The dependence structure and co-movement of Cryptocurrency based Bayesian approach

Anuphak Saosaovaphak Chukiat Chaiboonsri and Satawat Wannapan
Faculty of Economics, Chiang Mai University, Thailand
anuphak@gmail.com; chukiat1973@gmail.com; lionz1988@gmail.com

Abstract. Cryptocurrencies are unique and extra-ordinary currencies which to be econometrically forced into the linear model due to their systematic complexity and extreme movements. This paper was conducted to provide an alternative analysis as a solution for escaping the restrictions of traditional linear assumptions. Five predominant digital currencies such as Bitcoin (BTC), Stellar network (XLM), Litecoin (LTC), Ethereum Classic (ETC), and IOTA were chosen to be employed in the multiple processes based on Bayesian approaches. Market dominance and data regime classifications are the essential components that lead to successfully investigate the dependent structures and co-movements in the digital financial market. The empirical findings could assume that the modern time-series data was meticulously estimated by the flexible modern tool. Bayesian statistics and simulations have the sufficient potency as the suitable solution.

1. Introduction
The question is; why the digital currencies were chosen in this paper? Since the attempt to search for perfect competitive market with independent demands and supplies from presiding hands thus, cryptocurrencies seem to potentially make this concept becomes realistic [1]. This marketplace presented how digital currencies can efficiently connect buyers and sellers without any interfering authorities, and they could be continuously succeed [2], [3]. Inevitably, this transformation would be the highlight issue which considerably concerned in the world of finance.

Literarily, Bitcoin depends on the block-chain technology to account its transactions in a public ledger which is a distributed database that maintains a continuous growing data of records called “block sets” secured from tampering and revision. Also, the Stellar network (XLM) is based on block-chain technology. It is the decentralized network of servers which drives distributed ledgers. This communication process is referred as “consensus” [4]. Litecoin (LTC) is one of the most established networks in the crypto-space. In 2011, this kind of digital currency was created by an ex-Google and ex-Coinbase engineering to be an alternative for Bitcoin [5]. Ethereum Classic (ETC) is also purposed to provide the decentralized framework that drives decentralized applications which are exactly programed without any possibility of downtimes, censorships, frauds or third party interferences [6]. Differently, IOTA is unique currency from the other cryptocurrencies which relied on a non-blockchain data structure with a highly scalable approach for transactional confirmation that is not only mentioned as the digital money. Communication between machines could be defined as a protocol to have a part in the Internet of Things (IOT) [7]. From various studies in literature although the decentralized concept reasonably centralizes, most of digital currencies are under control which...
still unique and elusive to capture their behaviors. The linear relationship seems to be useless. It is possible to imply that it would be unnecessary to further search for their correlations. However, this issue became the most interesting point that led the authors provide some different conclusions.

To answer the question; why the linear does not work for this digital money data?, outliers are the key. Therefore, Fig.1 displays the linear assumption modeling in the original data type of the collected digital currencies. It looks suspicious if the conclusion is indicated that the line is suitably fitted for the data points. The doubtfulness is stronger as soon as the linear line was employed to explain the cumulative distribution details as represented in the Fig.2. Therefore, the line would be potentially useless. Moreover, in Fig.3, the graphical details presented the existence of outliers, which were indicated outside the interquartile range (IQR). Minimum and maximum outliers appeared far from the interval between Q1 and Q3. The observed digital currencies were implied that their relations seemed to be extreme and were supported by the scatter details as shown in Fig.4. In each digital money type, the black scatter points were revealed outside the interquartile box (the blue-line boxes). Consequently, this contribution was conducted to apply alternative econometric methods such as “non-linear data classification” and “structural dependence”. Additionally, the function of the alternatively statistical approach called “Bayesian statistics” was employed to avoid traditionally statistical assumptions for reliably seeking the deeply structural correlations and co-movements of cryptocurrencies, which was the novel solution of the paper.

2. Methodology and contributions

2.1 Descriptive information

In terms of general information, Table 1 showed five digital currencies chosen in this paper were sampled as a daily time-series data started from 13th June 2017 to 5th August 2019. 784 observations were transformed to the natural-log form which implied the collected variables were stationary. The major differences between maximum and minimum values implied data classifications were crucially
needed. The outcomes of the Jarque-Bera test represented the strong probabilistic evidence which implied the relation is not a normal distribution among five digital currencies.

