Aggregate Investor Attention and Bitcoin Return: The Long Short-term Memory Networks Perspective

Chen Wang

College of Management and Economics, Tianjin University, Tianjin, 300072, P.R. China
Email: saphirachen@tju.edu.cn

Dehua Shen [*]

School of Finance, Nankai University, Tianjin, 300350, P.R. China
Email: dhs@nankai.edu.cn

Youwei Li

Hull University Business School, University of Hull, HU6 7RX, UK
Email: Youwei.Li@hull.ac.uk

Acknowledgments:

This work is supported by the National Natural Science Foundation of China (72071141).

[*] **Corresponding author**

Dehua Shen, Associate Professor of Finance

School of Finance, Nankai University

No.38 Tongyan Road, Jinnan District, Tianjin, 300350, P.R. China

Email: dhs@nankai.edu.cn

Aggregate Investor Attention and Bitcoin Return: The Long Short-term Memory Networks Perspective

Abstract

Investor attention is a scarce cognitive resource which affects investment decisions, and recent studies suggest that investor attention also have impacts on asset prices. Although Bitcoin is found to be one of the most unpredictable cryptocurrencies with excessive volatilities, researchers are still looking for determinants of Bitcoin prices. In this study, we firstly adopt the Long Short-Term Memory Networks (LSTM) approach to evaluate the effect of investor attention on Bitcoin returns by constructing an aggregate investor attention proxy. We combine both direct and indirect proxies for investor attention, in addition to the Bitcoin trading variables as the LSTM inputs. Our empirical results suggest that the including of attention variables could effectively improve the LSTM's prediction accuracy of Bitcoin prices, whereas direct proxies (i.e. daily Google Trends to Bitcoin and Bitcoin tweets) contains more valuable information to further improve the LSTM's forecasting capacity.

Keywords: Investor attention; Bitcoin; Machine learning; LSTM; Social media

1. Introduction

Bitcoin is found to be one of the most unpredictable cryptocurrencies with excessive volatilities (Brauneis and Mestel, 2018). As Bitcoin is a recent phenomenon with a relatively short history, whether it is a real currency or a speculative asset is still controversial (Yermack 2015, Corbet et al., 2018b; Cheah and Fry 2015; Blau, 2017; Baur et al. 2018 etc.). However, the prediction of Bitcoin prices, mostly in asset-pricing settings, have still received extensive academic interests. Specifically, scholars have identified several factors which could determine Bitcoin prices from the empirical asset pricing point of view. These factors include the supply and demand of Bitcoin (Buchholz et al. 2012), trading volume (Balcilar et al. 2017), the spread between daily high and low prices (Baek and Elbeck, 2015), usage in trade (Kristoufek 2015), economic policy uncertainty index (Demir et al., 2018; Wang et al., 2020), market information, size, and momentum (Liu et al., 2019; Shen et al., 2020; Jia et al., 2021) etc. Most of these factors are considered as trading-based or technique indices which are closely related to the features of Bitcoin. Besides, existing studies find investors' attention or interest can also strongly forecast future Bitcoin returns (Bouoiyour and Selmi 2015; Kristoufek 2015; Ciaian et al. 2015; Liu and Tsyvinski 2021).

Investor attention, which refers to the limited attention that investors can devote to the information which might affect their investment decisions, is found to be a significant factor in determining asset prices (Seasholes and Wu, 2007; Da et al., 2011; Chen 2017; Gargano and Rossi 2018, etc.). The relevant empirical evidence is mainly obtained from the stock markets. In this context, some pioneer researchers introduce the attention-grabbing events as indirect proxies for investor attention on stocks such as the extreme returns (Barber and Odean, 2008), price limits (Seasholes and Wu, 2007), trading volumes (Gervais et al., 2001; Barber and Odean, 2008; Hou et al., 2009; Loh 2010), media coverage (Barber and Odean 2008; Yuan, 2015; Bajo et al., 2020) and advertising expenditure (Grullon et al., 2004; Chemmanur and Yan, 2009; Lou, 2014; Mayer, 2021), etc. However, there are intrinsic limitations for the indirect measures as these proxies are primarily based on critical assumptions that investors should have paid attention to these attention-grabbing events. This problem has been largely solved by the introduction of direct measures for investor attention relying on internet search queries or social media activities such as Internet search volume (Da et al., 2011; Vlastakis and Markellos 2012; Zhang et al., 2013; Zhang et al., 2021a), asset-specific stock tweets (Li et al., 2016); Wikipedia (Focke et al., 2020), etc., as these proxies capture actively expressed investor interests.

