universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Predictability of cryptocurrency returns"

verfasst von / submitted by

Dilshod Agishev

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

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 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

Contents

1. Introduction

In this thesis, I am going to explore the predictability of returns in the cryptocurrency market. Predictability of asset returns is important both for financial economists and investors. For the financial economists it is important, because, as pointed out by Roll (1988), the maturity of any science is reflected in its ability to predict important phenomena. And it is clear why it is important for investors: if returns can be predicted then one should be able to exploit predictability to earn a profit. It is especially important to test the predictability of returns in the cryptocurrency market, since, as I show later in the thesis, the cryptocurrency market exhibits huge volatility. Consequently, buy and hold strategy does not work on average in the crypto market, so investors need to engage in active day trading, where predictability of returns is important.

There are several contributions of this thesis to the literature on the predictability of cryptocurrency returns. First, I study the predictability of 467 cryptocurrencies, which is substantially higher than the sample sizes of previous studies on the predictability of cryptocurrency returns. This allowed me to explore the predictability of small-cap coins, which were dismissed by previous studies, and show that they are more predictable than large-cap coins. Second, I show that the degree of predictability of returns in the cryptocurrency market is not decreasing, as was shown in early studies, instead it is time-varying. Third, I explore the determinants of predictability of returns in the cryptocurrency market, which was never done before in the cryptocurrency literature. Fourth, I show that the cryptocurrency bubble in 2017-2018 had a significant effect on the degree of predictability in the market.

Before going into the details of my thesis, let me quickly introduce cryptocurrencies. In 2008 someone under the pseudonym of Satoshi Nakamoto, in an attempt to eliminate financial intermediaries from online transactions, introduced Bitcoin, a peer-to-peer digital currency that is completely decentralized. Nakamoto (2008) proposed an ingenious solution to secure transactions without a trusted third party – a distributed ledger called Blockchain, which contains the entire history of Bitcoin transactions and uses cryptographic encryption to ensure its safety. Shortly after the introduction of Bitcoin alternative coins were designed. Most of the coins copied the technology behind the Bitcoin. Other coins slightly changed the specifications, for example, Monero emphasized the anonymity of transactions. As of August 2019, more than

2000 cryptocurrencies were listed on coinmarketcap.com and their total market capitalization was more than $260 billion.

I have several research questions that I am going to address in the thesis. First, are returns in the cryptocurrency market predictable? Two main methodologies are used to test if asset returns are predictable: one is to use statistical tests to check if the returns are predictable from previous returns (an example is Lo and Mackinlay (1988)); another is to use regression to test if financial or macroeconomic factors can predict returns (an example is Fama and French (1988)). I am using the first methodology in this thesis. The tests are Automatic Portmanteau and Wild Bootstrapped Automatic Variance Ratio tests. I chose specifically these tests because both tests are robust to conditional heteroscedasticity and powerful in small samples.

Testing return predictability of cryptocurrencies is not a novel research direction, some papers did that, as shown in the survey by Kyriazis (2019). The majority of the papers report that returns of cryptocurrencies are predictable. But most of the papers focus on Bitcoin, or other top-10 cryptocurrencies, dismissing smaller ones. I fixed this shortcoming by testing the predictability of 467 largest cryptocurrencies by market capitalization. The results of the tests show that more than half of the cryptocurrencies are predictable for their entire trading period. Another major result of the tests is that small-cap cryptocurrencies are more predictable than large-cap coins.

The second research question is, how the predictability of cryptocurrency returns changes across time? Urquhart (2016) and Bariviera (2017) showed that the predictability of Bitcoin returns decreased with time. Khuntia and Pattanayak (2018) and Chu et. al. (2019) however showed that the predictability of Bitcoin returns is time-varying. To check how the predictability of returns changes across time in the cryptocurrency market, I divided the dataset of each cryptocurrency into monthly periods. I then applied Automatic Portmanteau and Wild Bootstrapped Automatic Variance Ratio tests on monthly datasets of daily returns. The results of the tests suggest that the predictability in the cryptocurrency market is time-varying, but there was a change in the degree of predictability after the year 2016. From the year 2013 until the first quarter of 2017 the share of predictable coins in a given month ranges between 10% and 18%. From the second quarter of 2017 onwards, following the 2017-2018 bubble in the cryptocurrency market, the degree of predictability in the market changed, it was ranging between 5% and 14%.

Since the degree of predictability is different among cryptocurrencies and across time, the third question is, what are determinants of cryptocurrency returns? This question was not addressed in the literature on the predictability of cryptocurrency returns before. I explore determinants of return predictability using panel data regression with fixed effects, where test statics of the abovementioned tests are used as dependent variables. Both variables indicate the degree of predictability of returns.

Regression results show that volume, volatility, and attention for the cryptocurrency market (represented by google search volume) can explain the predictability of returns. These results are in line with the literature on the predictability of stocks. The results of regression also suggest that the predictability of cryptocurrencies with a capitalization of less than $1 million is significantly different from other coins, even when controlling for other variables. Additionally, there is an interaction between cryptocurrency predictability and S&P 500 market.

At last, I tested if the interaction between cryptocurrency predictability and explanatory variables changed during and after the 2017-2018 cryptocurrency market bubble. To do that I divided the data used in the regressions into two periods: before the 2017-2018 bubble; during and after the 2017-2018 bubble. I then run the regressions separately for each period and compared the regressions coefficients of the corresponding periods. Overall, results indicate that there was a change in the market following the bubble. The surprising result here is that Google searches had no significant effect on the predictability of returns before the 2017-2018 bubble.

The structure of the thesis is as follows. Chapter 2 summarizes related literature on the predictability of cryptocurrency returns. In chapter 3, I elaborate on the data I used for analysis and some important characteristics of cryptocurrencies. Chapter 4 has a detailed description of the statistical tests I used for the analysis; I also outline the test results in the chapter. In chapter 5, I describe the regression models used to study determinants of predictability and their results. Finally, chapter 6 concludes the results of the thesis.

2. Related literature

The literature on the predictability of cryptocurrency returns is recent and not voluminous, compared to the literature on the predictability of other assets. There are three main categories of research on cryptocurrency predictability: 1) statistical tests of predictability, 2) profitability of technical analysis trading strategies, and 3) external factors affecting cryptocurrency returns. A survey by Kyriazis (2019) reviewed papers that tested the predictability of cryptocurrency returns. Out of 40 papers that tested the predictability of returns, only two conclude that returns are not predictable. Most of the papers focused on Bitcoin and other large-cap cryptocurrencies. It should be noted that almost all papers in the survey were presented as studies of market efficiency, when in fact they were testing predictability of returns. Consequently, Kyriazis also interpreted results as evidence for the rejection of the Efficient Market Hypothesis.

Efficient Market Hypothesis was developed in the 1960s by Samuelson (1965) and Fama (1965a). Fama (1965b) defined the efficient market as a market, where competition among a large number of rational profit-maximizing investors leads to the situation in the market, where all information is reflected in the market prices. As noted by Fama (1970) to test if the prices reflect the available information, one has to propose a model that describes prices. Then it becomes a test of join hypothesis: if the hypothesis is rejected, it could mean that either markets are inefficient, or the model is incorrect, or both. LeRoy (1973) and Lucas (1978) showed that, under EMH, returns should be unpredictable only if one additionally assumes that investors are risk-neutral. Under the risk-aversion assumption, the predictability of returns does not imply that EMH is violated. So, one should not treat the predictability of returns as evidence against EMH.

Urquhart (2016) was the first to investigate the predictability of Bitcoin. He used daily returns of Bitcoin from 1 August 2010 to 31 July 2016, divided into 2 sub-periods, and applied a battery of tests to check whether the returns are predictable. Urquhart reported that Bitcoin returns are predictable for the whole period. However, when he tested the second period (1/8/2013 – 31/7/2016) separately, some tests indicated randomness. Urquhart interpreted that as Bitcoin becoming less predictable through time. Nadarajah and Chu (2016) followed up on Urquhart by applying a similar methodology, but he used the odd-integer power of Bitcoin returns, arguing that it does not lead to any loss of information. Their result shows that the power transformation of Bitcoin returns is not predictable.

Bariviera (2017) looked into the long-term memory of Bitcoin, using a rolling sample of returns. He found out that returns exhibit persistence until 2014, while after 2014 behaviour of time series of returns was similar to white noise. Thus, he concluded that Bitcoin becomes less predictable with time. Caporale et. al (2018) tested the predictability of four cryptocurrencies (Bitcoin, Litecoin, Ripple, and Dash) in a similar fashion to Bariviera. They reported persistence in returns and suggested that technical analysis can be used to earn abnormal profits.

Khuntia and Pattanayak (2018) tested the predictability of Bitcoin by applying statistical tests on a rolling window of returns. Their results show that the degree of predictability of Bitcoin varies from one period to another. Chu et. al. (2019) applied a battery of predictability tests on Bitcoin and Ethereum returns and confirmed that predictability is time-varying. These results are contradictory to the results of Urquhart (2016), but they can be explained by Lo's (2004, 2005) Adaptive Market Hypothesis.

Adaptive Market Hypothesis (AMH) is based on the principles of evolutionary biology and bounded rationality (Lo (2004)). Lo highlighted the following components of the AMH: "*a) individuals act in their own self-interest, b) individuals make mistakes, c) individuals learn and adapt, d) competition drives adaptation and innovation, e) natural selection shapes market environment, f) evolution determines market dynamics*" (Lo (2005), p. 11). The implications of the AMH are as follows: 1) risk premium varies from one period to another, depending on the population of investors in the market and market conditions; 2) arbitrage opportunities appear and disappear from time to time as investors adapt to the market environment or population of investors changes, 3) different trading strategies might be suitable in different periods. In the context of return predictability that would mean that market predictability is not constant and varies from time to time, depending on market conditions. Kim et.al. (2011) and Lim et.al. (2013) using a rolling window sample of US stock returns showed that the predictability of US stock markets behaves according to Adaptive Market Hypothesis.

