Crypto price discovery through correlation networks

Crypto price discovery through correlation networks *

Paolo Giudici † Gloria Polinesi

Abstract

We aim to understand the dynamics of crypto asset prices and, specifically, how price information is transmitted among different bitcoin market exchanges, and between bitcoin markets and traditional ones.

To this aim, we hierarchically cluster bitcoin prices from different exchanges, as well as classic assets, by enriching the correlation based Minimum Spanning Tree method with a preliminary filtering method based on the Random Matrix approach.

Our main empirical findings are that: i) bitcoin exchange prices are positively related with each other and, among them, the largest exchanges, such as Bitstamp, drive the prices; ii) bitcoin exchange prices are not affected by classic asset prices, but their volatilities are, with a negative and lagged effect.

Key words: Bitcoin exchanges, Bitcoin price discovery, Correlation networks, Minimum Spanning Trees, Random matrix theory.

*Department of Economics, University of Pavia and Department of Economic and Social Sciences, Universitá Politecnica delle Marche

†Corresponding author: giudici@unipv.it; Via S. Felice 5, 27100 Pavia (PV) - Italy

1 Introduction

The research literature on crypto currencies is relatively new, but is constantly growing. After the seminal technical paper of Nakamoto (2008), the paper by Dwyer (2015) examines the economics and financial properties of cryptocurrencies, and the paper by Corbet et al. (2018b) provides a systematic review of the literature that has been developed after 2008 on cryptocurrencies as financial assets.

Within such literature, a relevant stream of research concerns the study of the dynamics of cryptocurrency market prices, either from an endogenous viewpoint or in relationship with other "classic" market prices. While some papers investigate this issue from a univariate statistical approach, focusing on bitcoin prices, very few consider a multivariate statistics viewpoint, which deals with the interconnectedness among crypto prices and between crypto prices and classic prices.

A noticeable exception is the paper by Corbet et al. (2018a), who analyses the relationships among alternative cryptocurrencies: Bitcoin, Litecoin and Ripple, and show that they are strongly interconnected, demonstrating similar patterns of returns and volatility. A related paper is Ciaian et al. (2018) who analyse the relationship between the bitcoin and sixteen alternative coin prices, and found that they are indeed interdependent, but independent from exogenous factors. Corbet et al. (2018a) also analyse interdependence between crypto prices and a variety of other financial assets such as gold, bonds and stocks. They found that the volatility of cryptoassets is substantially higher than that of traditional assets, and that cryptocurrencies are rather isolated from other assets, thus showing a diversification benefit. Dyhrberg (2016) and Bouri et al. (2017) reach similar conclusions, thus confirming that cryptocurrencies are rather isolated from classical assets. Note however that the same authors conclude that such isolation emerges in the short run, but not in the long run and, thus, the evidence on the "diversification benefit" is not conclusive.

Understanding price interconnectedness is important not only to describe

relationships between different asset prices, but also to understand whether prices in different markets quickly react to each other or, in other words, whether markets are efficient. The paper by Brandvold et al. (2015) is the first one that addresses this question, studying the price discovery process in bitcoin markets, by means of the econometric methodologies of Hasbrouck (1995) and Gonzalo & Granger (1995). Using data from seven exchanges, in the period from April 2013 to February 2014, they find that Mtgox (bankrupting shortly after the sampled period) and BTC-e are the price setters. Pagnottoni *et al.* (2018) extends their analysis to the period January 2014 to March 2017, and found an increased role of Chinese exchanges. A related work is the paper of Urquhart (2016) who specifically analyzes whether bitcoin markets are efficient, using price return data from August 2010 through July 2016: they cannot confirm the efficient market hypothesis. However, another study (Nadarajah & Chu, 2017) reveals that a power transformation of bitcoin returns can be concluded as "weakly efficient" and, thus, the evidence on bitcoin market efficiency is not conclusive.

Our contribution is to develop a novel multivariate statistical model to study cryptocurrency price dynamics, aimed at acquiring further empirical evidence on whether bitcoin prices from different exchanges are strongly interrelated, as in an integrated and efficient market, following the paper by Brandvold et al. (2015); but also whether such interactions are affected by "exogenous" prices of classical assets, as in the paper of Corbet et al. (2018a). In other words, we aim to answer, with the same multivariate statistical model, the question of whether the bitcoin, whose capitalisation is now substantial, is still an investment diversifier and the question of whether the markets where bitcoin are traded are efficiently integrated.