In recent years, the development of artificial intelligence partially overcomes the limitations of linear models by introducing machine learning approaches (Rather et al., 2015; Chen et al., 2017) such as the artificial neural network (ANN), support vector machine (SVM), and genetic algorithms (GA), which are used for pattern recognitions and nonlinear regressions. In this study, we adopt the long-short term memory networks (LSTM) to evaluate the effect of investor attention on Bitcoin returns. The LSTM, a variation of standard recurrent neural networks (RNN) which proposed by Hochreiter and Schmidhuber (1997), is also a special type of the ANN. The LSTM has been applied to advance the state-of-art for many challenging problems in diverse fields including language modeling and text generating, machine translation, speech recognition etc., and for its most important application, the time-series predictions. Still in the context of financial markets, existing studies use LSTM to forcast exchange rates and forex trading (Maknickienė and Maknickas, 2012; Islam and Hossain, 2020), gold price (Livieris et al., 2020), stock price (Nelson et al., 2017) and volatility (Kim and Won, 2018), etc. Whereas most studies make predictions by the LSTM only consider the basic features of the targeting assets such as the price, volume, volatility, or other trading-based variables, ignoring some of the informative factors which could directly affect the asset returns. Recently, Chen and Ge (2019) explore the attention mechanism in the LSTM to improve the forecasting accuracy of Hong Kong stock price movements. Similarly, Zhang et al. (2021b) embrace investor attention proxies from the internet containing the Baidu index's search volume and the number of news as additional inputs to the LSTM models, and find that attention proxies could improve the prediction accuracy, which therefore verify the predictability of investor attention on stock price. Thus, as discussed above, based on the empirical evidence from previous studies that investor attention could affect asset returns, and the LSTM is an effective tool in time series forecasts, we decide to embrace attention-based LSTM models to evaluate the effects of aggregate investor attention on Bitcoin returns.

The rest of the paper proceeds as follows: Section 2 introduces the data, variables and models. Section 3 discusses the main empirical results and section 4 concludes.

2. Data, Variables, and Models

2.1. Data

We extract the Bitcoin data covering 52 months from January 1, 2016 to May 1, 2020, including the daily opening price ($Opnprc$), closing price ($Clsprc$), highest price ($Hiprc$), lowest price ($Loprc$), and trading volume (Vol) of Bitcoin, which are derived from CoinMarketCap[1]. We employ these Bitcoin data as trading variables in our study.

We use Google Trends for "Bitcoin" (SVI) and Bitcoin Tweets ($Twitter$) as direct proxies for investor attention on Bitcoin. These data are obtained from BitInfoCharts[2] with the same sample period from 1 January 2016 to 1 May 2020. The main variables used in our study are defined in Table 1, where the Bitcoin data are classified as trading variables, Google trends and Bitcoin Tweets are the direct attention variables, and the calculations of traditional attention variables will be introduced in the following section.

(Please insert Table 1 here)

[1] Source of Bitcoin data: www. coinmarketcap.com

[2] Source of SVI and $Twitter$: www. bitinfocharts.com.

2.2. Traditional proxies for investor attention

Given attention is a scarce cognitive resource and individuals have limited capacity to process information (Kahneman, 1973; Pashler and Johnston, 1998), existing studies have shown that the allocation of attention, which is caused by the attention-grabbing events, would affect investors' decision making and asset pricing with empirical observations in stock markets (eg., Barber and Loeffler 1993; Seasholes and Wu 2007; Chen 2017; Mayer 2021, etc.). These attention-grabbing (or "stimuli") events have also been adopted as traditional indirect proxies to measure investor attention. In this study, we start with the selected four indirect proxies including extreme return (Koester et al., 2006; Seasholes and Wu 2007; Barber and Odean 2008; Yuan 2015), abnormal trading volume (Gervais et al., 2001; Barber and Odean, 2008), past return (Aboody et. al., 2010; Vozlyublennaia, 2014) and nearness to the 30-day high (Li and Yu, 2012).

Extreme return ($ERet$) is calculated by the daily return of Bitcoin over the average of 30-day absolute returns of Bitcoin as shown below:

$$ERet_t = \frac{Ret_t}{\frac{1}{30}\sum_{t=1}^{30} |Ret_{t-30}|} \tag{1}$$

where Ret_t is the return of Bitcoin on day t.

Abnormal trading volume ($AVol$) is the ratio of the daily trading volume to the average over the previous 30 days for Bitcoin calculated as follows:

$$AVol_t = \frac{Vol_t}{\frac{1}{30}\sum_{t=1}^{t=30} Vol_{t-30}} - 1 \tag{2}$$

where Vol_t is the daily trading volume of Bitcoin on day t.

Past return ($PRet$) is the daily cumulative return over the previous 30 for Bitcoin calculated as follows:

$$PRet_t = \left[\prod_{t=1}^{t=30} (1 + Ret_t)\right] - 1 \tag{3}$$

Nearness to the 30-day high ($30dH$) is the ratio of the current price at day t to its highest price over the previous 30 days for Bitcoin calculated as:

$$30dH_t = \frac{p_t}{\max_{t=1}^{t=30}(p_{t-30})} \tag{4}$$

where p_t is the close price of Bitcoin on day t.