The predictability of cryptocurrencies would not mean much to finance practitioners if it would not be possible to make a profit from it. Detzel et. al. (2018) tested the profitability of the use of technical analysis on Bitcoin. They found significant alphas and Sharpe ratio of the moving average rule compared to buy and hold strategy. Corbet et. al. (2019) also provide evidence that moving average strategies can be profitable when applied to Bitcoin returns.

Hudson and Urquhart (2019) went further, they tested five broad categories of technical trading rules, which include 15000 rules in total. They implement the rules on four coins: Bitcoin, Litecoin, Ethereum, and Ripple. They find out 1) that technical trading rules have high predictive power, 2) high Sharpe ratio, showing high risk-adjusted return of technical analysis, 3) profitability is not wiped out by transaction costs.

Even though Bitcoin was created to be used as a currency in online transitions, Glaser et. al. (2014) found that Bitcoin users treat it as an alternative investment rather than currency. That's why, it's important to understand what factors affect cryptocurrency returns, since coins do not have any fundamentals, such as cash flows or dividends. Sovbetov (2018) tested what factors affect the prices of five cryptocurrencies. He reported that volatility has a negative effect on prices of cryptocurrencies, while the trading volume, cryptocurrency attractiveness, and S&P500 returns have a positive effect on prices.

Liu and Tsyvinsky (2018) studied the risk and return profile of Bitcoin, Ripple, and Ethereum. They modelled the returns of the cryptocurrencies using traditional CAPM, multi-factor models, and other models. They concluded that the risk-return tradeoff of cryptocurrencies is different from those of stocks. Additionally, they found that returns cannot be predicted by stock markets, macroeconomic factors, or returns of commodities and currencies. The only variables that predict cryptocurrency returns are momentum and investor attention, which was represented by Google and Twitter search volumes.

Nasir et. al. (2019) specifically studied the effect of Google search on Bitcoin returns and volume. Their result shows that Google search volume has a positive effect on Bitcoin returns. A rapid increase in Google search has a greater positive effect, which lasts for about one week.

Bouri et. al. (2019) tested the effect of trading volume on returns of seven cryptocurrencies, using the copula-quantile causality test. They report that volume can be useful to predict extreme negative and positive returns of cryptocurrencies.

As you can see the literature on the predictability of returns is still in infancy, and most of the papers focus on a small sample of cryptocurrencies. That's why in this paper I expand the sample of cryptocurrencies being tested for the predictability. Next, several authors showed that the predictability of cryptocurrency returns is time-varying, I would like to see what affects the degree of predictability in the crypto market. The researches on factors affecting crypto returns would help me to choose appropriate variables.

3. Data

I downloaded prices of the top 500 coins by market capitalization as of August 24, 2019 from coinmarketcap.com. I dropped coins that had less than 3 months of data available, that left me with 467 cryptocurrencies. Prices of currencies on coinmarketcap.com are calculated as the weighted average of the prices on different exchanges, so it is not the actual price at which one could buy a coin. However, it is the best benchmark, since not all coins are traded on every cryptocurrency exchange. For each coin, daily prices for the entire period that was available were downloaded, that is why the date range for each currency varies, but the last date for all currencies is August 24, 2019. Figure 1 shows for how many coins data begins in a given month. Keeping in mind that all coins in the sample are among the top 500 coins by capitalization, it is interesting that the majority of coins in the sample were introduced after 2016.

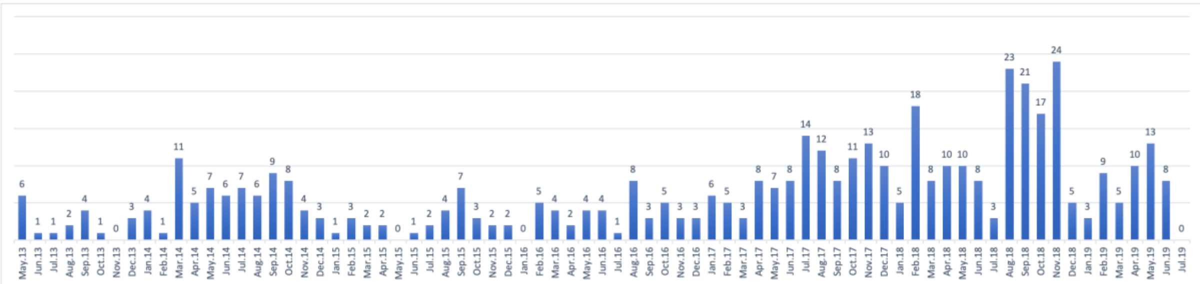

Figure 1: Number of coins per month, in which data of correspondent coin begins on coinmarketcap.com

Since this thesis is concerned with the predictability of returns, let me briefly discuss the main characteristic of cryptocurrency returns - volatility. I used daily log returns for the analysis, they are calculated as $R_t = \ln(\frac{P_t}{P_{t-1}})$, where P_t is a close price of a given coin in day t. Annualized volatility is calculated as square root of trading days in a year (365 days in the cryptocurrency market) times the standard deviation of returns in a given month. The blue line in Figure 2 depicts the average annualized volatility of daily returns in a given month for all cryptocurrencies in the sample. To put the volatility of cryptocurrencies into perspective I downloaded the prices of S&P 500 from Yahoo Finance and calculated the annualized volatility of daily S&P500 returns (red line). The difference in volatility between cryptocurrencies and S&P 500 is immense: annualized volatility of S&P 500 returns ranges between 4% and 30%, while the average annualized volatility of cryptocurrencies in the sample ranges between 109% and 485%.

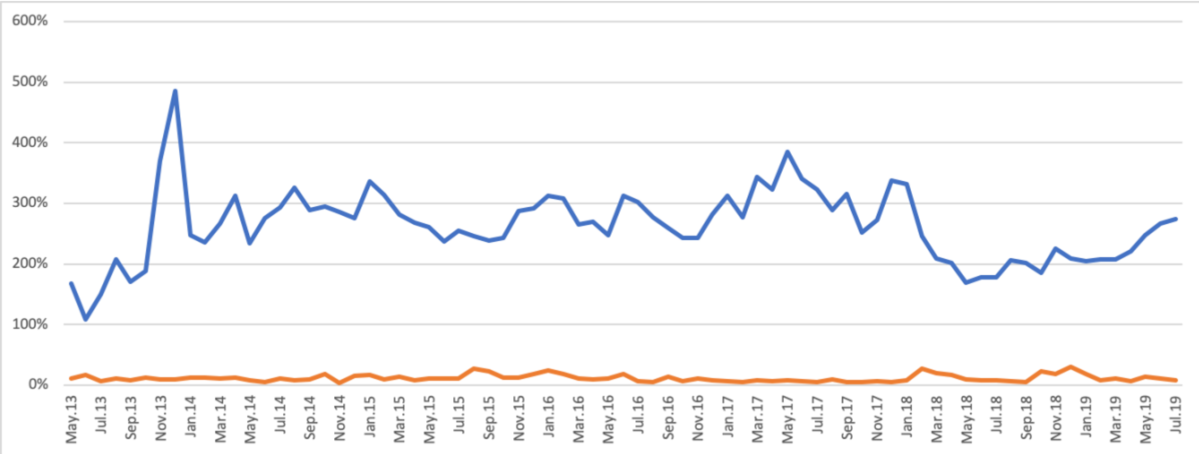

Figure 2: Average annualized volatility of daily cryptocurrency returns in the sample (blue line) and annualized volatility of daily S&P 500 returns (red)

The volatility of cryptocurrencies directly translates into their profitability. Figure 3 shows the distribution of currencies by their profit/loss for their entire holding period, $(\frac{P_{last\ day}}{P_{first\ day}} - 1)$. The main take away from the graph is that the crypto market is the market of extremes, only minority of coins have moderate profit or loss. More than half of the coins in the sample lost more than 50% of their value, and around a third of the coins had profit more than 100%. That implies that the buy and hold strategy would not be the best strategy in the crypto market unless you chose the right coins, then the only other option for investors is to engage in short term trading of cryptocurrencies. That is why studying the predictability of cryptocurrency returns is important: if buy and hold is not a proper strategy and crypto returns are not predictable, then trading cryptocurrencies is no different from gambling. On the other hand, if the returns in the market are predictable, then appropriate trading strategies could be implemented in the cryptocurrency market.

