University of Pretoria
Department of Economics Working Paper Series

Can Volume Predict Bitcoin Returns and Volatility? A Nonparametric Causality-in-Quantiles Approach

Mehmet Balcilar
Eastern Mediterranean University and University of Pretoria
Elie Bouri
Holy Spirit University of Kaslik
Rangan Gupta
University of Pretoria
David Roubaud
Montpellier Business School
Working Paper: 2016-62
August 2016

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

Can Volume Predict Bitcoin Returns and Volatility? A Nonparametric Causality-in-Quantiles Approach

Mehmet Balcilar[♣], Elie Bouri[♦], Rangan Gupta[•] and David Roubaud[♠]

Abstract

The objective of this paper is to employ the recently proposed nonparametric causality-in-quantiles test to analyse the predictability of returns and volatility of Bitcoin over the daily period of 19th December, 2011 to 25th April, 2016, based on information provided by traded volume. The causality-in-quantile approach allows us to test for not only causality-in-mean, but also causality that may exist in the tails of the joint distribution of the variables. In addition, we are also able to investigate causality-in-variance (volatility spillovers) when causality in the conditional-mean may not exist, yet higher order interdependencies might emerge. We motivate our analysis by employing tests for nonlinearity. These tests detect nonlinearity, as well as the existence of structural breaks in the Bitcoin returns, and in its relationship with volume, implying that the Granger causality tests based on a linear framework is likely to suffer from misspecification. Unlike the result of no predictability obtained under the misspecified linear set-up, our nonparametric causality-in-quantiles test indicated that volume predicts returns over the quantile range of 0.25 to 0.75, i.e., barring in the bear and bull regimes of the Bitcoin market. However, we could not detect any evidence of predictability emanating from volume for the volatility of Bitcoin returns at any point of the conditional distribution. Our results highlight the importance of our detecting and modeling nonlinearity when analyzing causal relationships between volume and return in the Bitcoin market.

Keywords: Bitcoin; Volume; Returns; Volatility; Quantile Causality.

JEL Codes: C22, G15.

[♣] Department of Economics, Eastern Mediterranean University, Famagusta, via Mersin 10, Northern Cyprus, Turkey and Department of Economics, University of Pretoria, Pretoria, 0002, South Africa; IPAG Business School, Paris, France. Email: mehmet@mbalcilar.net.

[♦] Corresponding author. USEK Business School, Holy Spirit University of Kaslik, PO BOX 446, Jounieh, Lebanon. Email: eliebouri@usek.edu.lb.

[•] Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. Email: rangan.gupta@up.ac.za.

[♠] Montpellier Business School, Montpellier, France. Email: d.roubaud@Montpellier-BS.com.

1. Introduction

Studying the volume–return relationship is important to an understanding how information is transmitted to the market and embedded in asset prices. It also helps in increasing the power of forecasting asset return and volatility. In times of stress, in particular, it is central to examine the return-volume relationship to better understand market booms and crashes (Marsh and Wagner, 2000).

While the volume–return relationship has been extensively uncovered in equities (Karpoff, 1987), bonds (Balduzzi et al., 2001), commodities (Chiarella et al., 2016), and interest rate and currency future (Puri and Philippatos, 2008), it remains unexplored in the Bitcoin market. The latter has recently attracted the attention of the media and scholars given the rising importance of Bitcoin not only as an electronic payment system but also as a financial and speculative asset (Kristoufek, 2014).

In a speculative market such as that of Bitcoin, understanding the volume–return is essential to shed lights on market efficiency and its potential implications on trading strategies. Practically, if the transaction volume in the Bitcoin market has predictive power on return, this provides evidence of weak-form inefficiency and thereby practitioners will be able to construct volume-based strategies to increase their profits (Chen et al., 2001). This is particularly important given that several traders and practitioners have been relied on technical analysis as an alternative tool to study Bitcoin prices because no reliable fundamental valuation techniques are available for quantifying the intrinsic value of Bitcoin. The fact that market technicians employ models and trading rules based the relation between return and volume further motivates a better understanding of the Bitcoin volume–return relationship.