2.3. Direct measures of investor attention

Comparing with indirect proxies with ex-post information, it is relatively difficult to adopt direct measures of investor attention in earlier empirical studies, however, the advances in information technology and growing popularity of social media provide more opportunities for constructing more ex-ante proxies of measuring investor attention. As searching activities could directly reflect the information demand of investors, following Da et al. (2011)'s study of introducing Google weekly search volume index (SVI) as a direct measure of investor attention to the stocks, we use daily Google Trends to Bitcoin (SVI) as one of direct attention proxies for Bitcoin (Philippas et al., 2019; Ibikunle et al., 2020; Liang et al., 2022; Aslanidis et al., 2022 etc.).

More microscopically, Li et. al (2016) use asset-specific stock tweets containing more temporal information as a direct measure of investor attention. In the context of Bitcoin markets, we adopt Bitcoin tweets ($Twitter$) (Shen et al., 2019; Choi, 2021; Li et al., 2021a, Li et al., 2021b etc.) as another additional attention proxy for Bitcoin which contains more direct and personal information.

2.4. Summary Statistics

As shown in Table 2, our sample covers 1,583 observations for all three groups of variables. There is a large difference between Min and Max of the direct attention proxy variables of Bitcoin ranging from 7.190 to 616.867 for SVI and 13,294 to 155,600 for $Twitter$, while the volatility is relatively high with a standard deviation of 51.991 for SVI and 20,487.455 for $Twitter$, which indicates a high volatility of investor attention on Bitcoin during the sample period.

(Please insert Table 2 here)

2.5. Models[3] and Experiment setup

We carry out comparative experiments with three sets of inputs for long short-term memory (LSTM) network to evaluate the predictive performance with different input variables. Firstly, we use the basic trading variables of Bitcoin including its daily opening price, highest price, lowest price, closing price and trading volume as the LSTM inputs. Secondly, in addition to the Bitcoin trading data, we include traditional attention proxies such as the extreme return, abnormal trading volume, past return and the nearness to the 30-day high into the LSTM inputs. After that, we further embrace the direct attention proxies including the daily Google Trends to Bitcoin and Bitcoin tweets as the LSTM input variables. As for parameter settings in LSTM, our main experiments with the three sets of

[3] For the detailed illustrations of the models, please refer to Zhang et al. (2021b) or ANN (Lek, 1996; Recknagel et al., 1997; Lion et al., 2000; Maier and Dandy 2000; Agatonovic-Kustrin and Beresford, 2000; Wang, 2003), RNN (Pearlmutter 1989),and LSTM (Hochreiter and Schmidhuber 1997), respectively.

inputs have 100 epochs, 10 neurons in hidden layer, the batch size of 10 and the length of windows set to be 10. All of our sample data are on a daily basis and share the same time span from January 1, 2016 to May 1, 2020, and the earlier 85% of the trading days are used as the training data set, whereas the remaining 15% are for testing. The experimental environment of LSTM is presented in Table 3.

(Please insert Table 3 here)

3. Empirical Results

In order to examine the performance of the LSTM predictions, firstly, we adopt the root mean squared error (RMSE), one of the most commonly used KPIs to compare the difference between the LSTM prediction values and the real (target) prices of Bitcoin. Formally, RMSE is defined as the square root of the average squared error calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{5}$$

where $\hat{y}_i$ is the predicted values and y_i is the real closing price of Bitcoin. The smaller the value of RMSE, the better the performance of LSTM forecasts.

The other evaluation criterion we adopted in our analysis to measure the forecast accuracy of LSTM is the mean absolute percentage error (MAPE), which is defined as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \tag{6}$$

MAPE provides the errors in terms for absolute percentages, and avoids the problem that positive and negative errors could cancel each other out. Still, the smaller the MAPE, the better the forecasts by LSTM.

(Please insert Figure 1 here)

Figure 1 illustrates the real (target) daily closing prices and the predicted values of Bitcoin using LSTM with all three types of input variables trained for 100 times. The red line represents the real (target) closing price of Bitcoin, while the blue and green lines represent the value predicted in training period and testing period respectively.

(Please insert Table 4 here)

The main results of LSTM with different groups of input variable to forecast Bitcoin prices are shown in Table 4. Here we adopt the LSTM with 10 neurons in the hidden layer, the batch size of 10 and the length of windows set to be 10. We find that adding up the traditional attention variables to Bitcoin trading variables as inputs to LSTM could improve the prediction accuracy, whereas combining all the three groups of input variables including the additional two direct attention proxies would generate even better performance of LSTM in predicting Bitcoin prices. The cost time of model running increases with the number of input variables, as the LSTM's computing time is proportional to the number of parameters. The best result is 394.058 for testing RMSE and 0.035 for testing MAPE under this parameter settings.

