Global Journal of Management and Business Research: C
Finance
Volume 20 Issue 3 Version 1.0 Year 2020
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4588 & Print ISSN: 0975-5853

Factor Model in Crypto Currency Market

By Saket Kumar, Mike Zeng & Ruinan Lu
University of California, Berkeley

Abstract- In our paper, we investigate the explanatory power to the crypto currency return premium of market factor and size factor. We tested both the value-weighted and the equally weighted market factor and a big minus small Fama-French size factor. We found the market and size together can explain 33% of the premium. We also used UMAP to find a non-linear transformation of the crypto returns to create two factors, who can explain over 80% of the premium in both training and testing periods. However, further analysis and research needs to be carried out to decipher what these two factors represent.

GJMBR-C Classification: JEL Code: G20

FACTORMODELINCRYPTOCURRENCYMARKET

Strictly as per the compliance and regulations of:

Factor Model in Crypto Currency Market

Saket Kumar [α], Mike Zeng [σ] & Ruinan Lu [ρ]

Abstract- In our paper, we investigate the explanatory power to the crypto currency return premium of market factor and size factor. We tested both the value-weighted and the equally weighted market factor and a big minus small Fama-French size factor. We found the market and size together can explain 33% of the premium. We also used UMAP to find a non-linear transformation of the crypto returns to create two factors, who can explain over 80% of the premium in both training and testing periods. However, further analysis and research needs to be carried out to decipher what these two factors represent.

I. Introduction

The cryptocurrency market has experienced rapid growth in the past decade. On an almost daily basis, new cryptocurrencies are being created, and the public is paying increasing attention to the new asset class. This market provides chances for companies to raise money without involving venture capitalists and to trade cryptos without being listed on stock exchanges. The set of coins in the crypto market ranges from the best-known cryptocurrency of the time, the Bitcoin, the prominent ones like Ripple, and Ethereum to several other obscure coins. There are over 1900 cryptos issued up to 2019, which resulted in a market of more than $850 billion. Many investment firms have been investing in and maintaining a portfolio of cryptos. Some even have specialized in crypto trading. More than 1,500 crypto currencies are being actively traded by individual and institutional investors worldwide across different exchanges. Over 170 cryptocurrencies focussed hedge funds, have emerged since 2017. Further, in response to increased institutional demand for trading and hedging, Bitcoin futures were launched in December 2017.

As is experienced with the appearance of any new technology, there is always an element of doubt during the initial phase along with differing points of view. Similarly, there are controversies surrounding the cryptocurrency market. Many people are struggling to understand what cryptocurrencies are or what is the exact mode of their operations. There is also a view that cryptocurrencies are the representative of some asset bubbles and fraud. The other perspective is that the blockcha in technology underlying the cryptos is a significant financial innovation and some of these cryptos could become major future technological assets. This belief system has led to major development in the overall crypto market. Thus, there is a need to analyse the cryptocurrency market from the empirical rule-based approach for at least two reasons. The first reason is to understand whether the returns of cryptocurrencies share similarities with other asset classes, most importantly, with equities. The second reason is that to assess and develop theoretical models of cryptocurrency, it is meaningful to build an empirical model to be used as stylized facts and inputs. Since there is no simple universal framework to construct a crypto portfolio unlike the equity market, we, therefore, propose to create a factor model for cryptocurrencies. The factor model has been traditionally used in the equity markets to decompose the assets return and risk. (e.g., CAPM, Fama-French, MSCI BARRA), so it could also provide a paradigm to analyze such patterns in the cryptocurrency market.

Therefore, in this paper, we have tested if there are stylized factors similar to the equity market, such as market, size, value, and momentum present in the crypto market. We have also used machine learning algorithms to look for the underlying factors in the market returns matrix. We have divided this paper into four sections, literature review, the data section, methodology and results section, the conclusion, and future directions.

II. Literature Review

Even since the advent of Bitcoin in around 2008, a lot of research have been conducted and corresponding literature has been published. With the surge of the cryptocurrencies from 2017 onwards, this attention has become more widespread. A particular focus of these studies has been to find out what is driving the cryptocurrency prices, be it the exogenous factors such as various economic and financial indicators or endogenous factors such as hash rate etc. Liu and Tsyvinski (2018) established that the risk-return trade-off of cryptocurrencies is distinct from those of stocks, currencies and commodities. Cryptocurrencies have no exposure to the most common stock market and macro-economic factors and are predictable based on the endogenous cryptocurrency market-based factors. Bhambhawani et al. (2019) found that endogenous fundamental indicators computing power (hash rate) and network (number of users) had a

Author α: Master of Financial Engineering, University of California, Berkeley. e-mail: saket_kumar@berkeley.edu
Author σ: Assistant General Manager at Reserve Bank of India, Mumbai. e-mail: mikezeng@berkeley.edu
Author: e-mail: ruinan.lu@berkeley.edu
The views expressed are personal and do not reflect of the institution

significant long-run relationship with the prices. Kakushadez (2018) proposed factor models for the cross-section of the daily crypto-asset returns and, based on empirical analysis, identified that three short horizon factors size, momentum, and intra-day volatility work well for crypto-assets. Momentum dominates as a factor for crypto-returns on short-horizons suggesting that the market is strong mean-reverting.

The momentum effect is a ubiquitous market phenomenon by which asset prices follow a trend for a rather long time. A large number of studies have been done about deciphering the momentum effect in the equity markets, and the controversy about its effect is not uncommon in the empirical equity literature. Momentum factor may be viewed as short volatility investing and historically has provided a long period of high returns with occasional large draw downs. The momentum factor is similarly being discussed and debated for the cryptocurrencies. Grobys and Sapkota (2019) investigated the existence of momentum implemented in the crypto market. They used the time series data on hundreds of cryptocurrencies in the period of 2014-2018 and implemented momentum strategies. They also checked the highest 30 market capitalizations cryptos for robustness. In their paper, they investigated the profitability of the momentum strategies in the cryptocurrency market on a portfolio level. Interestingly, they do not find any evidence for cross-sectional momentum in the cryptocurrency market. They also do not find strong evidence that supports the time-series momentum effects, even some of the strategies generate negative payoffs.

Liu, Tsyvinski, and Wu (2019) found a different point of view. They examined whether the factors that are considered prominent in the cross-section of equity returns are also significant in the cross-section of cryptocurrency returns, specifically cryptocurrency market, size, and momentum factors. They used 1707 crypto samples from the beginning of 2014 to the end of 2018, excluding coins with relatively small market capitalization. They found that the long-short strategy generated about 3% excess weekly returns. Additionally, the momentum effect is significantly greater in the larger coins. The momentum strategy in the below-median size group gives 0.6% weekly excess returns, while the momentum strategy in the above-median size group gives 4.2% weekly returns; both numbers are statistically significant. They conclude that momentum factors are significant in capturing the cross-section of cryptocurrency returns, similar to other asset classes.

Sobvetob (2018) examined the factors that most commonly influence the prices of the top five cryptocurrencies Bitcoin, Ethereum, Dash, Litecoin, and Monroe over 2010-2018 using weekly data. The study found that factors such as market beta, trading volume, and volatility appear to be significant returns determinant. However, there are limited works done on the market and size factor. The value factor may be viewed as ambiguous in the crypto market, even though, we can define it from a behavioural perspective, we do not delve into this aspect for now and focus on the market and size factors.

III. Data Sourcing and Analysis

a) Data Sourcing

We collect our data from Coin Gecko (https://www.coingecko.com/en). Coin Geckohas information on more than 6900 coins from over 400 exchanges and has daily data on prices, volume, and market capitalization (in dollar terms). Also, Coin Gecko also has community growth, open-source code development, major events, and on-chain metrics.

To be listed on Coin Gecko, a cryptocurrency needs to fulfil a list of criteria. These include, actuall trading on a public exchange such that the information matches the information in API to report the last traded price and the last 24-hour trading volume along with being liquid on at least one of the supporting exchanges in order for the price to be determined. We acquired a historical data of daily price, market capitalization, and trading volumes of 6682 cryptocurrencies over a time period of April 28th, 2013 to January 1st, 2020.

For each cryptocurrency on the website, the price is calculated based on the pairing available and is collected by Coin Gecko from various exchanges. The price shown on CoinGecko for a particular cryptocurrency is calculated based on a global volume-weighted average price formula. The trading volume for a cryptocurrency on Coin Gecko is the aggregate trading volume of all trading pairs of cryptocurrencies. The market capitalization of a cryptocurrency is the current cryptocurrency price in USD multiplied by its volume. We downloaded the data from the given API by the website, which further required heavy processing and wrangling to transform in a usable format. The data was processed into three categories of price data, volume data, and a cap-weighted market portfolio.

b) Data Analysis

As we introduced in the beginning, the number of cryptos boomed after 2018. We can see a change of the slope around the end of 2017, in *Figure 1.*

Global Journal of Management and Business Research (C) Volume XX Issue III Version I Year 2020

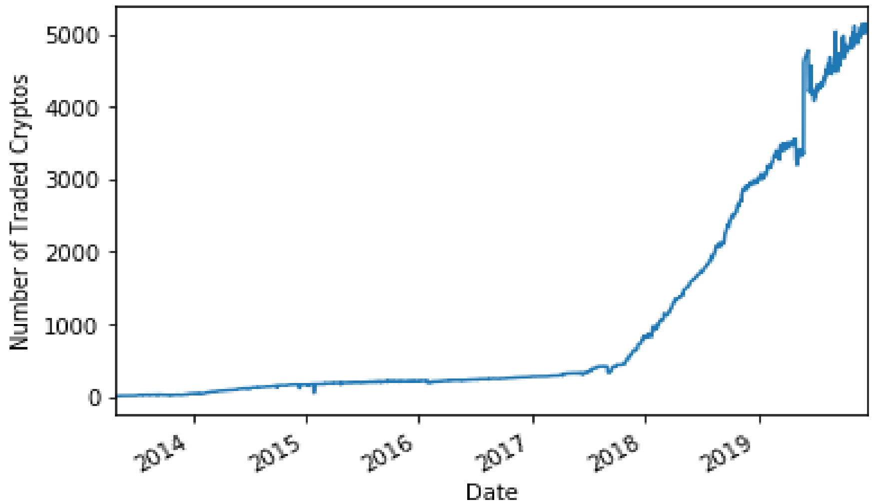


Figure 1: The Number of Cryptos Traded on Market

We can see that before 2017 trading cryptos was uncommon. But after 2017, new cryptos were issued every day.

c) Daily Return

Our first objective was to recover the daily percentage returns for all the cryptos from the price matrix. The above astronomical increase in the number of cryptocurrencies caused problems for our analysis. There are too many cryptos with few valid observations. We had to limit the cryptos to those with a long history. After a few trials and errors, we decided to use cryptos that were available before July 1st, 2017, as our sample, which is the elbow in *Figure 1.* This sample is reasonably stable. It contains 341 cryptos over 2423 days, as shown in Figure 2. We assume these cryptos have a stable behaviour compared to those trend-chasing new cryptos.

