Cryptoasset data analysis

KRISTIN M. SCHWARZMÜLLER

Master Student
Computer Science: Computer Graphics and Computer Vision
Munich University of Applied Sciences
k.schwarzmueller@hm.edu

Compiled July 23, 2021

Cryptoassets have gained central access to the digital financial markets today. They are characterized by higher volatility than conventional shares or capital investments. In part this is reflected in higher profits.
They behave differently from shares, so that a special analysis of their development over time, price, intraday spread, market capitalization, or correlation to each other is necessary.
Similar to share markets, they highly correlate with each other which is evidenced by the high correlation of many cryptoassets with the Bitcoin price.

1. INTRODUCTION

To make better predictions about their future it is important to sufficiently analyze, evaluate and understand complex issues. Financial markets have always been a popular target for data analysis due to a large amount of data available and the large profits from accurate predictions.

With the initial release of Bitcoin in 2009 a new kind of digital currency without a central bank or a single administrator has been introduced, the cryptocurrencies. To verify transactions, Blockchain technology is used within a peer-to-peer network.

The more general term cryptoassets ties into the fact that some are synonymous with currencies, but others refer to dept or smart contract settlement [1].

Cryptoassets act differently than stocks: they have a higher standard deviation and most of the cryptoassets have a high correlation to Bitcoin (section 3). But other cryptoassets with high market capitalization like Ripple and Litcoin can explain only a fragment of the Bitcoin course movement, since the price fluctuations are detached from each other [2].

2. DATASET

The dataset of cryptoassets is provided by Alpha Vantage [3] and spans 1000 days covering the period from 10/25/2018 to 07/20/2021. For some currencies the data was only available for some part of the time span. The dataset contains 102 cryptoassets like Bitcoin, Ethereum, Ripple, Dogecoin, Litecoin and Monero. A full list of all used currencies can be found in section 5. For every asset, the daily values of the opening, closing, high and low price with the volume and market capitalization in USD was provided.

In order to compare the changes of the prices of different cryptoassets, they were normalized by the opening price in USD of the last observation day of the regarded asset.

3. ANALYSIS

In the following chapter certain key performance indicators are analyzed to describe their interaction. This is useful to gain an understanding of the be-

haviour of cryptoassets and to make qualified predictions for the future.

A. Average profit per day and standard deviation

A complete dataset of the observation period was not available for every asset. Because of this the average profit per trading day is prevailed in comparison to the overall profit to grant comparability among the assets. The cryptoassets VeChain, Monaco and Binance-USD were omitted in this analysis as their normalized loss was more than a hundred times lower than the average profit of the other assets.

The vast loss for VeChain, a Chinese Blockchain application platform, is due to the small dataset of only 32 trading days at an unfortunate time. For Monaco a dataset over a longer period is available but the recording starts where the payment and cryptocurrency platform was high and end where it was low on 10/23/2020. Since then the course has once again risen excessively but has also fallen again back to a relatively stable value. Binance-USD had a vast price on one day during its lifetime and this was the first day of the dataset but remained stable ever since. The standard deviation of the opening price of every day is a key performance indicator for the volatility of an asset. A high volatility of an asset raises the chance of high-profit rates but it also raises the risk of high losses.

Linear regression with the standard deviation as the independent and the average profit per day as the dependent variable over all cryptoassets show, that this risk is rewarded in higher profit rates which are reflected in the positive slope in Figure 1.

