in volatility observed in financial data, as well as other important stylized facts, like lognormality and long-memory (Corsi, 2009). Features, that can be also observed for the Bitcoin-RV data used for this analysis, as shown in section 3.1. The HAR-RV model proposed by Corsi (2009) uses volatility cascades, leading to an autoregressive-type RV-model. According to Corsi (2009) this cascade-type structure of the RV-model is motivated by the heterogeneous market hypothesis, which recognizes the heterogeneous nature of traders. This heterogeneity can be explained by the differences that exist between traders in financial markets, such as variations in time horizons, distinctions within risk profiles, or disparities in access to relevant information. For Bitcoin, and the cryptocurrency market in general that would correspond to traders having different time horizons. On the one hand, there are intraday traders who buy and sell their positions multiple times within a day and try to profit on short term swings in prices. On the other hand, there are institutional traders who trade cryptocurrency with a longer time horizon, as it is the case for example with the cryptocurrency indices developed by the trading platform Bitpanda, which are rebalanced monthly.[14] As previously illustrated, the persistence and long-memory feature of Bitcoin volatility can be viewed in Figure 9 for realized variance, and especially in Figure 10 for its log-transformation by the plots of the corresponding autocorrelation functions.

Motivated by these volatility patterns and following Corsi (2009), the weekly and monthly realized variance are constructed as the average of the daily RV-components within these periods. The weekly realized variance is based on the seven-day (instead of five-day) average, and the monthly realized variance is based on the 28-day or four-week (instead of 21-day) average, since Bitcoin is traded - in contrast to more traditional financial assets like stocks - 24 hours a day, seven days a week, without any breaks on weekends or bank holidays. Equation (4) demonstrates how the weekly volatility component is constructed, where w = 7, and the monthly volatility component follows equation (5) where m = 28:

$$\mathrm{RV}_t^w = \frac{1}{w}\sum_{i=0}^{w} RV_{t-i} \tag{4}$$

$$\mathrm{RV}_t^m = \frac{1}{m}\sum_{i=0}^{m} RV_{t-i} \tag{5}$$

In combination with the parameter for daily variance RV^d, it is now possible to construct a simple HAR-RV regression model to forecast RV_t^d as shown in equation

[14] Cf. “The BCI25 consists of the top 25 cryptocurrencies […]. Every month, your portfolio will be rebalanced”. Available at https://www.bitpanda.com/en/prices/crypto-index/bci25. (Retrieved October 10th, 2022)

(8) with the three lagged volatility components. This model can be labelled as a HAR(3)-RV model and has an autoregressive structure that is able to capture the RV over various interval sizes (Corsi, 2009). As suggested by Christoffersen (2012), Lyócsa (2020) and others, I use the log transformation of the RV shown in equation (9) below because of the log normal property discussed earlier, therefore taking advantage of a more precise estimation of the parameters when using ordinary least squares (OLS). The weekly and monthly log-transformed volatility components are calculated as shown in equation (6) and equation (7).

$$\ln(\mathrm{RV}_{\mathrm{t}}^{\mathrm{w}}) = \frac{1}{w}\sum_{\mathrm{i}=0}^{\mathrm{w}} \ln(RV_{\mathrm{t-i}}) \tag{6}$$

$$\ln(\mathrm{RV}_{\mathrm{t}}^{\mathrm{m}}) = \frac{1}{m}\sum_{\mathrm{i}=0}^{\mathrm{m}} \ln(RV_{\mathrm{t-i}}) \tag{7}$$

$$\mathrm{RV}_t^d = \beta_0 + \beta_1 \mathrm{R}V_{t-1}^d + \beta_2 \mathrm{R}V_{t-1}^w + \beta_3 \mathrm{R}V_{t-1}^m + \epsilon_t \tag{8}$$

$$\ln\left(RV_t^d\right) = \beta_0 + \beta_1 \ln\left(\mathrm{R}V_{t-1}^d\right) + \beta_2 \ln(\mathrm{R}V_{t-1}^w) + \beta_3 \ln(\mathrm{R}V_{t-1}^m) + \epsilon_t \tag{9}$$

Model 1

4.2) Augmented HAR-RV Models

As previously discussed, this simple OLS-framework is open to augmentations with additional regressors. In this section, the different augmentations of the models used for my analysis to assess the effects of Musk’s Tweets on realized volatility are described.

Before estimating possible effects of Tweets on realized volatility, the HAR-RV model is expanded by adding two variables to the base HAR-RV model: The dummy variable, “weekend”, and the variable, “Google Trends”. In the following, I will justify why I chose to include these two variables.

By adding the dummy variable, “Weekend”, I want to control for a possible difference in volatility between weekdays and weekends. I expect there to be a difference based on a potential change in investors’ behaviour or activity between weekdays and weekends. Additionally, stock markets are typically not open on weekends, and possible spill-over effects between the stock market and the cryptocurrency market, as investigated by Iyer (2022), might play a role. This is also supported by the analysis of Wang et al. (2020), who find that return and volatility of Bitcoin is linked to the opening hours of European and US stock exchanges, with both variables being considerably higher during the week than on weekends. Similar evidence is presented in an online article on the cryptocurrency information platform Coindesk.com, in which the author Genç refers to an analysis undertaken by the crypto options analytics platform, Genesis Volatility, that indicates that there is less realized volatility on Saturdays and Sundays.[15]

By adding the variable daily Google Trends (denoted as GT_t^d) to the model I try to rule out spurious relations between Tweets of Elon Musk and realized volatility. For example, it could be the case that an already higher level of Bitcoin volatility draws more attention to Bitcoin, meaning that people are more likely to share their thoughts and comment about it on Twitter. To avoid neglecting the possibility that Musk may be more likely to publish posts on Twitter on days when there is news or shocks affecting Bitcoin, I added Google trends to the model as an independent variable. In doing so, I wanted to evidence changes in investors’ attention to Bitcoin separately from Musk’s Tweets, such as the impact on traders’ interest in the cryptocurrency because of news coverage, for example. The relationship between the attention of investors and internet search volume has been investigated by Da et al. (2011), who commend Google’s search volume tool as being a “a new and direct measure of

---

[15] Cf. Genç, E. on Coindesk. Available at https://www.coindesk.com/learn/is-there-a-best-time-to-trade-crypto-heres-what-the-data-says/ (retrieved April 22nd, 2023)

investor attention”, stressing the practicability of the use of this kind of data in finance. Additionally, the effectiveness of utilizing Google Trends when forecasting volatility has been demonstrated. For example, Dimpfl and Jank (2011) find that using search queries improves their volatility forecasts for four main stock market indices, and that volatility reacts positively to higher attention of investors.

After adding these two variables to the HAR-RV Model, the new augmented HAR-RV followed the equation shown in (10):

$$\ln\left(RV_t^d\right) = \beta_0 + \beta_1 \ln\left(\mathrm{R}V_{t-1}^d\right) + \beta_2 \ln(\mathrm{R}V_{t-1}^w) + \beta_3 \ln(\mathrm{R}V_{t-1}^m) + \beta_4 \mathrm{Weekend} + \beta_5 \mathrm{GT}_{\mathrm{t-1}}^{\mathrm{d}} + \epsilon_{\mathrm{t}} \tag{10}$$

Model 2

Model 2 was used as the benchmark model for the models containing Tweets by Elon Musk. So, the models with Tweets were compared to this model without the Tweets variable or variables. Having thus established a benchmark model for forecasting Bitcoin volatility, the variable of interest, Tweets, was then added to the model. The first approach was to simply include all Tweets as a dummy variable, as shown in (11).

$$\ln\left(RV_t^d\right) = \beta_0 + \beta_1 \ln\left(\mathrm{R}V_{t-1}^d\right) + \beta_2 \ln(\mathrm{R}V_{t-1}^w) + \beta_3 \ln(\mathrm{R}V_{t-1}^m) + \beta_4 \mathrm{Weekend} + \beta_5 \mathrm{GT}_{\mathrm{t-1}}^{\mathrm{d}} + \beta_6 \mathrm{Tweets}_{\mathrm{t-1}} + \epsilon_{\mathrm{t}} \tag{11}$$

Model 3

As there could have been a difference in the effect of Tweets on the volatility of Bitcoin depending on the number of interactions (for higher interactions, I expect Tweets to have a more noticeable impact) as described in 3.3, instead of a single dummy variable for all Tweets by Musk, the variable $\mathrm{Tweets}_{\mathrm{H}}$ and the variable $\mathrm{Tweets}_{\mathrm{L}}$ denoting Tweets with a higher (subscript “H”) and lower (subscript “L”) number of interactions was included. The regression formula for this model (referred to as Model 4 from now on), expanded by these two variables, is shown below (12).

$$\ln\left(RV_t^d\right) = \beta_0 + \beta_1 \ln\left(RV_{t-1}^d\right) + \beta_2 \ln(RV_{t-1}^w) + \beta_3 \ln(RV_{t-1}^m) + \beta_4 \mathrm{Weekend} + \beta_5 GT_{t-1}^d + \beta_6 \mathrm{Tweets}_{H,t-1} + \beta_7 \mathrm{Tweets}_{L,t-1} + \epsilon_t \tag{12}$$

Model 4

The independent variables Tweets as well as GT were obviously observed for the same day. Therefore, the possibility of unobserved interactions should be considered, due to there being the chance of an increased number of Google searches as a result of posts made by Musk on Twitter. To check for possible interactions between Tweets and the number of search queries for Bitcoin on that particular day, I include an interaction variable between Google Trends and Tweets ($GT_t^d * Tweets_{H,t}$ and $GT_t^d * Tweets_{L,t}$). This possible interaction can be captured by estimating the coefficients for the model (Model 5), as shown in equation (13).

