Quantitative Finance and Economics

QFE, 9(2): 479–505.
DOI: 10.3934/QFE.2025016
Received: 15 January 2025
Revised: 17 April 2025
Accepted: 20 May 2025
Published: 13 June 2025

https://www.aimspress.com/journal/QFE

Research article

Valuation of crypto assets on blockchain with deep learning approach

Xi Zhou[1], Yin Pang[1], Esther Ying Yang[2], Jing Rong Goh[3] and Shaun Shuxun Wang[1,*]

[1] School of Business, Southern University of Science and Technology, Shenzhen, China
[2] Risk Lighthouse LLC, New York, USA
[3] School of Economics, Singapore Management University, Singapore

* **Correspondence**: Email: wangsx@sustech.edu.cn.

Abstract: With the rapid expansion of the blockchain ecosystem, crypto asset valuation has become an essential area of study for investors and institutions. Here, we introduce a deep learning framework that was designed to predict the value index of crypto assets by integrating intrinsic value variables and decomposing market prices into value and sentiment components. The Crypto Asset Value-indexing Model (CAVM) was applied to Ethereum's cryptocurrency ETH to demonstrate its effectiveness. Four econometric tests were conducted to verify the informativeness, predictiveness, and reasonability of the generated value indices, as well as the efficiency of price decomposition. Our findings suggested that the value index can serve as a reliable proxy for the intrinsic value of crypto assets, offering a benchmark for investment decisions, consumption, financial reporting, and potential tax implications. Additionally, this research contributes to the literature on asset valuation by proposing a novel method that applies deep learning techniques to intangible assets traded in secondary markets. By utilizing the end-to-end nature and directed acyclic graph structure inherent to deep learning models, we enhance the modeling process with customized loss functions and regularization mechanisms.

Keywords: asset valuation; block chain; crypto asset; deep learning; market sentiment

JEL Codes: G17, G12

Abbreviations: CAVM: (Crypto Asset Value-indexing Model) is a deep learning framework designed to predict the value index of crypto assets by integrating intrinsic value variables and decomposing market prices into value and sentiment components. CAVM aims to provide a reliable proxy for the intrinsic value of crypto assets.

1. Introduction

Since the birth of Bitcoin, the market for crypto assets has been booming. Originally regarded as only a type of speculative investment, crypto assets have experienced a paradigm shift, formed an ecosystem, and gained increasing focus and investment. The number of unique DeFi wallets reached a peak of 7.5 million in late 2021, but this figure has since declined. By 2023, the total value locked (TVL) in DeFi was approximately $50 billion, down significantly from its 2021 high of $175 billion. This decline reflects broader market challenges, including regulatory pressures and competition from traditional financial instruments like U.S. Treasury bonds. The global NFT market size was valued at $26.9 billion in 2023 and is projected to grow at a compound annual growth rate (CAGR) of 34.5%, reaching $211.7 billion by 2030. In 2024, the market grew to $28.6 billion and is forecasted to reach $43.22 billion in 2025, driven by increased adoption in gaming, art, and asset tokenization. Furthermore, the Web 3.0 blockchain market was valued at $3.59 billion in 2023 and is expected to grow at a CAGR of 45.47%, reaching $104.04 billion by 2032. This growth is fueled by rising adoption of decentralized applications, enhanced security features, and user control over data. With the rapid growth of the number of participants, individuals and institutions require crypto asset valuation for the purpose of investment consideration, consumption, financial reporting, and taxation (Johnson et al., 2019).

Given that interests in crypto assets started to grow in recent years, the most accepted theoretical valuation model of crypto assets is based on Metcalfe's Law. Many researchers explored the relationship between transaction price and various variables of crypto assets (Barth et al., 2020; Liebi, 2022; Liu & Tsyvinski, 2020; Naeem et al., 2021; Zhang & Wang, 2020). These studies did not provide a pragmatic valuation approach on crypto assets that can be readily applied to real world transactional data, which is our aim of this paper.

Crypto asset's value is primarily determined by its intrinsic value variables. However, the intrinsic value of crypto asset is intangible and may be challenging to quantify precisely. Thus, we can generate a value index that integrates the information of key intrinsic value variables to be a proxy of the intrinsic value.

Numerous crypto assets are traded via the secondary market. The transaction price of a crypto asset is significantly affected by both the intrinsic value and the market sentiment. Hence, we designed a three-stage time-series deep learning model to decompose the transactional price into the value and sentiment components, in which the value variables are integrated to generate the value index, and the sentiment variables are integrated to generate a sentiment index. Once the model is trained for a crypto asset, it can be used to predict the future daily value index of the crypto asset. We name this model the Crypto Asset Value-indexing Model (CAVM).

To ensure real-world applicability of the CAVM, our paper applies it on a crypto asset: The Cryptocurrency ETH. Ethereum is one of the most used blockchain ecosystems globally and ETH is its cryptocurrency.

Furthermore, we seek to verify the informativeness, predictiveness, and reasonability of the value index via explainable econometric methods. Informativeness represents the extent to which the information provided by intrinsic value variables is captured by the value index, whereas predictiveness is used to judge if the predicted value index indeed contributes to the prediction of the market price of crypto assets, and reasonability postulates if the predicted value indices for ETH are consistent with the intuitions underlying crypto assets. In addition, we seek to test the goodness of decomposition of market price into the value index and the sentiment index.

Our paper makes three contributions to literature. First, we developed a value index to be the proxy of the intrinsic value of crypto assets. This may serve as a guide for the investment and consumption on crypto assets, and a proxy for a crypto asset's book value, which may contribute to a more appropriate accounting methodology of crypto assets on financial statements. Second, we propose a novel method in the field of asset valuation where we decompose the market price into the value and sentiment components and integrate the value and sentiment variables to the value and sentiment indices, respectively. This is especially useful in valuing intangible assets that are traded on secondary markets like crypto assets. Third, we contribute to the literature of deep learning where we provide a new application of deep learning models in asset valuation and pricing. We utilize key characteristics of deep learning models, including the end-to-end nature, the directed acyclic graph structure, and the modeling techniques like customized loss function and regularization mechanism.

In section 2, we review the related literature. In section 3, we demonstrate the framework of the CAVM and define and categorize the variables. In section 4, we develop four hypotheses to test informativeness, predictiveness, and reasonability of the value index as well as the decomposition efficiency of the CAVM. In Section 5, we describe the data used to train the model, demonstrate the process of training, validating, and testing sample sets, and conduct the hypothesis tests. In section 6, the constructed value indices and predicted prices for ETH are plotted, and the hypothesis testing results is discussed and evaluated. In Section 7, we discuss regulatory implications and practical adoption challenges. In Section 8, we provide future research and a conclusion.

2. Literature review

The most related research to our paper provides a theoretical framework for valuation of cryptocurrencies based on the Metcalfe's Law[1] (García-Monleón et al., 2021; Shapiro & Varian, 1999). However, some factors (e.g., value of the nodes) of this model are challenging to calibrate and quantify and thus restrict the applicability of the model in valuing real world crypto assets. Moreover, it has been demonstrated that the traditional valuation approaches, such as the market approach, income approach, cost approach and QTM, are not widely applicable in valuing crypto assets due to the decentralization nature of crypto assets (Johnson et al., 2019). Graham & Dood (2008) proposed that the stock prices may deviate from the intrinsic value due to irrational behaviors of investors, but this deviation will eventually be corrected over time and stock prices will revert to the intrinsic value. This indicates the common premise that the transactional price is mean-reverting back to its value and provides credence for our CAVM which works by decomposing the crypto-asset transactional price into the value and sentiment components.

Since the asset value and market price are time series data, it is essential to utilize proper time series prediction techniques, such as the traditional econometric timeseries models like ARCH, GARCH, and ARIMA (Yildirim & Bekun, 2023; Dama & Sinoquet, 2021).

As the deep learning approach can provide more freedom of variable inclusion and better nonlinear fitting ability, we decide to employ the deep learning approach in our CAVM. More comparisons between our approach and the traditional approach are tabulated in Section 3.

[1] The Metcalfe's Law states that the value of a network is proportional to the square of the number of nodes in the network. For networks like the blockchain network and the internet, the nodes are usually the users of the network.

2.1. Recent advancements in AI-driven asset pricing

There have been recent papers on applying artificial intelligence models to asset pricing. Most of these analyses have entailed returns of traditional assets in the US. For instance, Chen et al. (2024) used deep neural networks to estimate an asset pricing for individual stock returns. Gu et al. (2020) compared machine learning methods for predicting U.S. stock returns and demonstrated the benefits of flexible methods. Following a similar approach, Bianchi et al. (2021) compared machine learning methods for predicting bond returns. Avramov et al. (2023) highlighted that performance of machine learning portfolios may deteriorate when trading frictions are present.

Kelly et al. (2025) introduced transformer-based architectures into stochastic discount factors, enabling conditional pricing through cross-asset information sharing. Their Artificial Intelligence Pricing Model (AIPM) reduced pricing errors and improved Sharpe ratios compared to traditional machine learning models like gradient-boosted trees, demonstrating transformers' capacity to capture intricate market dynamics. This aligns with Chen et al. (2024), who showed deep neural networks outperform linear factor models in capturing latent risk premium.

There is an increasing amount of literature that is applying machine learning to analyzing returns on crypto assets. Akyildirim et al. (2021) found that support vector machines outperformed artificial neural networks in predictive accuracy for twelve liquid cryptocurrencies. Khan et al. (2023) applied various machine learning models to predict volatility of four cryptocurrencies and found that no single model performed uniformly across the four cryptocurrencies. Liu et al. (2023) forecasted crypto returns for 3703 cryptocurrencies and found that returns for small cryptocurrencies are more predictable than larger cryptocurrencies. In summary, the more advanced techniques such as the neural networks of Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and Transformer in deep learning field (Verma et al., 2024; Liu & Zhang, 2023; Ang et al., 2020) are well documented in crypto assets academic research literatures.

