Proceedings of the 3rd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/90/20241971

Bitcoin Return Prediction: Relevant Factors Determining Return

Canzhou He[1,a,*]

[1]*Wahaha Biningual School Hangzhou, Yaojiang Rd 5, ShangCheng District, Hangzhou, Zhejiang, China*
a. Charleyhecn@hotmail.com
**corresponding author*

Abstract: Cryptocurrencies with Bitcoin being the most well-known, are volatile in price. Being able to predict or speculate price movements in Cryptocurrencies would then be meaningful for researchers and lucrative for investors. This paper hypothesized relevant factors that may determine return of Bitcoin, namely Bitcoin price data, google trend and precious metal price data, and collects data for the past 6 years. After regrouping and processing the data, a regression model is used to test the strengths of predictability of each possible factor. The results reveal that Bitcoin price data is a valid predictor of daily Bitcoin return, google trend or search volume is valid only for weekly return predictions, and precious metal prices do not have a significant relationship with Bitcoin return.

Keywords: Cryptocurrency, Bitcoin, return prediction, search volume, precious metal price.

1. Introduction

Cryptocurrency, a digital or virtual currency that uses cryptography for security [1], has been a topic of high popularity. The idea of cryptocurrency came out aiming to address the inherent weakness of the trust-based transaction model, hoping to lower the transaction cost associated with an authoritative third party [2]. To put it more plainly, cryptocurrency is a payment in digital form that does not require verification from banks. They are stored in digital wallets, and anyone can send payments anywhere, and transactions are recorded on a public ledger named blockchain. The way cryptocurrency is created is called mining, where people use supercomputers to solve complex mathematical problems. Coins are generated for solving those problems. Although by bypassing financial controllers, transaction becomes fast and smooth, controversy remains about whether cryptocurrencies should be welcomed or treated with caution. On one hand, there are countries like El Salvador that, despite warnings from the world's monetary authorities including IMF and World bank, accepted Bitcoin, a major type of cryptocurrency as an official currency [3]. On the other hand, more people are aware of cryptocurrency's risks: its high volatility and failures to meet traditional functions of money [4]. This paper will use Bitcoin as the major topic for discussion and analysis because it is the most well-know kind of cryptocurrency and traded the most.

When talking about Bitcoin, investors and researchers are interested in two major aspects: the price determination of Bitcoin and price prediction of Bitcoin. One is to analyse what gives Bitcoin its value; the other is to find ways to guide financial decisions similar to those used in the stock market. There has been relatively more sufficient research on the issue with Bitcoin's price determination, so

DOI: 10.54254/2754-1169/90/20241971

people are fairly certain what gave Bitcoin its value. However, due to Bitcoin's high volatility, research on factors that could predict Bitcoin's price is still lacking. This paper collects data of factors that could impact Bitcoin's price and possibly be used for Bitcoin's price prediction, and makes suitable hypotheses and regrouping of the data. Finally, regression is used as the basis to support for correlational relationship and strength of prediction. Specifically, factors that are taken into account are Bitcoin's price data, search volume related to Bitcoin, and precious metal prices including gold, silver and platinum.

Section 2 summarizes important related research into cryptocurrencies' price prediction, with an emphasis on Bitcoin; section 3 lists the data used in the analysis and how each of them are obtained and processed; section 4 discusses the results of regression and insights into different factors examined, and section 5 concludes the contents discussed and leaves directions for further research.

2. Literature Review

There has been increasing research regarding Cryptocurrency, with most research related to Bitcoin. Yermack [4] studied the moneyness of Bitcoin and made a valid point about the high volatility of Bitcoin, that its dollar value varies significantly over a short period of time. This makes it challenging and tempting for researchers to find valid predictors that can offer guidance on Bitcoin's price prediction. Chen, Li and Sun [5] compared machine learning methods with statistical methods and found that statistical methods are more suited for low-frequency data with high-dimensional features while machine learning methods are more suited for high-frequency data. Sovbetov [1] looked specifically into possible factors that could be used for Cryptocurrencies' price determination and found that Bitcoin's trading volume and its attractiveness (shown by google search term frequency) have significance in predicting the price of different cryptocurrencies while macroeconomic indicators like the SP500 index (measurement of performance of 500 companies in the U.S.) are insignificance predictors of Cryptocurrencies' price movements. Deniz and Teker [6] finds that gold and oil prices have no significant effect on the price of Bitcoin, Ethereum and Ripple.