$$\ln\left(RV_t^d\right) = \beta_0 + \beta_1 \ln\left(RV_{t-1}^d\right) + \beta_2 \ln(RV_{t-1}^w) + \beta_3 \ln(RV_{t-1}^m) + \beta_4 \mathrm{Weekend} + \beta_5 \mathrm{GT}_{t-1}^d + \beta_6 \mathrm{Tweets}_{H,t-1} + \beta_7 \mathrm{Tweets}_{L,t-1} + \beta_8 \mathrm{GT}_{t-1}^d * \mathrm{Tweets}_{H,t-1} + \beta_9 \mathrm{GT}_{t-1}^d * \mathrm{Tweets}_{L,t-1} + \epsilon_t \tag{13}$$

Model 5

The variable Tweets is a dummy variable for all Tweets by Elon Musk and does not differentiate between Bitcoin and other cryptocurrencies. Model 6, which is shown in equation (14), was used to investigate whether there was a possible difference in the effect on volatility of Bitcoin between Tweets about Bitcoin specifically and Tweets about cryptocurrencies in general. The dummy variable $\mathrm{BTCTweets}$ was used for Musk's Tweets about Bitcoin, and the dummy variable $\mathrm{CryptoTweets}$ denotes Tweets relating to all other cryptocurrencies. To keep the number of independent variables lower, instead of using four interaction variables, I continued to use the interaction variables introduced above: An interaction between Google Trends and Tweets with higher interactions (both $\mathrm{BTCTweets}_H$ and $\mathrm{CryptoTweets}_H$) and the same but for

Tweets with lower interactions (both BTCTweets_L and CryptoTweets_L). The threshold for high and low interactions is the same for all independent variables.

$$\begin{aligned}\ln\left(RV_t^d\right) = {} & \beta_0 + \beta_1 \ln\left(\text{RV}_{t-1}^d\right) + \beta_2 \ln(\text{RV}_{t-1}^w) + \beta_3 \ln(\text{RV}_{t-1}^m) + \beta_4 \text{Weekend} \\ & + \beta_5 \text{GT}_{t-1}^d + \beta_6 \text{BTCTweets}_{H,t-1} + \beta_7 \text{BTCTweets}_{L,t-1} \\ & + \beta_8 \text{CryptoTweets}_{H,t-1} + \beta_9 \text{CryptoTweets}_{L,t-1} + \beta_{10} \text{G}T_{t-1}^d \\ & * \text{Tweets}_{H,t-1} + \beta_{11} \text{GT}_{t-1}^d * \text{Tweets}_{L,t-1} + \epsilon_t \end{aligned} \tag{14}$$

Model 6

5.) Model Comparison

To compare models and to be able to form a valid conclusion about the effect of Tweets on volatility, an F-Test and loss functions were used. This chapter gives an overview about the procedure and application.

5.1) F-Test

The F-Test for a regression analysis makes it possible to compare competing regression models. This F-Test, also referred to as an Overall-F-Test, compares a restricted model to an unrestricted model, testing for joint significance of the variables added to the restricted model. The null hypothesis states that the restricted and unrestricted model are equal, and the alternative hypothesis states that the unrestricted model has a better fit. By observing the p-value of the F-Test, we can check if the null hypothesis can be rejected. The result can be used for model selection and helps to assess whether to keep additional regressors in the regression model or not. The formula for the Overall-F-Test for joint significance follows equation (15), where SSR_1 is the sum of squared residuals for the restricted and SSR_2 is the sum of squared residuals for the unrestricted model. The letter q corresponds to the number of variables that are included in the unrestricted (but not in the restricted model), n is the sample size, and p is the number of regressors in the unrestricted model.

$$\mathrm{F} = \frac{\frac{SSR_1 - SSR_2}{q}}{\frac{SSR_2}{n-p-1}} \tag{15}$$

5.2) Loss-Functions

Loss functions are an important aspect of evaluating volatility models, especially when assessing the forecasting accuracy. They measure the deviation between actual and predicted values. Different loss functions weigh the deviations from the actual values differently, for example, some measure the deviation regardless the magnitude of the difference whereas some punish bigger outliers more than smaller ones. Comparing the loss functions of different models helps to assess the accuracy and effectiveness of models and forecasts. Below, the loss functions used for model comparison in this thesis are briefly outlined:

The mean absolute error (MAE) was calculated following equation (16) and is the sum of the absolute difference between the actual (y_i) and predicted values $(\widehat{y_i})$, scaled by the sample size.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \widehat{y_i}| \tag{16}$$

The mean absolute percentage error (MAPE) is calculated similarly to the MAE, with the formula following equation (17). Like with the previous measure, the absolute difference between actual and predicted values is calculated, with the difference that the MAPE gives the results in terms of percentage, making the result independent to the size of the variables.

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} * 100 \tag{17}$$

The mean squared error (MSE) is calculated following Equation (18), and is again similar to the MAE. The key distinction between MAE and MSE is that the differences between the actual and predicted values are squared, “punishing” higher differences our outliers more than smaller errors.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{18}$$

Another loss function that was used to evaluate the deviations of predicted versus actual values is QLIKE, which can be calculated using Equation (19) (Christoffersen, 2012). In contrast to the previous loss functions, where the loss function uses the differences between actual and predicted values, QLIKE relies on the relative error $\frac{\hat{y}i}{y_i}$ between predicted and actual values and punishes underestimation of volatility more strongly (Christoffersen, 2012).

$$QLIKE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}i}{y_i} - \log\left(\frac{\hat{y}i}{y_i}\right) - 1\right) \tag{19}$$

6.) Results

In this chapter, the estimation results will be discussed and interpreted. First, the in-sample modelling results will be presented, followed by an analysis of the OLS assumptions. To investigate these results further, an Overall F-test will be applied and the loss functions of both the in-sample and rolling window regression will be reported. The chapter is concluded by an interpretation of the results.

6.1) In-sample modelling Results

To investigate the effect of each of the independent variables on the realized variance, all the models described in the previous chapter were fitted with the full data sample. As heteroskedasticity often poses a problem in time series linear regressions, all estimations were performed using heteroscedasticity robust standard errors.

In Table 3, the OLS-regression results of the standard HAR-RV model (Model 1), together with the first augmented HAR-RV model (Model 2) and Model 3, are reported. In Table 4, the OLS-regression results of Model 4, Model 5 and Model 6 are shown. Both tables also provide the number of observations together with the values for R-Squared. By examining Table 3 and the standard HAR-RV model (1), it can be observed that the three lagged RV components show a positive relationship with RV_t^d, that is decaying in its lags. The first lagged RV variable shows the strongest positive relationship, followed by the lagged weekly RV, with both being highly significant. The coefficient of the monthly RV-component is smaller in comparison to the daily and weekly RV and is only significant at the 5%-level.

	Dependent Variable: RV_t^d		
	(1)	(2)	(3)
const	-0.582***	-0.674***	-0.696***
	(0.146)	(0.158)	(0.158)
RV_{t-1}^d	0.479***	0.445***	0.445***
	(0.030)	(0.034)	(0.034)
RV_{t-1}^w	0.355***	0.378***	0.378***
	(0.043)	(0.044)	(0.044)
RV_{t-1}^m	0.078**	0.084**	0.081**
	(0.039)	(0.039)	(0.039)
GT_{t-1}^d		0.110*	0.106*
		(0.060)	(0.059)
Weekend		-0.197***	-0.199***
		(0.034)	(0.034)
Tweets_{t-1}			0.1458*
			(0.087)
Observations	1,706	1,706	1,706
R^2	0.623	0.631	0.631
Note:		*p<0.1; **p<0.05; ***p<0.01	

Table 3, Regression Output for Model 1 – Model 3

The estimation results of Model 2 are shown in the second column of Table 3. The coefficients for Google Trends and the variable Weekend are estimated. It can be viewed that the coefficients for the RV components in the model show similar effects on the dependent variable as before and only change very little. The constant of the model does change in comparison to the previous model when Google Trends is added to the model, as Google Trends has an average value around one. Google Trends exhibit a positive effect on the next day's realized variance but is only significant at a 10% level. When observing the variable Weekend, it can be seen that the model indicates that there is less realized variance on weekends than on weekdays with its coefficient of -0.197. This variable is also highly significant. In Model 3, Tweets, is added to the previous model. The results are reported in the third column in Table 3 where it is possible to see that the variable Tweets_{t-1} has a positive coefficient, which indicates a positive effect on next day's Bitcoin RV. However, it can be viewed that this variable is only weakly significant (with a p-value of 0.094). The Tweet dummy variable, as previously described, includes all Tweets about Bitcoin and other cryptocurrencies. The Tweet variable exhibits a weakly positive significant effect, with a coefficient of 0.146. Looking at the other coefficients, it is visible that they do not change much and maintain approximately the same magnitude.

In Table 4, the results of the regressions of Model 4, Model 5 and Model 6 are reported in the respective 4th, 5th and 6th column. The model can now control for a possible difference in the effect of $Tweets_{t-1}$ on next day's realized variance, based on the number of the Tweet's interactions. Splitting the variable Tweets in two sub-groups – one with above the Tweets' mean interactions and one below Tweets' mean interactions, it is possible to see that the results of Model 4 display a positive coefficient of 0.354 for Tweets with higher interactions being significant at the 5%-level. This coefficient is higher than the coefficient for all Tweets in the previous model. The dummy variable for Tweets with a lower number of interactions did not show a significant effect.