The short history of Bitcoin has been characterized by sharp upward and downward price movements associated with high transaction volumes. On November 19, 2013, the price of

Bitcoin on Bitstamp, the largest European Bitcoin exchange, plunged 19.88% on the highest volume ever recorded (71,560 Bitcoins). Furthermore, on December 7, 2013, the Bitcoin price plunged 14.92% and recorded a new all-time high volume of 79,852 Bitcoins. Again, on December 18, 2013, Bitcoin price plunged 22.80% and hit new daily volume record high of 137,070 Bitcoins[1]. These features suggest a strong relationship between the magnitude of price movements and transaction volumes. However, no insightful work has been done so far to uncover this relationship in Bitcoin. To address this literature gap, we use a novel nonparametric causality-in-quantiles test of Balcilar et al. (2016a, b) to examine the predictability of Bitcoin returns and volatility based on trading volume. For our purpose, we use daily data covering the period of 19th December, 2011 to 25th April, 2016. The nonparametric causality-in-quantiles test combines elements of the test for nonlinear causality of k-th order developed by Nishiyama et al. (2011) with the causality-in-quantiles test developed by Jeong et al. (2012) and, hence, can be considered to be a generalization of the former. The causality-in-quantile approach has the following three novelties: Firstly, it is robust to misspecification errors as it detects the underlying dependence structure between the examined time series, which could prove to be particularly important as well known that stock returns display nonlinear dynamics (see Bekiros et al., forthcoming, for a detailed discussion in this regard); a fact we show below as well. Secondly, via this methodology, we are able to test not only for causality-in-mean (1st moment), but also for causality that may exist in the tails of the joint distribution of the variables, which in turn, is important if the dependent variable has fat-tails – something we show below to hold for Bitcoin returns and volume. Finally, we are also able to investigate causality-in-variance and, thus, study higher-order dependency. Such an investigation is important because, during some periods, causality in the conditional-mean may not exist while, at the same time, higher-order interdependencies may turn out to be significant.

[1] For a detailed explanation on the negative and positive bubbles in the Bitcoin market, the reader can refer to Fry and Cheah (2016).

Indeed, there are studies like Chuang et al., (2009), Chiang and Li (2012), Gebka and Wohar (2013), Lin (2013) and Chen et al., (2016) that have used quantile based methodologies to study the relationship between returns and volatility with volume of traditional stock indices of Pacific Basin and Asian countries. However, to the best of our knowledge, this is the first paper that analyzes the predictability of returns and volatility of Bitcoin returns due to its traded volume.
The overall results indicate that the Bitcoin volume–return relation under normal market condition differs from that in bull and bear periods, with volume predicting returns during the normal phase of the market. In addition, trading volume provides no valuable information that helps to predict price volatility.

The rest of the paper is organized as follows: Section 2 reviews the related literature on the finance and economics of Bitcoin. Section 3 presents the methodology, while Section 4 discusses the data and the results. Finally, Section 5 concludes.

2. Literature Review

Bitcoin is the first cryptocurrency to come to existence. In 2009, it was introduced by Nakamoto (2009) as an open source software based online payment system. Since then, its popularity among practitioners and economic actors has soared in response to the perceived failures of governments and central banks during the global financial crisis (GFC) of 2008 and the European sovereign debt crisis (ESDC) of 2010–2013. Unlike conventional currencies where central authorities and central banks guarantee them or have control over currencies, Bitcoin is fully decentralized and depends on a sophisticated protocol that uses only cryptography to control transactions, manage the supply, and prevent bad acting that may endanger the system. All transactions are stored digitally and recorded in a shared ledger data technology known as *blockchain*. While the algorithm behind Bitcoin represents a solid safeguard against counterfeiting, the system has proved to be vulnerable to illicit activities such as the massive theft of 350 million

USD worth of Bitcoin from the Mt.Gox exchange in February 2014. The principles of Bitcoin are explained in Dwyer (2014) and at bitcoin.org. While other cryptocurrencies such as Feathercoin, Peercoin exist, Bitcoin has managed to maintain its leading position in this particular market[2]. At the end of June 2016, Bitcoin market capitalization exceeded 10 billion USD (coinmarketcap.com) which represent more than 80% of the total market capitalization of all cryptocurrencies on the market.