3. Data

3.1. Bitcoin Price Data

Unlike other circumstances where there are exchange holidays and times when transactions are not possible, cryptocurrencies like Bitcoin is traded 24 hours a day, 7 days a week without planned closure. This allows for the direct usage of historical data from an open source. It is observed that price data from different exchange platforms differ, but they can be seen as similar and are all suffice in revealing correlational relationships. Price data is retrieved from Coinmarketcap, with the time interval being a total six years from 2018 to 2023. Because there are daily fluctuations of transaction prices, price data is unified to be the closing price of the day. The time interval is chosen so as to contain the time period when cryptocurrencies gained fullest attention from the public. Figure 1 shows the daily prices of Bitcoin.

DOI: 10.54254/2754-1169/90/20241971

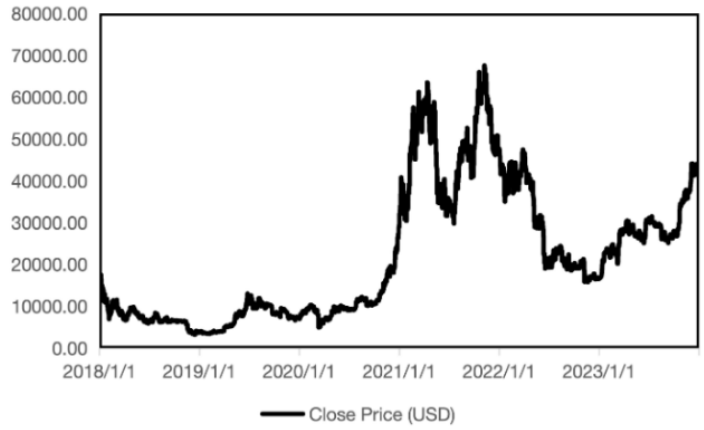

Figure 1: Daily closing prices for bitcoin during time period considered (Prices in USD)

3.2. Google Trend

To investigate public interest in Bitcoin, data is retrieved form google trend, a service provided by google that provides search volumes of terms on a scale of 0 to 100, with 100 representing the highest search volume within the time period selected and 0 being the lowest. A problem with google trend data is that it can only provide daily search volumes at a maximum of 8 months span when selecting a specific historical time interval. This brings up some difficulties to standardize the daily search volumes in 6 years. Specifically, in this paper, daily search volume is gathered every 8 months from 2018 to 2023, a total of 9 intervals. When retrieving the data, a one-day-overlap is intentionally retained to allow for the standardization of the data series. For instance, September 1st, 2018 is a 10 on the previous 8-month period but a 45 on the next period. So every data in the second interval is multiplied by the ratio 10/45 to reveal a search volume relative to the first 8 months. Likewise, every number on the third interval is multiplied by the ratio of two search volumes on the overlapping day between intervals 2 and 3 to get search volumes relative to interval 2, then multiplied by 10/45 again to get a standardized search volume relative to interval 1. The ratio goes longer and finally a standardized data set is obtained. Figure 2 shows the standardized daily search volume plotted by excel.

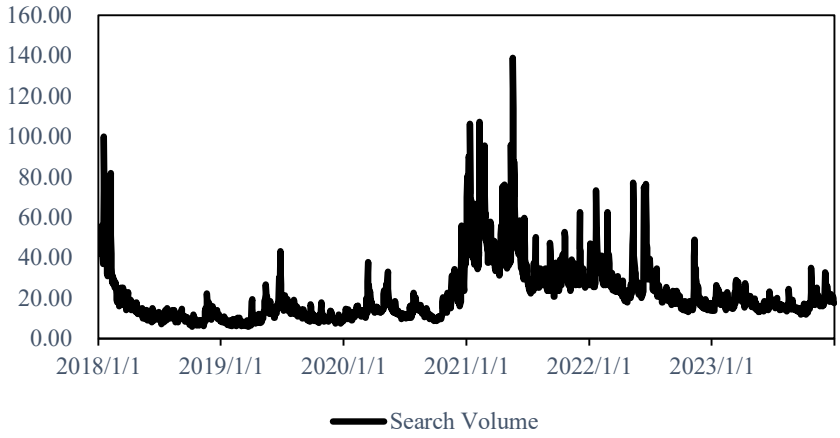

Figure 2: Standardized daily search volume by google trend

DOI: 10.54254/2754-1169/90/20241971

3.3. Precious Metal Price Data

Price data for precious metal, gold, silver and platinum to be specific, is obtained from MacroTrends, which is a research platform that enables users to screen and research stocks, commodities, precious metals, oil, gas, and global metrics. However, the market for precious metal has planned closures at weekends and in some exchange holidays. This leaves a considerable amount of blanks on daily prices within the 6 years considered. In order to keep data for other variables, which could be retrieved successfully on a daily bases, work needs to be done about the closing days of the market. Specifically, data for the previous day is continuously used for the second day if the market is closed. An IFERROR function together with an if clause is used in excel for the correction of the data set. Fig 3 shows the resulting daily prices of the three metals considered throughout the time period, with Fig 4 closely examining silver which is not distinguishable in Fig 3.

