The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1086-7376.htm

Received 12 April 2023
Revised 8 August 2023
5 September 2023
Accepted 14 September 2023

Drivers of the next-minute Bitcoin price using sparse regressions

Ikhlaas Gurrib
Department of Accounting and Finance, Faculty of Management, Canadian University Dubai, Dubai, United Arab Emirates

Firuz Kamalov
Department of Engineering, Faculty of Engineering and Architecture, Canadian University Dubai, Dubai, United Arab Emirates, and

Olga Starkova, Elgilani Eltahir Elshareif and Davide Contu
Department of Accounting and Finance, Faculty of Management, Canadian University Dubai, Dubai, United Arab Emirates

Abstract

Purpose – This paper aims to investigate the role of price-based information from major cryptocurrencies, foreign exchange, equity markets and key commodities in predicting the next-minute Bitcoin (BTC) price. This study answers the following research questions: What is the best sparse regression model to predict the next-minute price of BTC? What are the key drivers of the BTC price in high-frequency trading?

Design/methodology/approach – Least absolute shrinkage and selection operator and Ridge regressions are adopted using minute-based open-high-low-close prices, volume and trade count for eight major cryptos, global stock market indices, foreign currency pairs, crude oil and gold price information for February 2020–March 2021. This study also examines whether there was any significant break and how the accuracy of the selected models was impacted.

Findings – Findings suggest that Ridge regression is the most effective model for predicting next-minute BTC prices based on BTC-related covariates such as BTC-open, BTC-high and BTC-low, with a moderate amount of regularization. While BTC-based covariates BTC-open and BTC-low were most significant in predicting BTC closing prices during stable periods, BTC-open and BTC-high were most important during volatile periods. Overall findings suggest that BTC's price information is the most helpful to predict its next-minute closing price after considering various other asset classes' price information.

Originality/value – To the best of the authors' knowledge, this is the first paper to identify the covariates of major cryptocurrencies and predict the next-minute BTC crypto price, with a focus on both crypto-asset and cross-market information.

Keywords High frequency, Lasso, Ridge, Bitcoin price prediction, Cross-market information

Paper type Research paper

JEL classification – G12, G15, G17,
The authors thank the reviewers for their constructive comments.
Funding: Not applicable.
Availability of data and materials: The data sets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.
Competing interests: The authors report that they have no competing interests to declare.

1086-7376
DOI 10.1108/SEF-04-2023-0182

1. Introduction

Against the backdrop of the COVID-19 pandemic, which has sent shockwaves through the global financial landscape, understanding the intricate dynamics that have emerged becomes paramount. The surge in inflation to levels not witnessed in four decades and the doubling of consumer pessimism about the US economy compared to the early stages of the pandemic (Bloomberg, 2022; Charm *et al.*, 2022) underscore the unprecedented challenges that have unfolded.

As analysts seek to deepen their understanding of the complex interplay within the financial realm, recent studies have investigated the performance of diverse asset classes. These investigations include the stock market (Jawad *et al.*, 2022), foreign currency exchange (Narayan, 2020), commodities (Yu *et al.*, 2022) and the ever-evolving cryptocurrency market (Di and Xu, 2022). In this context, this study directs its attention to the latter, delving into cryptocurrencies with a specific focus on Bitcoin (BTC) and its intricate price relationships with other markets. While the scrutiny of individual asset classes undoubtedly provides valuable insights, it is the exploration of cross-market dynamics that promises a deeper understanding.

The literature highlights cross-market dynamics, evident in spillover effects between the crypto market and other markets, including equities, commodities and foreign currencies. This interplay presents diversification advantages, yielding reduced portfolio risks and losses, with cryptocurrencies acting as hedging instruments amid uncertain periods. For instance, Iyer (2022) used daily data to demonstrate the increasing interconnectivity between crypto and equity markets across economies. Notably, spillover effects from BTC to global equity markets carry significance, as it accounts for around 16% (9%) of the volatility (returns) variation. Hussain *et al.* (2020) determined that investors in G7 equity markets benefit from diversification via BTC inclusion. Matkovskyy *et al.* (2021) showcased that adding cryptos to equity portfolios amplifies portfolio returns. Additionally, Bouoiyour *et al.* (2019) and Bouri and Gupta (2019) underscored BTC's hedging role in mitigating equity losses during periods of electoral uncertainty.

Numerous studies have investigated relationships between BTC and various traditional and alternative assets within the context of portfolio management and financial markets. Ghabri *et al.* (2020) and Guesmi *et al.* (2019) demonstrated that integrating BTC into portfolios with existing components such as stocks, gold and oil decreases risk. Okorie and Lin (2020) evaluated the volatility between cryptocurrencies and crude oil prices, revealing spillover effects between the two domains. Moreover, Su *et al.* (2020) established that the BTC-oil price relationship aids in diversifying portfolio risk and optimizing returns via a balanced portfolio strategy. Furthermore, Dyhrberg (2016) affirmed that cryptos, including BTC, benefit financial markets and portfolio management because of shared features with gold and the US dollar. Bouri *et al.* (2018) indicated that gold and aggregate commodity price data can help predict BTC prices. Similarly, Elsayed *et al.* (2020) showed causal links between leading cryptocurrencies and major foreign currency markets. Finally, Panagiotidis *et al.* (2018) determined that exchange rates positively impact BTC daily returns.

The relationship between cryptocurrency prices and other markets is further complicated by cryptocurrencies' inherent high volatility and their sensitivity to macroeconomic news. Geuder *et al.* (2019) identified bubble behavior as a recurring trait in BTC. Kinateder and Papavassiliou (2021) suggested greater BTC volatility at the week's outset, accompanied by a reverse January effect. Afjal and Sajeev (2022) uncovered weak, time-varying associations between cryptocurrencies and energy market return volatility. Omri (2023) scrutinized the directional predictability and volatility spillover from equity markets to BTC, exposing a significant unidirectional volatility spillover from stock markets. Conversely, Ahmed *et al.* (2023) documented bidirectional causality between S&P500 returns and BTC returns. Furthermore, Gurrib *et al.* (2019) observed a feeble connection between global macroeconomic news and major cryptocurrency returns using daily data, suggesting the necessity for higher frequency data to grasp this relationship more comprehensively.

In the context of cryptocurrencies and their interactions with established markets, further complexities emerge, shaped by a range of influencing factors. One such facet involves the utilization of different variables or markets that lack synchronization, as highlighted by Forbes and Rigobon (2002). An additional factor is the diverse analyses conducted over differing timeframes and using disparate data frequency intervals. Finally, a noteworthy complexity surfaces when attempting to integrate distinct financial products characterized by their unique risk and return attributes. A case in point is the amalgamation of cryptocurrencies, known for their extreme volatility, with the relatively steadier landscape of equity-based portfolios (Gurrib *et al.*, 2019).

While the cryptocurrency market has demonstrated substantial historical price growth, it has been accompanied by considerable volatility. Accurately predicting the behavior of BTC is challenging because of the potential influence of multiple factors, including macroeconomic events, technological advancements, government policies and cross-market dynamics. Our focus lies in understanding these cross-market dynamics to uncover the driving forces behind BTC prices and to predict its price movements on a minute-by-minute basis. The development of a reliable regression model becomes crucial in assisting market participants to comprehend the factors shaping cryptocurrency prices and enable informed predictions. Several studies, such as Seabe *et al.* (2023), Gurrib and Kamalov (2021) and Wu *et al.* (2018), have aimed to forecast BTC prices using machine learning techniques such as long-short-term memory (LSTM), support vector machine (SVM), linear gated recurrent unit, linear discriminant analysis and bi-directional LSTM. However, these studies have limited their predictive models by relying on daily data.

In this research, we build upon methodologies used in existing literature in two distinctive ways. First, unlike most studies that rely on monthly and daily data, we use higher frequency data. In particular, high-frequency forecasting, such as next-minute predictions, remains largely unexplored in the existing literature. While applications of linear regression are common in financial forecasting methods (Chan-Lau, 2017), there seems to be a tendency to increase forecasting power by including more covariates. This leads to a minimization of bias with, however, a higher variance. To tackle the potential issue of including insignificant explanatory variables and overfitting our predictive model, our methodology uses the least absolute shrinkage selection operator (LASSO) regression (Tibshirani, 1996) and Ridge regression (Hoerl, 1962) [1]. The Lasso method provides the same performance as dynamic factor models and factor-augmented vector autoregressions, with the advantage of omitting difficulties in factor interpretation (Chan-Lau, 2017).