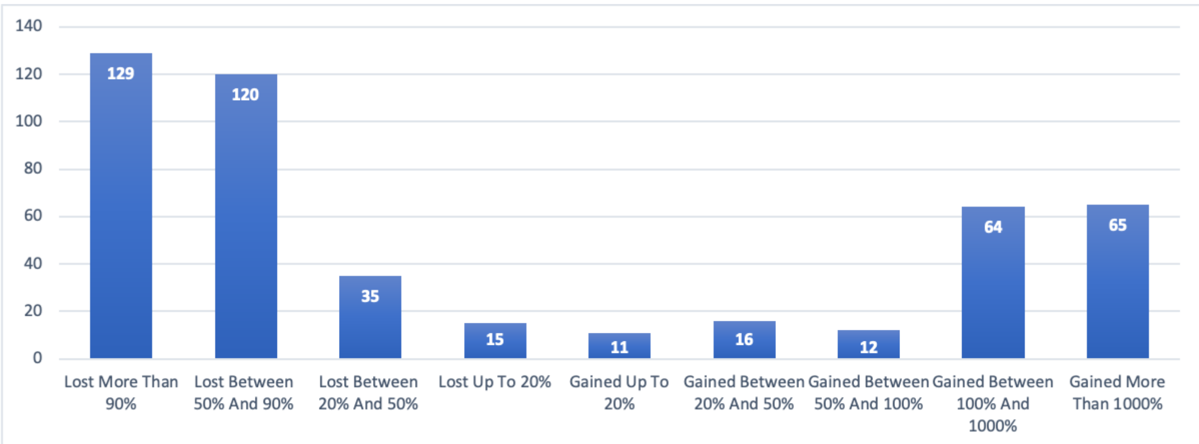

Figure 3: Distribution of cryptocurrencies by their profit/loss for their entire holding period

Daily market capitalization and volume of coins were downloaded from coinmarketcap.com alongside prices. The market capitalization (product of the price and circulating supply of a given coin) of all coins in the sample was $251 billion on August 24, 2019. Bitcoin alone represented 72% ($181 billion) of total capitalization and the top 10 coins made 94.4% of the sample capitalization. In Figure 4, which shows the time series of the total market capitalization of all cryptocurrencies, you can observe the 2017-2018 bubble: in the beginning of 2017, the market capitalization increased dramatically and peaked in January 2018, followed by a market decline for the entire year of 2018.

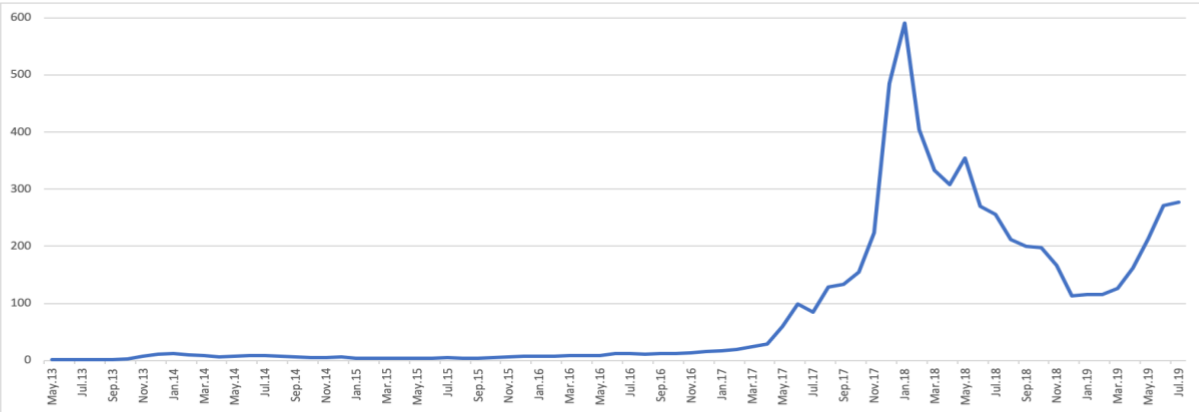

Figure 4: Total market capitalization (€ billion) of all cryptocurrencies in the sample in a given month

Figure 5 shows the total trading volume of all cryptocurrencies in the sample in a given year. As you can see, the trading volume similar to market cap increased significantly from 2016 to 2017, to be more specific it increased forty-threefold (43x). Unlike market capitalization, however, the trading volume continued to rise in subsequent years too. That shows that the 2017-2018 bubble had a strong effect on the cryptocurrency market, which is in line with the Adaptive Markets Hypothesis: bubbles and crashes change the behavior of existing investors in the market or/and changes the population of investors. This change potentially could lead to a change in the degree of predictability in cryptocurrency returns.

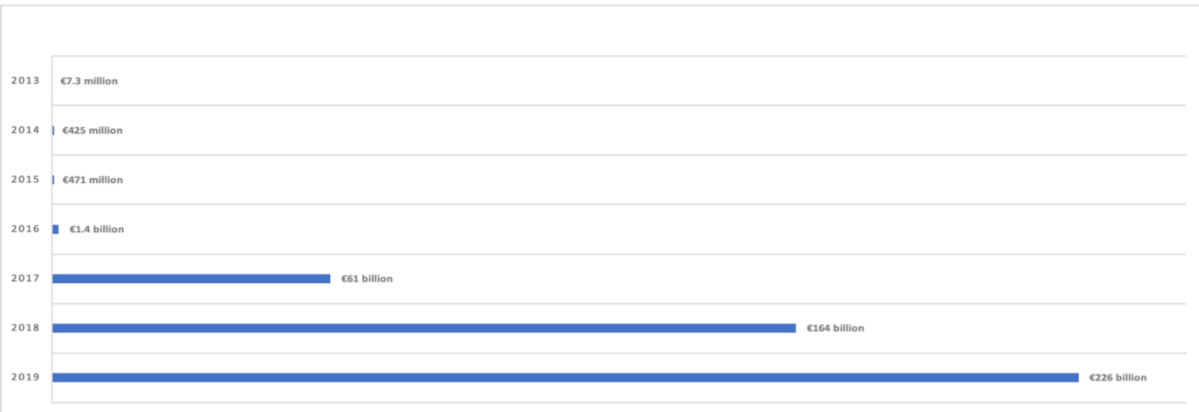

Figure 5: Total trading volume of all cryptocurrencies in the sample in a given year

Another way to observe the bubble in the cryptocurrency market by looking at Google Trends. Google does not provide an absolute number of searches for a given keyword; instead, they provide a normalized value (1-100) of search in a given period, where 100 represents the period with the highest volume of searches for a specific keyword. I downloaded the data from May 2013 until July 2019 for the search volume of the word "Cryptocurrency", which provides a fair benchmark for attention for the cryptocurrency market. Figure 6 depicts the time series of relative Google search volumes. One can instantly notice that the Google search volume is similar to the market capitalization graph (Figure 4): there was a huge increase in search volume starting from April 2017 and a drop in March 2018, which correspondent to the bubble in that period.

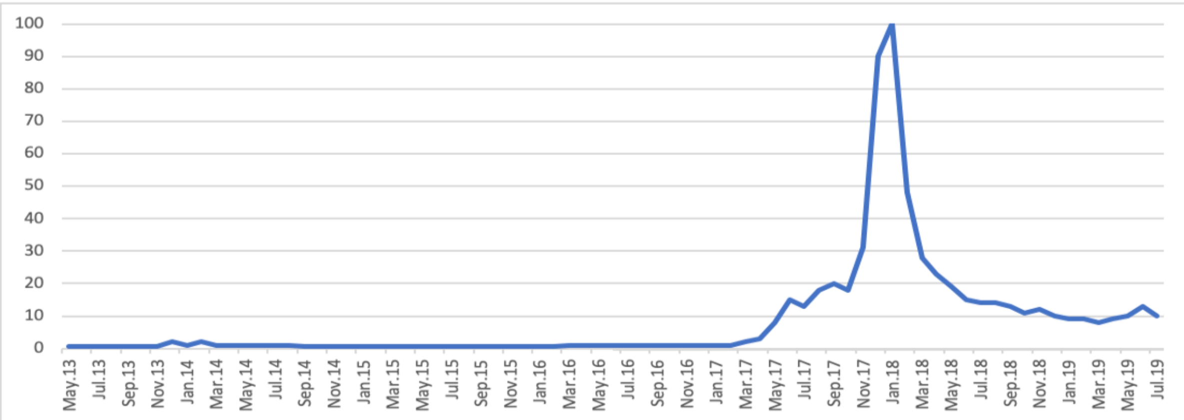

Figure 6: Relative (1-100) Google search volume for the word "Cryptocurrency". Source: Google Trends

The boom of Initial Coin Offerings (ICO) could explain the bubble in 2017. In an ICO company raises money to fund its operations, much like in Initial Public Offering in the stock market, but instead of receiving shares investor receives cryptocurrency issued by the company. In 2017 the market of ICO grew hundredfold (100x)[1], but more than half (52%) of all ICOs failed or turned out to be a fraud[2]. That led to countries introducing regulation on ICOs, with some countries completely banning it[3] and Google, Facebook and Twitter banning advertisement related to cryptocurrencies on their websites[4]. Another factor that might have caused a crash of cryptocurrencies is the fact that in December 2017 trading of Bitcoin futures on CBOE started[5]. Hale et. al. (2018) noted that the peak of Bitcoin price coincided with the start of the futures trading and the consequent drop in the price could have been caused by short-sellers on the derivative market.

1 Source
2 Source
3 Source
4 Source
5 Source

There are two main takeaways from this chapter. First, the volatility in the cryptocurrency market is huge. This fact makes the study of predictability of returns in the cryptocurrency market very important: since the buy and hold strategy is not profitable on average and an investor has to actively trade cryptocurrencies, one firstly needs to check if the returns are predictable to use some appropriate trading strategy. Second, the 2017-2018 bubble has changed the cryptocurrency market, which according to Adaptive Markets Hypothesis would lead to a change in the degree of predictability in the market. I will test if that is true in upcoming chapters.