Another critical advancement lies in hybrid models that integrate structured financial data with unstructured alternatives. Kim et al. (2021) developed a machine learning framework integrating blockchain transaction data and network activity to forecast Ethereum's price, demonstrating a 15% improvement over traditional models. Liu et al. (2023) demonstrated that transformer-CNN hybrids achieve superior Sharpe ratios in crypto markets by jointly processing temporal sequences and on-chain network graphs. Akyildirim et al. (2021) showed SVM-GARCH hybrids outperform pure neural networks in volatility prediction.

Extensive review of the literatures that predict crypto asset price with machine learning methods includes Dudek et al., 2024; Liu et al., 2023; Kim et al., 2021; Saad et al., 2020; Valencia et al., 2019. These literatures utilized several types of variables to predict crypto-asset price, which include the macro-economic development indices, global currency ratios, and generic blockchain information.

There is much literature regarding analyzing variables of crypto assets. These variables are generally categorized into four types: (i) The network variables like the active addresses (Bakhtiar et al., 2023a; Biais et al., 2020; Cong & He, 2019; Cong et al., 2020; Liebi, 2022; Pagnotta & Buraschi, 2018); (ii) the production variables related to the miners' problem (Cong et al., 2020; Sockin & Xiong, 2020); (iii) the market variables such as investor attention and Twitter Happiness Sentiment (Bakhtiar et al., 2023b; Liu & Tsyvinski, 2020; Naeem et al., 2021); and (iv) the exogenous variables (Thies et al., 2022; Jang & Lee, 2018; Poyser, 2019; Schilling & Uhlig, 2019). Apart from the exogenous variables, the other three types of variables can be reclassified into two categories: The variables related to the intrinsic value and that related to the market sentiment. The specific variable selection and categorization is described in section 3.3.

2.2. Sentiment analysis

Sentiment analysis (e.g., news impact on market sentiment) remains a cornerstone of behavioral finance research. With the rise of artificial models like natural language processing, it has made it increasingly tenable to investigate impact of news on stock returns in finance and economics (Xing et al., 2018, Wan et al., 2021). Some papers combined sentiment analyses with machine learning models to analyze sentiment impact on asset pricing. Costola et al. (2023) utilized COVID-19 event and documented that there is a significant relationship between sentiment scores and S&P 500 market returns. This is consistent with reported findings in Marty et al. (2020) that new flow of information is quickly absorbed in market prices.

In general, since the network effect plays a big role in crypto asset's acceptance, sentiment sometime can be a dominant factor in its valuation. For example, Google Trends-Sentiment hybrids (e.g., as reported on the website Quantpedia.com) achieved 41% annualized returns in BTC by combining search trends with price momentum.

To capture the distinct role that sentiment plays in cryptocurrencies' valuation, CAVM researches the value and sentiment variables, integrates the information of the value and sentiment variables, respectively, and predicts the value index and sentiment index.

While these advancements are promising, gaps persist in pragmatically linking sentimental indices to intrinsic value proxies, particularly for decentralized assets. Most models focus narrowly on price prediction than disentangling fundamental value from speculative noise - a challenge addressed in this paper through the Crypto Asset Value-indexing Model (CAVM). Additionally, few studies leverage blockchain-native data (e.g., active addresses, mean transaction fees) as intrinsic value indicators, a cornerstone of our methodology. By combining GRU-based temporal modeling with sentiment decomposition, our approach not only advances AI-driven valuation but also provides a scalable framework for regulatory and investment applications.

3. Methodology

3.1. Model design

The network structure of a deep learning model can be flexibly designed so long as it conforms to the structure of Directed Acyclic Graph. Thus, the model can be bifurcated and need not be strictly linear.

Figure 1 demonstrates the general framework of CAVM. It is a three-stage deep learning model, which includes a value indexing stage, a sentiment indexing stage, and a price prediction stage. Each stage is essentially a series of functions coupled with coefficients that need to be estimated. To determine these coefficients, we train our CAVM with the training data by fitting the price of crypto assets. Specifically, this training process involves many repeating epochs. Each of these epochs contains two steps: (i) A forward propagation step and (ii) a backward propagation step. During the forward propagation step, the value variables and the sentiment variables are used to determine the predicted price by utilizing coefficients from the previous epoch, and the associated error term[2] of this epoch is calculated by a loss function[3]. During the backward propagation step, the coefficients are updated to reduce the error of the next epoch. As the training process persists, the trained models with respect to each epoch gradually adjust from under fitting to over fitting. After training for a given number of epochs, we select a trained model that is appropriately fitted among the generated models. The selection process is described in the appendix.

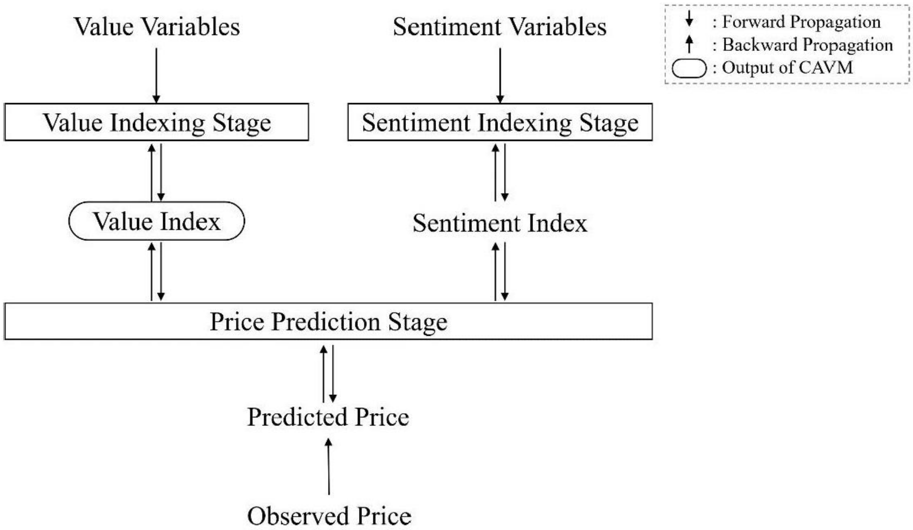

Figure 1. The general framework of CAVM.

Note: CAVM is a three-stage deep-learning model, which includes the value indexing stage, the sentiment indexing stage, and the price prediction stage. The samples of the value variables, the sentiment variables, and the observed price are used to train the CAVM. The forward and backward propagations for training the CAVM are represented by the arrows in the figure. Once the training process is completed, the value index, which is also the output of the CAVM, can be generated by the value indexing stage.

Once the training process is completed, the coefficients across the three stages are determined. Then, the value index, which is the output of the CAVM, can be generated by the value indexing stage.

[2] The error term represents the difference between the predicted price and the observed price of the current epoch.

[3] The loss function is used to calculate the error of the current epoch. There are many common regression loss functions, like Mean Squared Error (MSE) Loss, Mean Absolute Error (MAE) Loss, and Huber Loss. Since abnormal volatility caused by market sentiment is real, our model needs to capture the outliers of it. Therefore, the MSE Loss is the best choice for CAVM among the three loss functions, because MAE and Huber Loss will reduce the influence of outliers.

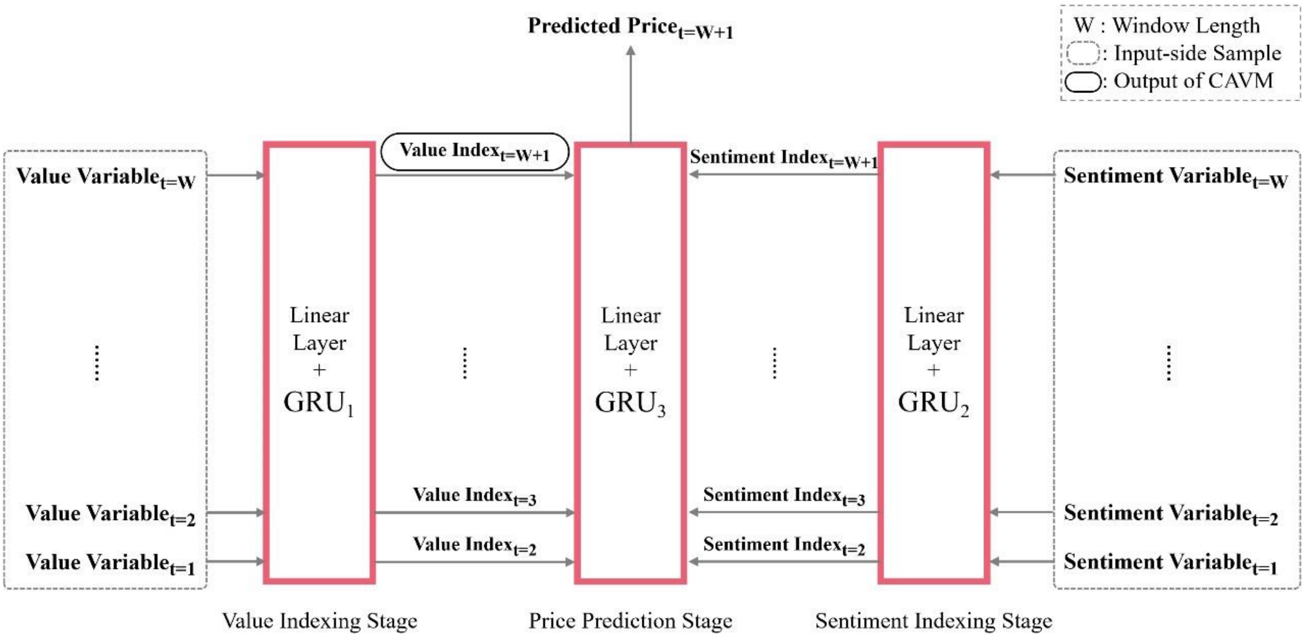

Figure 2. The time-series structure of the CAVM.