(Please insert Table 5 here)

(Please insert Table 6 here)

(Please insert Table 7 here)

(Please insert Table 8 here)

After that, we try to find the optimal architecture of LSTM in predicting Bitcoin prices with the combined input variables by tuning different parameters. For instance, as shown in Table 5, all else remains the same, we change the number of epochs ranging from 50 to 300, and find the LSTM forecasts perform the best with 100 training times as reported in Table 4. The computational cost also improves with the increase of training times. As for the length of windows, we find that the LSTM enables more accurate predictions of Bitcoin returns when the length of windows set at 20, with the smallest values of testing RMSE and MAPE (364.470 and 0.032 respectively). Moreover, as presented in Table 7 and Table 8, we find that the LSTM is optimal with 10 neurons in hidden layer and the batch size of 10, with the same minimum testing RMSE and MAPE values of 394.058 and 0.035 respectively as shown in Table 4. While the time cost is negative related to the change of batch size. In general, there is no linear relationship found between the change in parameter settings and LSTM prediction accuracy in our additional analysis.

By comparing the results of LSTM using different groups of inputs and parameter settings, we find that combining attention proxies and the initial trading variables in the input variables, especially adding up the two direct attention proxies (i.e. Google Trends to Bitcoin and Bitcoin tweets) would effectively improve the accuracy of the LSTM's prediction of Bitcoin prices. Moreover, these results suggest that the investor attention to Bitcoin carries more information which are valuable for analyzing and predicting Bitcoin price movements. Whereas the Google Trends to Bitcoin and Bitcoin tweets are effective direct attention proxies which can further improve the LSTM's forecasting capacity.

4. Conclusions

This paper examines the effect of aggregate investor attention on Bitcoin returns by adopting the attention proxies as inputs of the long-short term memory network (LSTM) to predict Bitcoin prices. Our empirical results indicate that including both direct and indirect attention variables in addition to the basic trading variables as the LSTM inputs could effectively improve its prediction accuracy of Bitcoin prices, which means that the aggregate investor attention on Bitcoin contains more valuable information than the historical Bitcoin data in generating Bitcoin price movements. Moreover, the two direct proxies of investor attention on Bitcoin (i.e., daily Google Trends to Bitcoin and Bitcoin tweets), measuring individual attention more directly, significantly improve the LSTM prediction accuracy with almost half of the RMSE and MAPE values comparing with the results only using trading variables as inputs. Therefore, we suggest that these two direct measures of investor attention could further enhance the prediction capacity of LSTM on Bitcoin. In most testing cases with different parameter settings, we find our LSTM model a promising architecture to forecast Bitcoin returns.

References

Aboody, D., Lehavy, R. and Trueman, B. 2010. Limited attention and the earnings announcement returns of past stock market winners. *Review of Accounting Studies*. **15** (2), pp.317–344.

Aslanidis, N., Bariviera, A. F. and López, Ó. G. 2022. The link between cryptocurrencies and Google Trends attention. *Finance Research Letters*, 102654.

Baek, C. and Elbeck, M. 2015. Bitcoins as an investment or speculative vehicle? A first look. *Applied Economics Letters*. **22**(1), pp.30-34.

Bajo, E., Chemmanur, T. J., Simonyan, K. and Tehranian, H. 2016. Underwriter networks, investor attention, and initial public offerings. *Journal of Financial Economics*. **122**(2), pp.376–408.

Balcilar, M., Bouri, E., Gupta, R. and Roubaud, D. 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*. **64**, pp.74-81.

Barber, B. and Loeffler, D. 1993. The "Dartboard" Column : Second-Hand Information and Price Pressure. *Journal of Financial and Quantitative Analysis*. **28**(2), pp.273-284.

Barber, B. M. and Odean, T. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*. **21**(2), pp.785–818.

Baur, D. G., Hong, K. and Lee, A. D. 2018. Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money*. **54**, pp.177-189.

Blau, B. M. 2017. Price dynamics and speculative trading in bitcoin. *Research in International Business and Finance*. **41**, pp.493-499.

Bouoiyour, J. and Selmi, R. 2015. What Does Bitcoin Look Like? *Annals of Economic and Finance*. **16**(2), pp.449-492.

Brauneis, A. and Mestel, R. 2018. Price discovery of cryptocurrencies: Bitcoin and beyond. *Economics Letters*. ***165***, pp.58-61.

Buchholz, M., Delaney, J., Warren, J. and Parker, J. 2012. Bits and Bets, Information, Price Volatility, and Demand for Bitcoin. *Economics*. **312**(1), pp.2-48.

Cheah, E.-T. and Fry, J. 2015. Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*. **130**, pp.32-36.

Chemmanur, T. J. and Yan, A. 2019. Advertising, Attention, and Stock Returns. *Quarterly Journal of Finance*. **9**(3), 1950009.