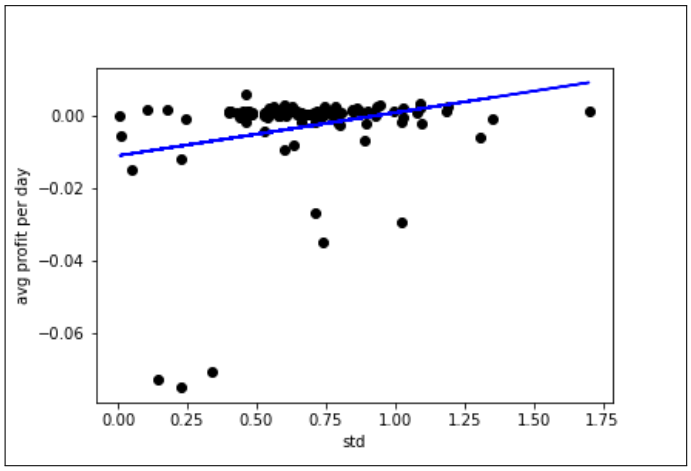

Fig. 1. Every black observation point refers to a cryptoasset so that the correlation between standard deviation of the normalized opening price in USD and the average normalized profit per day is shown. The slope of the linear regression (red line) is 0.011856.

B. Number of spread outliers and average profit per day

The spread is defined as the difference between the highest and the lowest price within a trading day. It is another measure to indicate an asset's volatility. In this test setup, an observation is specified as an outlier if its value deviates from the mean of an asset more than twice. Again, a positive correlation between volatility and profit per day is verifiable in Figure 2.

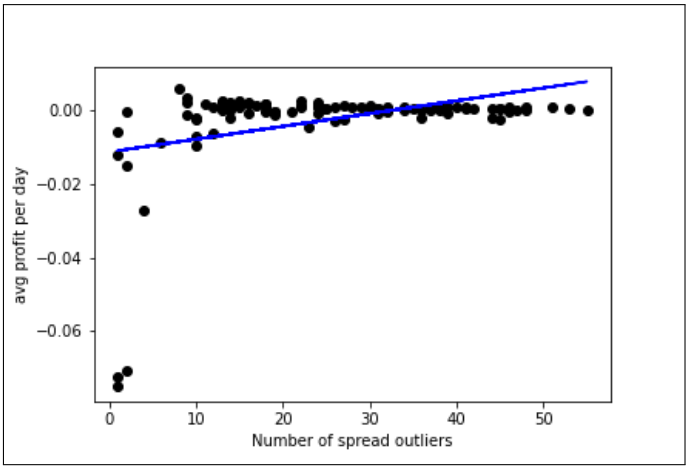

Fig. 2. Every black observation point refers to a cryptoasset so that the correlation between number of spread outliers and the average normalized profit per day in USD over the analyzed time period is shown. The slope of the linear regression (red line) is 0.000678.

C. Market capitalization and normalized standard deviation of the opening price

The market capitalization of an enterprise is the market value of a company's outstanding shares. It is the product of the number of shares times the share price. The market capitalization of a cryptoasset is the number of coins times their current market price [4]. One would expect that the higher the market capitalization is the less volatile the share price is. Regarding cryptoassets, Bitcoin is the oldest currency with the highest market capitalization. But due to an error in the compilation of the dataset, an analysis of this meter was not reasonable.

D. Correlation to Bitcoin price

Bitcoin is the best-known cryptoasset as it was the first cryptocurrency that used Blockchain technology in the world. Still, it is the most valuable cryptocurrency, widely used as an investment or means of payment and plays a major role in digital investment strategies. To get more information about the behaviour and development of cryptoassets, it is sensible to look at the relationship between an asset price and Bitcoin.

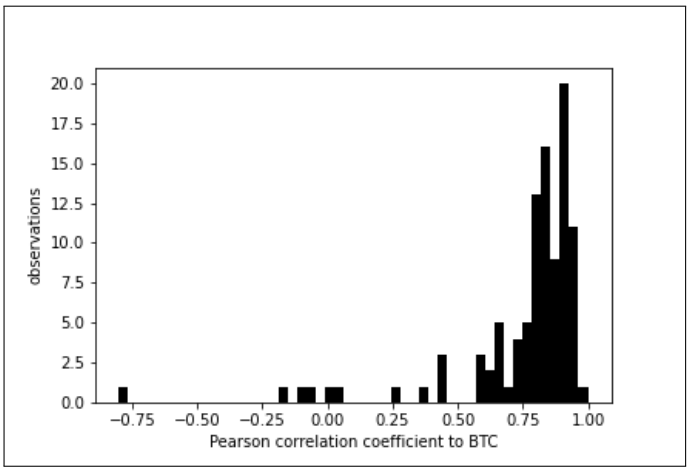

Fig. 3. Histogram of the Pearson correlation coefficient of a cryptoasset's opening price and Bitcoin.

