DEPARTMENT OF ECONOMICS AND FINANCE

MASTER’S THESIS IN
THEORY AND PORTFOLIO MANAGEMENT

CONTAGION DYNAMICS BETWEEN BITCOIN AND STOCK MARKETS: A NETWORK ANALYSIS APPROACH

CANDIDATE
Giorgio Campisano

SUPERVISOR
Prof. Nicola Borri

CO-SUPERVISOR
Prof.ssa Cecilia Susanna Jona Lasinio

Academic Year 2021/2022

The work of a financial analyst
falls somewhere in the middle
between that of a mathematician
and of an orator.

Benjamin Graham

Abstract

Bitcoin's behaviour as an asset class towards other asset classes has been mutating over the last decade, but research studies on this topic seem to be discontinuous. In this work, I utilise a Granger-causal Network Analysis to assess whether Bitcoin shows contagion dynamics with a selection of stock market indices. The analysis is performed using both a liquidity indicator and log-returns, and different time frames are considered. Results indicate that Bitcoin has both an active and passive role in the overall contagion system, depending on the specific analysis.

Table of contents

Preliminary Concepts on Bitcoin

In 2008, 31^{st} October, with the release of Bitcoin's whitepaper by Satoshi Nakamoto, a new technology came to light for the first time. As the author describes it, Bitcoin is "A purely peer-to-peer version of electronic cash"[1], and its original purpose, as intended by its creator, was to avoid all kind of brokerage involved in transactions between two parties while avoiding the risk of "double-spending" offered by some other peer-to-peer methods, eliminating the need for a trusted third party acting as a supervisor and guarantor for the transaction. Eventually, this leads to the elimination of all kind of transaction costs related to third-party brokerage.

Bitcoin's system relies on cryptographic proof instead of relying on trust in a third party as the standard electronic payment's scheme does, thus utilizing cryptographic techniques to make transactions "computationally impractical to reverse"[2]. In particular, the proof problem is addressed creating a peer-to-peer distributed timestamp server which serves as proof of the list of transactions by chronological order, making them "written in stone". With enough simplification, with such a public record of transactions all the clients in the network, i.e. the *nodes*, can verify the history of the coin, "agreeing" that it has not been double-spent. This "consensus" system is the core of Bitcoin's technology: transactions security is ensured as soon as the majority of nodes is "honest" and verifies the "true" history of the coin, determining which of the transactions arrived first to the server and considering it the only acceptable one. This is accomplished through hashing a block of transactions and publishing it on the server; the following block that arrives to the server will contain in its hash both its timestamp and the timestamp of the previous block, thus forming a chain, i.e. the *blockchain*. The more the chain grows, the more secure it gets: for an attacker to modify one of the blocks, thus modifying transactions in a fraud attempt for example, they would need to recreate the whole blockchain in order for the other nodes to accept it as the true one. S. Nakamoto (2008) proves in his work that the probability of this event drops exponentially with the length of the blockchain, being of $P = 1,2 \times 10^{-6}$ for a length $z = 10$ blocks; to date, Bitcoin's number of blocks is 767,061.

1 S. Nakamoto (2008)
2 S. Nakamoto (2008)

1 Introduction

This work's purpose is the investigation about the existence of contagion dynamics between cryptocurrency markets and other traditional financial markets, both in terms of asset liquidity and returns, with the support of a network analysis. Cryptocurrency markets are becoming widely more common among investors, and are starting to gain growing importance in the asset classes universe, attracting an increasing amount of money in the last years from both retail and institutional investors. More, recent events have shown how this asset class is still something not completely understood by investors, which can lead to market crashes and enormous money losses[3]. It is therefore more important now than ever to understand dynamics between such digital asset and other traditional assets.

When Satoshi Nakamoto invented Bitcoin, with the Bitcoin's whitepaper publication in 2008, their stated intent was to create an alternative form of currency which would have been free from third-party control and brokerage needs; but the cryptocurrencies environment that we can witness in our time seems far different from such intent. Today, we ask ourselves whether Bitcoin and other cryptocurrencies can be considered as alternatives to fiat money or they're some other sort of financial asset. Baur et al. (2015) investigated this question. In 2015, when their work was published, Bitcoin's situation was different compared to current days: Bitcoin was showing no correlation with traditional asset classes both in normal times and during market turmoils, the authors denote, but they also found that the cryptocurrency was mainly used in a speculative way for its high volatility and possibly consistent returns, while only a minority of users were treating Bitcoin as a medium of exchange. Moreover, its previous low correlation with other financial assets would have implied that Bitcoin could have served as a good diversification asset, and also that it would not present macroeconomic risk, containing possible speculative bubbles and crashes inside its environment without affecting other asset classes; this also due to its limited size in value, back in time.

But things have changed since that paper was published. Bitcoin's current market capitalization is 319B USD[4], more than 70 times bigger compared to the approximately 4.5B USD back in 2015, and in the last 7 years its price experienced a huge surge (over 68.000 USD in November 2021, bringing its market capitalization up to 1.280B USD) coming along

[3] I am here referring in particular to the FTX scandal that came out last November, 2022, where important institutional investors found themselves trapped in a multi-billion dollar scam attributable to the young entrepreneur Sam Bankman-Fried and carried out through its companies.

[4] as of 18^{th} Nov. 2022 (source: Coinmarketcap.com).

with the growing interest in the cryptocurrencies world from the public, especially young investors; it then experienced a huge crash recently, reminding everyone how risky it can be to invest in something that it is not well-known yet. So, researchers started to ask themselves again whether the absence of correlation between Bitcoin (and other cryptocurrencies) and other financial assets still holds, and this work is intended to contribute to such research area.

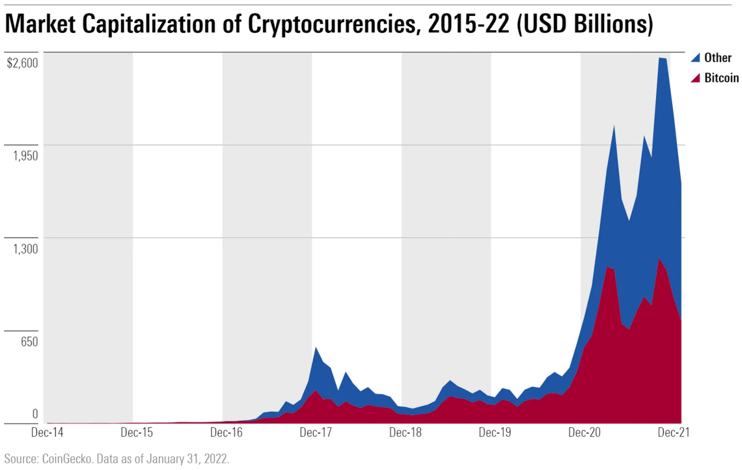

Figure 1. Cryptocurrencies market capitalization's growth in recent years.

1.1 Contagion Definition

What do I mean with the term "contagion"? In the current literature, this word appears to have multiple meanings. The first work about financial contagion is Allen, Gale (2000). The authors studied the propagation of crisis in the banking sector, and in their introduction the term "contagion" is mentioned for the first time as follows:

"One theory is that small shocks, which initially affect only a few institutions or a particular region of the economy, spread by *contagion* to the rest of the financial sector and then infect the larger economy."[5].

As we could expect, "contagion" is here intended in the same flavor of the medical term "contagion", i.e. the transmission of an infectious illness from an infected individual to an healthy but susceptible one, directly or indirectly. In the two author's work, infected individuals were distressed banks that, being linked to other financially healthy banks by various channels, were responsible for infecting those latter ones, thus propagating their distress to

[5] Allen, Gale (2000)

the broader financial sector and, eventually, to the whole economy: in short, "contagion" is intended as a form of systemic risk modeling[6].

After this pioneer work, the financial contagion research branch was born and many more working papers were produced by the literature; the focus was also enlarged from the sole banking system to the broader financial one, including capital markets. An important document in this sense is Pericoli, Sbracia (2003), where the authors traced the outline of a framework for understanding the possible transmission channels of financial shocks, investigating which channels exist, whether there are discontinuities in such channels and whether international investors and policy-makers should be concerned about those kind of interconnections, denoting for this last subject that: "If cross-country correlations of asset prices are significantly higher in periods of crisis [...] portfolio diversification may fail to deliver exactly when its benefits are needed most.".
But the authors denote that, at the time of writing, a lot of confusion was present on the term "contagion", as a theoretical or empirical identification of such phenomenon was missing. The term was usually linked to any shock with propagation effects outside the market that generated it, in a geographic sense, but there was no distinction between shocks generated by normal markets interconnection and shocks that generate a discontinuity in stock prices. They then provide different definitions of contagion, also depending on the specific financial crisis identified. Different kind of financial crises definition present in literature are: *currency crises*, referring to turmoils in the forex market characterised by pressures in exchange rates; *stock market crises*, referring to plunges in stock market indices or steep rises in stock prices volatility; and *banking crises*, referring to events like credit deterioration, failure or bailouts of important banking institutions or bank runs. The authors also provide their formal definition of financial crisis: "A crisis in country n at time t is an unexpected change in the distribution of Y_{t+1}^n that increases the risk of investing in country n."[7].

From these definitions, Pericoli, Sbracia (2003) find that five different predominant definitions of contagion can be found in the literature:

Definition 1. *Contagion is a significant increase in the probability of a crisis in one country, conditional on a crisis occurring in another country.*
This definition is usually linked to shocks in exchange rates that typically spread their ef-

[6] Interestingly enough, the authors, while explaining why the financial sector was (and, probably, still is) so fragile, introduced asset *decentralization* as a possible solution to enhance risk sharing and avoid distress in the system, but they demonstrate that such a model is not sustainable in the long run.
[7] see Pericoli, Sbracia (2003).

fects on multiple countries. Here, factors that drive propagation mechanisms are not directly investigated, so that an heterogeneous list of phenomena can be labeled as contagious - common shocks, trade links, irrationality or policy outcomes.

Definition 2. *Contagion occurs when volatility of asset prices spills over from the crisis country to other countries.*
As the authors denote, this definition relies on the typical stylized fact in international financial markets of an increase in assets volatility during periods of financial turmoil, thus identifying contagion as a volatility spillover between markets in different countries. It has to be specified, however, that this definition does not take into account shocks causes, not distinguishing between those shocks due to normal interdependence between markets and shocks attributable to structural breaks that affect more than one market.

Definition 3. *Contagion occurs when cross-country co-movements of asset prices cannot be explained by fundamentals.*
Such definition is in line with those theoretical frameworks that allow for multiple equilibria for a given coordination problem, the authors say. The definition itself is referring to the fact that, when a crisis occurs, a possible cause is an equilibrium switch: in such case, fundamentals alone are not enough to explain the entire variability experienced during the event. It is to underline that, in this context, fundamentals might be useful in predicting whether one country is susceptible for contagion or not: the authors bring forth the example of contagion effect via a liquidity crisis, where countries that have a low level of international reserves compared to short term liabilities denominated in foreign currency experience a higher risk of contagion with respect to those with higher level of reserves.
Also, multiple equilibria may not be arbitrary but event-driven: the authors denote that, mostly in standard models of currency crisis or bank runs, incomplete information (i.e. discontinuities in time series of data) ususally leads to multiple equilibria. In a similar situation, differences in behavior driven by different private information or operators uncertainty occur, thus producing unpredictable outcomes.

Definition 4. *Contagion is a significant increase in co-movements of prices and quantities across markets, conditional on a crisis occurring in one market or group of markets.*
This definition brings in a new concept in defining contagion: it is not a difference in *modalities* of effects, but a difference in *dimensions*, i.e. the intensity or different number of co-movements that occur during a financial crisis compared to the *same* co-movements that

exist during normal times. In my opinion, this is a new important concept. The focus is here moved from the possible different effects of a financial crisis in time series, often not easily observable, to a difference in magnitude of these effects. In this scenario, test for contagion would mean to understand the "normal" level of interconnections in markets, and then measuring differences from this found threshold, so to identify contagion whenever this normal quantity is exceeded. Of course, this leaves open the problem of finding a proper definition of "normal times".

Definition 5. *(Shift-)contagion occurs when the transmission channel intensifies or, more generally, changes after a shock in one market.*

This last definition acts as a synthesis of the previous two definitions, i.e. Definition 3 and Definition 4. As in Definition 3, no constraint is put on fundamentals: there can be other dynamics, like jumps in a situation of multiple equilibria, that explain some (or most) of the variability experienced during a crisis; and, as in Definition 4, the focus is on excess measures of some quantities - in this case, the Definition generally refers to the "transmission channel", which experiences a change (not necessarily an intensification) during contagion times. The authors denote also that such channels may start to exist due to the start of a financial crisis and cease to exist when the crisis ends, thus being "crisis-specific" channels.

Related to this set of definitions there are some measurement methods that the authors enumerate. It is interesting here to discuss some insights of co-movements in financial markets, related to Definition 4 above. In particular, let us consider the following data-generating process:

$$R_t = A + Bf_t + U_t, \quad U_t \sim (0, \Sigma_t) \tag{1}$$

$$\Sigma_t = C'C + D'\Sigma_{t-1}D + E'U'_{t-1}U_{t-1}E \tag{2}$$

where $R = [r_1, ...r_n]'$ is a vector containing rates of returns, $A = [\alpha_1, ..\alpha_n]'$ is a vector of constants, B is a $n \times k$ matrix of coefficients, $f = [f_1, ..., f_k]'$ is a vector of global (common) factors, $U = [u_1, ..., u_n]'$ is a vector of country-specific shocks (error terms) uncorrelated with f and with a covariance matrix Σ, with C, D and E being matrices of constants.

Assuming a single factor in f and constant variance in u, without loss of generality, and considering two elements i and j of the returns vector R:

$$r_i = \alpha_i + \beta_i \cdot f + u_i \tag{3}$$

$$r_j = \alpha_j + \beta_j \cdot f + u_j \tag{4}$$

with $\beta_i, \beta_j > 0$, the correlation between r_i and r_j is written as:

$$Corr(r_i, r_j) = \frac{1}{\left(1 + \frac{Var(u_i)}{\beta_i \cdot Var(f)}\right)^{1/2} \cdot \left(1 + \frac{Var(u_j)}{\beta_j \cdot Var(f)}\right)^{1/2}} \tag{5}$$

As Equation 5 shows, two main events for increase in sample correlation can occur: a decrease in $Var(u_i), Var(u_j)$, that is a contagion evidence; or an increase in $Var(f)$, that identifies interdependence rather than contagion, as the common factor affects both i and j. Since, during turmoils, both factors can experience large variations, a firm methodology to identify common factors variations compared to idiosyncratic factors variations is needed, thus being capable of distinguish between increased interdependence during a financial crisis and contagion dynamics.

The authors state that a way to tackle such problem is to assume a joint distribution for r_i and r_j and, then, find all the possible values of, say, r_i, and identify a subset C of values presented during a financial crisis. The objective is then to measure the sample correlation between r_i and r_j when $r_i \in C$ and compare it to the theoretical joint distribution to check for positive or negative excesses of co-movements.

All this being said, some caveats arise, as the authors report. First of all, for a statistical analysis, it is often needed to identify an initial shock in a market or a group of markets, but this may be not the case: a shock can belong to multiple markets or countries since its beginning, thus making its identification harder; this inaccuracy can limit the power of multiple tests and procedures used to measure contagion. Another important issue, closely related to co-movements in asset returns as discussed above, is data frequency selection. Pericoli, Sbracia (2003) correctly point out that there could be hidden contagion dynamics characterised by lags in the time series, making a correct data frequency selection hard: selecting intraday, daily, weekly data may not be appropriate due to said "infrequent but significant changes in asset prices that are correlated - with some lags - across markets"[8]. Also, these changes may lead to diverse effects - not necessarily higher correlations. Lastly, issues may be

[8] Pericoli, Sbracia (2003). The authors mention the example of a crisis in country i that leads investors to revise their expectations on productivity in country j, so that contagion in this scenario results in a one-time adjustment in the level of prices in country j.

found in selecting the correct quantities to analyse contagion: there may be other quantities more than prices that reflect the reaction of investors to a financial crisis in a country - like withdrawals of money. As it will be pointed out in Section 2.1, I addressed this last issue by choosing to measure contagion on a (il)liquidity measure that can take into account for both price and volumes movements, to try to capture the effect of withdrawals during crisis periods.

With respect to my analysis, considering all the definitions of contagion presented above, the definition that most adapts to my research intent is **Definition 5**. Such definition was proposed by Forbes, Rigobon (2002) in their work[9], where they measure stock market co-movements and try to understand whether these movements are due to interdependence or contagion. My choice is consistent with Gómez-Puig, Sosvilla-Rivero (2014), where the authors define contagion as *"an abnormal increase in the number or in the intensity of causal relationships, compared with that of tranquil period, triggered after an endogenously detected shock"*. Also, in the same way the authors of this last mentioned work did, the procedure I use to measure contagion not only takes into account the analysis of market correlations, but also tries to find causality in these co-movements, using a Granger-causality approach. Again, the usage of Granger causality tests to measure contagion is suggested by Forbes, Rigobon (2002): in particular, the authors state that such approach can be beneficial in all those cases "when the source of the crisis is less well identified and endogeneity may be more severe [...] to determine the extent of any feedback from each country in the sample to the initial crisis country."[10].