A notable paper using LASSO high-frequency return predictors for BTC is Huang and Gao (2021). In this study, open-high-low-close (OHLC) prices and volume data over the period 2012–2019 were used. As the results suggest that the BTC open price is a strong predictor, we borrow from this work by using OHLC prices as well. More importantly, we depart from the above-mentioned work by considering the value of cross-market information. Specifically, we investigate the relationship between BTC, major cryptocurrencies and other significant asset classes using high-frequency data over the period February 2020–March 2021. The decision to use minute-based data is also motivated by the inherent high volatility observed in cryptocurrency markets. By doing so, we aim to confirm whether the insights and findings reported in studies that rely on daily data still hold true at this finer time scale. Drawing on existing literature that supports potential relationships among cryptocurrencies, stock market indices, major foreign currency pairs, crude oil and gold, we incorporate a wide range of price information from various crypto and other asset classes. All in all, the goal of this study is to develop a predictive model for the next-minute BTC price, with a focus on the value of price and cross-market information.

This study contributes to the existing literature on the linkages between BTC, cryptocurrencies and other significant asset classes, assessing whether information across markets matters in BTC next-minute prediction. Specifically, it attempts to answer the following research questions:

RQ1. What is the best sparse regression model for predicting the next-minute price of Bitcoin?

RQ2. What are the drivers of the Bitcoin price in high-frequency trading?

To answer the above questions, we build and compare several regression models. The data considered in this study consists of 598,098-min observations from February 2, 2020, 8:01 until March 31, 2021, 23:59. A total of 108 variables, including the prices of major cryptos, gold, crude oil, exchange rates and stock market indices, were considered. To capture the possibility of structural breaks, we tested for potential structural break points. Because the data had a structural breakpoint on December 30, 2020, at 14:32, the whole sample was split into pre- and post-breakpoint parts, and we studied each part independently.

While it is outside the scope of this study to consider all possible drivers as suggested by existing literature, the inclusion of major cryptos, actively traded currency pairs, global stock market indices, West Texas intermediate (WTI) crude oil and gold prices allows us to use the Lasso and Ridge shrinkage methods as a variable selection tool. The resulting model is then used to predict the next-minute price of BTC. The selected cryptos represent about 82% of the total market capitalization value of the crypto market and can be proxied as representative of the broader crypto market. Similarly, the selected exchange rates are among the most actively traded USD-based foreign currency pairs globally. Finally, the chosen global market indices, WTI crude oil and gold price information, reflect the performance in global equity and commodity markets, respectively.

The rest of the paper provides additional background on the literature review regarding cross-market linkages with the stock market, foreign exchange and commodity markets. The data and research methodology sections then follow. Finally, the research findings and the conclusive remarks are presented and discussed, along with policy implications.

2. Literature review

2.1 Inter- and cross-market linkages

Inter-market linkages include the relationships between BTC and other cryptos in terms of risk and return. Cross-market linkages include the relationships between cryptocurrencies and the stock market (through global stock market indices); the foreign exchange market; and the commodity market (WTI crude oil and gold). Mensi *et al.* (2020) used hourly data to determine correlations between cryptocurrencies such as BTC, Ethereum (ETH), Litecoin (LTC) and Ripple (XRP). The study detected significant positive dynamic-conditional correlations among cryptocurrency markets. However, a portfolio with LTC and XRP had a weaker correlation compared to a portfolio with BTC and ETH, which had the highest correlation value. Li *et al.* (2021) identified a positive cross-sectional relationship between downside risk and future returns in the cryptocurrency market. Chuffart (2022) found that conditional correlations between cryptocurrency returns of BTC, ETH and LTC have increased since 2017.

Umar *et al.* (2020) examined the relationships between cryptos and major stock market indices. The study determined the presence of an asymmetric dynamic conditional correlation over 2013–2019. Although the detected correlation is weak in general, it can also be pointed out that correlations between stock market indices and cryptos are time-varying. Frankovic *et al.* (2021) investigated the connectedness between the cryptocurrency market and Australian-listed cryptocurrency-linked stocks (CLS). The study found significant unidirectional return spillovers and weak volatility spillovers from the crypto market to CLS. A stronger relationship was observed between CLS and the engagement of crypto and blockchain technologies. Findings support that investors make trading decisions about CLS based on price fluctuations of cryptos.

González *et al.* (2021) analyzed the link between twelve cryptocurrencies and gold prices during the COVID-19 period. A statistically significant strong correlation between the gold price and eleven cryptocurrency returns (excluding Tether) was detected. With BTC being connected to gold prices during the pandemic period, it can be used to hedge investment portfolios with constituents such as gold or gold-linked assets. In addition, the least connected cryptocurrencies, such as Tether, Tezos and Cardano, can be included in portfolios for diversification benefits or even as safe-haven tools. In the same vein, Hassan *et al.* (2021) identified gold as a stable and safe-haven asset in the presence of cryptocurrency risk. Additionally, Okorie and Lin (2020) suggested including crude oil assets in the investment portfolios as part of a diversification strategy and risk hedging when there are a substantial number of cryptos as constituents. Based on the findings of González *et al.* (2021), Hassan *et al.* (2021) and Okorie and Lin (2020), we include both gold and crude oil commodities as potential representatives of commodity markets in our analysis.

Chinthapalli (2021) used daily opening and closing prices for four cryptocurrencies (XRP, BTC, ETH and Tether) and seven foreign exchange (FX) rates (EUR, AUD, MXN, GBP, BRL, SAR and JPY) for computing volatility clusters. Five-day volatility forecasts support the idea that FX markets have significant volatility patterns compared with cryptos such as BTC. Similarly, the relationships between cryptos and traditional currencies and news announcements were captured in Rognone *et al.* (2020), where findings support the idea that BTC demonstrated a different behavior compared to traditional currencies. More importantly, negative news negatively affected the returns of traditional currencies, while positive news increased returns. Comparatively, BTC responds positively to both positive and negative news.

Colon *et al.* (2021) showed that the crypto market dynamically reacts to economic policy uncertainty and geo-political risks, with reactions to uncertainty being heterogeneous. Li and Huang (2020) report that cryptos have a strong connection with technological factors and do not have any link with fundamental factors, suggesting that they represent an isolated source of risk from traditional asset classes. However, other studies, such as those by Iyer (2022), found that spillover effects from BTC to global equity markets are significant. Similarly, Hussain *et al.* (2020) reported that investors in equity markets can benefit from diversification by including BTC.

2.2 Machine learning and cryptocurrency

The application of machine learning techniques, including regression models, SVMs and random forests, has been used to analyze historical data and identify patterns or indicators that can signal price movements of cryptocurrencies (Jaquart *et al.*, 2022; Smales, 2022). Additionally, deep learning methods, particularly recurrent neural networks, convolutional neural networks (CNNs) and LSTM networks, have gained popularity because of their ability to capture temporal dependencies in sequential data (Lahmiri and Bekiros, 2019; Alonso-Monsalve *et al.*, 2020; Zhong *et al.*, 2023; Oyedele *et al.*, 2023). However, deep learning methods might be limited by the non-stationarity potentially exhibited by cryptocurrency markets, possible overfitting and substantially higher computational requirements. More frequently, a combination of machine learning methods is being used in this stream of research (Ren *et al.*, 2022). Furthermore, studies have highlighted the need to use models to aid in the identification of the most important predictors of cryptocurrency volatility, such as the Lasso and Ridge regression. Besides helping with the prioritization of key explanatory variables, these models are better suited to prevent overfitting and deal with multicollinearity whilst presenting lower computational requirements (Nguyen *et al.*, 2020; Ciner *et al.*, 2022).

Based on the above existing literature on the linkage between BTC, major cryptocurrencies and other significant asset classes, the following research question was determined: does information across markets matter in BTC next-minute price prediction? In this study, we determine which variables serve as key drivers of BTC high-frequency trading and investigate the potential linkage between traditional assets and BTC next-minute price prediction using LASSO and Ridge regression. We use minute-based data to identify the covariate of major cryptocurrencies and use the selected model to predict the next-minute crypto price of BTC. In line with previous studies, we also assess the significance of crypto-based information such as OHLC and volume prices. Further, the use of LASSO and Ridge regression helps in both prioritizing key explanatory variables and preventing overfitting and dealing with multicollinearity whilst presenting lower computational requirements. Our study adds to the existing literature by determining the best sparse regression model to predict the next price of BTC and which variables serve as key drivers of BTC high-frequency forecasting.