As expected, the majority of the cryptoassets have a high Pearson correlation coefficient of over 0.5 to Bitcoin as shown in Figure 3. This means that a rising Bitcoin price is also reflected in rising asset prices, but this is not surprising as markets normally move similarly.

More astounding is the very left data point with a negative Pearson correlation coefficient of -0.80 of Stratis. Stratis offers Blockchain as a Service and is currently coded as STRAX instead of STRAT [5]. A closer look at the price development shows that this is due to the extremely volatile normalized opening price of Stratis compared to Bitcoin (Figure 4).

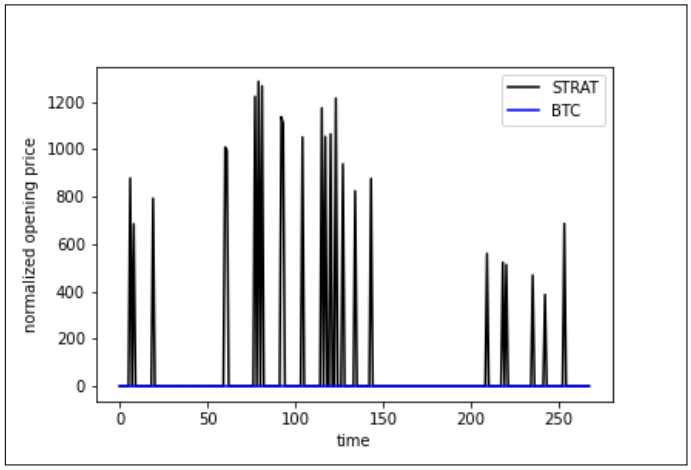

Fig. 4. Development of the opening price of Straits and Bitcoin over time whereas 0 refers to 12/11/2020 and 268 to 19/02/2020. The very low Pearson correlation coefficient of -0.80 between Straits and Bitcoin can be explained by the very high volatility of Straits.

4. OUTLOOK

The CSV files downloaded by atomized API requests from Alpha Vantage [3] show many errors as sometimes decimal separators are missing in the datasets. It is to be expected that with the improvements of the data basis e.g. by the use of another API more usable and more reliable results will be reproduced.

Furthermore, the data basis can cover a maximum of 12 years, as cryptoassets have not existed for longer. For a resilient macroeconomic data analysis, this is too short as it covers no bigger economical crises yet and no recurring wave patterns are recognized as one would expect from stock markets.

To gain more insights into the behaviour of cryptoassets, more key performance indicators need to be analyzed and statistical tests must check the findings obtained for their robustness.

5. CONCLUSION

The research field of cryptoassets is still new but knowledge from stock markets can partly be adapted. Yet cryptoassets are more volatile than stocks from companies with similar market capitalization. Based on the underlying data, it can be stated that more risky volatile cryptoassets correlate with higher profits.

Whether and which cryptocurrencies will be able to hold their own on the financial market in the long term is currently impossible to predict, as, in addition to technical details, the legal requirements and regulations must also be adapted to the new market environment.

ACKNOWLEDGEMENT

I would first like to thank my professor, Martin Mihajlov, whose expertise and passion for the subject was invaluable in gaining interest in the research questions and methodology. The insightful lecture pushed me to sharpen my thinking and brought my work to a higher level.

I would like to acknowledge Alpha Vantage [3] for providing me with a professional student account in a very short time that enabled me to build a dataset as a basis for my analytical work.

