“Are Cryptocurrencies’ returns predictable?”

The role of Momentum, Value, Carry and Investor Sentiment in the Cryptocurrencies World

Postgraduate Dissertation

Carlo Armillis

MSc Finance - Warwick Business School

Supervisor: Professor Daniele Bianchi

05/09/2018

"I declare that this work is entirely my own in accordance with the University's Regulation 11 and the WBS guidelines on plagiarism and collusion. All external references and sources are clearly acknowledged and identified within the contents. No substantial part(s) of the work submitted here has also been submitted by me in other assessments for accredited courses of study, and I acknowledge that if this has been done it may result in me being reported for self-plagiarism and an appropriate reduction in marks may be made when marking this piece of work.”

Acknowledgements

I would first like to express my sincere appreciation to Professor Daniele Bianchi for his support as a teacher, personal advisor, and dissertation supervisor. He consistently allowed this paper to be my own work, but steered me in the right direction whenever I needed it. His dedication to this area was crucial for the selection of my topic which I thoroughly enjoyed researching.

Many thanks to TM Lee, co-founder of Coingecko.com and all the staff working at the platform for providing and allowing me to use their data for this research. Best of luck to all of you.

A heartfelt thanks to my friends for making this year such an unforgettable experience and a special thanks to Macarena Martin and Stefano Murdaca for their patience and invaluable support.

Finally, I would like to express my deepest gratitude to my mother Anna, for her love and her encouragement since the first day I left Rome for my studies. Thank you for teaching me the good things that really matter in life. Thanks again, without you none of this would have been possible.

Abstract

The following study investigates whether there exist factors that have a degree of predictability when predicting the returns of cryptocurrencies. Using a daily dataset on 11 cryptocurrencies, this paper begins with analysing the role of Momentum Value and Carry in this market. It seems that these factors which in past literature have been reported to be effective when predicting traditional asset classes, are also positively associated with variations in cryptocurrencies' prices. Contemporaneously, another analysis is conducted on less conventional factors that are associated with investors sentiment. Evidences are found in favour of Technological Development, Public Interest, and Community. These indicators appear to affect, but evidence has not been found of their relation to weekly and monthly cumulative returns. Concluding, the study shows how investing in a passive portfolio which equally weights every cryptocurrency in the study, would have delivered better risk-adjusted returns than rebalancing the portfolio's weights using Momentum and Carry strategies.

Keywords: Cryptocurrencies, Bitcoin, Predictability, Momentum, Value, Carry, Sentiment, Public Interest, Community, Technological Development.

Table of Contents

Page

Section 1 : Introduction......1

Section 2 : Literature Review......3

Section 3 : Data......5

Subsection 3.1 : Data Description......5

Subsection 3.2 : Descriptive Statistics......7

Section 4 : Methodology......9

Subsection 4.1 : Factors Construction......9

Subsection 4.2 : Investor Sentiment Factors......11

Subsection 4.3 : Principal Component Analysis......12

Subsection 4.4: Predictive Regressions......12

Section 5 : Analysis and Discussion of Results......14

Subsection 5.1 : Predictability Analysis......14

Subsection 5.2 : Portfolios Analysis......17

Section 6: Conclusion......18

Bibliography......20

Appendices......22

List of Tables and Figures

Page

Table 1 : Descriptive Statistics of returns Currency by Currency............................8

Figure 1 : Bitcoin (BTC) Price evolution..8

Table 2 : Descriptive Statistics for variable and factors..9

Table 3 : Mom, Value and Carry Single Regressions..14

Table 4 : Community, Techdevelopment and Public Interest Joint Regression.15

Table 5 : Sentiment Single Regression...16

Table 6 : Mom, Value Carry and Sentiment Joint Regression............................17

Table 7 : Single Portfolios Analysis...18

Table 8 : Blended Portfolios Analysis...18

1. Introduction

Cryptocurrencies are becoming more and more popular since the introduction of Bitcoin in 2009 and are described as a peer-to-peer electronic cash system. One of their most important features is the decentralized nature of the system surrounding them. Hence, there is no controlling authority behind them nor a physical server.

Fama & French´s research (1992) brought to light the existence of some so-called "Factors" that can help in the predictability of conventional asset classes returns. In particular, Momentum, Value, and Size have demonstrated to be among the factors with the most effective predictability power in past researches.

The principal research question of this paper is whether, in the cryptocurrencies market, there also exists factors that can help to predict the returns of such volatile assets.
In particular, this paper will analyse three more conventional factors like Momentum, Value and Carry, alongside three indicators related to investors' sentiment: Technological Development, Community Activity, and Public Interest. While the factors will be formed using the closest definition that these predictors have in the main asset classes, as proposed in other relevant studies, the indicators will be recreated by a principal component analysis due to the diverse nature of their components.
In addition to the three factors and the three indicators, we take into consideration some of the characteristics related with each of the cryptocurrencies in the study, such as fees, active addresses, and median transaction value. The latter variables are included in order to make sure that what we are trying to capture with the analysis is purely driven by the factor chosen and not from other coins' specific characteristic.

As a second research subject, the three portfolios of Momentum, Value and Carry will be analysed in order to determine their performances on an annual basis, and whether or not they deliver enhanced performance on a risk-adjusted basis when compared to a benchmark portfolio. This benchmark portfolio will be created as a passive equally weighted portfolio which invests in every cryptocurrency included in this study. From there, the three strategies will be combined to find out whether there is any blend of the three which would deliver a portfolio with better characteristics.

Predictive Regression Analysis was used in order to address the research questions posed in this project. The analysis concentrates on the degree of predictability of these factors/indicators on the returns of 1 day after and the cumulative 7-days and 30-days returns, using a panel data with random effects (RE). The factors of Momentum, Value and Carry were created following past relevant literature where possible; however, some minor changes were needed due to the nature of the data. For the investor sentiment related factors, a principal component analysis was adopted due to the vast range of data taken into analysis.

The reason behind the choice of these research questions lies within the fact that Cryptocurrencies are still a relatively new phenomenon in the financial world and there are many questions that are yet to be answered. Cryptocurrencies came into our world for the first time in 2009 with the creation of Bitcoin and were firstly traded on an official cryptocurrency stock exchange in 2010. Its popularity increased over the years, but the academic literature on the topic is still quantified as incredibly small when compared to the existing literature on other asset classes. In addition, since cryptocurrencies have a lifespan of less than 10 years, with the majority of those having experienced a relatively low trading activity, the data available might not be considered enough to perform a robust analysis. This means that even just an additional year of data could influence the ultimate findings on this topic also due to the fact that the majority of the coins existent in the market have only a couple of years of life.

Moreover, the literature that tries to find what factors can be associated with price variations of cryptocurrencies or analyses to which extent these factors can actually help when predicting their returns, is very limited.

Among the few, the work of Hubrich, (2017) seems to be the closest to what ultimately is the goal of this paper. In his study, the author analyses the extent to which Momentum, Value, and Carry, are relevant when predicting cryptocurrencies returns both cross-sectionally and longitudinally. His results report a positive and statistically significant association of these factors to cryptocurrencies returns. Our findings for these factors are consistent with what was previously found by Hubrich (2017), reinforcing our hypothesis that Momentum Value and Carry play a significant role in explaining the price variation in the cryptocurrencies market.

Another study conducted by Wang and Vergne (2017), is related to the investor sentiment on these currencies and shows the effects of Technological Development, Public Interest and Community Activity on cryptocurrencies prices. Surprisingly their findings indicated that while technological innovation contributes to enhanced returns, both Public and Community interests are negatively and significantly associated with variances in cryptocurrencies prices. This, however, is not in line with the expectations of this research paper. Our belief is that Community and Public interest should be positively associated with returns especially now that the popularity of the cryptocurrencies is reaching unprecedented peaks. The more people involved and the more "buzz" around these coins should theoretically increase the sentiment that investors feel towards this phenomenon. In line with the hypothesis of this research and contradicting the findings of Wang and Vergne (2017), this study reports that all the three indicators of investor sentiment are positive and statistically significant when predicting the returns of the following day. However, mixed and insignificant signals were reported in the models analysing returns further in the future.