Besides shedding more light on the diversification and efficiency property of bitcoin prices, we extend Corbet et al. (2018a), Brandvold *et al.* (2015), and the related papers, by modelling price interconnectedness with correlation network models, as in the recent paper of Giudici and Abu-Hashish (2018). However, differently from the previous authors, instead of insert-

ing correlation networks into a Vector Autoregressive model, which requires strong distributional assumptions, we follow a non parametric clustering model, based on the minimum spanning tree (MST) approach proposed by Mantegna (1999). The MST approach will be extended with a preliminary random matrix filtering, that improves its interpretability.

The paper is organized as follows: Section 2 contains our proposed model; Section 3 presents the available data; Section 4 the empirical application of the proposed model to the data; Section 5 contains some concluding remarks.

2 Proposal

In this section we present our methodological contribution: a clustering method for market prices, based on the minimum spanning tree approach proposed by Mantegna, empowered by the random matrix approach.

Mantegna (1999) proposed the Minimum Spanning Tree (MST) to detect the hierarchical organization of stock prices in financial markets, using their correlation matrix. Spelta and Arayuo (2012) further qualified the MST as a network structure between a group of nodes, representing different time series, whose edges minimise the pairwise distances between each pair of nodes. In other words, an MST can be seen as a parsimonious representation of a network model, in which sparseness replaces completeness.

More formally, consider N financial assets, for which we observe the corresponding price time series: $(P_i, i = 1, \ldots, N)$, each of which is a vector of prices observed in T different time periods: $P_i = (P_i(t), t = 1, \ldots, T)$. From the price time series we can obtain N return time series, $(R_i, i = 1, \ldots, N)$, as follows:

$$R_i(t) = logP_i(t) - logP_i(t-1).$$

From the return time series we can calculate the correlation matrix C, whose elements C_{ij} are defined by:

$$C_{ij} = \frac{E(R_iR_j) - E(R_i)E(R_j)}{\sigma(R_i)\sigma(R_j)},$$

where $E()$ indicates the mean value and $\sigma()$ the standard deviation of each return time series.

From the correlation matrix C we can then calculate the distance between any two assset returns, d_{ij}, as follows:

$$d_{ij} = \sqrt{2 - 2\rho_{ij}},$$

a function which ranges between $(0, 2)$, with $d_{ij} = 0$ when $\rho_{ij} = 1$ and $d_{ij} = 2$ when $\rho_{ij} = -1$. It assumes that, for any pair of asset return time series, the higher the correlation, the lower the distance.

Let then $D = (d_{ij}, i = 1, \ldots, N; j = 1, \ldots, N)$ be a matrix which contains all pairwise distances. We can associate to the distance matrix a network $G = (V, W)$, with vertices V that correspond to the N asset return time series and with connection weights W between them which correspond to the $\frac{N(N-1)}{2}$ pairwise distances d_{ij}.

The Minimum Spanning Tree (MST) proposed by Mantegna (1999) is based on the distance matrix D. It reduces the number of weights that can connect the N nodes, from $\frac{N(N-1)}{2}$ to $N-1$. It does so through a hierarchical clustering algorithm which associates to each node only another one, that is minimally distant from it, under the constraint of avoiding loops between groups of nodes.

We remark that the network structure simplification induced by a minimum spanning tree may be too drastic, especially if based on random noise rather than on actual distances between nodes. To overcome this problem, in this paper we suggest to preprocess the correlation matrix and, therefore, the distance matrix, before applying the Minimum Spanning Tree method.

The necessity to improve the MST representation was pointed out by Tumminello et al. (2005), who introduced the Planar Maximally Filtered Graph (PMFG), which preserves the hierarchical structure of the MST, but with a more complex structure. Indeed, given a set of N time series, a MST contains $N - 1$ links whereas a PMFG contains $3(N - 2)$ links.

Here we aim to improve the MST without enriching its structure but, rather, working on its input: the distance matrix. To achieve this goal, we

employ the Random Matrix Theory approach (RMT), proposed by Onnela et al. (2004) and Tola et al. (2008), pre-processing the correlation matrix by removing the noise contained in it.

The rationale behind the random matrix theory approach is to employ each empirical eigenvalue $(\lambda_k, k = 1, \ldots, N)$ obtained from the correlation matrix C, as a test statistic for the null hypothesis that the correlation matrix is a random Wishart matrix $C' = \frac{1}{T}AA^T$, where A is a $N \times T$ matrix containing N time series of length T, whose elements are independent and identically distributed "white noise" random variables, with zero mean and unit variance.