Finally, I could not have completed this dissertation without the support of my partner, Lorenz Stortz, who provided stimulating discussions as well as happy distractions to rest my mind outside of my research.

FULL REFERENCES

1. S. Corbet, A. Meegan, C. Larkin, B. Lucey, and L. Yarovaya, "Exploring the dynamic relationships between cryptocurrencies and other financial assets," Econ. Lett. **165**, 28–34 (2018).
2. A. Berentsen and F. Schär, "A short introduction to the world of cryptocurrencies," (2018).
3. A. Vantage, "Api documentation | alpha vantage," https://www.alphavantage.co/documentation/#crypto-exchange (2021). (Accessed on 07/21/2021).
4. C. F. Institute, "Market capitalization," https://corporatefinanceinstitute.com/resources/knowledge/finance/what-is-market-capitalization/ (2021). (Accessed on 07/21/2021).
5. S. Blockchain, "Stratis blockchain," https://www.stratisplatform.com/ (2021). (Accessed on 07/21/2021).

APPENDIX

currency code	currency name
AION	Aion
ALGO	Algorand
ANT	Aragon
ARDR	Ardor
ATM	ATMChain
ATOM	Cosmos
AVAX	Avalanche
BAND	Band Protocol
BAT	Basic-Attention-Token
BCC	BitConnect
BCH	Bitcoin-Cash
BLZ	Bluzelle
BNB	Binance-Coin
BNT	Bancor-Network-Token
BTC	Bitcoin
BTG	Bitcoin-Gold
BTS	BitShares
BTT	BitTorrent
BUSD	Binance-USD
CAKE	PancakeSwap
COMP	Compound
CTXC	Cortex
CVC	Civic
DAI	Dai
DASH	Dash
DATA	DATAcoin
DCR	Decred
DENT	Dent
DGB	DigiByte
DNT	district0x

Fig. 5. List of all used currencies in the data analysis (1)

currency code	currency name
DOGE	DogeCoin
DOT	Polkadot
EGLD	Elrond
ENJ	Enjin-Coin
EOS	EOS
ETC	Ethereum-Classic
FIL	Filecoin
FTT	FTX Token
FUN	FunFair
GRT	Graph
GTC	Game
GTO	Gifto
GXS	GXShares
ICX	ICON
IOST	IOStoken
IOTA	IOTA
IOTX	IoTeX
KLAY	Klaytn
KMD	Komodo
KNC	Kyber-Network
KSM	Kusama
LEND	EthLend
LINK	ChainLink
LRC	Loopring
LSK	Lisk
LTC	Litecoin
LUNA	Terra
MANA	Decentraland
MATIC	Polygon
MCO	Monaco

Fig. 6. List of all used currencies in the data analysis (2)

currency code	currency name
MITH	Mithril
MKR	Maker
MLN	Melon
NANO	Nano
NEO	NEO
NMR	Numeraire
NPXS	Pundi-X-Token
NULS	Nuls
OMG	OmiseGo
ONT	Ontology
QTUM	Qtum
REP	Augur
RLC	RLC-Token
RUNE	THORChain
SC	Siacoin
SOL	Solana
SNX	Synthetix-Network-Token
STORJ	Storj
STORM	Storm
STRAT	Stratis
STX	Stox
THETA	Theta-Token
TRX	Tronix
TUSD	TrueUSD
UNI	Uniswap
UTK	UTrust
VEN	VeChain
VET	VeChain
WAN	Wanchain
WAVES	Waves

Fig. 7. List of all used currencies in the data analysis (3)

currency code	currency name
WTC	Walton
XEM	NEM
XLM	Stellar
XMR	Monero
XRP	Ripple
XTZ	Tezos
XVG	Verge
XZC	ZCoin
ZEC	Zcash
ZEN	ZenCash
ZIL	Zilliqa
ZRX	0x

Fig. 8. List of all used currencies in the data analysis (4)