Chemmanur, T.J. and Yan, A., 2019. Advertising, Attention, and Stock Returns. *Quarterly Journal of Finance*. **9**(3), 1950009.

Chen, S. and Ge, L. 2019. Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction. *Quantitative Finance*. **19** (9), pp.1507–1515.

Chen, T. 2017. Investor Attention and Global Stock Returns. *Journal of Behavioral Finance*. **18**(3), pp.358–372.

Chen, W., Pourghasemi, H.R., Kornejady, A. and Zhang, N. 2017. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt,and SVM machine learning techniques. *Geoderma*. **305***,* pp.314-327.

Choi, H. 2021. Investor attention and bitcoin liquidity: Evidence from bitcoin tweets. *Finance Research Letters*. **39**, pp.1-9.

Ciaian, P., Rajcaniova, M. and Kancs, D. A. 2016. The economics of BitCoin price formation. *Applied Economics*. **48**(19), pp.1799-1815.

Corbet, S., Lucey, B., Peat, M. and Vigne, S. 2018. Bitcoin Futures—What use are they? *Economics Letters*. **172**, pp.23-27.

Da, Z., Engelberg, J. and Gao, P. 2011. In Search of Attention. *Journal of Finance*. **66**(5), pp.1461–1499.

Focke, F., Ruenzi, S. and Ungeheuer, M. 2020. Advertising, Attention, and Financial Markets. *Review of Financial Studies*. **33**(10), pp.4676-4720.

Gargano, A. and Rossi, A. G. 2018. Does it pay to pay attention? *Review of Financial Studies*. **31**(12), pp.4595–4649.

Gervais, S., Kaniel, R. and Mingelgrin, D. H. 2001. The high-volume return premium. *Journal of Finance*. **56**(3), pp.877–919.

Grullon, G., Kanatas, G. and Weston, J. P. 2004. Advertising, breadth of ownership, and liquidity. *Review of Financial Studies*. **17**(2), pp.439–461.

Hou, K., Xiong, W. and Peng, L. 2009. A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum. *Working Paper*. Ohio State University and Princeton University.

Ibikunle, G., McGroarty, F. and Rzayev, K. 2020. More heat than light: Investor attention and bitcoin price discovery. *International Review of Financial Analysis*. **69**, 101459.

Islam, M. S. and Hossain, E. 2020. Foreign exchange currency rate prediction using a GRU-LSTM hybrid network. *Soft Computing Letters*. 100009.

Jia, B., Goodell, J. W., Shen, D. (2021). Momentum or reversal: Which is the appropriate third factor for cryptocurrencies? *Finance Research Letters*, 102139.

Kahneman, D. 1973. Attention and effort. *Englewood Cliffs*. **1063**, pp.218-226.

Kim, H.Y. and Won, C.H. 2018. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*. **103**, pp.25-37.

Koester, A., Lundholm, R. and Soliman, M. 2016. Attracting attention in a limited attention world: Exploring the causes and consequences of extreme positive earnings surprises. *Management Science*. **62**(10), pp.2871–2896.

Kristoufek, L. 2015. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLoS One*. **10**(4), e0123923.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J. and Aulagnier, S. 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling*. **90**, pp.39-52.

Li, J. and Yu, J. 2012. Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics*. **104**, pp.401–419.

Li, X., Hendler, J. A. and Teall, J. L. 2016. Investor Attention on the Social Web. *Journal of Behavioral Finance*. **17**(1), pp.45–59.

Li, Y., Goodell, J. W. and Shen, D. 2021a. Comparing search-engine and social-media attentions in finance research: Evidence from cryptocurrencies. *International Review of Economics & Finance*. **75**, pp.723-746.

Li, Y., W. Goodell, J. and Shen, D. 2021b. Does happiness forecast implied volatility? Evidence from nonparametric wave-based Granger causality testing. *The Quarterly Review of Economics and Finance*. **81**, pp.113-122.

Liang, C., Zhang, Y., Li, X. and Ma, F. 2022. Which predictor is more predictive for Bitcoin volatility? And why?. *International Journal of Finance & Economics*. **27**(2), pp.1947-1961.

Liu, Y. and Tsyvinski, A. 2021. Risks and Returns of Cryptocurrency. *The Review of Financial Studies*. **34**(6), pp.2689–2727.

Liu, Y., Tsyvinski, A. and Wu, X. 2019. Common Risk Factors in Cryptocurrencies. *NBER Working Paper NO. 25882*.

Livieris, L.E., Pintelas, E. and Pintelas P. 2020. A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications*. **32***,* pp.17351–17360.

Loh, R. K. 2010. Investor Inattention and the Underreaction to Stock Recommendations. *Financial Management*. **39**(3), pp.1223–1252.

Lou, D. 2014. Attracting investor attention through advertising. *Review of Financial Studies*. **27**(6), pp.1797–1829.