Table.1: Descriptive information for five collected cryptocurrencies

|       | XLM          | LTC          | IOTA         | ETC          | BTC          |
|-------|--------------|--------------|--------------|--------------|--------------|
| Max   | 0.666779     | 0.389338     | 0.383959     | 0.303521     | 0.225119     |
| Min   | -0.328337    | -0.395151    | -0.377046    | -0.435318    | -0.207530    |
| Jarque-Bera | 3038.805 | 1503.458    | 517.7475     | 1270.930     | 338.6739     |
| Prob  | 0.000000     | 0.000000     | 0.000000     | 0.000000     | 0.000000     |
| Obs   | 784          | 784          | 784          | 784          | 784          |

2.2 Distance correlation coefficient by Spearman’s rho based Bayesian approach

Fundamentally, the major theoretical framework is the Bayes’ theorem, we obtain the posterior density $\pi(\theta|y)$ that refers to a probability distribution of parameters $\theta$ as the parameters, and the complete-data likelihood function expressed in Equation (1) [8],

$$L(y|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left( -\frac{y_i^2}{2\sigma_i^2} \right).$$

where $y$ refers to the time-series data of n observations, $y = (y_1, y_2, ... y_n)$ and $\theta$ stands for the estimated parameters. According to [9] and [10], Bayesian statistics considers hypotheses regarding multiple parameters by adapting Bayes factors or Bayes computational comparisons.

First of all, this theorem is used to modify the Spearman’s rho calculation. Adapting from [11] and employing Markov Chain Monte Carlo (MCMC) simulations, the Bayesian approach is suitable tool for calculating distance correlation coefficient among variables. Expressly, the measure of distance evidence is the probability of $P(H_0 / x)$. $H_0$ is true with $X = x$, which is

$$P(H_0 / x) = P(\theta \leq 0 / x) = \int_{-\infty}^{0} f(x - \theta)\pi(\theta)\,d\theta.$$

The relation with distance covariance, $P(\theta \leq 0 / x)$, was examined using a Bayesian approach. In this section, a meta-analysis can be conducted to estimate the real correlation and determinants among digital currencies. The results of distance correlations for the market determinant investigation were detailed in Table.2.

Table.2: The comparison between parametric and non-parametric correlations for the market determinant investigation

| Parametric correlation (linear correlation) | Non-parametric correlation (non-linear correlation) |
|--------------------------------------------|-----------------------------------------------|
| Lower Pearson correlation                   | Lower Spearman’s rho                           |
| Upper Pearson correlation                   | Upper Spearman’s rho                           |
| IOTA                                       | 0.516                                        | 0.582                                        | 0.659                                      | 0.690                                      |
| BTC                                        | 0.513                                        | 0.671                                        | 0.575                                      | 0.684                                      |
| ETC                                        | 0.541***                                     | 0.694**                                     | 0.599**                                     | 0.702**                                    |
| XLM                                        | 0.455                                        | 0.635                                        | 0.580                                      | 0.683                                      |
| LTC                                        | 0.535**                                     | 0.685**                                     | 0.638***                                    | 0.734***                                   |
| Sampling method                            | Simple Bootstrap Specifications               | Simple Bootstrap Specifications               |
| Number of samples                          | 1000                                         | 1000                                         |

Note: *** implies the number one which is the most influenced factor. ** implies the number two which is the influenced factor.

Empirically, classical statistics, two predominant methods such as the Pearson correlation and spearman’s rho calculation were used to seek the market leader among five collected currencies based on the simple bootstrap specification. However, the details shown in Table.2 were the controversy results. In other words, the parametric correlation presented that the Ethereum Classic (ETC) was the
market leader while the non-parametric correlation presented that the Litecoin (LTC) was the market determinant. This contrast result was the problem to provide restricted points of these traditional methods.