The structure of this paper is organized as follow:

Section 2: *Literature Review.*

In this section, we discussed relevant literature of authors whose subject of study was similar or related to the one of this paper. This will provide the reader with a broad view of previous findings and the methodology that was used, together with the quality of their data. Moreover, previous findings will be compared with one another and with this study in order to address the reason of the discrepancies, if any.

Section 3: *Data.*

This contain every information to address the quality of the data used in the study, clarifying their source, their reliability, the period observed and the motivation over the selection of the observations. In addition, this section includes a statistical description of the returns and the factor included in the study.

Section 4: *Methodology:*

This section contains the statistical methodology behind this study. Firstly, it shows the methods around the construction of the factors and the generation of the indicators and it follows with the rationale behind predictive regressions and the test needed to be performed.

Section 5 : *Results*

This section contains the report, the analysis and the discussion of the results over a statistical and an economical point of view together with the comparison with previous literature.

Section 6 : *Conclusion*

Summary of the findings, implications and limitations of the study

2. Literature Review

There is evidence in the past literature regarding the crypto world that there exist factors or properties that are associated with the predictability of returns of these currencies. An example is provided by Wei (2018), who shows the role of liquidity and the efficiency in the market, reporting evidence that liquidity plays a significant role in the market efficiency of new cryptocurrencies. In particular, he states that this degree of predictability diminishes as liquidity increase. Another study conducted by Liew, Li and Budavari, (2018) show that less volatile cryptocurrencies are more predictable than currencies with higher volatility.

This paper analyses the effect of Momentum, Value and Carry, in addition to Technological development, Public Interest, and Community Interest, towards the predictability of crypto returns and mainly try to recreate the studies conducted by Hubrich (2017) and Wang and Vergne (2017).

Hubrich (2017) findings were in line with our hypothesis of a positive association of Momentum, Value and Carry with cryptocurrencies price variations. Following his findings, Momentum is reported as the best performing individual factor in the analysis, with the other two factors also being relevant at shorter term holding periods. In addition, after demonstrating that these three factors have little correlation among each other, the author constructed a blended multi-factor portfolio, momentum/value/carry, which reported better risk-adjusted returns when compared with individual portfolios, reflecting its high diversification properties. After having conducted our analysis, results report the three factors having coefficients of similar magnitude and strong significance on each of the three windows taken into consideration. However, in our 1-day model, Momentum shows some degree of reversal although being insignificant. Carry performs better than the other two factors, but again, the coefficients are all very close to each other.

A similar study has been conducted by Wang and Vergne (2017), using a panel regression with cryptocurrency fixed effects. The paper analyses whether a "buzz factor", computed as a proxy of investors' sentiment captured by indicators of Public Interest, Negative Publicity, and Community Activity, alongside a factor measuring cryptocurrencies' innovation potential, attract or detract investment in the cryptocurrencies market. His results show that while technological development is positive and significant, public and community interests have a negative association with cryptocurrencies returns. The latter is an interesting finding which has puzzled the authors themselves and is evidence against the hypothesis of our paper which aim to find a positive relationship between those two indicators and the prices of the currencies. Their explanation is that a sudden increase in Public or Community Interest would result in an increase in the volatility of future returns, which would make risk-averse investors less willing to hold a currency given the same expected mean returns. In addition, Negative Publicity has been reported as "not significant" when associated with returns. By using a daily and more complete database, including more cryptocurrencies, this paper aims to address these questions further and contribute to the literature on the topic on factor-based investing in the cryptocurrencies world with a more thorough emphasis on the data used.

Hubrich (2017), used daily data on 11 cryptocurrencies (although some were different from the one in this study), with some coins like Bitcoin an Litecoin having more than four years of observation and others

like Ethereum Classic and Zcash having less than one year of daily observation. This resulted in an unbalanced panel data set which could be inefficient when performing the analysis.
Wang and Vergne (2017) on the other hand, had 255 weekly observation for just five cryptocurrencies, forming a balanced panel dataset.

When selecting data for this study, importance was given to the trade-off between the number of cryptocurrencies and the time horizon of the observation collected. This analysis extends the data to 665 daily observations on 11 cryptocurrencies using a balanced panel dataset, to improve completeness and reliability of the findings.

From the remaining existing literature, we find that the association of momentum with cryptocurrencies is the most studied among the factors of this paper, with vast research such as Rohrbach et al. (2017), Stoffels (2017) and Yang (2018), supporting the hypothesis of momentum being a significant and positive risk-factor in the cryptocurrencies world.

One limitation that has been identified throughout the majority of the recent studies on cryptocurrencies is the scarce availability of the data due to either the length or to the number of currencies included in the study, that could lead to an unstructured datasets and less reliable findings.

The reason for the choice of the three main factors of Momentum, Value, and Carry, lies with the evidence that past literature reported on their effect on traditional asset classes. Momentum is the most widely used among these factors and a numerous number of authors like Jegadeesh and Titman (1993; 1999; 2012), Rouwenhorst (1997), Hong, Lim, Stein (2000), Griffin, Ji and Martin (2003), Moskowitz et. al (2012) and Fama and French (2012) who all found relevance of momentum effect in the stock markets.

Furthermore, the work of scholars such as Jeffers (1967) Hotelling (1957), Joliffe (1992) Campbell, Lo and Mackinlay (1997), and was used in order to perform the principal component analysis and its applications.

3. Data

3.1 Data Description

The data for the study was mainly collected from Coinmetrics.io, where data are pooled together from different exchanges. The raw data were collected on a daily basis for the following 11 cryptocurrencies: BTC, LTC, DOGE, VTC, DGB, DASH, XVG, ETH, PIVX, DCR, and ETC. Although some other currencies were available within our time period, we had to limit our selection to cryptocurrencies that would have allowed us to construct each factor. For example, some cryptocurrencies like Ripple (XRP) were omitted because of the absence of data about generated coins due to their fixed supply protocol.

When collecting the data, importance was given to the trade-off between the time horizon of the data and the number of different assets.
The data was collected in a time frame that goes from 18/08/2016 to 12/07/2018, but after creating variables and factors, the panel data was restricted to a time frame that goes from 16/09/2016 to 12/07/2018. Below is reported a short description of the data used:

- *Price* : denominated in USD, consists of the daily opening price of the currencies.

- *Transaction Volume* : denominated in USD, is the on-chain transaction volume simply reporting the total value of outputs on the blockchain on a given day;

- *Market Capitalization* : denominated in USD, consists of the total market capitalization at the beginning of the day calculated as the unit price in USD multiplied by the number of units in circulation.

- *Generated Coins:* number of newly generated coins on a given day through based on each different cryptocurrency protocol and mining process.

- *Fees* : total amount of fee in a single day, reported in native currency.

- *Active Addresses:* number of unique sending and receiving addresses participating in transactions on the given day.

- *Median Transaction Value:* value calculated over all the transaction made on the given day.

- *Standard Deviation:* simply calculated as the 30-days volatility of the returns.

Furthermore, more sophisticated data were collected from Coingecko.com; the website connect to different APIs of several social media and website and is therefore automatically and instantly updated.

<u>Community</u>

- Twitter Followers on each of the cryptocurrency page
- Reddit Subscribers on each of the cryptocurrency page
- Reddit posts averaged between the given day and the day before, mentioning each cryptocurrency.

- Reddit comments averaged between the given day and the day before, mentioning each cryptocurrency.
- Reddit active accounts mentioning the cryptocurrencies in question averaged between the given day and the day before.