1.2 Network Analysis and Granger Causality

To analyse contagion, I take advantage of a procedure that estimates Granger-causal relations in panel data and creates a network based on the Granger interconnections found (see Section 2.3 for details). But what is intended for "network analysis"? In general, a network applied to some sort of data is a graphical interaction model. Simply put, network analysis is a way to represent relations present in panel data through a mathematical graph, where vertices are the variables in panel data and edges between vertices are a given kind of interconnection between the variables.

Many monographs and working papers have been published in the last century on the

[9] "This paper defines contagion as a significant increase in cross-market linkages after a shock to one country (or group of countries).", Forbes, Rigobon (2002). It is possible to denote the parallelism with Definition 5 provided by Pericoli, Sbracia (2003).

[10] Forbes, Rigobon (2002).

application of graphical interaction models to many kinds of data for various researches. It is here worth mentioning the work from Eichler (2007): the intent of the author was to provide a framework that allows the application of methods for causal inference to multivariate time series data - so far, such kind of application was only available for cross-sectional data. As Eichler says, graphical approach presents a set of advantages; principally, graphs can be easily read as they are a concise representation of the phenomenon object of research, and allow for a simple representation even if the panel is composed of a large number of variables.
The author demonstrates that some *"path diagrams"* that represent the autoregressive structure of a multivariate time series can be associated to a Granger-causality structure. Considering a multivariate time series X with an autoregressive representation such as follows:

$$X(t) = \sum_{u=1}^{\infty} \Phi(u) \cdot X(t-u) + \varepsilon(t) \tag{6}$$

where $\varepsilon(t)$ is a white noise process, a path diagram associated with X is a graph $G = (V, E)$, being $V = \{1, ..., d\}$ the set of vertices and E the set of edges. This path diagram presents the following two characterizations:

1. $a \to b \notin E \iff \Phi_{ba}(u) = 0 \;\; \forall \; u \in \mathbb{N}$
2. $a — b \notin E \iff \Sigma_{ab} = 0$

with the first one characterizing *directed* edges[11], while the second one describes *undirected* edges.
Related to the autoregressive representation is the Granger-causality probabilistic concept, and Eichler in his work refers to a linear specification of such concept[12]. Formally, let us consider a subprocess of X, denoted X_A, corresponding to a certain $A \subseteq V$. It is possible to define a certain information set $I_{X_A} = \{I_{X_A}(t), t \in \mathbb{Z}\}$ as a sequence of closed subspaces generated by the past and present values of X_A at time t, denoted as $\overline{X}_A(t) = \{X_A(s), \; s \leq t\}$. From here, we can consider two subsets A and B of $S \subseteq V$. With a "negative" definition, the process X_A is Granger-noncausal for the process X_B considering the information set I_{X_S} if:

$$X_A \nrightarrow X_B \; [I_{X_S}] \coloneqq \; X_B(t+1) \perp \overline{X}_A(t) \mid \overline{X}_{S \setminus A}(t) \quad \forall t \in \mathbb{Z} \tag{7}$$

Similarly, the processes are contemporaneously uncorrelated if:

$$X_A \nsim X_B \; [I_{X_S}] \coloneqq \; X_A(t+1) \perp X_B(t+1) \mid \overline{X}_S(t) \quad \forall t \in \mathbb{Z} \tag{8}$$

[11] note that the author specifies that directed edges are associated to a temporal ordering. This will be true also in my application - for further details, see Section 2.3
[12] I will stress on this concept later on in this work, in particular in Section 2.3, as it is one of the fundamentals of my network analysis. For more details, see Eichler (2007).

These definitions can be linked to the path diagram defined above considering the following two statements:

1. $a \rightarrow b \notin E \iff X_A \nrightarrow X_B \; [I_{X_S}]$
2. $a - b \notin E \iff X_A \nsim X_B \; [I_{X_S}]$

In this way, it is possible to draw the Granger-causality system of the variables through a graph, generating a Granger-causality path diagram which, ultimately, can be used to understand causal dynamics among the variables of the system.

Another important work in this field is the one from Diebold, Yilmaz (2014), where the authors study connectedness measures based on variance decomposition, where the forecast error variance of a certain variable in a set of variables is decomposed into different parts attributable to each other variable in the same set, and find that these measures are strictly related to network theories. A variance decomposition matrix D^H is defined as:

$$D^H = [d_{ij}^H], \quad i, j = 1, ..., N \tag{9}$$

where d_{ij}^H represents the fraction of i's H-step forecast error variance caused by shocks in variable j. It is important to stress out that connectedness measures are based on all those d_{ij}^H where $i \neq j$. Note as well that this matrix is not symmetric, i.e. $d_{ij}^H \neq d_{ji}^H$, since variance caused by j on i is usually not the same of that caused by i on j. It is then possible to define the *pairwise directional connectedness* from j to i simply as:

$$C_{i \leftarrow j}^H = d_{ij}^H \tag{10}$$

and, as said before, in general $C_{i \leftarrow j}^H \neq C_{j \leftarrow i}^H$, so that there may exist up to $N^2 - N$ different pairwise directional connectedness measures. Finally, it is possible to define the *total directional connectedness from others to* i as:

$$C_{i \leftarrow \bullet}^H = \sum_{j=1}^{N} d_{ij}^H, \quad j \neq i \tag{11}$$

and the *total directional connectedness to others from* j as:

$$C_{\bullet \leftarrow j}^H = \sum_{i=1}^{N} d_{ij}^H, \quad i \neq j \tag{12}$$

From this point, Diebold and Yilmaz went further, noticing that a variance decomposition matrix can be interpreted as a network adjacency matrix, thus being capable of producing

a network. First, they define the mathematical structure of a general network as a $N \times N$ adjacency matrix A made of zeros and ones, $A = [A_{ij}]$: $A_{ij} = 1$ if nodes are connected, i.e. related; otherwise $A_{ij} = 0$. With this definition, A is symmetric since, if i and j are connected, j and i are connected as well. So, all the network properties are embedded in the adjacency matrix A.
It becomes important then, for comparison and other study purposes, to understand how much the network described by A is connected, and some connectedness measure is needed. A first measure for connectedness in a network is the study of nodes *degree*. The degree of node i is defined as follows:

$$\delta_i := \sum_{j=1}^{N} A_{ij} = \sum_{j=1}^{N} A_{ji} \tag{13}$$

It is possible to collect the degrees of all the network nodes under a univariate discrete distribution with support $\{0, ..., (N-1)\}$, and its moments define some network characteristics. In particular, for network connectedness, the distribution mean is the commonly used benchmark: the larger the mean, the greater the connectedness of the network.

But what if the network is generated by a variance decomposition matrix, i.e. if $A = D^H$? In this case, some changes compared to a general network occur. First, the adjacency matrix A is not just filled with zeros and ones, since the elements of a variance decomposition matrix are *weights* that indicate a share of the total variance of that variable. Second, as pointed out above, A is no longer symmetric, and for how each entry d_{ij}^H is defined the matrix generates a *directed network*, where an edge between two variables indicates a causal connection from one to another. Also, diagonal elements of A are no longer 0, and each row has the constraint that all elements must sum to 1. Finally, the degree measure can be split into "to-degree" and "from-degree" (or "out-degree" and "in-degree") measures, that sum all the edges that nodes have depending whether the edge arrow points inwards or outwards with respect to a node. In particular, these correspond to column sums for the "to-degree" measure:

$$\delta_i^{to} = \sum_{i=1}^{N} A_{ij}, \quad i \neq j \tag{14}$$

and row sums for the "from-degree" one:

$$\delta_i^{from} = \sum_{j=1}^{N} A_{ij}, \quad j \neq i. \tag{15}$$

From these definitions, "to-degree" and "from-degree" probability distribution can be found. Notice that these degrees are identical to the total directional connectedness measures defined above, i.e. $\delta_i^{to} = C^H_{\bullet\leftarrow j}$ and $\delta_i^{from} = C^H_{i\leftarrow\bullet}$.

For what concerns my analysis, I will take advantage of a Granger-causal network estimation algorithm for panel data, called NETS, developed by Barigozzi, Brownlees (2018), and all the concepts presented above find place in the authors' work. In particular, their application is specifically meant for sparse systems of data, where interconnections are rare. In this sense, the NETS algorithm is an ideal tool to be used to distinguish between interconnection and contagion, as discussed in Section 1.1: in fact, after taking into account for interconnectedness in a system much less variability remains, making the interconnection system sparse. More details on the model are provided in Section 2.3 of this work.

1.3 Work Inspiration

Besides my personal knowledge and curiosity over this topic, many works from many authors have inspired me for the creation of this work; I will report here the ones that had a major impact on the genesis of this work.

The first inspiration for my work comes from Allen, Gale (2000): as said before, their work is the first one about financial contagion, and as soon as I understood its deep meaning I decided to conduct a similar systemic risk analysis on capital markets instead of the credit market; so, this working paper has been the first input.
The second inspiration for my work comes from Borri (2018), where part of the author's work is gaining insights about co-movements among some cryptocurrencies and between these and US Equities and Gold: my intention with my work was to create a sort of continuation to this paper, expanding on the co-movement analysis and thus producing a contagion-focused study that is valid for cryptocurrencies in general and Bitcoin in particular.
Third, I've always been fascinated by network analysis. I strongly agree with Eichler (2007) in saying that graphical analysis is a powerful tool that can provide a friendly way to understand more complex relationships that may happen in a system of random variables; this being considered, the work from Barigozzi, Brownlees (2018) seemed the most obvious choice for my research objective.
Fourth, but most importantly, my work talks about cryptocurrencies; so I feel safe in saying that Satoshi Nakamoto (2008) inspired me as well. Bitcoin's creation can ignite a true revolution for the global economic system, but the complete and correct understanding of what cryotocurrencies are and how these instruments can be implemented is crucial to improve our

welfare - it is my personal opinion that this technology can bring tons of positive changes in the overall system. Being Bitcoin a "young" asset, its market experiences a continuous and fast-paced development, so that available studies on it may result outdated after few years.

1.4 Contribution to the Literature

While looking for references, I noticed that available literature about cryptocurrencies' properties as an asset class is still relatively thin, particularly on contagion studies between these and other asset classes - which I feel as a hot topic, considering recent year's events. My intent is then to provide some contribution to this research area, for such a growing-interest market, with a work that could be beneficial for portfolio structuring and optimization, strategy and risk-management purposes.

Matkovskyy, Jalan (2019) point out that their work is the first one involved in the investigation of contagion effects from equity markets to Bitcoin markets; following their work, a relatively low number of working papers on the same topic has been produced by the literature. To mention some of these works, H. Wang et al. (2022) and X. Guo et al. (2021) are the most similar to the two author's work in their research intent, and thus to mine as well. In particular, the latter has the same research intent investigated with the same network modeling approach as I do; the tools that the authors utilize, however, are different from the one I rely on, as well as datasets and one of the contagion metrics.

It is useful to denote that a very limited amount of the works in the current literature utilizes a network modeling tool and, as far as I know, my work is the sole that utilizes the NETS algorithm in such research topic; most of these works, however, stress on Granger causality tests for their analysis. Lastly, it is important to point out that, to the best of my knowledge, my work is the first one analyzing contagion stressing on a liquidity measure in this specific research topic, as in my daily data analysis.

2 Empirical Analysis

To understand if different stock markets and cryptocurrency markets show contagion dynamics between each other, a simple correlation analysis would be ineffective. As denoted in Section 1.1, many different definitions of contagion are available in the current literature. It was thereby specified that in this work I will refer to contagion in the same way Gómez-Puig, Sosvilla-Rivero (2014) referred to[13].

As specified by the two authors in their work, "[...] if two markets show a high degree of co-movement during periods of stability, even if they continue to be highly correlated after a shock to one market, this may not constitute contagion, but only the outcome of the "interdependence" that has always been present in the markets.". Thus, in order to capture contagion dynamics a different approach is needed, which will be discussed in the following sections.

2.1 Contagion Metric

The original idea of this work was to test contagion on the base of an illiquidity measure. As stated in Ranaldo, Santucci de Magistris (2022), distressed markets are characterized by sudden drops in liquidity: in such context, using an illiquidity measure[14] can be useful to detect when trading volumes have an high price impact on a broad range of financial instruments. Ideally, this measure should be capable of detecting the so-called "panic selling" phenomenon.

The most famous illiquidity measure comes from Amihud (2002), defined as:

$$ILLIQ = \frac{1}{D} \sum_{t=1}^{D} \frac{|R_t|}{VOLD_t} \tag{16}$$

where D is the number of days considered, $|R_t|$ is the daily absolute return of the security and $VOLD_t$ is the respective daily trading volume in dollars.

In this work, as in Ranaldo, Santucci de Magistris (2022), a realized version of the Amihud's measure is used. For a daily illiquidity estimator ($D = 1$), using the classical definition of Amihud's measure could lead to an underestimation of the illiquidity measure: this because the Amihud's measure numerator is represented by the absolute value of the daily return of the security, $|R_t|$, which is computed using the close prices in $t-1$ and t, thus not taking

13 "[...] an abnormal increase in the number or in the intensity of causal relationships, compared with that of tranquil period, triggered after an endogenously detected shock", Gómez-Puig, Sosvilla-Rivero (2014), Page 14.

14 in the sense of Amihud (2002) illiquidity measure, capable of measuring the price impact of trading volumes.

into account the daily volatility of the security. To avoid this possible underestimation, the original numerator has been replaced by a daily volatility measure.
The ratio of a volatility measure and trading volume represents a measure belonging to the wide family of "Volatility-over-Volume" (VoV) indicators, defined by Fong et al. (2018) as a class of liquidity proxies that present the following form:

$$VoV = \frac{a\sigma^b}{V^c} \tag{17}$$

where *a, b, c* are strictly positive. Here, the volatility statistic (σ) is expected to capture the daily volatility of the security. For this reason, the choice fell on Parkinson's high-low range (Parkinson, 1980)[15].

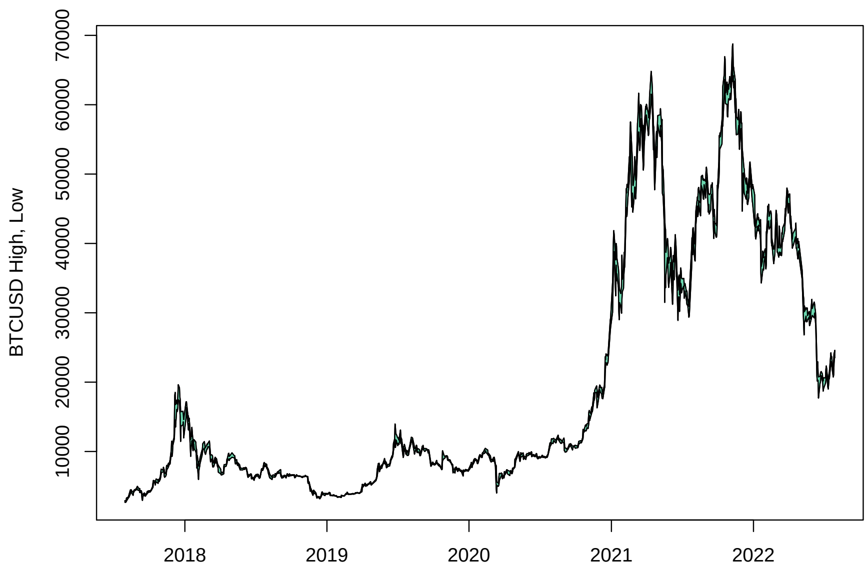

Figure 2. an example of the effect captured by Parkinson's Volatility: the high-low daily volatility is represented by the shaded area.

For a time range of $T = 1$, thus indicating daily volatility, Parkinson's measure is defined as follows:

$$\tilde{\sigma_{it}}^2 = \frac{1}{4\log 2}(p_{it}^{high} - p_{it}^{low})^2 \tag{18}$$

where p_{it}^{high} and p_{it}^{low} denote, respectively, the maximum and minimum log price of security

[15] this choice is consistent with Barigozzi, Brownlees (2018).

i on the day t.

Given this definition of volatility, the VoV measure used for the empirical analysis is defined as:

$$VoV_{it} = \frac{\tilde{\sigma_{it}}^2}{V_{it}} \tag{19}$$

where $\tilde{\sigma_{it}}^2$ is the realized Parkinson Volatility for security i at time t, and V_{it} is the respective daily trading volume. Note that, with respect to Fong et al. (2018) definition, for this measure results that $a = 1, b = 2$ and $c = 1$[16]. This metric is the one used for the first contagion analysis presented later.