To account for possible interactions between Google searches on days when a post by Elon Musk on Twitter about Bitcoin or cryptocurrencies was published, the model is expanded by adding interaction variables between these two regressors for both high-interactions and low-interaction Tweets. The results are reported in column 5 in Table 4. The changes in the variables' coefficients are noteworthy for the Tweet regressors. Still, only the $\mathrm{Tweets}_{H,t-1}$ is significant and its coefficient more than doubles to 0.9830. The interaction variable of $\mathrm{Tweets}_{H,t-1}$ with GT^{d}_{t-1} is significant as well and has a negative coefficient, that counteracts the strong positive effect of the above-mean interaction tweets. The combined effect of $\mathrm{Tweets}_{H,t-1}$ and its interaction term is 0.355 and has a similar magnitude to the coefficient of $\mathrm{Tweets}_{H,t-1}$ in the previous model (0.354).[16] Looking at the coefficients for the other explanatory variables, it can be seen that the lagged realized variance components hardly change. Furthermore, $\mathrm{Tweets}_{L,t-1}$, the variable for Tweets with below-average interactions, still did not exhibit a significant effect and neither its interaction variable.

[16] Calculation: 0.9830 + 1.06679*(-0.5889) = 0.35467, where 1.06679 is the average Google trends value for days with high-interaction tweets

	Dependent Variable: RV_t^d		
	(4)	(5)	(6)
const	-0.711***	-0.712***	-0.712***
	(0.159)	(0.158)	(0.158)
RV_{t-1}^d	0.444***	0.446***	0.446***
	(0.034)	(0.033)	(0.034)
RV_{t-1}^w	0.382***	0.382***	0.382***
	(0.044)	(0.044)	(0.044)
RV_{t-1}^m	0.077**	0.074*	0.074*
	(0.039)	(0.039)	(0.039)
Weekend	-0.196***	-0.196***	-0.196***
	(0.034)	(0.034)	(0.034)
GT_{t-1}^d	0.108*	0.106*	0.106*
	(0.060)	(0.058)	(0.058)
$\text{Tweets}_{H,t-1}$	0.354**	0.983***	
	(0.176)	(0.352)	
$\text{Tweets}_{L,t-1}$	0.031	-0,576	
	(0.088)	(0.446)	
$\text{BTCTweets}_{H,t-1}$			1.087***
			(0.366)
$\text{BTCTweets}_{L,t-1}$			-0.595
			(0.444)
$\text{CryptoTweets}_{H,t-1}$			0.983***
			(0.348)
$\text{CryptoTweets}_{L,t-1}$			-0.572
			(0.453)
$GT_{t-1}^d * \text{Tweets}_{H,t-1}$		-0.589**	-0.607***
		(0.236)	(0.227)
$GT_{t-1}^d * \text{Tweets}_{L,t-1}$		0.548	0.553
		(0.423)	(0.421)
Observations	1,706	1,706	1,706
R^2	0.632	0.634	0.634
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 4, Regression Output for Model 4 – Model 6

In column six of Table 4, the results of Model 6 are reported. For the estimation of this model, the variable Tweets was split up further into Tweets made by Musk about the cryptocurrency Bitcoin and Tweets regarding all other cryptocurrencies but Bitcoin. In doing so, the possible differences in the effect of the content of these Tweets on the realized volatility could be investigated. Both groups continued to be divided by their interactions like in the previous model, splitting up Tweets into a total of four dummy variables. The estimation results further demonstrate that Tweets with above-mean interactions do show a significant positive effect on next day's realized volatility, regardless of their content. The coefficient for Tweets specifically about Bitcoin is slightly higher than the coefficient for the Tweets variable regarding all other cryptocurrencies. Moreover, the interaction variable is still significant. In line with the estimation results of the previous models, Tweets with lower interactions still did not display any significant effect on the next day's RV. The average total combined effect of $\text{CryptoTweets}_{H,t-1}$ is 0.3355 and the average total combined effect of $\text{BTCTweets}_{H,t-1}$ is 0.4394, indicating that Tweets about Bitcoin exhibit a slightly stronger effect.

Looking at R-squared, a statistical measure that describes the proportion of the variance of the dependent variable that is explained by the independent variables' variance, it can be concluded that Model 5 and Model 6 perform best among these six models according to this measure.

6.2) Further Analysis

In this sub-chapter, the assumptions of this time series regression are checked first to proceed confidently with testing and the interpretation of the results obtained. To further investigate the effect of Tweets of Elon Musk on RV, an F-Test, loss functions, and a rolling-window-estimation are used. I will then compare the results of the loss functions for the whole in-sample modelling results, followed by a comparison for the rolling-window regression. Therefore, a training period of 730 days (or two years) is set, and that window is moved day by day to predict the next day's log realized variance. So, each daily prediction of the log realized variance is based on the fitted model to the preceding 2-years' worth of data.

6.2.1) OLS-Assumptions

To investigate the properties of the OLS estimators of this time series regression, the classical OLS assumptions needed to be checked. Following Wooldridge (2013), the assumptions for time series OLS estimators (based on the classical assumptions) are linearity in parameters (TS-1), no perfect collinearity (TS-2), zero conditional mean (TS-3), homoskedasticity (TS-4) as well as no serial correlation (TS-5) and normality (TS-6). When the first three assumptions hold, the OLS estimators are said to unbiased, and under TS-1 to TS-5, the Gauss-Markov conditions for OLS to be BLUE (Best Linear Unbiased Estimators) are being fulfilled (Wooldridge, 2013). If TS-6 would hold as well, “exact statistical inference” could be drawn from the estimations, regardless the size of the sample (Wooldridge, 2013).

When dealing with time series, some of these assumptions tend to be violated, especially when using autoregressive models. For example, the third (TS-3) assumption cannot hold when the model includes lagged dependent variables, as it is the case in my HAR-RV models (Wooldridge, 2013). As this violation of TS-3 alone would cause the estimators to be biased, it was necessary to turn to the asymptotic properties of OLS, utilising the properties of large samples. These assumptions are based on the assumptions exhibited briefly before but are weaker. The assumptions were tested on all the six models, plots shown here are based on Model 6.

The first asymptotic assumption requires not only linearity in parameters but also stationarity and weak dependence of the time series of the data used (Wooldridge, 2013). To investigate the linearity requirement, scatter plots between the dependent variable and the independent variables were checked, and I could conclude that this requirement is fulfilled. By applying an Augmented Dickey-Fuller-Test, the null hypothesis of a presence of a unit root in these time-series can be rejected for all the time series data used in the estimations (see Table 7 in the appendix). Therefore, the alternative, that these samples are stationary, can be accepted. When two observations of a time series tend to become less correlated the further the distance in time between them, the time series can be characterized as weakly dependent. Due to the autoregressive form of the models with the decreasing effect of the lagged regressors on the dependent variable as well as the decaying autocorrelation of Log-RV, I consider the requirement for weak dependence to be fulfilled. Thus, the first asymptotic assumption holds. The second asymptotic assumption of no perfect collinearity could be tested using the VIF or Variance Inflation Factor (Wooldridge, 2013). The VIF provides a measure of multicollinearity of the regressors used and in the case for the estimations used in this thesis, either no or only moderate multicollinearity was detected.

The third assumption of zero conditional mean requires "contemporaneous exogeneity" of the regressors, meaning that the expected values of the residuals conditional on the regressors in one period is zero (Wooldridge, 2013). To investigate whether this assumption holds, I looked at the covariance between the regressors and the residuals. The covariance between the independent variables and the error term is quite close to zero (values for the different models are between -0.082 and 0.002), and therefore I can conclude that the residuals appear to be uncorrelated with the regressors.

The fourth assumption is the one of contemporaneously homoscedasticity (Wooldridge, 2013). This assumption is weaker because conditional on the independent variables, the variance of the error term must be constant only at their respective points in time and not for the whole sample (Wooldridge, 2013). One way to detect homoscedasticity is by a visual inspection. Analysing the scatter plots of the residuals, I cannot detect a pattern in the residuals, that would be evidence for heteroscedasticity in general. For both plots (Figure 14 and Figure 15), the distribution of the residuals appears to be constant among both time (Figure 15) and predicted values (Figure 14), apart from a small number of positive outliers.

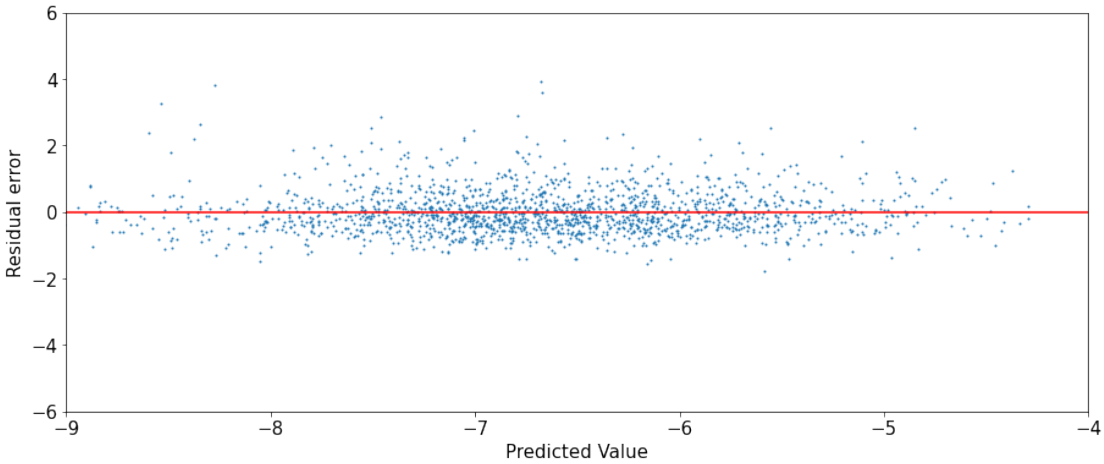

Figure 14, Residuals plotted against the Predicted Values (Model 6)

Another way to test for heteroscedasticity is the Goldfeld-Quandt-test. This test splits the sample into two parts and compares the variance of the residuals. Applying this test, I cannot reject the null hypothesis of presence of homoscedasticity.

