Proceedings of ICFTBA 2025 Symposium: Data-Driven Decision Making in Business and Economics
DOI: 10.54254/2754-1169/2025.BL30746

The Predictive Power of Bitcoin Return for American Major Stock Indexes Return

Yu Gao[1*†], Bowen Liu[2†], Yishan Yang[3†], Junxiong Huang[4†]

[1]SAIS, Johns Hopkins University, Washington, USA
[2]College of Arts and Science, Boston University, Boston, USA
[3]Weatherhead School of Management, Case Western Reserve University, Cleveland, USA
[4]Department of Information & Computer Science, University of California Irvine, Irvine, USA
**Corresponding Author. Email: bjyugao97@163.com*
[†]These authors contributed equally to this work and should be considered as co-first author.

Abstract. Virtual currency has become one of the most sought-after alternative assets in the past decade with bitcoin being a leading example. value leapt from its starting price of \$0.0025 to increase by more than 40 million times that amount, creating one of the greatest rises in value in the entire history of finance. In the past few years, many academic studies show that even though Bitcoin runs independently from traditional finance, but still there is a high correlation between Bitcoin and stock market. In particular, following the introduction of Bitcoin options back in 2017, Bitcoin now appears more predictive of stock return movements than before. Research by Afees A. Salisu and his coworkers display that a solitary Bitcoin price prediction model using an optimized predictive regression framework notably surpasses older ones. but don't say how long this goes on Therefore this research will go to try and determine the time frame when Bitcoin is better at predicting the future of the stock market as opposed to stock options. Also, we'll use machine learning techniques to train machine learning models to predict the movements of the stock market and see if they work.

Keywords: Asset Pricing, Cybercurrency, Stock Indexes' Pricing, Emerging Market, Market Correlation

1. Introduction

Bitcoin has gone from a kind of sideshow curiosity to an actual liquid, institutional asset that is very frequently moving on macro news before others. With its 24/7 trading, volatile prices and increasing participation of both retail and institutional investors, there is no reason for us to assume that crypto prices do not include information about global risk appetite before local (European) equity prices are fully reflected. This observation motivates the question at the heart of this paper: Does Bitcoin have incrementally practical, usable, predictive information for regional equity indices, and if so, over which time horizons, for which markets?

There is a suggestion in the existing research but no complete answer. Early studies characterized Bitcoin as a diversifier with low correlation to equities; more recent work documents state-

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2025.BL30746

dependent co-movement that intensifies during stress or risk-on regimes. On prediction, parsimonious regressions for crypto sometimes beat naïve benchmarks, hinting that Bitcoin may carry broader forecasting content. Yet three gaps persist. First, most papers do not compare multiple forecast horizons within a unified framework, making it hard to judge how long any lead-lag effects last. Second, evidence on regional heterogeneity is sparse: effects may differ between developed and emerging markets, or between tech-tilted and broad indices. Third, few studies bridge identification and usability—linking structural, economically interpretable effects to genuine out-of-sample performance—while carefully handling mixed-frequency data and overfitting risks.

This paper addresses those gaps with a goal-driven, data-disciplined design. We assemble daily series for Bitcoin and a panel of representative regional equity indices, align macro controls with principled down-/temporal-disaggregation, and document all joins, filters, and diagnostics. For identification, we estimate a VARX specification to quantify the dynamic transmission from Bitcoin to equities while conditioning on standard market determinants. We summarize the magnitude and persistence of effects using impulse-response functions (IRFs) and horizon-specific variance decompositions. For prediction, we convert IRF summaries and related statistics into engineered features for a supervised support-vector machine (SVM) that forecasts next-period direction. Validation uses strict train/test splits, cross-validation, and robustness checks (stationarity tests, alternative horizons, regime subsamples, and outlier handling). Evaluation focuses on direction, AUC, and simple U-based metrics that talk about portfolios.

Here we make three contributions: (1) We deliver a multi-horizon map of Bitcoin's transmission to equities, disciplining crypto shock claims. (2) We offer a regional comp, revealing where the signal is strongest (developed vs. emerging; tech-tilted vs. broad), informing tactical tilts and hedge overlays. (3) We link mechanism to usability: showing when VARX/IRF effects are economically interpretable & translate into out-of-sample gains and when not. the bigger signifcance is practically: we do not declare that crypto "leads" equities everywhere, but rather where and when and how does bitcoin add to forecast value, and the cases in which one should be ignoring BTC. This gives investors something to direct their limited research capacities to, calibrate risk budgets by, and design clear, testable analytics that actually do improve decisions without promising more than the numbers can stand up for.

2. Literature review

Academic discussion about the connection between Bitcoin and financial market has changed greatly during the last decade. Early studies tended to say that Bitcoin wasn't much connected with regular asset types and could count as something that would vary. Like Bitcoin has weak correlations compared to equities and bonds, Baur, Dimpfl, and Kuck [1] found this showing that Bitcoin is more speculative rather than a hedge. Similarly, Corbet et al. [2], point out that cryptocurrencies were operating separately from the global financial system and thus not much predictive for equity markets

In literature it is seeming that bitcoin has increased its involvement with the stock markets, especially when there has been more volatility. Bouri, Salisu et Gupta [3] demonstrate that bitcoin prices can augment the forecast of realized volatility in U.S. stock sectors, especially during tumultuous times. The authors from LópezCABARCOS et al [4] further prove that Bitcoin has volatility due tothe S&P 500 and the VIX, investor sentiment is a very important transmission pathway. The researchers in Eross et al. [5] study intraday market linkages and conclude that Bitcoin's volatility becomes more intense when U.S. and European equities overlap, which emphasizes trading windows and market microstructure.

Bitcoin for predicting future events due for being acknowledged thus more advanced forms of prediction have been investigated Early approaches made use of econometric models like GARCH, but more recent studies use machine learning techniques to capture the nonlinear features of cryptocurrency markets. Catania and Sandholdt [6] compare GARCH - type models with machine learned alternatives and find the later often do better. Dudek [7] also mentions that nonlinear models such as the Support Vector Machine (SVM) model forecast the cryptocurrency volatility better than the traditional statistical model

Another widely used machine learning methods is Random Forest (RF). It has also been used for cryptocurrency prediction task. Kristjanpoller and Minutolo [8] present a hybrid volatility forecast model combining GARCH, neural networks, and random forests to identify that ensembles are superior. Chen, Zhu, and Wang [9] tested different ensemble learning approaches of Bitcoin volatility and confirmed that using machine learning algorithms together produces a good forecast model compared to using a single model. This shows how very concrete and useful something RF can be in terms of explaining things that are hard to explain such as lag return, volume, Macro-economic variable.

Others contribute by applying machine learning to broader prediction tasks. Mallqui and Fernandes [10] use machine learning classifiers like SVM to predict the direction of daily Bitcoin returns and show that these models can capture short-term momentum and reversal patterns. Lahmiri and Bekiros [11] utilize deep - learning models and find that nonlinear methods are especially effective for predicting highly-volatile cryptocurrency market. To get this all done these studies have increasing agreements that machine learning methods do better with curvature, regime change and noise in the digital assets markets.