In order to have a more precise result, a similar analysis with intraday data would have been carried out. Unfortunately, due to unavailability of intraday volumes data, it was not possible to compute the same metric above described. Still, since an intraday analysis can give better insights about how the two markets are causally interrelated, a single exercise is then performed using log-returns as the contagion metric, where log-returns are defined as:

$$r_{i,t} = log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \tag{20}$$

P_i being the prices of security i at times $t, t-1$.

2.2 Data Selection

Since two different analyses are performed, two different datasets are considered: one composed of daily data, whose time range is 5 years, from 31/07/2017 to 31/07/2022, thus leading to a panel of 1080 observations; and a second one consisting of intraday 5-minutes data, whose time range is 10 days, from 15/03/2020 to 25/03/2020, thus leading to a panel of 2710 observations, not taking into account market closures among different markets. This data has been synchronized using GMT as the common timestamp. The rationale for the time window selection is the desire to analyse contagion dynamics during a time frame characterized by a high level of stress on financial markets. In particular, the selected time window is around the 16/03/2020, day referred to as "Black Monday II"[17]. During that day, due to the COVID-19 pandemic outbreak, most of the global financial markets reported losses of around

[16] a value of $a = 1$ is consistent with the findings of Fong et al. (2018) for their $VoV(\lambda)$. Note as well that the specification used here is just the square of this latter measure. In fact, Fong et al. $VoV(\lambda)$ is defined as:

$$VoV(\lambda) = \frac{a\sigma}{V}$$

where $a = 1$.

[17] from *"2020 stock market crash"*, Wikipedia, last edited on 3^{rd} Sep. 2022.

12-13% (as an example, the Dow Jones Industrial Average Index registered a loss of -12.93%).

Data for stock markets is represented by a selection of global stock indices. A specific panel is therefore composed by high and low prices and trading volumes for the daily resolution dataset; by last prices for the intraday dataset. Note that, for stock indices, trading volume refers to the sum of the trading volumes in dollar of all the components of a single index, while prices are differently computed depending on the specific index, starting from the components prices. This choice was made in order to have the widest possible indicators for the broader stock market[18]. Also, where possible, Tech-specific sector indices have been used. This comes from the fact that, according to various sources[19], it seems that Bitcoin's price moves in the same way tech stocks move.
For the two datasets, a different series of stock markets data has been considered. For the 5-year dataset, due to sparsity and multicollinearity problems[20], only a selection of 7 stock indices was made, representing different geographic areas. Instead, for the high-frequenxy dataset, a collection of 29 stock indices was selected. Stock indices data comes from Bloomberg for the 5-year dataset, and Bloomberg and Barchart Premier for the high-frequency one. The selected indices are summarized below:

Regarding cryptocurrency data, different options were possible. Nowadays, a wide variety of cryptocurrencies exist, but it has to be underlined that a great number of these are projects born due to the recent years' euphoria around crypto-assets in general, and some of these are traceable to scams or fraud attempts; thus choosing the correct crypto-asset is crucial. Bloomberg has defined, in May 2018, the "Bloomberg Galaxy Crypto Index" (BGCI) which serves as a capped market capitalization-weighted performance indicator for the largest cryptocurrencies traded in US Dollars[21]. But, for the purpose of this work, this said indicator presents two issues: first, no data is available before its launch date (3^{rd} May, 2018); second, to date Bitcoin's market capitalization is more than double of Ethereum's one[22], while in the BCG Index they are equally weighted[23]. This difference in weighting could have led to a distorsion in results. Moreover, as denoted by many (for example, Katsiampa et al. (2022)), interconnectedness among different crypto-assets is very high. For this reason, paired to the

[18] alternatively, ETFs replicating the stock indices could have been used. However, this choice would have led data to be influenced by passive investors only.
[19] for example, the arcticle *"Bitcoin Is Increasingly Acting Like Just Another Tech Stock"*, The New York Times, 11^{th} May 2022.
[20] see Subsection 2.4.1 for further details.
[21] source: Bloomberg LP, Bloomberg Digital Asset Indices.
[22] according to Coinmarketcap.com, BTC's market capitalization stands around 319B USD, while ETH's one waves around 148B USD, as of 18^{th} Nov. 2022.
[23] each of them accounts for 30% of the index total weight (source: BCGI Factsheet, Feb. 2020)

Daily dataset		
Variable	Description	Source
F3TECHS	UK 350 Tech. Index	Bloomberg
FCHI	France 40	Bloomberg
GDAXI	Germany 30	Bloomberg
N225	Nikkei 225	Bloomberg
NDXT	Nasdaq-100 Tech. Index	Bloomberg
NIFTYIT	Nifty IT Index	Bloomberg
OSPTX	S&P/TSX Index	Bloomberg

5 min dataset		
Variable	Description	Source
AEX	Amsterdam Exchange	Bloomberg
ATX	Austrian Traded Index	Bloomberg
BEL20	Belgium 20	Bloomberg
BIST100	Istanbul 100	Bloomberg
BUX	Budapest Stock Exchange	Bloomberg
CAC40	France 40	Bloomberg
DAX	Germany 30	Bloomberg
FTSE100	UK 100	Bloomberg
FTSEMIB	Italy 40	Bloomberg
HSI	Hang Seng Index	Bloomberg
IBEX35	Spain 35	Bloomberg
JCI	Jakarta Composite Index	Bloomberg
KOSPI	Korea Composite Index	Bloomberg
MOEX	Russia Index	Bloomberg
NKY	Nikkei 225	Bloomberg
NSEIT	Nifty IT Index	Bloomberg
OMX	Stockholm 30	Bloomberg
SASEIDX	Tadawul All Share	Bloomberg
SMI	Swiss Market Index	Bloomberg
SPASX200	S&P/ASX 200	Bloomberg
SSE	Shanghai Composite Index	Bloomberg
TA35	Tel-Aviv 35	Bloomberg
TWSE	Taiwan Stock Index	Bloomberg
WIG	Warsaw Stock Exchange	Bloomberg
DJ	Dow Jones	Barchart Prem.
NDXT	Nasdaq-100 Tech. Index	Barchart Prem.
PSI20	Lisbon 20	Barchart Prem.
SPTSX	S&P/TSX Index	Barchart Prem.
SPX	S&P500 Index	Barchart Prem.

TABLE I. Presentation of the two Stock Indices datasets.

fact that Bitcoin is the first born, most traded and with highest market capitalization cryptocurrency, it was decided to summarize the cryptocurrencies market in the Bitcoin itself. Thus, data for the BTCUSD rate was obtained from CoinCodex (for the 5-year daily data) and Barchart Premier (for the 5-minutes data).

2.3 NETS Algorithm

To estimate causal interconnectedness a network-based approach is used, as described in Section 1.2. In particular, the Network Estimation for Time Series (NETS) algorithm from Barigozzi, Brownlees (2018) was selected. In the following sections, a summarized description of the model and the estimation algorithm are presented[24].

2.3.1 Model Description

With the NETS algorithm, the authors presented a novel network estimation technique for the analysis of multivariate time series. In particular, the objective is to model a panel of time series as a vector autoregression, with the assumption that the autoregressive matrices and the inverse covariance matrix of the system innovations are assumed to be sparse, i.e.

[24] The two subsequent sections take directly from Barigozzi, Brownlees (2018).

with most of their elements equal to zero. The resulting system can be represented by two different kind of graphs: a directed graph, representing predictive Granger relations smong the time series; and an undirected graph, representing their contemporaneous partial correlations. In both cases, the vertices of the graph represent the variables in the panel, and an edge between two vertices represents some kind of interrelation.

To estimate the model, the authors' innovative NETS algorithm is used. It consists of a LASSO regressions-based procedure, and it is able to contemporaneously estimate the autoregressive matrices and the concentration matrix of the system, instead of splitting the estimation in two different steps. More details on the estimation model are presented in the next section.

For the estimation purpose, the panel data is modeled as the following VAR:

$$y_t = \sum_{k=1}^{p} A_k y_{t-k} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d(0, C^{-1}) \tag{21}$$

where the autoregressive matrices A_k and the concentration matrix C^{-1} are $n \times n$ matrices and, as said before, assumed to be sparse.
To assess dynamic interdependance[25] among time series a multivariate version of the Granger causality notion is used. Formally:

$$E[(y_{it+k} - E(y_{it+k}|\{y_{1t}, ..., ynt\}))^2] = E[(y_{it+k} - E(y_{it+k}|\{y_{1t}, ..., y_{nt}\} \backslash y_{jt}))^2] \tag{22}$$

meaning that y_{jt} does not Granger cause y_{it} if adding the former as a predictor does not improve the mean squared forecast error of y_{it+k} for any $k > 0$. This implies that, if $a_{k(i,j)} = 0 \ \forall k, \ a_k \in A_k$, then y_{jt} does not Granger cause y_{it}. Note that this approach allows to spot contagion dynamics even in case these are characterized by lags, as pointed out in Section 1.1 talking about contagion measurement issues.

Based on this relations, a network is then produced. The Granger causality network is defined as a graph $N_G = (V, E_G)$, V being the set of vertices $\{1, ..., n\}$ and E_G the set of edges, subset of $N \times N$ such that the pair $(i, j) \in E_G$ if and only if i and j are linked by an edge, i.e. are Granger causality-related. Since the Granger network is a directed network, the presence of an edge from i to j means that i Granger causes j in the sense of Equation

25 For the purpose of this work, we are only going to consider dynamic interdependance. For details on contemporaneous interdependance, here not specified for brevity, see Barigozzi, Brownlees (2018)

22. In formal notation:

$$E_G = \Big\{(i,j) \in V \times V \iff a_{k(i,j)} \neq 0, for\ at\ least\ one\ k \in \{1,..,p\}\Big\} \tag{23}$$

2.3.2 Model Estimation

Goal of the estimation is to find all the non-zero entries of each autoregression matrix, A_k, and of the concentration matrix C. Since the direct estimation on the latter from the residuals of the former has some issues[26], authors created a procedure to estimate both sets of parameters jointly. For this purpose, the VAR model is re-parametrized as follows:

- a n^2p-dimensional vector α containing a set of coefficients α_{ijk}, corresponding to autoregressive coefficients a_{kij};
- a $n(n-1)/2$-dimensional vector ρ containing a set of partial correlations[27] ρ^{ij};
- a n-dimensional c vector containing a set of c_{ij} coefficients, corresponding to the diagonal of the concentration matrix C.

The resulting VAR and contemporaneous equations are the following:

$$y_{it} = \sum_{k=1}^{p}\sum_{j=1}^{n} \alpha_{ijk} y_{jt-k} + \epsilon_{it}, \quad i = 1,..,n \tag{24}$$

$$\epsilon_{it} = \sum_{h=1}^{n} \rho^{ih} \sqrt{\frac{c_{hh}}{c_{ii}}} \epsilon_{ht} + u_{it}, \quad i = 1,..,n \quad h \neq i \tag{25}$$

From these two equations it is possible to derive a LASSO-type estimator for the parameters of the model. Indeed, y_{it} can be written as a function of the lags of the time series and the contemporaneous representation of all the other series (regression representation):

$$y_{it} = \sum_{k=1}^{p}\sum_{j=1}^{n} \beta_{ijk} y_{jt-k} + \sum_{h=1}^{n} \gamma_{ih} y_{ht} + e_{it}, \quad h \neq i \tag{26}$$

e_{it} being an estimation error term. This equation can be rewritten with β_{ijk} and γ_{ih} expressed as a function of α_{ijk}, ρ^{ih} and c_{ii} parameters:

$$y_{it} = \sum_{k=1}^{p}\sum_{j=1}^{n} \left(\alpha_{ijk} - \sum_{l=1}^{n} \rho^{il} \sqrt{\frac{c_{ll}}{c_{ii}}} \alpha_{ljk}\right) y_{jt-k} + \sum_{h=1}^{n} \rho^{ih} \sqrt{\frac{c_{hh}}{c_{ii}}} y_{ht} + u_{it}, \quad l,h \neq i \tag{27}$$

[26] see Barigozzi, Brownlees (2018) for further details.
[27] defined as $\rho^{ij} = corr(\epsilon_{it}, \epsilon_{ij} | \{\epsilon_{kt} : k \neq i,j\})$

where

$$\beta_{ijk} = \alpha_{ijk} - \sum_{l=1}^{n} \rho^{il} \sqrt{\frac{c_{ll}}{c_{ii}}} \alpha_{ljk}$$

$$\gamma_{ih} = \sum_{h=1}^{n} \rho^{ih} \sqrt{\frac{c_{hh}}{c_{ii}}}.$$

Note that errors e_{it} and u_{it} are the same.

Denoting θ the vector of parameters of interest $(\alpha', \rho')'$ of dimension $m = n^2 p + n(n-1)/2$, the following quadratic loss function can be associated to the estimation problem, conditional on c:

$$\mathcal{L}(\theta; y_t, c) = \sum_{i=1}^{n} \left(y_{it} - \sum_{k=1}^{p} \sum_{j=1}^{n} \left(\alpha_{ijk} - \sum_{l=1}^{n} \rho^{il} \sqrt{\frac{c_{ll}}{c_{ii}}} \alpha_{ljk} \right) y_{jt-k} + \sum_{h=1}^{n} \rho^{ih} \sqrt{\frac{c_{hh}}{c_{ii}}} y_{ht} \right)^2 \quad (28)$$

Thus, considering a sample of T observations for y_t, the LASSO-type estimator for θ is the following:

$$\hat{\theta}_T = \underset{\theta \in \mathbb{R}^m}{argmin} \left[\frac{1}{T} \sum_{t=1}^{T} \mathcal{L}(\theta; y_t, \hat{c}_T) + \lambda_T^G \sum_{k=1}^{p} \sum_{i,j=1}^{n} \frac{|\alpha_{ijk}|}{|\tilde{\alpha}_{Tijk}|} + \lambda_T^C \sum_{l,h=1}^{n} \frac{|\rho^{lh}|}{|\tilde{\rho}_T^{lh}|} \right], \quad l > h \quad (29)$$

where λ_T^G and λ_T^C are the LASSO tuning parameters and $\tilde{\alpha}_{Tijk}, \tilde{\rho}_T^{lh}$ and $\hat{c}_T$ are pre-estimators for α, ρ and c coefficients. With T sufficiently large, a pre-estimator for α is the least squares estimator of the VAR autoregressive matrices, while a pre-estimator for ρ is the partial correlation estimator obtained from the sample covariance of the VAR residuals. Moreover, since autocorrelation is supposed in and across y_t components, an adaptive LASSO penalty is adopted.

Since Equation 28 is not the standard quadratic loss function of a linear regression model, the standard LASSO algorithms cannot be applied. Thus, the authors propose an iterative coordinate descent algorithm to minimize the objective function. This is the NETS algorithm. For brevity, I will avoid to report the detailed calculations here[28]. Intuitively speaking, each iteration updates one component of the parameters vector $\theta = (\alpha', \rho')'$, and a residual estimate is computed. In each iteration, the parameter is updated using a combination of two different set of variables: a first set of regressors corresponding to the coefficient being updated; and a second vector containing the partial residuals of the model with respect to all

[28] an extensive mathematical derivation of the procedure, with its theoretical foundations as well as an argumentation on estimation and selection consistency matters, can be found in Barigozzi, Brownlees (2018). Here it is sufficient to report that, after a simulation study conducted by the authors, the procedure "performs satisfactorily, and that the gains with respect to the traditional estimator can be large for sparse VAR systems.".

the parameters besides the coefficient being currently updated. The algorithm then proceeds with the updates until convergence, which is checked at the end of each full cycle of updates of θ. The first full cycle is based on a pre-estimation of c; then, after an estimation of θ is available, the estimation of c is updated as well. This two last steps are then iterated until convergence, and the result is the full set of $\hat{\theta}_T$.

2.4 Analysis Details and Results

In this section, a description of the analysis process as well as empirical results are given. These are split in two different subsections, each regarding a unique analysis on one of the two aforementioned datasets.

2.4.1 Daily Data Analysis

For this first analysis, the daily data dataset is used. First of all, a synchronization of data has been performed. Since Bitcoin is traded every day without interruptions, it was necessary to take into account weekends and other market closure days for the selected stock exchanges. Thus, all data has been joined on BTCUSD data and NAs have been omitted[29].
Then, data has been split in two different categories: "analysis" data, comprehending daily high prices, daily low prices and daily volumes for BTCUSD, NDXT, FCHI, GDAXI, N225, F3TECHS, OSPTX, NIFTYIT; and "market" data, comprehending daily high prices, daily low prices and volumes for XLK. The reason for this separation sits in the sparsity hypothesis of the NETS algorithm, which asks for the data to be sparse and, thus, to be cleansed from common factors. This step will be described later in this section. For each subset of data, Parkinson Volatility has been computed according to Equation 39 in Subsection 2.1. Then, the liquidity VoV measure has been calculated, as in Equation 19. Descriptive statistics for the VoV measure[30] as well as a simple correlation analysis can be found below.