3. Research methodology

While the use of ordinary least squares (OLS) remains popular, two important issues arising out of OLS are prediction accuracy, where OLS tends to overfit the data, resulting in low bias with increased variance, and interpretation of results with too many explanatory variables, leading to a complex model with inefficient predictions. While the second issue can be tackled with information criteria such as Schwarz and Bayesian information criteria (SIC and BIC), the first issue can be dealt with shrinkage methods such as Lasso and Ridge, which can be used by introducing a small amount of bias and drastically reducing the variability in predictors. Lasso and Ridge regressions can be summarized as follows:

Lasso regression:

$$\min_{\beta_0,\beta_j} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{i=1}^{z} \beta_j w_{ij} \right)^2 \text{subject to } \|\beta\|_1 \leq k \tag{1}$$

Ridge regression:

$$\min_{\beta_0,\beta_j} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{i=1}^{z} \beta_j w_{ij} \right)^2 \text{subject to } \|\beta\|_2 \leq k \tag{2}$$

where y is the vector of observations for the BTC closing price, w are the covariates *and* β are the corresponding coefficients. $\|\beta\|_1$ and $\|\beta\|_2$ are the L_1 and L_2 norms and k is a user-specified parameter. In Langranian form, the Lasso and Ridge regression can be expressed as:

Lasso regression:

$$\min_{\beta_0,\beta_j} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{i=1}^{z} \beta_j w_{ij} \right)^2 + \alpha \|\beta\|_1 \tag{3}$$

Ridge regression:

$$\min_{\beta_0,\beta_j} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{i=1}^{z} \beta_j w_{ij} \right)^2 + \alpha \|\beta\|_2 \tag{4}$$

where α (alpha) represents the Lasso and Ridge penalties. As α increases, the slope of the regression line becomes more horizontal, and the regression itself becomes more robust to changes in the independent variables. The distinctiveness of Lasso is that it can shrink the bias term, which is the absolute value of the coefficients, to zero, compared with Ridge regression, which can only shrink it asymptotically close to zero. Setting α to zero in equation (3) results in an OLS regression. While the least squares estimation relates to an unconstrained minimization problem, the Lasso regression adds a convex but non-smooth L_1 constraint. The Ridge regression adds a convex and smooth L_2 constraint. For this study, we compare our results from the Lasso and Ridge regressions. An optimal alpha is sought while ensuring that the lowest mean squared error (MSE) is achieved while testing for different alphas. All 108 variables from the different crypto, FX, equity and commodity markets are first analyzed simultaneously in the Lasso and Ridge regressions. Further, because of the presence of a significant break in the data around December 2020, as found earlier, the analysis is also conducted on the pre- and post-break dates. This allows us to gauge if the same variables or specific markets (crypto, FX, equity or commodity) affect the model predictions during the stable (pre-break period) and more volatile (post-break period).

The data is normalized to have a mean of 0 and a standard deviation of 1. Normalization allows us to compare the coefficients of the variables and improves the optimization of the model. As is customary, in all of our numerical experiments, the data is divided into train and test sets. Because it is time-series data, the train set will comprise the initial 80% of the data, and the last 20% is used for model testing. Thus, for each model, the train set is used to fit the regression coefficients. Afterward, the trained model is evaluated on the test set. We calculate the MSE both on the train and test sets.

Both the Lasso and Ridge models are regularized versions of the basic OLS linear regression. The Lasso imposes an L_1 penalty on the model. Concretely, the cost function in Lasso consists of the MSE and the sum of the absolute values of the coefficients. The relative importance of the Lasso penalty is controlled by the hyperparameter alpha. A low value of alpha means less penalty and increased freedom on the values of the model coefficients, while a high value of alpha leads to a restricted model with sparse coefficients. To determine the optimal value of alpha, a grid search is applied. Lasso is often used as a variable selection model. As a conditional optimization problem, Lasso has a hypercuboid boundary that leads to sparse solutions. Comparatively, Ridge is a popular extension of OLS linear regression with an L_2 penalty. The cost function in Ridge consists of the MSE and the sum of the squared values of the coefficients. The main difference between Ridge and Lasso is the shape of the boundary over which the values of the coefficients are optimized. While in Lasso the boundary is a hypercuboid, in Ridge it is a hypersphere.

4. Data and empirical results

4.1 Data

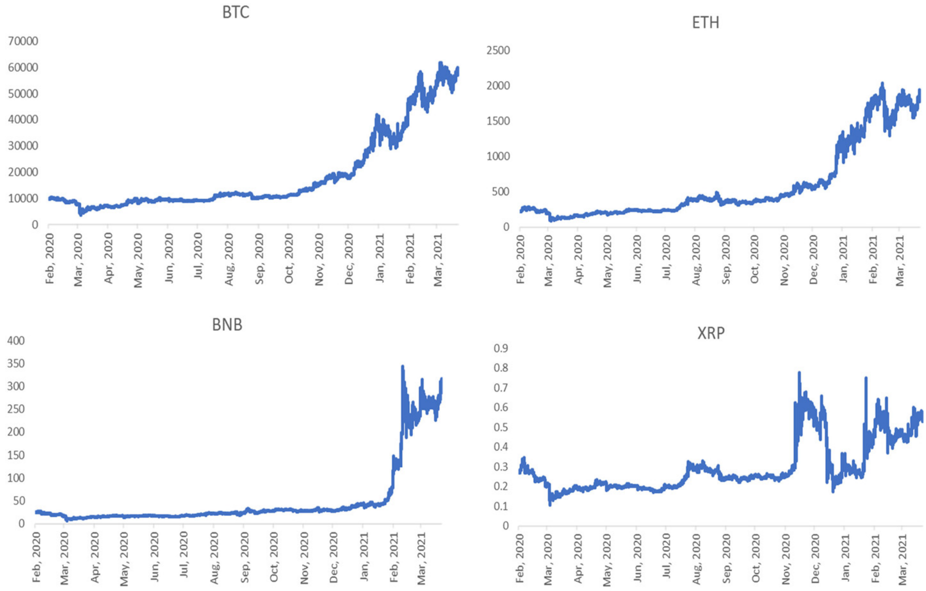
We initially chose 12 cryptos, based on their market capitalization values relative to the total crypto market capitalization values, to represent the crypto market. However, because of the period of study (February 10, 2020–March 31, 2021), four of the cryptos, namely, Polygon (MATIC), USD Coin (USDC), Solano (SOL) and Polkadot (DOT), were disregarded because they were introduced in the crypto market later than February 2020. Nonetheless, the selected ones were deemed appropriate because they represented about 82% ($654bn) of the total market capitalization values ($797bn) of the crypto market as of September 18, 2022, as per CoinMarketCap. We opted for the period February 2020–March 2021, as this period is subjected to various events such as COVID-19 and the first US vaccination against COVID-19, which affected global markets. Further studies can tap into a more recent period of study. The selected cryptos and their relative market capitalization value weights are presented in Table 1. The opening price (open), highest price (high), lowest price (low), closing price (close), volume and number of trades for each period (trade count) for the cryptos were sourced from Binance at each minute interval. BTC was trading just below $19,000 in late September 2022, an amount far from its peak of over $65,000 (November 2021). Nonetheless, BTC still holds nearly 60% of the total crypto market capitalization value, followed by ETH with a 27% weight.

We selected USD-based foreign currency pairs, in line with BIS (2019), where 88% of all trades involved the US dollar. Specifically, we chose the seven most liquid currency pairs, namely, the euro (EUR/USD), British pounds (GBP/USD), Japanese yen (USD/JPY), Australian dollar (USD/AUD), Canadian dollar (USD/CAD), Swiss francs (USD/CHF) and New Zealand dollar (USD/NZD). These currency pairs have maintained their rank as the most liquid ones since 2016. Selected stock market indices represent developed equity markets, namely, the S&P 500 (USA), NIKKEI 225 (Japan), FTSE 100 (UK), DAX 30 (Germany), ASX 200 (Australia) and NASDAQ 100 (USA). To represent the commodity market, the WTI crude oil and gold prices are used. OHLC data were compiled for the foreign exchange markets, equity market indices and commodities from Histdata. For consistency purposes, all data is set to Eastern Standard Time. For any missing data, the last available minute data is used. The whole dataset consists of 108 variables scattered over the cryptocurrency, foreign exchange, equity and commodities markets. Figure 1 (Panels A, B and C) displays the performance of the eight cryptos, stock market indices, foreign currency pairs, WTI crude oil and gold price. Because of the established strongly positive

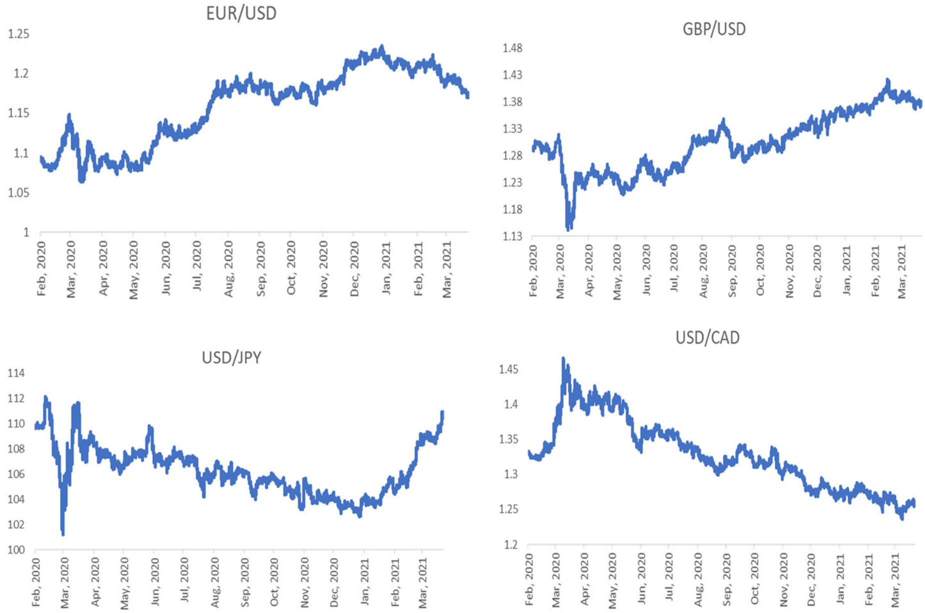
Crypto	Trading symbol	Weight (%)
Cardano	ADA	2
BNB	BNB	7
Bitcoin	BTC	58
Dogecoin	DOGE	1
Ethereum	ETH	27
Chainlink	LINK	1
Litecoin	LTC	1
Ripple	XRP	3