To actually implement the test, we need a statistical distribution. Marchenko and Pastur (1967) showed that, under the null hypotheses, $\lambda_1 = \ldots = \lambda_N = \lambda$, and that the asymptotic density of λ, for a fixed $Q = \frac{T}{N} \geq 1$, as $N \to \infty$ and $T \to \infty$, is given by:

$$f(\lambda) = \frac{T}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}$$

where $\lambda \in (\lambda_-, \lambda_+)$, with $\lambda_+ = \sigma^2 + \frac{1}{Q} + \sqrt{\frac{1}{Q}}$ and $\lambda_- = 1 + \frac{1}{Q} - \sqrt{\frac{1}{Q}}$.

From the above density, it follows that, when $\lambda_k \geq \lambda_+$, the null hypotheses is rejected, as the k-th empirical eigenvalue cannot be an eigenvalue from a random Wishart matrix.

From an operational viewpoint, if the eigenvalues are ordered from the largest to the smallest, we can retain only those that exceed λ_+ and reconstruct the correlation matrix, through singular value decomposition, using only the eigenvectors corresponding to them. Doing so, as suggested by Plerou et al. (2002) we "filter" the correlation matrix.

From an empirical viewpoint, Miceli and Susinno (2004) show that, when the random matrix approach is applied as outlined before, the minimum spanning tree leads to a grouping of assets that better correspond to "typical" investment strategies. Our aim is different: we would like to verify whether the application of the RMT on the correlation matrix between bitcoin exchange and classical market prices produces a minimum Spanning Tree

that can shed light on what drives bitcoin prices: endogenous or exogenous factors.

3 Data

In this Section, we describe the analysed data.

We consider, without loss of generality, the most important cryptocurrency: the bitcoin, whose relative price will be taken with respect to the US dollar. With no further loss of generality, and to reduce volatility issues, we consider daily prices, obtained at the end of the day.

Our first research question is to assess whether bitcoin prices in different exchange markets are correlated with each other, thus exhibiting "endogenous" price variations. To understand this question, we have chosen a set of representative exchange markets, for which price data is available, in a sufficiently long period of time. Specifically, we have selected eight exchange markets, representative of different geographic locations, which represent about 60% of the total daily volume trades. They are reported in Table 1, along with the corresponding market shares. For each exchange market, we have collected daily data for a time period that goes from May, 18th, 2016, to April 30th, 2018.

[Table 1 about here]

Our second research question is to understand whether bitcoin price variations can also be explained by exogenous classical market prices. To evaluate this issue, we have obtained daily data on some of the most important asset prices: Gold, Oil and SP500; as well as on the exchange rates USD/Yuan and USD/Eur. Similarly to what done for bitcoin prices, we have considered, as daily price, the market closing price. When jointly considering bitcoin and "standard" markets, one issue to be solved is that, while Bitcoins are traded 24 hours per day and 7 days per week, standard markets have closing times

and days. We have overcome this issue keeping standard market prices constant at the last closing time, during market closure.

Figure 1 presents the time evolution of the Bitcoin prices, in the considered time period.

[Figure 1 about here]

From Figure 1 note the well known 2017 rise in Bitcoin prices, from a minimum of about 430 dollars per bitcoin to a maximum of almost 20,000 dollars, followed by a high volatility in 2018. Note the slight differences between prices, which shows that bitcoin prices in different market exchanges are not perfectly aligned. To better understand the latter finding, some summary statistics on the considered data are presented in Table 2.

[Table 2 about here]

Table 2 confirms the slight differences in bitcoin prices along the considered market exchanges: the means and the standard deviations are slightly different, and more so are the extreme statistics. With respect to classical assets, such as Gold and Oil, the volatility of bitcoin prices is much higher: respectively, about 80 and 1400 times higher. Even with respect to SP500, the volatility of Bitcoin prices is about 20 times higher. Instead, exchange rates are, as well known, much less volatile than bitcoin prices. These results are in line with the available literature (see e.g. Corbet et al., 2016b). Finally, looking at the last column in Table 2 note that bitcoin prices have values of kurtosis quite similar among each other, and lower than those of the classical assets,

4 Empirical findings

The aim of this section is to apply our proposed model to verify whether bitcoin prices from different exchanges are strongly interrelated with each other and whether such "endogenous" interactions are affected by "exogenous" prices of classical assets.

Figure 2 presents, by means of a heatmap, all pairwise correlations between the considered asset prices, in the considered time period. Positive correlations are marked in blue, and negative correlations in red, with stronger colors indicating higher correlations (in absolute values).