3. Methodology

3.1. Data

3.1.1. Data acquisition and description

On the dependent variable side, we choose a group of different stock index types for calculation. They would be drawn from the main financial markets which accept cryptos: the US. The US markets data will be collected from the large-cap and small-cap stock indices to validate the model. Independent variable is Bitcoin 5-day moving average trading price or Bitcoin's daily median price. For control variables, will include a few macroeconomic measures. Including inflation and interest rates.

And secondly, taking into account that traditional macroeconomic series are usually updated every month, which does not meet the frequency requirements for our analysis, we intend to introduce daily data that reflect the economy's macro expectations and market sentiment. Just as in the examples for the VIX index, and of the prices on other goods like gold and crude.

As for further elaboration, U.S. equities index data are the Dow Jones Industrial Average, the S&P 500, the NASDAQ index, the S&P Small-Cap 600 Index and Russell 2000 Index. As I have a curiosity about Chinese assets, so Golden Dragon Index was taken into consideration. As well as the main indices, the following macroeconomic variables were collected for each region corresponding: Consumer Price Index (CPI), Producer Price Index (PPI), overnight inter-bank lending rate, implied volatility of major equity index options such as the VIX in U.S., crude oil and gold prices. These macroeconomic variables are included to remove the overall market and economic expectation

DOI: 10.54254/2754-1169/2025.BL30746

impact on stock return, so that this model will properly show how the liquidity changes due to price changes of Bitcoin is affecting the stock markets.

All data were sourced from the Bloomberg Terminal. Considering that Bitcoin futures trading began on December 11, 2017—which enabled institutional investors to enter the market, making trading more rational and structured—our sample period starts from December 11, 2017, and ends on September 6, 2025.

Additionally, since the transition from LIBOR to SOFR (Secured Overnight Financing Rate) in the U.S. market started on April 1, 2018, data for the U.S. market analysis also begin from this date.

3.1.2. Data preprocessing

Due to the differences between Bitcoin and stock market trading, this study will standardize the data using the following methods. First, since Bitcoin is traded seven days a week while stocks are only traded on weekdays, Five-day Moving average of Bitcoin daily trading price will be used to analysis the model. For the missing stock index trading data on holidays and weekends, the data from the most recent previous trading day will be adopted to fill in the gaps. The monthly macro data will be disaggregated directly into daily frequency—an approach that, in theory, preserves its economic interpretability.

This research will conduct ADF tests on all variables, regardless of their frequency, to determine whether they are stationary (I(0)). Given that stock prices and Bitcoin prices are typically non-stationary and integrated of order one (I(1)), we will apply differencing to these series. For example:

$$\mathrm{ReturnBTC} = \frac{\mathrm{BTC\_price_t\ \text{-}\ BTC\_price_{t\text{-}1}}}{\mathrm{BTC\_price_{t\text{-}1}}}$$

Following this processing, a certain degree of correlation between stock index returns and Bitcoin returns can be observed upon preliminary examination, as illustrated by the case of the S&P 500. There’re some correspondences in extreme shifts of bitcoin returns as well, like the highfrequency swings in both assets in 2020. But in terms of what can be seen with the naked eye, Bitcoin has a lot more volatility than the S&P 500: The actual relationship as well as the ability of Bitcoin returns to predict, will be tested again through the second model.

DOI: 10.54254/2754-1169/2025.BL30746

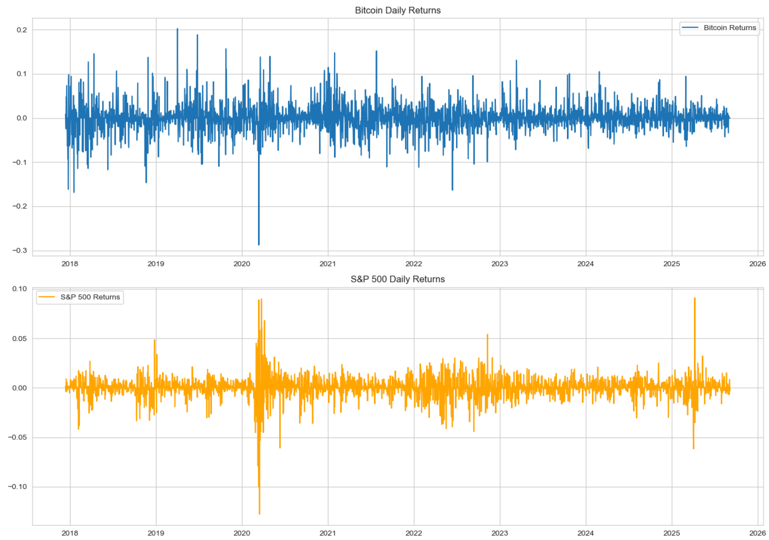

Figure 1. The comparison of bitcoin daily return and S&P 500 daily return

3.1.3. Data visualisation

First of all, based on the trend chart, we have some general impression on the relations among variables: According to theory, both Bitcoin and all of the different stock indices should have experienced similar macroshocks like changes to the US risk-free interest rates or change in investors macroeconomic expectations. On a fundamental level though, Bitcoin is not really an investment asset because there is no future cash flow – a huge difference compared to shares. As shown in figure 2, the normalized price trends are easily seen, with Bitcoin showing far more volatility compared to the major U.S. stock indices. On the contrary, there appears a degree of co-movement when it comes to fluctuation, but their trends have diverged as far as upward trends goes, Nasdaq Index grew at an impressively faster rate.

In line with Gil-Alana et al.'s conclusion, the cryptocurrency market still holds very little dependency to traditional financial markets which means diversification may still be possible. So this means that Bitcoin could have neither much relationship (or correlation) nor ability (or predicting power) to tell what's going on in regular, normal stock markets. The Correlation Heat map on appendix, the relation between bitcoin and stock index level has a correlation, however when looking at the returns of the same the strength of the correlation has greatly reduced. these 2 returns' correlation coefficient is like zero, about 0.1.

There is a difference, and that difference could be accounted for by observing an upward trend that occurred over our time period, including in 2019 with the occurrence of the COVID-19 pandemic, where financial assets were affected all around the world. Noteworthy as well is that Bitcoin had a spectacular and nearly irreversible rally soon thereafter. Even after a 2022 correction, it quickly bounced back in less than a year time before continuing to trend upwards. It could be related with a lot of liquidity around the world or it might be because people see Bitcoin as a safe place to keep their money.

These observations indicate that should global investors take Bitcoin into their portfolio calculations, it could impact their choice of traditional financial markets at the same time. In turn, this channel can give Bitcoin some kind of predictive capacity on equity index movements.