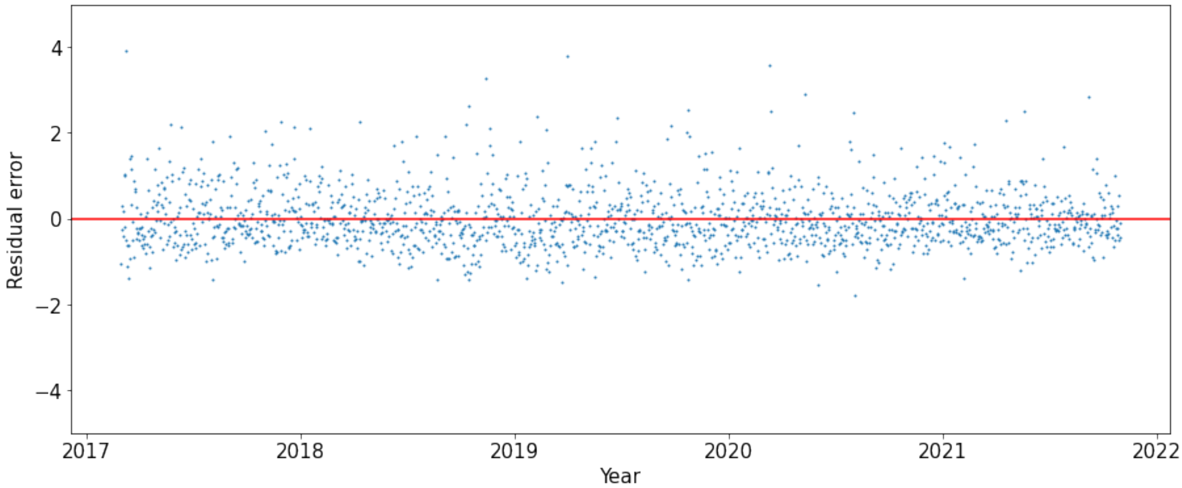

Figure 15, Residuals plotted against Time (Model 6)

The fifth assumption requires no correlation of the error term conditional on the independent variables in the same period (Wooldridge, 2013). The Durbin Watson statistic makes it possible to measure whether autocorrelation within in the residuals is present. A value of 2 would indicate that there is no autocorrelation, and a value above or below would demonstrate a negative or positive autocorrelation. The test statistic is part of the regression output tables displayed in Python, and I can conclude that for my models, this assumption holds as well. The test statistic takes values between 1.988 and 2.005. Furthermore, the autocorrelation function (see Figure 16) of the error term hardly shows any significant autocorrelation.

Under these five asymptotic assumptions for OLS in time series, the estimators are "asymptotically normally distributed" and the "usual OLS standard errors, t statistics, F statistics, and LM statistics are asymptotically valid" (Wooldridge, 2013). The classic assumptions for multiple linear regressions do not hold by the construction of the HAR-RV models used in this thesis, and therefore it cannot be demonstrated that OLS estimators are unbiased or BLUE.

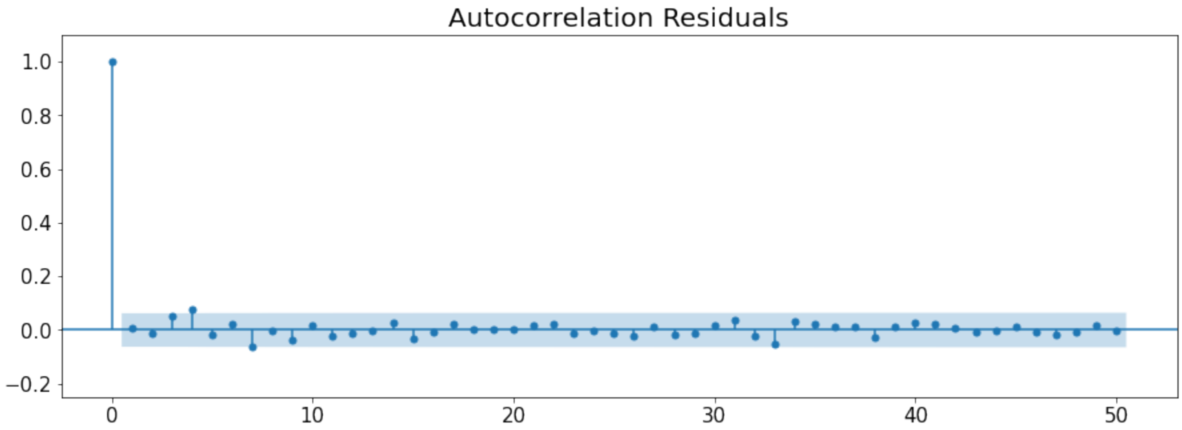

Figure 16, Autocorrelation function of Residuals with confidence interval (alpha = 0.01)

However, by checking these asymptotic assumptions, it can be shown, that the OLS estimators are consistent and "the usual inference procedures are valid" under the previously stated five assumptions (Wooldridge, 2013). Consistent estimators get more precise with an increasing sample size.

6.2.2) F-Test

Using an Overall-F-Test, it can be investigated whether adding the different variations of the Tweets variable to the base model significantly improves the regression model using the in-sample modelling results. The test is performed to compare all the models in which Tweets variables are included (Model 3, Model 4, Model 5 and Model 6) to the base model (Model 2) – the HAR-RV model augmented with the Weekend and Google Trends variables. The test results show that the variables Tweets and Tweets_H as well as Tweets_L are (jointly) significant when added to model 2 - at a 10 per cent level for Model 3 and at a 5 per cent level for Model 4. When expanding the base model with the Tweets variables together with the interaction variables (in Model 5 and Model 6), the test shows that these variables show joint significance as well (at a one per cent level for Model 5 and at a 5 per cent level for Model 6). These results suggest that models that are augmented with the Tweets variables can explain the variance of Log-RV better. The detailed results of the F-Tests are shown in Table 8 in the Appendix.

6.2.3) Loss Functions

In this section, the results of the loss functions are discussed. The loss functions are calculated according to the formulas laid out in chapter 5. First, the results of the in-sample estimations are reported, followed by the results of the rolling-window estimation.

6.2.3.1) In sample

To compare the models between each other and to assess the estimation as well as the forecasting accuracy, loss functions were used. In this section, the loss functions of the model using the whole sample period (in-sample) are reported, with the "best" results highlighted in Table 5. After comparing all measures of the loss function, it can be seen that Model 5 and Model 6 appear to be the superior models, having the smallest numbers for the loss functions. When comparing the MAPE, all models have the same first 14 digits after the decimal point, with Model 5 performing slightly better. This suggests that in-sample, Model 5 and Model 6 have the best accuracy modelling the RV of Bitcoin and that augmenting the base model with the Tweets variable, together with the interaction variables with Google Trends, modelling errors decrease, in comparison with the base model. Furthermore, all models (Model 3 -

Model 6), extended with variations of the Tweets variable, show smaller values for the loss functions, in contrast to Model 1 and the augmented base model (Model 2).

Model	MAE	MSE	QLIKE	MAPE
1	0.5038	0.4602	0.006923	0.1412
2	0.4972	0.4501	0.006802	0.1412
3	0.4972	0.4501	0.006787	0.1412
4	0.4968	0.4490	0.006776	0.1412
5	0.4957	0.4470	**0.006727**	**0.1412**
6	**0.4956**	**0.4469**	0.006728	0.1412

Table 5, Loss functions for in-sample regression

6.2.3.2) Rolling Window Regression

To evaluate the forecast performance of the regression models further, a rolling window regression was performed. Therefore a 2-year-rolling-window (730 days) was used for a daily estimation of the model parameters for the last year of the sample data (from 30th October 2020 until 30th October 2021). To evaluate the forecasting abilities of the models, the predicted values are used to obtain the results of the loss functions. The results are reported in Table 6, with the "best" values highlighted again. Model 6 performs best among all loss functions, followed by Model 5. Again, all models that use Tweets to forecast realized variance show smaller loss functions than the standard HAR-RV model and the base model, except for the MAE of Model 3, which shows smaller values than Model 2. Here, R-squared is also included, a measure that is reported within the previously reported regression output tables for the in-sample modelling results. These numbers underline that the errors in forecasting the realized variance of Bitcoin can be reduced by expanding the models with variables about Tweets by Elon Musk about Bitcoin and cryptocurrencies in general.