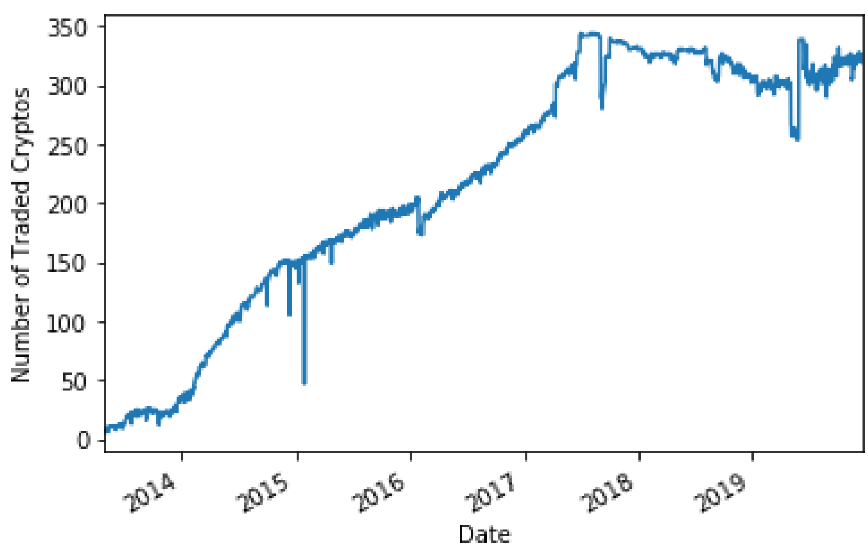


Figure 2: The Number of Long-lived Cryptos Traded on Market

There are four usual dips in the data.

However, we do see that there are four abnormal dips in our plot. The first dip happened on January 28th and 29th, 2015. Since these are only two days, we believe it to be a data error, and we fixed the same by linear interpolating the data.

The second dip happened in February 2016. For about a week, the prices of about 20 cryptos' were missing, as shown in *Figure 3.*

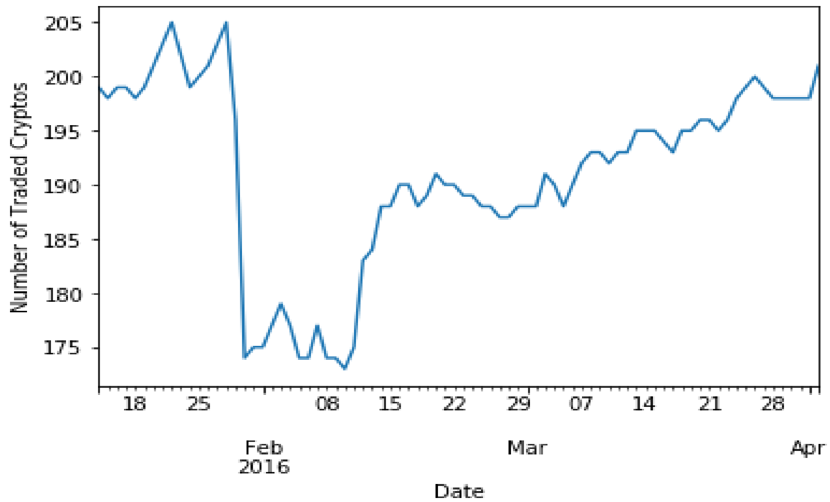


Figure 3: The Unusual Dip in February 2016

We can observe that these missing data points are rather systematic. We analyzed the details and have presented the name of the cryptos in *Table 1* below.

Table 1: The name of Disappeared Cryptos

bluecoin	bottlecaps	cachecoin	casinocoin	crave	deutsche-emark	fastcoin
gambit	goldcoin	ixcoin	nyancoin	opal	phoenixcoin	ratecoin
salus	sexcoin	supercoin	terracoin	tittiecoin	unitus	whitecoin

The reason for this dip is unknown as these cryptos seem to be uncorrelated. We suspect there was a shock to the market that caused liquidity to decrease. Before this, all these cryptos were already very illiquid, as shown in *Figure 4.*

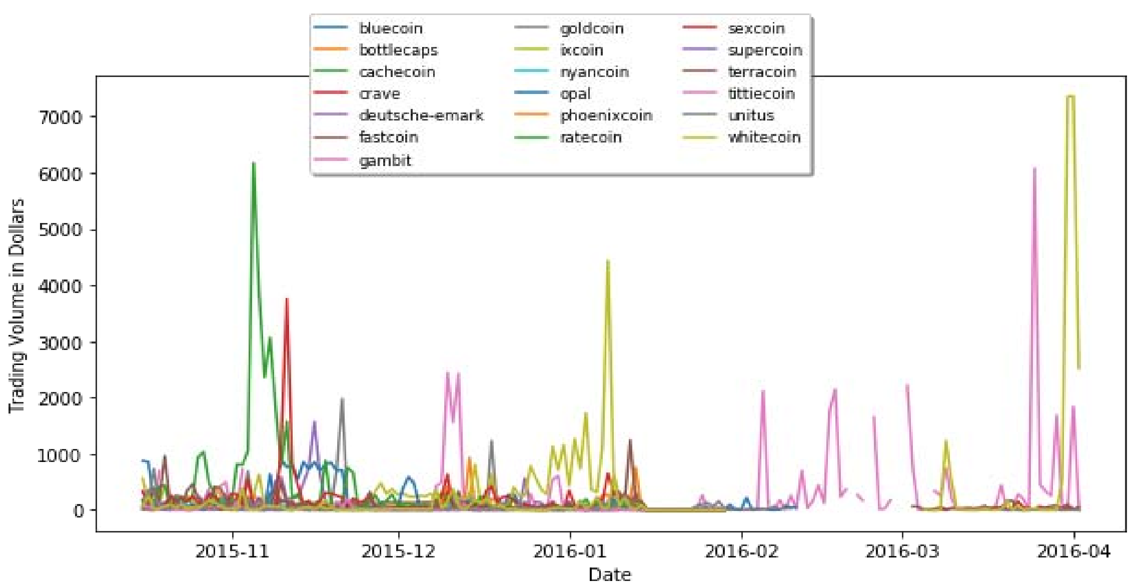


Figure 4: The Trading Volumes of Missing Coins

Most missing coins were very illiquid around February. Therefore, the missing is likely to be a market event rather than a data error.

The third dip happened in September 2017. This time about 60 cryptos were missing and then gradually recovered in the following month as shown in *Figure 5.*

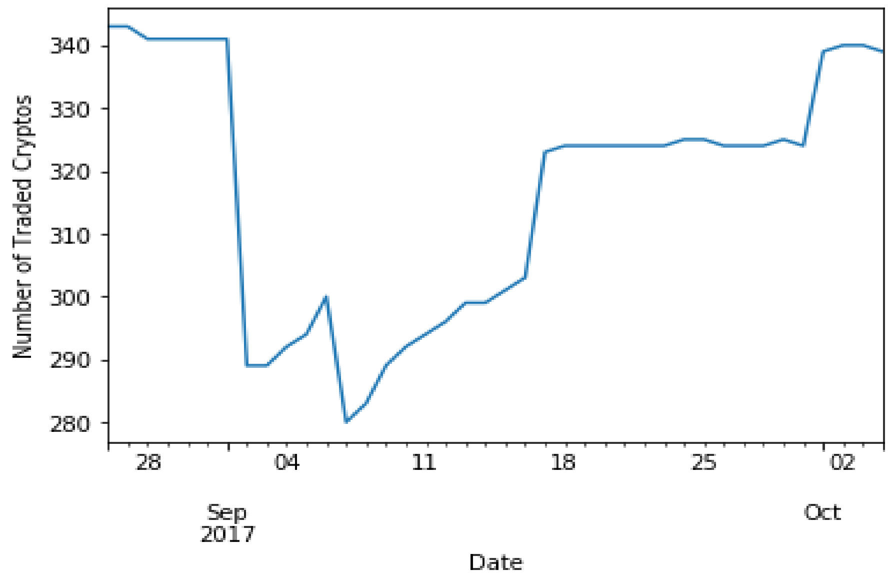


Figure 5: The Unusual Dip in September 2017

The possible reason for this dip is attributed to the fact that in September 2017, the Chinese Government banned all cryptos and Initial Coin Offerings (ICOs) in China and issued a warning to the crypto exchanges. This event may have likely triggered some China-specific cryptos to stop trading. The ownership pattern of these cryptos reveal that a majority of them as China based.

The last dip happened in May 2019 as shown in Figure 6. Due to the unnatural behaviour, we believe it's a data error like the dip 1, which can be fixed by linear interpolation. Given these, the errors are local and minor and should not cause any significant errors.

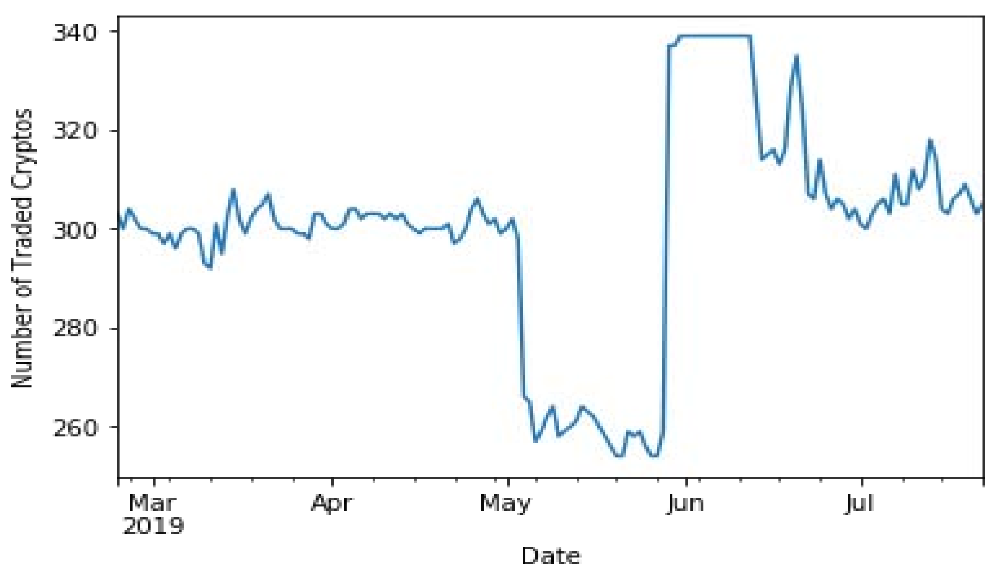


Figure 6: The Unusual Dip in May 2019

The flat top and two vertical jumps suggest it's likely a data error.

d) Market Capitalization

To create a market factor, we looked into the evolution of market capitalization's distribution. The market cap's distribution, in general, has three modes and a few outliers. The market can be separated into small-cap (< 250,000 USD), mid-cap (1 ~ 200 million USD), and large-cap (> 300 million USD). The few outliers are Bitcoin, Ethereum, and Ripple (all with a cap greater than 30 billion). *Figure 7* shows a typical distribution of market cap suggesting size as a good factor.