DOI: 10.54254/2754-1169/2025.BL30746

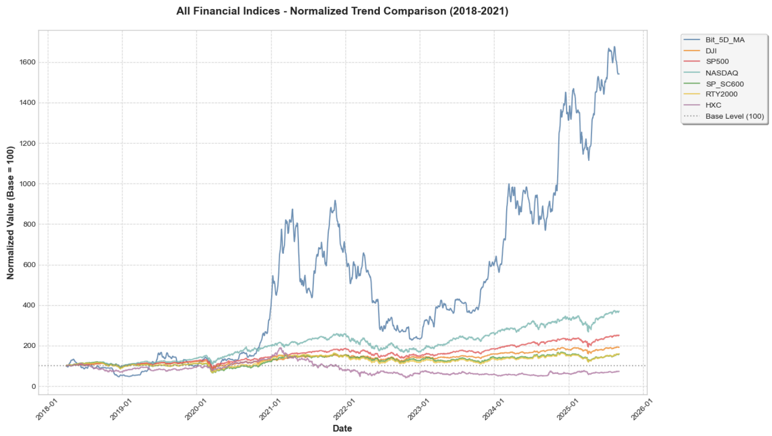

Figure 2. Normalized price trends (all assets)

When we put in our control variables, we see for every variable, its ups and downs remained pretty much the same, none had the super steep, near 18x vertical spike like Bitcoin did. NASDAQ performs best, which shows the market highly recognizes and has long-term commitment to technology stocks. This is also indirectly proof that Bitcoin could have such a huge surge within 2 years. It's also worth knowing that old-fashioned safe-have assets like gold, which usually go in reverse direction compared to high-risk assets such as tech or growth shares have also moved more aligned upwards from 2024 onwards. It might indicate there's a certain market bubble, or it could mean most of these asset values are mainly driven by liquidity. As for performance of small-cap stock indices though, they've not been delivering all those great results, which means US equity investors over the past few years still prefer bigger firms on the list rather than smaller companies. But it still cannot be denied that both large-cap and small-cap have the same extent of volatilities towards the macroeconomic shocks, the different patterns we saw are mostly due to the differences in index composition which we have to look out in the next analysis for controlling for macroeconomic factors. Prior to analyzing the experimental results, it can be anticipated that Bitcoin may exhibit stronger predictive power over large-cap stocks, as investors may engage in cross-market allocation strategies—such as realizing gains from equities and reallocating into Bitcoin to maintain liquidity and growth potential.

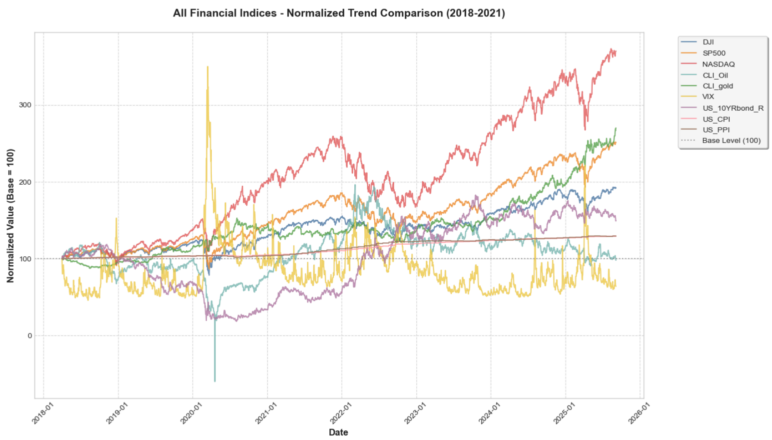

Figure 3. Normalize the trends of the control variables relative to the large-cap stock indexes

DOI: 10.54254/2754-1169/2025.BL30746

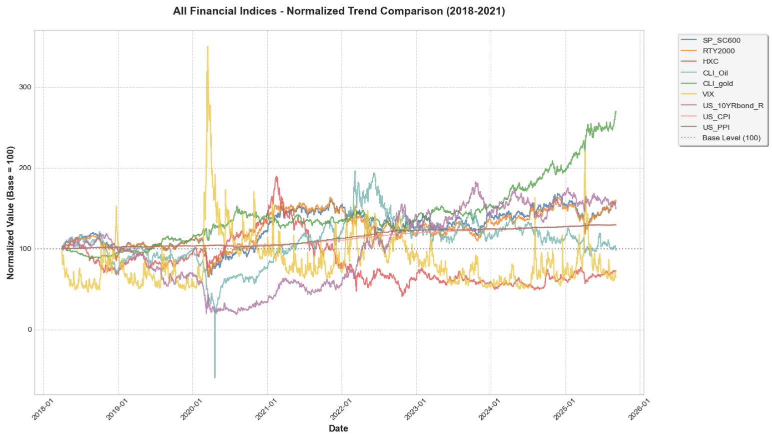

Figure 4. Normalize the trends of the control variables relative to the small-cap stock indexes

Both bit coins and stock indices have returns with a normal distribution, so it seems quite random Apart from the HXC series which is slightly right skewed, all other return series are slightly left skewed meaning that there will be an occurrence of extremely negative returns. This means that even if the overall performance of HXC is poor, it is still relatively stable, but other upward trending assets (especially BTC) carry a big risk of a large drop in the drawdown. Also Bitcoin is different from most traditional financial assets as we do not see significant fat-tailed distributions, and instead more resemble a Gaussian normal distribution

3.2. Model

3.2.1. VARX model

The Vector Autoregression with Exogenous Variables (VARX) model can be formally specified as:

$$Y_t = A + \sum_{i=1}^{p} \Phi_i Y_{t\text{-}i} + \sum_{j=1}^{s} B_j X_{t\text{-}j} + DC_t + \epsilon_t$$

Y_t : A vector of endogenous dependent variables (e.g., returns of the Russell 2000 Tech Index, S&P 500 etc.).

X_t : A vector of exogenous independent variables (Lagged Bitcoin returns).

C_t : A vector of control variables: Macroeconomic factors (e.g., changes in interest rates, CPI inflation)

A: A vector of constants.

Φ_i : Matrices of parameters to be estimated.

ϵ_t : A vector of white noise error terms.

IRFs are used to trace the dynamic effect of a 1 std deviation innovation in one variable (Bitcoin return in our case) on all other variables in the system now and in the future. To find the structural shocks, we use the Cholesky decomposition, so we have to pick an ordering for our variables under some assumptions about what is exogenous.

Macroeconomic variables behave as the major drivers of the market, in accordance with the research design and theoretical economics. So,the transmission mechanism comes from these

DOI: 10.54254/2754-1169/2025.BL30746

fundamental ones. Then it flows into the different Bitcoin market variables, which are returns, volatility, and trading volume, before the shock has an effect on the regional equity indices. Variable ordering accordingly is defined as:

Macroeconomic Variables → Bitcoin Metrics → Regional Equity Indices.

This will enable us to then order these volatilities in order to see how the volatility that was transmitted through the Bitcoin market on a macroeconomic level impacts the equity markets.

3.2.2. Predictive validation with machine learning (SVM)

The insights derived from the VARX and impulse response function (IRF) analyses will directly guide the construction of the feature set for the supervised learning model. The target variable is binary, showing if the equity index moves positively or flat, it takes a 1, whereas negative or zero takes 0. Predictive features will be built based on IRF result, with Bitcoin return, volatility and trading volume's lagged value. the number of lags (dd) will be decided by the persistence period seen from the IRF analysis, for instance, if the effect is on for five periods, the lags 1-5 would be part of it. In addition, the equity index's own return's lagged return will be included as well if there was any momentum or mean reversion effect:

It's interesting that there will be specific characteristics for every equity index. For indices with a high level of sensitivity toward Bitcoin changes in the IRF analysis (Russell 2000 Tech), more weight will be given to Bitcoin-based features. Conversely, as indices where bitcoin is not the most influential, the feature set will be more autoregressive terms.