Maier, H.R. and Dandy, G.C. 2000. Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environmental Modelling and Software*. **15**, pp.101–124.

Maknickienė, N. and Maknickas, A. 2012. Application of Neural Network for Forecasting of Exchange Rates and Forex Trading. *The 7th International Scientific Conference "Business and Management*. pp.10-11.

Mayer, E. J. 2021. Advertising, investor attention, and stock prices: Evidence from a natural experiment. *Financial Management*. **50**(1), pp.281-314.

Nelson, D.M., Pereira, A.C. and De Oliveira, R.A. 2017. Stock market's price movement prediction with LSTM neural networks. *2017 International Joint Conference on Neural Networks* (IJCNN). pp.1419-1426.

Pashler, H. and Johnston, J.C. 1998. Attentional Limitations in Dual-task Performance. *Attention*. pp.155-189.

Philippas, D., Rjiba, H., Guesmi, K. and Goutte, S. 2019. Media attention and Bitcoin prices. *Finance Research Letters*. **30**, pp.37-43.

Rather, A.M., Agarwal, A. and Sastry, V.N. 2015. Recurrent neural network and a hybrid model for prediction of stock returns. *Expert Systems with Applications*. **42** (6), pp.3234–3241.

Recknagel, F., French, M., Harkonen, P. and Yabunaka, K.I. 1997. Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling*. **96**, pp.11–28.

Seasholes, M.S. and Wu, G. 2007. Predictable behavior, profits, and attention. *Journal of Empirical Finance*. **14**, pp.590–610.

Shen, D., Urguhart, A. and Wang, P. 2019. Does twitter predict Bitcoin? *Economics Letters*. **174,** pp.118-122.

Shen, D., Urquhart, A., Wang, P. (2020). A three-factor pricing model for cryptocurrencies. *Finance Research Letters*, **34**, 101248.

Vlastakis, N. and Markellos, R.N. 2012. Information demand and stock market volatility. *Journal of Banking and Finance*. **36**, pp.1808–1821.

Vozlyublennaia, N. 2014. Investor attention, index performance, and return predictability. *Journal of Banking and Finance*. **41**, pp.17–35.

Wang, P., Li, X., Shen, D., Zhang, W. (2020). How does economic policy uncertainty affect the bitcoin market? *Research in International Business and Finance*, **53**, 101234.

Wang, S.C. 2003. Artificial Neural Network. *Interdisciplinary Computing in Java Programming*. **743**, pp.81-100.

Yermack, D. 2015. Is Bitcoin a real currency? An economic appraisal. *Handbook of Digital Currency*. pp.31-43.

Yuan, Y. 2015. Market-wide attention, trading, and stock returns. *Journal of Financial Economics*. **116**(3), pp.548–564.

Zhang, W., Shen, D., Zhang, Y., & Xiong, X. (2013). Open source information, investor attention, and asset pricing. *Economic Modelling*, **33**, 613-619.

Zhang, W., Yan, K., & Shen, D. (2021a). Can the Baidu Index predict realized volatility in the Chinese stock market? *Financial Innovation*, **7**(7), 1-31.

Zhang, Y., Chu, G. and Shen, D. 2021b. The role of investor attention in predicting stock prices: The long short-term memory networks perspective. *Finance Research Letters*. **38**, 101484.