Note: The three stages of the CAVM are essentially three distinct time-series deep-learning models. Each model includes a connection of a GRU network and a linear function. The time series sample data are generated via the rolling window method with window length of W. The input-side samples are the time series observations of the value variables and the sentiment variables over W days and the output-side samples are the realized price of the crypto asset on day W+1. Thus, the output of CAVM is the value index on day W+1.

Figure 2 provides a detailed description of the time-series structure of the CAVM (forward propagation step only). The three stages of the CAVM are essentially three distinct models that implement deep learning methods on time series data. Each model includes a connection of a Gated Recurrent Unit (GRU) network and a linear function. In the literature, there are other classical deep learning networks that are applicable for time-series data, such as RNN, LSTM and Transformer (Ang et al., 2020). For our data, GRU is more appropriate than the others because it do not suffer from the issue of vanishing gradient relative to RNN (DiPietro & Hager, 2020), and it exhibits better performance than LSTM and Transformer on smaller datasets (Chung et al., 2014; Gruber & Jockisch, 2020) as it has fewer model parameters.

Our time series sample data are generated via the rolling window method with the length (in days) of W. The input-side samples are the time series observations of the value variables and the sentiment variables over W days and the output-side samples are the realized prices of the crypto asset on day W+1. Furthermore, the output of the CAVM is the value index on day W+1. This implies that the constructed value index on day t is fundamentally a predicted value.

The GRU (Gated Recurrent Unit) model outperforms traditional time series models like AR, VAR, ARCH, GARCH, ARIMA, and SARIMA in handling complex, nonlinear, and high-dimensional data. Unlike these models, which assume linearity or stationarity, GRUs can capture intricate temporal dependencies and long-term patterns through their gating mechanisms by learning dynamic relationships without strict assumptions. A much more direct comparison is given in Table 1.

Table 1. The Comparisons between deep learning and traditional means.

Dimension	CAVM	Traditional Time Series Models (AR/VAR/ARCH/GARCH/etc.)
Assumptions	No assumptions	Linearity/stationarity required
Nonlinearity	Built-in capture	Manual extensions needed
Multivariate	Handles multiple variables (value + sentiment)	Limited support
Long-term Patterns	Effective via stacked GRUs	Short-term focus
High-dim Data	Robust (implicit dimension reduction)	Struggles (parameter explosion)
Noise Handling	Resilient (multi-layer filtering)	Sensitive to outliers

Compared to traditional methods, CAVM's innovations lie in: (i) Integrating multi-dimensional time-series signals (value + sentiment); (ii) capturing non-linear temporal dependencies via GRU; and (iii) enabling high-frequency rolling predictions to adapt to the high volatility of cryptocurrency markets.

3.2. Tuning the CAVM

The basic model design described in section 3.1 is insufficient in generating a reliable value index, because it may result in a value index that (i) fluctuates excessively, (ii) is distant from the observed price, and (iii) does not move in tandem with the price of the crypto asset. The first issue suggests that useful information and signals cannot be extracted from the value index; the second issue violates our paper's premise where transactional price exhibits mean-reverting behavior; and the third issue implies that the value index may not be readily usable in predicting price. To mitigate these issues, we introduce two additional tuning mechanisms to refine the CAVM, which are shown in Table 2.

First, the loss function of the basic model design only incorporates the error between predicted price and observed price. We customize the loss function by incorporating the error between the value index and the observed price. Specifically, the weight of the error derived from the value index in the total error, termed the "weight of the value index loss" and "weight of the sentiment index loss", become a hyper-parameter of the CAVM. We expect this tuning mechanism to mitigate issue (i) by moving the value index closer to the price, and to mitigate issue (ii) by better aligning the movement of the value index and the price.

Second, we seek to mitigate issue (i) by introducing L2 regularization mechanism into the value indexing stage of the CAVM. Regularization refers to techniques that lower the complexity of a neural network during training (Kukačka et al., 2017), among which the L2 regularization is the most common one (Goodfellow et al., 2016). The incorporated L2 regularization will smoothen the value index and reduce its volatility (Georga et al., 2018). In doing so, we have generated another hyper-parameter in CAVM, termed the "weight decay coefficient". These two additional tuning mechanisms helped mitigate the issues faced by the basic model design and led to a model that has stronger predictive power. The detailed explanation about the numerical value selection of the hyperparameters is in section 6.1.

Table 2. Introducing hyperparameters to tuning CAVM.

Hyperparameters	Mechanism
Weight of the sentiment index loss	(i) By moving the predicted price closer to the price
Weight of the value index loss	(ii) By better aligning the movement of the value index
Weight decay coefficient	(iii) By introducing L2 regularization mechanism into the value indexing stage of the CAVM

3.3. Categorizing variables

To ensure that the CAVM can achieve a successful price decomposition, it is critical to appropriately classify and differentiate between the value variables and the sentiment variables. We classify independent variables as either a value variable or a sentiment variable based on their underlying intuition and market perspective. While historical prices do reflect aggregated market information, they primarily capture short-term trends and noise influenced by speculative trading and sentiment-driven fluctuations. The intrinsic value of crypto assets, however, is fundamentally driven by network-specific factors such as transactional activity, network utility, and supply dynamics (e.g., active addresses, transaction fees, and inflation rates). The definition and classification of them are given in Table 3.

DVOL is introduced in detail by deribit.com, and the detailed description of the other variables can be seen from docs.coinmetrics.io. All the independent variables are not in units of U.S. Dollars. The variables are categorized as either a value variable or a sentiment variable based on their definitions, underlying intuition, and market perspectives. The detailed description for the value variables can be seen from the websites of CoinMetrics and Deribit (Fernando, 2021). With sentiment factors ETH DVOL and BTC DVOL, both bull and bear markets, CAVM can effectively evaluate the underlying crypto currency. Market sentiment fluctuates greatly during bull and bear markets, and the DVOL variable can effectively reflect the sentiment swing.

All the independent variables we selected are not in units of U.S. Dollars. The purpose of the value indexing stage and the sentiment indexing stage of the CAVM is to capture information representing the intrinsic value and the market sentiment. If the independent variables are in units of U.S. Dollars, the CAVM will also capture information on the trend of price, which can be a source of noise for the results.

Table 3. The definitions and categories of the independent variables.

Raw Variable Name	Variable Name	Definition	Category
AdrActCnt	Active Addresses	The daily number of unique addresses that participate in activities on blockchain.	Value
FeeMeanNtv	Mean Transaction Fee	The mean transaction fee per byte of all blocks that interval in native units.	Value
TxTfrCnt	Transfers	The number of movements of ETH from one entity to another in a day.	Value
FlowTfrFromExCnt	ExTxFlow	Count transfers from Bitfinex addresses to non-Bitfinex addresses in the interval.	Value
IssContPctAnn	AnnInfRate	The percentage of new native units (continuous) issued on that day, extrapolated to one year (i.e., multiplied by 365), and divided by the current supply on that day. Sometimes referred to as the annual inflation rate.	Value
FlowOutExNtv	Exchange Withdrawals	The sum withdrawn from an exchange that day.	Value
SER	SmallToTop	Ratio of supply held by small addresses to supply held by top 1% addresses.	Value
ETH DVOL	ETH DVOL	Deribit Implied Volatility Index of ETH, utilizing the implied volatility smile of expiry of relevant options to proxy for the 30-day annualized implied volatility. It can be referred as “fear and greed gauges” (Fernando, 2021).	Sentiment
BTC DVOL	BTC DVOL	Deribit Implied Volatility Index of BTC, utilizing the implied volatility smile of expiry of relevant options to proxy for the 30-day annualized implied volatility. It can be referred as “fear and greed gauges” (Fernando, 2021).	Sentiment

4. Hypothesis development

We verify the quality of the value index and the efficiency of the CAVM via econometric and statistical testing methods. We develop our key hypotheses in this section.

4.1. Informativeness

Fundamentally, the value index integrated information provided by the value variables. We consider a value index as informative when it contains significant information provided by each of the value variables. An informative value index implies that it can be an effective proxy of the crypto asset's intrinsic value.

Our paper assumes that the crypto asset's intrinsic value on day t is primarily determined by the value variables on day t. However, the CAVM directly predicts the value index on day t via the value variables from day $(t-W)$ to day $(t-1)$. This can be illustrated intuitively via a three-step process (as shown in figure 3): (i) Predict value variables on day t using value variables from day $(t-W)$ to day $(t-1)$; (ii) determine the crypto asset's value on day t using the predicted value variables of day t; and (iii) generate the value index of day t using the determined value on day t.