Technological Development

- Developer Forks: number of time a developer takes a copy of the source code, starting its own separate and independent project
- Github Stars: Starring is referred to the Github feature that allows bookmarking a developer repository, which is simply a storage space when you can access the developer projects and its files. This can show an approximate level of interest in the repository.
- Developer Forks: number of times a developer takes a copy of the source code, starting its own separate and independent project. In other words, is a copy of a repository and it allows to adopt changes to the repository without affecting the initial information.
- GitHub Subscribers: number of subscribers on issues or pull request.
- GitHub developer Total issues: number of total issues raised by the community about the code
- GitHub developer Closed Issues: number of Total Issues fixed by the developers
- Pull Requests merged: number of proposals merged in the core codebase
- Pull requests contributors: number of unique collaborators contributing codes to the project.

Public Interest

- Alexa Rank is the rank of each cryptocurrency's official website, based on the overall website traffic over the internet, for each of the days in our sample.
- Bing searches of a given cryptocurrency on a given day.

3.2 Descriptive Statistics

Cryptocurrencies

Table 1 below reports some descriptive data on each of the 11 cryptocurrencies' daily returns analysed in the study. It is important to show how these cryptocurrencies have significant differences in average return, volatility and especially the min/max value of their daily returns. From the table it can be observed how the cryptocurrency Verge (XVG) has the highest daily average returns, highest standard deviation and the second highest maximum value among all the returns, while also having the lowest drawback of them all.

Table 1

	Mean	St.dev	Skewness	Kurtosis	Min	Max
BTC	0.5%	4.6%	0.28	6.95	-18%	25%
LTC	0.7%	7.4%	2.43	19.38	-32%	68%
DOGE	0.7%	8.2%	1.62	14.35	-38%	62%
VTC	1.0%	10.9%	2.15	17.07	-32%	99%
DGB	1.4%	13.7%	6.69	96.96	-33%	217%
DASH	0.7%	6.9%	1.45	10.17	-22%	46%
XVG	2.8%	20.7%	2.66	17.10	-50%	160%
ETH	0.7%	6.4%	0.90	6.70	-24%	33%
PIVX	1.6%	12.2%	2.06	14.44	-32%	102%
DCR	1.0%	9.6%	1.54	8.07	-27%	57%
ETC	0.7%	7.9%	1.16	10.7	-35%	60%

Descriptive Statistics on daily returns from 12/06/2016 to 11/07/2018 (665 Observations per currency)

During last December only, it gained the astonishing returns of around 900% in just one week. On the opposite side, we have Bitcoin (BTC) who reports the lowest volatility among the 11 cryptocurrencies. However, the latter is still considered a risky investment when compared to traditional assets. Figure 1 shows the price evolution of Bitcoin during the time span of the study and Appendix 1 reports the evolution of the prices for the remaining coins. In the time span of three months, from the 17/09/2017 to 17/12/2017, date in which Bitcoin find its "all-time high" the currency increased 440% before collapsing to the value it has today. Due to the highly correlated nature of the cryptocurrencies, similar increase and drawbacks were reported in the same periods. Appendix 2 shows a correlation matrix of the 11 cryptocurrencies in this study.

Figure 1

Price of Bitcoin (BTC) between 16/09/2016 and 12/07/2018

Variables/Factors

Table 2 show a summary of the statistics for each of the factors/variables in our study. Data on the variable "Returns" of our Panel data report an expected return of 1.05% daily and a 10.8% Standard Deviation.

Our momentum, value and carry portfolios have similar volatility but a higher average return of 1.33% and 2.49% for the first two and 0.79% for the last. The indicators of Community Technological development and Public Interest have been created with a Principal component analysis, and show a mean of zero and a relatively high standard deviation/variance. Same applies for Sentiment which has been formed following another PCA on the previous three indicators. Pooled returns of the 11 cryptocurrencies analysed appear to be not normally distributed with skewness of 3.95 and kurtosis of 51.35. The latter results are consistent with the findings in the study conducted by Chan, Chu, Nadarajah and Osterrieder (2017) which also shows how the returns in the cryptocurrencies world clearly deviate from normality.

Table 2

	Mean	*St.dev*	*Skewness*	*Kurtosis*	*Min*	*Max*	*Obs.*
Returns	1.05%	0.108	3.94	51.35	-50%	217%	7315
Momentum	1.33%	0.112	2.16	15.35	-33%	87%	7315
Value	2.49%	0.110	2.42	15.09	-25%	84%	7315
Carry	0.79%	0.100	0.36	8.30	-48%	68%	7315
Sentiment	0%	0.894	1.31	4.67	-3.21	5.53	6584
Community	0%	0.78	0.77	5.37	-4.75	4.31	6866
Tech.Develop.	0%	0.42	-0.09	5.44	-1.67	1.39	7003
Public Int.	0%	0.90	3.09	14.74	-0.59	8.47	7315

Descriptive Statistics for variables and factors in the study.

From the table above, it can also be observed how pooled daily returns go from a minimum of negative 50% to a maximum of 216%. Those figures are just outliers, but it can be seen from the QQ plots in Appendix 3 that the number of those outliers is considerably high. Additionally, 120 times the positive returns are equal or exceed 30%, and 27 times the negative return are worse or equal to -30% on a single day. However, excluding XVG from this count, those numbers become 73 and 11 respectively. Appendix 4 shows a more detailed analysis of this matter.

4. Methodology

4.1 Factors Construction

Momentum: Momentum is defined as one of the strongest and most analysed asset pricing anomalies. It follows the belief that past winners will keep outperforming past losers in the next trading period. Momentum is widely used in various past literature and there is evidence of several different ways to construct the factor. In our case, the factor construction will consist of a 30-days look back period with only a 2-days holding period, contrarily to other literature which tends to use a week or a month, due to the rapid reversal nature of the cryptocurrencies market. The first step taken was ranking the cumulative past month returns of 11 cryptocurrencies, and forming a zero-cost portfolio which longs the top 5 cryptocurrencies and shorts the bottom 5 cryptocurrencies, not including the median ranked value. The weights were allocated proportionally to their rank such that the top-ranked cryptocurrency will have the same weight but opposite sign of the lowest ranked, and so on. After constructing our portfolio, the performance that the latter would achieve two days after that date was calculated, to avoid short-term reversal which is commonly the case in the cryptocurrencies world. The returns of this portfolio would finally be our momentum factor.

Value: Traditionally, the value factor in past literature has been constructed to capture excess returns to stocks that have low prices relative to their fundamental value. For this reason, several measures have been used as B/M, P/E or more in general, a valuation of their cash flows. In our case, cryptocurrencies do not produce any sort of cash flows and therefore would be impossible to replicate such literatures exactly. In order to assess the problem, we adopt a notion of "value" already used by Hubrich (2017). He believes that the $-value of on chain transaction volume of a cryptocurrency can be seen as a good proxy for economic activity within the currency's blockchain and its product or service. Reflecting those considerations, the raw value metric was constructed as the ratio of current market capitalization and the trailing 7-day average on-chain transaction volume in USD of for each of the currencies. Before ranking the signals, the values were de-meaned to get more accurate estimates, and after ranking those value a portfolio was formed. These portfolios will be re-balanced on a daily basis and is generated in such a way that the currency with the lowest Market/Volume ratio would be top-ranked and so on, once again creating a zero-cost portfolio that goes long on the top 5 ranked currencies and short on the bottom 5. The performance of this portfolio on the following day will form our value factor.

Carry: All the cryptocurrencies in this study, generate new coins on a daily basis, and the amount of those new coins differs between each currency due to differences in the protocols of each blockchain and their mining process. This carry factor will reflect a proxy for the measure of "inflation", given by the number of new coins generated. It is reasonable to believe that as new coins will be generated, the total supply will increase and the currency would lose its nominal value. However, in order to give a meaning to this factor one must assume that the demand for a currency does not change. In addition, "inflation" would by definition be positive, since the protocol of each blockchain does not allow to reduce the number of coins already in existence*. Following the methodology reported by (Hubrich, 2017), this paper will define "high carry" as a low coin issuance reported by a cryptocurrency. Reflecting the abovementioned assumptions, the carry factor is constructed as the negative of the sum of total coin generated over the previous 7 days, divided by the total coin outstanding at the beginning of that 7-day period. Here, the signals are once again ranked and the portfolio is formed in the same way as for momentum and value.