Model	MAE	MSE	QLIKE	MAPE	R-Squared
1	0.4298	0.3445	0.00568	0.1082	0.4778
2	0.4175	0.3319	0.00552	0.1082	0.4969
3	0.4199	0.3251	0.00537	0.1081	0.5071
4	0.4149	0.3138	0.00518	0.1081	0.5244
5	0.4051	0.3040	0.00498	0.1081	0.5392
6	**0.4048**	**0.3039**	**0.00497**	**0.1081**	**0.5394**

Table 6, Loss functions for the rolling-window-regression

6.3) Interpretation

Overall, it can be concluded that the estimation results support the hypothesis that Tweets by Elon Musk about cryptocurrencies do have an effect on next day's Bitcoin volatility, although with some limitations. Before interpreting the regression results for our variable of interest, Tweets, the effects of all other independent variables on log realized variance are being discussed. The interpretation of the results is based on the numbers reported previously in Table 3 and Table 4.

The lagged RV components show a decaying positive effect on the RV of Bitcoin prices. The daily lagged RV component drives next day's realized variance the most, whereas the monthly RV component exhibits just a small effect and is only significant at a 5 or 10 per cent level, depending on the model. This can be explained by the construction of the lagged RV components, where the lagged daily RV component is already, to some extent, included in the weekly and monthly realized variance variables and that the latest movements in prices seem to have a higher effect. This decaying effect has also been demonstrated by others for Bitcoin and for various asset classes.

The variable weekend shows a highly significant negative effect on RV, with a coefficient of around -0.196 for all models. Keeping everything else constant, this effect corresponds to a 17.88 percent decrease in Bitcoin RV on weekends, in comparison to the average realized variance on weekdays. This result is in line with previous research, as mentioned above, finding that volatility of Bitcoin appears to be lower on weekends. This can be explained by a decrease in trading activity on weekends by institutional traders and fewer spillover effects of big changes in the stock market on Bitcoin. Google Trends in general shows a weakly significant positive effect on the next day's volatility among all models that it is used in, which indicates that the Google Trends measure used for the keyword Bitcoin can to some extent help to model and forecast Bitcoin's volatility. The effect of this variable on next day's RV is higher, when the number of Google searches for 'Bitcoin' increases relative to the previous periods' averages.

Tweets, regardless of their number of interactions or their content, only show a weakly significant effect on next day's RV. In Model 3, the regression coefficient for Tweets is only significant at a 10 per cent level. Therefore, based on the regression result of this model, it cannot be concluded with confidence that Tweets are a significantly strong driver of next day's realized variance. Splitting the Tweet variable up according to their number of interactions (retweets) changes that: The estimation in Model 4 now shows that Tweets with an above-average number of interactions are a highly significant driver of next day's realized variance. The coefficient of 0.354 for $\mathrm{Tweets}_{H,t-1}$ would correspond to an average increase of next day's log realized variance by 42.52 per cent compared to the average daily log RV, if all other values were kept constant. This result indicates that including interactions with a post on Twitter can help to improve the assessment of the effect of posts on Twitter on the RV of Bitcoin prices in these models. Interactions in these estimations act as a proxy for attention of users and the reach of Tweets. So, a high reach and a higher attention on Tweets is associated with a higher next day's volatility. In contrast, low-interaction Tweets do not exhibit any significant effect at all. These results show that including the number of interactions can be useful when using posts on social media to model or forecast volatility.

The interpretation of the coefficients of the models with the interaction variables of Google Trends and Tweets is less trivial. The interaction term tries to capture the interaction between Tweets by Musk and Google Trends. In Model 5 and Model 6, this interaction variable is introduced to the regression models, both for Tweets with a higher and with a lower number of interactions. Again, for both models, Tweets with below-average interactions as well as their respective interactions term with Google Trends are not significant.

The variables of Tweets with high interactions in Model 5 and Model 6 have an even higher coefficient. However, the combined effect of $\text{Tweets}_{H,t-1}$ together with its interaction term in Model 5 would be similar to the Tweets variable in Model 4 and would correspond to an average increase of next day's log realized variance by around 43 per cent, everything else held constant. It is surprising that the regression results of Model 6 show that realized variance is not only affected by Tweets by Musk about Bitcoin, but also by Tweets about cryptocurrencies in general. It might be expected that certain Tweets by Musk, such as posts about Dogecoin, would not affect the next day's bitcoin RV, but these estimation results strongly suggest otherwise. An explanation for that can be given by the extreme interconnection of the cryptocurrency market. For example, Ferroni (2022) finds that only a small fraction of movements in prices can be explained by "idiosyncratic characteristics" of cryptocurrencies and that shifts in prices are mainly caused by market spill overs. However, Tweets about Bitcoin have a slightly higher coefficient than Tweets about all other cryptocurrencies, indicating an even stronger effect. It is important to stress, however, that the sample size of Tweets categorized by their content decreases the sample size further for each regressor, so conclusions drawn from this model should be taken with some caution.

The coefficient of the interaction term of Tweets with either high or low interactions, and Google Trends for both Model 5 and Model 6, is negative and highly significant. The negative effect counteracts the strong positive effect of Tweets on the realized variance. This negative effect on next day's RV is stronger with the higher the number of google searches on days when these Tweets were published. The interpretation of this negative coefficient might not be intuitive at first glance, but it appears that a lower number of Google searches for Bitcoin on a day when Musk publishes a Tweet about cryptocurrencies, strengthens the effect of Tweets on Bitcoin volatility. Therefore, I would argue that these results indicate that Tweets by Musk have a bigger impact on realized variance when they are posted on days with less investors or media’s attention and therefore go to some extent against the current “mainstream”. On days where there is already a higher attention for Bitcoin, Tweets might not be that surprising for investors and therefore the effect of these Tweets is smaller (as the negative effect by the interaction variable coefficient increases). Also, this counter-effect might control for the effect of important news or events for the Bitcoin or cryptocurrency community on days when Musk posted a Tweet.

The analysis in section 6.2 supports the validity of my findings that were previously discussed. The Overall-F-Test shows that the Tweets variables, added to the base model, show joint significance, and improve the base RV-model. Also, the loss functions confirm that the errors are reduced when modelling realized volatility when the regression models are expanded by the Tweets variables. These results are strong evidence that Bitcoin volatility is being significantly affected by Tweets by Elon Musk on cryptocurrencies. It also shows that it is possible – at least for the period chosen – to improve volatility models when augmenting these HAR-RV-models with the variable Tweets. In addition to that, the loss functions for the rolling-window-regression, that provide an out-of-sample one-day-ahead forecast are being improved when the standard HAR-RV model is augmented with the Tweets variables. These results show that these Tweets variables that were used improve the model, even if the model is not trained on the full sample.

Finally, I also want to come back to the two examples I gave in the introduction. These two tweets by Musk were about Tesla accepting Bitcoin as a mean of payment and then reversing this very decision. By looking at the plots, it can also be seen that the models and forecasts for Bitcoin log RV do improve for these examples as well. In Figure 14 and Figure 15 the estimated realized variance is plotted together with the predicted values for RV from the different models, calculated with the full sample data and by the rolling window approach with a 730 days window, as described before. For both Tweets, the deviation between the actual and predicted value for log RV is smaller for model 6 – the model augmented with Tweets variables – than the base model 2 (the model without any tweet variable) on the day after the Tweet was posted.

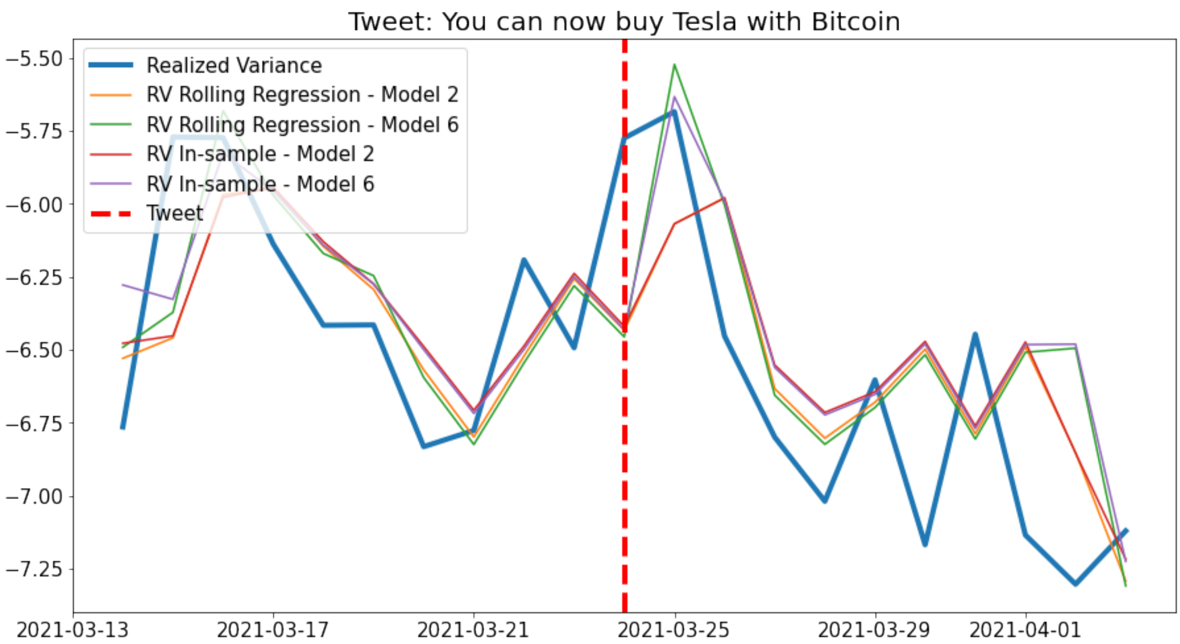

Figure 17, Predicted and actual values for log RV before & after the tweet "You can now buy Tesla with Bitcoin".