4. Test methodology and results

There are two major ways to test the predictability of asset returns: 1) use statistical tests to check if the returns are completely random, or there is some predictable element: i.e. testing if returns are predictable based on previous returns; 2) using regression, test if financial, macroeconomic or other factors can predict returns. While I am not going into details of the later methodology, since I am not using this methodology, it is worth noting that US stock returns can be predicted using interest rates (Fama and Schwert (1977), Campbell (1987)), price to earnings ratios (Basu (1977)) and dividend yield (Fama and French (1988)). As for the cryptocurrency market, Liu and Tsyvinsky (2018) showed that returns of three major cryptocurrencies cannot be predicted by stock market returns, macroeconomic factors, and traditional currency returns, but can be predicted by the Google and Twitter search volume.

The most popular way to test return predictability from past returns is to test whether prices follow a random walk: if they follow a random walk then returns are not predictable, if the random walk is rejected then the returns can be predicted. Campbell et. al. (1997) distinguished three versions of the random walk. In the first version (RW1), the most restrictive version, the assumption is that the returns are independently identically distributed. To test if the time series follows RW1 one can use runs test: compare the number of realized runs - sequences of returns with the same positive/ negative sign - to the expected number of runs. If the number of actual runs is significantly different from the expected number of runs, then we reject RW1. The second version (RW2) relaxes the assumption of identically distributed returns and assumes only independent returns. RW2 could be tested using non-parametric correlation rank tests, such as Spearman test, but Campbell et. al. (1997) argue that these tests are not always appropriate. Another way to test RW2 is to use filter rules trading strategy - purchase the asset if the price increases by x%, sell the asset if price decreases by x%. If the strategy is significantly profitable, it would indicate predictability in returns.

Even though RW1 and RW2 were tested most frequently in early predictability literature (see Fama (1965a)), the assumption of independent (and identically) distributed returns is not realistic for financial markets. The third version of random walk (RW3) relaxes the independence of returns assumption and assumes only uncorrelated returns, which means that the correlation between returns at all lags is equal to zero. The two most widely used tests for RW3 are the Portmanteau test and the Variance ratio test. Since RW3 has the most realistic assumptions for financial time series, I am testing against this particular version of the random walk model.

It is important to use statistical tests that are robust to conditional heteroscedasticity, since asset returns time series exhibit conditional heteroscedasticity, as shown by Akgiray (1989). That is why I am following the methodology of Kim et. al. (2011) and Lim et. al. (2013), where they use tests that are robust to heteroscedasticity. The first test is an Automatic Portmanteau (AQ) test for serial correlation, introduced by Escanciano and Lobato (2009). The second test is the Wild Bootstrapped Automatic Variance Ratio (AVR*) test proposed by Kim (2009).

4.1 Automatic Portmanteau test

Box-Pierce and Ljung-Box tests are used most frequently to test the predictability of returns of cryptocurrencies (Kyriazis (2019)). Box and Pierce (1970) test statistics formally defined as:

$$Q_k = n \sum_{k=1}^{K} \widehat{\rho_k}^2 \tag{1}$$

with $\widehat{\rho_k}$ is a sample serial correlation of order *k*. Its null hypothesis is that there is no autocorrelation for all K.

Box-Pierce Q test has two issues: 1) it is not robust to heteroscedasticity; 2) the choice of k is arbitrary. Lobato et al. (2001) tackled the first issue by designing a Q statistic that is robust to conditional heteroscedasticity:

$$Q_K^* = n \sum_{k=1}^{K} \widetilde{\rho_k}^2 \tag{2}$$

with $\widetilde{\rho_k}^2 = \widehat{\gamma_k}^2 / \widehat{\tau_k}$, where $\widehat{\gamma_k}$ is sample autocovariance and $\widehat{\tau_k} = \frac{1}{n-k} \sum_{t=1+k}^{n} (Y_t - \bar{Y})^2 (Y_{t-k} - \bar{Y})^2$.

Escanciano and Lobato (2009) used the robustified test statistic Q* and made the choice of k automatic to deal with the second issue. Automatic Portmanteau Test defined as:

$$AQ = Q_{\tilde{k}}^{*} = n \sum_{k=1}^{K} \widetilde{\rho_{\tilde{k}}}^{2} \tag{3}$$

where $\tilde{k}$ is data-driven and is chosen using a combination of Akaike Information Criterion and Bayesian Information Criterion. The automatic Portmanteau test follows Chi distribution with one degree of freedom. Its null hypothesis is that $\tilde{\rho}_{\tilde{k}} = 0$. Escanciano and Lobato (2009) showed that the AQ test has more empirical power compared to its alternatives.

4.2 Wild Bootstrapped Automatic Variance Ratio test

In their seminal paper, Lo and Mackinlay (1988) proposed a variance ratio (VR) test of a random walk hypothesis. The variance ratio is defined as:

$$VR(k) = \frac{\sigma^2(k)}{k\sigma^2(1)} \tag{4}$$

where $\sigma^2(k)$ is the variance of k-period log-returns and $\sigma^2(1)$ variance of 1-period log returns, the null hypothesis is that *VR(k)* is equal to one.

Lo and Mackinlay (1988) applied the variance ratio tests to weekly CRSP index returns ranging from 1962 to 1985 and found a positive serial correlation, thus rejecting the random walk hypothesis.

The variance ratio test quickly became one of the most popular tests to test against random walk in finance, but it has some limitations. Charles and Darne (2009) review all the variations of the test that were developed to improve the test. One such improvement was proposed by Choi (1999), where he used quadratic spectral kernel to optimally choose k to estimate variance ratio since in literature k was chosen as some arbitrary number. Choi defined variance ratio estimator as:

$$VR(k) = 1 + 2 \sum_{i=1}^{K-1} m\left(\frac{i}{k}\right) \hat{\rho}(i) \tag{5}$$

where $\hat{\rho}(i)$ is sample autocorrelation and $m\left(\frac{i}{k}\right)$ is the quadratic spectral kernel. The standardized Automatic Variance Ratio test statistic has the following form:

$$AVR(k) = \frac{\sqrt{\frac{T}{k}}\left[\widehat{VR}(k) - 1\right]}{\sqrt{2}} \tag{6}$$

AVR(k) converges in distribution to standard normal distribution as $T \to \infty$, $k \to \infty$, and $T/k \to \infty$. The null hypothesis of the test is the absence of serial correlation. Rejecting the null hypothesis would indicate the predictability of returns.

Unfortunately, as pointed out by Kim (2009), AVR is not robust to conditional heteroscedasticity if the sample size is not large enough. Kim (2009) suggested that the wild bootstrapping AVR test could solve the conditional heteroscedasticity problem. Mammen's (1993) wild bootstrap method is usually used in financial data, where conditional heteroscedasticity is present. In the case of wild bootstrapping the AVR the procedure is as follows: first, a statistical software creates a bootstrap sample and applies AVR to the bootstrap sample; then the software repeats this procedure N times (in case of this thesis N = 500). This way the software creates N different values of AVR statistics, which is called bootstrap distribution - $\{AVR^*(k;\ j)\}_{j=1}^{N}$.

Based on bootstrap distribution we get the p-value, which is calculated as a percentage of absolute values of AVR test statics from the bootstrap distribution that is greater than the absolute value of the original AVR test statistic (Kim (2009). If the p-value is smaller than some chosen level of significance, we reject the null hypothesis of no serial correlation and conclude that returns are predictable. Kim compared Wild Bootstrapped AVR (AVR*) to other alternative variance ratio tests, including Choi's AVR. Kim reported that AVR* "*shows desirable small sample properties under conditional heteroscedasticity with correct size and higher power than its competitors*" (Kim (2009), p. 184).

4.3 Test results

I applied Automatic Portmanteau (AQ) and Wild Bootstrapped Automatic Variance Ratio (AVR*) tests to the daily returns of all cryptocurrencies in the sample, Table 1 summarizes the test results. First, the tests were applied to the entire trading period of each cryptocurrency. The null hypothesis of the AQ test was rejected for 252 coins out of 467 at a 5% level of significance. There are similar results for the AVR* test, for 259 coins the null hypothesis of the random walk was rejected at a 5% significance level. That means that more than half of all coins are indicated as predictable by AQ and AVR* over the entire period.

As I mentioned in the literature review, Urquhart (2016) reported that Bitcoin is becoming less predictable with time. I want to check whether that is true for the entire market of cryptocurrencies, which is well represented by the sample of coins used in this thesis. Using the same methodology as Urquhart, I divided the data of each coin into two equal periods. Then I applied AQ and AVR* to each of those halves, results of the tests are presented in Table 1. It seems that Urquhart's conclusion does not hold for the cryptocurrency market. As you can see, more coins were predictable in the second half of their trading period. For AQ percentage of predictable coins increased from 38% to 48% and for AVR* the percentage increased from 41% to 50%.

	Number of Currencies	Predictable (AQ)	Predictable (AVR*)	% Predictable (AQ)	% Predictable (AVR*)
Entire Period		252	259	54%	55%
1st half	467	177	190	38%	41%
2nd half		222	233	48%	50%

Table 1: Automatic Portmanteau (AQ) and Wild Bootstrapped Automatic Variance Ratio (AVR) Tests results. First, the tests were applied to the entire trading period of each cryptocurrency: the results are presented in the "Entire period" row. Then the data was divided into two equal periods for each cryptocurrency, AQ and AVR* were applied to each of those halves: the results are presented in the 1st half and 2nd half rows. Predictable columns show for many coins H0 was rejected at 5% of significance for a correspondent test. %Predictable columns show the percent of coins that were indicated predictable for a correspondent test.*

To have a clearer picture of results in Table 1, I divided the sample of coins into baskets by capitalization: Basket 1 includes coins with capitalization greater than $100 million; Basket 2 has coins with capitalization between $10 and $100 million; in Basket 3 coin capitalization between $1 and $10 million; Basket 4 has capitalization smaller than $1 million. In Table 2 you can see some summary of statistics of baskets. Currencies in the top basket have the longest average trading period and the currencies in the lowest basket have the shortest. The volume drops significantly, as we go from Basket 1 to Basket 4, while volatility increases. So, the coins with higher capitalization are more mature and (relatively) less volatile, which could imply that they are less predictable than coins in other baskets.