[Figure 2 about here]

From Figure 2 note that the correlations between different exchange prices are quite high, revealing that markets are highly correlated and synchronized, resulting in a strong endogenous driver of price variation. On the other hand, correlations with "real" asset prices, such as gold and oil, are low, a result in line with the literature, that considers bitcoins as potential diversification assets (see e.g. Corbet et al., 2018a). However, the correlation with the SP500 index is positive and those with the exchange rates are negative, a result that seems to conflict with the reference literature.

To better understand the implications of Figure 2, Giudici and Abu-Hashish (2018) analysed similar bitcoin price data using partial correlation networks. This because pairwise correlation may be inflated by correlations that may arise from a common relationship with third variables. Their empirical findings show that bitcoin prices on one hand, and "classic" asset prices on the other hand, form two rather distinct clusters of connections, which are are highly interconnected inside. They also show the high centrality of two of the largest bitcoin exchanges: Bitfinex and Bitstamp, which thus emerge as "price setters". They also find that the link between the two clusters is given by the Hitbtc exchange, which is affected both by standard asset prices and by other exchange market prices.

Here we take a different approach to improve the empirical findings that can be obtained from the correlation matrix in Figure 2. We derive the Minimum Spanning Tree of the correlation matrix, introduced in Section 2. The obtained MST is shown in Figure 3.

[Figure 3 about here]

From Figure 3 note that Bitfinex and Bitstamp have a pivotal role, as found in Giudici and Abu-Hashish (2018). However, the MST reveals more insights. For example, it shows that, while Bitfinex is "closer" to real financial assets, such as Gold, SP500 and Oil, Bitstamp is more related with exchange rates. In addition, Hitbtc separates Bitfinex from Gold. Note also that smaller exchanges are more peripheral.

Indeed, the advantage of MST models, with respect to correlation network models, is that they provide a "hierarchical" split of the prices (nodes), showing them in order of distances (weights), calculated from their correlations. Table 3 reports the weights corresponding to the application of the MST algorithm to the considered data.

[Table 3 about here]

From Table 3, note that the "closest" nodes are those between bitcoin price exchanges, as expected: their pairwise connections correspond to the first seven edges of the MST. The following edge is placed between the two exchange rates, then between SP500 and oil. Last, the procedure finds three edges that break the "separation" between crypto and classical asset prices: the first one relates Hitbtc with Gold; the second one Gemini with SP500; the last one UsdYuan with Kraken. These latter results are quite meaningful, as they characterise the "local" behaviour of specific exchanges, a phenomena already found in Giudici and Abu-Hashish (2018).

We now verify whether the application of the Random Matrix Theory approach, before implementing the Minimal Spanning Tree, can extract further empirical findings from the correlation matrix. The results are shown in Figure 4.

[Figure 4 about here]

From figure 4 note that the filtered MST provides a graphical structure that is simpler then that obtained in 3, without the application of RMT

filtering. On one hand, bitcoin exchange prices form a "star" configuration, with Bitstamp at the center, confirming its role of price setter; while Bitfinex, probably because of its relatively high volatility, is not found to be central. On the other hand, all classic asset prices are separated from bitcoin prices, pointing towards a "diversification benefit" of bitcoins with respect to them, a result fully in line with the existing literature. Note also that the MST well separates the role of "real" assets, such as SP500, Oil and Gold, from "financial" assets such as the exchange rates.

To summarise, filtering the correlation matrix with the random matrix approach leads to a Minimum Spanning Tree that, with respect to the unfiltered one, is simpler and which leads to empirical findings that: i) do not indicate a significant correlation between crypto prices and exogenous price drivers, from classical markets, consistently with the literature; ii) indicate that exchange prices have a strong endogenous source, which specifically come from the largest and least volatile exchanges, such as Bitstamp.

We can draw more interpretation examining the distance weights corresponding to the joint application of the RMT and MST, in Table 4.

[Table 4 about here]

Comparing Table 4 with Table 3, the previously discussed findings are confirmed. Again the seven closest pairs of nodes concern bitcoin exchange prices, indicating a strong presence of endogeneus price variation; in addition, in Table 4, all pairs contain the Bitstamp node, indicating its centrality. A further difference is that the connections between crypto prices and classic prices reduce to two, and they both involve Hitbtc. This result is more in line with what obtained in Giudici and Abu-Hashish (2018) about the role of Hitbtc as a "separator" between classic and crypto assets.

To assess the robustness of our empirical findings, we now verify whether the found tree structure is stable over time. For this purpose, Figure 5 shows the MST obtained in each of nine one-year rolling periods, after the application of RMT. The first one starts from 18/05/2016, the following are shifted ahead by one month, until the eigth one which starts on 18/02/2017.