Notes: Table 1 represents eight selected cryptocurrencies. Weights are estimated as the proportion of each individual crypto's market capitalized value relative to the total crypto market capitalized value. The period of study is February 10, 2020–March 31, 2021
Source: Authors' own creation

Table 1.
Crypto specifications



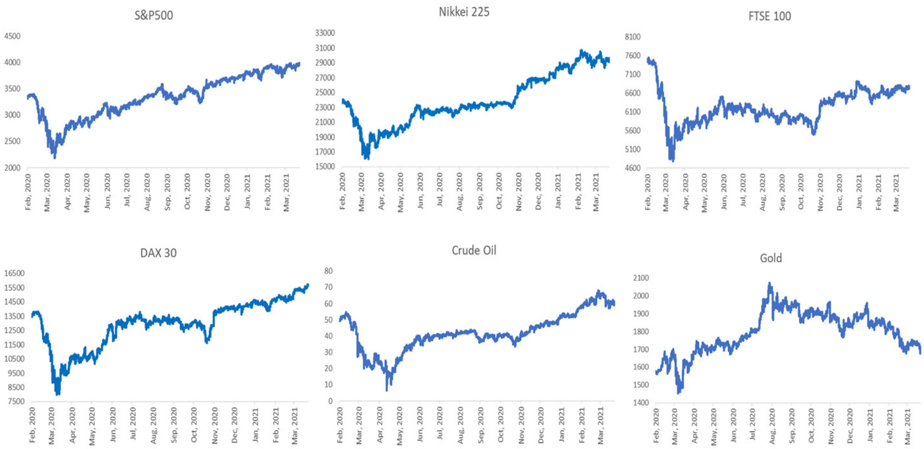
(a)



(b)

(continued)

Figure 1.
Minute-based performance of select cryptos, foreign currency pairs, equity market indices and commodities



(c)

Notes: (a) Represents the minute-based performance of selected crypto prices. These are Bitcoin (BTC), Ethereum, Binance Coin (BNB) and Ripple (XRP). The period of study was February 10, 2020–March 31, 2021. (b) represents the minute-based performance of selected USD-based foreign currency pairs. These are the euro (EUR/USD), British pounds (GBP/USD), Japanese yen (USD/JPY) and Canadian dollars (USD/CAD). The period of study was February 10, 2020–March 31, 2021. (c) represents the minute-based performance of selected major cryptos, foreign currency pairs, equity market indices and commodity markets. Selected equity markets are S&P500 (USA), FTSE 100 (UK), Nikkei 225 (Japan) and DAX 30 (Germany). Selected commodity markets (WTI crude oil and gold). The period of study was February 10, 2020–March 31, 2021. (a) Performance of selected crypto prices; (b) performance of selected USD-based foreign currency pairs; (c) performance of selected equity market indices and commodity markets
Source: Authors' own creation

Figure 1.

relationship between BTC and most major cryptos, the positive relationship among major USD-based foreign currency pairs and globally linked stock market indices, we only report the closing prices of the top four cryptos based on their market capitalization values relative to others in the list as per CoinMarketCap (BTC, ETH, BNB and XRP) in Panel A of Figure 1. The most actively traded foreign currency pairs, as reported by BIS (2019), are EUR/USD, GBP/USD, USD/JPY and USD/CAD in Panel B. The stock exchanges with the highest market capitalization of listed companies, as per Statista (2022), are S&P500, Nikkei 225, DAX 30 and FTSE100. WTI crude oil and gold prices represent commodity markets in Panel C.

As observed in Panel A of Figure 1, all the cryptocurrencies witnessed a similar price trend over the period 2020–2021. This can be explained by the dominance of BTC in the crypto market and prior studies that support strong positive correlations between BTC and major cryptos. Compared with the other cryptos, XRP experienced price hikes relatively later, around early 2021, compared with late 2020 for most other cryptos. This can be explained by the US Security Exchange Commission (SEC) event where the financial regulator filed a legal complaint against Ripple, causing the XRP price to drop from \$0.7 to \$0.2 approximately. Panel B, which reports the performance of the most actively USD-based currency pairs, supports that all the major currencies gained value against the US dollar during the 2020–2021 period. Noticeably, during the early COVID-19 impact of February–March 2020, the Canadian dollar, euro and British pounds depreciated against the US dollar, while the Japanese yen appreciated against the US dollar. The relatively early appreciation of the yen against the US dollar can be explained by the spread of infections in the USA and Europe, heightening the risk-aversion appetite and causing a flight to the Japanese yen relative to the dollar. Panel C supports the globalized effect among leading financial markets, with leading market indices such as S&P500, FTSE 100, Nikkei 225 and DAX 30 witnessing the same short-term drop in values (because of early COVID-19 impact in February–March 2020), followed by an uptrend in stock market performance. This was also observed in the crude oil market, with a downtrend noticed in the gold market from around August 2020 onwards. This can be explained by investors flying into more risky asset classes, including stocks, compared to safe-haven commodities such as gold.

Because of the rather "exponential" trend in the price of BTC, as observed in Figure 1 Panel A, it is suspected that there could be a significant break in the data. Resultingly, to test for possible multiple structural breaks at unknown dates, this study conducts Bai-Perron tests of $L+1$ vs L sequentially determined breaks (Bai and Perron, 1998). A significant break is found on December 30, 2020 (2:32 p.m.). The remaining part of the study breaks the whole data sample into pre- and post-structural break periods, with the pre-break period being February 10, 2020–December 30, 2020 and the post-break period being December 30, 2020–March 31, 2021. The decomposition of the pre- and post-break into the analysis allows us to better understand whether BTC's next-minute price can be forecasted with the same level of accuracy during a more stable (pre-break) compared to a more volatile (post-break) period. The pre-break period can be characterized as relatively more stable than the post-break period, which witnessed more price fluctuation.

4.1.1 Descriptive statistics. We captured 598,098-min observations of 416 daily prices for each crypto. Correlation values fluctuate from 0.62 to 0.98 among cryptocurrencies. As reported in Table 2, negative correlation values were found between all eight cryptos and USDCHF, USDCAD and USDJPY. Strong correlations were found among cryptos and equity market indices, as well as WTI crude oil. The only exception was for FTSE100, which had a moderate correlation with selected cryptocurrencies. Weak and negative correlation values were identified between cryptos and gold prices. Crpyto prices ranged from \$0.001 for DOGE to \$61,811 for BTC, with average prices stretching from \$0.01 for DOGE to \$18,652 for BTC. While DOGE had the smallest risk value with a standard deviation of \$0.02, BTC had the highest risk with a value of \$15,062. All cryptocurrencies were positively skewed, with values ranging between 0.95 and 2.48. Compared with all markets under study, cryptos exhibited the highest level of skewness. Equity, foreign currency and commodity markets exhibited both positive and negative skews. Except for the FTSE100, all foreign currencies, stock market indices and commodities observed negative kurtosis. Comparatively, except for LINK, all cryptos exhibited positive kurtosis, with ADA and BNB witnessing leptokurtic distributions. Gold and crude oil exhibited both negative skew and kurtosis.