In order to mitigate the problem of look-ahead bias, a rolling-window or expanding-window training scheme will be used to ensure that the prediction accuracy is out of sample. We will use the standard metrics of Accuracy, Precision, Recall, F1-score to evaluate the model. The F1-Score is important because of class imbalance, like in a long bull market. The SVM model's accuracy will also be compared to a simple strategy, such as always predicting an upward move, to check if it has any real-world use.

SVM models' comparative performance in different regional indexes will also be assessed, which is mainly conducted through F1-Score. If we have a better ability to forecast the previously identified indices which are sensitive to shocks of bitcoin, this will greatly support the results obtained through the econometric analysis initially. So far this is an integrated methodology to provide a strong assessment of the statistical as well as practical predictive ability of Bitcoin across the world equity markets.

3.2.3. Logistic regression model

In addition to the above two models, this paper introduces a baseline model, which is a logistic regression model, using whether Bitcoin will rise or fall in the short term to predict whether major U.S. equity indices will rise or fall. Logistic regression is selected as an interpretable model and one for predicting binary outcomes so we use it as a simple but effective baseline before introducing more nonlinear methods.

The model focuses solely on the U.S. market. The dependent variable Y_t represents whether a given stock index—DJI, S&P500, NASDAQ, S&P SmallCap600, Russell 2000, or NASDAQ Golden Dragon China (HXC)—increased relative to the previous day. Specifically, $Y_t = 1$ if the daily return is positive and $Y_t = 0$ otherwise. The main explanatory variable is Bitcoin's five-day moving average return ($Bit_5D\_MA\_Return$), While the rest of the macro and financial variables are control variables. Specifically, the oil return, the return on Gold, short term interest rate (SOFR),

DOI: 10.54254/2754-1169/2025.BL30746

volatility index (VIX), U.S. 10-year treasury, and CPI/PPI growth. I have these values aligned by keeping them on a daily basis by removing the missing value so the data is consistent.

The logistic regression model yields an estimate of the probability of a market index rising as:

$$P\left(Y_t = 1\right) = \frac{1}{1 + e^{-\left(\beta_0 + \beta_1 X_{BTC,t} + \beta_2 X_{controls,t}\right)}}$$

where $X_{BTC},_t$ denotes Bitcoin's five-day moving average return, and $X_{controls},_t$ represents the vector of control variables. Index are modeled separately using the same predictors. Each index is predicted by itself using the same predictors. dataset is split chronologically as 70% for a training sample and 30% for testing, parameters are estimated using the principle of maximum likelihood. Model performance will be presented via Accuracy, Precision, Recall, F1-Score and ROC – AUC in results part to show how well model performs in predicting the market direction.

3.2.4. Long short memory model

To supplement the VARX, SVM, and logistic regression modeling frameworks, an LSTM neural network is applied here for the first time to capture complex non-linear temporal effects of Bitcoin relative to major U.S. Equity indices. LSTM networks are particularly well-suited to financial time series prediction since they're able to learn long-term dependencies (Hochreiter & Schmidhuber, 1997; Gers et al., 2000), and therefore capture sequential patterns that traditional econometrics models might ignore.

$$\mathrm{Y}_{\mathrm{t+1}} = \mathrm{f\_LSTM}(\mathrm{x\_\{t-n+1:t\}};\ \theta)$$

For LSTM arch, we address the issues of vanishing gradient found in standard RNNs, allowing it to be effective in capturing the dynamic relationship between crypto and US market equities and equity indices respectively. The model uses the same focus on the U.S. equities market and binary classification framework as the baseline approaches. The dependent variable Y is the same six major US equity index directional movement, DJI, S&P 500,NASDAQ, S&P SmallCap600, Russell 2000, and NASDAQ Golden Dragon China (HXC)—where Y = 1 indicates a positive daily return and Y = 0 otherwise. The input features include Bitcoin's five-day moving average return (Bit_5D_MA_Return) as the primary predictor, along with the same comprehensive set of control variables: oil and gold returns, short-term interest rate (SOFR), volatility index (VIX), U.S. 10-year Treasury yield, and CPI/PPI growth. To leverage the LSTM's sequential learning capabilities, the model incorporates a lookback window of 10 trading days, allowing it to capture temporal patterns and momentum effects in the relationship between Bitcoin and equity markets.

The input sequence is processed by the LSTM model through several layers, including a dropout regularization (with a rate of 0.2) to avoid overfitting and enhance generalization. The network architecture is composed of two LSTM layers, with 64 and 32 units respectively, followed by a dense layer that uses ReLU activation. The final layer is a sigmoid output layer designed for binary classification tasks. The model is trained using the Adam optimizer, which employs a learning rate of 0.001, along with the binary cross-entropy loss function. Following the methodology akin to logistic regression, the dataset is divided chronologically into 70% for training and 30% for testing.

DOI: 10.54254/2754-1169/2025.BL30746

To mitigate overfitting, early stopping is applied, guided by the validation loss during the training process. The performance of the model will be assessed using a set of comprehensive metrics—namely accuracy, precision, recall, F1-score, and ROC-AUC—to facilitate a direct comparison with baseline approaches. This evaluation framework aims to quantify the added value of deep learning techniques in forecasting equity market directionality.

4. Result

4.1. VARX model result

4.1.1. Overview

Before the formal model estimation procedure, VIF test is carried out to check whether there might be any multicollinearity problem among the control variable. Using a threshold value of 10, the criterion was achieved for all variables. Afterwards, stationarity of all variables is checked with the ADF test. In terms of the ADF results, it is shown that the soft rate as well as the us 10yr treasury rate are non-stationary hence we are taking first differences of both series for further testing. After the AIC showed we can use 9 lags in our VARX model to get the best fit:

In terms of model estimation, due to computer restrictions, we adopt OLS to get the VARX model. To derive Impulse Response Functions (IRFs) and Orthogonalized Impulse Response Functions (OIRFs), the model was re-estimated as a standard VAR by excluding the control variables.

The model results for large capitalization equities are promising and earlier lags are more statistically significant. The opposite is true for the returns of small-cap stocks which are not as obvious and variable. That said, we could still conclude that Bitcoin return does have predictive strength, and that there is a stronger link between the returns of Bitcoin and returns on other traditional assets, although that does not mean that shocks to Bitcoin are immediately transmitted into the economy—the market will require a certain amount of time to absorb them. The same can be said from Granger causality test carried out on Bitcoin returns, except for the first lag, the subsequent lags have high significance, meaning Bitcoin returns are good predictors for stock index movements.