* Sometimes coins are "burned" upon decision of the developer but we decided not to consider this as a potential disruption for our studies due to unavailability of data. However, this is of an unusual nature for the currencies included in this study.

For the weighting and normalization of the three factors, the procedure adopted by Bianchi (2018), was followed:

$$Weight_{it} = z\left(rank(S_{it}) - \frac{N+1}{2}\right)$$

Where:

- i is the cross-sectional dimension and t is the time-series dimension;
- S is the signal associated to the factor;
- N is the number of cryptocurrencies;

Portfolios for each signal will be constructed allowing negative weights. It is known how cryptocurrencies cannot be shorted yet in current markets and this would make these strategies not implementable in the real world. However, the idea of is that these strategies would capture risk factor and help to explain the dynamics of the returns in the cryptocurrencies world.

Control Variables

In order to enhance the reliability of our results and minimize a potential omitted variable bias, we include Fee, Active Addresses, Median Transaction Value and Standard Deviation as control variables, therefore allowing the results to report with more accuracy the effect of our main independent variables on the dependent variable.

4.2 Investor Sentiment Factors

This study adopted the same raw data as Wang and Vergne (2017) did in their study about the "buzz" surrounding cryptocurrencies. Yet, due to their confidentiality agreement with Coingecko.com, accurate detail of the weighting scheme of each of those factors could not be shared. However, common knowledge is that Coingecko.com weighted each of the indicators relative to their importance and more weight was given to indicators that would be difficult to manipulate. This paper tries to recreate the following factors based on the data collected and the knowledge of the cryptocurrencies world.

Public interest includes the Alexa rank for the official coin website and web searches on Bing, with the aim to recreate the general popularity of the coin;

Community includes discussions and popularity of a coin from social media and forum content on Reddit, Facebook, and Twitter. Its importance lays in the belief that Cryptocurrencies tend to grow tremendously when there is a strong and active community backing them;

Technological Development includes development activity on public source code repositories on Git Hub and Bitbucket. Some coins are no longer maintained by the developer and would appear very unlikely to progress over time without a constant innovation. This would allow us to identify "Shell" cryptocurrencies to which no development was made to their original code, (Wang and Vergne, 2017).

4.3 Principal Components Analysis

In order to select an accurate weighting scheme for each of the three above indicators, this paper performs a Principal Components Analysis (PCA).

The aim of this analysis is the reduction of the dimensions of the observation space in which the variables are studied, without losing excessive information in the covariance matrix, as shown by Campbell, Lo and Mackinlay (1997). In practice, this method determines the linear combinations with the maximum variance among the set of initial variables, taking this component as a new variable. This is needed because of the different scales that the initial parameters have in addition to their substantial difference in mean and variance despite the fact of being highly correlated information. This method would not deliver only one principal component, but it will report as many components as there are variables in the analysis. In order to select which components to take into consideration this study base the choice both on their eigen-values and their cumulative explanatory power. This paper will only take into consideration components that have a cumulative explanatory power which is greater than 80%, with a eigen-value of 1 or higher. Due to the high similarity in the data and their correlation, in each of the three cases is the second principal component to meet these requirements and these are therefore used to create indicators for Community Technological Development and Public Interest.

To give a better understanding of the second principal component, it is important to underline the following statement:

"The second principal component is the normalized linear combination with maximum variance of all combinations orthogonal to the first principal component", (Campbell, Lo and Mackinlay 1997, p.236).

Subsequently, an additional PCA is performed with these three indicators as inputs, and the same methodology is used to form an overall Sentiment indicator.

4.4 Predictive Regressions

The analysis started with a panel regression with Fixed Effects (FE) in order to control for variables that are not observable, accounting for individual heterogeneity. FE would enable to remove the effect of time-invariant characteristics giving the possibility to capture the true effect of the predictors on the dependent variable. However, a Hausman test was performed in order to approve the adoption of FE in our analysis against the H_0 of Random effect (RE) adoption as a preferred model as suggested by Wooldridge (2013). The estimates of both models have been stored and used to perform the test, which essentially tests the correlations of the unique errors with the regressors.

Hausman Test:

$$H_0: errors\ are\ not\ correlated\ with\ the\ regressors$$

$$H_1: errors\ are\ correlated\ with\ the\ regressors$$

The failed to reject the null which led to the RE adoption and from here the analysis proceeded with the specified model.
The rationale behind RE model is that the variation across entities is assumed to be random and uncorrelated with the independent variables included in the model, Greene (2007).

Panel Regression with Random Effects

The regressions will be estimated as follows:

$$r_{i,t} = \alpha + \sum_k \beta_k x_{i,k,t-1} + u_{it} + \varepsilon_{i,t}$$

Where:

- i is the cross-sectional dimension and t is the time-series dimension
- $x_{i,k,t-1}$ is k^{th} predictor for i^{th} currency, lagged by one period except for some of the characteristic which we will specify upon implementation.
- u is the between the entity error
- $\varepsilon_{i,t}$ is the within entity error;

White standard errors (robust), will be used to have heteroscedasticity-consistent t-statistics as suggested by White, (1980).

These predictive regressions will then be performed on returns of three different windows. The first model will analyse the predictability power on the 1-day ahead returns, and the other two will examine the degree of predictability on 7-days cumulative and 30-days cumulative returns. This procedure has been chosen because factors might not show evidence of association with the returns of the following day but they might be significant when regressed on aggregate returns.

5. Analysis and Discussion of Results

5.1 Predictability Analysis

In order to enrich the analysis, the regressions were first performed with each factor separately and then were input together in supplementary regressions. Starting with the three main factors Momentum, Value and Carry, the study shows models explaining 1-day, 7-days and 30-days ahead cumulative returns. Table 3 shows the results of the models in which the factors are regressed one by one on the different step ahead returns. Taken singularly, Value and Carry result significant and with a positive coefficient when explaining each of the three steps and in particular they seem to work better the further in time the models try to predict returns. Same reasoning for the Momentum factor except for the 1-day model for which the coefficient resulted negative although insignificant, suggesting a possible reversal in the cryptocurrencies returns. However, the latter seems to disappear when regressing the three factors together although remaining non-significantly different from zero (see appendix 5).

Table 3

Mom Value and Carry single regressions

	1-Day	7-Days	30-Days
Mom	-0.0005 (0.0110)	0.4088** (0.1798)	0.6380** (0.2504)
Value	0.1939** (0.0808)	0.4252*** (0.1399)	1.8394*** (0.5836)
Carry	0.2904*** (0.0738)	0.5437*** (0.1010)	1.5869*** (0.5268)
Obs.	7304	7238	6985

*Where ***= p<1%, **=p< 5% *=p< 10%*

The findings of this study for Momentum Value and Carry are in line with those of the similar research done by Hubrich (2017) who found all three factors playing an important role in "Cryptoland". However, he reports Momentum to have the strongest explanatory power among the three while our analysis found this spot to be held by Carry in the 1-day and 7-days models and by Value in the 30-days model. Differences with his results could be due to different time periods or to the fact that in the analysis reported by Hubrich (2017), weekly returns were used as raw data, hence all the calculation and factor construction were made upon those observations. In addition, some minor changes were applied to the factor construction methodology but being these differences being only minor, we found our results to be very close to his.

Moving onto the analysis of the investor's sentiment indicators, the study shows the importance of Community Activity, Technological Development, and Public Interest, in attracting investors towards the cryptocurrency market. Table 4 below shows the results of regressions performed using the three indicators and the same control variables adopted for the previous factors.