	Basket 1	Basket 2	Basket 3	Basket 4
Number of coins	41	73	146	207
Average Trading period (days)	1075	917.2	948.8	903.4
Average Volume ($)	306 200 000	1 804 186	136 736	92 134
Average Volatility	151%	174%	230%	331%

Table 2: Summary statistics of coins by capitalization baskets on August 24, 2019. Basket 1: coins with capitalization greater than $100 million; Basket 2: coins with capitalization between $10 and $100 million; Basket 3: coins with capitalization between $1 and $10 million; Basket 4: coins with capitalization smaller than $1 million. Average Trading period (days): average number of days within basket since the start of trading until August 24, 2019. Average Volume ($): average daily trading volume within a basket. Average Volatility: average annualized volatility within a basket

In Table 3 you can see test results from Table 1 broken down into capitalization baskets. Looking at “Entire period” rows, one can see that the share of predictable currencies increases as we go from Basket 1 to Basket 4. In Basket 1 only 12% of currencies are predictable according to the AQ test and 24% according to AVR* test, while in Basket 4 whopping 75% of coins are marked predictable by AQ and AVR*. This difference can be explained by the difference in volatility and trading volume among baskets. Additionally, there might be some other factors, like investor attention to a given currency, that affect the difference in predictability among baskets.

If we look at how the level of predictability changed in Table 3, we can observe that in Baskets 1 and 2 fewer coins are predictable in the second period, which is in line with Urquhart (2016). However, when we look at Baskets 3 and 4, we can see that more coins in those baskets are predictable in the second half. In the Appendix, you can take a look at the summary of test results at a 1% level of significance. One could notice that the overall pattern of predictability is similar to one in Table 3: more coins are predictable in the second period in Baskets 3 and 4. That is an interesting result, which implies that small-cap cryptocurrencies behave differently than large-cap cryptocurrencies. As previously, it might be explained by the level of volatility, volume, and investor attention, but I think there is also some inherent difference between small-cap and large-cap cryptocurrencies. I will test that hypothesis in Chapter 5.

	Number of Currencies	Predictable (AQ)	Predictable (AVR*)	% Predictable (AQ)	% Predictable (AVR*)
Basket 1					
Entire Period		5	10	12%	24%
1st half	41	5	8	12%	20%
2nd half		5	5	12%	12%
Basket 2					
Entire Period		19	21	26%	29%
1st half	73	14	18	19%	25%
2nd half		12	13	16%	18%
Basket 3					
Entire Period		73	72	51%	50%
1st half	146	54	56	38%	39%
2nd half		65	67	45%	47%
Basket 4					
Entire Period		155	156	75%	75%
1st half	207	104	108	50%	52%
2nd half		140	148	67%	71%

Table 3: Automatic Portmanteau (AQ) and Wild Bootstrapped Automatic Variance Ratio (AVR) Tests results. Test results are broken down into baskets by capitalization: Basket 1 - coins with capitalization greater than $100 million; Basket 2 - coins with capitalization between $10 and $100 million; Basket 3 - coins with capitalization between $1 and $10 million; Basket 4 - coins with capitalization smaller than $1 million. First, the tests were applied to the entire trading period of each cryptocurrency: the results are presented in the "Entire period" row. Then the data was divided into two equal periods for each cryptocurrency, AQ and AVR* were applied to each of those halves: the results are presented in the 1st half and 2nd half rows. Predictable columns show for many coins H0 was rejected at 5% of significance for a correspondent test. %Predictable columns show the percent of coins that were indicated predictable for a correspondent test.*

It seems that Urquhart's (2016) conclusion is valid at least for mature cryptocurrencies, i.e. their level of predictability decreased with time. However, according to the Adaptive Markets Hypothesis, the return predictability in a financial market is time-varying, which should apply to the cryptocurrency market as well because the market itself is not in a constant state. As we saw in Figure 2, volatility is time-varying; the same applies to volume. In Figure 7, which shows what share of cryptocurrencies in the sample were in a given basket at a given point of time, you can see how dynamic the market is.

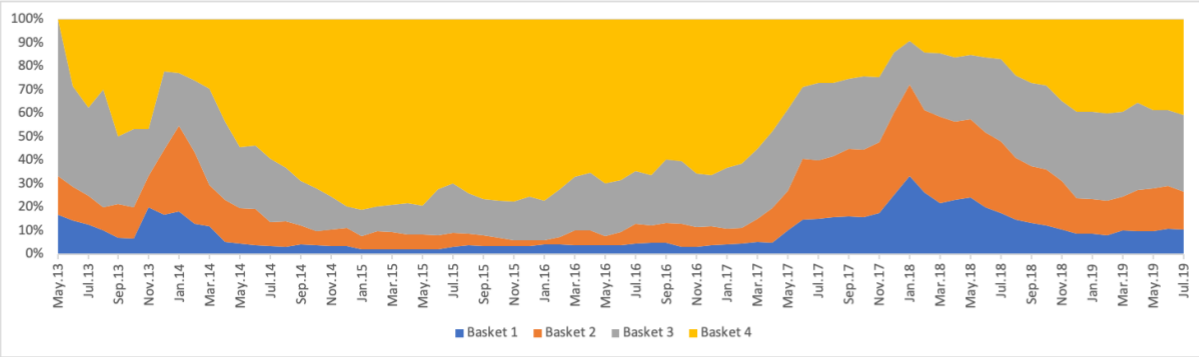

Figure 3: Percentage of cryptocurrencies in a given capitalization basket in a given month. Basket 1 (blue): coins with capitalization greater than $100 million; Basket 2 (orange): coins with capitalization between $10 and $100 million; Basket 3 (grey): coins with capitalization between $1 and $10 million; Basket 4 (yellow): coins with capitalization smaller than $1 million

To test AMH in the cryptocurrency market, I broke down the dataset of each coin into complete months (e.g. January, February, etc.). If a month was incomplete, i.e. days of January, March, etc. ≠ 31; days of April, June, etc. ≠ 30; days of February ≠ 28, then the data for that month would be dropped. For all of the coins first and last months were incomplete since the coins are traded 24/7. That left me left with 466 coins and 13355 monthly periods of data. For each of the periods, I applied AQ and AVR* tests.

Figures 8 and 9 summarize the results of AQ and AVR* tests respectively. The blue bars show the number of coins that were indicated predictable at a 5% level of significance in a given month and the red line shows the percentage of coins that were predictable in a given month. One can observe that the absolute number of predictable coins (blue bar) increased throughout the entire period, but that is attributed to the overall increase in the number of coins in the market.

Looking at the red line, one can see that degree of predictability in the market is time-varying. The period before mid-2014 has too few coins to draw any clear conclusion, but even then we observe that degree of predictability changes from one period to another. Starting from the 2nd quarter of 2014 there is clear cyclicality, i.e. predictability increases for some period (up to 8 months) and then decreases. This confirms that market predictability is time-varying, which is in line with AMH.

Next, one can see that share of predictable coins in the market ranges between 10% and 18% after the 2nd quarter of 2014 up until the start of the 2017 bubble. Then during and after the

2017-2018 bubble the share of predictable coins decreased and was ranging between 5% and 14%. That shows that the 2017-2018 bubble indeed had an effect on the predictability of cryptocurrency returns, as was hypothesized in Chapter 3. I will explore how significant that change was in Chapter 5.

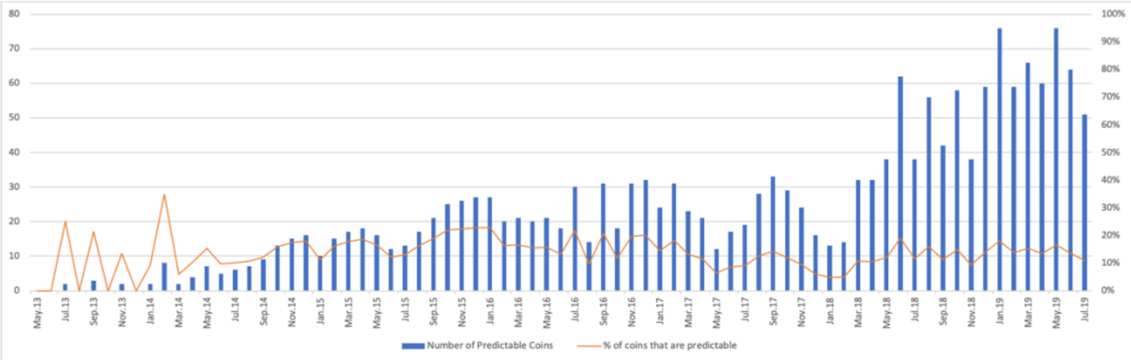

Figure 4: AQ Test results by month for the entire sample of coins. Blue bars show how many coins were indicated predictable by the AQ test in a given month at a 5% level of significance (left axis). The red line shows the percent of coins that were indicated predictable by AQ (right axis).