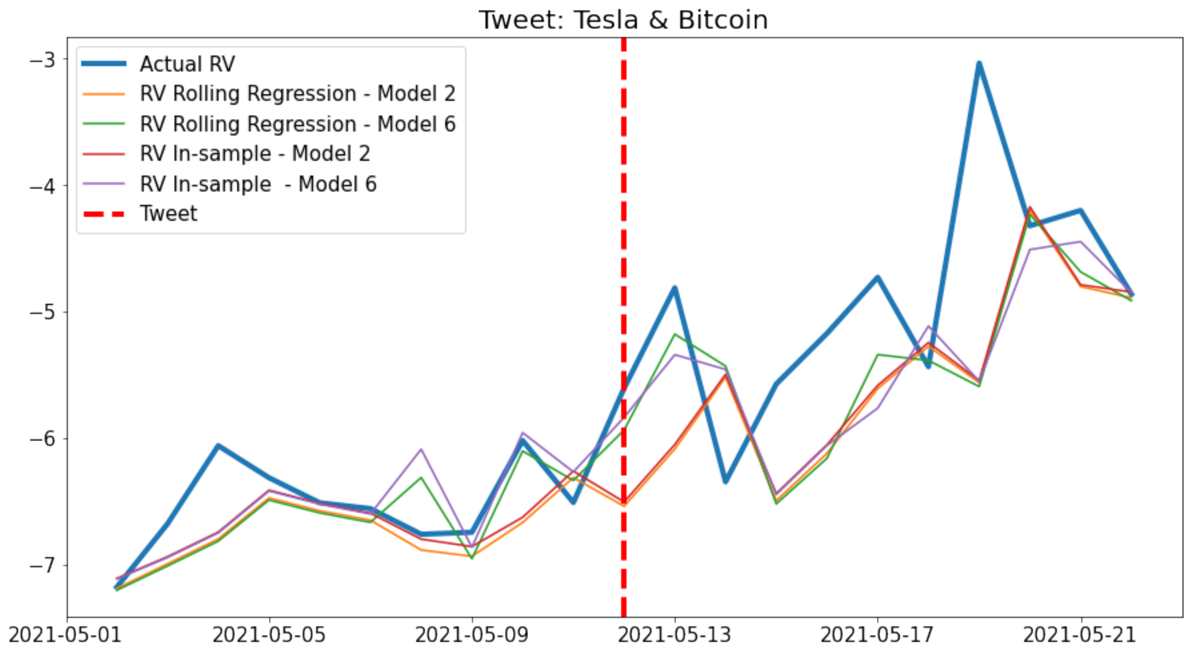

Figure 18, Predicted and actual values for log RV before & after the tweet "Tesla & Bitcoin"

7.) Conclusion

This thesis investigates the influence of Tweets by Elon Musk on Bitcoin volatility. To assess whether there is a causality, I gathered a data set of all Tweets between January 2017 and October 2021 by Musk about Bitcoin and other cryptocurrencies and used high-frequency Bitcoin data to estimate the realized volatility of Bitcoin prices.

I utilize a HAR-RV model following Corsi (2009) augmented by Tweets by Musk, a regressor that incorporates the number of Google searches for the word Bitcoin and another regressor to differentiate between weekdays and weekends. By doing so, I find evidence that Tweets by Musk affect the volatility of Bitcoin. Tweets with an above average number of interactions significantly drive next day’s volatility, whereas Tweets with a low number do not show any effect in the models used.

The effect of these high-interaction Tweets about cryptocurrencies (both about Bitcoin and all other cryptocurrencies) by the Tesla-CEO on next day’s Bitcoin volatility amounts to an average increase of around 43 per cent compared to the average daily realized volatility. I also find that the effect is not only strong for Tweets about Bitcoin in specific, but also for Tweets about all other cryptocurrencies with above-average interactions, measured by the number of “retweets”.

Using Tweets by Musk about cryptocurrencies does not only show the effect of these posts on volatility, but also helps to improve modelling and forecasting Bitcoin RV. This is shown by analysing the loss functions of the regressions results, for both modelling RV for the whole sample, but also when applying a rolling-window approach. Models augmented by Tweets variables show smaller losses for the loss functions used in comparison to the benchmark model, as well as the base HAR-RV model without the Tweets variables.

As already demonstrated in previous analyses, Bitcoin RV is highly persistent and the three lagged volatility components exhibit a decaying effect on next day’s RV. Additionally, the models used confirm prior research that Bitcoin volatility is significantly lower on weekends compared to weekdays. On the other hand, the positive effect of Google searches for the term Bitcoin on volatility is only weakly significant. For the Google Trends variable, there might be room for improvement on how to standardize or apply this variable for the models. Also, the search terms could be expanded to other words or phrases beyond just “Bitcoin”.

After undertaking this analysis, I am confident to conclude that it can be beneficial to incorporate content from social media to volatility models. Especially such as in this case, for a volatile asset like Bitcoin. It also leaves room for further research, for example to investigate the effect of Tweets about Bitcoin or cryptocurrencies with high interactions, regardless of the posts’ authors, and investigate whether they also have a significant impact on Bitcoin volatility. Moreover, other social media platforms like Facebook, Instagram or Reddit could be considered when investigating the effect of social media content on Bitcoin volatility.

These findings raise questions as well, for example on how to deal with the issue of influential people single-handedly impacting the volatility of such big assets with just a few words in a single tweet. On the one hand, people have the freedom to express themselves, but on the other hand they could harm investors by doing so. Some posts made by Musk on Twitter now serve as evidence in a lawsuit against the multi-billionaire, that claims he intentionally made certain statements on Twitter to push and inflate the price of the cryptocurrency Dogecoin. Although some might argue that his Tweets are just fun or his own opinion, the findings in this thesis show that Tweets by Musk about cryptocurrencies can cause changes in volatility, regardless of the intentions behind the posts.

8.) References

Aalborga, H. A., Molnár, P. & de Vries, J. E., 2019. What can explain the price, volatility and trading volume of Bitcoin?. *Finance Research Letters*, Volume 29, pp. 255-265. Available at: https://doi.org/10.1016/j.frl.2018.08.010

Ante, L., 2023. How Elon Musk's Twitter activity moves cryptocurrency markets. *Technological Forecasting and Social Change*, 186(A). Available at: https://doi.org/10.1016/j.techfore.2022.122112

Baur, D., Lee, A. D. & Hong, K., 2018. Bitcoin: Medium of exchange or speculative assets?. *Journal of International Financial Markets, Institutions and Money*, Volume 54, pp. 177-189. Available at: https://doi.org/10.1016/j.intfin.2017.12.004

Bergsli, L. Ø., Lind, A. F., Molnár, P. & Polasik, M., 2022. Forecasting volatility of Bitcoin. *Research in International Business and Finance*, p. Article 101540. Available at: https://doi.org/10.1016/j.ribaf.2021.101540

Bijl, L., Kringhaug, G., Molnár, P. & Sadvik, E., 2016. Google searches and stock returns. *International Review of Financial Analysis*, Volume 45, pp. 150-156. Available at: https://doi.org/10.1016/j.irfa.2016.03.015

Christoffersen, P. F., 2012. Elements of Financial Risk Management, 2nd Edition. pp. 93-110.

Corsi, F., 2009. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), p. 174–196. Available at: http://dx.doi.org/10.1093/jjfinec/nbp001

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

Available at: https://econpapers.repec.org/RePEc:bla:jfinan:v:66:y:2011:i:5:p:1461-1499

Dimpfl, T. & Jank, S., 2011. Can Internet Search Queries Help to Predict Stock Market Volatility?. *University of Tübingen Working Papers in Economics and Finance,* Volume 18.

Available at: https://www.econstor.eu/bitstream/10419/52239/1/671705237.pdf

Ferroni, F., 2022. How interconnected are cryptocurrencies and what does this mean for risk measurement?. *Chicago Fed Letter*, March, Issue 466, pp. 1-5. Available at: https://ideas.repec.org/a/fip/fedhle/93855.html

Hu, J., Kuo, W. & Härdle, W. K., 2019. Risk of Bitcoin Market: Volatility, Jumps, and Forecasts. *IRTG 1792 Discussion Papers*, 2019(024). Available at: https://ideas.repec.org/p/zbw/irtgdp/2019024.html

Iyer, T., 2022. Cryptic Connections: Spillovers. *Global Financial Stability Notes*, 2022(001), pp. 1-10. Available at: https://www.imf.org/-/media/Files/Publications/gfs-notes/2022/English/GFSNEA2022001.ashx

Kim, N., Lučivjanská, K., Molnár, P. & Villa, R., 2019. Google searches and stock market activity: Evidence from Norway. *Finance Research Letters*, Volume 28, pp. 208-220. Available at: https://doi.org/10.1016/j.frl.2018.05.003

Liu, L., Patton, A. J. & Sheppard, K., 2015. Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes. *Journal of Econometrics*, 187(1), pp. 293-311.