[Figure 5 about here]

Figure 5 shows that the MST configuration is quite stable over time, particularly from the third period onwards, as all graphs show a configuration similar to the "static" one in Figure 4. From a theoretical viewpoint, we remark that, when random matrix theory is applied before the application of the minimal spanning tree the results are stabilized, as RMT filters out noise. In fact, comparing the different time periods in Figure 5, the spanning trees do not change sensibly, even during bubble periods. We have indeed applied the test for crypto bubbles suggested in Cheah and Fry (2015) and Hafner (2018), obtaining that the december 2017 period shows a significant bubble. However, Figure 5 shows that the two correlation networks at the bottom of Figure 5, which fully contain the bubble period, do not show an evident structural change.

We have conducted a further robustness test on the time dynamics of our results. From Table 2 the kurtosis observed for the bitcoin prices is smaller than that of classical assets, and this may justify the use of an unconditional variance. We have however assumed that the unconditional variance is different from the realised one and we have calculated pairwise correlations not among returns, as before, but among volatilities, to see what could drive the volatility dynamics, rather than the price dynamics.

In particular, we have postulated the existence of a negative correlation between the realised macroeconomic volatility and the realised volatility of bitcoin prices, as suggested by Conrad, Custovic and Ghysels (2018). These authors report that the two months lagged SP500 realised volatility may be a useful predictor for the bitcoin volatility. Following this suggestion, we have calculated the pairwise correlations between all bitcoin exchange volatilities, and all classical assets volatilities (lagged by two months), and reported them in Figure 8.

[Figure 8 about here]

From Figure 8, note that the correlation between classic assets and bitcoins is constant across different exchanges. In particular, the bitcoin volatility is negatively correlated with that of the SP500 index. This indicates a further, important, empirical finding: the volatility of classic asset prices negatively affects the volatility of bitcoin prices, with a delay.

This result, that confirms Conrad, Custovic and Ghysels (2018) can be better seen in Figure 9, which reports in the same graph the realised volatilities of the Bitcoin Bitstamp price and for the SP500 index In the figure, both voltilities have been normalised, being the volatility of the bitcoin price much higher (10 times more on average).

[Figure 9 about here]

From Figure 9 the two months lagged effect of the volatility of the SP500 index on the bitcoin price volatility of the Bitstamp exchange, is evident. Similar results hold for all other exchange prices, consistently with the found price setter nature of the Bitstamp exchange.

As a last robustness exercise on our empirical findings, we compare, on the same data, our method with the Planar Maximally Filtered Graph and with the Granger causality network, suggested in Billio et al. (2012). Figure 8 and 9 give the results from the application of these two methodologies.

[Figure 8 about here]

[Figure 9 about here]

Figure 9 and 8 show that, as expected, the Granger causality network graph and the Planar maximally filtered graph are more connected than our Minimal Spanning tree graphs. They also show that the connections found with the MST are also significant present in the Planar Maximally Filtered Graph and in the Granger causality network graph. Upon comparison with Giudici and Abu-Hashish (2018) the same connections are also present in their partial correlation graph. All these findings lead to the conclusion

that the relationships found by our MST graphs are consistently found using other methods and, therefore, the interpretation drawn upon their findings are quite robust.

5 Conclusions

We have proposed a new statistical model for the explanation of what drives the bitcoin prices. The model is based on the correlation matrix between the observed returns, which is first filtered from noise, applying the Random matrix theory and, then, employed to derive a clustering structure among prices, applying the Minimum Spanning Tree.

Our main methodological contribution consists in the combination of the random matrix theory approach with the minimum spanning tree approach, and in their application to the determination of the bitcoin price drivers.

Our empirical findings show that bitcoin prices from different exchanges have a strong endogenous driver of variation: they are highly interrelated, as in an efficiently integrated market, consistently with the literature. In addition, we found that the largest and least volatile exchanges (such as Bitstamp) are the most important price setters. Our results also confirm the literature in showing that bitcoin prices are unrelated with classical market prices, thus bringing further support to the "diversification benefit" property of crypto assets. In addition, we found that the volatility of classic assets affects negatively, and with a time lag, the volatility of bitcoin prices.

Finally, our empirical findings are robust, with respect to the consideration of different time periods, that also include bubbles, and are consistent with those obtained from different methodologies aimed at measuring interconnectdness between market prices.