4.2 Pre-breakpoint analysis

We first consider the pre-break point data from February 10, 2020, 8:01 to December 30, 2020, 14:32, which consists of 466,663-min observations. The graph of BTC over this period is given in Figure 2. We also consider the task of predicting the BTC close price at time $t+1$

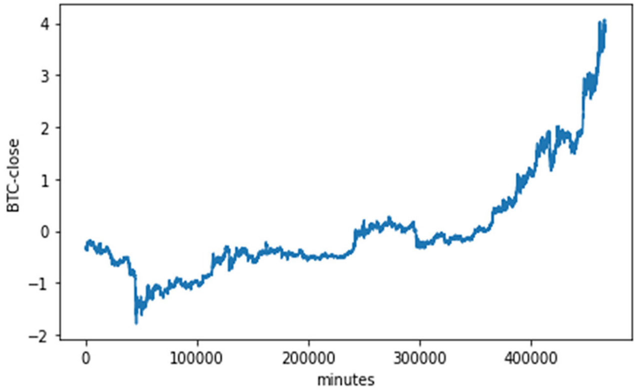
Markets	Mean	SD	Skewness	Kurtosis
Cryptos				
ADA	0.24	0.34	2.14	3.11
BNB	49.84	71.51	2.48	4.55
BTC	18,652.04	15,062.59	1.46	0.77
DOGE	0.01	0.02	2.08	2.68
ETH	584.30	530.08	1.39	0.45
LINK	11.70	8.45	0.95	−0.04
LTC	81.46	53.20	1.34	0.42
XRP	0.30	0.13	1.26	0.40
Major foreign currencies				
USD/AUD	0.71	0.05	−0.46	−0.58
EUR/USD	1.16	0.05	−0.39	−1.25
GBP/USD	1.30	0.06	−0.04	−0.69
USD/NZD	0.66	0.04	−0.18	−0.92
USD/CAD	1.33	0.05	0.37	−0.72
USD/CHF	0.93	0.03	0.27	−1.20
USD/JPY	106.34	1.97	0.42	−0.37
Equity indices				
ASX 200	6,160.43	551.31	−0.37	−0.55
DAX 30	12,944.84	1,590.49	−0.75	0.01
FTSE 100	6,231.88	450.06	0.22	0.50
NASDAQ 100	10,992.49	1,765.82	−0.39	−0.88
NIKKEI 225	23,942.18	3,432.60	0.18	−0.83
S&P 500	3,349.52	402.48	−0.41	−0.50
Commodities				
Gold	1,796.81	117.73	−0.34	−0.48
Crude oil	42.06	11.59	−0.13	−0.11

Notes: Table 2 captures all the financial instruments in the crypto, foreign exchange, equity and commodity markets used in the study. Selected cryptos are Cardano (ADA), BNB, Bitcoin (BTC), Dogecoin (DOGE), Ethereum (ETH), Chainlink (LINK), Litecoin (LTC) and Ripple (XRP). Selected US-based foreign currency pairs are the Australian dollar (USD/AUD), euro (EUR/USD), British pounds (GBP/USD), New Zealand dollar (USD/NZD), Canadian dollar (USD/CAD), Swiss francs (USD/CHF) and Japanese yen (USD/JPY). Selected global equity market indices are ASX 200 (Australia), DAX 30 (Germany), FTSE 100 (UK), Nasdaq 100 (USA), Nikkei 225 (Japan) and S&P 500 (USA). Select commodity markets are gold and West Texas Intermediate crude oil
Source: Authors' own creation

Table 2.
Descriptive statistics

using the information from time t. The input variables consist of all the variables in the dataset, including the BTC values at time t.

We first use Lasso to model the BTC price by optimizing equation (3). To determine the optimal value of the parameter alpha, we perform a grid search over a range of possible values. For each value of alpha, we train Lasso as stated in equation (3) and evaluate the model on the test set. To evaluate the models, we calculate the train and test MSE for each value of alpha. The results are presented in Figure 3, Panel A. Note that for the sake of enhanced visual comprehension, the horizontal axis represents −Log(alpha). Thus, the alpha values are scaled and reversed. The left subfigure shows that MSE decreases as the value of −Log(alpha) increases. The minimum test MSE is 0.001818, which is achieved when −Log(alpha) is III. This indicates that to achieve optimal results, the value of alpha



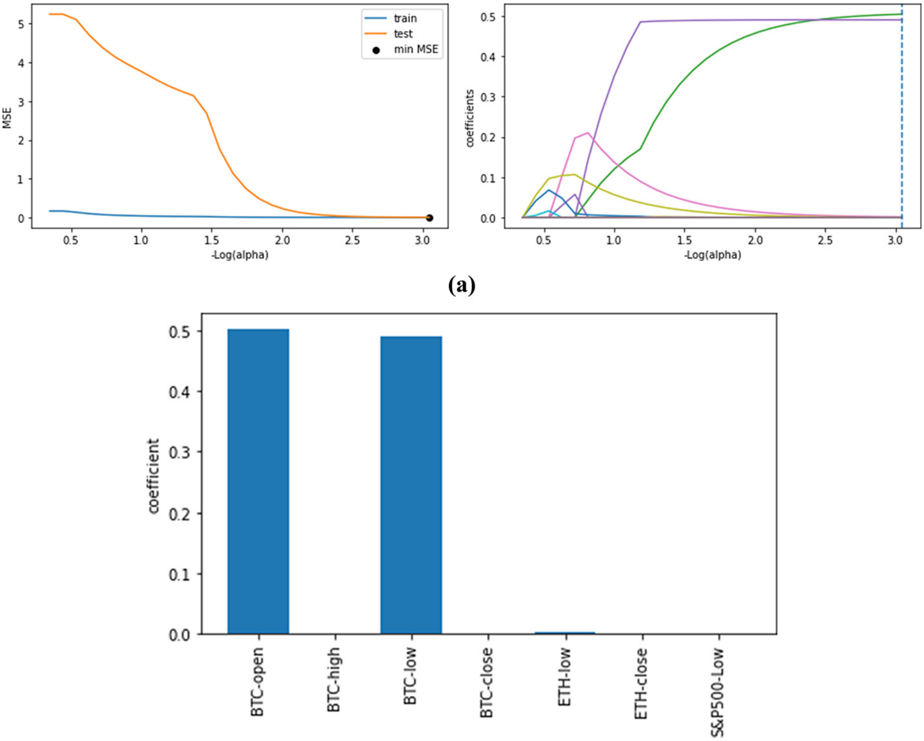
Notes: Figure 2 displays the closing price of BTC price based on the pre-breakpoint analysis. Bai–Perron tests of L+1 vs L sequentially determined breaks (Bai and Perron, 1998) were used to determine the pre-and post-breakpoint periods. The pre-breakpoint period covers February 10, 2020–December 30, 2020. There were 466,663-min observations
Source: Authors' own creation

Figure 2.
Pre-breakpoint BTC price

must be small, i.e. no regularization. On the other hand, the test MSE changes very little beyond −Log(alpha) values of 2.5. This suggests that −Log(alpha) of 3 can be considered as an acceptable optimization point. This implies that by keeping a limited amount of regularization, we improve the generalizability of the model.

In the same vein, the right subfigure shows the Lasso coefficients for each value of alpha. It can be seen that at the optimal value of alpha, as indicated by the dashed line, there are only two significant covariates. The corresponding nonzero coefficients of the model at the optimal value of alpha are presented in Figure 3, Panel B. As can be seen from the results in Figure 3 Panel B, the most significant drivers of the BTC closing price at time $t + 1$ are the values of BTC-open and BTC-low at time t. It indicates that the remaining 106 variables in the dataset are redundant. Although the original variables may influence the price of BTC, their effect is canceled given the variables BTC-open and BTC-low. This outcome is both surprising and expected. It is expected that BTC-based covariates will affect the BTC price. However, it is surprising that the two covariates are responsible for most of the model accuracy and that the other 106 covariates are essentially redundant.

Next, we consider the Ridge regression model. We also perform a similar experiment as above by training and testing the Ridge model for different values of alpha and calculating the corresponding MSE values. The coefficients of variables under Ridge are also determined for each value of alpha. The results are presented in Figure 4. We observe a situation that is similar to Lasso. The MSE decreases as the value of −Log(alpha) increases. The minimum test MSE is 2.41003e−05, which is achieved with a −Log(alpha) value of 5. We conclude that the best performance is achieved with minimum regularization. On the other hand, as in the case of Lasso, the change in the test MSE is insignificant beyond a −Log(alpha) value of 2.5, so it could be accepted as a valid optimization level. There are two important differences with Lasso. First, Ridge achieves much better test accuracy than Lasso. Second, the Ridge model has a single high-value coefficient (BTC Close) and many



(a)

(b)

Figure 3. MSE and coefficients from Lasso regression (pre-breakpoint)

Notes: Figure 3 displays the mean square error (MSE) and alpha from the Lasso regression model. For each value of alpha, the model is trained. To evaluate the models, we estimate the train and test MSE for each value of alpha. Note that for the sake of enhanced visual comprehension, the horizontal axis represents -Log(alpha). Alpha values are scaled and reversed. (a) Mean squared error (MSE), alpha and coefficients from Lasso regression; (b) non-zero coefficients from Lasso with optimal alpha
Source: Authors' own creation

low-value covariates. Thus, the Ridge model is quite different than the Lasso model, although both respond similarly to different levels of regularization.

4.3 Post-breakpoint analysis

We use a similar approach as above to investigate the data post-breakpoint from December 30, 2020, 14:32 to March 31, 2021, 23:59. The graph of the BTC close price over the second period is provided in Figure 5. It can be seen that the price is significantly more volatile than before the breakpoint.