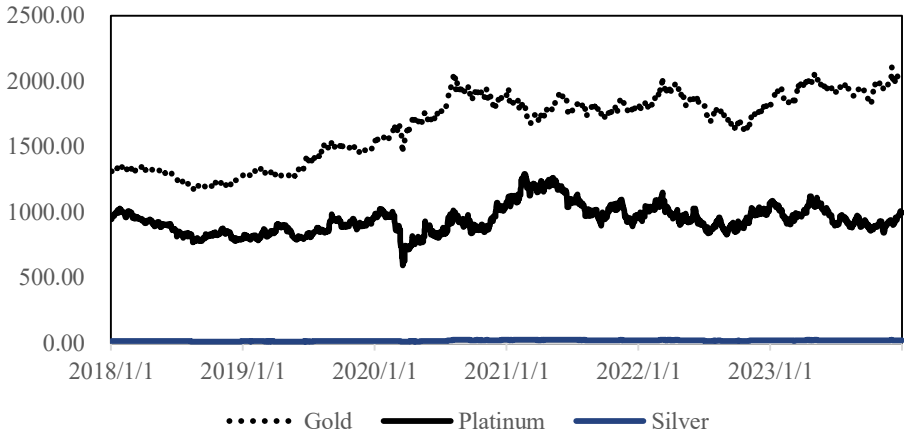

Figure 3: Daily price data for the three metals (Prices in USD)

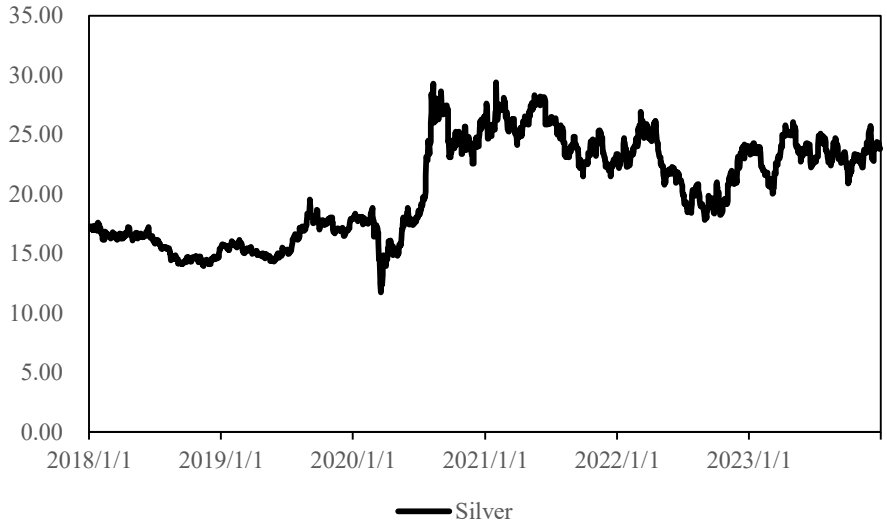

Figure 4: Daily price data for silver (Prices in USD)

DOI: 10.54254/2754-1169/90/20241971

4. Analysis and Discussion

4.1. Bitcoin Return

Measuring momentum, or specifically using the previous day's return to predict the next day's return, is a common methodology used for price prediction. It is reasonable because price for bitcoin can be seen as following a trend. Causes for increases in demand or supply would not just affect the bitcoin market for one single day. The velocity of price increase can suggest the strength of bitcoin. It is worthy to note that momentum has been seen as a stronger indicator in rising markets then falling markets throughout history. In the data collected, a correlational relationship is seen with t-statistic being -2.088 and coefficient -0.445. There is therefore sufficient evidence to prove the correlational relationship between the previous day's return and the next day's return. A step further was taken to investigate how predictive this evidence is. Specifically, the return on day t is predicted using the return on days t-2, t-3, t-4 and up to t-5. The summarized results of coefficient and t-statistics are shown in table 1.