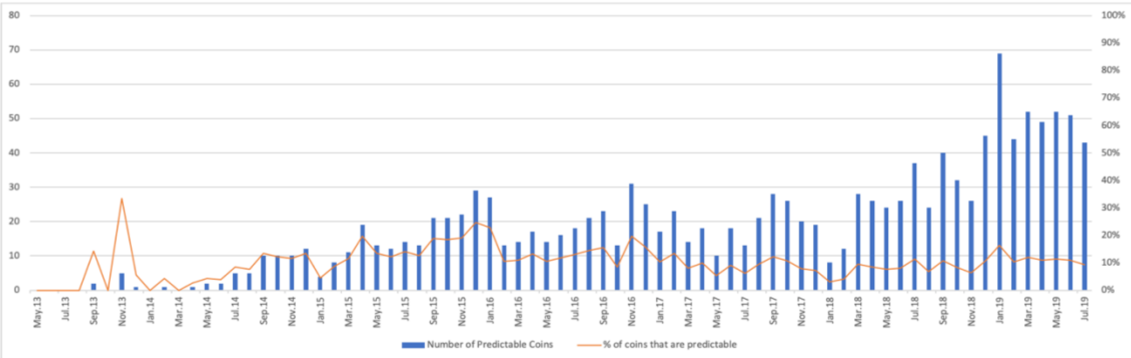

Figure 5: AVR test results by month for the entire sample of coins. Blue bars show how many coins were indicated predictable by the AVR test in a given month at a 5% level of significance (left axis). The red line shows the percent of coins that were indicated predictable by AVR (right axis)*

5. Determinants of predictability

5.1 Regression

As I showed in the previous chapter, the degree of predictability is time-varying. Additionally, the degree of predictability varies among cryptocurrencies too, which was shown by the difference in predictability among capitalization baskets. That is why, in this section, I want to find out what affects this difference, or putting it in another way, I will study determinants of return predictability in the cryptocurrency market. In a similar fashion to Kim et. al. (2011), where they explored determinants of predictability in the US stock market, I will use the AQ test statistic and the absolute value of AVR test statistic as dependent variables. Both indicate the degree of market predictability: the larger (absolute) value of test statistics indicates a higher level of predictability.

I am going to use the fixed effects panel data model because there is a difference in predictability across time and among cryptocurrencies. There are 466 coins in the sample, the dataset of each coin was broken down into complete months (e.g. January, February, etc.), consequently there are 13355 monthly periods. Each coin was labeled with numeric id and each month was labeled with *t*. For each of the periods, I applied AQ and AVR tests. Dependent variables are notated as $AQ_{i,t}$ and $|AVR_{i,t}|$, where *i* is an id of a coin and *t* is a month, for example, January 2015. It should be noted that the panel data is unbalanced since the starting date of data for each coin is different.

First, I would like to see how volatility affects the predictability of returns, that is why annualized volatility of returns of a given coin in a given month ($volatility_{i,t}$) is included in the model. $Volatility_{i,t}$ is calculated as square root of trading days in a year (365 days in the cryptocurrency market) times the standard deviation of returns in a given month.

There is no clear agreement on the effect of volatility on the predictability of returns in the literature. Kim et. al. (2011) and Lim and Kim (2011) report a positive relationship of volatility with the predictability of stock returns, i.e. higher level of volatility is associated with a higher level of predictability. Pesaran and Timmermann (1995) similarly reported that there was no excess return based on forecasting stock returns and trading during calm periods, but during volatile periods there were gains from forecasting. LeBaron (1992), on the other hand, reports that the serial correlation of returns is lower when volatility is higher. Todea and Pleşoianu (2013) also reported that the relation between volatility and predictability is negative. In Chapter 4 I showed that small-cap coins, which have a higher level of volatility, have a higher degree of predictability, that is why I expect the coefficient of $volatility_{i,t}$ to be positive.

The second control variable is the average daily trading volume of a given coin in a given month ($volume_{i,t}$). The relation between volume and predictability is rather straightforward in the literature. Campbell et. al. (1993) theoretically and empirically showed that a higher level of volume leads to a lower degree of predictability, Lim and Kim (2011) and Todea and Pleşoianu (2013) showed the same results. So I expect that increase in $volume_{i,t}$ will be associated with a lower degree of return predictability.

The next independent variable that I wanted to include in the regression was the average market capitalization of a coin in a given month because Lo and Mackinlay (1988) showed that small-cap stocks are more predictable than large-cap stocks. But the inclusion of average capitalization caused a multicollinearity problem since capitalization was highly correlated with volume. That is why I am using dummy variables of capitalization baskets introduced in Chapter 4: $basket1_{i,t}$- capitalization is greater than 100 million; $basket2_{i,t}$ - capitalization is between 10 million and 100 million; $basket3_{i,t}$ - capitalization is between 1 million and 10 million; $basket4_{i,t}$ - capitalization is less than 1 million. In Chapter 4 I showed that there is a difference in predictability among baskets, I hypothesized that it might be caused by a difference in volume and volatility among baskets. The inclusion of a baskets dummy is a perfect opportunity to test if that is true.

Liu and Tsyvinsky (2018) found that Bitcoin, Ripple, and Ethereum do not have exposure to the stock market and that returns can be predicted using Google search volumes. I want to test if their finding applies to other coins. The following variables were included to do so:

- Average daily S&P 500[6] log returns, which were calculated using Close price, for a given month (sp_return_t)
- The average volume of S&P500 in a given month (sp_volume_t)
- Average VIX index[7], which indicates volatility in S&P 500 index, for a given month (vix_t)
- Relative Google search volume of the word “Cryptocurrency”, which takes value from 1 to 100 in a given period ($google_t$)

Additionally, the dummy variables for months (Jan_t, Feb_t, etc.) were included in the model, to test if coins have a different level of predictability in different months.

The functional forms of the models are:

$$AQ_{i,t} = \alpha + \beta_1 volume_{i,t} + \beta_2 volatility_{i,t} + \beta_3 basket2_{i,t} + \beta_4 basket3_{i,t} + basket4_{i,t} + \gamma_1 google_t + \gamma_2 sp\_return_t + \gamma_3 sp\_volume_t + \gamma_4 vix_t + \sum_{j=2}^{12} \eta_{j-1} Month\_j_t + u_{i,t}$$

$$|AVR|_{i,t} = \alpha + \beta_1 volume_{i,t} + \beta_2 volatility_{i,t} + \beta_3 basket2_{i,t} + \beta_4 basket3_{i,t} + \beta_5 basket4_{i,t} + \gamma_1 google_t + \gamma_2 sp\_return_t + \gamma_3 sp\_volume_t + \gamma_4 vix_t + \sum_{j=2}^{12} \eta_{j-1} Month\_j_t + u_{i,t}$$

6 Source for S&P 500 prices and volume: Yahoo Finance
7 Source for VIX index prices: Yahoo Finance

To check if the fixed effects model is the right model for AQ and |AVR|, models with random effects were estimated. Then the Hausman test was used to test if the difference in coefficients between fixed effects and random effects models is systematic. The null hypothesis that the difference in coefficients is not systematic was rejected with a p-value of zero.

At first, the models were estimated with the assumption of homoscedasticity. To test if models suffer from conditional heteroscedasticity, I ran a modified Wald test for groupwise heteroskedasticity in the fixed effect regression model. The null hypothesis of homoscedasticity was rejected for AQ and |AVR| with a p-value of zero. To deal with heteroscedasticity robust standard errors are used.

5.2 Regression results

The results of the full sample regressions with robust standard errors are presented in Table 4. Results for both regressions are similar, which should not come up as a surprise, since both dependent variables indicate the same thing. R_2 is 4.2% and 9.9% for AQ and |AVR| models respectively.

Higher volume in a given period is linked with a lower level of predictability, but the result is significant only in AQ regression. Negative relation of volume and predictability of stock returns was recently documented in Lim and Kim (2011) and Todea and Pleşoianu (2013). These results are consistent with the theoretical model developed by Campbell et. al. (1993).

An increase in volatility of returns is associated with a higher level of predictability in cryptocurrency returns, holding all other variables constant. That means that during more volatile periods returns are more predictable than during less volatile periods. This result is very contra intuitive, but Pesaran and Timmermann (1995), Lim and Kim (2011), and Kim et. al. (2011) found the same relation between predictability and volatility in stock markets.

If a coin were in Basket 4 by capitalization, then its predictability is significantly higher than if it were in other baskets, holding all other variables constant. The regression suggests that there is no significant difference in predictability among baskets one, two, and three. That means that difference in the predictability of returns among Baskets 1, 2, 3 can be explained by other variables, such as volatility and volume. However, control variables cannot explain the higher level of predictability of coins with a capitalization of less than $1 million. So, there must be some inherent characteristic of coins in Basket 4 that is not captured by the model.

As for the stock market variables, S&P Volume is insignificant, but S&P 500 returns variable is positive and significant at a 1% level of significance, meaning that cryptocurrency returns are more predictable when S&P prices are increasing, which might indicate that when the stock market performs well there is less interest in cryptocurrencies, consequently the predictability increases. The VIX variable is significant and positively associated with the predictability of returns too. These results contradict the Liu and Tsyvinsky (2018) results, namely that cryptocurrencies do not have exposure to stock markets.

An increase in Google searches is associated with a lower level of predictability, even when controlling for volume and volatility, so there is a distinct effect of investor (or more likely, general public) attention to the crypto market on the predictability of cryptocurrencies, which is not captured by other variables in the models.

The coefficients of month dummy variables are mixed in AQ and |AVR| models, the only month that has a significantly different level of predictability compared to January in both models is December. The level of predictability is higher in December compared to January.