As it is possible to see in Figure 3, a medium-high degree of positive correlation is shown among different stock indices series, reaching the value of 0.90 between FCHI and GDAXI. For this reason, as mentioned in Section 2.2, it was not possible to select more indices for the analysis: the NETS algorithm would have not be able to compute the concentration matrix C, being the error term matrix singular. For what concerns BTCUSD instead, it presents small negative correlation with almost all the stock indices.

[29] it was here decided not to interpolate data in order to avoid an undesired smoothing effect.
[30] for descriptive statistics computation, the VoV measure has been multiplied for a factor of 10^{12}.

VoV measure descriptive statistics								
	Variables							
	BTCUSD	NDXT	FCHI	GDAXI	N225	F3TECHS	OSPTX	NIFTYIT
Mean	0.17	1.48	0.79	0.87	0.07	16.11	484.02	4.17
SD	0.50	1.71	1.01	1.07	0.09	30.50	656.65	8.04
Median	0.02	0.86	0.46	0.53	0.04	9.83	236.06	2.78
Min	0.00	0.07	0.04	0.03	0.00	0.58	4.11	0.14
Max	7.57	16.28	9.00	11.09	1.14	782.88	6377.16	230.37
Skewness	7.32	3.16	3.83	4.05	4.74	16.01	3.67	21.09
Kurtosis	76.13	15.84	20.21	24.58	34.24	373.67	19.26	577.92

TABLE II. Descriptive statistics for the VoV "analysis" data.

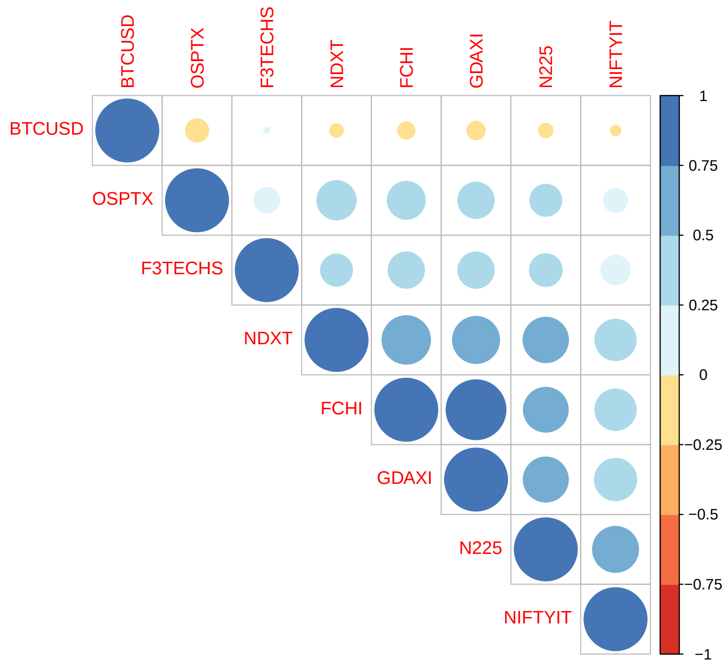

Figure 3. Correlation heatmap for the VoV "analysis" data.

As it was mentioned before, the main NETS algorithm working assumption is the underlying sparsity of the model. But, as it can be seen by the statistics reported in the table above, the system is definitely not sparse. For this reason, data must be treated in a proper way, i.e. common factors must be removed[31]. For this goal, a preliminary regression is specified as follows:

$$VoV_{i,t} = \alpha_i + \beta_i VoV_{M,t} + \varepsilon_{i,t} \tag{30}$$

where VoV_i denotes a list of VoV measures for the time series i in the "analysis" dataset and VoV_M denotes the same measure for a list of market-wide factors.

[31] note that this is the same treatment carried out by Barigozzi, Brownlees (2018) in their application of the NETS algorithm.

It is important to note that this preliminary regression should depurate data from common market-wide factors and common sector-specific factors. For this purpose, two different factors were selected: a VoV measure constructed from data of S&P500 index (SPX), and a VoV measure constructed on the SPDR Technology sectorial index (XLK). After constructing the two indicators, different preliminary sets of regressions have been tested: a regression set with both SPX and XLK factors; a regression set with SPX only; and a regression set with XLK only. Analyzing regression results, the choice fell on the regression including the XLK factor only, for this factor alone is the one that captures most of the models variance. Note that the regression including both factors was affected by multicollinearity, since the SPY and XLK factors have a strong degree of correlation during the selected time frame. Thus, the chosen preliminary set of regressions is specified as follows:

$$VoV_{i,t} = \alpha_i + \beta_i VoV_{XLK,t} + \varepsilon_{i,t} \tag{31}$$

Of this series of regressions, estimated with OLS method, residuals have been stored. These residuals are the one fed to the NETS algorithm in order to estimate the model. Descriptive statistics for the residuals as well as their charts are presented below:

Residuals descriptive statistics								
	Variables							
	BTCUSD	NDXT	FCHI	GDAXI	N225	F3TECHS	OSPTX	NIFTYIT
Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SD	0.50	0.95	0.83	0.89	0.08	29.77	597.88	7.26
Median	-0.14	-0.10	-0.16	-0.17	-0.02	-4.66	-159.02	-0.44
Min	-0.22	-4.01	-3.77	-3.94	-0.27	-49.84	-2254.25	-21.56
Max	7.35	9.47	5.63	8.00	0.73	768.41	5013.42	195.85
Skewness	7.33	2.08	2.28	2.58	3.49	16.87	3.31	18.40
Kurtosis	76.56	16.90	11.60	15.82	21.93	412.49	18.27	489.67

TABLE III. Descriptive statistics for the preliminary regression residuals data.

It is useful to note that, although a common factor has been removed, some kind of variability persists in residuals, whose dynamics will have to be explained through the NETS algorithm. Moreover, something interesting can be denoted from the BTCUSD residuals chart (Figure 4): after a certain point in time, residuals start to be much lower in absolute value[32] becoming comparable to residuals of the two Technology indices here considered, F3TECHS and NIFTYIT (besides for some outliers). This means that the BTCUSD series, after a certain date, can more accurately be explained by the XLK factor, i.e. BTCUSD

[32] note that the VoV measure, being a ratio of two positive numbers, is always a positive number.

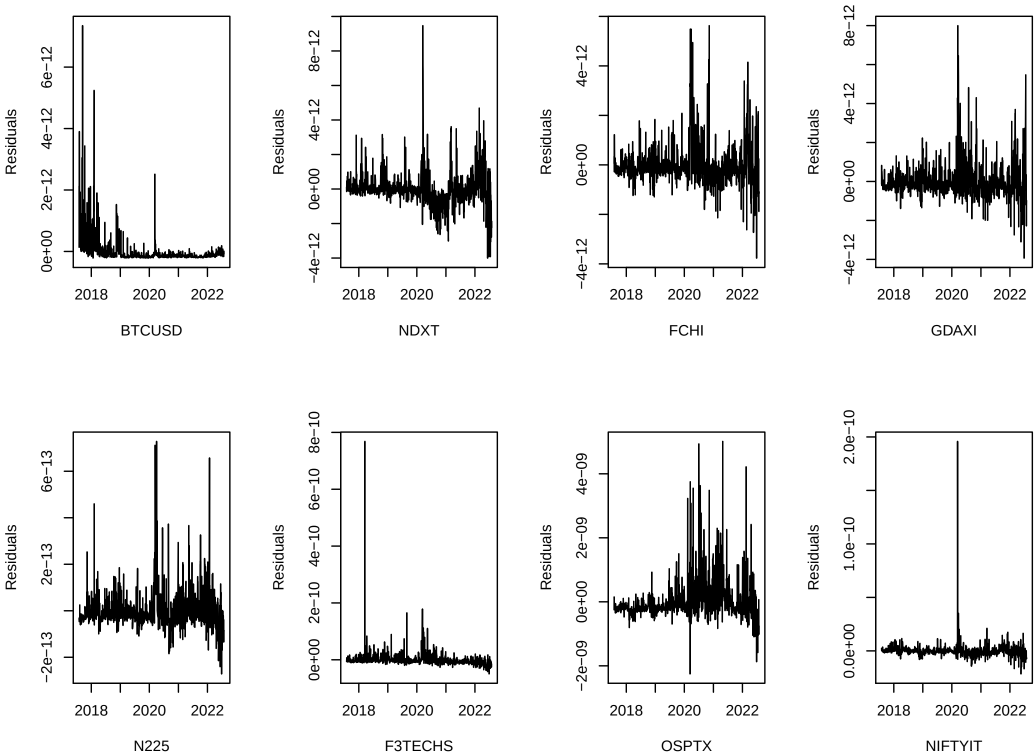

Figure 4. Preliminary regression residuals for the "analysis" data.

behaves almost in the same way the Technology market behaves, as hypothesized in Section 2.2. This result is consistent with the findings of Matkovskyy, Jalan (2019). In their work, the two authors studied the contagion effect between financial markets and Bitcoin market with co-movements in asset returns, and their result showed contagion between the two, in contrast with the prevailing literature. Their explanation, supported by some research articles[33], was the introduction of Bitcoin futures on 18^{th} December 2018 which, according to the literature, improved efficiency in the Bitcoin market. Among authors, Hale et al. (2018) denoted that the introduction of Bitcoin futures led investors to behave in the same manner following the introduction of futures in asset markets: with this kind of contracts it was possible for investors with pessimistic views on Bitcoin to short-sell the cryptocurrency, causing its price to fall. This is what was observed on the market by the authors, with a sharp drop in BTC price after the launch of futures contracts on the said date.

This set of data is then fed as input to the NETS algorithm. For what is said before, it is useful to check for contagion from the date of introduction of Bitcoin futures, i.e. 18^{th} De-

[33] see Matkovskyy, Jalan (2019) for further details on according and contrasting literature.

cember 2018, corresponding to the 304^{th} observation in the sample out of 1080 observations total, thus feeding to the NETS algorithm a panel made of 8 variables and 777 observations. Other inputs for the algorithm are: a $\lambda(\lambda_1, \lambda_2)$ set of parameters, the LASSO tuning parameters described in Section 2.3.2, set as $\lambda(0.5, 0)$[34]; a p parameter, indicating the number of lags underlying the VAR model, set in three different values ($p = 1, p = 5, p = 10$) for three different model estimations; and a `iter.it` parameter, for the maximum number of iterations allowed, set to 200. The resulting Granger Network graphs are presented in Figures 9, 10, 11 in the Appendix. It should be no surprise that not all the stock markets are related: the analysis is focused on causal relationships, but this does not mean that there may have been co-movements between non causally-related time series. It is important to remember that we took into account a market common factor as well, thus removing a great part of co-movements. Below, the matrices associated with the graphs are presented. Note that these are the three different E_G matrices described in Subection 2.3.1, each for a single model estimation with a unique lag p, where a value of 1 determines the presence of and edge **from** x_i (row i) **to** x_j (column j):

$$\begin{array}{r} BTCUSD \\ NDXT \\ FCHI \\ GDAXI \\ N225 \\ F3TECHS \\ OSPTX \\ NIFTYIT \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Granger Network matrix for $p = 1$

$$\begin{array}{r} BTCUSD \\ NDXT \\ FCHI \\ GDAXI \\ N225 \\ F3TECHS \\ OSPTX \\ NIFTYIT \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Granger Network matrix for $p = 5$

$$\begin{array}{r} BTCUSD \\ NDXT \\ FCHI \\ GDAXI \\ N225 \\ F3TECHS \\ OSPTX \\ NIFTYIT \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Granger Network matrix for $p = 10$

Figure 5. Granger Network adjacency matrices for different specifications of the model. Variables in matrices columns are in the same order as in matrices rows, from left to right.

Since the contagion metric here used is a liquidity (price impact) metric, each edge can be interpreted as a price impact movement in a market that causes other price impact movements

[34] note that these values of λ are the one that minimize the RSS.

in other markets. In other words, the liquidity drainage that happens on a single market on a specific day affects, after 1, 5, or 10 trading days, the drainage of liquidity in some other markets, as drawn by the three graphs. Looking at both graphs and matrices, it is possible to understand the evolution of such relationships with the passage of time. In particular, there are three stock market indices that causes liquidity movements in all the other vertices: the FTSE 350 SuperSector Technology Index (F3TECHS), the S&P/TSX Composite Index (OSPTX), and the Nifty IT Index (NIFTYIT). In the first plot (Figure 9) it is possible to see that, after one trading day, these three markets are interrelated in the Granger Causality sense, and they generate liquidity spillovers on the other markets, including Bitcoin (BT-CUSD) which is Granger-caused by OSPTX. In the second plot (Figure 10), after 5 trading days, interrelations find a new asset but the only market causing spillovers on BTCUSD is OSPTX still. Finally, in the third plot (Figure 11) it is possible to note that, after 10 trading days, both OSPTX and F3TECHS are responsible for spillovers on BTCUSD. Note that, as testified by the graph matrices, with the passage of time the overall system tends to be more interrelated, with the matrix $E_{G,p=10}$ having more non-zero elements than both $E_{G,p=1}$ and $E_{G,p=5}$. Also, imposing $p > 10$ did not produce different results compared to Figure 11. Thus, this last plot can be considered the final liquidity network for the considered variables. Below a summary table of the three models:

Daily models specifications			
Model	1	2	3
p	1	5	10
RSS	3.369048	3.144677	2.998165
N. Parameters	43	83	133

TABLE IV. models specification and statistics summary.

In conclusion, considering liquidity movements measured with the described proxy, the daily empirical estimation for the selected time period suggests that BTCUSD is exposed to spillovers from stock markets, in particular from the english Technology stock market and the Canadian stock market. This result can be compared to the results from Saiedi et al. (2021)[35]. With their work, the authors investigated the worldwide spread of infrastructures needed to maintain of Bitcoin as a system and its growth, i.e. Bitcoin Nodes, and infrastructures for the effective usage of Bitcoin as a mean of payment, i.e. Bitcoin merchants. Their result can be summarized with the following maps:

[35] even if data from this paper may be outdated, the quality of the study is desirable for a meaningful

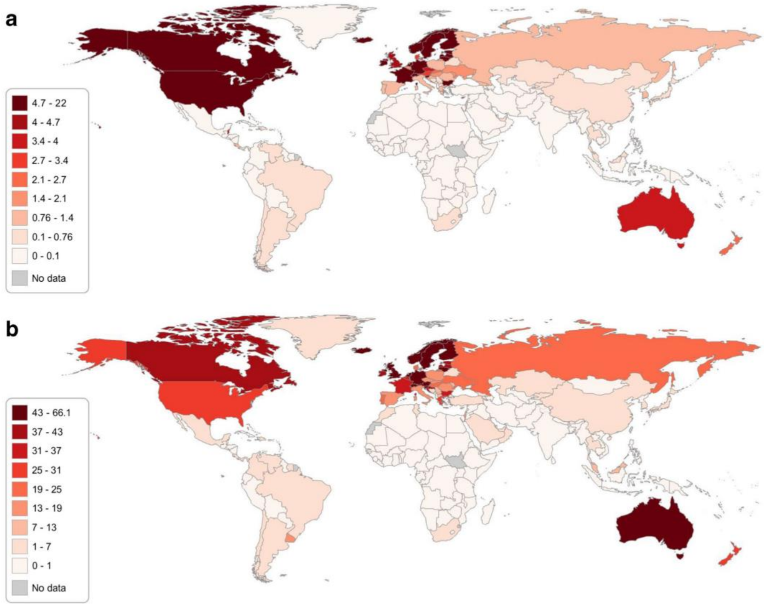

Figure 6. **a.** Global map of *bitnode intensity*[36]. **b.** Global map of *unique bitnodes* in a country. Image from Saiedi et al. (2021).

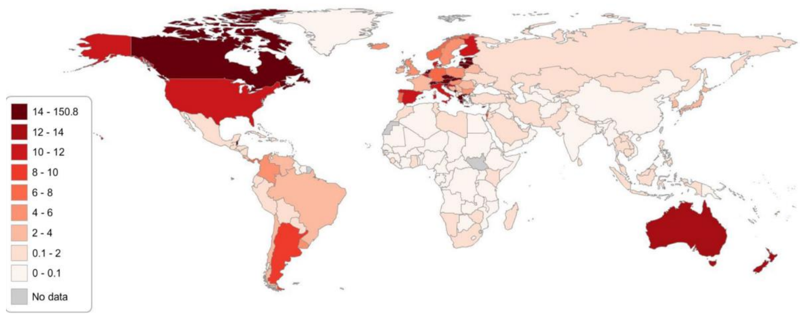

Figure 7. global map of total *Bitcoin merchants* per million users. Image from Saiedi et al. (2021).

As the maps show, Canada seems to have top scores regarding *bitnode intensity* and *Bitcoin merchants* metrics and a very high score regarding *unique bitnodes* count; UK has a top score in *unique bitnodes* and a high score in *bitnode intensity*, but a much lower score in terms of *Bitcoin merchants*. These metrics may then give a preliminary explanation of my results; anyway, the reasons behind these findings are beyond the scope of this work. Still, I can conclude that a causal relation between stock markets and Bitcoin market exists in terms of liquidity, and thus the research hypothesis is confirmed.

comparison. For more details, see Saiedi et al. (2021).