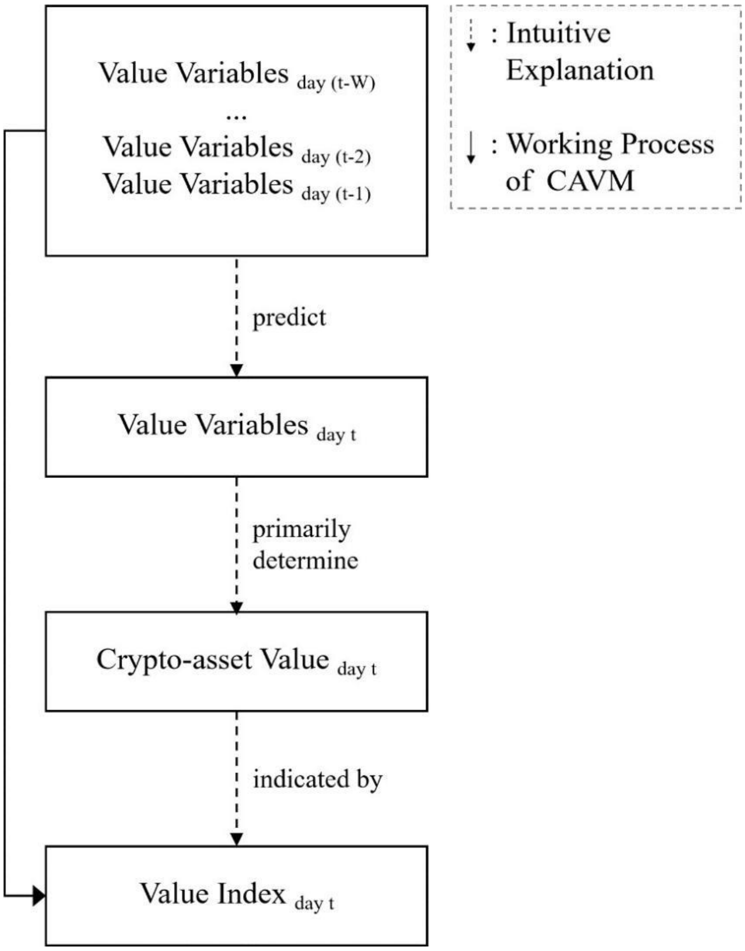

Figure 3. The working process of the value indexing stage of CAVM and its intuitive explanation.

Note: The value indexing stage of the CAVM predicts the value index on day t by the value variables from day $(t-W)$ to day $(t-1)$. However, based on the assumption that the crypto asset's value on a given day is primarily determined by the value variables on the same day, we provide an intuitive explanation underlying the working process of the value indexing stage to a three steps process illustrated in this figure. This intuition underlies the reasonability of Hypothesis 1.

Taking together, the value index is informative if the value index on day t significantly contains information provided by the observed value variables on day t. This leads us to Hypothesis 1.

Hypothesis 1: The value index of day t is significantly related with the value variables of day t.

4.2. Decomposition efficiency

The CAVM decomposes the crypto asset's price into value component and sentiment component. An efficient decomposition implies that the value index and the sentiment index should be sufficiently independent from one another. This gives rise to the following hypothesis.

Hypothesis 2a: There is an insignificant correlation between the value index and the sentiment index.

Additionally, an efficient decomposition implies that the crypto asset's price can be significantly explained by the joint explanatory power of the value index and the sentiment index. This gives rise to the following hypothesis.

Hypothesis 2b: There is significant joint explanatory power of the value index and the sentiment index on the crypto asset's price.

4.3. Predictiveness

The CAVM construct the value index on day t using data from day $t-W$ to day $t-1$. This implies that the value index on day t is fundamentally a predicted value. We expect the value index on day t to be predictive of the crypto asset's price on day t. This gives rise to the hypothesis below.

Hypothesis 3: The constructed value index on day t is significantly related to the crypto asset's price on day t.

4.4. Reasonability

Our CAVM and constructed value index can be regarded as reasonable if the outputs of the CAVM are consistent with the fundamental nature and market intuition of the underlying crypto assets. We apply our CAVM on the ETH. To ascertain the reasonability of our CAVM and the constructed value index, it is imperative that the outputs align with the intrinsic characteristics and market intuition of the underlying crypto assets. In this study, we focus on Ethereum's native cryptocurrency, ETH, to evaluate the validity of our model and index.

Ethereum serves as a foundational blockchain infrastructure, facilitating a myriad of cryptocurrencies and decentralized applications. ETH, as the medium of exchange within this ecosystem, is pivotal to the functioning and value realization of the entire network. Its value is fundamentally tied to the transactional activities and the network's expansion and utility.

The variables selected for ETH in our CAVM are meticulously chosen to reflect its essential nature and market dynamics. Active Addresses, for instance, capture the daily engagement of unique participants on the blockchain, directly correlating with the level of activity and thus the perceived value of ETH. A higher number of active addresses suggest a more vibrant and valuable network, which is intuitively expected to positively influence ETH's value index.

Mean Transaction Fee, measured in native units per byte of all blocks within a given interval, reflects the cost of utilizing the Ethereum network for transactions. This fee is indicative of the network's demand and the value placed on its transactional capabilities. As the demand for Ethereum's transactional services increases, so too does the mean transaction fee, which in turn should positively contribute to ETH's value index, aligning with market intuition.

Transfers, representing the daily movements of ETH between entities, are a direct measure of the currency's utility and liquidity. A robust transfer activity signifies a healthy market for ETH, where it is actively being used as a medium of exchange. This activity is expected to have a significant positive correlation with ETH's value index, as it reflects the fundamental role of ETH in facilitating transactions within the Ethereum ecosystem.

ExTxFlow, which counts transfers from Bitfinex addresses to non-Bitfinex addresses, provides insight into the movement of ETH between different market segments. This variable can indicate the flow of funds from centralized to decentralized platforms or vice versa, reflecting the market's confidence and the perceived value of ETH in different trading environments. A positive ExTxFlow is likely to be associated with a positive movement in ETH's value index, as it suggests a healthy and dynamic market for ETH.

AnnInfRate, or the annual inflation rate, is a crucial factor in assessing the long-term value of ETH. A lower inflation rate implies a more stable and potentially appreciating asset, as it suggests a controlled supply in the face of increasing demand. This variable is expected to have a negative correlation with ETH's value index, as lower inflation is generally perceived as more favorable for the asset's value preservation and growth.

Exchange Withdrawals, denoting the sum withdrawn from exchanges on a given day, reflect the market's sentiment towards holding ETH off-exchange. Although a higher volume of withdrawals might indicate a preference for long-term holding or usage in off-exchange activities, which could be seen as a positive sign for ETH's value, but from market sentiment point of view, it could also mean selling pressures. This variable is anticipated to have a slight negative correlation with ETH's value index, as it could suggest short-term selling pressure and user confidence in ETH's value beyond short-term trading.

SmallToTop, the ratio of supply held by small addresses to that held by top 1% addresses, offers a glimpse into the distribution of ETH holdings. A more equitable distribution, indicated by a higher SmallToTop ratio, could suggest a broader base of support for ETH's value, as it implies that the asset is not overly concentrated in the hands of a few large holders. This variable is expected to positively influence ETH's value index, as a diverse holder base is generally considered more stable and resilient to market manipulations.

Last, ETH DVOL, the Deribit Implied Volatility Index of ETH, serves as a sentiment indicator, often referred to as a "fear and greed gauge." While high volatility might initially seem detrimental, it can also attract traders and investors seeking to capitalize on price movements, thereby increasing the overall market activity and potentially the value of ETH. However, this relationship is complex and may not always be straightforward, as excessive volatility can also deter some investors. Nonetheless, ETH DVOL is an important variable to consider in the context of ETH's value index, as it reflects the market's perception of risk and opportunity associated with ETH.

In conclusion, the selected valued variables for ETH in our CAVM are intricately linked to the fundamental nature of ETH as a transactional medium and a key component of the Ethereum ecosystem. These variables are expected to coincide with the intuitive expectation based on ETH's characteristics, leading to the following hypothesis:

Hypothesis 4: The selected valued variables of ETH coincide with the intuitive expectation based on the nature of ETH

5. Data description

5.1. Sample construction and descriptive statistics

The ETH DVOL Index and BTC DVOL Index are available from 24 March 2021. We collected the 1343-day observed data starting from 24 March 2021 to 25 November 2024. The data of ETH DVOL and BTC DVOL is extracted from the API console of Deribit, and the data of other variables is downloaded from CoinMetrics. The descriptive statistics for ETH are shown in table 4.

Table 4. The descriptive statistics of ETH data.

	mean	std	min	25%	50%	75%	max
Dependent Variable							
Price	2422.17	857.60	992.79	1715.31	2296.36	3095.84	4811.16
Independent Variable							
Active Addresses	$5.45*10^5$	$8.67*10^4$	$3.30*10^5$	$4.90*10^5$	$5.33*10^5$	$5.89*10^5$	$1.51*10^5$
Mean Transaction Fee	0.004	0.004	0.00	0.002	0.003	0.01	0.07
Transfers	$9.46*10^5$	$9.38*10^4$	$6.36*10^5$	$8.88*10^5$	$9.47*10^5$	$1.00*10^6$	$1.60*10^6$
ExTxFlow	$7.58*10^4$	$2.79*10^4$	$2.55*10^4$	$5.65*10^4$	$7.23*10^4$	$9.04*10^4$	$2.72*10^5$
AnnInfRate	2.22	1.90	0.47	0.66	0.78	4.53	4.68
Exchange Withdrawals	$2.94*10^5$	$1.92*10^5$	$5.77*10^4$	$1.67*10^5$	$2.51*10^5$	$3.69*10^5$	$1.93*10^6$
SmallToTop	0.27	0.13	0.08	0.15	0.23	0.39	0.49
ETH DVOL	78.00	26.14	30.69	61.06	71.98	100.46	193.34
BTC DVOL	66.98	18.67	32.38	53.30	62.66	81.08	156.20

The data is from 22 April 2021 to 25 November 2024, and the table includes the mean, the standard deviation, the minimum, the maximum, and three percentiles of both the dependent and independent variables.

From table 4, we observe that the dependent and independent variables have different scales. Using the unscaled data, the independent variables would not contribute effectively to training the deep learning model, and the regularization mechanism described in section 3.2 will not function to expectation. To mitigate these issues, we rescale the features of the original data, which is described in section 5.2.

The correlation matrices of variables for ETH are shown in table 5. From table 5, we can see that all the absolute correlation coefficients between Price and each of the independent variables are less than 0.74. If Price is highly correlated with any of the independent variables, the trend of this independent variable may be similar with that of the price, and the CAVM may over capture information about the trend of this independent variable in the training process, resulting in the similar noise issue as described in section 3.3.