The results of the Lasso regression are presented in Figure 6, Panel A. The test MSE decreases as −Log(alpha) increases. The minimum test MSE is 0.00012009, which is achieved when −Log(alpha) is 3. While the test MSE does decrease, it changes very little beyond a −Log(alpha) value of 1.5. This suggests it could be considered an acceptable

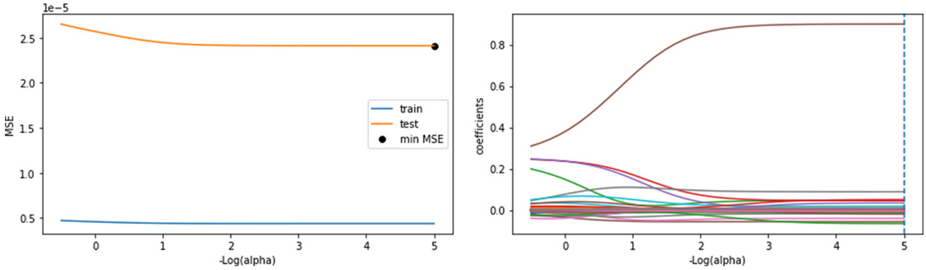


Figure 4. MSE, alpha and coefficients from Ridge regression (pre-breakpoint)

Notes: Figure 4 displays the mean square error (MSE) and alpha from the Ridge regression model. For each value of alpha, the model is trained. To evaluate the models, we estimate the train and test MSE for each value of alpha. Note that for the sake of enhanced visual comprehension, the horizontal axis represents -Log(alpha). Alpha values are scaled and reversed
Source: Authors' own creation

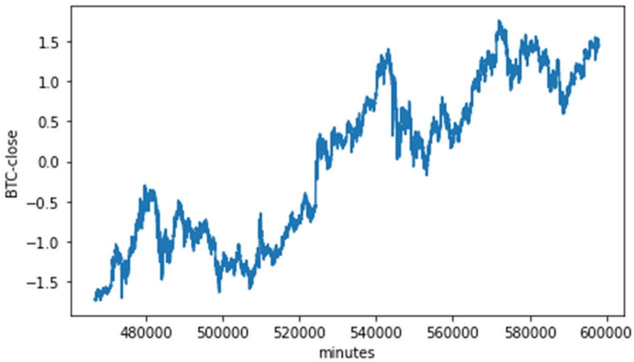
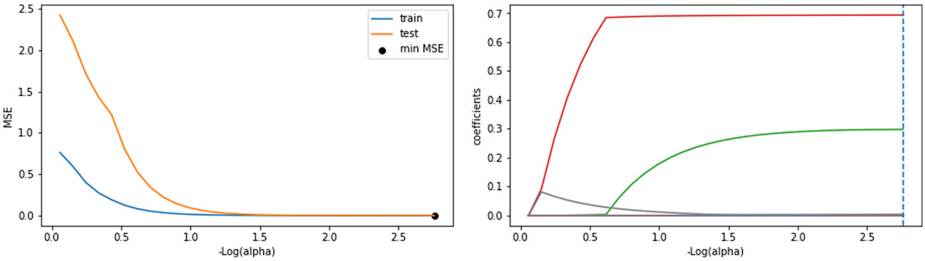


Figure 5. Post-breakpoint BTC price

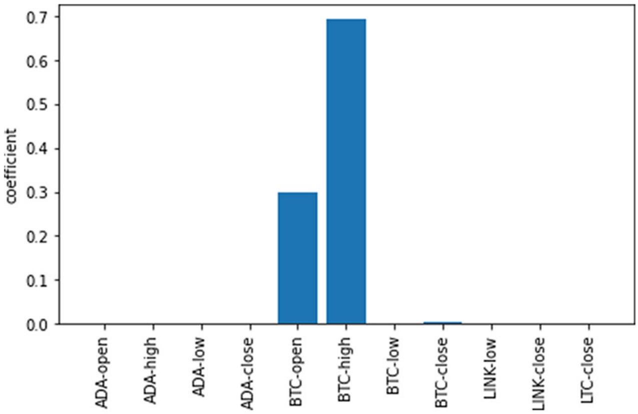
Notes: Figure 2 displays the closing price of BTC price based on the pre-breakpoint analysis. Bai–Perron tests of L+1 vs L sequentially determined breaks (Bai and Perron, 1998) were used to determine the pre-and post-breakpoint periods. The pre-breakpoint period covers December 30, 2020–March 31, 2021
Source: Authors' own creation

optimization point. The right subfigure indicates that there are only two nonzero covariates in the Lasso model based on the optimal alpha. The corresponding nonzero covariates in the optimal model are shown in Figure 6, Panel B. It reports that BTC-open and BTC-high are the driving variables in the Lasso model. It is not surprising that BTC-based variables are the most significant. However, it is unexpected to see that the rest of the 108 covariates play essentially no role in determining the price of BTC, given BTC-open and BTC-high.

The results of the Ridge regression using the post-break point period are presented in Figure 7. It can be seen that the test MSE remains essentially flat beyond a −Log(alpha) value of 1.5, which indicates that additional relaxation of the penalty does not improve the



(a)



(b)

Notes: Figure 6 displays the mean square error (MSE) and alpha from the Lasso regression model. For each value of alpha, the model is trained. To evaluate the models, we estimate the train and test MSE for each value of alpha. Note that for the sake of enhanced visual comprehension, the horizontal axis represents -Log(alpha). Alpha values are scaled and reversed. (a) MSE, alpha and coefficients from Lasso regression; (b) non-zero coefficients from Lasso with optimal alpha
Source: Authors' own creation

Figure 6.
MSE and coefficients from Lasso regression (post-breakpoint)

model performance. The minimum test MSE is 0.000040, which is achieved at around a −Log(alpha) value of 0.9. The right subfigure indicates that the optimal model consists of a single dominant covariate and multiple less significant variables.

4.4 Discussion

The above findings lead to the conclusion that the most effective regression model for predicting the next-minute BTC price is the Ridge regression based on BTC-related covariates, such as BTC-open, BTC-high and BTC-low, with a moderate amount of regularization. These findings have practical implications for traders and financial regulators. First, for traders engaged in 1-min trades, the non-significance of non-BTC-related covariates from other asset classes, such as equity, commodity and foreign exchange markets, indicates that they should consider BTC's price information (open, high and low) as a potential driver of the next-minute BTC closing price. Moreover, the study revealed that

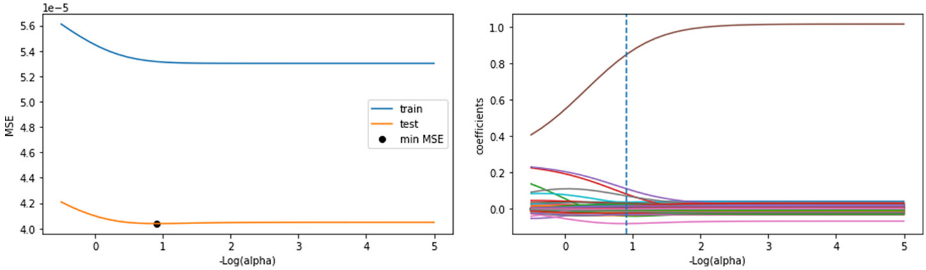


Figure 7.
MSE, alpha and coefficients from Ridge regression (post-breakpoint)

Notes: Figure 7 displays the mean square error (MSE) and alpha from the Ridge regression model. For each value of alpha, the model is trained. To evaluate the models, we estimate the train and test MSE for each value of alpha. Note that for the sake of enhanced visual comprehension, the horizontal axis represents -Log(alpha). Alpha values are scaled and reversed
Source: Authors' own creation

BTC trade counts and volume were redundant variables and did not provide significant predictive power.

Second, the analysis also indicates that the forecasting of BTC prices in the next minute is driven by different aspects of BTC's price information, depending on market conditions. The pre-break period results suggest that open and low prices were the two significant variables in predicting the next-minute closing price for BTC. In contrast, the post-break period results suggest that open and high prices became the two most important drivers of BTC's price. This highlights that during more volatile periods, traders should pay greater attention to open and high prices as key indicators.

Finally, financial regulators, such as the Commodity Futures Trading Commission, SEC and derivatives exchanges such as the Chicago Mercantile Exchange, which offer innovative financial products such as BTC futures contracts, can benefit from using open and high prices to gain insights into where the BTC market is heading in the next minute, especially during periods of heightened volatility. This information can guide regulators in better controlling market activity involved in BTC price speculation and contribute to market stability.

5. Concluding remarks

Despite the recent fintech literature that suggests the need for accuracy in BTC price prediction using cross-market linkage (Iyer, 2022; Umar *et al.*, 2020; Frankovic *et al.*, 2021), the variability in BTC prediction has not yet received significant attention. Hence, this study attempts to close the gap by using shrinkage regression techniques (Lasso and Ridge), which help eliminate some of the highly collinear variables by introducing a small amount of bias and drastically reducing the variability in predictors.