Table 4

Community Techdevelopment and Public Interest Joint Regression

	1-Day	7-Days	30-Days
Community	0.0046** (0.0019)	0.0126 (0.0100)	-0.1474 (0.1111)
Techdevelopment	0.0013* (0.0008)	-0.0222 (0.0158)	-0.0667 (0.0944)
Public Interest	0.0030** (0.0014)	0.0050 (0.0168)	0.2067 (0.2089)
Obs.	6577	6511	6258

*Where ***= p<1%, **=p< 5% *=p< 10%*

For the three indicators, the coefficients show a significant and positive association with the variations in prices of cryptocurrencies only when trying to predict the 1-day returns. The findings show that an increase in each of indicators of Community Activity, Technological Development, and Public Interest does indeed affect the returns of cryptocurrencies, but only when analysing the returns of the day after the increase in the value of the indicators is reported. When we take into account the cumulative 7 days after the increase in the indicators, Technological Development turns negative with Community and Public Interest maintaining a positive coefficient. One could suppose that, with the everchanging technology behind these coins, an issue fixed today would not influence returns of a week after the issue was closed, or perhaps would not take into account coding issues or downgrades that happened between that day and the 7 days after. However, none of the three coefficients is statistically significant therefore these findings are not considered reliable. Similar reasoning is applied for the 30-step ahead model; The coefficients turn out to be non-significant therefore any interpretation here would again be misleading.

Wang and Vergne (2017), proposed a similar study, concentrating on Technological Development, Public Interest, Liquidity, Negative Publicity and Community Interest.

When comparing the findings of this study with theirs, consistency with the latter is found only for Technological Development, showing evidence that an increase in Technological Development does attract investors and is therefore associated with an increase in returns of cryptocurrencies. However, the model analysed by Wang and Vergne (2017), report Public Interest to be negatively associated with

Cryptocurrencies returns. They claim that greater visibility in the society would not enhance the returns of these coins but instead would lower them. Contrarily, our findings support the idea that the "buzz" around cryptocurrencies, created by media and advertisements, does indeed contribute to higher returns.

Once again, the discrepancies in findings can be explained by different dataset (number of currency or timespan) different timespan of observations (Daily vs Weekly) or different indicators construction although having the same source for the raw data.

Furthermore, the study tries to create an overall proxy for investor sentiment which would include the three indicators above, and a further principal component analysis was applied. We called this indicator, Sentiment. Table 5 below shows an additional predictive regression with the Sentiment indicators and the usual control variables of *Fee, Active Addresses, Median Transaction Value, and Standard deviation.*

Table 5

Sentiment single Regression

	1-Day	7-Days	30-Days
Sentiment	0.0033*** (0.0012)	-0.0010 (0.0101)	-0.1112 (0.0898)
Obs.	6577	6511	6258

*Where ***= p<1%, **=p< 5% *=p< 10%*

The Sentiment factor having components from each of the three indicators has itself positive and significant coefficient when explaining the 1-day ahead returns. This means that an overall increase in Community, Technological Development and Public Interest is associated with enhanced returns only one day after that increase has happened

The Sentiment factor having components from each of the three indicators, has itself positive and significant coefficient when explaining the 1-day ahead returns. This means that an overall increase in Community, Technological Development, and Public Interest is associated with enhanced returns only one day after that increase has happened. Once again, the 7-step ahead and the 30-step ahead models did not show any statistical significance. This confirms how these indicators are not helpful when trying to predict cumulative returns far ahead in time.

Additionally, a predictive regression was performed by adding the variable Sentiment to our first basic model with Momentum Value and Carry. The differences here are small when compared to the basic model (Table 3) and the single Sentiment model (Table 5). Table 6 shows the regression results of our final (Full) model.

Table 6

Mom, Value, Carry and Sentiment Joint Regression

	1-Day	7-Days	30-Days
Momentum	0.01274 (0.0112)	0.4262** (0.1840)	0.7133** (0.3067)
Value	0.1800** (0.0860)	0.4012*** (0.1410)	1.8059*** (0.6058)
Carry	0.2760*** (0.0776)	0.5545*** (0.1035)	1.6202*** (0.5939)
Sentiment	0.0028*** (0.0008)	-0.0053 (0.0129)	-0.1411 (0.1087)
Obs.	6577	6511	6258

*Where ***= p<1%, **=p< 5% *=p< 10%*

While the Sentiment Factors exhibited a small decrease in coefficients for each of the windows, the Momentum factor in the one-day model switch from negative to positive although remaining insignificant. Perhaps here the model could be trying to tell that the reversal effect of momentum is picked up by our Sentiment variable. The results also report slightly enhanced coefficients for 7-step ahead and the 30-step ahead models both for Momentum and Value, maintaining the same degree of significance. Surprisingly the variable Carry reports small decreases in coefficient probably due to the "Inflation" or generation of coins being related with one of the components of Sentiment, Technological Development. However, all the results maintain the significance they showed in the previous models and reports only minimal changes in coefficients.

5.2 Portfolios Analysis

In this part of the study, a step away was taken from the primary aim of the paper in order to analyse the three factor's portfolios created, Momentum, Value and Carry.

Firstly, the three portfolios were compared to a benchmark passive portfolio which was calculated by equally weighting the 11 cryptocurrencies throughout the entire period under review. The portfolios were ranked based on their risk-adjusted returns assuming that no returns would be generated when investing in a risk-free asset. The Value portfolio generates an annualised return of 908% that in "crypto jargon" correspond to a yearly "9x". Although the passive portfolio has a lower annualised volatility, the latter is not low enough to compensate for the excess return that the value portfolio would generate. However, the Passive portfolio performs better on a risk-adjusted return basis than the Momentum and the Carry portfolios. The returns are annualised by multiplying the daily figures of expected returns by the number of days in a year (365). The volatility is annualised by dividing it by the square root of the sample period and then multiplying it by the square root of 365.

Table 7 exhibits the annualised expected return of the four portfolios and their annualised volatility together with their ranking based on their risk-adjusted ratios.

Table 7

Single Portfolios Analysis

	Ann. E[r]	Ann.σ	Risk-Adj Ratio	Rank
Passive Portf.	384%	4.42%	86.93	2
Mom	486%	8.30%	58.61	3
Value	908%	8.17%	111.34	1
Carry	289%	7.42%	39.00	4

The risk-adjusted ratios are simply calculated as a sort of sharp ratio, assuming that the risk-free asset is zero. The value of the ratios is only assumed to be used to rank the different portfolios among each other, and it should not be considered as a real proxy for performance.

Successively we want to see if any combination of the Momentum, Value and Carry portfolios could beat the single risk-adjusted level of any single portfolio and we proceed in constructing blended portfolios with our three factors. Table 9 shows the results obtained by a battery of blended portfolios. When allocating equal weight to the three factors (A), the portfolio shows a similar risk-adjusted ratio provided by investing in the Value portfolio alone. In light of these findings, the study decided to create a portfolio which is more biased towards Value with the aim of getting an enhanced performance. A portfolio of portfolios that invest 2/3 in Value, 1/6 in Momentum and 1/6 in Carry (B), is found to deliver a risk-adjusted ratio of 122.38 which exceed the single Value portfolio. Subsequently, a constrained optimization was performed to find the allocation of weights which would maximize the efficiency of a portfolio. The portfolio obtained resulted having an allocation of weights extremely close to the one in portfolio B, with Momentum (24%), Value (57%) and Carry (19%), delivering a ratio of 123.73, and we denominated this portfolio (C).

Table 8

Blended Portfolios Analysis

	Ann. E[r]	Ann.σ	Risk-Adj Ratio	Rank
A	493%	4.44%	111.01	3
B	700%	5.72%	122.38	2
C	641%	5.18%	123.73	1

Where :

(A) is an Equally weighted portfolio of Momentum, Value and Carry;

(B) is a portfolio which invests 2/3 in Value, 1/6 in Momentum and 1/6 in Carry;

(C) has the following weights: Mom (24%), Value (57%) and Carry (19%) given by a constrained optimization;

6. Conclusion

This research project analysed 11 cryptocurrencies covering a time horizon of approximately two years. The analysis clearly shows how Momentum, Value and Carry play a significant role when trying to predict the variations in cryptocurrencies prices. Using a panel data predictive regression analysis, with Random effects, regressions on three different windows of 1-day, 7-day, and 30-day cumulative returns were performed. All of the three above mentioned factors resulted as a positive and significant predictor and these results were in line with those of past literature on the topic. Moreover, the analysis proceeded with examining the effect of more unconventional indicators present in the cryptocurrencies market: Technological Development, Community Activity, and Public Interest. Same process was applied to these predictive regressions which delivered evidence regarding the association of these indicators with the coins they belong to, and these evidences were in line with our initial hypothesis. If one would think of cryptocurrencies as technologies, it is clear how the improvement behind them should be crucial for the coins itself. Newly issued coins might be considered as a vanguard for the market in terms of the technology they adopt, but if these coins are not continuously developed, they could lose the advantage they had in the first place. As expected Technological Development has proven to be positive and statistically significant when predicting the one-day ahead returns. However, the coefficients for the 7 and the 30-step ahead models did not show evidence of statistical significance. This result is not surprising considering how fast these technologies evolve and any issue fixed today might recur in other forms in the same week or month and new problems are encountered on a daily basis.