Available at: https://doi.org/10.1016/j.jeconom.2015.02.008

Lyócsa, Š., Mólnar, P., Plíhal, T. & Širanová, M., 2020. Impact of macroeconomic news, regulation and hacking. *Journal of Economic Dynamics & Control*, Band 119, pp. 1-21. Available at: https://doi.org/10.1016/j.jedc.2020.103980

Shen, D., Urquhart, A. & Wang, P., 2018. Does twitter predict Bitcoin?. *Economics Letters*, Band 174, pp. 118-122.

Available at: https://doi.org/10.1016/j.econlet.2018.11.007

Wang, J.-N., Liu, H.-C. & Hsu, Y.-T., 2020. Time-of-day periodicities of trading volume and volatility in Bitcoin exchange: Does the stock market matter?. *Finance Research Letters*, Band 34, pp. 1-10. Available at: https://doi.org/10.1016/j.frl.2019.07.016

Wooldridge, J. M., 2013. *Introductory Econometrics: A Modern Approach.* 5 ed. Mason, Ohio: South-Western Cengage Learning.

9.) Appendix

9.1) Tables

Test statistics Augmented Dickey-Fuller Test for log returns.

Test statistic:	-72.1377
P-value:	0.0

Test statistics Augmented Dickey-Fuller Test for log RV.

Test statistic:	-6.49925478560394
P-value:	1.173004443196496e-08

Table 7, ADF-Tests

Overall F- Test

Model	Df Resid	SSR	Df Diff	SSR Diff	F	Pr(>F)
Model 2	1699	769.355999	0			
Model 3	1698	767.910618	1	1.445381	3.19602	0.073996
Model 2	1699	769.355999	0			
Model 4	1697	766.039077	2	3.316922	3.673975	0.025577
Model 2	1699	769.355999	0			
Model 5	1695	762.636765	4	6.719234	3.733462	0.004963
Model 2	1699	769.355999	0			
Model 6	1693	762.583810	6	6.772189	2.505805	0.020365

Table 8, Overall-F-Test

9.2) Tweets

Below all Tweets used for this Thesis are listed. The columns BTC & Crypto denote whether the Tweet is about Bitcoin or cryptocurrencies in general, the content-column contains the written text (without pictures). The columns Likes & RTs (short for Retweets) show the number of interactions and in the final column "Link", the direct link to all the tweets can be found.

Date	BTC	Crypto	Content	Likes	RTs	Link
28.11.2017 02:08	1	0	Not true. A friend sent me part of a BTC a few years, but I don't know where it is.	3660	575	Link
23.12.2017 03:46	1	0	Never heard of it	9767	999	Link
29.01.2018 01:27	0	1	But wait, there's more: the flamethrower is sentient, its safe word is "cryptocurrency" and it comes with a free blockchain	103136	16807	Link
22.02.2018 09:16	1	0	Not sure. I let @jack know, but it's still going. I literally own zero cryptocurrency, apart from .25 BTC that a friend sent me many years ago.	1136	195	Link
05.05.2018 23:09	0	1	Cryptocandy	49361	7460	Link
08.07.2018 18:54	0	1	I want to know who is running the Etherium scambots! Mad skillz …	5069	540	Link
28.08.2018 16:50	0	1	At this point, I want ETH even if it is a scam	5072	2144	Link
22.10.2018 23:51	1	0	Wanna buy some Bitcoin? 😉😉	23864	8677	Link
21.02.2019 10:44	1	0	That said, I still only own 0.25 BTC, which a friend sent me several years ago. Don't have any crypto holdings.	42875	2300	Link
21.02.2019 09:50	1	0	Whoever owns the early BTC deserves a Nobel prize in delayed gratification	6195	1067	Link
02.04.2019 09:24	0	1	Dogecoin might be my fav cryptocurrency. It's pretty cool.	12093	4092	Link
02.04.2019 19:40	0	1	Uh oh	15241	531	Link
02.04.2019 20:16	0	1	Dogecoin rulz	126046	17158	Link

03.04.2019 20:38	0	1	Dogecoin value may vary	19644	2136	Link
13.04.2019 16:22	0	1	Cryptocurrency is my safe word	8257	1918	Link
14.04.2019 22:15	0	1	Just use this handy guide	1762	173	Link
30.04.2019 01:15	0	1	Ethereum	101299	11059	Link
10.01.2020 06:53	1	0	Bitcoin is *not* my safe word	114095	9132	Link
01.02.2020 23:47	0	1	The crypto scam level on Twitter is reaching new levels. This is not cool.	3059	633	Link
01.05.2020 22:44	1	0	How much for some anime Bitcoin?	17476	2592	Link
15.05.2020 22:03	1	0	Pretty much, although massive currency issuance by govt central banks is making Bitcoin Internet Geist money look solid by comparison	14640	3322	Link
15.05.2020 22:51	1	0	I still only own 0.25 Bitcoins btw	9327	1105	Link
18.07.2020 00:53	0	1	Excuse me, I only sell Doge!	15698	1018	Link
18.07.2020 00:58	0	1	It's inevitable	608279	85340	Link
16.11.2020 21:02	1	0	Toss a bitcoin t your Witcher	63448	3852	Link
17.11.2020 17:45	0	1	Dojo/Doge	1038	134	Link
20.12.2020 08:21	1	0	Bitcoin is my safe word	232588	22349	Link
20.12.2020 08:45	1	0	Bitcoin Meme	385921	49908	Link
20.12.2020 09:24	1	0	Bitcoin is almost as bs as fiat money	136057	13275	Link
20.12.2020 09:30	0	1	One word: Doge	216630	32050	Link
25.12.2020 16:47	0	1	Merry Christmas & happy holidays! 🎁	228746	14899	Link
09.01.2020 05:28	1	0	Me neither	52789	4320	Link
28.01.2021 10:47	0	1	Picture: DOGUE	431452	78913	Link
04.02.2021 07:35	0	1	Doge	254647	51628	Link
04.02.2021 07:57	0	1	ur welcome	976380	162650	Link
04.02.2021 08:15	0	1	Dogecoin is the people's crypto	540934	13640	Link

04.02.2021 08:27	0	1	No highs, no lows, only Doge	754988	113537	Link
06.02.2021 04:02	0	1	Much wow!	351837	29475	Link
06.02.2021 04:51	0	1	The future currency of Earth	255944	67270	Link
06.02.2021 08:42	0	1	Dogecake	9489	1535	Link
07.02.2021 06:24	0	1	It’s the most fun crypto!	22390	3902	Link
07.02.2021 07:41	0	1	So … it’s finally come to this …	673950	87722	Link
07.02.2021 22:25	0	1	🎶 Who let the Doge out 🎶	768146	122171	Link
08.02.2021 00:09	0	1	The people have spoken …	218984	34787	Link
08.02.2021 01:13	0	1	Ð is for Ðogecoin! Instructional video.	253760	55131	Link
08.02.2021 05:27	0	1	Doge appears to be inflationary, but is not meaningfully so (fixed # of coins per unit time), whereas BTC is arguably deflationary to a fault. Transaction speed of Doge should ideally be a few orders of magnitude faster.	15144	3989	Link
10.02.2021 07:18	1	0	This is true power haha	190149	12213	Link
10.02.2021 15:08	0	1	Bought some Dogecoin for lil X, so he can be a toddler hodler	535385	66185	Link
11.02.2021 09:08	0	1	Frodo was the underdoge, All thought he would fail, Himself most of all.	242266	33315	Link
11.02.2021 22:27	0	1	Doge is underestimated	27851	9318	Link
14.02.2021 22:25	0	1	If major Dogecoin holders sell most of their coins, it will get my full support. Too much concentration is the only real issue imo.	322335	39391	Link
14.02.2021 22:33	0	1	An acceptable percentage. Doge is much more concentrated	5720	733	Link
19.02.2021 02:01	1	0	Tesla’s action is not directly reflective of my opinion. Having some Bitcoin, which is simply a less dumb form of liquidity than cash, is adventurous enough for an S&P500 company.	26269	3591	Link
20.02.2021 06:07	0	1	An email saying you have gold is not the same as having gold. You might as well have crypto. Money is just data that allows us to avoid the inconvenience of barter. That	83155	11219	Link

			data, like all data, is subject to latency & error. The system will evolve to that which minimizes both.			
20.02.2021 06:02	1	0	That said, BTC & ETH do seem high lol	18269	2939	Link
20.02.2021 08:42	0	1	I just set up some little Doge mining rigs with my kids. It was fun.	54568	7392	Link
20.02.2021 20:52	0	1	Heard a rumor some crypto coin was pegging the dollar 🤣🤣	267662	4325	Link
20.02.2021 21:06	0	1	Or did they say pegged to the dollar? Something like that …	105635	4192	Link
20.02.2021 23:42	1	0	Cryptocurrency explained	149958	25361	Link
21.02.2021 21:27	0	1	Dojo 4 Doge	318765	35273	Link
24.02.2021 13:00	0	1	Literally	563233	74066	Link
01.03.2021 19:57	0	1	Doge meme shield (legendary item)	278268	27931	Link
02.03.2021 08:36	0	1	2022: Dogecoin is dumb	15088	2766	Link
02.03.2021 17:50	0	1	Scammers & crypto should get a room	192784	18345	Link
06.03.2021 04:40	0	1	Doge spelled backwards is Egod	399767	42020	Link
12.03.2021 18:58	1	0	BTC (Bitcoin) is an anagram of TBC(The Boring Company) What a coincidence!	238704	19249	Link
13.03.2021 23:40	0	1	Doge day afternoon	171703	19903	Link
13.03.2021 23:46	0	1	Origin of Doge Day Afternoon: The ancient Romans sacrificed a Dogecoin at the beginning of the Doge Days to appease the rage of Sirius, believing that the star was the cause of the hot, sultry weather.	73336	7854	Link
13.03.2021 23:51	0	1	Why are you so dogematic, they ask	240172	20265	Link
14.03.2021 03:54	0	1	I’m getting a Shiba Inu #resistanceisfutile	244861	23797	Link
15.03.2021 07:52	0	1	I’m selling this song about NFTs as an NFT	236986	36524	Link
15.03.2021 23:11	0	1	420M Doge	78230	5390	Link
18.03.2021 21:17	0	1	Sometimes it’s about Doge	48244	3732	Link
24.03.2021 07:02	1	0	You can now buy a Tesla with Bitcoin	876903	133660	Link