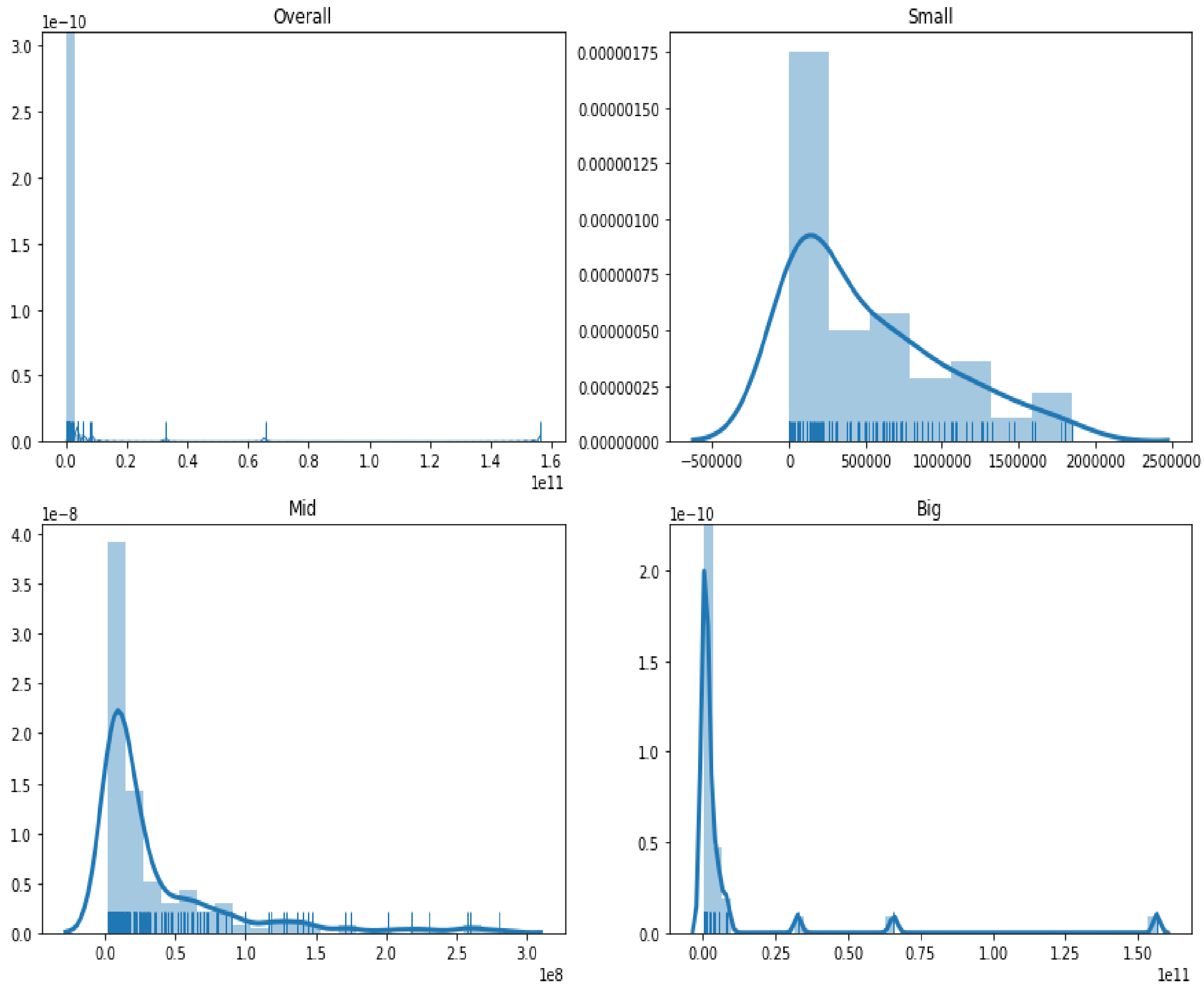


Figure 7: Market Cap Distribution on May 1st, 2018

The top left is the overall kernel estimated density plot. The top right is the kernel estimated small-cap distribution. The bottom left is the kernel estimated mid-cap distribution. The bottom right is the kernel estimated big-cap distribution. Each group has 106, 184, and 30 cryptos, respectively. We see a clear separation of big-cap, mid-cap, and small-cap. (Look at the Appendix for more plots.)

e) Excess Return

For traditional assets, the excess return is defined in terms of a risk-free rate generally taken as the 10- year Treasury bond rate yield. Considering cryptocurrency as an investment asset, it makes more sense to look at the cryptocurrency return in comparison to the risk-free rate, even though traditionally crypto prices is not correlated with interest rate or monetary policy (Benigno, 2019). We have used the universally used US 10-year Treasury bond rate yield of the same time window as our risk-free rate. We forward filled the weekend values to accommodate the crypto market. After we calculated the premium returns, we found that many cryptos have big outliers due to illiquidity as shown in *Figure 8.*

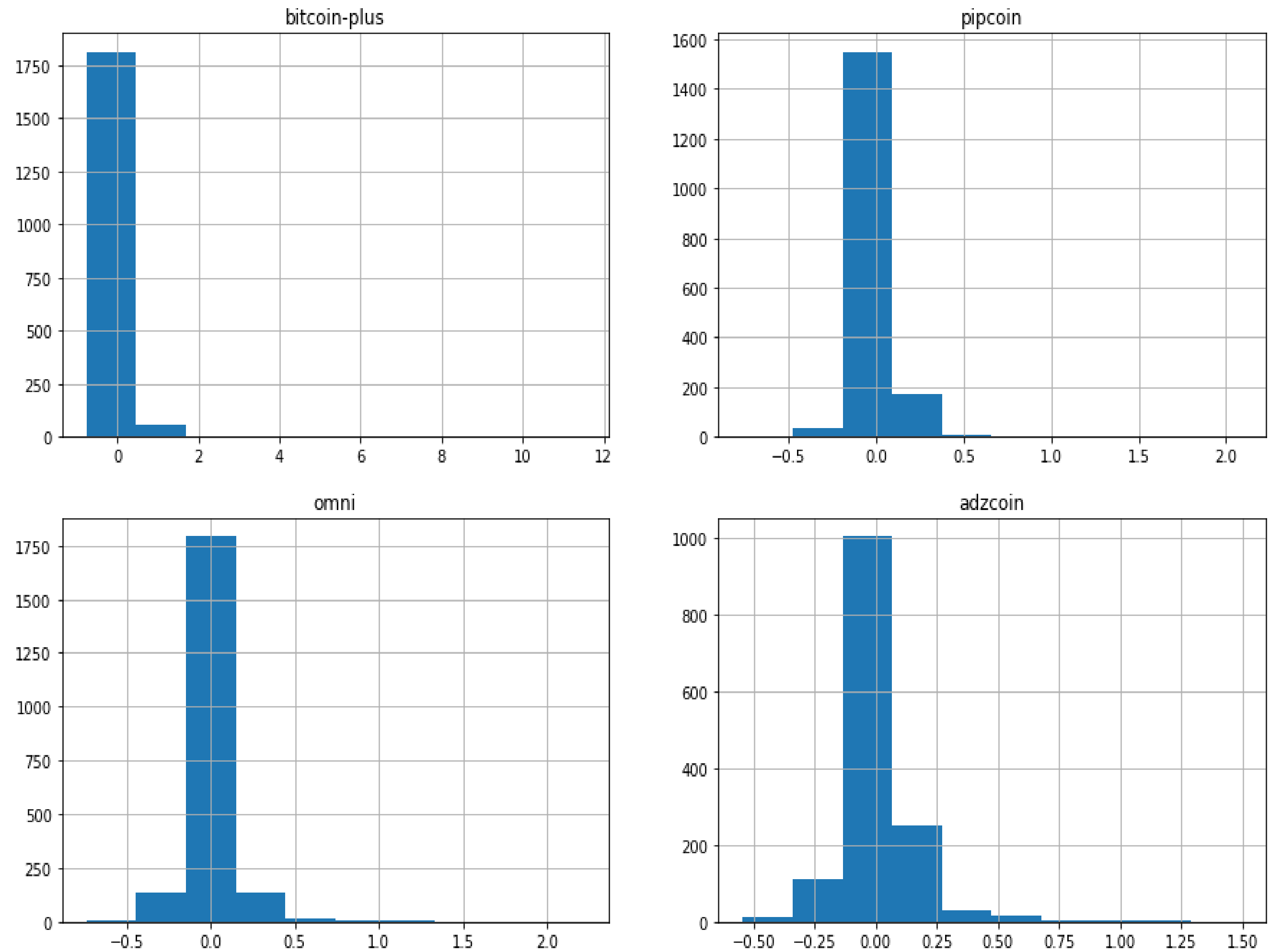


Figure 8: The Raw Return Distribution of Four Cryptos

We can observe outliers in all four of them, which suggest we should further process the data for better behavior.

A well-established method is to winsorize the data. We transform the statistics by limiting the extreme values to reduce the effects of possibly spurious out liersby trial and error and replace the extremes with 0.1 percentile and 99 percentiles. After the winsorization transformation, the distribution looks close to a normal distribution as shown in *Figure 9.*