The solution of the market determinant investigation was found by employing Bayesian inference and simulations to Spearman’s rho calculation. Since the outcomes of posteriors were random parametric estimations, correlation parameters were presented as interval levels; 25%, 50%, and 75%, respectively. LTC and ETC were the predominant variables which were paired with others. The conclusion was stated by the total means represented in Table 3. ETC was chosen to be the market leader since its total average correlation overcame the counterpart.

| Pairs          | LTC is the market leader. | 25% | 50% | 75% | Average |
|----------------|---------------------------|-----|-----|-----|---------|
| 1              | LTC & IOTA                | 0.694| 0.738| 0.822| 0.751   |
| 2              | LTC & BTC                 | 0.650| 0.687| 0.745| 0.694   |
| 3              | LTC & ETC                 | 0.645| 0.682| 0.732| 0.687   |
| 4              | LTC & XLM                 | 0.725| 0.753| 0.799| 0.760   |
|                | **Total**                 | **0.723** |     |     |         |

| ETC is the market leader. | 25% | 50% | 75% | Average |
|----------------------------|-----|-----|-----|---------|
| 1  ETC & IOTA              | 0.752| 0.798| 0.826| 0.792   |
| 2  ETC & BTC               | 0.751| 0.818| 0.885| 0.819   |
| 3  ETC & ETC               | 0.712| 0.781| 0.821| 0.771   |
| 4  ETC & XLM               | 0.642| 0.688| 0.717| 0.683   |
| **Total**                  | **0.766** |     |     |         |

2.3 Data classifications by Naive Bayes Classifier (NB classifiers)

Another crucial part is the Bayesian application in data classification. Like the characteristic trends displayed in Fig.5, Naive Bayes Classifier (NB classifiers) is the flexible tool for computing the conditional a-posterior probabilities of a categorical class variable given independent predictor variables relied on the Bayes’ rule. In literature, NB classifiers were mentioned by [12] and applied to classify cross-sectional samples from decision makers in the domain of stock market exchange [13]. However, this paper is different since the classifiers are based on termed prior probability that reflects the most probable guess on the outcome without additional evidence [14]. Moreover, the NB classification employed in this research is to fix and avoid disadvantages of the linear model from previous works such as [15],[16], and [17]. Expressly, it is started with the assumption in the NB classifier that features $X_1, X_2, \ldots, X_n$ are conditionally independent of each other given the class is

$$P(C | X) = \frac{P(C) \prod_{i=1}^{n} P(X_i | C)}{P(X)}.$$  \hspace{1cm} (3)

In classification problems, Equation (3) is satisfactory to predict the most probable class given a test observation. The variables $W_i, i = 1,2,\ldots,n$, is introduced for the probability $P(1)$ which indicates a high regime, meaning that $\rho$, and since $P(1) + P(-1) = 1$. This provides $1-W_i$ for the probability $P(-1)$ which implies a low regime [18]. The model can be written as the equation 4, and the result of divided regimes was represented in Table.4,

$$\rho(W_i; C) = \begin{cases} W_i & \text{if } C = 1 \\ 1-W_i & \text{if } C = -1 \end{cases}$$  \hspace{1cm} (4)
Figure 5. The daily trends of five collected cryptocurrencies during 2017 to 2019

Table 4. Data classifications by Naïve Bayesian classifiers

| Regimes      | Number of classified dates |
|--------------|---------------------------|
| High regimes | 422 days                  |
| Low regimes  | 362 days                  |

2.4 Bayesian inference under known copula families

The development of the Bayesian inference scheme for the case where the parametric copula families are known is employed to structurally investigate the marginal relations of factors of cryptocurrencies. The Metropolis-Hastings within-Gibbs algorithm for the 1-factor copula model is firstly depended on setting initial values for random parameters, \( \theta^0 \), \( \tau^0 \), and \( \phi^0 \). The random stages of initial parameters are equivalent to sampling from the three blocks due to the conditional independence structure that rises from the factor copula model. The conditional posterior densities for each of the parameters are given from the joint posterior density by [19].

\[
p(f | \theta, \tau, \gamma) \propto \prod_{j=1}^{m} p(f_i | \theta, \tau, \gamma) \prod_{i=1}^{m} c_j \left( \phi_i \left( y_i \right) \middle| f_i, \tau_j \right) p_\phi \left( y_i \right)
\]

\[
\times \prod_{j=1}^{m} \left[ C_j \left( \phi_j \left( y_j \right) \right) \right] p(f_i, \tau_j - C_j \left( \phi_j \left( y_j - 1 \right) \right) \middle| p(f_i; \tau_j) \right] \times 1(0 < f_i < 1) \tag{5}
\]