[36] *Bitnode intensity* has been measured as the number of active nodes in a country multiplied the number of hours of activity of each node. For more details, see Saiedi et al. (2021).

2.4.2 Intraday Data Analysis

For this second analysis, the intraday (5 minutes) dataset is used. Also in this case, a synchronization of data has been performed, taking into account common market closures and omitting those observations. Then, as before, an “analysis” dataset and a “market” dataset are created, the latter containing data from the S&P500 Index (SPX) while the former containing all the remaining data. As in the previous analysis, the “market” data will serve to cleanse data from common factors. Here, data to compute Parkinson Volatility was not available, since 5-minutes data only comes with a Last Price value. For this reason, 5-minutes log-returns have been computed for each time series, both in the “analysis” and in the “market” datasets, as described in Section 2.1. These returns will be the starting data for the analysis.

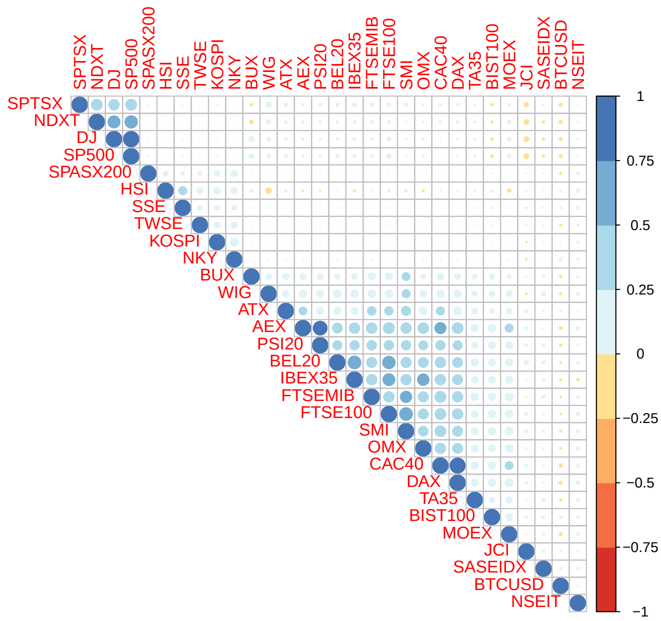

Figure 8. Correlation heatmap for log-returns.

A preliminary correlation analysis on returns can be found in Figure 8 below. As shown in the figure, there seems to be some kind of correlation in log-returns for indices belonging to the same geographic area.

Next, to ensure sparsity of data for the NETS algorithm assumptions, also in this analysis a preliminary set of regressions is performed on data. The general preliminary regression model is specified as follows:

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t} \tag{32}$$

where r_i is the series of returns for variable i and r_M indicates a collection of series of returns from common factors for the "analysis" data. Here, the same multicollinearity problem presented, so different specifications were tested: a first specification with S&P500 Index as the only market factor; a second specification with XLK; and a third specification with both factors. The specification that was able to capture most of the variance was the one with the SPX factor only. Thus, the final specification of the preliminary regressions model is the following:

$$r_{i,t} = \alpha_i + \beta_i r_{S\&P500,t} + \varepsilon_{i,t} \tag{33}$$

Of this set of regressions, estimated with OLS, residuals have been stored: these are the data fed to the NETS algorithm. Reporting here complete descriptive statistics for residuals is superfluous. It has to be noted that, as in the previous analysis, variability persists among data even after correcting for common factors.

This dataset is then fed to the NETS algorithm. The other set of NETS parameters are: a $\lambda(\lambda_1, \lambda_2)$ set of parameters set as $\lambda(0.5, 0)$[37]; the p parameter set in three different values ($p = 1, p = 2, p = 6$)[38] for three different model estimations; the `iter.it` parameter set to 200. The resulting Granger Network graphs are presented in Figures 12, 13, 14 in the Appendix. Here, reporting the complete E_G matrices would be superfluous, as there are many relations between all the variables of the system. For this work purpose it is useful to report BTCUSD interactions with other indices. A summary table with sub-matrices from $E_{G,p=1}, E_{G,p=2}$ and $E_{G,p=6}$ is presented in Table V.

Here, the contagion metric used is securities log-returns. In this case, edges can be interpreted as a causal relation in return movements from a variable to another. As the graphs show, during the so-called "Black Monday II" the overall interconnectedness between markets has been high in the number of causal relations. The only market that seems non-responsive within a 5-minute lag period is the S&P/ASX 200 (SPASX200), being the only vertex not connected to the system in Figure 12. With the passage of time, as in the previous analysis, system's interconnectedness increases, w.r.t. Figures 13 and 14. Looking at both graphs and Table V, it is clear that, in this particular event, BTCUSD had a significant impact on the overall system: 27 out of 28 markets suffered from BTCUSD spillover effects, with the AEX index (AEX) being the only market not influenced by the cryptocurrency return movements in a 30-minute range. It is possible to note that the most influencial market index in the system has been the FTSE/MIB index (FTSEMIB).

[37] note that these values of λ are the one that minimize the RSS.
[38] note that, in this case, each lag corresponds to a 5-minute lag in the time series. Thus, $p = 1$ is a 5 minutes lag, $p = 2$ is a 10 minutes lag and $p = 6$ is a 30 minutes lag.

Spillover effects of BTCUSD						
	BTCUSD to x			x to BTCUSD		
	$p = 1$	$p = 2$	$p = 6$	$p = 1$	$p = 2$	$p = 6$
AEX	1	0	0	0	0	0
ATX	0	1	1	1	1	1
BEL20	1	1	1	0	0	0
BIST100	1	1	1	0	0	0
BUX	0	0	1	0	0	0
CAC40	0	1	1	0	0	0
DAX	1	1	1	0	0	0
FTSE100	0	0	1	0	0	0
FTSEMIB	1	1	1	0	0	0
HSI	0	1	1	0	0	0
IBEX35	1	1	1	0	0	0
JCI	0	1	1	0	0	0
KOSPI	0	1	1	0	0	0
MOEX	0	1	1	0	0	0
NKY	1	1	1	0	0	0
NSEIT	0	1	1	0	0	1
OMX	0	0	1	0	0	0
SASEIDX	0	0	1	0	0	0
SMI	1	1	1	0	0	0
SPASX200	0	1	0	0	0	0
SSE	0	0	1	0	0	0
TA35	0	1	1	0	0	0
TWSE	1	1	1	0	0	0
WIG	1	1	1	0	0	0
DJ	1	1	1	0	0	0
NDXT	1	1	1	0	0	0
PSI20	0	0	1	0	0	0
SPTSX	0	1	1	0	0	0

TABLE V. Spillover effects from BTCUSD to market x (left three columns) and from market x to BTCUSD (right three columns).

For what concerns contagion from stock markets to BTCUSD, not many dynamics can be found. In particular, the only two stock indices directly affecting BTCUSD performance have been the Austrian Traded Index (ATX) and the Nifty IT index (NSEIT), the latter showing contagion dynamics only with a 30 minutes lag while the former affecting BTCUSD in each moment of the system. No influence seem to have either the London Stock Exchange or the Canadian Stock Exchange as found in the previous analysis, which maybe need more time to propagate their effects on the cryptocurrency market[39].

Looking at Figures 6 and 7 in the previous section, the metrics proposed by Saiedi et al. (2021) seem to give a good explanation for the ATX influence on BTCUSD, while regarding NSEIT they seem to give little explanation of the phenomenon: in fact, India is in the lower part of the classifications for *bitnode intensity*, *unique bitnodes* and *Bitcoin merchants* all. It can be denoted, however, that the fact that a stock market of a certain country is the cause

[39] note that for $p > 6$ the computation time for estimation starts to be material, due to the high number of parameters of the model.

of spillovers in the global stock market can be unrelated with internal investors behavior: capital outflows may be caused by foreign investors, leading to a significant effect when the total stock detained by foreigners is a large percentage of the total equity listed[40].
Below, a summary of the three model specifications:

Intraday models specifications			
Model	1	2	3
p	1	2	6
RSS	0.0404924	0.03925473	0.03462707
N. Parameters	525	628	999

TABLE VI. model specifications and statistics summary.

Trying to explain this wide set of interrelations can be challenging, and it is out of the scope of this work. Here, it can be concluded that BTCUSD has had causal relationships, both inwards and outwards, with a large number stock markets during "Black Monday II". This outcome confirms the research hypothesis that Bitcoin and stock markets show contagion dynamics between each other.

[40] according to the article *Who owns the Indian stock exchanges?* from MarketToday, 13^th Oct. 2022, "Foreign investors have an overall stake of 44.6% in NSE".

3 Conclusions

My research goal was to test for contagion dynamics between a set of stock markets and Bitcoin, with the definition of contagion being specified in Section 1.1. To accomplish this, I used a Network Estimation tool from Barigozzi, Brownlees (2018), called NETS, that builds a Granger-causality Network starting from panel data, where variables linked by an arrow indicate that x_i directly Granger-causes x_j if and only if there exists a directed edge originated by x_i and pointing towards x_j; while x_i indirectly Granger-causes x_j if and only if there is a set of directed edges that, combining their starting and ending point, have their origin in x_i and their end in x_j. The analysis has been conducted both in terms of liquidity, with a liquidity measure build upon daily data, and in terms of log-returns, calculated with intraday 5-minutes resolution data. Detailed graphs of the generated Granger Networks can be found in the Appendix.

As it is possible to denote from the graphs, my conclusion is that Bitcoin (BTCUSD) shows direct contagion dynamics with various stock markets; in particular, for what concern the liquidity analysis based on daily data, Bitcoin (BTCUSD) is affected by direct contagion spillovers from the S&P/TSX Index (OSPTX) for $p = \{2, 5, 10\}$ and from the UK 350 Technology Index (F3TECHS) for $p = 10$. It is also possible to note that there is a contemporaneous contagion effect betweeen F3TECHS and OSPTX for $p = 10$. For what concerns the log-returns analysis based on intraday data, during the so-called "Black Monday II", Bitcoin (BTCUSD) directly affected a large number of stock markets[41], while it was affected by direct contagion spillovers from the Austrian Traded Index (ATX) for $p = \{1, 2, 6\}$ and from the Nifty IT Index (NSEIT) for $p = 6$. With this evidence, I can conclude that Bitcoin shows direct contagion effects, in a Granger-causal sense, both inwards and outwards; the research hypothesis is then confirmed with a positive outcome.

In general, my conclusions are in line with all the literature that worked on my same research topic and that was produced from the work from Matkovskyy, Jalan (2019) on. In fact, this mentioned work seems to be a pivot point in the literature: before such document, the prevailing results showed no or meaningless existing relations between cryptocurrency markets and stock markets; while research papers produced after this mentioned work showed intensified relations between these two asset classes. Again, as described in Section 2.4.1, the main reason for this behavior switch seems to be the introduction of Bitcoin Futures on 18^{th} December 2018.

[41] see Table V in Section 2.4.2 for details.

More specifically, my conclusions are in line with that from Matkovskyy, Jalan (2019). The authors in their work showed that investors, during risk-off periods, moved from two famous Bitcoin markets (Bitmap USD and GDAX GBP) towards assets listed on two stock markets, i.e. the NASDAQ Index (which is also a technology index) and the NIKKEI225 Index; and also, investors moved from Euro GDP/Bitcoin markets towards European assets. These movements can be interpreted as a "fly to stability" move coming from risk-averse investors: selling Bitcoin in favor of other financial assets would lead to a liquidity drainage in Bitcoin markets, caused by a risk-off period. In this sense, my results from the daily analysis of liquidity drainage in Bitcoin following liquidity drainage in stock markets show the same dynamic highlighted by the two authors. The logical connection between the two results may be the following: the start of a financial crisis causes liquidity drainage from stock markets inducted by the behavior of more risk-averse investors (who would not have any holding in Bitcoin); the same crisis would lead less risk-averse investors (who have holdings in Bitcoin) to move their liquidity from Bitcoin, a more volatile asset, to less volatile assets as the stock markets ones.

My conclusions, with respect to my intraday analysis, are also in line with X. Guo et al. (2021): using a methodology highly comparable to the one I used, the authors find that, from the COVID-19 pandemic outbreak on, Bitcoin showed a rising contagion effect with gold and stock markets. In particular, the authors find that "[...] European market has a dominant role in transmitting risks and information, that it is the net contagion source to Bitcoin and contagion intermediary in the infection chain of US market.". This is comparable to my results, where the predominant stock index in the Granger network graph for the intraday contagion analysis was the FTSE-MIB Index (FTSEMIB), with the difference being that, in my case study, Bitcoin was the origin of contagion towards this market, as described by the Table V in Section 2.4.2. A possible explanation for this difference may be the data length and resolution characteristics, mine being much shorter in length, including just a few days around a specific event, i.e. the "Black Monday II" market crash.

The said switch in Bitcoin's behavior (and, consequently, other cryptocurrencies' one) modifies the possible real-world applications of literature's findings: as an example, before 2019, Bitcoin could have been used for portfolio diversification and optimization purposes in a completely different way it would have to be done today, knowing these new findings. For this reason, it is important to expand the existing literature in this research area, the current one being of limited depth.

3.1 Tips for Further Research

This work provides meaningful insights on contagion dynamics of Bitcoin with respect to the other chosen variables; in this sense, it can be a starting point for future research. There are, by the way, some caveats that need to be addressed in future works.

First of all, the liquidity analysis would be more meaningful with more appropriate data. This would ideally mean building a price impact indicator, similar to the one I used, that possibly captures price impact effects for each trade during the trading day, and also takes into account for after-market trades to avoid different behaviors based on trading hours. In order for this to be possible, high-resolution data with execution price and trade volumes information would have to be available for each trade, for each variable in the system, thus the researcher being able to build a high-frequency price impact measure. While this is currently true for (most) single stock data, it is not true for market indices: these would have then to be reconstructed from each index component's data. Again, using ETFs would not give the correct insights, in my opinion: as pointed out in Section 2.2, investors in ETFs are typically passive, thus reacting slower to mutating market conditions than other investors. Potentially, their share in the ETF could also not be moved during market turmoils, so not figuring in the overall liquidity drainage.

Second, to use the same procedure I used a solid data cleansing process is needed. In this particular case, it would mean to being able to take into account for high interrelation among variables in the system, as they are all market indices that share a great amount of variance. In this sense, a researcher would have to find a series of factors that ensure to capture the most common variance possible among the selected indices, and this task results harder than it looks like: it is crucial, in fact, that the variables' idiosyncratic variance stays untouched, as it is the fundamental information on which the whole procedure is based.

Third, given the relations described through the Granger-causal network, the most natural follow-up would be to find the economic and logic meanings of these relations. This is most likely the hardest task, also due to the fact that cryptocurrencies' environment is continuously developing, then constantly requiring up-to-date data. In Section 2.4 I tried to give a first interpretation of my results, stressing on how Bitcoin and its related technology are present around the world; but many more variables should be taken into account.

A non-comprehensive list of these possible subjects of study would include:

- overall holdings of Bitcoin holders divided by asset class, and their behavior with respect to all their holdings during market turmoil;
- overall status of the Bitcoin blockchain, i.e. miners profit, hash rate, etc.;
- more precise data on Bitcoin's presence in the selected world countries.

Note that some of this data is either not available yet, or it is difficult to process in a way that is useful for a thorough analysis; so a pivotal topic for future research would be how to collect and review this kind of data.

Everything considered, I am firmly convinced of the power of network analysis in econometrics and investment problems, and that it should be used more consistently for various topics. The construction of a relational graph from the variables of a system is a perfect way to highlight relations between those variables, and the contemporaneous use of different graphs based on different relations can drive to deep conclusions on the entire system's behavior; also, a network can more precisely indicate how to build an impulse analysis to conduct stress tests over the selected environment and simulate different scenarios to derive possible backup strategies.

To end, finding relations between economic variables has always represented the greatest and most interesting challenge for financial researchers; trying to explain our complex financial system through mathematical and statistical models revealed a much harder task than expected during the years. Nowadays, sophisticated computing techniques such as Machine Learning and Deep Learning are available, and their first applications in financial markets are seeing light for the first time. Results are promising: computers can find relations that we cannot find, due to our biologic dimensional limitations, and in a much faster way. However, it is critical for the successful usage of such technology to gain a deep understanding of the original problem: a misuse of these advanced tools could lead to enormous losses in all terms. Computers will start to gain a growing degree of autonomy to execute their tasks, and soon end up in a completely autonomous form; but we, as their creators, should not step back from understanding the problems that we ask machines to solve. That, in my opinion, is the only way to grant success from the human-machine collaboration.