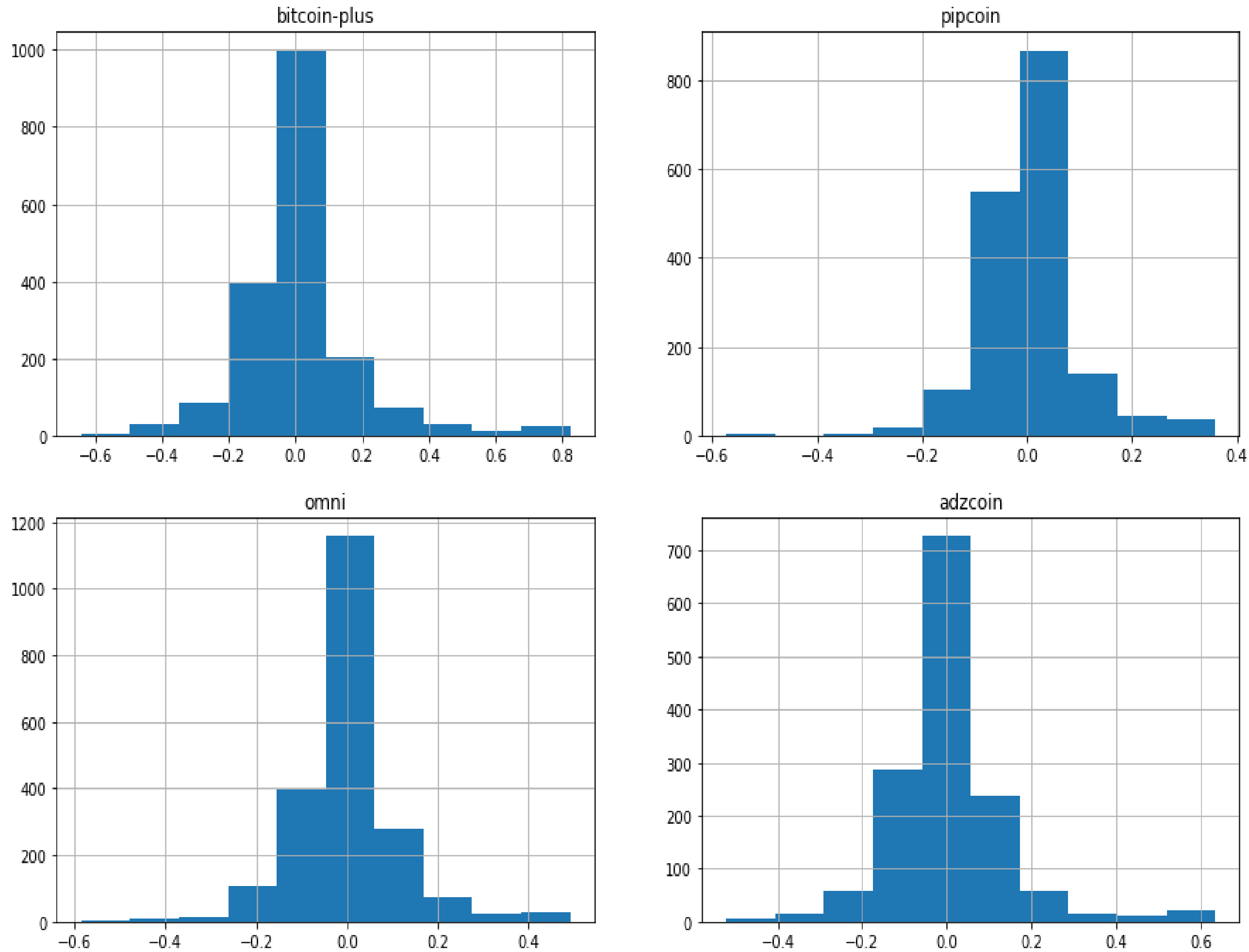


Figure 9: The Winsorized Return of Four Cryptos

We see a better bell-shape distribution

IV. Methodology and Results

We have used several different ways to construct traditional equity factors, namely market, and size. Additionally, we have also tried unsupervised machine learning techniques to uncover the low dimensional representation of the crypto model. We used three years of data from 2017 April to 2020 January to construct the factors of size and market.

a) The Market Factor

We tried three ways to create a market factor, cap-weight, equal-weight, and cap-weight of the most liquid 100 cryptos.

i. *Cap-weighted Market Factor*

We used the total market cap to divide each crypto's market cap to get the appropriate weights. Then we took the weighted average return as the market return, which is negative on average. *Figure 10* shows a bell-shape distribution and autocorrelation of the factor. The market is not auto correlated.

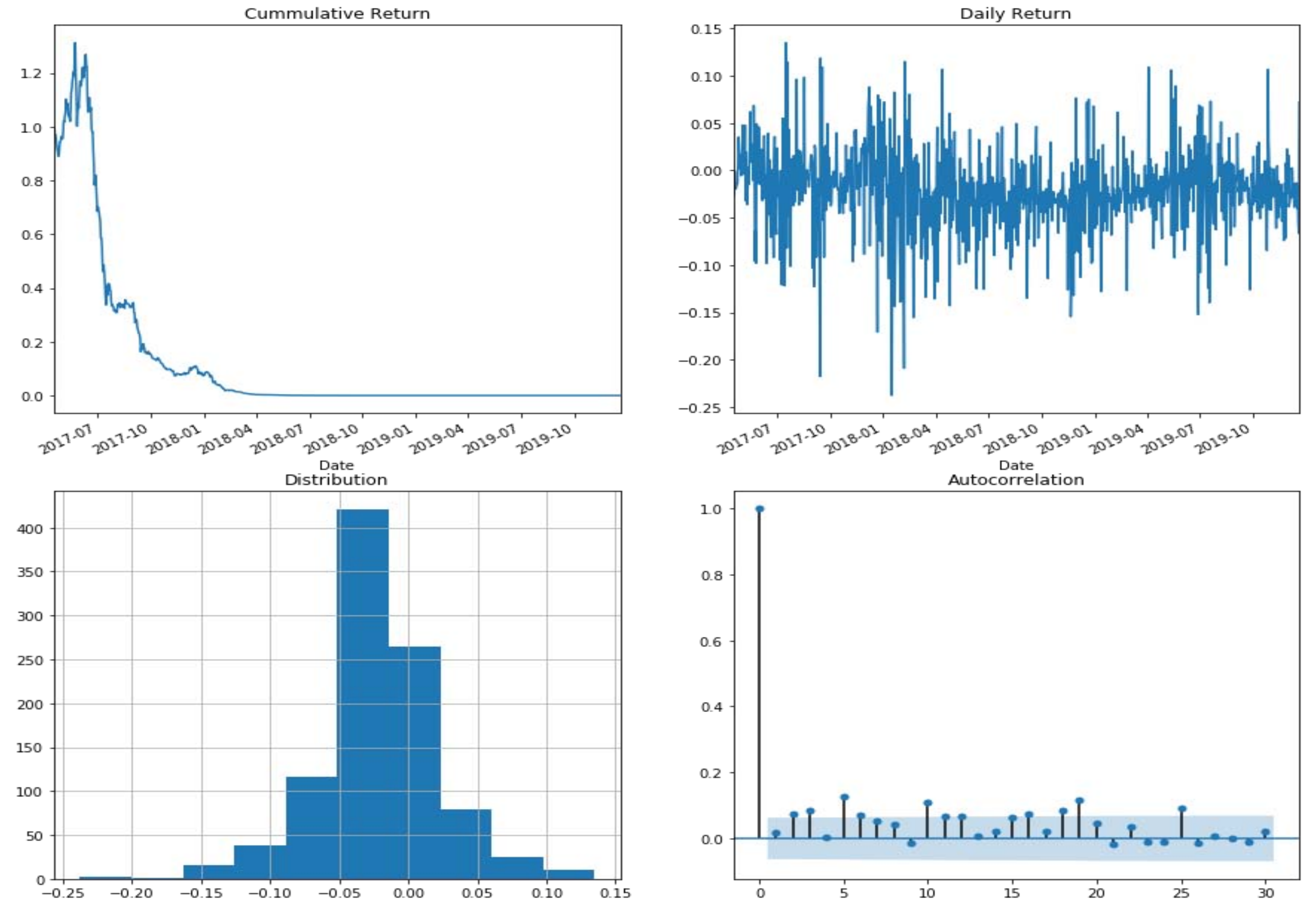


Figure 10: Summary Plots of Value-weighted Market Factor

a) The top left is the cumulative return. We see the market is declining overall.
b) The top right is the daily return plot. We observe many large returns, both positive and negative.
c) The bottom left is the histogram of returns. We see a heavy tail bell shape.
d) The bottom right is the autocorrelation plot. The market cannot predict itself, which suggests a mostly efficient market.

We present a summary statistics of this value-weighted market factor in *Table 2 below.* We observe that the overall mean returns are -0.2%, with a standard deviation of 4.3%. The variability is also large as the minimum value is -23.71%, while the maximum is 13.49%.

Table 2: Summary Statistics of Value-Weighted Market Factor

mean	std	min	25%	50%	75%	max
-0.02122	0.04307	-0.23771	-0.03881	-0.021178	0.001463	0.134928

To test this factor, we ran regressions between each crypto and the factor to get the exposures (Beta). The distribution of market exposure is presented below in *Figure 11.* The average beta is about 0.8, with a standard deviation of 0.32. We also would like to know how much premium this factor explains. So we ran a regression between beta and risk premium. It turns out the cap-weighted market does not explain the premium at all. *Table 3* shows that the R-squared is close to zero.

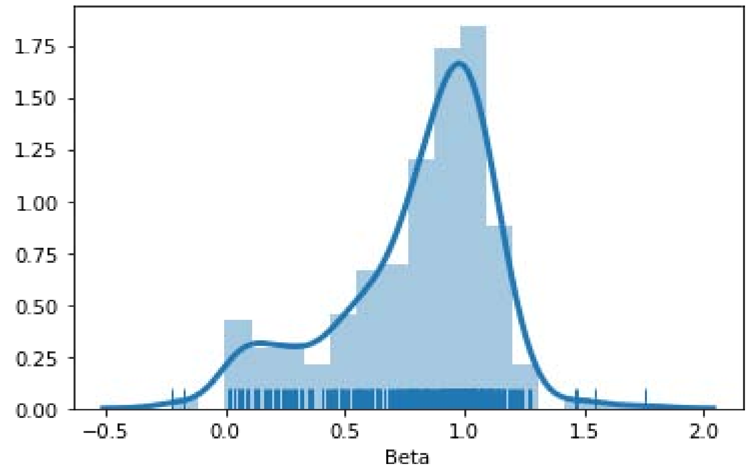


Figure 11: The Distribution of Market Exposure

Table 3: The Regression Results of VW Market Factor

Dep. Variable:	y	**R-squared:**	0.004
Model:	OLS	**Adj. R-squared:**	0.001
Method:	Least Squares	**F-statistic:**	1.301
Date:	Sat, 14 Mar 2020	**Prob (F-statistic):**	0.255
Time:	20:44:14	**Log-Likelihood:**	807.74
No. Observations:	341	**AIC:**	-1611.
Df Residuals:	339	**BIC:**	-1604.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
x1	0.0043	0.004	1.141	0.255	-0.003	0.012
const	-0.0186	0.003	-5.668	0.000	-0.025	-0.012

Omnibus:	394.598	**Durbin-Watson:**	1.935
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	17940.762
Skew:	5.316	**Prob(JB):**	0.00
Kurtosis:	36.907	**Cond. No.**	5.21

X1 is market exposure

ii. *Equal-Weight Market Factor*

We now create an equally weighted market factor following the same steps as above but with equal weight. The equal-weighted factor has very similar behaviour to the cap-weighted factor with negative average return. The same is represented in *Figure 12* below.