Table 5: The computation of the conditional a-posterior probabilities of a categorical class variable given independent predictor variables using the Bayes rule.

| Pairs      | ETC is the market leader in normal cases. | 25%    | 50%    | 75%    | Average |
|------------|----------------------------------------|--------|--------|--------|---------|
| (1/A)      | ETC & LTC                              | 1.5125 | 2.000  | 2.4875 | 2.00    |
| (1/B)      | ETC & IOTA                             | 0.775  | 1.500  | 2.225  | 1.50    |
| (1/C)      | ETC & BTC                              | 0.650  | 1.250  | 1.850  | 1.25    |
| (1/D)      | ETC & XLM                              | 1.2625 | 1.500  | 1.7375 | 1.50    |

**Copula type**

Clayton copula

| Pairs      | ETC is the market leader in bull periods. | 25%    | 50%    | 75%    | Average |
|------------|----------------------------------------|--------|--------|--------|---------|
| (2/E)      | ETC & LTC                              | 0.2625 | 0.500  | 0.7375 | 0.50    |
| (2/F)      | ETC & IOTA                             | 0.5125 | 1.000  | 1.4875 | 1.00    |
| (2/G)      | ETC & BTC                              | 0.3875 | 0.750  | 1.1125 | 0.75    |
| (2/H)      | ETC & XLM                              | 0.2375 | 0.450  | 0.6625 | 0.45    |

**Copula type**

Normal copula

| Pairs      | ETC is the market leader in bear periods. | 25%    | 50%    | 75%    | Average |
|------------|----------------------------------------|--------|--------|--------|---------|
| (3/I)      | ETC & LTC                              | 1.5125 | 2.000  | 2.4875 | 2.00    |
| (3/J)      | ETC & IOTA                             | 1.3875 | 1.750  | 2.1125 | 1.75    |
| (3/K)      | ETC & BTC                              | 1.130  | 1.250  | 1.3700 | 1.25    |
| (3/L)      | ETC & XLM                              | 1.7625 | 2.000  | 2.2375 | 2.00    |

**Copula type**

Clayton copula
To consider the findings represented in Table.5, the first cases; Archimedean copula and Clayton copula employed to explain normal data trends without the classification. The Clayton copula model was chosen to provide an initially structural parameter for setting the prior and give a co-movement tree. The empirical result of the dependent structure and tree were displayed in Fig.6. ECT was the market leader which surrounded by four digital currencies with positive copula parameters. This meant the dependent structure of the digital money market was a parallel movement. ETC and LTC were the strongest dependent structures (Pair 1A) among 4 connections of Clayton copula estimations.

The second case; Elliptical copula and Normal copula used to explain high (bull) regimes. The reason behind choosing Normal copula was the joint distribution of those cryptocurrencies symmetrically depended on its distribution shape. The joint details were graphically shown in Fig.7. The empirical result of the dependent structure and co-movement tree were displayed in the left-hand-side picture in Fig.7. When the cryptocurrency market is in the expansion periods, the co-movement trend is still a positive parallel motion, but there are some quite weakly dependent structures among five selected currencies. ETC and IOTA were predominant to be chosen for forecasting (Pair 2F).

The last case; Archimedean copula and Clayton copula applied to describe low (bear) regimes in collected digital currencies. The reason behind choosing the Clayton copula is the joint distribution of those cryptocurrencies asymmetrically depended on its distribution shape. Interestingly, the empirical result of the dependent structure and co-movement tree displayed in Fig.7 (right-hand-side) indicates strongly dependent structures, when those of cryptocurrencies face downsizing periods. In particular, two pairs (Pair 3I and 3L) are ETC connecting LTC and ETC related to XLM are the strongest structures.