Furthermore, the analysis found Community Activity and Public Interest to deliver similar results. They are both positive and significant in the one-day ahead model but results insignificant if regressed on the cumulative returns of a week and or month ahead. Here again, social media and the internet play an important role of diffusion in the society and their coefficient were expected to behave accordingly. While the origination of a "hype" behind these cryptocurrencies could benefit them in the short term, they might already be "old news" after a week or a month, due to the fast and everchanging environment of cryptocurrencies.

Concluding, the study shows how a passive portfolio which invests an equal weight in every currency in the study, performs better on risk-adjusted return basis than some of the portfolios initially created. This might be due to the fact that cryptocurrencies experienced altogether a "boom" in the recent period which saw the prices of these coins increase massively. In addition, creating blended strategies with Momentum Value and Carry, will enhance the performance of the portfolio due to diversification benefits.

If time would have allowed us, we could have taken a step further and separately analysed the market behaviour before and after the cryptocurrencies "boom", and possibly the extent to which fraudulent news and negative publicity affect this market.

All in all, the researcher hypothesises that there might exist unidentified external factors other than the ones analysed in this study, that play a role in the crypto world, some of which had generated and will generate anomalies in the market. Unfortunately, these factors have yet to be identified but we are only at the beginning of this phenomenon and there is still room for improvement and for new "mechanisms" to be discovered.

Bibliography

Campbell, J., Lo, A. and MacKinlay, A. (1997). *Econometrics of Financial Markets*. 2nd ed. New Jersey: Princeton University Press, pp.236-239.

Chan, S., Chu, J., Nadarajah, S. and Osterrieder, J. (2017). A Statistical Analysis of Cryptocurrencies. *Journal of Risk and Financial Management*, 10(2).

Coin Metrics. (2018). *Data downloads - Coin Metrics*. [online] Available at: https://coinmetrics.io/data-downloads/ [Accessed 12 Jul. 2018].

Coin Metrics. (2018). *Network Value to Transactions Ratio - Coin Metrics*. [online] Available at: https://coinmetrics.io/nvt/#assets=btc [Accessed 7 Jul. 2018].

CoinGecko. (2018). *CoinGecko: 360 Degree Overview of Cryptocurrencies Chart*. [online] Available at: https://www.coingecko.com/en [Accessed 1 Jul. 2018].

Coingecko.com. (2018). *CoinGecko API (V3)*. [online] Available at: https://www.coingecko.com/api/docs/v3 [Accessed 15 Jul. 2018].

Fama, E. and French, K. (1992). 'The Cross-Section of Expected Stock Returns'. *The Journal of Finance*, 47(2).

Greene, W. (2007). Fixed and Random Effects Models for Count Data. *SSRN Electronic Journal*.

Griffin, J., Ji, X. and Martin, J. (2003). Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole. *The Journal of Finance*, 58(6), pp.2515-2547.

Hong, H., Lim, T. and Stein, J. (2000). Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *The Journal of Finance*, 55(1), pp.265-295.

Hotelling, H. (1957). THE RELATIONS OF THE NEWER MULTIVARIATE STATISTICAL METHODS TO FACTOR ANALYSIS. *British Journal of Statistical Psychology*, 10(2), pp.69-79.

Hubrich, S. (2017). 'Know When to Hodl Em, Know When to Fodl Em': An Investigation of Factor Based Investing in the Cryptocurrency Space. *SSRN Electronic Journal*.

Jeffers, J. (1967). Two Case Studies in the Application of Principal Component Analysis. *Applied Statistics*, 16(3).

Jegadeesh, N. and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), pp.65-91.

Jegadeesh, N. and Titman, S. (2002). Momentum. *SSRN Electronic Journal*.

Joliffe, I. (1992). Principal component analysis and exploratory factor analysis. *Statistical Methods in Medical Research*, 1(1), pp.69-95.

Moskowitz, T., Ooi, Y. and Pedersen, L. (2011). Time Series Momentum. *SSRN Electronic Journal*.

Rohrbach, J., Suremann, S. and Osterrieder, J. (2017). *Momentum and trend following trading strategies for currencies and bitcoin*. Zurich: University of Applied Sciences, School of Engineering.

Rouwenhorst, K. (1997). International Momentum Strategies. *SSRN Electronic Journal*. Yale ICF Working Paper

Stoffels, J. (2017). *Asset Pricing of Cryptocurrencies and Momentum based Patterns*. Rotterdam: Erasmus School of Economics.

Titman, S. and Jegadeesh, N. (1999). Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *SSRN Electronic Journal*.

Torres-Reyna, O. (2007). *Panel Data Analysis Fixed and Random Effects*. [online] Dss.princeton.edu. Available at: http://dss.princeton.edu/training/ [Accessed 5 Jul. 2018].

Wang, S. and Vergne, J. (2017). Buzz Factor or Innovation Potential: What Explains Cryptocurrencies' Returns?. *PLOS ONE*.

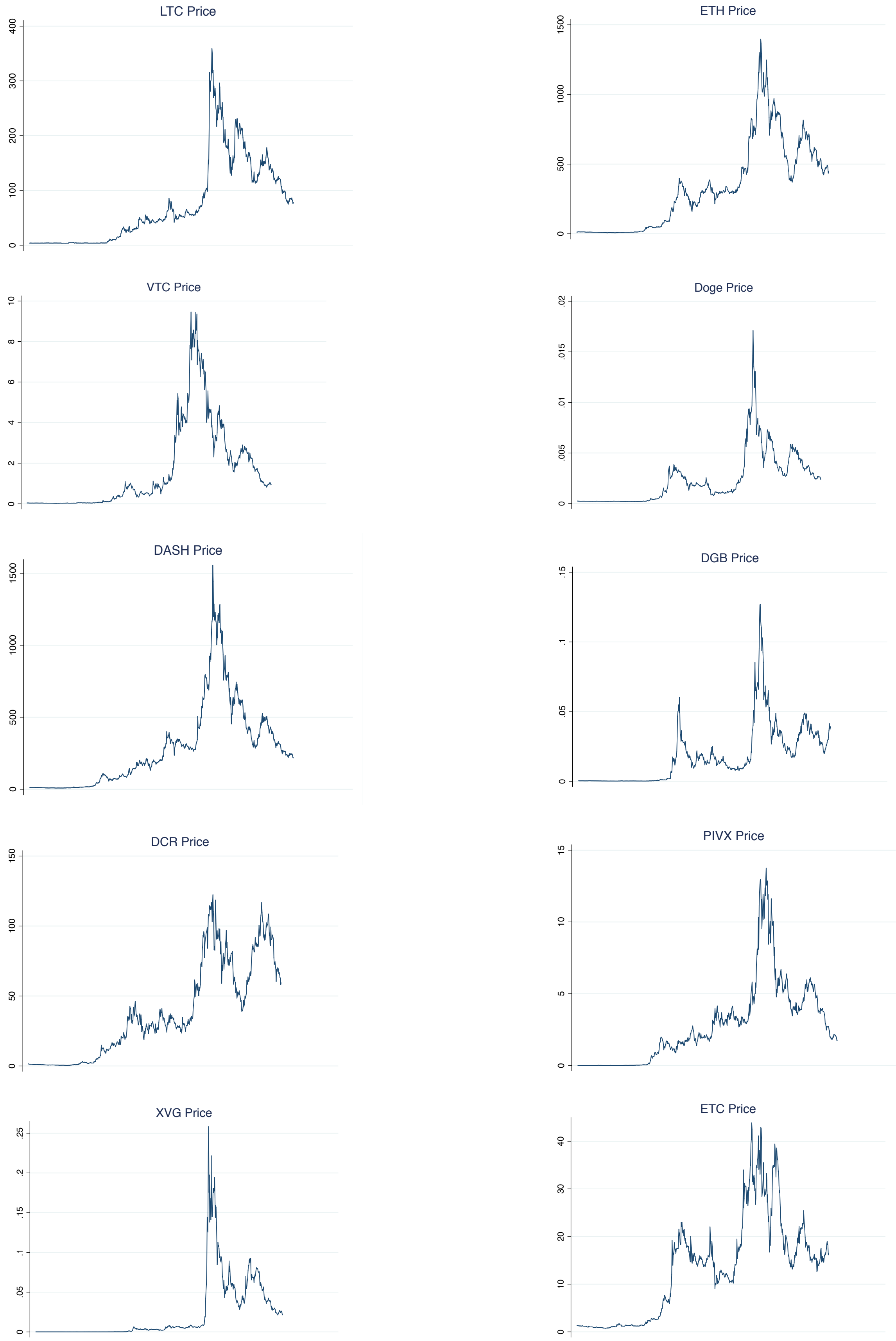
White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4).