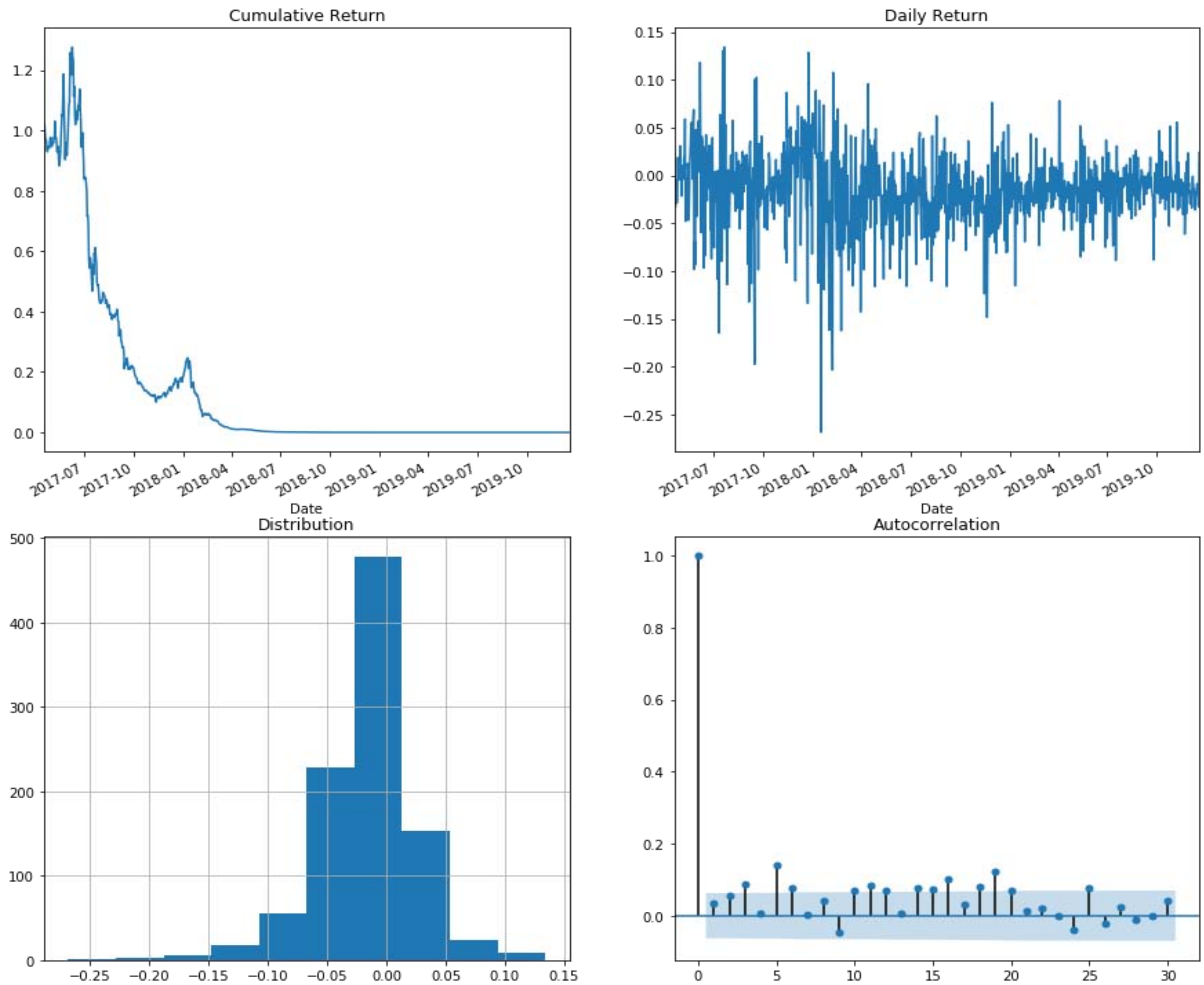


Figure 12: Summary Plots of Equally Weighted Market Factor

a) The top left is the cumulative return. We see the market is declining overall.
b) The top right is the daily return plot. We observe many large returns, both positive and negative.
c) The bottom left is the histogram of returns. We see a heavy tail bell shape.
d) The bottom right is the autocorrelation plot. The market cannot predict itself, which suggests a mostly efficient market.

Table 4: Summary Statistics of Equally Weighted Market Factor

mean	std	min	25%	50%	75%	max
-0.015772	0.0399	-0.268749	-0.033485	-0.014128	0.004247	0.134217

We tested this factor in the same way as above. The average beta, is about 1, with a standard deviation of 0.44. It explains about 30% of the premium. We also examined the correlation between the two factors. We found the two factors volume-weighted and equal-weighted correlate 87%. And they are both strongly correlated with Bitcoin and Ethereum. The distribution of market exposure can be seen in *Figure 13* below.

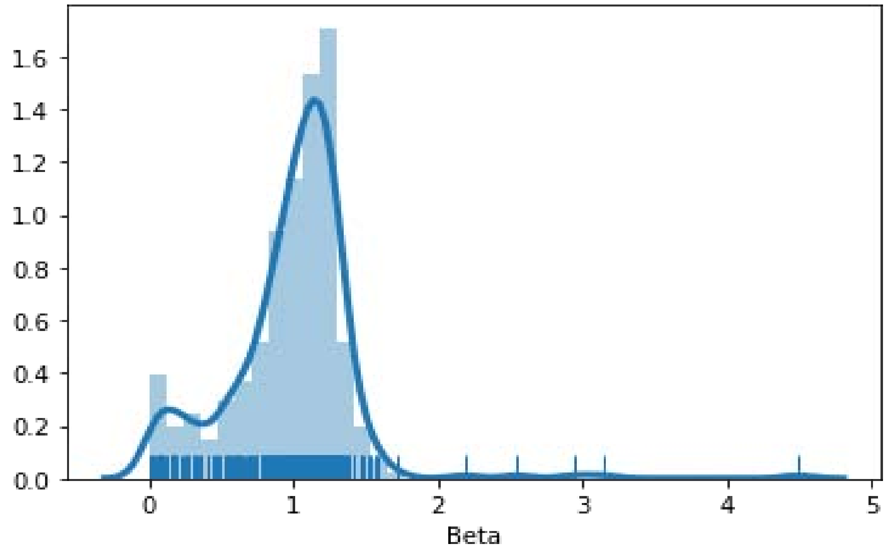


Figure 13: The Distribution of Market Exposure

Table 5 shows that the regression result of the equal-weighted market factor, the R square is about 30%.

Table 5: The Regression Results of EW Market Factor

Dep. Variable:	y	**R-squared:**	0.299
Model:	OLS	**Adj. R-squared:**	0.297
Method:	Least Squares	**F-statistic:**	144.5
Date:	Sat, 14 Mar 2020	**Prob (F-statistic):**	5.74e-28
Time:	20:53:32	**Log-Likelihood:**	867.63
No. Observations:	341	**AIC:**	-1731.
Df Residuals:	339	**BIC:**	-1724.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
x1	0.0277	0.002	12.021	0.000	0.023	0.032
const	-0.0426	0.003	-17.012	0.000	-0.048	-0.038

Omnibus:	279.191	**Durbin-Watson:**	2.063
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	5020.193
Skew:	3.351	**Prob(JB):**	0.00
Kurtosis:	20.561	**Cond. No.**	4.65

X is market exposure

iii. *Cap-weighted Market Factor (100 most liquid cryptos)*

We factored liquidity into consideration by constructing a cap-weighted market factor with 100 most liquid cryptos. However, there was no observed significant improvements over the existing two factors. Therefore, we conclude that an equally weighted portfolio is a better measure of the market factor.

b) The Size Factor

As we mentioned in section 3.2.1, we have segregated the whole market into three sizes of large-cap, mid-cap, and small-cap. We look at the size factor in a manner synchronous to the Fama-French style. We have sorted the market cap into ten bins every day, then we use the biggest minus smallest cap to create a Big minus Small (BMS) size factor. We have plotted the cumulative returns of all portfolios in *Figure 14.*

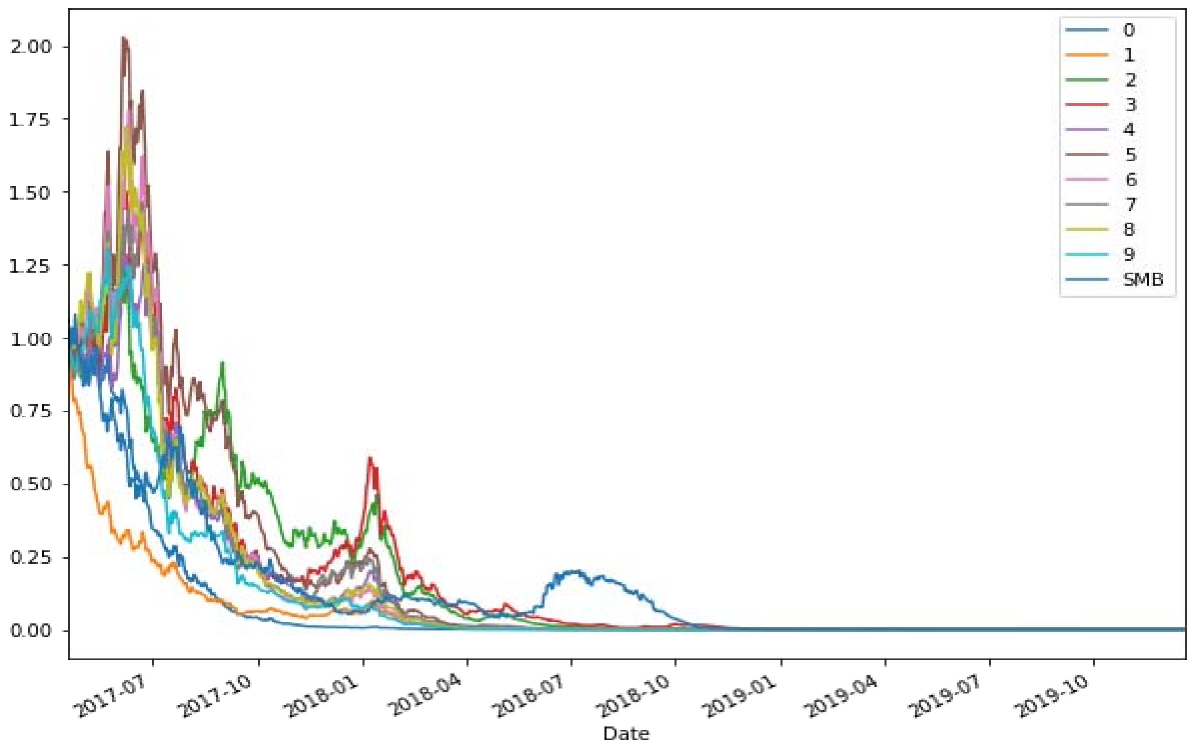


Figure 14: The Cumulative Returns of All Size Portfolios

All portfolios have negative average returns

Next, we analyze the BMS factor. As shown in Figure 15. below, the BMS factor has a positive average return and grows exponentially over the period. The size factor has almost zero correlation with the market factors.