The results of the study allow us to address the main research questions raised at the beginning of the paper. First, in terms of the best regression model for predicting the next-minute price of BTC, we find that Ridge significantly outperforms Lasso. The minimum test MSE achieved by the Lasso model on the pre- and post-break point data was 0.000120 and 0.001818, while for Ridge, it was 0.000024 and 0.000040, respectively. The Ridge model is particularly more accurate than Lasso on the post-break point data. In both Lasso and Ridge, the optimal results were achieved around a −Log(alpha) value of 2, beyond which the results changed very little. Second, in terms of the drivers of the BTC price in high-frequency trading (1 min), the BTC -based covariates BTC-open and BTC-low were the most important variables in predicting the BTC closing price during the first period, while BTC-open and BTC-high were the most important in the second period. It is not surprising to observe that BTC-related covariates play a significant role in predicting its next-minute price. However, it is unexpected to find that the remaining 105 variables play almost no role in the models. The results of the study suggest that the most effective regression model for predicting the next-minute BTC price is Ridge based on BTC-related covariates such as BTC-open, BTC-high and BTC-low with a moderate amount of regularization. Future avenues of research should explore how prior *x* min can be used to forecast the next 1-min price of BTC.

In terms of policy implications, stakeholders, e.g. traders, should consider adopting the Ridge regression model over LASSO when crafting their prediction strategies. By incorporating Ridge regression techniques into trading models, it is possible to enhance the reliability of price forecasts and make more informed trading decisions. Given the significance of BTC-related covariates such as BTC-open, BTC-high and BTC-low in predicting next-minute BTC prices, traders should prioritize these factors when designing their trading strategies. Finally, analysts can experiment with incorporating shorter term historical data to potentially gain insights into short-term price movements. This approach might enable traders to capitalize on intraday market dynamics more effectively.

Note

1. As a covariate shrinkage method, Lasso constraints eliminate some of the highly collinear variables, thereby leading to a sparser model with a higher forecasting power. Further, while various authors used factor models to reduce the problem of multivariate regressions (e.g. Stock and Watson, 2002), Li and Chen (2014) supported that Lasso models tend to outperform factor-based models while removing the challenge of interpreting factor and principal components. By applying the Lasso method, Nguyen *et al.* (2020) found that crypto-based portfolios are less sensitive to downside risk and experience a strong performance during positive market anticipations. Similarly, Huang and Gao (2021) used the Lasso method to select the strongest predictive variables in BTC trading. Wang and Ngene (2020) conducted an intraday analysis to identify that BTC can forecast, in a non-linear fashion, the performance of other cryptocurrencies as it includes predictive information. Ridge regressions are included in our analysis, as they have the advantage of not requiring unbiased estimators. Instead, it adds bias to estimators to reduce the standard errors. Lasso (Ridge) tends to do well if there are a small (large) number of significant parameters.

References

Afjal, M. and Sajeev, K.C. (2022), "Interconnection between cryptocurrency and energy market: an analysis of volatility spillover", *OPEC Energy Review*, Vol. 46 No. 3, doi: 10.1111/opec.12227.

Ahmed, M.Y., Sarkodie, S.A. and Leirvik, T. (2023), "Mutual coupling between stock market and cryptocurrencies", *Heliyon*, Vol. 9 No. 5, doi: 10.1016/j.heliyon.2023.e16179.

Alonso-Monsalve, S., Suarez-Cetrulo, L., Cervantes, A. and Quintana, D. (2020), "Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators", *Expert Systems with Applications*, Vol. 149, p. 113250, doi: 10.1016/j.eswa.2020.113250.

Bai, J. and Perron, P. (1998), "Estimating and testing linear models with multiple structural changes", *Econometrica*, Vol. 66 No. 1, pp. 47-78, doi: 10.2307/2998540.

BIS (2019), "Triennial central bank survey foreign exchange turnover in April 2019", Bank of International Settlements, 16 September 2019, available at: www.bis.org/statistics/rpfx19_fx.pdf (accessed on 22nd July 2022).

Bloomberg (2022), "US consumer sentiment slumps to record low on rapid inflation, Bloomberg", available at: www.bloomberg.com/news/articles/2022-06-10/us-consumer-sentiment-slumps-to-record-low-on-inflation-woes (accessed on 15th July 2022).

Bouoiyour, J., Selmi, R. and Wohar, M.E. (2019), "Safe havens in the face of presidential election uncertainty: a comparison between bitcoin, oil and precious metals", *Applied Economics*, Vol. 51 No. 57, pp. 6076-6088, doi: 10.1080/00036846.2019.1645289.

Bouri, E. and Gupta, R. (2019), "Predicting bitcoin returns: comparing the roles of newspaper- and internet search-based measures of uncertainty", *Finance Research Letters*, Vol. 38, doi: 10.1016/j.frl.2019.101398.

Bouri, E., Gupta, R., Lahiani, A. and Shahbaz, M. (2018), "Testing for asymmetric nonlinear short- and long-run relationships between bitcoin, aggregate commodity and gold prices", *Resources Policy*, Vol. 57, pp. 224-235, doi: 10.1016/j.resourpol.2018.03.008.

Chan-Lau, J.A. (2017), "Lasso regressions and forecasting models in applied stress testing", pp. 1-34, International Monetary Fund, available at: www.imf.org/en/Publications/WP/Issues/2017/05/05/Lasso-Regressions-and-Forecasting-Models-in-Applied-Stress-Testing-44887

Charm, T., Saavedra, J.R., Robinson, K. and Skiles, T. (2022), "Growth, marketing and sales practice: the great uncertainty: US consumer confidence and behavior during inflationary times", McKinsey and Company, available at: www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-great-uncertainty-us-consumer-confidence-and-behavior-during-inflationary-times

Chinthapalli, U.R. (2021), "A comparative analysis on probability of volatility clusters on cryptocurrencies, and FOREX currencies", *Journal of Risk and Financial Management*, Vol. 14 No. 7, pp. 1-23, doi: 10.3390/jrfm14070308.

Chuffart, T. (2022), "Interest in cryptocurrencies predicts conditional correlation dynamics", *Finance Research Letters*, Vol. 46, pp. 1-7, doi: 10.1016/j.frl.2021.102239.

Ciner, C., Lucey, B. and Yarovaya, L. (2022), "Determinants of cryptocurrency returns: a LASSO quantile regression approach", *Finance Research Letters*, Vol. 49, p. 102990, doi: 10.1016/j.frl.2022.102990.

Colon, F., Kim, C., Kim, H. and Kim, W. (2021), "The effect of political and economic uncertainty on the cryptocurrency market", *Finance Research Letters*, Vol. 39, pp. 1-7, doi: 10.1016/j.frl.2020.101621.

Di, M. and Xu, K. (2022), "COVID-19 vaccine and post-pandemic recovery: evidence from Bitcoin cross-asset implied volatility spillover", *Finance Research Letters*, Vol. 50, p. 103289, doi: 10.1016/j.frl.2022.103289.

Dyhrberg, A.H. (2016), "Bitcoin, gold and the dollar – a GARCH volatility analysis", *Finance Research Letters*, Vol. 16, pp. 85-92, doi: 10.1016/j.frl.2015.10.008.

Elsayed, A.H., Gozgor, G. and Lau, C.K.M. (2020), "Causality and dynamic spillovers among cryptocurrencies and currency markets", *International Journal of Finance and Economics*, Vol. 27 No. 2, pp. 2026-2040, doi: 10.1002/ijfe.2257.

Forbes, K.J. and Rigobon, R. (2002), "No contagion, only interdependence: measuring stock market comovements", *The Journal of Finance*, Vol. 57 No. 5, pp. 2223-2261.

Frankovic, J., Liu, B. and Suardi, S. (2021), "On spillover effects between cryptocurrency-linked stocks and the cryptocurrency market: evidence from Australia", *Global Finance Journal*, Vol. 54, pp. 1-14, doi: 10.1016/j.gfj.2021.100642.

Geuder, J., Kinateder, H. and Wagner, N.F. (2019), "Cryptocurrencies as financial bubbles: the case of Bitcoin", *Finance Research Letters*, Vol. 31, doi: 10.1016/j.frl.2018.11.011.

Ghabri, Y., Guesmi, K. and Zantour, A. (2020), "Bitcoin and liquidity risk diversification", *Finance Research Letters*, Vol. 40, p. 101679, doi: 10.1016/j.frl.2020.101679.

González, M., de la, O., Jareño, F. and Skinner, F.S. (2021), "Asymmetric interdependencies between large capital cryptocurrency and gold returns during the COVID-19 pandemic crisis", *International Review of Financial Analysis*, Vol. 76, pp. 1-16, doi: 10.1016/j.irfa.2021.101773.

Guesmi, K., Saadi, S., Abid, I. and Ftiti, Z. (2019), "Portfolio diversification with virtual currency: evidence from bitcoin", *International Review of Financial Analysis*, Vol. 63, pp. 431-437, doi: 10.1016/j.irfa.2018.03.004.

Gurrib, I. and Kamalov, F. (2021), "Predicting bitcoin price movements using sentiment analysis: a machine learning approach", *Studies in Economics and Finance*, Vol. 39 No. 3, pp. 347-364, doi: 10.1108/SEF-07-2021-0293.

Gurrib, I., Kweh, Q., Nourani, M. and Ting, I. (2019), "Are cryptocurrencies affected by their asset class movements or news announcements?", *Malaysian Journal of Economic Studies*, Vol. 56 No. 2, pp. 201-225, doi: 10.22452/MJES.vol56no2.2.