Wooldridge, J. (2013). *Introductory econometrics: A Modern Approach*. 5th ed. Mason, Ohio: Thomson/South-Western, pp.495-496.

Yang, H. (2018). *Momentum and Reversal in Cryptocurrency Markets*. Zurich: Department of Banking and Finance, University of Zurich.

Appendices

Appendix 1 : Price of the of the cryptocurrencies between 16/09/2016 and 12/07/2018



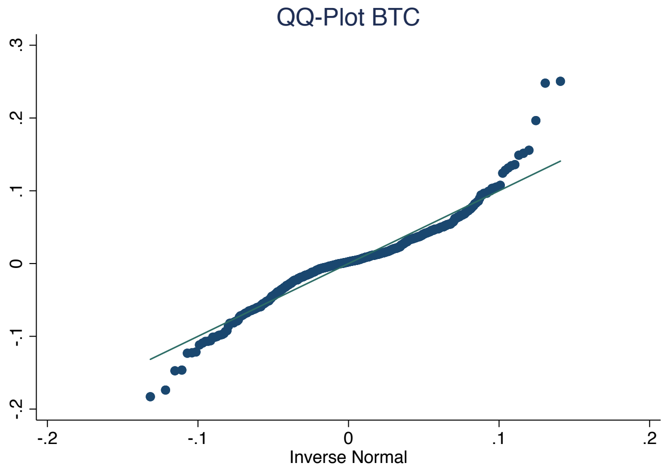
Appendix 2 :

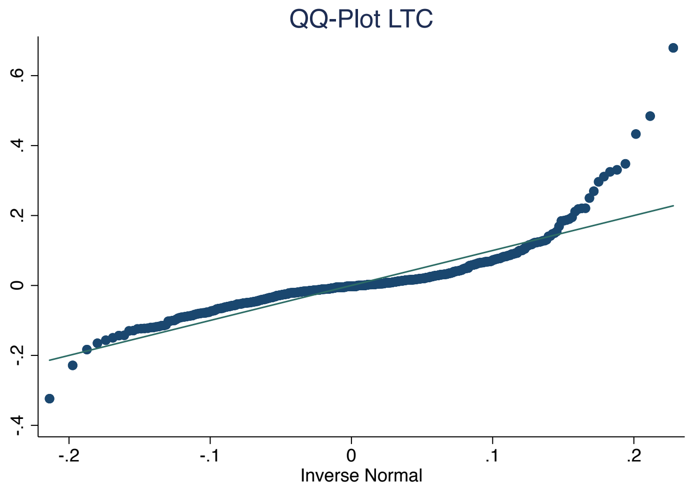
This Correlation Matrix shows the highly correlated nature of cryptocurrencies.

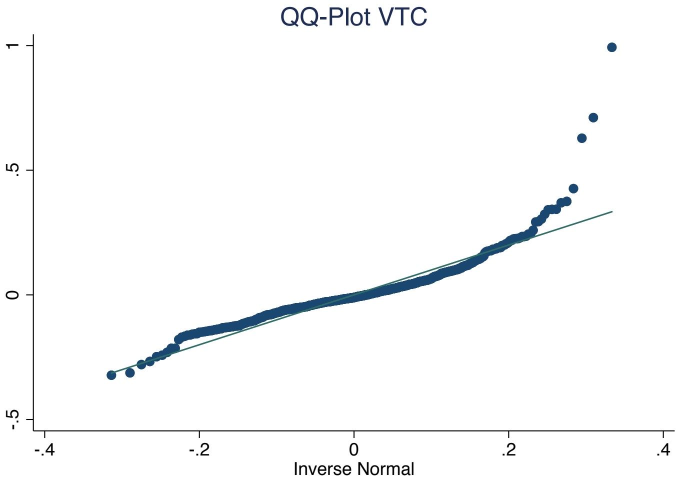
	BTC	LTC	DOG	VTC	DGB	DASH	XVG	ETH	PIVX	DCR	ETC
BTC	100%										
LTC	95%	100%									
DOG	83%	88%	100%								
VTC	93%	84%	72%	100%							
DGB	78%	82%	96%	65%	100%						
DASH	95%	93%	86%	93%	81%	100%					
XVG	77%	84%	90%	68%	88%	82%	100%				
ETH	90%	92%	91%	75%	89%	88%	81%	100%			
PIVX	90%	91%	92%	83%	88%	95%	89%	92%	100%		
DCR	87%	88%	86%	70%	87%	82%	78%	94%	86%	100%	
ETC	89%	89%	87%	82%	83%	90%	70%	91%	87%	85%	100%

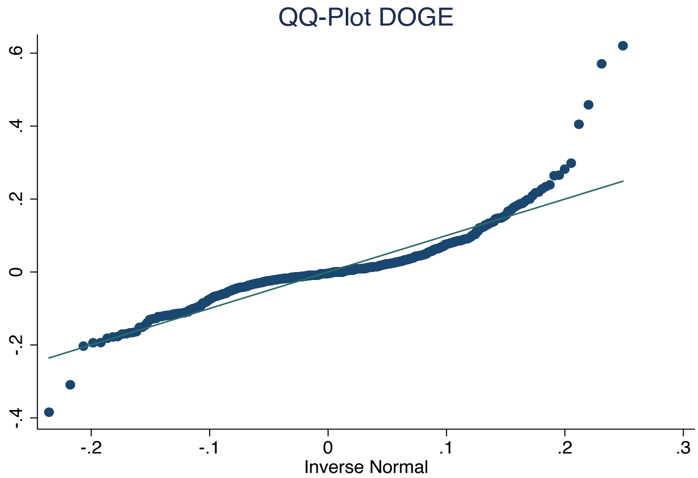
Appendix 3 :

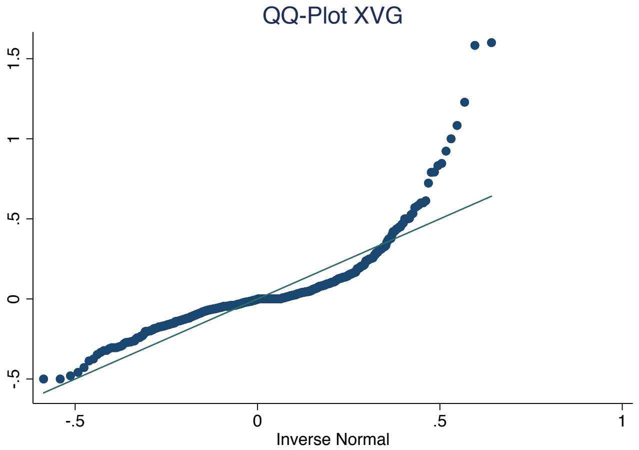
QQ plot reporting the Quantiles of returns over the quantiles of the normal distribution, showing departure from normality.