4.1.2. The coefficient from regression

Different stock index's reaction to Different Bitcoin's lag term return has a significant difference As for the Dow Jones Industrial Average, the fourth and fifth lag terms of Bitcoin return shows to have considerable influence on the index return, with the fifth term displaying especially strong significance. This shows the relatively slower response of Dow-Jones to shocks caused by the bitcoin, this might be due to the fact that Dow consists of only large-cap blue-chip stocks making it more stable and slower to react to external impulses.

For the S&P 500 Index, there are large effects from both the first and fifth lag of Bitcoin returns, which could be interpreted as segmented shock transmission. For lag1, the coefficient – -0.0693, and for lag5 the coefficient – -0.0736 are both very negative, and also very close in value – meaning that a positive return on Bitcoin "pulls" the S&P 500 down. The same negative impact is seen on the NASDAQ Composite but just the first lag of Bitcoin returns is statistically significant. NASDAQ is made up mainly of technology and growth companies, and that may be an explanation for its heightened sensitivity to the ups and downs of developing assets such as Bitcoin. also, the

DOI: 10.54254/2754-1169/2025.BL30746

index being market- cap – weighted, Tesla and Nvidia these heavyweight techs, play an important part and these companies are usually related to these emerging thematic assets such as Bitcoin.

On the contrary though the returns of the S & P Small Cap 600 Index do not show much reaction to any of the laggered Bitcoin Return Shocks. Because of the concentration of ownership among small-cap firms, these companies may be more likely to avoid being affected by the external market. However, this does not disagree with the Granger Causality tests, which show that Bitcoin returns Granger cause the index.

Similar to NASDAQ, Russell 2000 Index has immediate response to Bitcoin return shocks as well; only the first lag demonstrates statistical significance. Positive Returns of 1 day of Bitcoin has a slight negative effect on the Russell 2000's Returns.

The Chinese gold dragon index (HXC) stands out as a special case because these are enterprises that derive income from China yet are traded on the US market. Here's the 7th lag of the Bitcoin return showing a positive and significant effect on HXC. This is a quite interesting phenomenon, a positive impact on the Bitcoin returns, it sends a positive impulse to the Chinese enterprise index one week later. An explanation might be that Chinese equities in the US market as well as Bitcoin have similar roles in the investor's portfolio for them to diversify their assets and allocate the surplus liquidity.

Besides the effect on Bitcoin returns, there are some relationships between stock index returns themselves. But as they're not central aspects of this study, I won't delve into them now.

4.1.3. IRF and OIRF

The IRF plots show a lot of difference in the path of response for the different stock indices after a shake-up in Bitcoin returns.

Dow Jones Industrial Average: The initial IRF is slightly negative but then turns positive before becoming strongly negative. The OIRF suggests that although there was an initial shock that was definitely negative, the impact that that had on the Dow was mainly positive and that it would still be quite strong later on.

Regarding the S&P 500 Index case, the IRF presents a negative response until lag 6, then there is a strong reverse towards a quickly increasing positive impulse. This response trajectory is confirmed by OIRF.

As for the influence of Bitcoin on the NASDAQ Composite, there's a clear path of events that takes place, first comes a positive reaction, after that things turn negative, and then that influence fades towards zero over time. Both IRF and OIRF plots show this.

Regarding to the S&P SmallCap 600 index: IRF shows constantly negative and finally decay to zero. But the OIRF shows a more complicated path - first goes up, then goes down, ends up at zero.

Looking at both the IRF and OIRF plot for the Russell 2000 Index, the shock from Bitcoin returns are all positive and die out over time, eventually vanishing after around the 9th trading session.

The same result also can be found with China Golden Dragon index(HXC), where the IRF and OIRF converge as well: Here, the negative effect continues to increase slowly, reaches its turning point around the 7th period and then decays back toward zero.

DOI: 10.54254/2754-1169/2025.BL30746

4.2. SVM model result

4.2.1. Overview

This section reports the predictive performance for the support vector machine model fitted to the U.S. market data with S&P 500 daily return as the target variable The model selects RBF kernel so as to capture nonlinear correlation between Bitcoin dynamics and equity market movement Bitcoin features include: 5 day moving average returns lagged volatility and trading volumes plus the macro variables, interest rate, inflation, VIX, GOLDE oil returns. Implementation of a rolling window training scheme to maintain out of sample validity and prevent look ahead bias was made.

4.2.2. Classification Performance

The SVM classifier has a general accuracy of about 56.65%, which is close to the 50% benchmark for random guessing of the 5-class classifier. Table 1 Summary of main performance metrics

Table 1. SVM performance results

Class	Precision	Recall	F1-score	Support
0 (Down)	0.52	0.16	0.24	372
1 (Up/Stable)	0.57	0.89	0.70	477
Overall	0.55 (macro avg)	0.52	0.47	849

The results indicate that the SVM model is more sensitive to upward or stable market movements, with a high recall of 0.89 for class 1, correctly identifying most bullish days. In contrast, its ability to detect downward market movements is relatively weak (recall = 0.16), reflecting class imbalance and the general upward drift in financial time series.

4.2.3. Confusion matrix analysis

An inspection of the confusion matrix further reveals the classifier's behavior. The model correctly classified only 58 of the 372 downward-movement samples, mislabeling the remaining 314 as upward movements. Conversely, it performed much better on the 477 upward-movement samples, correctly identifying 423 of them. This clear asymmetry indicates that the SVM is biased toward predicting upward movement, likely because the underlying dataset contains predominantly up-days or days with minimal change.

This type of directional bias frequently occurs in financial time-series classification, underscoring the need for methods that explicitly address class imbalance, such as weighted loss functions or re-sampling strategies.

4.2.4. Interpretation and implications

Moderate accuracy and biased recall indicate that Bitcoin has some weak predictive signals but non-zero about U.S Equity index movements. Its info seems more bullish or stable-market-leaning rather than bearish-market-oriented. It matches past papers that showed Bitcoin was a leading indicator during risk-on periods where markets were doing well.

From a practical standpoint, while the SVM's performance may not be good enough to make precise daily trading decisions. High recall on upward moves could be used as a directional filter or

DOI: 10.54254/2754-1169/2025.BL30746

in an ensemble strategy [16]. Such as combining SVMs with other macros/sentiment indicators to increase robustness and decrease false positives.

4.2.5. Limitations and next steps

The performance gap between upward and downward classifications suggests several directions for enhancement:

• Incorporating cost-sensitive training to penalize false upward predictions [15].

• Testing alternative kernels (e.g., polynomial, sigmoid) and feature selection strategies [18].

• Applying ensemble approaches that combine SVM with logistic regression or deep learning architectures to capture complementary signals [17].

These refinements may improve classification balance and overall out-of-sample predictability.

4.3. Logistic regression results

4.3.1. Overview

This section presents the empirical results of the logistic regression model used to examine whether Bitcoin's short-term price dynamics can predict daily directional movements in major U.S. equity indices. The dependent variables represent the binary outcomes of index returns ('up' = 1, 'down' = 0), while the explanatory variables include Bitcoin's five-day moving-average return and a series of macro-financial controls (oil and gold returns, SOFR, VIX, 10-year Treasury yield, CPI and PPI growth). All models were trained and tested using balanced time-series samples, and their out-of-sample classification performance was evaluated via standard metrics such as accuracy, precision, recall, and F1-score.