We believe that the main beneficiaries of our results may be regulators and supervisors, aimed at preserving financial stability, as well as investors of crypto assets, who should be protected against the negative sides of fintech innovations (higher risks) while keeping their positive sides (lower costs and

better user experience). For a general discussion of this point see also Giudici (2018).

Future work requires acquiring more data, on other bitcoin exchanges, and on other crypto assets, to further assess the validity of the obtained conclusions, and possibly obtain further findings. From a methodological viewpoint, it may be worth considering modelling assets returns with generalised extreme value distrutions (as in Calabrese and Giudici, 2015), which can take high volatility into account; or with Bayesian models (as in Figini and Giudici, 2011), which can incorporate expert information into the model. It would also be importna to consider the implications of our results in terms of asset allocation.

Acknowledgements

We acknowledge support from the European Union's Horizon 2020 research and innovation programme, under grant agreement No 825215 (Topic: ICT-35-2018 Type of action: CSA), and from the Universitá Politecnica delle Marche. We also thank two anonymous referees, for the provided suggestions which have helped improving the paper.

References

BOURI, E., AZZI, G., & DYHRBERG, A.H. 2017. On the return-volatility relationship in the Bitcoin market around the price crash of 2013. *Economics: The Open-Access, Open-Assessment E-Journal*, **11**(2017-2).

BRANDVOLD, M., MOLNR, P., VAGSTAD, K., VALSTAD, A., & OLE, C. 2015. Price discovery on Bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money*, **36**(C), 18–35.

CALABRESE, R., & GIUDICI, P. 2015. Estimating bank default with generalised extreme value regression models. *Journal of the Operational Research Society*, 1–10.

CHEAH, E.T., & FRY, J. 2015. Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, **130**, 32–36.

CIAIAN, P., & RAJCANIOVA, M. 2018. Virtual relationships: Short-and long-run evidence from bitcoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money*, **52**, 173–195.

CONRAD, C., CUSTOVIC A., & GHYSELS, E. 2018. Long- and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis. *Journal of Risk and Financial Management*, **to appear**.

CORBET, S., LUCEY, B., URQUHART, A., & YAROVAYA, L. 2018a. Cryptocurrencies as a financial asset: a systematic analysis. *Working Paper*.

CORBET, S., MEEGAN, A., LARKIN, C., LUCEY, B., & YAROVAYA, L. 2018b. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, **165**, to appear.

DWYER, G.P. 2015. The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, **17**(apr), 81–91.

Dyhrberg, A.H. 2016. Bitcoin, gold and the dollar: A GARCH volatility analysis. *Finance Research Letters*, **16**, 85 – 92.

Figini, S., & Giudici, P. 2011. Statistical merging of rating models. *Journal of the Operational Research Society*, **62**, 1067–1074.

Giudici, P. 2018. Fintech Risk Management: A Research Challenge for Artificial Intelligence in Finance. *Frontiers in Artificial Intelligence*, **1**.

Giudici, Paolo, & Abu-Hashish, Iman. What determines bitcoin exchange prices? a network VAR approach. *Finance Research Letters*.

Gonzalo, J., & Granger, C. 1995. Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business & Economic Statistics*, **13**(1), 27–35.

Hafner, C.M. 2018. Testing for bubbles in cryptocurrencies with time-varying volatility. *Journal of Financial Econometrics*, **to appear**.

Hasbrouck, J. 1995. One Security, Many Markets: Determining the Contributions to Price Discovery. *The Journal of Finance*, **50**(4), 1175–1199.

Mantegna, Rosario N. 1999. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, **11**(1), 193–197.

Marchenko, V.A., & Pastur, L.A. 1967. Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, **114**, 507–536.

Miceli, Maria-Augusta, & Susinno, Gabriele. 2004. Ultrametricity in fund of funds diversification. *Physica A: Statistical Mechanics and its Applications*, **344**(1-2), 95–99.

Nadarajah, S., & Chu, J. 2017. On the inefficiency of Bitcoin. *Economics Letters*, **150**(6-9).

Nakamoto, Satoshi. 2008. Bitcoin: A peer-to-peer electronic cash system. *www.bitcoin.org*.

Onnela, J-P, Kaski, Kimmo, & Kertész, Janos. 2004. Clustering and information in correlation based financial networks. *The European Physical Journal B*, **38**(2), 353–362.

Pagnottoni, P., Dirk, B., & Dimpfl, T. 2018. Price Discovery on Bitcoin Markets. *Working Paper*.

Plerou, Vasiliki, Gopikrishnan, Parameswaran, Rosenow, Bernd, Amaral, Luis A Nunes, Guhr, Thomas, & Stanley, H Eugene. 2002. Random matrix approach to cross correlations in financial data. *Physical Review E*, **65**(6), 066126.