Appendix

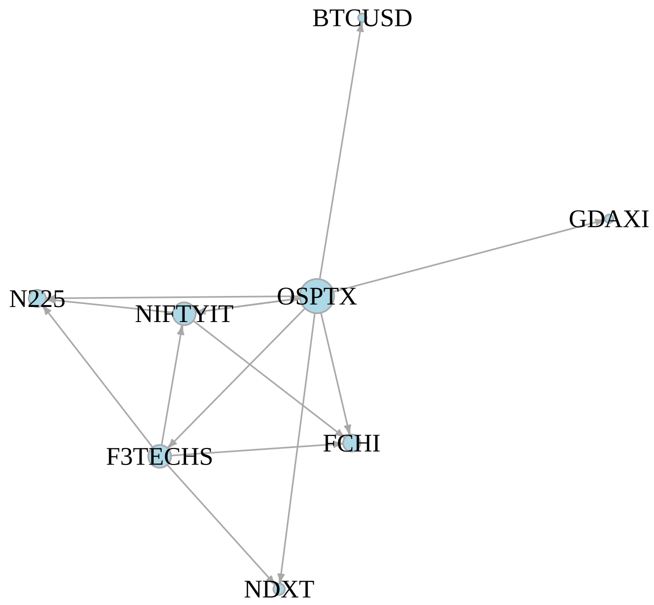

Figure 9. Granger Network estimation, p=1.

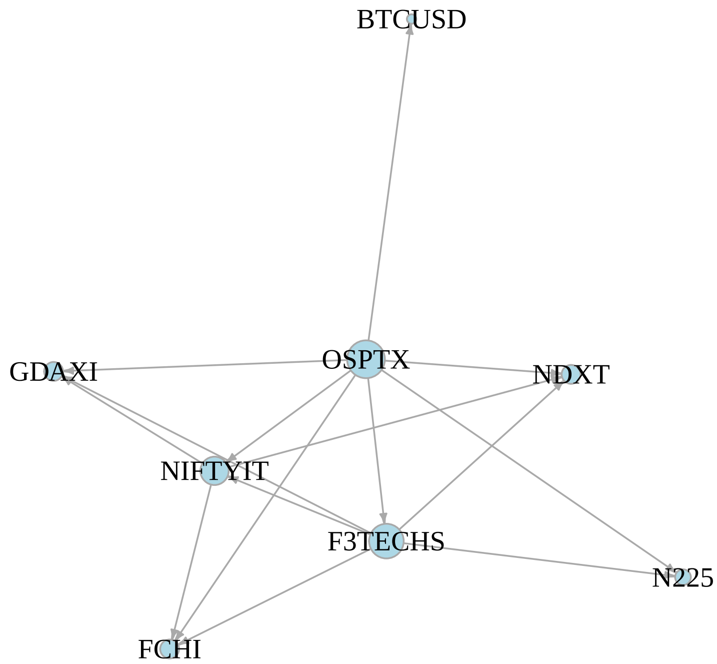

Figure 10. Granger Network estimation, p=5.

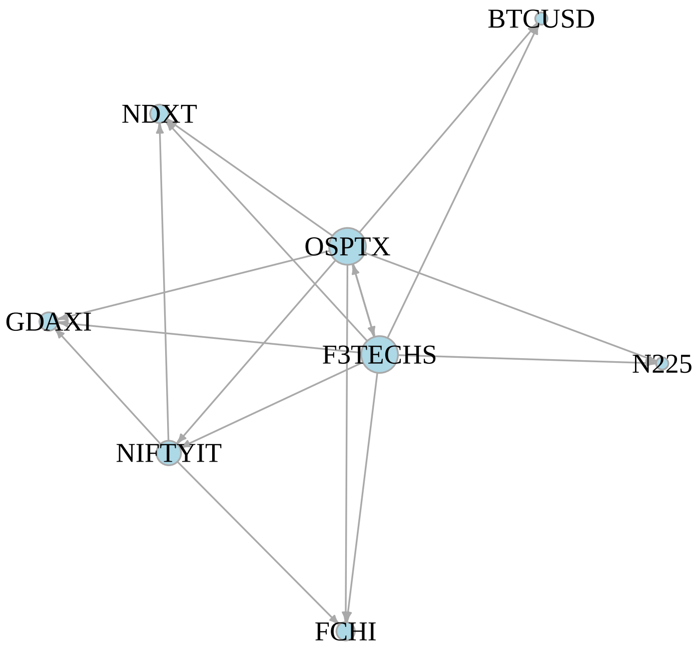

Figure 11. Granger Network estimation, p=10.

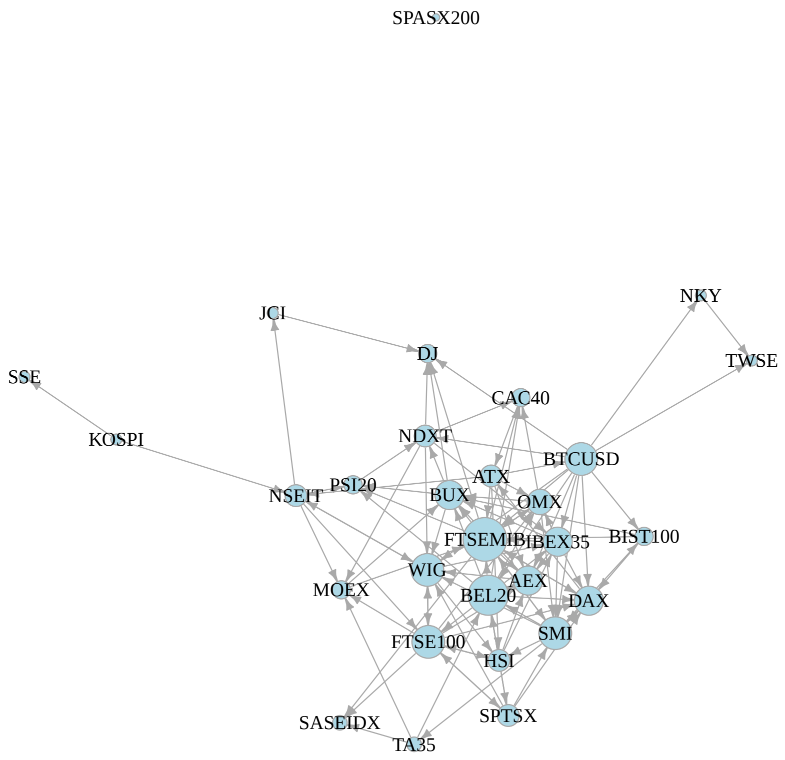

Figure 12. Granger Network estimation, p=1.

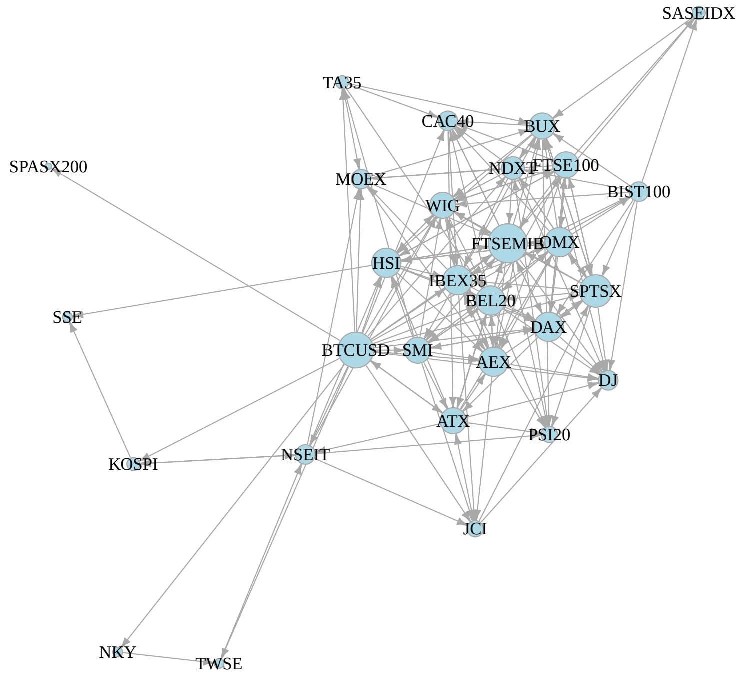

Figure 13. Granger Network estimation, p=2.

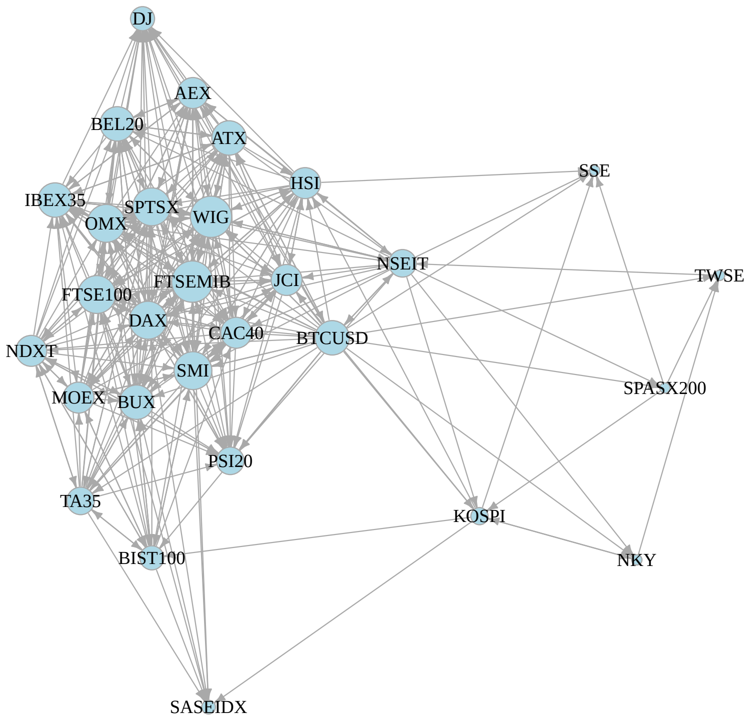

Figure 14. Granger Network estimation, p=6.

References

Allen, F., Gale, D. (2000): "Financial Contagion", *Journal of Political Economy, Volume 108, No. 1*

Amihud, Y. (2002): "Illiquidity and stock returns: cross-section and time-series effects", *Journal of Financial Markets, Volume 5, 31–56*

Barigozzi, M., Brownlees, C. (2018): "NETS: Network estimation for time Series", *Journal of Applied Econometrics, Volume 34, No. 3, 347-364*

Baur, D.G., Hong, K.J., Lee, A.D. (2015): "Bitcoin: Currency or Asset?", *Melbourne Business School, 2016 Financial Institutions, Regulation & Corporate Governance (FIRCG) Conference*

Borri, N. (2018): "Conditional Tail-Risk in Cryptocurrency Markets", *Journal of Empirical Finance, Forthcoming*

Dahlahlus, R. (2000): "Graphical interaction models for multivariate time series", *Metrika, 51, 157–172*

Diebold, F.X., Yilmaz, K. (2014): "On the network topology of variance decompositions: Measuring the connectedness of financial firms", *Journal of Econometrics, Volume 182, 119–134*

Eichler, M. (2007): "Granger causality and path diagrams for multivariate time series", *Journal of Econometrics, Volume 137, 334–353*

Fong, K.Y.L., Holden, C.W., Tobek, O. (2018): "Are Volatility Over Volume Liquidity Proxies Useful For Global Or US Research?", *Kelley School of Business Research Paper No. 17-49*

Forbes, K.J., Rigobon, R. (2002): "No Contagion, Only Interdependence: Measuring Stock Market Comovements", *The Journal of Finance, Volume LVII, No. 5*

Gómez-Puig, M., Sosvilla-Rivero, S. (2014): "Causality and contagion in EMU sovereign debt markets", *International Review of Economics and Finance, Volume 33, 12–27*

Guo, X., Lu, F., Wei, Y. (2021): "Capture the contagion network of bitcoin – Evidence from pre and mid COVID-19", *Research in International Business and Finance, Volume 58, 101484*

Katsiampa, P., Yarovaya, L., Zięba, D. (2022): "High-frequency connectedness between Bitcoin and other top-traded crypto assets during the COVID-19 crisis", *Journal of International Financial Markets, Institutions & Money, Volume 79, 101578*

Matkovskyy, R., Jalan, A. (2019): "From financial markets to Bitcoin markets: A fresh look at the contagion effect", *Finance Research Letters, Volume 31, 93–97*

Pericoli, M., Sbracia, M. (2003): "A primer on financial contagion", *Journal of Economic Surveys, Volume 17, No. 4*

Ranaldo, A., Santucci de Magistris, P. (2022): "Liquidity in the global currency market", *Journal of Financial Economics, Volume 146, 859–883*

Saiedi, E., Broström, A., Ruiz, F. (2021): "Global drivers of cryptocurrency infrastructure adoption", *Small Bus Econ 57, 353–406*

Wang, H., Wang, X., Yin, S., Ji, H. (2022): "The asymmetric contagion effect between stock market and cryptocurrency market", *Finance Research Letters, Volume 46, Part A, 102345*

Summary

This work's purpose is the investigation about the existence of contagion dynamics between cryptocurrency markets and other traditional financial markets, both in terms of asset liquidity and returns, with the support of a network analysis. Cryptocurrency markets are starting to be more common among investors, and to gain growing importance in the asset classes universe, attracting an increasing amount of money in the last years from both retail and institutional investors. It becomes therefore important to understand dynamics between such digital asset and other traditional assets. Today, we ask ourselves whether Bitcoin and other cryptocurrencies can be considered as alternatives to fiat money or they're some other sort of financial asset. Baur et al. (2015) investigated this question; Bitcoin was showing no correlation with traditional asset classes both in normal times and during market turmoils, the authors denote, and this implied that Bitcoin could have served as a portfolio diversification asset and that it would not present macroeconomic risk, containing possible speculative bubbles and crashes inside its environment without affecting other asset classes; but they also found that the cryptocurrency was mainly used in a speculative way, while only a minority of users were treating Bitcoin as a medium of exchange. But things have changed since that paper was published: Bitcoin's current market capitalization is 319B USD[42], more than 70 times bigger compared to the approximately 4.5B USD back in 2015. So, researchers started to ask themselves again whether the absence of correlation between Bitcoin (and other cryptocurrencies) and other financial assets still holds, and this work is intended to contribute to such research area.

Contagion Definition

In the current literature, the term 'contagion' appears to have multiple meanings. The first work ever about financial contagion is Allen, Gale (2000), focusing on contagion within the banking system, and the term was intended as the transmission of an "infectious illness" from an "infected individual" to an "healthy" but susceptible one, directly or indirectly; in short, "contagion" is a form of systemic risk modeling.

After this work, many more working papers were produced by the literature; the focus was also enlarged from the sole banking system to the broader financial one, including capital markets. But Pericoli, Sbracia (2003) denote that, at the time of writing, a lot of confusion

[42] as of 18^{th} Nov. 2022 (source: Coinmarketcap.com).

was present on the term “contagion”, as a theoretical or empirical identification of such phenomenon was missing. In particular, there was no distinction between shocks generated by normal markets interconnection and shocks that generate a discontinuity in stock prices. They then find that five different predominant definitions of contagion can be found in the literature, each of them focusing on a different measure for the contagion effect. Among these definitions, that can be found in the detailed thesis document, the most complete (and more widely used) one is the following:

Definition 5. *(Shift-)contagion occurs when the transmission channel intensifies or, more generally, changes after a shock in one market.*
In this definition the focus is on excess measures of some quantities - in this case, the Definition generally refers to the “transmission channel”, which experiences a change (not necessarily an intensification) during contagion times. The authors denote also that such channels may start to exist due to the start of a financial crisis and cease to exist when the crisis ends, thus being “crisis-specific” channels.

When coming to how to measure such events some caveats arise, as the authors report. First of all, for a statistical analysis, it is often needed to identify an initial shock in a market or a group of markets, but this may be not the case: a shock can belong to multiple markets or countries since its beginning, thus making its identification harder. Another important issue, closely related to co-movements in asset returns as discussed above, is data frequency selection. Pericoli, Sbracia (2003) correctly point out that there could be hidden contagion dynamics characterised by lags in the time series, making a correct data frequency selection hard: selecting intraday, daily, weekly data may not be appropriate due to said “infrequent but significant changes in asset prices that are correlated - with some lags - across markets”[43]. Also, these changes may lead to diverse effects - not necessarily higher correlations. Lastly, issues may be found in selecting the correct quantities to analyse contagion: there may be other quantities more than prices that reflect the reaction of investors to a financial crisis in a country - like withdrawals of money. As it will be pointed out later, I addressed this last issue by choosing to measure contagion with a (il)liquidity measure that can take into account for both price and volumes movements, to try to capture the effect of withdrawals during crisis periods.

[43] Pericoli, Sbracia (2003). The authors mention the example of a crisis in country i that leads investors to revise their expectations on productivity in country j, so that contagion in this scenario results in a one-time adjustment in the level of prices in country j.