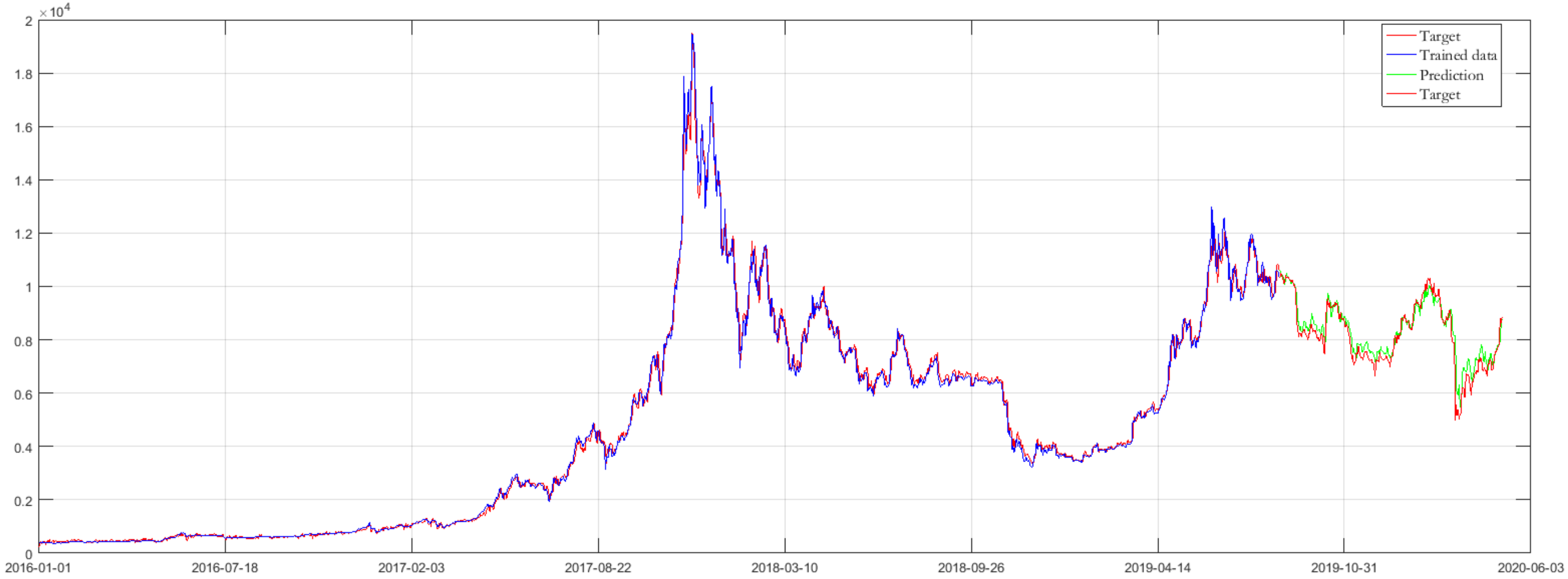

Figure. 1. Real (Target) and predicted values of Bitcoin using trading/traditional attention/direct attention proxies with 100 epochs

Table 1.
Variable definitions
This table reports the three groups of main variables used as inputs of the long short-term memory network (LSTM), which includes the trading variables, traditional attention proxies and the direct attention proxies.

Variables	Definitions
Trading variables	
$Opnprc$	The daily opening price of Bitcoin
$Hiprc$	The daily highest price of Bitcoin
$Loprc$	The daily lowest price of Bitcoin
$Clsprc$	The daily closing price of Bitcoin
Vol	The daily trading volume of Bitcoin
Traditional attention variables	
$ERet$	The extreme return of Bitcoin
$Avol$	The abnormal trading volume of Bitcoin
$PRet$	The past return of Bitcoin
$30dH$	The nearness to the 30-day high of Bitcoin
Direct attention variables	
SVI	The daily Google search trend volume, which normalized by the value on January 1, 2012.
$Twitter$	The number of Twitter posts of Bitcoin

Table 2. Summary Statistics

This table reports the descriptive statistics including the number of observations (Obs.), mean, median, standard deviation (S.D), maximum (Max), and minimum (Min) for each of the main variables. Panel A reports the descriptive statistics of five trading variables: daily opening price ($Opnprc$), highest price ($Hiprc$), lowest price ($Loprc$), closing price ($Clsprc$), and trading volume (Vol) of Bitcoin. Panel B reports the descriptive statistics of the four traditional attention variables: extreme return ($ERet$), abnormal volume($AVol$), past return ($PRet$), nearness to the 30-day high ($30dH$). Panel C reports the descriptive statistics of the two direct attention proxies: Google search trend volume (SVI) and Bitcoin tweets ($Twitter$). All variables are winsorized at 1% level.

	Obs.	Mean	Median	S.D	Max	Min
Panel A: Trading Variables						
$Opnprc$	1583	5118.88	5061.20	3902.33	19475.80	365.07
$Hiprc$	1583	5256.33	5235.19	4032.40	20089.00	374.95
$Loprc$	1583	4971.53	4919.49	3746.92	18974.10	354.91
$Clsprc$	1583	5124.06	5064.49	3901.70	19497.40	364.33
Vol	1583	8672082055	4273640000	11373334437	74156772075	28514000
Panel B: Traditional Attention Variables						
$ERet$	1583	0.107	0.118	1.570	12.039	-12.868
$Avol$	1583	0.077	-0.011	0.408	2.497	-0.691
$PRet$	1583	0.063	0.034	0.276	1.563	-0.635
$30dH$	1583	0.896	0.933	0.108	1.000	0.397
Panel C: Direct Attention Variables						
SVI	1583	45.136	35.677	51.991	616.867	7.190
$Twitter$	1583	33652.292	25882	20487.455	155600	13294

Table 3. System environment

This table reports the experimental environment of Long-short term memory network (LSTM), and the experiment results in this paper are all based on this hardware and software environment.

CPU	Intel® Core™ i5-6500 CPU @3.2GHz 3.2GHz
RAM	48.0G
GPU	NVIDIA Quadro K620
System	Windows 10.0
Python version	Python 3.7.2
Keras version	Keras 2.0.8
Sklearn version	Sklearn 0.4.0
Tensorflow version	Tensorflow 1.2.1

Table 4. Comparative results using different input variables

This table reports the results of Long-short term memory network (LSTM), including the values of running time, RMSE and MAPE for training and testing sets with 100 training times.