Table 5. The correlation matrix of variables for ETH.

	Price	Active Addresses	Mean Transaction Fee	Transfers	AnnInfRate	ExTxFlow	Exchange Withdrawals	SmallToTop	DVOL_ ETH	DVOL_ BTC
Price	1.00									
Active Addresses	0.36	1.00								
Mean Transaction Fee	0.35	0.17	1.00							
Transfers	0.45	0.29	0.08	1.00						
AnnInfRate	0.33	0.42	0.46	−0.05	1.00					
ExTxFlow	0.74	0.32	0.36	0.49	0.12	1.00				
Exchange Withdrawals	0.12	0.24	0.19	0.03	0.18	0.30	1.00			
SmallToTop	0.03	−0.37	−0.47	0.32	−0.78	0.14	−0.12	1.00		
DVOL_ETH	0.19	0.47	0.26	−0.17	0.74	0.09	0.42	−0.68	1.00	
DVOL_BTC	0.26	0.47	0.34	−0.16	0.76	0.19	0.42	−0.68	0.95	1.00

The data is from 24 Mar 2021 to 18 May 2022, and downloaded from coinmetrics.io/community-network-data and deribit.com/api_console.

5.2. Training, validation, and testing samples

As described in section 3.1, our time-series data consists of three types of 1343-day variables: The crypto asset's price, the set of value variables, and the set of sentiment variables. We then proceed to generate our sample observations by applying the rolling window method. Specifically, each observation consists of the crypto asset's price on day t, and the set of value variables and set of sentiment variables on day $t{-}W$, …, $t{-}2$, $t{-}1$, where W is the window's length. Using one month period as our rolling window length, our fundamental objective is to model the value of cryptocurrencies. Therefore, using an excessively small sliding window would be inappropriate. Conversely, the sentiment indexing phase aims to capture market sentiment, and using an overly large sliding window would be inadvisable. We choose a sliding window length of 30 days (i.e., W=30). If the sliding window length is significantly shortened (e.g., to 5 days) or lengthened (e.g., to 6 months), the validity of both value indexing and sentiment indexing would be compromised. Minor adjustments to the sliding window length have a minimal impact on the results. Then, we arrive at a sample size of 1372 windows.

To train the CAVM, we split the samples into three sets: Training set, validation set, and test set. Only the training set is used to fit the model parameters when training the CAVM. Thus, we regard the combination of the validation and test sets as the out-of-sample set and the training set as the within-sample set.

Given the need to conduct a hypothesis test on the informativeness of the value index for both within-sample and out-of-sample data, we require the lengths of the within-sample set and the out-of-sample set to be relatively long, because the informativeness of value variables persists beyond short-term. Hence, we set the observations before 17 August 2024 to be the training set, which contains 1214-day observations. Thus, there are 100-day observations remaining. We set the first 60 of the remaining observations to be the validation set, and the last 40 of the remaining observations to be the test set.

Finally, we apply the Z-score standardization[4] on all variables. Therefore, the outputs of CAVM are also in Z-score standardized format. The Z-score standardization is a data rescaling method that will enable the independent variables to contribute proportionately while training the CAVM and enable the L2 regularization mechanism function to expectation. Furthermore, the gradient descent will converge much more quickly (Ioffe & Szegedy, 2015).

6. Results

In this chapter, we analyze the outputs from the CAVM and discuss the results of the hypotheses tests for ETH.

6.1. The predicted value indices

The hyperparameters, “weight of the value index loss” and “weight decay coefficient” need to be determined when training the CAVM. Tuning the “weight of the value index loss” hyperparameter to 40% for ETH and setting the “weight of the sentiment index loss” to half of the “weight of the value index loss” results in the “weight of the sentiment index loss” at 20%.

The “weight decay coefficient” is tuned to 0.5 while training the CAVM for ETH. The hyperparameter should not be zero as the fluctuation of the value index needs to be managed (as described in section 3.2). Furthermore, it should not be too high as this may result in a weak informativeness of the value index.

In the training process of CAVM, the samples in the training set are first utilized to fit the model. The updated model parameters are saved on every epoch during the training process. Next, we apply the saved CAVMs on the validation set and generate loss curves for both the training set and the validation set. Thereafter, we select the best one among the saved CAVMs according to the loss curves. The specific selection process is documented in the appendix. Finally, we administer the selected CAVM on the test set.

Using the trained CAVM, we derive the value index for ETH on all the three sample sets. The sentiment index and the predicted price for ETH are also output. After restoring the CAVM’s outputs from the Z-score standardized format to the realized price’s scale, we plot the curves of observed price and predicted value index for ETH in Figure 4.

[4] Z-score Standardization makes the values of each variable in the data have zero-mean and unit-variance. It involves two steps: subtracting the variable’s mean from this variable, and then dividing the variable by its standard deviation.

From the figure, we observe that the value indices are less volatile than the crypto assets' observed price. This trend appears to be persisted over the longer term, which is consistent with the notion that the trend of intrinsic value is generally more stable than that of the price. In addition, the observed price exhibits a mean-value reverting behavior in the figure. Thus, it can be inferred that our value index plays an important role in determining the crypto asset's price. Finally, we observe that the price movements tend to lag the value index during most of the mean-value reverting behavior. This lag implies that our value index may serve as a signal of subsequent price movements and may be helpful for crypto assets' investors.

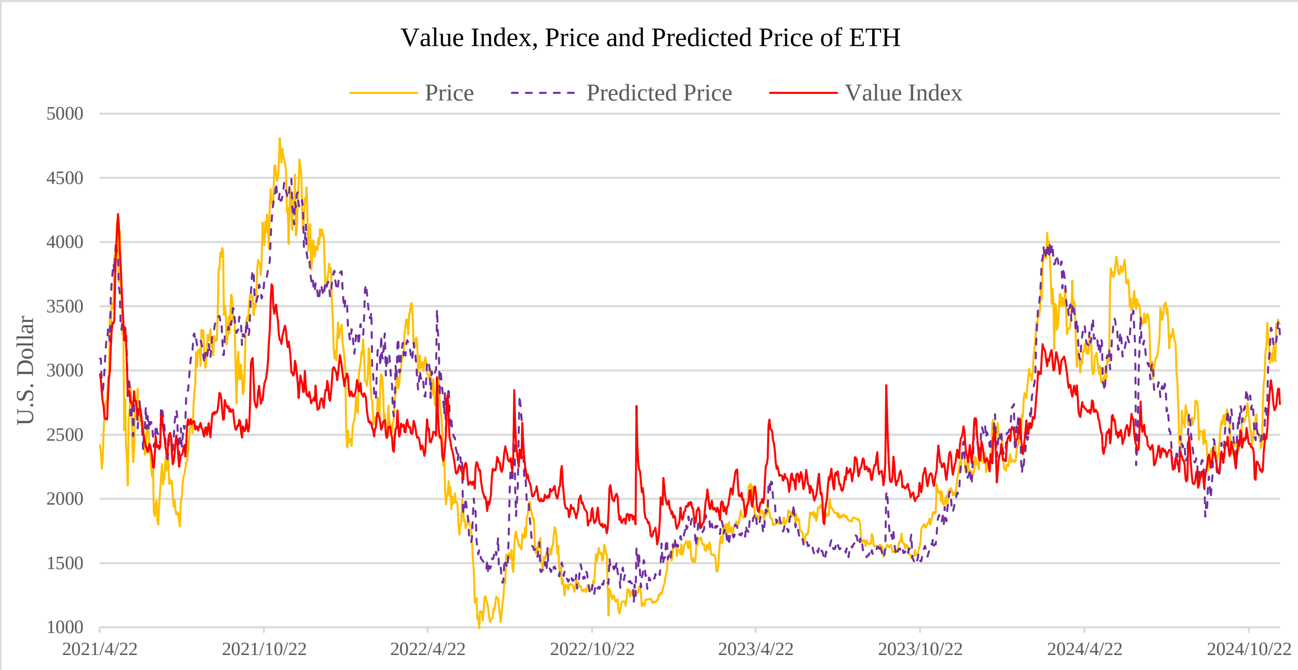

Figure 4. The curves of the value index, the predicted price, and the observed price of ETH from 22 April 2021 to 25 November 2024.

Note: Both the value index and the predicted price are the outputs of the trained CAVM. The orange curve is the observed price, the purple curve is the predicted price, and the red curve is the value index. The training set is up to 17 August 2024, and the remaining are the validation set and the test set.

6.2. Hypothesis testing results

To ensure tractability, we use the Z-score standardized data in our hypothesis tests, noting that this rescaling will not affect the significance of the test results.

6.2.1. Informativeness test results

We implement an Ordinary Least Squares (OLS) regression model to test Hypothesis 1. We build the OLS model. Specifically, we regress the predicted value index on day t against the observed value variables on day t. The regression model for ETH is given in equation (1),

$$VI_t = \beta_0 + \beta_1 \cdot Adr_t + \beta_2 \cdot \text{MeanTrFee}_t + \beta_3 \cdot \text{Tr}ant_t + \beta_4 \cdot \text{Ex}TF_t + \beta_5 \cdot \text{A}InfR_t + \beta_6 \cdot \text{E}xWith_t + \beta_7 \cdot STT_t + \epsilon \quad (1)$$

where VI_t denotes the predicted Value Index of ETH on day t; Adr_t denotes Active Addresses of ETH on day t; MeanTrFee_t denotes Mean Transaction Fee of ETH on day t; Tran_t denotes Transfers, the number of movements of ETH from one entity to another on day t; ExTF_t denotes ExTxFlow, count transfers from Bitfinex addresses to non-Bitfinex addresses in the interval on day t.; AInfR_t denotes AnnInfRate, the percentage of new units issued on day t, scaled to a yearly figure and divided by the current supply; ExWith_t denotes Exchange Withdrawals, the sum withdrawn from an exchange on day t; and STT_t denotes SmallToTop, ratio of supply held by small addresses to supply held by top 1% addresses. The results of the OLS regression for ETH are listed in table 6.