Table 1: Summarized results of regression using return t-1 to t-5

	Coefficients	t-Statistics
t-1	-0.044536232	-2.088150112
t-2	0.043655502	2.042866336
t-3	0.006061659	0.283984203
t-4	0.007147184	0.334435392
t-5	0.025205377	1.18000215

In a nutshell, the predictability of bitcoin return can reach as far as the second day, so when using the return on day t-2 to predict the return on day t, the result is significant. There are two interesting observations: first, the correlation between returns changed from a negative correlation between return t and t-1 to a positive relationship between return t and t-2. Second, the significance of the correlation between returns fell steadily after t-2 but showed a sharp twist on return t-5, becoming close to a significant relationship. To interpret the first observation, it could be hypothesized that people expect the price of bitcoin to fall after it rose the other day. Or, it is possible that after making a profit out of bitcoin, people might act safely in order to retain the profit without any risks, so they would sell bitcoin. More supply of bitcoin may then result in a decrease in price, or at least a decrease in return. The third day, after seeing a decrease in return, people might act opposite and return would then rise. In terms of the second observation, it is possible that there are two different kinds of investors associated with Bitcoin. One of them analyzes price data and are more serious in the issue with Bitcoin. These people will probably be influenced by the previous day's price data when making their purchase decisions the next day. Another group of people might base their purchase decisions on the news or articles online. It takes time for information or news about bitcoin to travel online through the noises and reach people's ear. That may explain why for these people, a price data would guide their purchase decisions five or more days later. This is also in alignment to the hypothesis made later in the section about search volume, in which weekly return, instead of daily return, is being predicted.

DOI: 10.54254/2754-1169/90/20241971

4.2. Search Volume

It is also hypothesized that the search volume of bitcoin, retrieved from google trend, would be another indicator for return prediction. This is because search volume can be regarded as a proxy for attention. Attention may be assumed to have a positive correlation with increases in demand and thus a rise in price. When computing the significance of the relationship, however, it is shown that the relationship between google search volume and bitcoin's daily return is rather insignificant. A possible explanation for this is that reactions take time, so there is a time lag between price changes and attention. Also, it is highly possible that there are noises on the internet which interfere or act as camouflage to change in bitcoin price. Based on these explanations, instead of a daily return, correlation between search volume and weekly return is computed with the t-statistic being -1.978 and coefficient -0.0003, signaling sufficient evidence to justify the correlation.

4.3. Precious Metal Price

Being a type of cryptocurrency, bitcoin is not backed up by any commodity or precious metal, but it is still reasonable to assume that the pattern of bitcoin price changes would match with the changes in other commodities or precious metals. If the correlation stands, it is possible to tie the volatile cryptocurrency with certain stable forms of value. However, with further inspections using gold, silver and platinum, none of the price of precious metal seem to be in correlation with bitcoin return. The t-statistics are 1.190, 1.052 and -0.062 respectively, revealing insufficient evidence to conclude a relationship.

4.4. Application of Expanding Window

The analysis before was inspecting the data taking a total of six years. For the long-term valid predictors, a step further was taken to analyse the short-term predictability and reveal how many data points need to be selected for the relationship to be significant. Bitcoin return was analyzed selecting different time intervals beginning by 90 days. The relationship is insignificant at first and gradually increases significance with the more data points selected. Table 2 represents the result of the expanding window.

Table 2. Summarized regression of expanding window

Size of window	Coefficients	t-Statistics
90	-0.01464	-0.13707
180	-0.04522	-0.60485
270	-0.02667	-0.43851
360	-0.02294	-0.43443
450	-0.04344	-0.92359
540	-0.0195	-0.45353
630	-0.04449	-1.11837
720	-0.02866	-0.76996
810	-0.06821	-1.92293
900	-0.06671	-2.00651
990	-0.0682	-2.15182
1080	-0.0647	-2.13178

DOI: 10.54254/2754-1169/90/20241971

Table 2: (continued).

1170	-0.05299	-1.81516
1260	-0.06685	-2.37752
1350	-0.06502	-2.39386
1440	-0.05714	-2.17068
1530	-0.04965	-1.94232
1620	-0.05621	-2.2668
1710	-0.04987	-2.06526
1800	-0.04942	-2.09998
1890	-0.04704	-2.04791
1980	-0.04406	-1.96319
2070	-0.04416	-2.0109
2160	-0.04485	-2.08762
2192	-0.04454	-2.08815

4.5. Application of Moving Window

Expanding windows are more accurate when the underlying logic of the data remains the same. In other words, if bitcoin price data does follow a specific underlying pattern, the expanding window would give us a more accurate insight because it takes into account more data points. However, when given the high volatility of Bitcoin, a rolling window of data points is also applied. The size of the rolling window is the lowest number of data points that reveal a significant predictability according to the expanding window, that is, an interval of 810 days. By using a rolling window, the coefficient of the regression can be directly compared to reveal whether Bitcoin momentum is a stable predictor.