4.3.2. Classification performance

Table 2 summarizes the predictive performance of the logistic regression across six major indices: the Dow Jones Industrial Average (DJI), S&P 500, NASDAQ Composite, S&P SmallCap 600, Russell 2000, and NASDAQ Golden Dragon China Index (HXC). Overall, model accuracies range between 53% and 57%, indicating slightly better-than-random predictive power. Indices with broader sectoral coverage such as the DJI and S&P 500 show marginally higher accuracies (≈56%), whereas small-cap and China-exposure indices yield more variable outcomes.

Table 2. Logistic regression performance by indexes

Index	Accuracy	Precision (Up)	Recall (Up)	F1 Score
DJI_UP	0.57	0.55	0.91	0.69
SP500_UP	0.55	0.54	0.88	0.67
NASDAQ_UP	0.53	0.53	0.95	0.68
SP_SC600_UP	0.54	0.51	0.87	0.65
RTY2000_UP	0.54	0.52	0.91	0.66
HXC_UP	0.53	0.49	0.48	0.49

From here on, it is clear that there exists a recall-precision trade-off. Recall across all indices for up is greater than 0.85, meaning the model catches the vast majority of upward movement days. but the precision values(≈0.5 - 0.55) suggests that quite some of the predicted 'ups' were in fact false positives. This means that if the dynamics of Bitcoin show a change in the market, we have the signal but not the way to predict the day by day the movement of the market.

4.3.3. Confusion-matrix analysis

Appendix A1 have the detailed confusion matrices for each index Take the DJI_UP model as an example, where true positives (389) far outnumber true negatives (74), indicating that the predicted results tend to be upwards movement, which is also the nature of all financial time series with an increase in the drift. Similar pattern is also found in the S&P 500 and NASDAQ model, HXC is the most unbalanced with upward and downward classifications. This constancy amongst the indices further affirms that it appears to be the momentum within the price of Bitcoin itself which has given the faint but real prediction signal.

4.3.4. ROC-curve evaluation

Furthermore, the ROC of these findings can be seen in Appendix A2. Area under the curve (AUC) mostly in the 0.55-0.62 range means they have somewhat predictive power but not much more than guessing randomly. As for the market capitalizations (market caps) of stocks included in the indices, as the market cap gets larger (like in DJI, S&P 500), the index tends to give more distinction, which implies that Bitcoin may affect the larger-cap U.S. stocks more than the smaller-cap ones or the specific regions.

4.3.5. Discussion

The logistic regression results indicate that in Bitcoin's 5day moving-average return, there is some statistically detectable information about U.S. equity-market directions, albeit small. Higher recalls indicate to bullish momentum sensitive and lower precision shows to volatility spillovers and noise in crypto-equity link. These findings are in line with prior work (Bouri et al., 2023, Dudek, 2021) that documents limited but nonzero cross-market predictability. Future extension- Non-Linear Machine Learning models (RFC, SVM) can perform better in capturing Asymmetric dependencies and in overall classification accuracy.

4.4. Long short memory model results

4.4.1. Overview

This section provides the actual results that came out from the research conducted using the LSTM neural network in terms of predicting if small term price fluctuations of Bitcoin would be able to predict the direction of a daily price change for several large U.S. indices. The LSTM model is a kind of sequential learning model, which can learn the time-dependent factors and the nonlinear patterns that are missing from the traditional econometric models. Dependent variables are the binary index returns: 'up' = 1 and 'down' = 0, explanatory variables include Bitcoin's 5-day MA return and the same comprehensive set of macro-financial controls, namely oil and gold returns, SOFR, VIX, 10-year Treasury yield, CPI and PPI growth. The model looks back 10 days to pick up on momentum effects and time patterns in Bitcoin equity. All models were trained with chronologically balanced time series samples and were regularized using early stopping, the out of

DOI: 10.54254/2754-1169/2025.BL30746

sample classification accuracy for all the model was measured and compared using the standard metrics like accuracy, precision, recall and F1-Score.

4.4.2. Classification performance

Table 3 presents the prediction performance of the LSTM model on several major stock indices from four countries. The model has strong predictive power, with accuracy being from 48%–74% and AUC between 0.48–0.79, beating support vector machine and logistic regression.

Among all markets, the one with the highest predictive accuracy (74%) and AUC (0.79), which is the HXC index market, shows that the LSTM model can capture more consistent temporal dependencies in this market. On the contrary is SP500 (SP500) which yielded the weakest prediction results (accuracy 48%, AUC 0.48), it may be inferred that its prices are harder to model with the help of features drawn from Bitcoin.

In general, it reached a accuracy around 60%,the F1-score is arround 0.68, also maintained precision higher than 0.7 for most indices This shows that the LSTM is able to learn the meaningful temporal features from the Bitcoin price movements, therefore cross market is possible.

Table 3. LSTM model performance by indexes

Index	Accuracy	AUC	Precision (Up)	F1 Score
DJI	0.54	0.54	0.57	0.57
SP500	0.48	0.48	0.46	0.67
NASDAQ	0.57	0.63	0.73	0.72
SP_SC600	0.0.57	0.51	0.59	0.59
RTY2000	0.60	0.51	0.75	0.75
HXC	0.74	0.79	0.80	0.80

4.4.3. ROC and feature importance

The ROC curves show that it has meaningful discrimination power w.r.t. U.S stock indices rather than random classifiers: Looking at feature importance, MACD is the biggest contributor with 0.235, followed by Trend with 0.193 and Volume with 0.175, while volatility has 0.090 and RSI is 0.112. This ranking tells us that Bitcoin’s speed and how much it moves in price give the most helpful hints for figuring out how stock indexes change.

4.4.4. Trading and performance evaluation

If we use the LSTM alone, as a simple long-short trading strategy, we get a cumulative return of 258.5%, CAGR of 47.4%, Sharpe ratio of 2.33, and a maximum drawdown of 17.8%. Yearly breakdowns show high consistency, with positive annual returns in all years tested (2021- 2025). The equity curve and drawdown plot show a steady compounding with not much downside volatility. In contrast to traditional models, LSTM can better adapt to changes in the market and thereby allow for better signal extraction of the high-frequency Bitcoin price movements.

DOI: 10.54254/2754-1169/2025.BL30746

4.4.5. Discussion

Overall, on the LSTM side of things, it looks like Bitcoin's short-term movement does have some measurable predictive information about US stock indices, although more so in periods of heightened volatility or macro uncertainty. The model is non-linear, which can pick up on the asymmetries that the linear regression can't. And although its predictive strength (AUC ≈ 0.66) is still moderate, there appears to be a positive trading record, so such a model may still have practical uses for a portfolio trader or market hedger. These findings correspond to current literature (for example [2-3]) highlighting that crypto - equity interaction is an episodic phenomenon but it is increasingly significant with growing integration of digital assets in global finance.

5. Suggestion and future prospect

5.1. Suggestion to investors and financial institutions

And it repeats the results from one that we've talked about for ages: Bitcoin, is top of the pile cryptocurrency, it has become essential in the global financial system. As with this paper specifically highlighting the predictive ability of returns on Bitcoin with respect to some major U.S. stock markets, it is worth noting that using any sort of cryptocurrency information for a quantitative trading indicator would not be advisable.