The **Definition 5** presented above is the one I selected to be used in my analysis. My choice is consistent with Gómez-Puig, Sosvilla-Rivero (2014), where the authors define contagion as *"an abnormal increase in the number or in the intensity of causal relationships, compared with that of tranquil period, triggered after an endogenously detected shock"*. Also, in the same way the authors of this last mentioned work did, the procedure I use to measure contagion not only takes into account the analysis of market correlations, but also tries to find causality in these co-movements, using a Granger-causality approach. The usage of Granger causality tests to measure contagion is suggested by Forbes, Rigobon (2002): in particular, the authors state that such approach can be beneficial in all those cases "when the source of the crisis is less well identified and endogeneity may be more severe [...] to determine the extent of any feedback from each country in the sample to the initial crisis country."[44].

Network Analysis and Granger Causality

In general, a network applied to some sort of data is a graphical interaction model. Simply put, network analysis is a way to represent relations present in panel data through a mathematical graph, where vertices are the variables in panel data and edges between vertices are a given kind of interconnection between the variables. Many monographs and working papers have been published in the last century on the application of graphical interaction models to many kinds of data for various researches. As Eichler (2007) says, graphical approach presents a set of advantages; principally, graphs can be easily read as they are a concise representation of the phenomenon object of research, and allow for a simple representation even if the panel is composed of a large number of variables.
The author demonstrates that some *"path diagrams"* that represent the autoregressive structure of a multivariate time series can be associated to a Granger-causality structure. Considering a multivariate time series X with an autoregressive representation such as follows:

$$X(t) = \sum_{u=1}^{\infty} \Phi(u) \cdot X(t-u) + \varepsilon(t) \tag{34}$$

where $\varepsilon(t)$ is a white noise process, a path diagram associated with X is a graph $G = (V, E)$, being $V = \{1, ..., d\}$ the set of vertices and E the set of edges. This path diagram presents the following two characterizations:

1. $a \rightarrow b \notin E \iff \Phi_{ba}(u) = 0 \;\; \forall\, u \in \mathbb{N}$
2. $a — b \notin E \iff \Sigma_{ab} = 0$

[44] Forbes, Rigobon (2002).

with the first one characterizing *directed* edges[45], while the second one describes *undirected* edges.

Related to the autoregressive representation is the Granger-causality probabilistic concept. Formally, let us consider a subprocess of X, denoted X_A, corresponding to a certain $A \subseteq V$. It is possible to define a certain information set $I_{X_A} = \{I_{X_A}(t), t \in \mathbb{Z}\}$ as a sequence of closed subspaces generated by the past and present values of X_A at time t, denoted as $\overline{X}_A(t) = \{X_A(s),\ s \leq t\}$. From here, we can consider two subsets A and B of $S \subseteq V$. With a "negative" definition, the process X_A is Granger-noncausal for the process X_B considering the information set I_{X_S} if:

$$X_A \nrightarrow X_B\ [I_{X_S}] \coloneqq\ X_B(t+1) \perp \overline{X}_A(t) \mid \overline{X}_{S \setminus A}(t) \quad \forall t \in \mathbb{Z} \tag{35}$$

Similarly, the processes are contemporaneously uncorrelated if:

$$X_A \nsim X_B\ [I_{X_S}] \coloneqq\ X_A(t+1) \perp X_B(t+1) \mid \overline{X}_S(t) \quad \forall t \in \mathbb{Z} \tag{36}$$

These definitions can be linked to the path diagram defined above considering the following two statements:

1. $a \rightarrow b \notin E \iff X_A \nrightarrow X_B\ [I_{X_S}]$
2. $a — b \notin E \iff X_A \nsim X_B\ [I_{X_S}]$

In this way, it is possible to draw the Granger-causality system of the variables through a graph, generating a Granger-causality path diagram which, ultimately, can be used to understand causal dynamics among the variables of the system.

Contagion Metric

As stated in Ranaldo, Santucci de Magistris (2022), distressed markets are characterized by sudden drops in liquidity: in such context, using an illiquidity measure[46] can be useful to detect when trading volumes have an high price impact on a broad range of financial instruments. Ideally, this measure should be capable of detecting the so-called "panic selling" phenomenon.

The most famous illiquidity measure comes from Amihud (2002), defined as:

[45] note that the author specifies that directed edges are associated to a temporal ordering. This will be true also in my application.

[46] in the sense of Amihud (2002) illiquidity measure, capable of measuring the price impact of trading volumes.

$$ILLIQ = \frac{1}{D}\sum_{t=1}^{D}\frac{|R_t|}{VOLD_t} \tag{37}$$

where D is the number of days considered, $|R_t|$ is the daily absolute return of the security and $VOLD_t$ is the respective daily trading volume in dollars.

In this work, as in Ranaldo, Santucci de Magistris (2022), a realized version of the Amihud's measure is used. For a daily illiquidity estimator ($D = 1$), using the classical definition of Amihud's measure could lead to an underestimation of the illiquidity measure: this because the Amihud's measure numerator is represented by the absolute value of the daily return of the security, $|R_t|$, which is computed using the close prices in $t-1$ and t, thus not taking into account the daily volatility of the security. To avoid this possible underestimation, the original numerator has been replaced by a daily volatility measure.

The ratio of a volatility measure and trading volume represents a measure belonging to the wide family of "Volatility-over-Volume" (VoV) indicators, defined by Fong et al. (2018) as a class of liquidity proxies that present the following form:

$$VoV = \frac{a\sigma^b}{V^c} \tag{38}$$

where *a, b, c* are strictly positive. Here, the volatility statistic (σ) is expected to capture the daily volatility of the security. For this reason, the choice fell on Parkinson's high-low range (Parkinson, 1980)[47]. For a time range of $T = 1$, thus indicating daily volatility, Parkinson's measure is defined as follows:

$$\tilde{\sigma_{it}}^2 = \frac{1}{4\log 2}(p_{it}^{high} - p_{it}^{low})^2 \tag{39}$$

where p_{it}^{high} and p_{it}^{low} denote, respectively, the maximum and minimum log price of security i on the day t.

Given this definition of volatility, the VoV measure used for the empirical analysis is defined as:

$$VoV_{it} = \frac{\tilde{\sigma_{it}}^2}{V_{it}} \tag{40}$$

where $\tilde{\sigma_{it}}^2$ is the realized Parkinson Volatility for security i at time t, and V_{it} is the respective daily trading volume. Note that, with respect to Fong et al. (2018) definition, for this measure results that $a = 1, b = 2$ and $c = 1$[48]. This metric is the one used for the first

[47] this choice is consistent with Barigozzi, Brownlees (2018).

[48] a value of $a = 1$ is consistent with the findings of Fong et al. (2018) for their $VoV(\lambda)$. Note as well that the specification used here is just the square of this latter measure. In fact, Fong et al. $VoV(\lambda)$ is

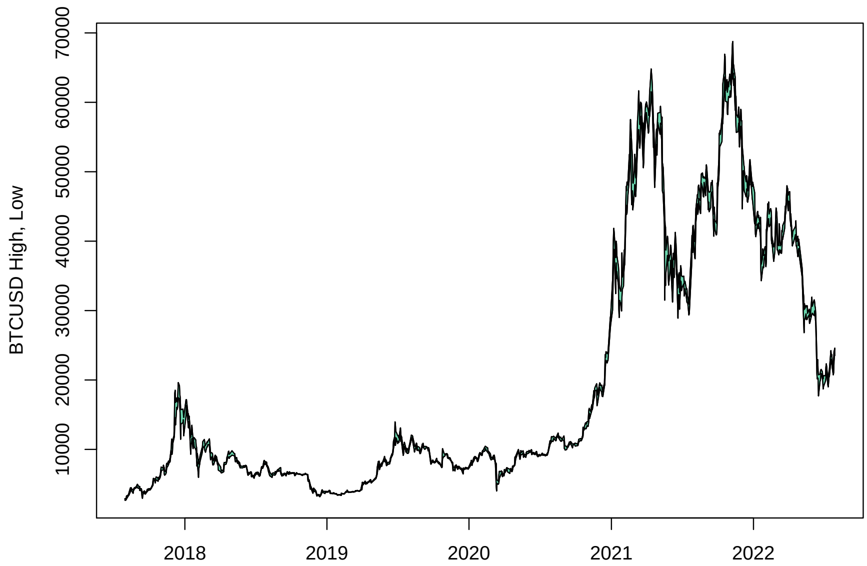

Figure 15. an example of the effect captured by Parkinson's Volatility: the high-low daily volatility is represented by the shaded area.

contagion analysis presented later.

In order to have a more precise result, a similar analysis with intraday data would have been carried out. Unfortunately, due to unavailability of intraday volumes data, it was not possible to compute the same metric above described. Still, since an intraday analysis can give better insights about how the two markets are causally interrelated, a single exercise is then performed using log-returns as the contagion metric, where log-returns are defined as:

$$r_{i,t} = log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \tag{41}$$

P_i being the prices of security i at times $t, t-1$.3

Data Selection

Since two different analyses are performed, two different datasets are considered: one composed of daily data, whose time range is 5 years, from 31/07/2017 to 31/07/2022, thus leading

defined as:

$$VoV(\lambda) = \frac{a\sigma}{V}$$

where $a = 1$.

to a panel of 1080 observations; and a second one consisting of intraday 5-minutes data, whose time range is 10 days, from 15/03/2020 to 25/03/2020; the selected time window is around the 16/03/2020, day referred to as "Black Monday II"[49], thus leading to a panel of 2710 observations, not taking into account market closures among different markets. This data has been synchronized using GMT as the common timestamp.

Data for stock markets is represented by a selection of global stock indices. A specific panel is therefore composed by high and low prices and trading volumes for the daily resolution dataset; by last prices for the intraday dataset. This choice was made in order to have the widest possible indicators for the broader stock market[50]. Also, where possible, Tech-specific sector indices have been used. This comes from the fact that, according to various sources[51], it seems that Bitcoin's price moves in the same way tech stocks move.
For the two datasets, a different series of stock markets data has been considered. For the 5-year dataset, due to sparsity and multicollinearity problems, only a selection of 7 stock indices was made, representing different geographic areas. Instead, for the high-frequenxy dataset, a collection of 29 stock indices was selected. Stock indices data comes from Bloomberg for the 5-year dataset, and Bloomberg and Barchart Premier for the high-frequency one. The selected indices are summarized in Table VII.

Regarding cryptocurrency data, a wide variety of cryptocurrencies exist nowadays, but a great number of these are projects born due to the recent years' euphoria around crypto-assets in general, and some of these are traceable to scams or fraud attempts; thus choosing the correct crypto-asset is crucial. Bloomberg has defined, in May 2018, the "Bloomberg Galaxy Crypto Index" (BGCI) which serves as a capped market capitalization-weighted performance indicator for the largest cryptocurrencies traded in US Dollars[52]. But, for the purpose of this work, this said indicator presents two issues: first, no data is available before its launch date (3^{rd} May, 2018); second, to date Bitcoin's market capitalization is more than double of Ethereum's one[53], while in the BCG Index they are equally weighted[54]. This difference in weighting could have led to a distorsion in results. Moreover, as denoted by many (for example, Katsiampa et al. (2022)), interconnectedness among different crypto-assets is very

49 from *"2020 stock market crash"*, Wikipedia, last edited on 3^{rd} Sep. 2022.
50 alternatively, ETFs replicating the stock indices could have been used. However, this choice would have led data to be influenced by passive investors only.
51 for example, the arcticle *"Bitcoin Is Increasingly Acting Like Just Another Tech Stock"*, The New York Times, 11^{th} May 2022.
52 source: Bloomberg LP, Bloomberg Digital Asset Indices.
53 according to Coinmarketcap.com, BTC's market capitalization stands around 319B USD, while ETH's one waves around 148B USD, as of 18^{th} Nov. 2022.
54 each of them accounts for 30% of the index total weight (source: BCGI Factsheet, Feb. 2020)

Daily dataset		
Variable	Description	Source
F3TECHS	UK 350 Tech. Index	Bloomberg
FCHI	France 40	Bloomberg
GDAXI	Germany 30	Bloomberg
N225	Nikkei 225	Bloomberg
NDXT	Nasdaq-100 Tech. Index	Bloomberg
NIFTYIT	Nifty IT Index	Bloomberg
OSPTX	S&P/TSX Index	Bloomberg

5 min dataset		
Variable	Description	Source
AEX	Amsterdam Exchange	Bloomberg
ATX	Austrian Traded Index	Bloomberg
BEL20	Belgium 20	Bloomberg
BIST100	Istanbul 100	Bloomberg
BUX	Budapest Stock Exchange	Bloomberg
CAC40	France 40	Bloomberg
DAX	Germany 30	Bloomberg
FTSE100	UK 100	Bloomberg
FTSEMIB	Italy 40	Bloomberg
HSI	Hang Seng Index	Bloomberg
IBEX35	Spain 35	Bloomberg
JCI	Jakarta Composite Index	Bloomberg
KOSPI	Korea Composite Index	Bloomberg
MOEX	Russia Index	Bloomberg
NKY	Nikkei 225	Bloomberg
NSEIT	Nifty IT Index	Bloomberg
OMX	Stockholm 30	Bloomberg
SASEIDX	Tadawul All Share	Bloomberg
SMI	Swiss Market Index	Bloomberg
SPASX200	S&P/ASX 200	Bloomberg
SSE	Shanghai Composite Index	Bloomberg
TA35	Tel-Aviv 35	Bloomberg
TWSE	Taiwan Stock Index	Bloomberg
WIG	Warsaw Stock Exchange	Bloomberg
DJ	Dow Jones	Barchart Prem.
NDXT	Nasdaq-100 Tech. Index	Barchart Prem.
PSI20	Lisbon 20	Barchart Prem.
SPTSX	S&P/TSX Index	Barchart Prem.
SPX	S&P500 Index	Barchart Prem.

TABLE VII. Presentation of the two Stock Indices datasets.

high. For this reason, paired to the fact that Bitcoin is the first born, most traded and with highest market capitalization cryptocurrency, it was decided to summarize the cryptocurrencies market in the Bitcoin itself. Thus, data for the BTCUSD rate was obtained from CoinCodex (for the 5-year daily data) and Barchart Premier (for the 5-minutes data).

NETS Algorithm

To estimate causal interconnectedness a network-based approach is used, as previously mentioned. In particular, the Network Estimation for Time Series (NETS) algorithm from Barigozzi, Brownlees (2018) was selected. In the following sections, a summarized description of the model and the estimation algorithm are presented[55].

The objective is to model a panel of time series as a vector autoregression, with the assumption that the autoregressive matrices and the inverse covariance matrix of the system innovations are assumed to be sparse, i.e. with most of their elements equal to zero. The resulting system can be represented by two different kind of graphs: a directed graph, representing predictive Granger relations smong the time series; and an undirected graph, representing their contemporaneous partial correlations. The authors' innovative NETS algo-

[55] The following dissertation takes directly from Barigozzi, Brownlees (2018).

rithm for the estimation consists of a LASSO regressions-based procedure: its main feature is the contemporaneous estimation of the autoregressive matrices and the concentration matrix of the system, instead of splitting the estimation in two different steps.
In this model, panel data is modeled as a VAR:

$$y_t = \sum_{k=1}^{p} A_k y_{t-k} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d(0, C^{-1}) \tag{42}$$

where the autoregressive matrices A_k and the concentration matrix C^{-1} are $n \times n$ matrices and assumed to be sparse.
To assess dynamic interdependance[56] among time series a multivariate version of the Granger causality notion is used. Formally:

$$E[(y_{it+k} - E(y_{it+k}|\{y_{1t}, ..., ynt\}))^2] = E[(y_{it+k} - E(y_{it+k}|\{y_{1t}, ..., y_{nt}\} \backslash y_{jt}))^2] \tag{43}$$

meaning that y_{jt} does not Granger cause y_{it} if adding the former as a predictor does not improve the mean squared forecast error of y_{it+k} for any $k > 0$. This implies that, if $a_{k(i,j)} = 0 \ \forall k, \ a_k \in A_k$, then y_{jt} does not Granger cause y_{it}. Note that this approach allows to spot contagion dynamics even in case these are characterized by lags, as pointed out before.

Based on this relations, a network is then produced. The Granger causality network is defined as a graph $N_G = (V, E_G)$, V being the set of vertices $\{1, ..., n\}$ and E_G the set of edges, subset of $N \times N$ such that the pair $(i, j) \in E_G$ if and only if i and j are linked by an edge, i.e. are Granger causality-related. Since the Granger network is a directed network, the presence of an edge from i to j means that i Granger causes j in the sense of Equation 43. In formal notation:

$$E_G = \{(i, j) \in V \times V \iff a_{k(i,j)} \neq 0, for\ at\ least\ one\ k \in \{1, .., p\}\} \tag{44}$$

Goal of the estimation is to find all the non-zero entries of each autoregression matrix, A_k, and of the concentration matrix C. Since the direct estimation on the latter from the residuals of the former has some issues[57], authors created a procedure to estimate both sets of parameters jointly. For this purpose, the VAR model is re-parametrized as follows:

- a n^2p-dimensional vector α containing a set of coefficients α_{ijk}, corresponding to autoregressive coefficients a_{kij};

[56] For the purpose of this work, we are only going to consider dynamic interdependance. For details on contemporaneous interdependance, here not specified for brevity, see Barigozzi, Brownlees (2018)
[57] see Barigozzi, Brownlees (2018) for further details.