Spelta, Alessandro, & Araújo, Tanya. 2012. The topology of cross-border exposures: beyond the minimal spanning tree approach. *Physica A: Statistical Mechanics and its Applications*, **391**(22), 5572–5583.

Tola, Vincenzo, Lillo, Fabrizio, Gallegati, Mauro, & Mantegna, Rosario N. 2008. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, **32**(1), 235–258.

Tumminello, Michele, Aste, Tomaso, Di Matteo, Tiziana, & Mantegna, Rosario N. 2005. A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(30), 10421–10426.

Tumminello, Michele, Di Matteo, Tiziana, Aste, Tomaso, & Mantegna, Rosario N. 2007. Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B*, **55**(2), 209–217.

Urquhart, A. 2016. The inefficiency of Bitcoin. *Economic Letters*, **148**(80-82).

Figures and Tables

Table 1: Exchange markets by daily trading volumes

Exchange	Market share
Bitfinex	42%
Coinbase	6%
Bitstamp	5%
Hitbtc	3%
Gemini	2%
itBit	1%
Kraken	0.5%
Bittrex	0.5%

Note: The considered market exchanges, by daily trade volume market shares. Data taken from https://cryptocoincharts.info/markets/info

Table 2: Summary statistics

Price	Mean	St. Dev.	Min	Max	Kurtosis
Bitfinex Bitcoin	3899.56	4274.46	435.61	19187.12	3.22
Coinbase Bitcoin	3919.05	4318.98	438.38	19650.01	3.22
Bitstamp Bitcoin	3899.04	4286.02	439.62	19187.78	2.50
HitBtc Bitcoin	3916.19	4297.17	436.36	19095.30	3.77
Gemini Bitcoin	3910.38	4306.36	437.57	19475.90	2.90
ItBit Bitcoin	3907.13	4300.32	438.61	19357.97	2.67
Kraken Bitcoin	3890.18	4272.55	433.50	19356.91	2.05
Bittrex Bitcoin	3893.83	4269.89	421.11	19261.10	2.53
Gold	1275.57	52.34	1128.42	1366.38	7.02
Oil	48.67	3.16	39.51	54.45	18.98
SP500	2414.78	212.308	2000.54	2872.87	11.86
USDYuan	6.67	0.19	6.26	6.96	4.85
USDEur	0.88	0.04	0.80	0.96	4.53

Note: Summary statistics for bitcoin and classic asset prices. Means, standard deviations, minimum and maximum values are all expressed in dollars.

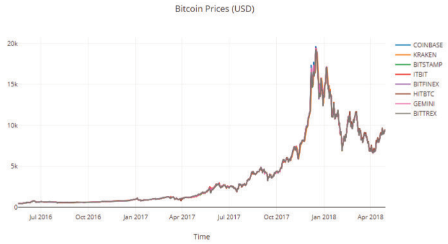


Figure 1: The time series plot of Bitcoin prices.

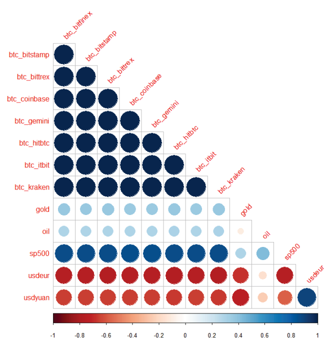


Figure 2: Correlation matrix between prices.

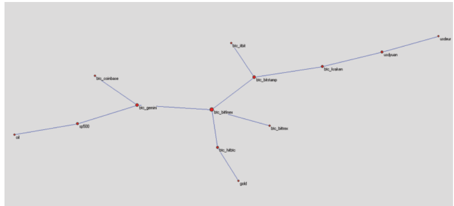


Figure 3: Minimum spanning tree between prices.