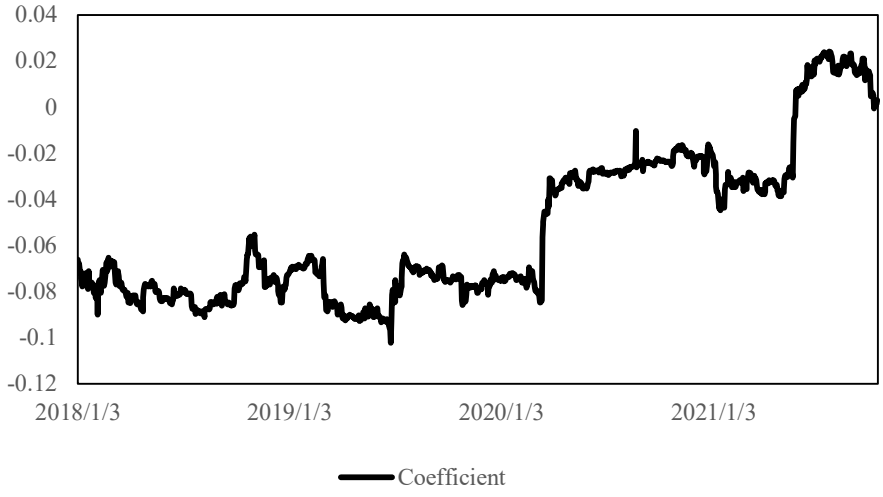

Figure 5: Coefficients of 810-day-period rolling window

However, the result shows that the coefficients of the relationship obtained by running the successful regressions vary significantly depending on the start date. With the same number of data

DOI: 10.54254/2754-1169/90/20241971

points selected, when exactly is the start date seems to play an important role in the price prediction process. This again shows the high volatility of Bitcoin, and makes the prediction using bitcoin momentum relatively risky.

5. Conclusion and Further Research

In conclusion, Bitcoin's price data, tracing back to as far as two days before, could be used to offer guidance to Bitcoin's return on the next day. Bitcoin's search volume, a proxy for attention, also has a significant relationship with its return. By contrast, Bitcoin's return seems to be independent from outside factors like prices of gold, platinum and silver. Of course, these relationships are obtained using only a simple regression model, which means that the research has room for improvements. Specifically, machine learning models identified by Chen, Li and Sun [5] as being more suitable for Bitcoin could be applied. Also, it is interesting to look into Bitcoin's relationship and possible interdependence with other kinds of cryptocurrency. Strengths of relationships between predictive factors and Bitcoin return can also be examined in the sense whether they show robustness against outside shocks like that seen during the Covid, which is something more recent and has not been touched upon as much.

References

[1] *Sovbetov Y. Factors Influencing Cryptocurrency Prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero [Internet]. Ssrn.com. 2018 [cited 2024 Jul 11]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3125347.*

[2] *Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System [Internet]. 2008. Available from: https://assets.pubpub.org/d8wct41f/31611263538139.pdf*

[3] *PWC. El Salvador's law: a meaningful test for Bitcoin [Internet]. 2021. Available from: https://www.pwc.com/gx/en/financial-services/pdf/el-salvadors-law-a-meaningful-test-for-bitcoin.pdf*

[4] *Yermack D. Chapter 2 - Is Bitcoin a Real Currency? An Economic Appraisal [Internet]. Lee Kuo Chuen D, editor. ScienceDirect. San Diego: Academic Press; 2015. p. 31–43. Available from: https://www.sciencedirect.com/science/article/abs/pii/B9780128021170000023*

[5] *Chen Z, Li C, Sun W. Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics [Internet]. 2020 Feb;365:112395. Available from: https://sciencedirect.com/science/article/pii/S037704271930398X*

[6] *Deniz EA, Teker D. Crypto currency applications in financial markets: factors affecting crypto currency prices. Pressacademia. 2020 Jul 30;11(1):34–7. Available from: https://dergipark.org.tr/en/download/article-file/1221538*