24.03.2021 07:09	1	0	Tesla is using only internal & open source software & operates Bitcoin nodes directly. Bitcoin paid to Tesla will be retained as Bitcoin, not converted to fiat currency.	175528	23130	Link
24.03.2021 07:10	1	0	Pay by Bitcoin capability available outside US later this year	135283	9228	Link
01.04.2021 10:25	0	1	SpaceX is going to put a literal Dogecoin on the literal moon	534373	59770	Link
17.04.2021 01:37	0	1	Everything to the moon!	547633	70187	Link
24.04.2021 08:45	0	1	What does the future hodl?	451641	34431	Link
26.04.2021 23:15	1	0	No, you do not. I have not sold any of my Bitcoin. Tesla sold 10% of its holdings essentially to prove liquidity of Bitcoin as an alternative to holding cash on balance sheet.	96609	15638	Link
28.04.2021 06:20	0	1	The Dogefather SNL May 8	452983	95734	Link
07.05.2021 04:24	0	1	Cryptocurrency is promising, but please invest with caution!	221253	42247	Link
07.05.2021 21:05	0	1	Guest starring …	755879	121744	Link
09.05.2021 22:41	0	1	SpaceX launching satellite Doge-1 to the moon next year – Mission paid for in Doge – 1st crypto in space – 1st meme in space To the mooooonnn!!	533243	129589	Link
11.05.2021 08:13	0	1	Do you want Tesla to accept Doge?	395210	112445	Link
12.05.2021 22:06	1	0	Tesla & Bitcoin	498171	125524	Link
13.05.2021 09:54	1	0	Energy usage trend over past few months is insane https://cbeci.org	89061	12411	Link
13.05.2021 21:11	1	0	To be clear, I strongly believe in crypto, but it can't drive a massive increase in fossil fuel use, especially coal	366238	42600	Link
13.05.2021 22:45	0	1	Working with Doge devs to improve system transaction efficiency. Potentially promising.	541347	101994	Link
16.05.2021 01:20	0	1	Ideally, Doge speeds up block time 10X, increases block size 10X & drops fee 100X. Then it wins hands down.	55471	17725	Link
16.05.2021 02:03	0	1	Only if Doge can't do it. Big pain in the neck to create another one.	22875	4879	Link

16.05.2021 18:17	1	0	Bitcoin is actually highly centralized, with supermajority controlled by handful of big mining (aka hashing) companies.	39842	10375	Link
16.05.2021 18:48	1	0	Indeed	32877	7201	Link
16.05.2021 19:37	1	0	Screenshot of Tweet	17620	3510	Link
17.05.2021 05:56	1	0	To clarify speculation, Tesla has not sold any Bitcoin	100188	24633	Link
19.05.2021 14:42	1	0	Tesla has 💎 🙌	458877	74897	Link
20.05.2021 17:56	0	1	Yeah, I haven’t & won’t sell any Doge	51753	14000	Link
20.05.2021 18:00	1	0	I agree that this *can* be done over time, but recent extreme energy usage growth could not possibly have been done so fast with renewables. This question is easily resolved if the top 10 hashing orgs just post audited numbers of renewable energy vs not.	8781	1047	Link
22.05.2021 10:25	0	1	The true battle is between fiat & crypto. On balance, I support the latter.	65002	14571	Link
24.05.2021 05:25	0	1	He fears the …	64978	16584	Link
24.05.2021 19:42	1	0	Spoke with North American Bitcoin miners. They committed to publish current & planned renewable usage & to ask miners WW to do so. Potentially promising.	324716	48776	Link
24.05.2021 19:49	0	1	If you’d like to help develop Doge, please submit ideas on GitHub & http://reddit.com/r/dogecoin/	191127	35560	Link
24.05.2021 20:29	0	1	Someone suggested changing Dogecoin fees based on phases of the moon, which is pretty awesome haha	127317	12646	Link
25.05.2021 05:37	0	1	Doge has dogs & memes, whereas the others do not	48376	9120	Link
25.05.2021 20:16	0	1	Please note Dogecoin has no formal organization & no one reports to me, so my ability to take action is limited	33132	4774	Link
01.06.2021 23:28	0	1	…	39764	8085	Link
04.06.2021 01:07	1	0	#Bitcoin 💔	214070	31769	Link

04.06.2021 02:49	1	0	*Meme*	174213	22893	Link
04.06.2021 03:45	0	1	Nice	28171	3742	Link
04.06.2021 03:57	0	1	*Meme*	48888	8295	Link
05.06.2021 08:21	0	1	I pretty much agree with Vitalik	24382	3980	Link
05.06.2021 08:39	0	1	💨🚀 –> 🌜	305153	36125	Link
13.06.2021 17:46	1	0	This is inaccurate. Tesla only sold ~10% of holdings to confirm BTC could be liquidated easily without moving market	52698	12918	Link
25.06.2021 02:02	1	0	Bicurious?	10386	11918	Link
25.06.2021 02:10	1	0	How many Bitcoin maxis does it take to screw in a lightbulb?	100161	10268	Link
25.06.2021 02:11	1	0	“That’s not funny!” – Bitcoin maxis	58705	3085	Link
01.07.2021 08:43	0	1	Release the Doge!	130420	17906	Link
01.07.2021 09:24	0	1	Baby Doge, doo, doo, doo, doo, doo, Baby Doge, doo, doo, doo, doo, doo, Baby Doge, doo, doo, doo, doo, doo, Baby Doge	278953	57038	Link
02.07.2021 03:09	0	1	But can I pay in Doge	16854	3746	Link
02.07.2021 13:20	0	1	*Meme*	206797	202394	Link
09.07.2021 07:15	0	1	BTC & ETH are pursuing a multilayer transaction system, but base layer transaction rate is slow & transaction cost is high. There is merit imo to Doge maximizing base layer transaction rate & minimizing transaction cost with exchanges acting as the de facto secondary layer.	26971	8078	Link
12.07.2021 21:32	0	1	It’s true, x is a toddler hodler. Never once has he said sell!	24233	3623	Link
17.07.2021 16:53	0	1	Lil X is hodling his Doge like a champ. Literally never said the word “sell” even once!	14901	3072	Link
25.07.2021 04:23	0	1	And finally	269666	35385	Link
28.07.2021 23:41	1	0	We don’t have that many Bitcoin, but it’s close	25973	1510	Link
15.08.2021 15:02	0	1	I've been saying that for a while	26165	4673	Link

22.09.2021 20:50	0	1	Super important for Doge fees to drop to make things like buying movie tix viable	23615	4477	Link
21.10.2021 02:41	1	0	*Meme*	759779	96149	Link

Abstract:

This master's thesis investigates the question of whether posts by Elon Musk on Twitter (now: X) can affect Bitcoin volatility. Since some claim Musk is able to influence the Bitcoin price with a single tweet, I investigate this issue by modelling and forecasting the volatility of Bitcoin. I apply a heterogeneous autoregressive model of realized volatility (HAR-RV) and utilise high-frequency intraday data of Bitcoin prices. The model is augmented not only by variables that indicate Tweets by Musk, but also by several other variables such as a variable that distinguishes between weekdays and weekends and a variable that reflects the number of Google searches for the word Bitcoin. Among other things, I find that tweets by Elon Musk with interactions above the average have a strong significant effect on realised volatility. Also, adding tweets variables to the HAR-RV model can help improve the modelling and forecasting of Bitcoin's realised volatility.

Zusammenfassung:

Diese Masterarbeit untersucht, ob Posts von Elon Musk auf Twitter (jetzt: X) die Bitcoin-Volatilität beeinflussen können. Einige meinen, dass Musk in der Lage sei, den Bitcoin-Kurs mit einem einzigen Tweet zu beeinflussen. Deshalb untersuche ich diese Frage, indem ich die Volatilität von Bitcoin modelliere und prognostiziere. Dafür verwende ich ein heterogenes autoregressives Modell der realisierten Volatilität (HAR-RV) basierend auf Hochfrequenz-Daten von Bitcoin-Preisen. Das Modell erweitere ich nicht nur durch Variablen, die für die Tweets von Musk stehen, sondern auch durch andere. Beispielsweise eine Variable, die zwischen Wochentagen und Wochenenden unterscheidet und eine Variable, die die Häufigkeit der Google-Suchen nach dem Wort Bitcoin widerspiegelt. In der Masterarbeit zeige ich, dass Tweets von Elon Musk, die Interaktionen über dem Durchschnitt aufweisen, einen starken signifikanten Effekt auf die realisierte Volatilität haben. Außerdem zeigt sich, dass das Hinzufügen der Tweets-Variablen zum HAR-RV-Modell dazu beiträgt, die Modellierung und Vorhersage der Volatilität von Bitcoin zu verbessern.