As shown in table 6, the coefficient for Transfers, Active Addresses, Mean Transaction Fee, AnnInfRate, ExTxFlow, and SmallToTop are positive, while the coefficient for Exchange Withdrawals is negative. More particularly, the coefficients of these are significant at the 0.1% level. Furthermore, the F-statistic is 1140.00 and it is significant at the 0.1% level, which suggests that there exists explanatory power in the value variables on the predicted value index of ETH. These results provide support for Hypothesis 1 of ETH where the predicted value index is informative as it is significantly related to all the value variables.

Table 6. The OLS regressions' results of the value index informativeness test for ETH.

Independent Value Variable Whole Sample	Coefficient	t-value	p-value	R2 / Adjusted R2	F-statistic / Prob(F-statistic)
Constant	−0.03	−6.79***	0.000***		
Active Addresses	0.09	14.67***	0.000***		
Mean Transaction Fee	0.05	7.56***	0.000***		
Transfers	0.03	4.67***	0.000***	0.859	1140.00***
AnnInfRate	0.19	24.43***	0.000***	0.859	0.000***
ExTxFlow	0.27	41.44***	0.000***		
Exchange Withdrawals	−0.02	−4.47***	0.000***		
SmallToTop	0.09	9.86***	0.000***		

Note: The explained variables of the regressions are the whole-sample value indices of ETH, which are output by the trained CAVM of ETH. The explanatory variables are the intrinsic value variables of ETH. There are 1343 observations for the regression. The asterisk superscripts represent the significance level of the regression results, in which three asterisks represent the result is significant at the 0.1% level, and one asterisk represents the result that is significant at the 5% level.

The test results support our Hypothesis 1, and we conclude that the constructed value index for crypto assets is informative. This implies that the value index can significantly capture the information provided by the value variables, and it can be an effective proxy for the crypto asset's intrinsic value. Our results also suggest that the CAVM may perform well on the future unobserved data in generating an informative value index.

Table 7. The Pearson correlation coefficients and their p-values between the value and sentiment indices of ETH.

Crypto Asset	Pearson Correlation Coefficient	p-value
ETH	0.69***	0.000***

The indices are output by the trained CAVM of ETH. The samples for the calculation of the correlations are the 1343-day value and sentiment indices from 22 April 2021 to 25 November 2024. The three-asterisk superscripts represent the correlations are significant at the 0.1% level.

6.2.2. Decomposition test results

A successful decomposition of crypto asset's price into the value index and the sentiment index implies that there should not be a strong correlation between the two indices. To test the decomposition efficiency, we combine the training, validation, and test sets because we conclude in section 6.2.1 that the value index is informative throughout the sample. We document in table 7 that the Pearson correlation coefficients between the value index and sentiment index are 0.69 for ETH. Furthermore, the p-value of the correlation for ETH is 0.00. For ETH, the correlation coefficients are significant at the 0.1% level and are relatively small. This suggests that the value index and the sentiment index are weakly correlated with each other, which is the evidence for Hypothesis 2a of ETH.

Additionally, a successful decomposition implies that there should be significant joint explanatory power of the value index and the sentiment index on the crypto asset's price.

Hence, we run three regressions according to equations (3), (4), and (5) for ETH.

$$\mathrm{Pr}ice_{\mathrm{t}} = \beta_0 + \beta_1 \cdot SI_t + \varepsilon_t \tag{3}$$

$$\mathrm{Pr}ice_{\mathrm{t}} = \beta_0 + \beta_1 \cdot VI_t + \varepsilon_t \tag{4}$$

$$\mathrm{Pr}ice_{\mathrm{t}} = \beta_0 + \beta_1 \cdot VI_t + \beta_2 \cdot SI_t + \varepsilon_t \tag{5}$$

where $Price_t$ is the crypto asset's observed price on day t, VI_t is the predicted value index on day t, and SI_t is the predicted sentiment index on day t. Equations (3) and (4) are separate regressions and equation (5) is the joint regression model.

Next, we calculate the Joint F-test statistic (Blackwell, 2008) using the sum squared residuals (SSRs) calculated from each of the three regression models above, and the formulas are given in equation (6) and equation (7).

$$\text{Joint } F_{VI} = \frac{(SSR_{VI} - SSR_{VS})/q}{SSR_{VS}/rdf} \tag{6}$$

$$\text{Joint } F_{SI} = \frac{(SSR_{SI} - SSR_{VS})/q}{SSR_{VS}/rdf} \tag{7}$$

where SSR_{VI}, SSR_{SI}, and SSR_{VS} are the SSRs calculated from regression models (3), (4), and (5), respectively; $Joint\ F_{VI}$ is the Joint F-test statistic for the value index, and $Joint\ F_{SI}$ is the Joint F-test statistic for the sentiment index; rdf is the Residual Degrees of Freedom of regression model (5); and q is the cost of rdf s between regression models (3) and (5) or between regression models (4) and (5). We summarize the key results of the above regressions and Joint F-test statistics for ETH in Table 8.

Table 8. The results of the OLS regressions of value index on price, sentiment index on price, combination of value and sentiment indices on price, and the joint F-tests among these regressions for ETH.

Independent Value Variable	Coefficient	t-value	p-value	R2 / Adjusted R2	F-statistic / Prob(F-statistic)
Regression (3) of Value Index on Price					
Constant	0.07	4.21^{***}	0.000^{***}	0.669	2657.00^{***}
Value Index	1.86	51.54^{***}	0.000^{***}	0.669	0.000^{***}
Regression (4) of Sentiment Index on Price					
Constant	−0.18	-11.04^{***}	0.000^{***}	0.679	2777.00^{***}
Sentiment Index	1.00	52.70^{***}	0.000^{***}	0.679	0.000^{***}
Regression (5) of both Value and Sentiment Indices on Price					
Constant	−0.07	-5.14^{***}	0.000^{***}	0.796	2555.00^{***}
Value Index	1.08	27.37^{***}	0.000^{***}	0.796	0.000^{***}
Sentiment Index	0.60	28.49^{***}	0.000^{***}		
Joint Testing Results					
SSR of (3):	428.13	SSR of (4): 41.14		SSR of (5):	272.41
Residual Degrees of Freedom of (5): 1311.0					
Joint F-statistic		Value Index	749.38^{***}	p-value	0.000^{***}
		Sentiment Index	811.95^{***}		0.000^{***}

Note: The regressions (3), (4), and (5) in the table are according to formulas (3), (4), and (5). The value and sentiment indices are output by the trained CAVM of ETH. There are 1343-day observations for all three regressions. The three-asterisk superscripts represent the result as significant at the 0.1% level, and the t-values and p-values without asterisk are insignificant.

In Table 8, we document that the price of ETH can be significantly explained by the combination of the value index and the sentiment index. First, we note that in the joint regression model given by equation (5), the coefficient of the value index and the sentiment index are 1.08 and 0.60, respectively. These coefficients are significant at the 0.1% level. Furthermore, the goodness-of-fit F-test statistic is 2555.00, which is also significant at the 0.1% level. This suggests that the combined model provides significant explanatory power on ETH's price. In addition, by conducting two joint F-tests according to equations (6) and (7), we document the joint F-test statistics of 749.38 for the value index and 811.95 for the sentiment index, and they are significant at the 0.1% level. This implies that the explanatory power of the combination of the value index and sentiment index is significantly better than regressing them individually. Hence, the results support our hypothesis 2b that there is significant joint explanatory power of the value index and the sentiment index on the price of ETH.

6.2.3. Predictiveness test results

We hypothesize that the predicted value index on day t is significantly related to the crypto asset's price on day t. To test this hypothesis, we look at the results of equation (3) where we regress the crypto asset's price on the constructed value index. Specifically, in table 8, we document a coefficient value of 1.86 which is significant at the 0.1% level. This provides evidence that the constructed value index for ETH on day t is significantly positively related to ETH's price on day t. Our results provide support for Hypothesis 3.

Our results are pragmatic and of interest to crypto assets' investors because our value index and the CAVM are predictive of crypto assets' prices. The CAVM is trained with historical observed data, and it can be applied to predict the daily value index in the future with value variables from the past W-th day to the past day. Therefore, our CAVM may be used to guide the trading decisions of crypto assets' investors.

6.2.4. Reasonability test results

Following Hypothesis 4, we propose that the intrinsic value of ETH is primarily determined by the value factors we intentionally selected. In table 6, the coefficients of all the listed variables are significant at the 0.1% level, which supports the idea of Hypothesis 1. Apart from the significance, analyzing how those chosen variables contribute to the intrinsic value of the underly asset shows the reasonability. Reasonability can be described as intuitive expectation based on the nature of ETH.

From Table 6, within those factors, Exchange Withdrawals contributes to the intrinsic value of ETH in a negative way, the coefficient is −0.02. While this variable represents the sum withdrawn from an exchange that day, this negative coefficient suggests that an increase in the amount of ETH withdrawn from exchanges is associated with a decrease in the intrinsic value of ETH. This could be interpreted as indicating that when more ETH is withdrawn from exchanges, it may lead to a reduction in the perceived value of ETH in the market, possibly due to increased selling pressure or a shift in market sentiment. This could be because when ETH is held off-exchange, it might be perceived as a positive sign for its value, indicating a preference for long-term holding or usage in off-exchange activities. We also suspect this variable's negative coefficient may have something to do with the dataset used being from a bear market.

Other variables like Active Addresses: The positive coefficient of 0.09 suggest that an increase in the number of unique addresses participating in activities on the blockchain is associated with an increase in ETH's value index. This aligns with the expectation that network activity is a positive indicator of value.