- a $n(n-1)/2$-dimensional vector ρ containing a set of partial correlations[58] ρ^{ij};
- a n-dimensional c vector containing a set of c_{ij} coefficients, corresponding to the diagonal of the concentration matrix C.

The resulting VAR and contemporaneous equations are the following:

$$y_{it} = \sum_{k=1}^{p} \sum_{j=1}^{n} \alpha_{ijk} y_{jt-k} + \epsilon_{it}, \quad i = 1, .., n \tag{45}$$

$$\epsilon_{it} = \sum_{h=1}^{n} \rho^{ih} \sqrt{\frac{c_{hh}}{c_{ii}}} \epsilon_{ht} + u_{it}, \quad i = 1, .., n \quad h \neq i \tag{46}$$

From these two equations it is possible to derive a LASSO-type estimator for the parameters of the model. I'm not going to report all the calculations here, that can be found in the detailed thesis document as well as in Barigozzi, Brownlees (2018). It is here sufficient to report that, for the estimation, the authors propose an iterative coordinate descent algorithm to minimize the objective function. Intuitively speaking, each iteration updates one component of the parameters vector $\theta = (\alpha', \rho')'$, and a residual estimate is computed. In each iteration, the parameter is updated using a combination of two different set of variables: a first set of regressors corresponding to the coefficient being updated; and a second vector containing the partial residuals of the model with respect to all the parameters besides the coefficient being currently updated. The algorithm then proceeds with the updates until convergence, which is checked at the end of each full cycle of updates of θ. The first full cycle is based on a pre-estimation of c; then, after an estimation of θ is available, the estimation of c is updated as well. This two last steps are then iterated until convergence, and the result is the full set of the estimated parameters $\hat{\theta}_T$.

Daily Data Analysis

Data has been split in two different categories: "analysis" data, comprehending daily high prices, daily low prices and daily volumes for BTCUSD, NDXT, FCHI, GDAXI, N225, F3TECHS, OSPTX, NIFTYIT; and "market" data, comprehending daily high prices, daily low prices and volumes for XLK. For each subset of data, Parkinson Volatility has been computed according to the equation presented above. Then, the liquidity VoV measure has been calculated, as in Equation 40. Descriptive statistics for the VoV measure[59] as well as a simple correlation analysis can be found in the detailed document.

[58] defined as $\rho^{ij} = corr(\epsilon_{it}, \epsilon_{ij} | \{\epsilon_{kt} : k \neq i, j\})$
[59] for descriptive statistics computation, the VoV measure has been multiplied for a factor of $10^1 2$.

A medium-high degree of positive correlation is shown among different stock indices series, reaching the value of 0.90 between FCHI and GDAXI. For what concerns BTCUSD instead, it presents small negative correlation with almost all the stock indices.
For the NETS algorithm sparsity assumption, common factors must be removed. For this goal, a preliminary regression is specified as follows:

$$VoV_{i,t} = \alpha_i + \beta_i VoV_{M,t} + \varepsilon_{i,t} \tag{47}$$

where VoV_i denotes a list of VoV measures for the time series i in the "analysis" dataset and VoV_M denotes the same measure for a list of market-wide factors. The reasons for selecting the SPDR Technology sectorial index (XLK) as the market factor are presented in the detailed document.

Of this series of regressions, estimated with OLS method, residuals have been stored[60]. After the pre-processing, some kind of variability persists in residuals. Moreover, from the BTCUSD residuals chart can be denoted that the series start to be much lower in absolute value[61] becoming comparable to residuals of the two Technology indices F3TECHS and NIFTYIT, i.e. the BTCUSD series, after a certain date, can more accurately be explained by the XLK factor. This result is consistent with the findings of Matkovskyy, Jalan (2019). Their explanation, supported by some research articles[62], was the introduction of Bitcoin futures on 18^{th} December 2018 which, according to the literature, improved efficiency in the Bitcoin market.

This set of data is then fed as input to the NETS algorithm. The resulting Granger Network graphs are presented in Figures 9, 10, 11 in the Appendix. It should be no surprise that not all the stock markets are related: the analysis is focused on causal relationships, but this does not mean that there may have been co-movements between non causally-related time series. It is important to remember that we took into account a market common factor as well, thus removing a great part of co-movements. Below, the matrices associated with the graphs are presented, where a value of 1 determines the presence of and edge **from** x_i (row i) **to** x_j (column j).

Since the contagion metric is a liquidity measure, each edge can be interpreted as a price impact movement in a market that causes other price impact movements in other markets

[60] Residuals' statistics and charts can be found in the detailed document of the thesis.
[61] note that the VoV measure, being a ratio of two positive numbers, is always a positive number.
[62] see Matkovskyy, Jalan (2019) for further details on according and contrasting literature.

$$\begin{array}{r} BTCUSD \\ NDXT \\ FCHI \\ GDAXI \\ N225 \\ F3TECHS \\ OSPTX \\ NIFTYIT \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Granger Network matrix for $p = 1$

$$\begin{array}{r} BTCUSD \\ NDXT \\ FCHI \\ GDAXI \\ N225 \\ F3TECHS \\ OSPTX \\ NIFTYIT \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Granger Network matrix for $p = 5$

$$\begin{array}{r} BTCUSD \\ NDXT \\ FCHI \\ GDAXI \\ N225 \\ F3TECHS \\ OSPTX \\ NIFTYIT \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Granger Network matrix for $p = 10$

Figure 16. Granger Network adjacency matrices for different specifications of the model. Variables in matrices columns are in the same order as in matrices rows, from left to right.

after 1, 5, or 10 trading days, as drawn by the three graphs. Looking at both graphs and matrices, it is possible to understand the evolution of such relationships with the passage of time. In particular, there are three stock market indices that causes liquidity movements in all the other vertices: the FTSE 350 SuperSector Technology Index (F3TECHS), the S&P/TSX Composite Index (OSPTX), and the Nifty IT Index (NIFTYIT). Overall, these three markets are interrelated in the Granger Causality sense, and they generate liquidity spillovers on the other markets, including Bitcoin (BTCUSD). Note that, as testified by the graph matrices, with the passage of time the overall system tends to be more interrelated. Also, imposing $p > 10$ did not produce different results compared to Figure 11. Thus, this last plot can be considered the final liquidity network for the considered variables.

A preliminary explanation of my results, even if it's out of the scope of my work, can be found in the graphs from Saiedi et al. (2021)[63]: both F3TECHS and OSPTX have high scores concerning *bitnode intensity*, *unique bitnodes* and *Bitcoin merchants*.

[63] see Figures 6 and 7 in the detailed document.

Intraday Data Analysis

For this second analysis, the intraday (5 minutes) dataset is used. As before, an "analysis" dataset and a "market" dataset are created, the latter containing data from the S&P500 Index (SPX) while the former containing all the remaining data. Here, data to compute Parkinson Volatility was not available, since 5-minutes data only comes with a Last Price value. For this reason, 5-minutes log-returns have been computed for each time series, both in the "analysis" and in the "market" datasets. These returns will be the starting data for the analysis. A preliminary correlation analysis on returns can be found in the detailed document. Overall, there seems to be some kind of correlation in log-returns for indices belonging to the same geographic area.
Next, to ensure sparsity of data for the NETS algorithm assumptions, also in this analysis a preliminary set of regressions is performed on data. The general preliminary regression model is specified as follows:

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t} \tag{48}$$

The SPX factor has been selected to be the market factor[64]. Of this set of regressions, estimated with OLS, residuals have been stored: these are the data fed to the NETS algorithm. Reporting here complete descriptive statistics for residuals is superfluous. It has to be noted that, as in the previous analysis, variability persists among data even after correcting for common factors.

This dataset is then fed to the NETS algorithm. The resulting Granger Network graphs are presented in Figures 12, 13, 14 in the Appendix[65]. Here, reporting the complete network matrices would be superfluous, as there are many relations between all the variables of the system. For this work purpose it is useful to report BTCUSD interactions with other indices. A summary table with selected sub-matrices is presented in Table V in the detailed document.

Here, the contagion metric used is securities log-returns. In this case, edges can be interpreted as a causal relation in return movements from a variable to another. As the graphs show, during the so-called "Black Monday II" the overall interconnectedness between markets has been high in the number of causal relations. The only market that seems non-responsive within a 5-minute lag period is the S&P/ASX 200 (SPASX200), being the only vertex not connected to the system in Figure 12. With the passage of time, as in the previous analysis,

[64] details on this factor's selection can be found in the detailed document.
[65] note that, in this case, each lag corresponds to a 5-minute lag in the time series. Thus, $p = 1$ is a 5 minutes lag, $p = 2$ is a 10 minutes lag and $p = 6$ is a 30 minutes lag.

system's interconnectedness increases, w.r.t. Figures 13 and 14. Looking at both graphs and Table V, it is clear that, in this particular event, BTCUSD had a significant impact on the overall system: 27 out of 28 markets suffered from BTCUSD spillover effects, with the AEX index (AEX) being the only market not influenced by the cryptocurrency return movements in a 30-minute range. It is possible to note that the most influencial market index in the system has been the FTSE/MIB index (FTSEMIB).

For what concerns contagion from stock markets to BTCUSD, the only two stock indices directly affecting BTCUSD performance have been the Austrian Traded Index (ATX) and the Nifty IT index (NSEIT), the latter showing contagion dynamics only with a 30 minutes lag while the former affecting BTCUSD in each moment of the system. No influence seems to have either the London Stock Exchange or the Canadian Stock Exchange as found in the previous analysis, which maybe need more time to propagate their effects on the cryptocurrency market[66].

Looking at the graphs from Saiedi et al. (2021)[67], the metrics proposed by the authors seem to give a good explanation for the ATX influence on BTCUSD, while regarding NSEIT they seem to give little explanation of the phenomenon: in fact, India is in the lower part of the classifications for *bitnode intensity*, *unique bitnodes* and *Bitcoin merchants* all. It can be denoted, however, that the fact that a stock market of a certain country is the cause of spillovers in the global stock market can be unrelated with internal investors behavior: capital outflows may be caused by foreign investors, leading to a significant effect when the total stock detained by foreigners is a large percentage of the total equity listed[68].

Trying to explain this wide set of interrelations can be challenging, and it is out of the scope of this work. Here, it can be concluded that BTCUSD has had causal relationships, both inwards and outwards, with a large number stock markets during "Black Monday II". This outcome confirms the research hypothesis that Bitcoin and stock markets show contagion dynamics between each other.

Conclusions

My research goal was to test for contagion dynamics between a set of stock markets and Bitcoin, with the definition of contagion being above specified. To accomplish this, I used a Network Estimation tool from Barigozzi, Brownlees (2018), called NETS, that builds a

[66] note that for $p > 6$ the computation time for estimation starts to be material, due to the high number of parameters of the model.

[67] see Figures 6 and 7 in the detailed document.

[68] according to the article *Who owns the Indian stock exchanges?* from MarketToday, $13^t h$ Oct. 2022, "Foreign investors have an overall stake of 44.6% in NSE".

Granger-causality Network starting from panel data, where variables linked by an arrow indicate that x_i directly Granger-causes x_j if and only if there exists a directed edge originated by x_i and pointing towards x_j; while x_i indirectly Granger-causes x_j if and only if there is a set of directed edges that, combining their starting and ending point, have their origin in x_i and their end in x_j. The analysis has been conducted both in terms of liquidity, with a liquidity measure build upon daily data, and in terms of log-returns, calculated with intraday 5-minutes resolution data. Detailed graphs of the generated Granger Networks can be found in the Appendix.

With the evidence from the Granger network graphs, I can conclude that Bitcoin shows direct contagion effects, in a Granger-causal sense, both inwards and outwards; the research hypothesis is then confirmed with a positive outcome. In general, my conclusions are in line with all the literature that worked on my same research topic and that was produced from the work from Matkovskyy, Jalan (2019) on. In fact, this mentioned work seems to be a pivot point in the literature: before such document, the prevailing results showed no or meaningless existing relations between cryptocurrency markets and stock markets; while research papers produced after this mentioned work showed intensified relations between these two asset classes. The main reason for this behavior switch seems to be the introduction of Bitcoin Futures on 18^{th} December 2018.

More specifically, my conclusions are in line with that from Matkovskyy, Jalan (2019). The authors in their work showed that investors, during risk-off periods, moved from two famous Bitcoin markets (Bitmap USD and GDAX GBP) towards assets listed on two stock markets, i.e. the NASDAQ Index (which is also a technology index) and the NIKKEI225 Index; and also, investors moved from Euro GDP/Bitcoin markets towards European assets. These movements can be interpreted as a “fly to stability” move coming from risk-averse investors: selling Bitcoin in favor of other financial assets would lead to a liquidity drainage in Bitcoin markets, caused by a risk-off period. In this sense, my results from the daily analysis of liquidity drainage in Bitcoin following liquidity drainage in stock markets show the same dynamic highlighted by the two authors. The logical connection between the two results may be the following: the start of a financial crisis causes liquidity drainage from stock markets inducted by the behavior of more risk-averse investors (who would not have any holding in Bitcoin); the same crisis would lead less risk-averse investors (who have holdings in Bitcoin) to move their liquidity from Bitcoin, a more volatile asset, to less volatile assets as the stock markets ones.

My conclusions, with respect to my intraday analysis, are also in line with X. Guo et al. (2021): using a methodology highly comparable to the one I used, the authors find that, from the COVID-19 pandemic outbreak on, Bitcoin showed a rising contagion effect with gold and stock markets. In particular, the authors find that "[...] European market has a dominant role in transmitting risks and information, that it is the net contagion source to Bitcoin and contagion intermediary in the infection chain of US market.". This is comparable to my results, where the predominant stock index in the Granger network graph for the intraday contagion analysis was the FTSE-MIB Index (FTSEMIB), with the difference being that, in my case study, Bitcoin was the origin of contagion towards this market. A possible explanation for this difference may be the data length and resolution characteristics, mine being much shorter in length, including just a few days around a specific event, i.e. the "Black Monday II" market crash.

The said switch in Bitcoin's behavior (and, consequently, other cryptocurrencies' one) modifies the possible real-world applications of literature's findings: as an example, before 2019, Bitcoin could have been used for portfolio diversification and optimization purposes in a completely different way it would have to be done today, knowing these new findings. For this reason, it is important to expand the existing literature in this research area, the current one being of limited depth.

Tips for Further Research

This work provides meaningful insights on contagion dynamics of Bitcoin, and it can be a starting point for future research. There are, by the way, some caveats that need to be addressed in future works.

First of all, the liquidity analysis would be more meaningful with more appropriate data. This would ideally mean building a price impact indicator, similar to the one I used, that possibly captures price impact effects for each trade during the trading day, and also takes into account for after-market trades to avoid different behaviors based on trading hours. In order for this to be possible, high-resolution data with execution price and trade volumes information would have to be available for each trade, for each variable in the system, thus the researcher being able to build a high-frequency price impact measure.

Second, to use the same procedure I used a solid data cleansing process is needed. In this particular case, it would mean to being able to take into account for high interrelation among variables in the system, as they are all market indices that share a great amount of variance. In this sense, a researcher would have to find a series of factors that ensure to capture the most common variance possible among the selected indices, and this task results harder than it looks like: it is crucial, in fact, that the variables' idiosyncratic variance stays untouched, as it is the fundamental information on which the whole procedure is based.

Third, given the relations described through the Granger-causal network, the most natural follow-up would be to find the economic and logic meanings of these relations. This is most likely the hardest task, also due to the fact that cryptocurrencies' environment is continuously developing, then constantly requiring up-to-date data. I tried to give a first interpretation of my results, but many more variables should be taken into account.

A non-comprehensive list of these possible subjects of study would include:

- overall holdings of Bitcoin holders divided by asset class, and their behavior with respect to all their holdings during market turmoil;
- overall status of the Bitcoin blockchain, i.e. miners profit, hash rate, etc.;
- more precise data on Bitcoin's presence in the selected world countries.

Note that some of this data is either not available yet, or it is difficult to process in a way that is useful for a thorough analysis; so a pivotal topic for future research would be how to collect and review this kind of data.

Everything considered, I am firmly convinced of the power of network analysis in econometrics and investment problems, and that it should be used more consistently for various topics. The construction of a relational graph from the variables of a system is a perfect way to highlight relations between those variables, and the contemporaneous use of different graphs based on different relations can drive to deep conclusions on the entire system's behavior; also, a network can more precisely indicate how to build an impulse analysis to conduct stress tests over the selected environment and simulate different scenarios to derive possible backup strategies.