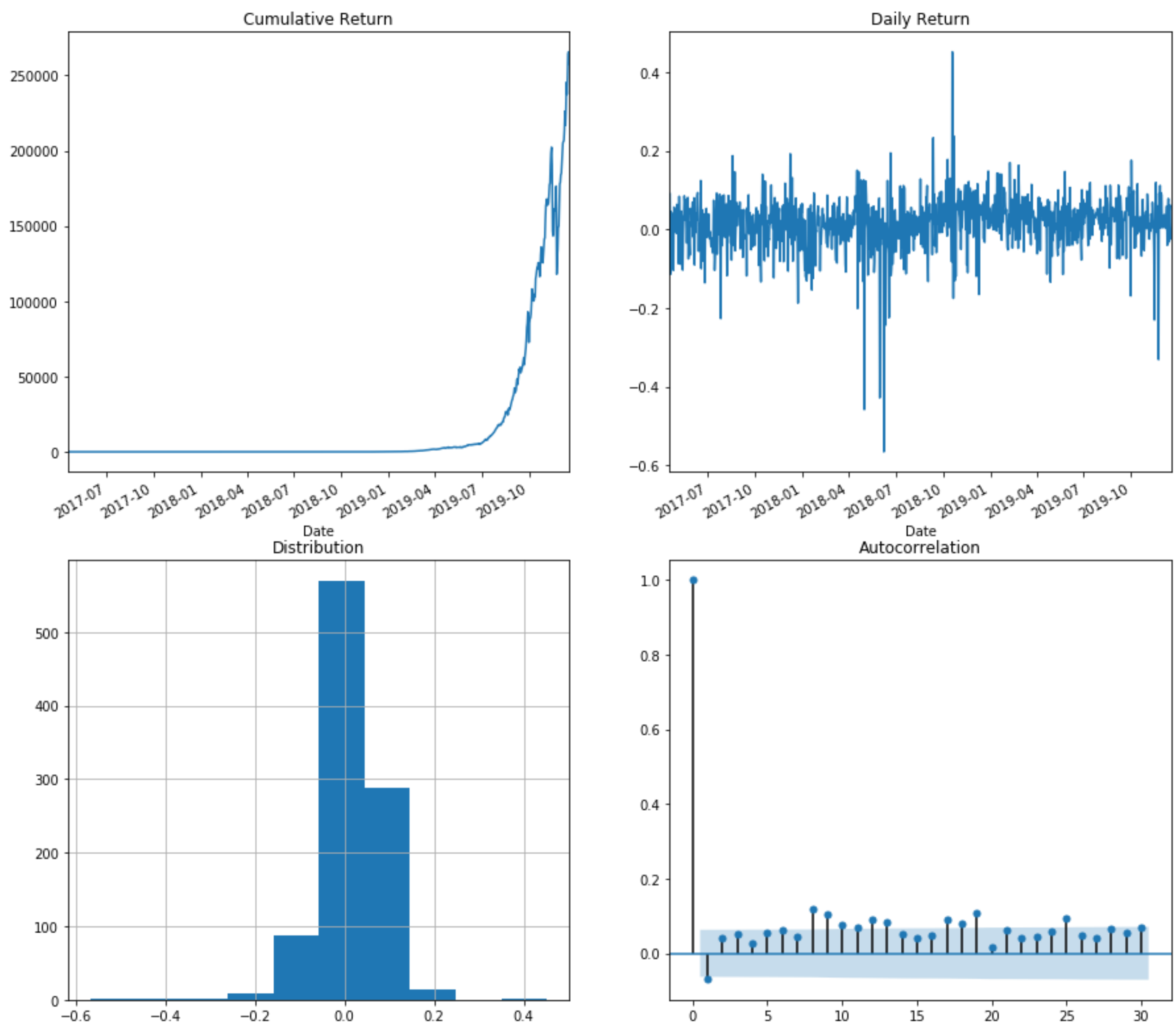


Figure 15: Summary Plots of Big-Minus-Small Size Factor

a) The top left is the cumulative return. We see the market is increasing overall.
b) The top right is the daily return plot. We observe many large returns, both positive and negative.
c) The bottom left is the histogram of returns. We see a heavy tail bell shape.
d) The bottom right is the autocorrelation plot. The BMS cannot predict itself.

A summary statistics of this value-weighted market factor is given in *Table 6 below.* We observe that the overall mean returns are 1.54%, with a standard deviation of 6.86%. The variability in returns is also large and greater than the market factor as the minimum return experienced has been -56.6% while the maximum return is 45.13%.

Table 6: Summary Statistics of BMS Size Factor

mean	std	min	25%	50%	75%	max
0.01546	0.06864	-0.56584	-0.01540	0.021854	0.053298	0.451382

We then tested the factor the same way as done above. The BMS factor explains only a marginal part of the premium. However, if we combine the BMS factor with the equally weighted market factor, they can explain 33.3% of the premium. The detailed results are shown in *Table 7* below.

Table 7: The Regression Results of EW Market Factor and BMS Size Factor

Dep. Variable:	y	**R-squared:**	0.333
Model:	OLS	**Adj. R-squared:**	0.329
Method:	Least Squares	**F-statistic:**	84.53
Date:	Sat, 14 Mar 2020	**Prob (F-statistic):**	1.71e-30
Time:	22:25:44	**Log-Likelihood:**	876.24
No. Observations:	341	**AIC:**	-1746.
Df Residuals:	338	**BIC:**	-1735.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
x1	-0.0268	0.006	-4.466	0.000	-0.039	-0.015
x2	0.0299	0.002	12.930	0.000	0.025	0.034
const	-0.0448	0.003	-17.896	0.000	-0.050	-0.040

Omnibus:	298.427	**Durbin-Watson:**	2.048
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	6570.561
Skew:	3.616	**Prob(JB):**	0.00
Kurtosis:	23.252	**Cond. No.**	8.70

The x1 is the exposure to the BMS size factor. The x2 is market exposure. Exposure to the market generates positive returns, while to size negative returns.

c) Unsupervised Machine Learning

After trying the two traditional equity factors, we now turn to the machine learning approach. We try Independent component analysis (ICA), Partial component analysis (PCA), and Probabilistic PCA. While the PCA works with maximizing the variance, the ICA focusses on independent components. They both separate a multivariate signal into additive subcomponents. After running the analysis over our dataset, we found out that all the three methods above, perform roughly the same as the two-factor model of market and size discussed above. To improve upon this, we used Uniform Manifold Approximation and Projection (UMAP) to find a better two-factor representation of the market. UMAP technique can be used for visualization similarly to t- Distributed Stochastic Neighbour Embedding (t-SNE), but also for general nonlinear dimension reduction. UMAP also works on dimension reduction[1]. To avoid over fitting, we split the data into two windows, test and train. The train window is about two years, followed by a one-year test window.

Through a grid search cross-validation, we found that by looking at 160 neighbors and using the Chebyshev matrix with the minimum distance of one, we can explain as much as 80% of the premium with two statistical factors. The two factors have a negative 60% correlation. The values of the two factors are given in *Figure 16 and Figure 17.*

[1] https://umap-learn.readthedocs.io/en/latest/

Global Journal of Management and Business Research (C) Volume XX Issue III Version I Year 2020

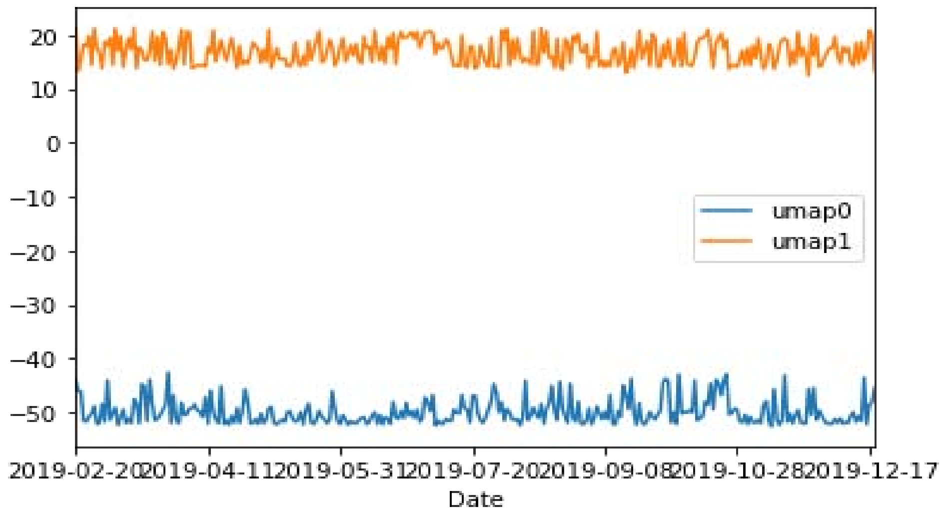


Figure 16: The Values of Two Factors

The magnitude of the factors is no longer returns but it doesn't change the explanatory power.

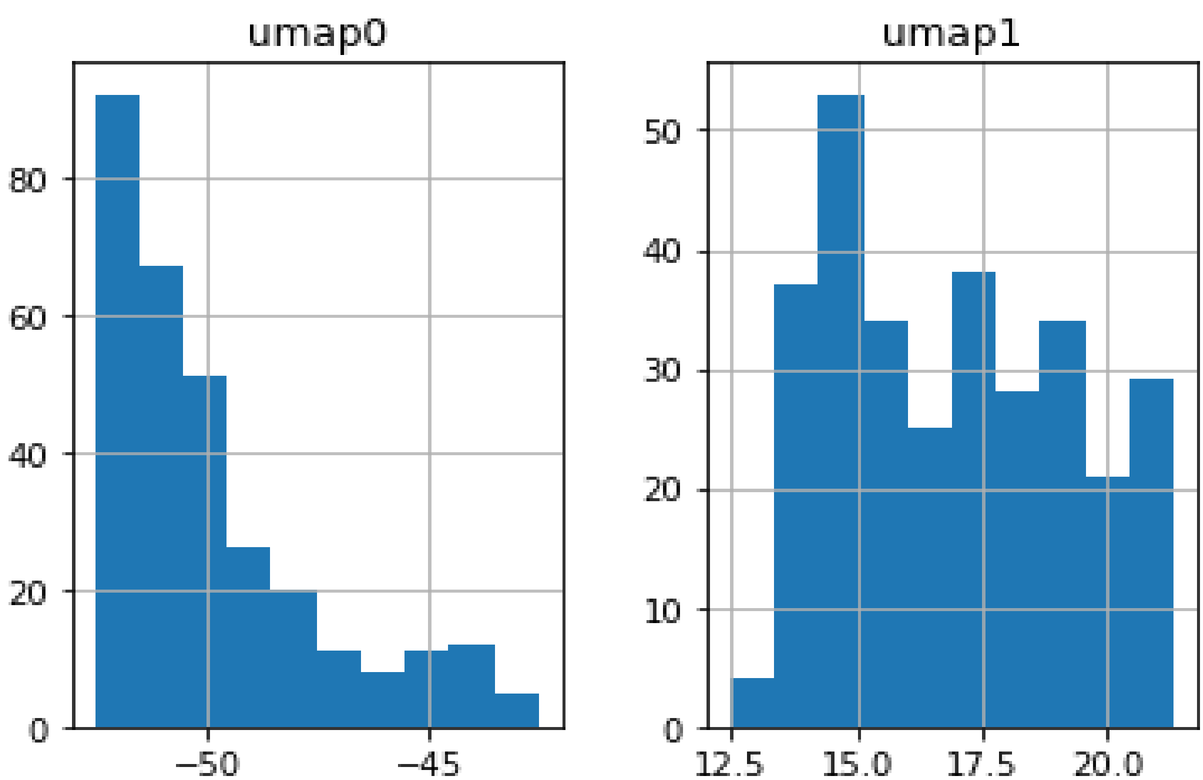


Figure 17: The distribution of two factors

UMAP0 is very right-skewed, while UMAP1 is more uniformly spread.