Since the predictability is not constant in the cryptocurrency market, both across time and individual coins, results reported in Table 4 might help active crypto investors to use forecasting models of cryptocurrencies more efficiently. For example, when a coin exhibits higher volatility and/or the trading volume of a cryptocurrency decreased, there is on average a higher degree of predictability compared to other periods, so a forecasting model or technical trading strategy (hypothetically) should perform better in such periods. Second, small-cap cryptocurrencies have a higher degree of predictability compared to large-cap coins, so traders should pay closer attention to those smaller coins.

	AQ		|AVR|	
	Coefficients	P-Value	Coefficients	P-Value
volume	-2.34E-11	0.01	-2.94E-12	0.28
volatility	0.056	0.01	0.044	0.00
Basket 2	-0.027	0.76	0.001	0.97
Basket 3	0.153	0.23	0.061	0.06
Basket 4	0.515	0.00	0.136	0.00
google	-0.008	0.00	-0.003	0.00
spreturn	57.325	0.00	14.182	0.00
spvolume	1.99E-11	0.87	2.50E-11	0.36
vix	0.023	0.02	0.005	0.03
Feb	-0.088	0.32	-0.043	0.08
Mar	0.138	0.22	0.023	0.33
Apr	-0.045	0.66	-0.013	0.61
May	0.055	0.61	-0.014	0.63
Jun	0.215	0.05	0.016	0.52
Jul	-0.078	0.50	-0.014	0.64
Aug	0.260	0.07	0.042	0.21
Sep	0.180	0.21	0.051	0.09
Oct	0.075	0.49	0.050	0.08
Nov	-0.123	0.26	-0.042	0.10
Dec	0.328	0.01	0.070	0.01
cons	1.097	0.01	0.388	0.00
Number of coins (N)	466		466	
Observations (NT)	13355		13355	

Table 4: Results of regression models with fixed effects, where dependent variables are AQ test statistic and the absolute value of the AVR test statistic. The results are robust to conditional heteroscedasticity.

Next, I want to check if the 2017-2018 bubble had a significant effect on the interaction between predictability and its determinants. As depicted in Figures 8 and 9, there was a change in the degree of predictability in the cryptocurrency market following the 2017-2018 bubble. According to the Adaptive Markets Hypothesis that could lead to change in the interaction of independent variables with the level of predictability. To see how the interaction between predictability and independent variables behaved in corresponding periods, I sliced the data in two periods: 1) before the 2017-2018 bubble (May 2013-March 2017), and 2) during and after the bubble (April 2017-July 2019). The functional form of the models is the same; the only thing that needed adjustment is $google_t$ variable since it is relative to the corresponding period. The Google search volumes for two corresponding periods were used in regressions.

Table 5 shows the regression results for two periods: 1) before the 2017-2018 bubble, 2) during and after the bubble. One can see that coefficients for volume and volatility have the same sign in both periods in AQ and |AVR| models, meaning that interaction between predictability and volume, and predictability and volatility did not change. Nonetheless, volatility had no significant effect on AQ before the bubble, while volume had no significant effect on AVR. So, the effect of volume and volatility on overall predictability before the bubble is not that clear. During and after the bubble, however, both volume and volatility have significant coefficients in AQ and |AVR| models.

With regards to capitalization baskets, there is no significant difference in predictability among baskets one, two, and three in AQ regression in both periods, while being in Basket 4 significantly increases the predictability in both periods. As for the |AVR| model, we can see that there is no significant difference in predictability among all four baskets in the first period, but in the second period, predictability was significantly different in Basket 4.

The surprising result is that Google volume was insignificant with a positive coefficient before the bubble, meaning that an increase in searches lowered the level of predictability, but during and after the bubble the variable is significant with a negative coefficient. So, there was no effect of Google search on the predictability of returns before the bubble, that cannot be explained out by lower absolute volumes of searches before 2017, since the relative search volumes for corresponding periods were used.

The relation between the S&P 500 index and predictability also changed after the bubble. The *AQ* model coefficients of S&P return, volume, and VIX changed their signs during and after the bubble. In |*AVR*| S&P volume and VIX changed their signs. This means that the interaction between S&P500 and the cryptocurrency market changed completely after the crypto bubble in 2017.

We can also see a complete change in the difference in the degree of predictability across months of the year. Before the bubble, (almost) all the coefficients for months have positive signs, meaning that January had the lowest degree of predictability. After 2016 though we see a complete reversal in signs of coefficients, which means that January was the month with the highest degree of predictability. The month effect was also more significant during and after the bubble because more months have p-values lower than 5% in the second period.

So, as you can see, the hypothesis, that there was a major change in the degree of predictability in the cryptocurrency market after 2016, is confirmed. What does that mean in practice? First of all, that might imply that trading strategies, which were profitable before the bubble, could have become not profitable after the year 2016. Second, the forecasting models that were used before the bubble might have lost their predictive power. So, what that means for practitioners is that they should have re-evaluated their trading strategies and forecasting models in the aftermath of the 2017-2018 bubble. However, these claims should be tested in further research.

	AQ: before 2017 bubble		AQ: during and after 2017 bubble		|AVR|: before 2017 bubble		|AVR|: during and after 2017 bubble	
	Coefficients	P-Value	Coefficients	P-Value	Coefficients	P-Value	Coefficients	P-Value
volume	-2.74E-09	0	-3.00E-11	0.09	-2.49E-10	0.57	-1.33E-11	0
volatility	0.009	0.84	0.072	0.07	0.035	0	0.05	0
Basket 2	0.147	0.55	-0.068	0.44	-0.049	0.63	-0.008	0.79
Basket 3	0.369	0.2	0.104	0.42	0.036	0.72	0.037	0.3
Basket 4	0.8	0.01	0.447	0.01	0.13	0.23	0.086	0.04
google	0.001	0.97	-0.013	0	0.001	0.88	-0.004	0
spreturn	37.592	0.4	-45.287	0.14	-1.562	0.85	-4.059	0.52
spvolume	2.64E-10	0.23	-9.02E-10	0	9.98E-11	0.03	-2.72E-10	0
vix	-0.005	0.89	0.02	0.14	-0.002	0.74	0.008	0.01
Feb	0.102	0.55	-0.441	0	-0.044	0.27	-0.052	0.11
Mar	0.191	0.3	-0.392	0.04	0.014	0.73	-0.064	0.14
Apr	0.068	0.74	-0.898	0	0.022	0.67	-0.219	0
May	0.319	0.1	-0.879	0	0.043	0.43	-0.209	0
Jun	0.185	0.36	-0.232	0.19	0.032	0.46	-0.081	0.04
Jul	0.424	0.03	-1.208	0	0.1	0.03	-0.293	0
Aug	0.125	0.48	-0.662	0.03	0.125	0.01	-0.266	0
Sep	0.574	0.03	-0.838	0	0.083	0.06	-0.153	0.01
Oct	0.198	0.24	-0.655	0	0.115	0.01	-0.106	0.04
Nov	0.397	0.06	-0.744	0	0.087	0.07	-0.167	0
Dec	0.602	0.01	-0.175	0.33	0.134	0	0.007	0.87
cons	0.253	0.7	5.233	0	0.208	0.25	1.585	0
Period	May 2013-March 2017		April 2017-July 2019		May 2013-March 2017		April 2017-July 2019	
Number of coins (N)	174		466		174		466	
Observations (NT)	4179		9176		4179		9176	

Table 5: Results of regression models with fixed effects for two periods - 1) before the bubble: May 2013-March 2017, 2) during and after the bubble: April 2017-July 2019. Dependent variables are AQ test statistic and the absolute value of the AVR test statistic in corresponding periods. The results are robust to conditional heteroscedasticity.

6. Conclusion

So, are returns of cryptocurrencies predictable? Yes, but not each coin and not always. I tested the predictability of coins at their entire trading period and showed that around 50% of 467 coins were indicated predictable by the tests at a 5% level of significance. Small-cap coins are more predictable than large-cap coins: around 64% of small-cap cryptocurrencies (capitalization less than $10 million) were indicated predictable by the tests at a 5% level of significance, while only 25% of large-cap cryptocurrencies (capitalization more than $10 million) are predictable at 5% level of significance. Then I tested the predictability of daily returns in a given month to understand how predictability changes throughout time. Overall, results suggest that the degree of predictability in the market is time-varying, i.e. it changes from one period to another. The share of predictable coins in a given month was varying anywhere between 5% and 18% between the years 2013 and 2019.

To understand, what are the determinants of varying degree of predictability, I used panel data regression with fixed effects. Regression models suggest that volatility has significant positive interaction with the degree of predictability, while trading volume and Google search volume have a negative correlation. Models also indicate that there is some interaction between the S&P 500 and the predictability of cryptocurrency returns, despite previous research showing that there is no effect of stock markets on the crypto market. I also showed that the predictability of returns in the cryptocurrency market changed following the cryptocurrency bubble. That is represented by the change in the interaction between predictability and independent variables in the regression models.

There are both theoretical and practical implications of this thesis. From the theoretical perspective, results in Chapters 4 and 5 imply that the return predictability of cryptocurrencies behaves in accordance with the Adaptive Market Hypothesis, which suggests that the degree of predictability in a market is time-varying. Adaptive Market Hypothesis also implies that after some major event in a market there should be a change in behavior of asset prices and/or change in the predictability of their returns, that is what happened following the bubble in the crypto market. From a practical standpoint, the time-varying degree of predictability implies that investors should periodically re-evaluate their trading strategies and forecasting models. While results in Chapter 5 can help to use the trading strategies more efficiently.