Although our LSTM model reached a prediction accuracy of 65.9% for the US market and could potentially be improved, such as by excluding minor fluctuations in the model training process, the model's accurate prediction cannot be said to be a correct strategy in the long run. The underlying cause is the Bitcoin being too unstable to invest in. The feeling among investors about the Bitcoin has changed.

Firstly, people only viewed bit coin as a novel digital token. Nobody paid attention to it and no one treated it as a real asset. Decentralized ecosystems together with the metaverse, its price has shot up drastically, and is increasingly seen as a speculative or risky asset. As of right now Bitcoin is often viewed as being in competition with gold for the designation of safe haven asset and is absorbing a lot of extra liquidity in US dollars.

Such evolutionary nature makes it hard to define a consistent Bitcoin based-Indicator's definition or behavior. but it is still worth continuing to pay attention to this new type of assets. It is already visible in its effect on traditional financial markets and is only going to increase in the future.

5.2. Suggestion to regulator and policy makers

Looking ahead, regulation of markets will probably become more heterogeneous with a huge part of that due to Bitcoin and other cryptocurrencies. In this paper, the returns of Bitcoin can be demonstrated to have created several interconnections with major stock indices return. such sort of statistical relationships also shows how investors' minds have been shifting.

The Fed traditionally relies mainly on three key instruments. But Bitcoin can seriously damage the force of money policy. When the fed raises the fed funds rate, a pertinent question would be can the freed liquidity be quickly absorbed by the Bitcoin market? Bitcoin is growing, it is almost unregulated, and liquid. So as a de centralized asset, it may be enough to give a rival of the Fed's role, that of being the sole arbitrator of dollar liquidity. In this system, we have an unmanaged levee causing some unexpected effects.

DOI: 10.54254/2754-1169/2025.BL30746

5.3. Suggestion to researchers and future prospect

Models in this study continue to be rather simple. New modeling frameworks could be introduced by other scholars in future studies or the present LSTM based model could be improved upon. Promising directions include: rolling window validation, more complex models, or better selection of data features to include. And it can also be argued that the minor oscillations in the stock index returns will probably be swallowed up by the fluctuations in Bitcoin, thus reducing the influence of the model. A somewhat more accurate predictive model, including a Bitcoin return term, has potential practical applications, too: it may be helpful even if not for direct trading purposes (to institutions, etc., in their own risk management). Looking even further ahead, as a distinct emerging asset class, the implications is that there now are some deeper questions about Bitcoin. It's actually playing what part? What kind of logics are valued on? Bitcoin, if it's highly tied to conventional assets, could its indicators act like sentiment indicators? Similar to implied vol:

6. Conclusion

The study looks closely at how much more predicting power Bitcoin has over region's equities index using an overly strict many part way to do so. Moving away from the early literature that presented Bitcoin as an isolated asset or just a diversifier, the paper frames Bitcoin as a liquid, institutional asset whose price dynamics are an economically significant, and actionable leading indicator of short-run equity market movements given certain states.

The main contributions and results are the following:

First, reveals heterogeneity & multi-horizon cross-market transmission: By utilizing the VARX and IRF models, this paper is able to precisely map out the pathways by which Bitcoin return shocks flow to equity markets. The results show significant regional and index differences; tech-heavy indices like the NASDAQ and emerging market proxies (e.g., the China Golden Dragon Index) are more and quicker to react to Bitcoin shocks, traditional blue-chip indices (e.g., the Dow Jones) have a slow response. And it's got differences across markets with regard to how long lasting these effects are, so Bitcoin's info is cleared up multi-horizon.

Second, this paper connects identification and practical usability: Step outside of standard econometric associations alone, and turn those dynamic relations found in VARX/IRF into crafted features used with supervised ML models. Out-of sample predicitive validation of SVM and Logistic Regression confirm limited but practical usefulness in predicting future equity index direction by Bitcoin signals. Additionally, an LSTM neural network performed best (such as 65.9% accuracy for the U.S. market), showing that deep learning is better at capturing the complex and nonlinear time dependence between the cryptocurrency and equity markets.

Third, it specifies the boundaries and conditions of predictive value: It refutes the claim that bitcoin is a leading universal index. Its predictive power, instead, depends greatly on the market type (developed vs. emerging), the index composition (tech-heavy vs. broad), and the market regime. Model's performance showed considerable assymetry, it was considerably superior when it came to spotting upward vs. downward shifts in the market. This serves as a caution to investors that while Bitcoin does contain forward-looking information, there is enough signal for high frequency tactical trading, but probably some value to be gained for informing strategic risk budgets and tactical allocation tilts.

In conclusion, this paper shows that there does exist the existence of incremental predictive information for regional equity indices within Bitcoin, however it is conditional, heterogeneous and asymmetric. The research creates some kind of structure and something testable - starting from recognition of structure, through out-of sample testing - and makes “crypto leading equities” stand on ground with some testable evidence, and it clearly shows where and when and how much the Bitcoin helps to improve forecasts - and also when it can be ignored.

Acknowledgement

Yu Gao, Bowen Liu, Yishan Yang and Junxiong Huang contributed equally to this work and should be considered co-first authors.

Keywords:Asset Pricing, Cybercurrency, Stock Indexes’ Pricing, Emerging Market, Market Correlation