Table 8: The Regression Results of UMAP Factors

Dep. Variable:	y	R-squared:	0.803
Model:	OLS	Adj. R-squared:	0.801
Method:	Least Squares	F-statistic:	683.3
Date:	Sun, 15 Mar 2020	Prob (F-statistic):	4.00e-119
Time:	06:12:03	Log-Likelihood:	981.22
No. Observations:	339	AIC:	-1956.
Df Residuals:	336	BIC:	-1945.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
x1	1.6004	0.046	34.522	0.000	1.509	1.692
x2	0.6229	0.103	6.073	0.000	0.421	0.825
const	-0.0157	0.001	-21.319	0.000	-0.017	-0.014

Omnibus:	262.172	Durbin-Watson:	1.976
Prob(Omnibus):	0.000	Jarque-Bera (JB):	3162.024
Skew:	3.265	Prob(JB):	0.00
Kurtosis:	16.462	Cond. No.	141.

The x1 and x2 are the exposure to the UMPA0 and UMAP1. They both generate positive returns.

These two factors are no longer portfolio returns now. They are a low-rank representation of the market. But they serve the same purpose as the factors and can systematically explain the risk premium and have a very low correlation with the market and the size, as shown in Table 9 below.

Table 9: The Correlation between Four Factors

	Market	Size
umap0	0.16663	-0.0133
umap1	0.03975	0.01545

We see the two UMAP factors are essentially uncorrelated with market and size.

Implementation costs

An important consideration for any trading strategy is its implementation costs. Like other asset classes, trading cryptocurrency also entails costs. Cryptocurrency exchanges charge fees based on a tiered approach with a flat fee per transaction and a proportional fee based on the thirty-day trading volume for an account, which essentially means, higher the activity, greater will be the trading costs. However, one aspect is that based on the signals, a sudden big buy above the thirty-day average traded volume would entail a comparative less cost as compared to staggered buy. However this may be constrained by the general cryptocurrency market liquidity conditions. As the cryptocurrency exchanges are not regulated, hence there is no standardized fee pattern, and the respective exchanges charge the fees as per their discretion. The return from any trading strategy thus will depend on the crypto traded and the exchange chosen for the execution.

Another consideration is that some exchanges charge costs only in terms of specific cryptos, and any pay-out through the use of any fiat currency for deposit and withdrawal entails additional fees. Further, even the most well-known exchanges do not offer access to all cryptos. Some of the costs available in public domain hints that the trading costs are generally higher than those with other asset classes. A trading fees of about 0.1% to 0.2%, with fiat currency deposit fees of about 0.8%, and withdrawal fees of about 0.4%. There are also maker fees ranging between 0.01% to 0.06%.

Therefore, an empirical assessment of the trading costs to the trading strategies will be variable and would be contextual to a particular trade.

V. Conclusion

In our paper, we found that an equally weighted market factor can explain about 30% of the return premium and the size factor BMS can explain another 3% of the premium. Overall, the traditional equity market factors are not as powerful in the crypto market as compared to the equity market.

The unsupervised machine learning approaches turned out to be better in explaining the returns. Using UMAP, we successfully found two factors that can explain over 80% of the premium and are very much uncorrelated to the market and size.

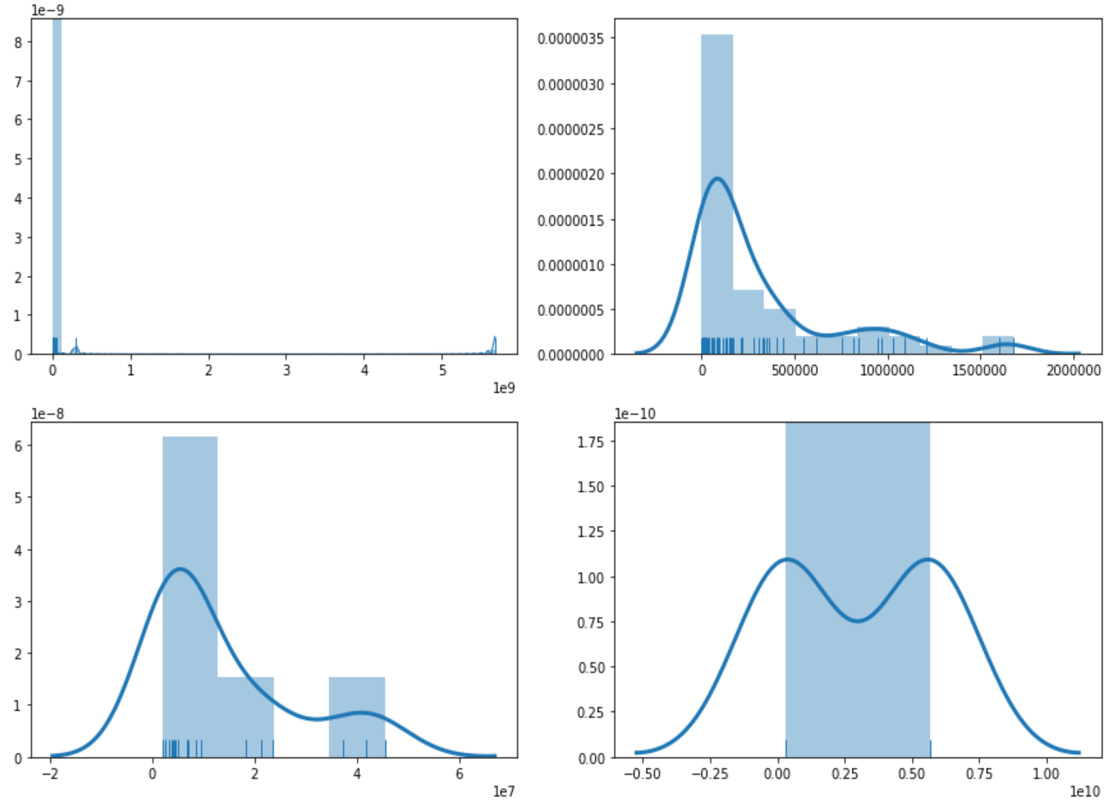
Our findings may have a considerable impact on trading cryptos. One can build their portfolio risk profile in terms of these two factors. However, our method is not flawless. Given that the cryptos market is still under development, we can only use a small sample, 341 cryptos. Because of the short time-series data, we had to conduct most analyses in-sample. It would be optimal if we could test the same factors in two years with more data. The above findings are only a starting point of our crypto factors research.

There are a few future directions we would like to take. First, since the traditional equity factors mostly failed in the crypto market, we can look at some crypto features, such as stock-to-flow ratio and mining cost. Second, finding UMAP factors is good, but figuring out what they represent may be desirable. Third, as we mentioned that the market is growing rapidly, we need to find if the new cryptos also obey the patterns we found here. Fourth, a particular consideration in this regard would be the implementation costs of these strategies and the residual premium catering in for the trading costs. We would love to do more research on this in the future.

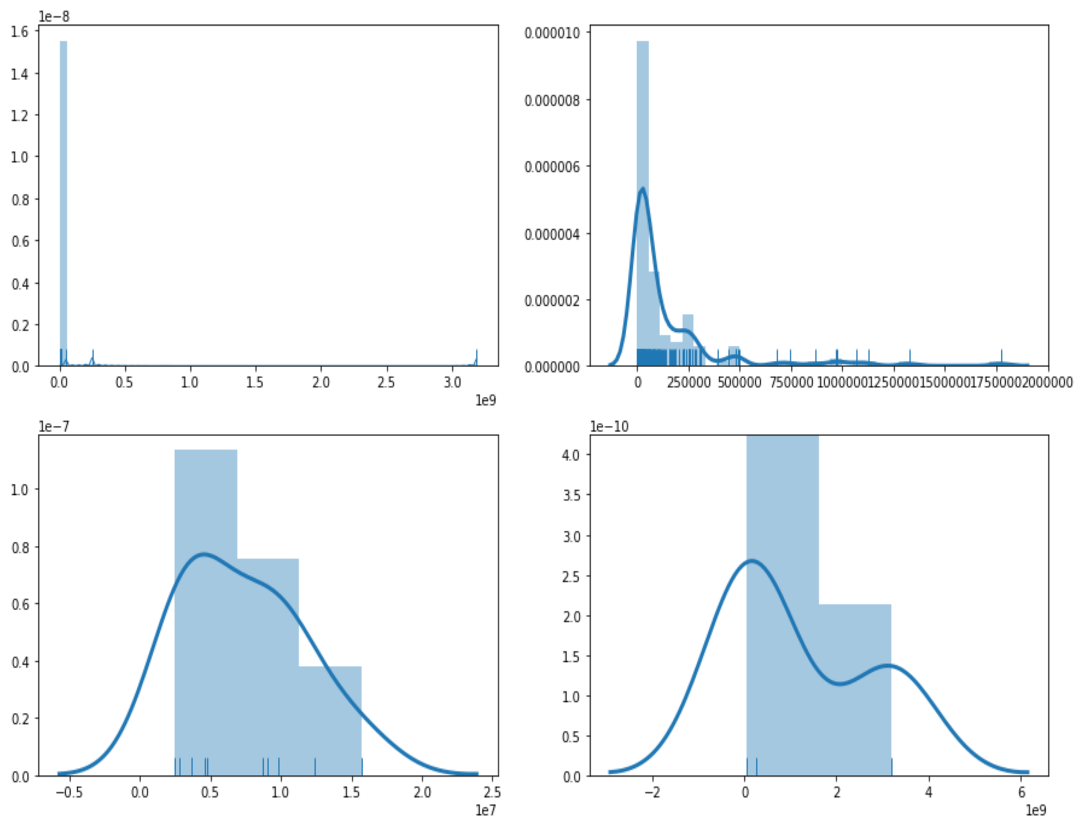
References Références Referencias

1. Benigno, P (2019), *Monetary policy in a world of cryptocurrencies*. CEPR Discussion Paper 13517.
2. Chu. J, Chan. S, Zhang. Y (2019). High frequency momentum trading with cryptocurrencies. *Research in International Business and Finance*, Volume 52,
3. ElBahrawy. A, Alessandretti. L, Kandler. A (2017). *Evolutionary dynamics of the cryptocurrency market.* Academic Press.
4. Liu. Y, Tsyvinski. A, Wu. X (2019). *Common Risk Factors in Cryptocurrency.* Academic Press.
5. Grobys. K, Sapkota. N, (2019). Cryptocurrency and Momentum. *Economics Letters*, Volume 180, 6-10
6. Khamisa. Z (2019). *An analysis of the factors driving performance in the cryptocurrency market. Do these factors vary significantly between cryptocurrencies?* Academic Press.
7. Sovbetov. Y (2018). Factors Influencing Cryptocurrency Prices: Evidence from Bitcon, Ethereum, Dash, Litcoin, and Monero. *Journal of Economics and Financial Analysis,* 2(2), 1-27.
8. Y. Yao, J. Y. (2018). Predictive Analysis of Cryptocurrency Price Using Machine Learning. *International Journal of Engineering & Technology*, 7 (3.27) (2018) 258-264.
9. Siddharth M. Bhambhwani, Stefanos Delikouras and George M. Korniotis, Do Fundamentals
10. Drive Cryptocurrency Prices?, 2019, SSRN-id3342842.pdf
11. Yukun Liu and Aleh Tsyvinski, Risk and Returns of Cryptocurrency, 2018, Yale University.
12. https://www.investopedia.com/tech/how-much-does-it-cost-buy-cryptocurrency-exchanges