Mean Transaction Fee: The positive coefficient of 0.05 indicates that higher transaction fees are correlated with a higher value index for ETH. This is reasonable as it reflects the demand for using ETH for transactions on the Ethereum network.

Transfers: The coefficient of 0.03 is positive, suggesting that more transfers of ETH between entities are associated with a higher value index. This is consistent with the idea that liquidity and utility of ETH contribute positively to its value.

AnnInfRate: The positive coefficient of 0.19 might seem counterintuitive at first, as we perceived in section 4.4, so one might expect a higher inflation rate to decrease value. However, in the context of ETH, it could reflect the market's anticipation of future demand or the relative scarcity in the context of total supply, which can be perceived as positive for value.

ExTxFlow (External Transaction Flow): The strong positive coefficient of 0.27 implies that an increase in transfers from Bitfinex addresses to non-Bitfinex addresses is significantly associated with an increase in ETH's value index. This could be interpreted as a sign of increased market activity and confidence in ETH's value.

Ratio of Supply Held by Small Addresses to Top Addresses (SmallToTop): The positive coefficient of 0.09 indicates that a more equitable distribution of ETH holdings is associated with a higher value index. This aligns with the expectation that a broader base of support for ETH's value is a positive factor.

Each variable's coefficient and significance level in the regression results should be interpreted in the context of the variable's role in the Ethereum ecosystem and the market's perception of its impact on ETH's value. The positive coefficients for most variables suggest that they contribute positively to ETH's value index, which is in line with the intuitive expectations based on the nature of ETH. The negative coefficient for Exchange Withdrawals also aligns with the interpretation that holding ETH off-exchange might be seen as a positive sign for its value.

Thus, our paper provides significant support for the intuitive expectation based on the nature of ETH. This suggests that our constructed value index is reasonable.

7. Regulatory implications and practical adoption challenges

7.1. Regulatory implications

The CAVM's ability to decompose crypto asset prices into value and sentiment indices offers significant potential for enhancing financial regulation and taxation frameworks. Key implications include:

Fair Market Valuation for Taxation and financial reporting: The value index provides a stable proxy for intrinsic value, reducing reliance on volatile market prices for tax assessments. This could mitigate disputes during tax audits. By anchoring tax liabilities to value indices, regulators can ensure consistency, even during speculative bubbles or crashes. Under IFRS and GAAP, intangible assets like crypto require periodic revaluation. The value index could serve as an auditable benchmark for "fair value" measurements. The auditability of the benchmark requires significant amounts of data collections and standardization on the blockchain.

Anti-Manipulation Surveillance: Sentiment indices derived from CAVM (especially those behavioral/sentiment metrics) could help regulators detect market manipulation (e.g., pump-and-dump schemes) by identifying abnormal deviations between price and intrinsic value. For instance, the SEC's Cyber and Emerging Technology Unit could leverage such models to flag suspicious trading patterns.

7.1.1. Basel III and risk weighting

Banks holding crypto assets may use the value index to compute risk-weighted capital buffers. A stable intrinsic value metric would align with Basel III's emphasis on prudential risk management for volatile assets.

Under the Basel III framework, banks are required to maintain capital buffers proportional to the riskiness of their assets. Crypto assets, classified as high-risk due to their volatility, face punitive risk weights (e.g., 1250% under Basel III's "Group 2" crypto classification). The CAVM's value index, a stable proxy for intrinsic value, can refine this approach by providing a closer to intrinsic value risk assessment grounded in fundamental metrics than speculative market prices.

Basel III mandates that risk weights reflect an asset's inherent risk. The value index enables banks to: Compute the volatility (standard deviation) of the value index instead of market prices. For example, if ETH's value index exhibits 20% annualized volatility versus 80% for its market price, the risk weight could be calibrated to the lower volatility, reducing capital requirements.

Furthermore, regulators might impose a conservative haircut on the value index to account for residual uncertainty (e.g., 15–30%), like adjustments for illiquid assets. The haircut ensures prudential safeguards while acknowledging the index's stability.

With the value index approach, one can align with Basel III's Prudential Principles better: The value index decomposes speculative (sentiment) from fundamentals, aligning with Basel III's push for risk-sensitive models. By tethering risk weights to intrinsic value, banks avoid overcapitalizing during irrational market swings (e.g., crypto "fear and greed" cycles). Furthermore, the decomposition adds insight to understanding risk sensitive to different market drivers.

Regulators could mandate stress scenarios where the value index is adjusted for systemic shocks (e.g., blockchain forks, DeFi collapses) for stress testing purposes. This ensures capital adequacy even if intrinsic value metrics degrade. The value index's stability aids in projecting reliable cash flows for LCR compliance, as it reflects durable network activity (e.g., transaction fees, active addresses) than transient sentiment.

7.1.2. The Basel III regulatory adoption pathway

For implementation of the value index approach, regulators need to first define eligibility criteria, specifically, define which value indices (e.g., CAVM, there are various indicators that can be included in the indices) meet Basel III's validation standards for robustness and transparency; increase data governance, e.g., blockchain data sources (e.g., CoinMetrics) require certification to ensure integrity; and mandate granular reporting of value index inputs (e.g., active addresses, transaction volume) to prevent model manipulation.

Regulators must validate the CAVM's methodology (the model) to prevent general undercapitalization and require banks to back test against historical crises (e.g., 2022 crypto winter) to ensure resilience. Thus, Basel Committee guidance would harmonize cross-border adoption, avoiding regulatory arbitrage.

7.2. Practical adoption challenges

Despite its promise, implementing the CAVM in regulatory and institutional settings faces hurdles, as we suggest in the Basel III adoption pathway section.

The adoption challenge list below is far from a comprehensive one:

1. Data Standardization: The model relies on granular blockchain metrics (e.g., active addresses, transaction fees). However, data availability varies across exchanges and blockchains. Regulatory mandates for standardized on-chain reporting (e.g., MiCA in the EU) will be critical for scalability.

2. Computational Complexity: Deep learning models require significant computational resources. Smaller tax authorities or institutions in emerging markets may lack infrastructure for real-time index generation, necessitating cloud-based solutions or regulatory sandboxes.

3. Interpretability and Trust: While CAVM's GRU architecture improves performance on small datasets, its "black-box" nature could hinder adoption among regulators accustomed to transparent models. Explainable AI techniques, such as SHAP values, must be integrated to clarify variable contributions.

4. Cross-Jurisdictional Coordination: Crypto markets are global, but regulations remain fragmented. A value index adopted in one jurisdiction (e.g., Monetary Authority of Singapore) may

conflict with another’s approach (e.g., India’s punitive crypto tax policies). Harmonization by regulatory bodies will be essential.

5. Industry Resistance: Exchanges and funds profiting from volatility may oppose intrinsic value benchmarks. For CAVM to become the crypto valuation standard, regulators must compel adoption through laws, deadlines, and penalties, mirroring the global effort to replace LIBOR. The LIBOR analogy underscores that transformative financial framework, even objectively better ones, rarely succeed without regulatory muscle.

By bridging crypto’s unique risks with regulatory principles, the CAVM could evolve into a cornerstone of crypto asset regulation and transform speculative assets into tractable instruments for prudential regulation, fostering financial stability without stifling innovation.

8. Conclusion and future research

In this paper, we construct value indices for crypto asset ETH by implementing a three-stage time-series deep learning model (CAVM) to decompose the crypto assets’ price to value index and sentiment index. In the model, the value variables’ information is integrated to generate the value index, and the sentiment variables’ information is integrated to generate the sentiment index. The CAVM is trained with historical data and can be applied to predict the daily value index in the future. Furthermore, we provide empirical support for various inferences that are consistent with the underlying intuition of ETH. Therefore, we postulate that the value index can be an effective proxy for the crypto asset’s intrinsic value. Thus, our value index may serve as an effective benchmark to the investment, consumption, and financial reporting on crypto assets.

To assess variable stability, our OLS regression coefficients (Table 6) quantify the relative importance of each variable value. For example, **Active Addresses** (β=0.32, p<0.01) and **Transaction Fees** (β=0.28, p<0.01) exhibit high sensitivity, indicating their critical role in ETH’s value index. Future work may incorporate variance decomposition.

Future research directions include extending the CAVM framework to other cryptocurrencies such as Bitcoin (BTC) or Solana (SOL) to validate its external validity. While we focus on ETH due to its unique ecosystem characteristics and data availability, the modular design of the CAVM enables adaptation to other assets by adjusting value variables specific to their network fundamentals. For instance, BTC’s scarcity metrics (e.g., block reward halving events) could be incorporated into the value index to assess cross-asset applicability.

While our primary focus of this paper is on valuation theory, the value index’s predictive power (as demonstrated by the **predictiveness test** in Section 6.2.3) suggests potential applications in trading strategies. For instance, deviations between the market price and the value index could signal overvaluation or undervaluation, enabling mean-reversion strategies. However, practical implementation requires further analysis of transaction costs, latency, and regulatory constraints, which are beyond the scope of this theoretical framework.