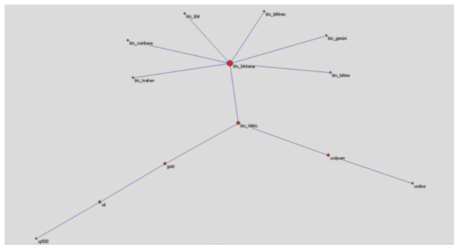
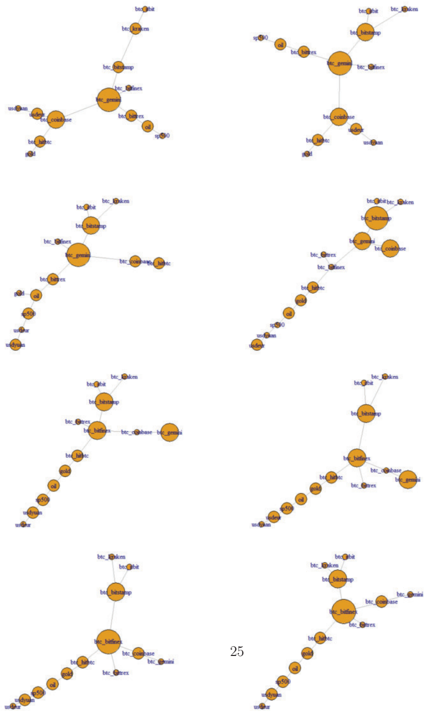


Figure 4: Minimum spanning tree between prices, after random matrix theory filtering.

			weight
1	btc_coinbase	btc_gemini	0.207968897069462
2	btc_bitfinex	btc_gemini	0.252298438331244
3	btc_bitstamp	btc_itbit	0.279684526027965
4	btc_bitstamp	btc_bitfinex	0.314896774631822
5	btc_bitfinex	btc_hitbtc	0.380165409259575
6	btc_bitfinex	btc_bittrex	0.458640635735809
7	btc_kraken	btc_bitstamp	0.596065889298915
8	usdeur	usdyuan	0.989859235886413
9	sp500	oil	1.23032152766612
10	gold	btc_hitbtc	1.35470479294313
11	btc_gemini	sp500	1.3728902044797
12	usdyuan	btc_kraken	1.3951809819423

Table 3: Adjacency matrix from MST.

Figure 5: Time evolution of the the minimum spanning tree.



			weight
1	btc_gemini	btc_bitstamp	0.288433056210555
2	btc_bitstamp	btc_bitfinex	0.292594189179681
3	btc_coinbase	btc_bitstamp	0.292645290798001
4	btc_bitstamp	btc_itbit	0.359144767683509
5	btc_hitbtc	btc_bitstamp	0.39147736801806
6	btc_bitstamp	btc_bittrex	0.440520251858754
7	btc_kraken	btc_bitstamp	0.600090764365521
8	usdeur	usdyuan	0.759135861596123
9	sp500	oil	1.14667657820831
10	oil	gold	1.2914392737851
11	gold	btc_hitbtc	1.35598797995048
12	usdyuan	btc_hitbtc	1.39185686202477

Table 4: Adjacency matrix from RMT+MST.

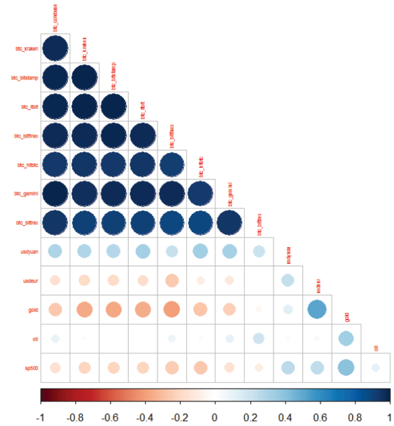


Figure 6: Volatility correlation plot.

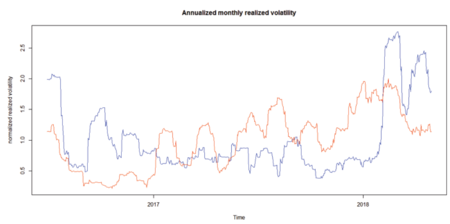


Figure 7: Realised volatilities for SP500 (blue) and Bitcoin Kraken (red), normalised.

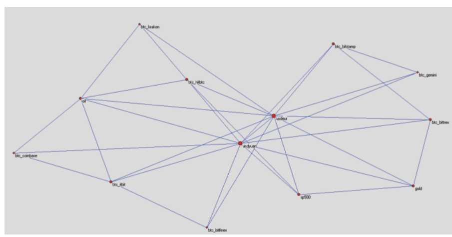


Figure 8: Planar maximally filtered graph

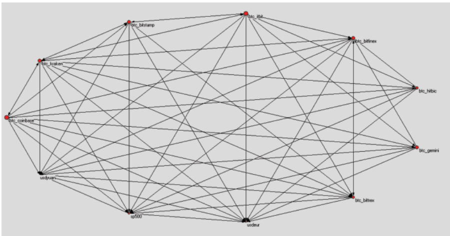


Figure 9: Interconnectedness graph obtained from the application of Granger causality networks.