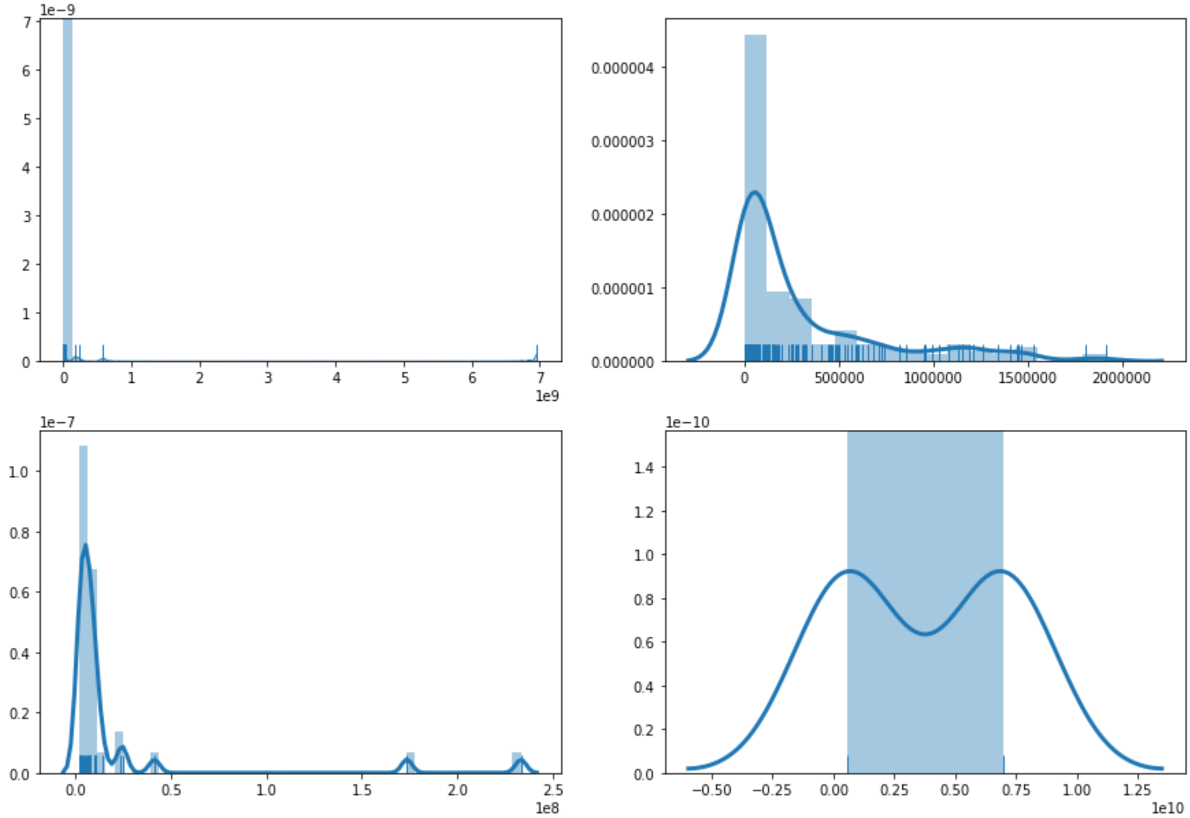
Appendix



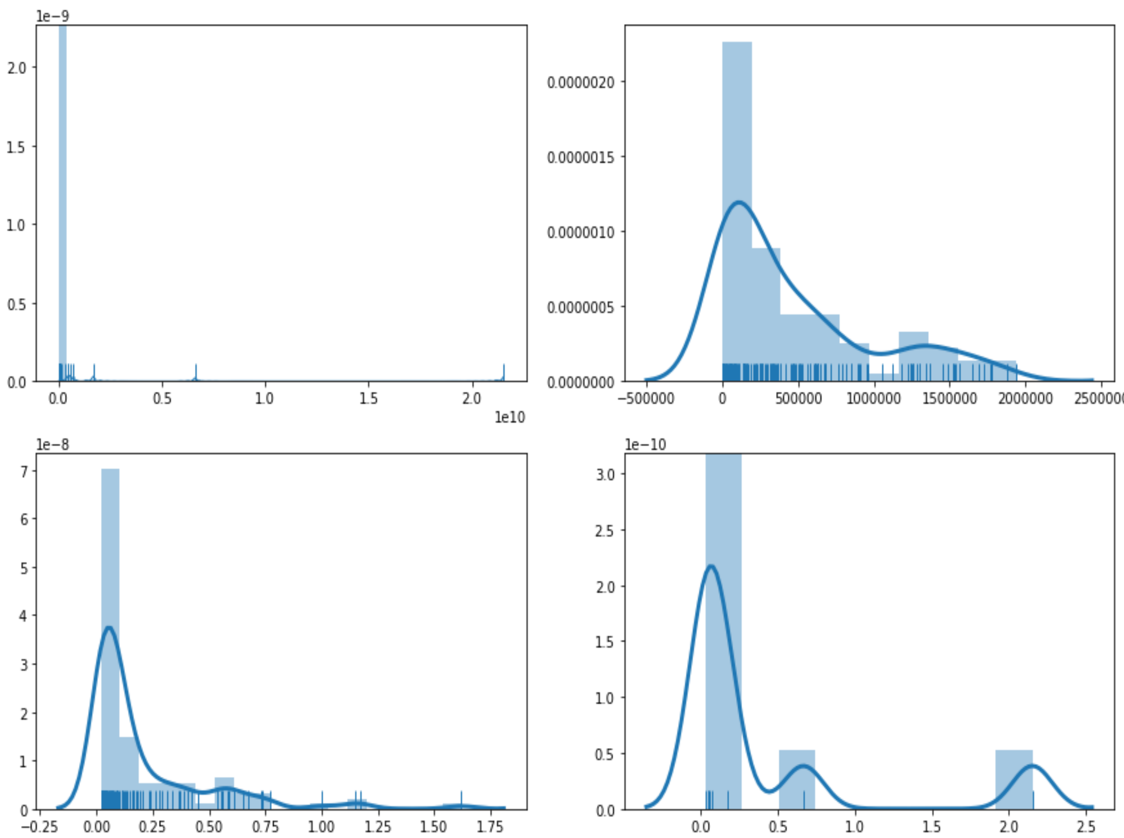
Market Cap Distribution on May 1st, 2015



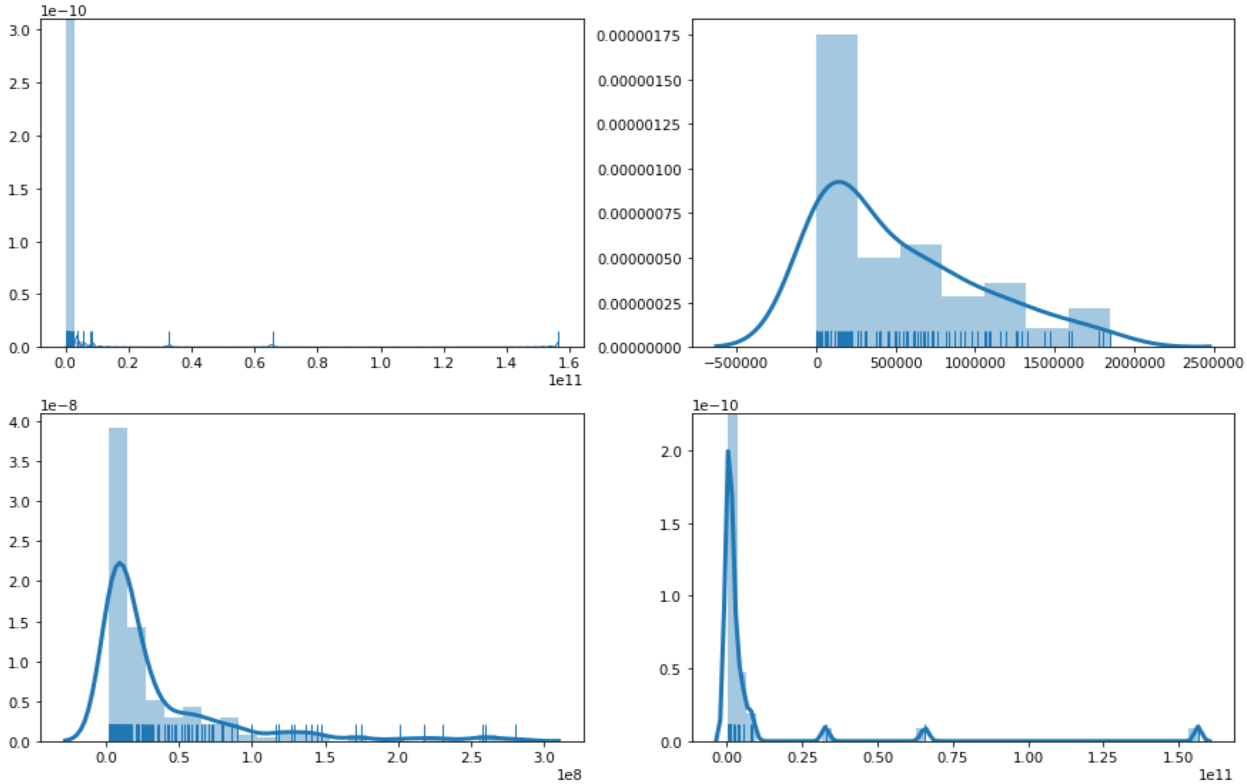
Market Cap Distribution on May 1st, 2016



Market Cap Distribution on May 1st, 2017



Market Cap Distribution on May 1st, 2018



Market Cap Distribution on May 1st, 2019