References

[1] Baur, D.G., Dimpfl, T. & Kuck, K. (2018) 'Bitcoin, gold and the US dollar – A replication and extension’, Finance Research Letters, 25, pp. 103–110. https: //doi.org/10.1016/j.frl.2017.10.012
[2] Corbet, S., Meegan, A., Larkin, C., Lucey, B. & Yarovaya, L. (2019) 'Exploring the dynamic relationships between cryptocurrencies and other financial assets’, Economics Letters, 165, pp. 28–34. https: //doi.org/10.1016/j.econlet.2018.12.004
[3] Bouri, E., Salisu, A.A. & Gupta, R. (2023) 'The predictive power of Bitcoin prices for the realized volatility of U.S. stock sector returns’, Financial Innovation, 9(64). https: //doi.org/10.1186/s40854-023-00464-8
[4] López-Cabarcos, M.Á., Pérez-Pico, A.M., Piñeiro-Chousa, J. & Šević, A. (2021) 'Bitcoin volatility, stock market and investor sentiment’, Finance Research Letters, 38, p. 101428. https: //doi.org/10.1016/j.frl.2020.101428
[5] Eross, A., McGroarty, F., Urquhart, A. & Wolfe, S. (2019) 'The intraday dynamics of bitcoin’, Research in International Business and Finance, 49, pp. 71–81. https: //doi.org/10.1016/j.ribaf.2019.02.008
[6] Catania, L. & Sandholdt, M. (2019) 'Forecasting Bitcoin volatility: GARCH models vs. machine learning’, Finance Research Letters, 32, p. 101368. https: //doi.org/10.1016/j.frl.2018.12.020
[7] Dudek, G. (2021) 'Forecasting cryptocurrency volatility with machine learning’, Financial Innovation, 7(58). https: //doi.org/10.1186/s40854-021-00272-2
[8] Kristjanpoller, W. & Minutolo, M.C. (2018) 'A hybrid volatility forecasting framework integrating GARCH, artificial neural networks, and random forests’, Expert Systems with Applications, 104, pp. 38–48. https: //doi.org/10.1016/j.eswa.2018.03.009
[9] Chen, Y., Zhu, S. & Wang, G.J. (2021) 'Forecasting Bitcoin volatility using ensemble learning’, Annals of Operations Research, 299(1–2), pp. 3–21. https: //doi.org/10.1007/s10479-019-03475-7
[10] Mallqui, D.C.A. & Fernandes, R.A. (2019) 'Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques’, Applied Soft Computing, 75, pp. 596–606. https: //doi.org/10.1016/j.asoc.2018.11.038
[11] Lahmiri, S. & Bekiros, S. (2019) 'Cryptocurrency forecasting with deep learning chaotic neural networks’, Chaos, Solitons & Fractals, 118, pp. 35–40. https: //doi.org/10.1016/j.chaos.2018.11.014
[12] Mthembu, N., Sanusi, K. A., & Eita, J. H. (2022). Do stock market volatility and cybercrime affect cryptocurrency returns? Evidence from South African economy. Journal of Risk and Financial Management, *15*(12), 589. https: //doi.org/10.3390/jrfm15120589
[13] Salisu, A. A., Isah, K., & Akanni, L. O. (2019). Improving the predictability of stock returns with Bitcoin prices. North American Journal of Economics and Finance, *48*, 857–867. https: //doi.org/10.1016/j.najef.2018.08.010
[14] Lahiani, A., Jeribi, A., & Jlassi, N. B. (2021). Nonlinear tail dependence in cryptocurrency-stock market returns: The role of Bitcoin futures. Research in International Business and Finance, 56, 101351. https: //doi.org/10.1016/j.ribaf.2020.101351
[15] Hitam, Nor Azizah, Amelia Ritahani Ismail, and Faisal Saeed. "An optimized support vector machine (SVM) based on particle swarm optimization (PSO) for cryptocurrency forecasting." Procedia computer science 163 (2019): 427-433.
[16] Karasu, Seçkin, et al. "Prediction of Bitcoin prices with machine learning methods using time series data." 2018 26th signal processing and communications applications conference (SIU). IEEE, 2018.
[17] Khedr, Ahmed M., et al. "Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey." Intelligent Systems in Accounting, Finance and Management 28.1 (2021): 3-34.

DOI: 10.54254/2754-1169/2025.BL30746

[18] Ferdiansyah, Ferdiansyah, et al. "A lstm-method for bitcoin price prediction: A case study yahoo finance stock market." 2019 international conference on electrical engineering and computer science (ICECOS). IEEE, 2019.

Appendix

Appendix 1

Table A1. Description of variables

Variable name	Description
Bit_5D_MA	Five day moving average price of Bitcoin
DJI	Dow Jones Industrial Average index (DJIA)
SP500	Standard & Poor's 500 Index (S&P 500)
NASDAQ	NASDAQ Composite Index (NASDAQ)
SP_SC600	Standard & Poor's small cap 600 Index
RTY2000	Russell 2000 Index
HXC	NASDAQ Golden Dragon China Index
US_CPI	Consumer price index of the U.S.
US_PPI	Producer price index of the U.S.
sofr	Secured Overnight Financing Rate
vix	CBOE Volatility Index
US_10YRbond_R	Ten year treasury bond yield of the U.S.
CLI_gold	Gold Price on the Chicago Mercantile Exchange
CLI_Oil	Crude Oil Price on the Chicago Mercantile Exchange
XXX_Return	Simple daily return of XXX
XXX_growth	Monthly growth rate of XXX

Appendix 2: distribution of returns and control variables

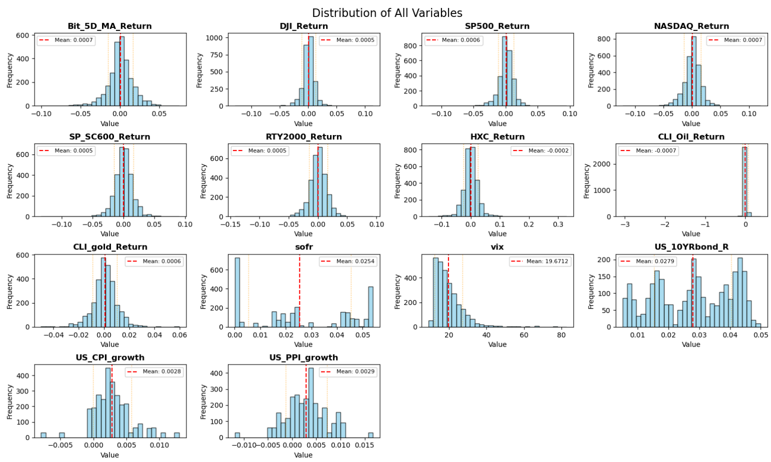

Figure A1. Distribution of all variables

DOI: 10.54254/2754-1169/2025.BL30746

Appendix 3: correlation matrix

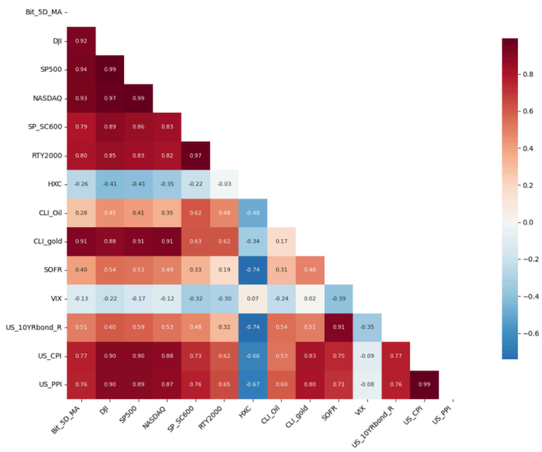

Figure A2. Correlation matrix for prices

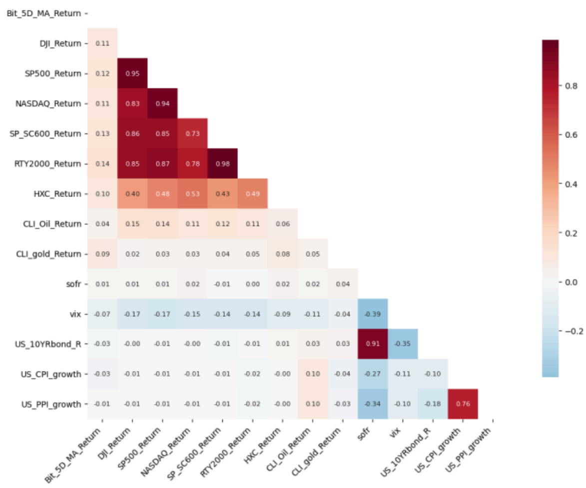

Figure A3. Correlation matrix for returns