Hassan, M.K., Hasan, M.B. and Rashid, M.M. (2021), "Using precious metals to hedge cryptocurrency policy and price uncertainty", *Economics Letters*, Vol. 206, pp. 1-5, doi: 10.1016/j.econlet.2021.109977.

Hoerl, A.E. (1962), "Application of Ridge analysis to regression problems", *Chemical Engineering Progress*, Vol. 1958, pp. 54-59.

Huang, W. and Gao, X. (2021), "Lasso-based high-frequency return predictors for profitable Bitcoin investment", *Applied Economics Letters*, Vol. 29 No. 12, pp. 1-5, doi: 10.1080/13504851.2021.1908512.

Hussain, S., Bouri, E., Roubaud, D. and Kristoufek, L. (2020), "Safe haven, hedge and diversification for G7 stock markets: gold versus bitcoin", *Economic Modelling*, Vol. 87, pp. 212-224, doi: 10.1016/j.econmod.2019.07.023.

Iyer, T. (2022), "Cryptic connections: spillovers between crypto and equity markets", Global Financial Stability Notes: No. 2022/01, *IMF*, available at: www.imf.org/-/media/Files/Publications/gfs-notes/2022/English/GFSNEA2022001.ashx

Jaquart, P., Köpke, S. and Weinhardt, C. (2022), "Machine learning for cryptocurrency market prediction and trading", *The Journal of Finance and Data Science*, Vol. 8, pp. 331-352, doi: 10.1016/j.jfds.2022.12.001.

Jawad, M., Naz, M., Maroof, Z., Waheed, N. and Rashid, T. (2022), "Impact of stock investment on economic performance: a comparative study of on developed and developing economies", *Quality and Quantity*, doi: 10.1007/s11135-022-01447-0.

Kinateder, H. and Papavassiliou, V.G. (2021), "Calendar effects in Bitcoin returns and volatility", *Finance Research Letters*, Vol. 38, p. 101420, doi: 10.1016/j.frl.2019.101420.

Lahmiri, S. and Bekiros, S. (2019), "Cryptocurrency forecasting with deep learning chaotic neural networks", *Chaos, Solitons and Fractals*, Vol. 118, pp. 35-40, doi: 10.1016/j.chaos.2018.11.014.

Li, J. and Chen, W. (2014), "Forecasting macroeconomic time series: lasso-based approaches and their forecast combinations with dynamic factor models", *International Journal of Forecasting*, Vol. 30 No. 4, pp. 996-1015, doi: 10.1016/j.ijforecast.2014.03.016.

Li, S. and Huang, Y. (2020), "Do cryptocurrencies increase the systemic risk of the global financial market?", *China and World Economy*, Vol. 28 No. 1, pp. 122-143, doi: 10.1111/cwe.12314.

Li, W.Z.Y., Xiong, X. and Wang, P. (2021), "Downside risk and the cross-section of cryptocurrency returns", *Journal of Banking and Finance*, Vol. 133, pp. 1-18, doi: 10.1016/j.jbankfin.2021.106246.

Matkovskyy, R., Jalan, A., Dowling, M. and Bouraoui, T. (2021), "From bottom ten to top ten: the role of cryptocurrencies in enhancing portfolio return of poorly performing stocks", *Finance Research Letters*, Vol. 38, p. 101405, doi: 10.1016/j.frl.2019.101405.

Mensi, W., Al-Yahyaee, K.H., Al-Jarrah, I.M.W., Vo, X.V. and Kang, S.H. (2020), "Dynamic volatility transmission and portfolio management across major cryptocurrencies: evidence from hourly data", *The North American Journal of Economics and Finance*, Vol. 54, pp. 1-14, doi: 10.1016/j.najef.2020.101285.

Narayan, P.K. (2020), "Has COVID-19 changed exchange rate resistance to shocks?", *Asian Economics Letters*, Vol. 1 No. 1, p. 17389, doi: 10.46557/001c.17389.

Nguyen, L.H., Chevapatrakul, T. and Yao, K. (2020), "Investigating tail-risk dependence in the cryptocurrency markets: a Lasso quantile regression approach", *Journal of Empirical Finance*, Vol. 58, pp. 333-355, doi: 10.1016/j.jempfin.2020.06.006.

Okorie, D.I. and Lin, B. (2020), "Crude oil price and cryptocurrencies: evidence of volatility connectedness and hedging strategy", *Energy Economics*, Vol. 87, p. 104703, doi: 10.1016/j.eneco.2020.104703.

Omri, I. (2023), "Directional predictability and volatility spillover effect from stock market indexes to Bitcoin: evidence from developed and emerging markets", *The Journal of Risk Finance*, Vol. 24 No. 2, pp. 226-243, doi: 10.1108/JRF-06-2022-0130.

Oyedele, A.A., Ajayi, A.O., Oyedele, L.O., Bello, S.A. and Jimoh, K.O. (2023), "Performance evaluation of deep learning and boosted trees for cryptocurrency closing price prediction", *Expert Systems with Applications*, Vol. 213, p. 119233, doi: 10.1016/j.eswa.2022.119233.

Panagiotidis, T., Stengos, T. and Vravosinos, O. (2018), "On the determinants of bitcoin returns: a Lasso approach", *Finance Research Letters*, Vol. 27, pp. 235-240, doi: 10.1016/j.frl.2018.03.016.

Ren, Y.S., Ma, C.Q., Kong, X.L., Baltas, K. and Zureigat, Q. (2022), "Past, present, and future of the application of machine learning in cryptocurrency research", *Research in International Business and Finance*, Vol. 63, p. 101799, doi: 10.1016/j.ribaf.2022.101799.

Rognone, L., Hyde, S. and Zhang, S. (2020), "News sentiment in the cryptocurrency market: an empirical comparison with forex", *International Review of Financial Analysis*, Vol. 69, pp. 1-17, doi: 10.1016/j.irfa.2020.101462.

Seabe, P.L., Moutsinga, C.R.B. and Pindza, E. (2023), "Forecasting cryptocurrency prices using LSTM, GRU, and Bi-directional LSTM: a deep learning approach", *Fractal and Fractional*, Vol. 7 No. 2, p. 203, doi: 10.3390/fractalfract7020203.

Smales, L.A. (2022). Investor attention and cryptocurrency price crash risk: a quantile regression approach. *Studies in Economics and Finance*, Vol. 39 No. 3, pp. 490-505, doi: 10.1108/SEF-09-2021-0371.

Statista (2002), "Tracking stock market developments through composite indices", available at: www.statista.com/topics/1009/global-stock-exchanges/#editorsPicks (accessed 6 April 2023).

Stock, J. and Watson, M. (2002), "Forecasting using principal components from a large number of predictors", *Journal of the American Statistical Association*, Vol. 97 No. 460, pp. 1167-1179, available at: www.jstor.org/stable/3085839

Su, C.-W., Qin, M., Tao, R. and Umar, M. (2020), "Financial implications of the fourth industrial revolution: can bitcoin improve prospects of energy investment?", *Technological Forecasting and Social Change*, Vol. 158, p. 120178, doi: 10.1016/j.techfore.2020.120178.

Tibshirani, R. (1996), "Regression shrinkage and selection via the Lasso", *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 58 No. 1, pp. 267-288, available at: www.jstor.org/stable/2346178

Umar, M., Hung, N.T., Chen, S., Iqbal, A. and Jebran, K. (2020), "Are stock markets and cryptocurrencies connected?", *The Singapore Economic Review*, pp. 1-20, doi: 10.1142/S0217590820470050.

Wang, J. and Ngene, G.M. (2020), "Does Bitcoin still own the dominant power? An intraday analysis", *International Review of Financial Analysis*, Vol. 71, pp. 1-12, doi: 10.1016/j.irfa.2020.101551.

Wu, C.-H., Lu, C.-C., Ma, Y.-F. and Lu, R.-S. (2018), "A new forecasting framework for Bitcoin price with LSTM", *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, Singapore, 2018, pp. 168-175, doi: 10.1109/ICDMW.2018.00032

Yu, Y., Guo, S. and Chang, X. (2022), "Oil prices volatility and economic performance during COVID-19 and financial crises of 2007–2008", *Resources Policy*, Vol. 75, p. 102531, doi: 10.1016/j.resourpol.2021.102531.

Zhong, C., Du, W., Xu, W., Huang, Q., Zhao, Y. and Wang, M. (2023), "LSTM-ReGAT: a network-centric approach for cryptocurrency price trend prediction", *Decision Support Systems*, Vol. 169, p. 113955, doi: 10.1016/j.dss.2023.113955.

Further reading

Lee, J.H., Shi, Z. and Gao, Z. (2022), "On Lasso for predictive regression", *Journal of Econometrics*, Vol. 229 No. 2, pp. 322-349, doi: 10.1016/j.jeconom.2021.02.002.

Corresponding author

Ikhlaas Gurrib can be contacted at: ikhlaas@cud.ac.ae