Author contributions

Esther Ying Yang: Conceptualization, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. Xi Zhou: Data curation, Investigation, Resources, Visualization, Writing – original draft, Writing – review & editing. Yin Pang: Formal analysis, Software, Validation, Writing – original draft, Writing – review & editing. Jing Rong Goh: Data curation, Investigation, Resources, Visualization, Writing – original draft, Writing – review & editing. Shaun Shuxun Wang: Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in creating this article.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Akyildirim E, Goncu A, Sensoy A (2021) Prediction of cryptocurrency returns using machine learning. *Ann Oper Res* 297: 3–36. https://doi.org/10.1007/s10479-020-03575-y

Ang JS, Ng KW, Chua F (2020) Modeling Time Series Data with Deep Learning: A Review, Analysis, Evaluation and Future Trend. *2020 8th International Conference on Information Technology and Multimedia (ICIMU),* 32–37. https://doi.org/10.1109/ICIMU49871.2020.9243546

Avramov D, Cheng S, Metzker L (2023) Machine Learning vs. Economic Restrictions: Evidence from Stock Return Predictability. *Manage Sci* 69: 2587–2619. https://doi.org/10.1287/mnsc.2022.4449

Bakhtiar T, Luo X, Adelopo I (2023a) Network effects and store-of-value features in the cryptocurrency market. *Technol Soc* 74: 102320. https://doi.org/10.1016/j.techsoc.2023.102320

Bakhtiar T, Luo X, Adelopo I (2023b) The impact of fundamental factors and sentiments on the valuation of cryptocurrencies. *Blockchain-Res Appl* 4: 100154. https://doi.org/10.1016/j.bcra.2023.100154

Barth JR, Herath HSB, Herath TC, et al. (2020) Cryptocurrency valuation and ethics: a text analytic approach. *J Manag Anal* 7: 367–388. https://doi.org/10.1080/23270012.2020.1790046

Biais B, Bisiere C, Bouvard M, et al. (2020) Equilibrium bitcoin pricing. *J Financ* 78: 967–1014. https://doi.org/10.1111/jofi.13206

Bianchi D, Büchner M, Tamoni A (2021) Bond Risk Premiums with Machine Learning. *Rev Financ Stud* 34: 1046–1089. https://doi.org/10.1093/rfs/hhaa062

Blackwell M (2008) Multiple hypothesis testing: The F-test. *Matt Blackwell Res*, 1–7. Available from: https://www.mattblackwell.org/files/teaching/ftests.pdf.

Chen L, Pelger M, Zhu J (2024) Deep Learning in Asset Pricing. *Manage Sci* 70: 714–750. https://doi.org/10.1287/mnsc.2023.4695

Chung J, Gulcehre C, Cho K, et al. (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling, In: *NIPS 2014 Workshop on Deep Learning, December 2014*.

Cong LW, He Z (2019) Blockchain Disruption and Smart Contracts. *Rev Financ Stud* 32: 1754–1797. https://doi.org/10.1093/rfs/hhz007

Cong LW, He Z, Li J (2020) Decentralized Mining in Centralized Pools. *Rev Financ Stud* 34: 1191–1235. https://doi.org/10.1093/rfs/hhaa040

Costola M, Hinz O, Nofer M, et al. (2023) Machine learning sentiment analysis, COVID-19 news and stock market reactions. *Res Int Bus Financ* 64: 101881. https://doi.org/10.1016/j.ribaf.2023.101881

Dama F, Sinoquet C (2021) Time Series Analysis and Modeling to Forecast: a Survey. Available from: https://arxiv.org/abs/2104.00164

DiPietro R, Hager GD (2020) Deep learning: RNNs and LSTM, In: *Handbook of Medical Image Computing and Computer Assisted Intervention,* 503–519, Elsevier. https://doi.org/10.1016/B978-0-12-816176-0.00026-0

Dudek G, Fiszeder P, Kobus P, et al. (2024) Forecasting cryptocurrencies volatility using statistical and machine learning methods: A comparative study. *Appl Soft Comput* 151: 111132. https://doi.org/10.1016/j.asoc.2023.111132

Fernando S (2021) DVOL-Deribit Implied Volatility Index. Available from: https://insights.deribit.com/exchange-updates/dvol-deribit-implied-volatility-index/.

García-Monleón F, Danvila-del-Valle I, Lara FJ (2021) Intrinsic value in crypto currencies. *Technol Forecast Soc* 162: 120393. https://doi.org/10.1016/j.techfore.2020.120393

Georga EI, Fotiadis DI, Tigas SK (2018) Methodology for Developing a Glucose Prediction Model, In: E. I. Georga, D. I. Fotiadis, & S. K. Tigas (Eds.), *Personalized Predictive Modeling in Type 1 Diabetes,* 37–59, Academic Press. https://doi.org/10.1016/B978-0-12-804831-3.00003-0

Goodfellow I, Bengio Y, Courville A (2016) *Deep Learning.* MIT Press.

Graham B, Dood D (2008) *Security Analysis* (6th Edition), McGraw Hill.

Gruber N, Jockisch A (2020) Are GRU Cells More Specific and LSTM Cells More Sensitive in Motive Classification of Text? *Front Artif Intell* 3: 40. https://doi.org/10.3389/frai.2020.00040

Gu S, Kelly B, Xiu D (2020) Empirical Asset Pricing via Machine Learning. *Rev Financ Stud* 33: 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Ioffe S, Szegedy C (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International Conference on Machine Learning*.

Jang H, Lee J (2018) An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. *IEEE Acces*s 6: 5427–5437. https://doi.org/10.1109/ACCESS.2017.2779181

Johnson R, Bufton J, Danie JG (2019) The Valuation of Crypto-assets. Ernst and Young. Available from: https://www.ey.com/en_qa/financial-services--emeia-insights/the-valuation-of-crypto-assets.

Kelly BT, Kuznetsov B, Malamud S, et al. (2025) Artificial Intelligence Asset Pricing Models (January 2025). NBER Working Paper No. w33351. Available from: https://ssrn.com/abstract=5103546.

Khan FU, Khan F, Shaikh PA (2023) Forecasting returns volatility of cryptocurrency by applying various deep learning algorithms. *Futur Bus J* 9: 1–11. https://doi.org/10.1186/s43093-023-00200-9

Kim HM, Bock GW, Lee G (2021) Predicting Ethereum prices with machine learning based on Blockchain information. *Expert Syst Appl* 184: 115480. https://doi.org/10.1016/j.eswa.2021.115480

Kukačka J, Golkov V, Cremers D (2017) Regularization for deep learning: A taxonomy.

Liebi LJ (2022) Is there a value premium in cryptoasset markets? *Econ Model* 109: 105777. https://doi.org/10.1016/j.econmod.2022.105777

Liu Y, Zhang L (2023) Cryptocurrency Valuation: An Explainable AI Approach, In: Arai, K (eds) *Intelligent Computing*, SAI 2023, Lecture Notes in Networks and Systems, 711, Springer, Cham. https://doi.org/10.1007/978-3-031-37717-4_51

Liu Y, Li Z, Ramzi N, Jahangir S (2023) Forecasting cryptocurrency returns with machine learning. *Res Int Bus Financ* 64: 101905. https://doi.org/10.1016/j.ribaf.2023.101905

Liu Y, Tsyvinski A (2020) Risks and Returns of Cryptocurrency. *Rev Financ Stud* 34: 2689–2727. https://doi.org/10.1093/rfs/hhaa113

Marty T, Vanstone B, Hahn T (2020) News media analytics in finance: a survey. *Account Financ* 60: 1385–1434. https://doi.org/10.1111/acfi.12466

Naeem MA, Mbarki I, Shahzad SJ (2021) Predictive role of online investor sentiment for cryptocurrency market: Evidence from happiness and fears. *Int Rev Econ Financ* 73: 496–514. https://doi.org/10.1016/j.iref.2021.01.008

Pagnotta ES, Buraschi A (2018) An Equilibrium Valuation of Bitcoin and Decentralized Network Assets. *SSRN Electron J.* https://doi.org/10.2139/ssrn.3142022

Poyser O (2019) Exploring the dynamics of Bitcoin's price: a Bayesian structural time series approach. *Eur Econ Rev* 9: 29–60. https://doi.org/10.1007/s40822-018-0108-2

Saad M, Choi J, Nyang D, et al. (2020) Toward Characterizing Blockchain-Based Cryptocurrencies for Highly Accurate Predictions. *IEEE Syst J* 14: 321–332. https://doi.org/10.1109/JSYST.2019.2927707

Schilling L, Uhlig H (2019) Some simple bitcoin economics. *J Monetary Econ* 106: 16–26. https://doi.org/10.1016/j.jmoneco.2019.07.002

Shapiro C, Varian H (1999) *Information Rules: A Strategic Guide to The Network Economy,* Harvard Business School Press, 1999.

Sockin M, Xiong W (2020) A Model of Cryptocurrencies. *Manage Sci* 69: 6684–6707. https://doi.org/10.3386/w26816

Thies F, Wallbach S, Wessel M, et al. (2022) Initial coin offerings and the cryptocurrency hype - the moderating role of exogenous and endogenous signals. *Electron Mark* 32: 1691–1705. https://doi.org/10.1007/s12525-021-00460-9

Valencia F, Gómez-Espinosa A, Valdés-Aguirre B (2019) Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning. *Entropy* 21: 589. https://doi.org/10.3390/e21060589

Verma A, Walia R, Kumar V (2024) Advanced predictive analytics for cryptocurrency valuation: A comparative study of machine learning and deep learning approaches, In: *Computational Methods in Science and Technology*, 1st Edition, CRC Press. https://doi.org/10.1201/9781003501244-60

Wan X, Yang J, Marinov S, et al. (2021) Sentiment correlation in financial news networks and associated market movements. *Sci Rep* 11: 3062. https://doi.org/10.1038/s41598-021-82338-6

Xing FZ, Cambria E, Welsch RE (2018): Natural language based financial forecasting: a survey. *Artif Intell Rev* 50: 49–73. https://doi.org/10.1007/s10462-017-9588-9

Yildirim H, Bekun FV (2023) Predicting volatility of bitcoin returns with ARCH, GARCH and EGARCH models. *Future Bus J* 9: 75. https://doi.org/10.1186/s43093-023-00255-8

Zhang W, Wang P (2020) Investor attention and the pricing of cryptocurrency market. *Evol Inst Econ Rev* 17: 445–468. https://doi.org/10.1007/s40844-020-00182-1