3. Conclusion
The novelty of finding true structural dependencies inside cryptocurrencies’ co-movements was accomplished in this paper. Bayesian approaches were the powerful tool that applied to avoid the restrictions of linear assumptions. Inside the relationship between the selected digital currencies, it was necessary to clarify the predominant determiner among the set of data. ETC was computationally chosen through Bayesian Spearman’s rho approach. Next, the Naïve Bayesian classifier was the key for the regime categorization providing. In terms of three scenarios; bull situations, downsizing periods, and original trends were classified. This section was the evidence to confirm that only linear assumption was not technically and practically able to cover the data which extremely difficult to clarify the true distribution. Finally, the elusive question to deeply seek for the correlations among unique digital currencies was solved through the Bayesian approach in copula models. In each
scenario, the empirical result showed that the Clayton copula was fitted for the correlated distributions in the cases of normal trends and recessions. In bull periods, the Normal copula model was appropriate. In conclusion, these findings approved modern time-series data especially the digital money information could not be investigated, estimated, and predicted in any restriction and forcing actions by traditional linear models.

References
[1] Kasiyanto S. 2016 Bitcoin's potential for going mainstream. Journal Of Payments Strategy & Systems 10 pp. 28-39.
[2] Salman A, and Razzag M G A. 2017 Bitcoin and the world of digital currencies. IntechOpen Publisher. DOI: 10.5772/intechopen.71294.
[3] DeVries P D. 2016 An analysis of cryptocurrency, bitcoin, and the future. International Journal of Business Management and Commerce 1 pp. 1-9.
[4] AVA Trade Ltd. 2019. What is Stellar?. [2019-08-22]. Retrived from: https://www.avatrade.com/forex/cryptocurrencies/stellar.
[5] Russell J. 2017 Litecoin founder Charlie Lee has sold all of his LTC. [2019-08-22]. Retrived from: https://techcrunch.com/2017/12/20/litecoin-charlie-lee-conflict-of-interest/.
[6] Back M. 2017 Into the ether with Ethereum Classic: the store-of-value commodity to power the Internet of thing. Grayscale Investment, LLC publisher. Retrived from: https://ethereumclassic.github.io/assets/etc-thesis.pdf.
[7] Divya M, and Biradar N B. 2018 IOTA-next generation block chain. International Journal Of Engineering And Computer Science 7 pp. 23823-23826.
[8] Moral-Benito E. 2007 Determinants of economic growth: a Bayesian panel data approach. CEMFI Working Paper No. 0719. CEMFI. Casado del Alisal 5; 28014 Madrid.
[9] Takaishi T 2010 Bayesian inference with an adaptive proposal density for GARCH models Journal of Physics: Conference Series 221.
[10] Kass R E and Raftery A E 1995 Bayes factors Journal of the American Statistical Association 90 pp. 773-795.
[11] Bhattacharjee A. 2014 Distance correlation coefficient: an application with Bayesian approach in clinical data analysis. 13 pp. 354-366.
[12] Abdulsattar G, Alkubaisi J, Kamaruddin S S, and Husin H. 2018 Conceptual framework for stock market classification model using sentiment analysis on twitter based on Hybrid Naïve Bayes classifiers. International Journal of Engineering & Technology 11 pp. 52-64.
[13] Zhang Z. 2016 Naive Bayes classification in R. Annals of Translational Medicine 4 pp. 1-10.
[14] Wannapan S, Chaiboonsri C, and Sriroonchitta S. 2018 Application of the Bayesian dsge model to the international tourism sector: Evidence from Thailand's economic cycle. WIT Transactions on Ecology and the Environment 226 pp. 257-268.
[15] Chaitip P, and Chaiboonsri C. 2016 Dependence modelling of Malaysian Ringgit (MYR) and Thai Baht (THB): the Markov switching model with dynamic copula approach (DCA) and bivariate extreme value approach. International Journal of Computational Economics and Econometrics 6 pp. 138-155.
[16] Moreira R, Chaiboonsri C, and Chaitip P. 2014 Analysing monetary policy's transmission mechanisms through effective and expected interest rates: an application of MS-models, Bayesian VAR and cointegration approaches for Brazil. International Journal of Monetary Economics and Finance 7 pp. 1-12.
[17] Taheri S, and Mammadov M. 2013 Learning the Naïve Bayes classifier with optimization models. Int. J. Appl. Math. Comput. Sci. 23 pp. 787-795.
[18] Tan B K, Panagiotelis A, and Athanasopoulos G. 2017 Bayesian inference for a 1-factor copula copula. working paper 06/2017.Department of Econometrics and Business Statistics, Monash Business School.