Inputs	No. of Epochs	Number of neurons in hidden layer	Batch size	Length of windows	Time Cost	Training RMSE	Training MAPE	Testing RMSE	Testing MAPE
Trading	100	10	10	10	96.316	473.180	0.089	777.295	0.082
Trading/Traditional attention	100	10	10	10	143.656	489.989	0.120	719.614	0.076
Trading/Traditional attention/Direct attention	100	10	10	10	176.717	327.527	0.094	394.058	0.035

Table 5. Comparative results using different number of Epochs

This table reports the results of Long-short term memory network (LSTM), including the values of running time, RMSE and MAPE for training and testing sets, with different number of Epochs ranging from 50 to 300.

Inputs	No. of Epochs	Number of neurons in hidden layer	Batch size	Length of windows	Time Cost	Training RMSE	Training MAPE	Testing RMSE	Testing MAPE
Trading/Traditional attention/Direct attention	50	10	10	10	86.539	326.818	0.056	432.300	0.039
Trading/Traditional attention/Direct attention	100	10	10	10	176.717	327.527	0.094	394.058	0.035
Trading/Traditional attention/Direct attention	150	10	10	10	253.126	363.102	0.057	789.248	0.088
Trading/Traditional attention/Direct attention	200	10	10	10	343.602	267.187	0.045	651.084	0.065
Trading/Traditional attention/Direct attention	250	10	10	10	403.097	243.868	0.064	647.876	0.063
Trading/Traditional attention/Direct attention	300	10	10	10	558.894	229.890	0.060	622.903	0.060

Table 6. Comparative results using different length of windows

This table reports the results of Long-short term memory network (LSTM), including the values of running time, RMSE and MAPE for training and testing sets, with different length of windows ranging from 5 to 30.

Inputs	No. of Epochs	Number of neurons in hidden layer	Batch size	Length of windows	Time Cost	Training RMSE	Training MAPE	Testing RMSE	Testing MAPE
Trading/Traditional attention/Direct attention	100	10	10	5	178.751	327.256	0.064	417.091	0.036
Trading/Traditional attention/Direct attention	100	10	10	10	176.717	327.527	0.094	394.058	0.035
Trading/Traditional attention/Direct attention	100	10	10	15	175.327	347.525	0.056	383.687	0.034
Trading/Traditional attention/Direct attention	100	10	10	20	180.885	274.124	0.046	364.470	0.032
Trading/Traditional attention/Direct attention	100	10	10	25	179.301	250.067	0.046	530.759	0.051
Trading/Traditional attention/Direct attention	100	10	10	30	185.066	281.769	0.050	442.339	0.043

Table 7. Comparative results using different number of neurons in hidden layer

This table reports the results of Long-short term memory network (LSTM), including the values of running time, RMSE and MAPE for training and testing sets, with different number of neurons in hidden layer ranging from 5 to 30.

Inputs	No. of Epochs	Number of neurons in hidden layer	Batch size	Length of windows	Time Cost	Training RMSE	Training MAPE	Testing RMSE	Testing MAPE
Trading/Traditional attention/Direct attention	100	5	10	10	167.581	314.683	0.120	600.150	0.060
Trading/Traditional attention/Direct attention	100	10	10	10	176.717	327.527	0.094	394.058	0.035
Trading/Traditional attention/Direct attention	100	15	10	10	184.967	306.649	0.050	649.053	0.065
Trading/Traditional attention/Direct attention	100	20	10	10	184.777	303.985	0.048	457.952	0.042
Trading/Traditional attention/Direct attention	100	25	10	10	179.972	326.915	0.081	442.733	0.039
Trading/Traditional attention/Direct attention	100	30	10	10	194.061	331.604	0.114	659.968	0.071

Table 8. Comparative results using different batch size

This table reports the results of Long-short term memory network (LSTM), including the values of running time, RMSE and MAPE for training and testing sets, with different batch size ranging from 5 to 30.

Inputs	No. of Epochs	Number of neurons in hidden layer	Batch size	Length of windows	Time Cost	Training RMSE	Training MAPE	Testing RMSE	Testing MAPE
Trading/Traditional attention/Direct attention	100	10	5	10	322.484	330.590	0.077	426.314	0.038
Trading/Traditional attention/Direct attention	100	10	10	10	176.717	327.527	0.094	394.058	0.035
Trading/Traditional attention/Direct attention	100	10	15	10	137.521	327.854	0.108	421.700	0.039
Trading/Traditional attention/Direct attention	100	10	20	10	108.772	327.001	0.099	442.407	0.042
Trading/Traditional attention/Direct attention	100	10	25	10	87.034	327.811	0.055	417.113	0.038
Trading/Traditional attention/Direct attention	100	10	30	10	74.441	335.970	0.075	436.551	0.040