There are several limitations to this study that I should point out. First, the sample of cryptocurrencies used suffers from survivorship bias, since I was not able to include coins that failed or turned out to be a fraud. Second, the R^2 of the models is low, so I must have not included some other significant variables that could explain the varying degree of predictability. Third, I did not test some claims that I made in Chapter 5. For example, I speculated that after the bubble, forecasting models, which were used prior to the bubble, should have lost their predictive power and profitable trading strategies could have become not profitable.

In further research on this topic, I would suggest exploring other variables that explain the predictability of returns, since the R^2 of models was relatively low in my case, so there must be some other variables that can explain predictability. Since I could not explain the higher level of predictability of coins with capitalization less than $1 million, I would also suggest having a closer look at their predictability in the future. Next, I would recommend testing if different trading strategies perform better than others in the cryptocurrency market and how their performance is linked to the time-varying degree of predictability. Last but not least, I would check if the performance of forecasting models used in the crypto market prior to 2017 still have predicting power.

References

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of business*, 55-80.

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

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. The journal of Finance, 32(3), 663-682.

Bouri, E., Lau, C. K. M., Lucey, B., & Roubaud, D. (2019). Trading volume and the predictability of return and volatility in the cryptocurrency market. Finance Research Letters, 29, 340-346.

Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American statistical Association, 65(332), 1509-1526.

Campbell, J. Y. (1987). Stock returns and the term structure. Journal of financial economics, 18(2), 373-399.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. The Quarterly Journal of Economics, 108(4), 905-939.

Campbell, J., Lo, A., & MacKinlay, A. (1997). The Econometrics of Financial Markets. Princeton, New Jersey: Princeton University Press. doi:10.2307/j.ctt7skm5

Caporale, G. M., Gil-Alana, L., & Plastun, A. (2018). Persistence in the cryptocurrency market. Research in International Business and Finance, 46, 141-148.

Charles, A., & Darné, O. (2009). Variance-ratio tests of random walk: an overview. Journal of Economic Surveys, 23(3), 503-527.

Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. Journal of Applied Econometrics, 14(3), 293-308.

Chu, J., Zhang, Y., & Chan, S. (2019). The adaptive market hypothesis in the high frequency cryptocurrency market. International Review of Financial Analysis, 64, 221-231.

Corbet, S., Eraslan, V., Lucey, B., & Sensoy, A. (2019). The effectiveness of technical trading rules in cryptocurrency markets. Finance Research Letters, 31, 32-37.

Detzel, A. L., Liu, H., Strauss, J., Zhou, G., & Zhu, Y. (2018). Bitcoin: predictability and profitability via technical analysis. SSRN Electronic Journal.

Escanciano, J. C., & Lobato, I. N. (2009). An automatic portmanteau test for serial correlation. Journal of Econometrics, 151(2), 140-149.

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

Fama, E. (1965b). Random Walks in Stock Market Prices. Financial Analysts Journal, 21(5), 55-59.

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

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. Journal of financial economics, 22(1), 3-25.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. Journal of financial economics, 5(2), 115-146.

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

Hale, G., Krishnamurthy, A., Kudlyak, M., & Shultz, P. (2018). How Future Trading Changed Bitcoin Prices. FRBSF Economic Letter.

Hudson, R., & Urquhart, A. (2019). Technical trading and cryptocurrencies. Annals of Operations Research, 1-30.

Khuntia, S., & Pattanayak, J. K. (2018). Adaptive market hypothesis and evolving predictability of bitcoin. Economics Letters, 167, 26-28.

Kim, J. H. (2009). Automatic variance ratio test under conditional heteroskedasticity. Finance Research Letters, 6(3), 179-185.

Kim, J. H., Shamsuddin, A., & Lim, K. P. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long US data. Journal of Empirical Finance, 18(5), 868-879.

Kyriazis, N. A. (2019). A survey on efficiency and profitable trading opportunities in cryptocurrency markets. Journal of Risk and Financial Management, 12(2), 67.

LeBaron, B. (1992). Some relations between volatility and serial correlations in stock market returns. Journal of Business, 199-219.

LeRoy, S. F. (1973). Risk aversion and the martingale property of stock prices. International Economic Review, 436-446.

Lim, K. P., & Kim, J. H. (2011). Trade openness and the informational efficiency of emerging stock markets. Economic Modelling, 28(5), 2228-2238.

Lim, K. P. & Luo, W. & Kim, J. H. (2013). Are US stock index returns predictable? Evidence from automatic autocorrelation-based tests. Applied Economics. 45. 953-962.

Liu, Y., & Tsyvinski, A. (2018). Risks and returns of cryptocurrency (No. w24877). National Bureau of Economic Research.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. The review of financial studies, 1(1), 41-66.

Lo, A. W. (2004). The adaptive markets hypothesis. The Journal of Portfolio Management, 30(5), 15-29.

Lo, A. W. (2005). Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis. Journal of investment consulting, 7(2), 21-44.

Lobato, I., Nankervis, J. C., & Savin, N. E. (2001). Testing for autocorrelation using a modified box-pierce Q test. International Economic Review, 42(1), 187-205.

Lucas Jr, R. E. (1978). Asset prices in an exchange economy. Econometrica: journal of the Econometric Society, 1429-1445.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. The annals of statistics, 255-285.

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

Nasir, M. A., Huynh, T. L. D., Nguyen, S. P., & Duong, D. (2019). Forecasting cryptocurrency returns and volume using search engines. Financial Innovation, 5(1), 2.

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. The Journal of Finance, 50(4), 1201-1228.

Roll, R. (1988). R2. The Journal of Finance, 43(3), 541-566. doi:10.2307/2328183

Samuelson, P. A. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly. IMR; Industrial Management Review (pre-1986), 6(2), 41.

Todea, A., & Pleşoianu, A. (2013). The influence of foreign portfolio investment on informational efficiency: Empirical evidence from Central and Eastern European stock markets. Economic Modelling, 33, 34-41.

Sovbetov, Y. (2018). Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero. Journal of Economics and Financial Analysis, 2(2), 1-27.

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

Appendix

	Currencies	Predictable (AQ)	Predictable (AVR*)	% Predictable (AQ)	% Predictable (AVR*)
All					
Whole Period		177	194	38%	42%
1st Half	467	116	126	25%	27%
2nd Half		150	155	32%	33%
Basket 1					
Whole Period		2	4	5%	10%
1st Half	41	1	1	2%	2%
2nd Half		0	1	0%	2%
Basket 2					
Whole Period		13	14	18%	19%
1st Half	73	11	10	15%	14%
2nd Half		9	8	12%	11%
Basket 3					
Whole Period		45	52	31%	36%
1st Half	146	37	42	26%	29%
2nd Half		43	39	30%	27%
Basket 4					
Whole Period		117	124	56%	59%
1st Half	207	67	73	32%	35%
2nd Half		98	107	47%	51%

Automatic Portmanteau (AQ) and Wild Bootstrapped Automatic Variance Ratio (AVR) Tests results. Test results are broken down into baskets by capitalization: Basket 1 - coins with capitalization greater than $100 million; Basket 2 - coins with capitalization between $10 and $100 million; Basket 3 - coins with capitalization between $1 and $10 million; Basket 4 - coins with capitalization smaller than $1 million. First, the tests were applied to the entire trading period of each cryptocurrency: the results are presented in the "Entire period" row. Then the data was divided into two equal periods for each cryptocurrency, AQ and AVR* were applied to each of those halves: the results are presented in the 1st half and 2nd half rows. Predictable columns show for many coins H0 was rejected at 5% of significance for a correspondent test. %Predictable columns show the percent of coins that were indicated predictable for a correspondent test.*

Abstract

The goal of this thesis is to study the predictability of returns in the cryptocurrency market. I applied Automatic Portmanteau and Wild Bootstrapped Automatic Variance Ratio tests to daily returns of 467 cryptocurrencies to test if returns are predictable. Test results show that half of the coins in the sample are predictable and that small-cap coins are more predictable than large-cap coins. Further, I show that the degree of predictability in the cryptocurrency market is time-varying and follows the Adaptive Market Hypothesis. To understand, what are the determinants of varying degree of predictability, I used panel data regression with fixed effects, where test statistics of the abovementioned tests are used as dependent variables. Regression models show that trading volume, volatility, and attention to cryptocurrencies can explain some variation of predictability.

Keywords: cryptocurrencies, predictability of returns, determinants of predictability, Adaptive Market Hypothesis

Zusammenfassung

Das Ziel dieser Arbeit ist die Vorhersagbarkeit der Renditen auf dem Markt der Kryptowährung zu studieren. Ich habe Automatic Portmanteau- und Wild Bootstrapped Automatic Variance Ratio-Tests auf tägliche Renditen von 467 Kryptowährungen angewendet, um zu testen, ob Renditen vorhersehbar sind. Die Testergebnisse zeigen, dass die Hälfte der Münzen in der Probe vorhersehbar ist und dass Small-Cap-Münzen vorhersehbarer sind als Large-Cap-Münzen. Ferner zeige ich, dass der Grad der Vorhersagbarkeit auf dem Kryptowährungsmarkt zeitlich variiert und der Adaptive Market Hypothesis folgt. Um zu verstehen, was die Determinanten für einen unterschiedlichen Grad an Vorhersagbarkeit sind, habe ich die Paneldatenregression mit fixen Effekten verwendet, wobei Teststatistiken der oben genannten Tests als abhängige Variablen verwendet werden. Regressionsmodelle zeigen, dass Handelsvolumen, Volatilität und Aufmerksamkeit für Kryptowährungen einige Variationen der Vorhersagbarkeit erklären können.

Schlüsselwörter: Kryptowährung, die Vorhersagbarkeit der Renditen, Determinanten an Vorhersagbarkeit, Adaptive Market Hypothesis