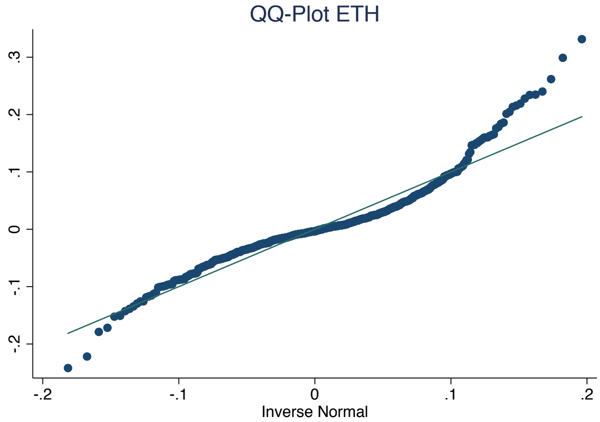




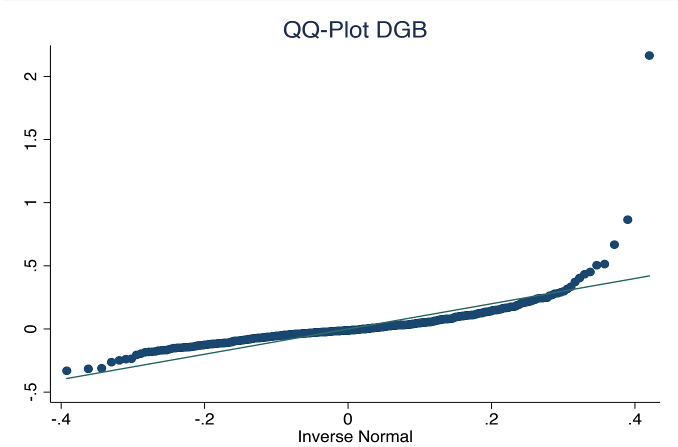


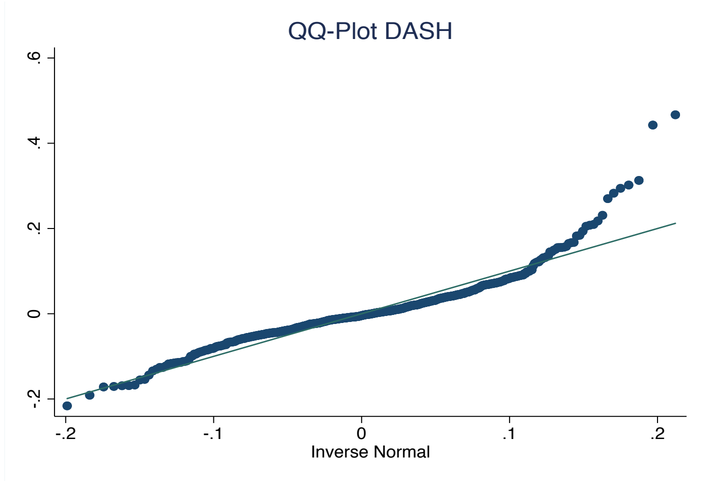


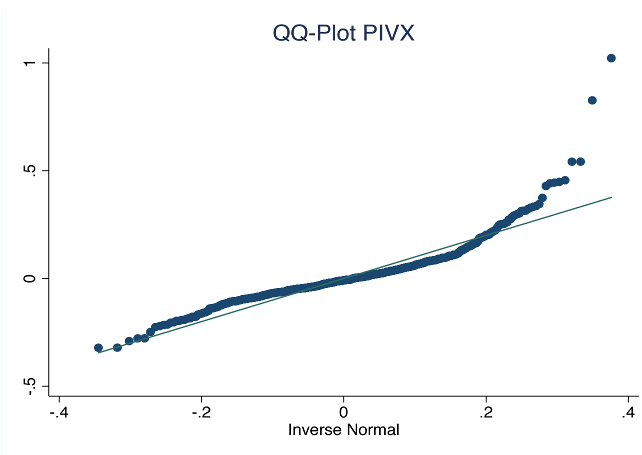


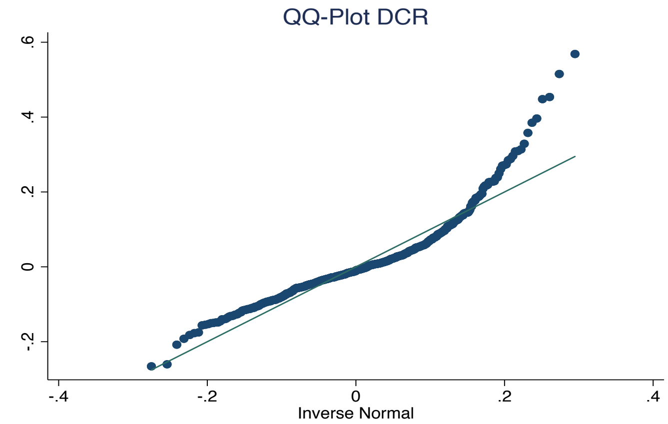


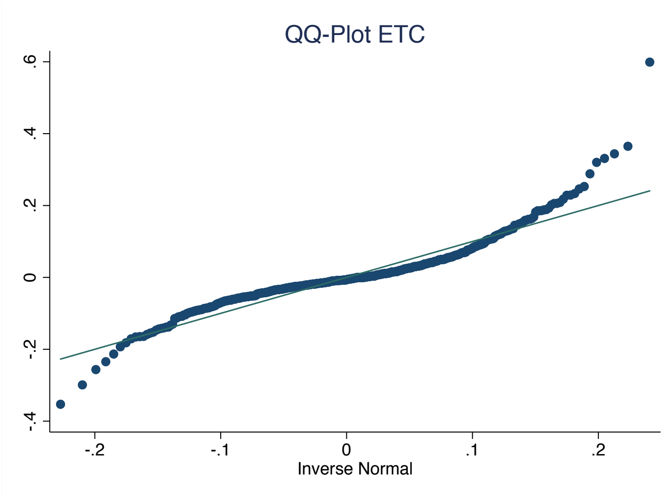
Appendix 3 Continued:











Appendix 4 :

This table shows the number of times that a return exceeded 30% 20% 10% or -30% -20% -10% on any single day.

	BTC	LTC	DOGE	VTC	DGB	DASH	XVG	ETH	PIVX	DCR	ETC	**SUM**
> 30%	0	7	4	11	11	4	47	1	19	11	5	**120**
< -30%	0	1	2	2	3	0	16	0	2	0	1	**27**
> 20%	2	15	16	31	33	13	72	13	44	32	17	**288**
< -20%	0	2	3	9	8	1	40	2	12	3	5	**85**
> 10%	17	47	56	91	101	42	134	44	97	85	66	**780**
< -10%	15	24	47	68	67	25	112	19	61	50	30	**518**

Appendix 5 :

These are the result of a regression in which Momentum Value and Carry were regressed together. When comparing them to the regression of the three single factors, which can be found on table 4 in the main body, we see one major difference: Momentum turns positive in the 1-day returns model, probably correcting the reversal effect reported when regressed singularly. The remaining results do not see significant discrepancies if analysed by their own or jointly

Mom Value and Carry Joint Regression

	1-Day	7-Days	30-Days
Mom	0.0120 (0.0117)	0.4346** (0.1859)	0.7229*** (0.2688)
Value	0.1805** (0.0824)	0.4104*** (0.1417)	1.7860*** (0.5691)
Carry	0.2801*** (0.0753)	0.5340*** (0.1008)	1.5869*** (0.5002)
Obs.	7304	7238	6985

*Where ***= p<1%, **=p< 5% *=p< 10%*