In addition to the early large literature on the technical and legal aspects of Bitcoin, the economics and finance debate on Bitcoin have recently mounted. Kristoufek (2014) argues that Bitcoin represents a unique asset possessing properties of both a standard financial asset and a speculative one. Conversely, Popper (2015) considers Bitcoin to be a digital gold. Regardless of whether Bitcoin is a financial or a speculative asset, a digital gold or a commodity, some studies have been interested in the "moneyness" of Bitcoin. Yermack (2013) argue that Bitcoin has no intrinsic value but behaves more like a speculative investment than a currency, because its market capitalization is high compared to the economic transactions it facilitates. The author also concludes that Bitcoin volatility adversely affects its usefulness as a currency. Glaser et al. (2014) find that most of the interest in Bitcoin is due to its "asset" aspect and not its currency aspect. Hanley (2013) also indicates that Bitcoin has no fundamental value to support its pure market valuation against conventional currencies. In contrast, Woo et al. (2013) argue that Bitcoin has some fair value due to its money-like properties. Garcia et al. (2014) and Hayes (2016) show that the cost of producing a Bitcoin via mining adds some fundamental value for Bitcoins.

Other studies have been examined the price formation in the Bitcoin market. Bouoiyour and Selmi (2014) illustrate the significant role of lagged Google search for the word "Bitcoin" in explaining Bitcoin price, while the velocity of Bitcoin measured by data transaction fails to explain Bitcoin price. Similar results regarding the roles of the two abovementioned variables

[2] By the end of June 2016, there were more than 700 cryptocurrencies traded in the market.

(the volume of daily search for Bitcoin on the internet and the number of Bitcoin transactions) in explaining Bitcoin price are reported by Polasik et al. (2014). Within the same research subject, Kristoufek (2014) finds that the trade-exchange ratio plays an essential role in driving Bitcoin price fluctuations in the long-run. Bouoiyour et al. (2014) examine the relations between Bitcoin price and transactions proxied by the exchange-trade ratio. The authors find that Bitcoin price Granger causes exchange-trade ratio in the short- and the medium-run.

Like Kristoufek (2014), they also find that the increasing use of Bitcoin in exchange-trade ratio expands Bitcoin price in the long-run and then there is a significant link that runs from exchange-trade ratio to Bitcoin price, which becomes stronger in the long term. An interesting paper by Ciaian et al. (2016) focuses on the determinants of BitCoin price fluctuations. It shows that the total number of unique Bitcoin transactions per day – a demand side variable – has more impact on BitCoin price than the number of BitCoins – a supply side variable.

We argue that the above literature presents an incomplete picture on the role of trading volume in predicting the Bitcoin returns because the Bitcoin volume–return relationship at the tails may be different from that near the mean of the return distribution. Furthermore, prior studies have overlooked the dependency between the second moment of Bitcoin return and trading volume. These issues suggest the appropriateness of using the nonparametric causality-in-quantiles test of Balcilar et al. (2016 a, b).

3. Methodology

By building on the framework of Nishiyama *et al.*, (2011) and Jeong *et al.*, (2012) we use a novel methodology as advanced by Balcilar *et al.*, (2016a, b), a method that is useful in detecting nonlinear causality through a hybrid approach. The returns on Bitcoin is designated as y_t while the traded volume is designated as x_t. Based on Jeong *et al.,* (2012), we define the quantile-based causality as follow[3]: x_t does not cause y_t in the θ-quantile with regards to the lag-vector of $\{y_{t\text{-}1}, \dots, y_{t\text{-}p}, x_{t\text{-}1}, \dots, x_{t\text{-}p}\}$ if

$$Q_\theta\left(y_t \middle| y_{t\text{-}1}, \dots, y_{t\text{-}p}, x_{t\text{-}1}, \dots, x_{t\text{-}p}\right) = Q_\theta\left(y_t \middle| y_{t\text{-}1}, \dots, y_{t\text{-}p}\right) \quad (1)$$

x_t is presumably cause of y_t in the θ-th quantile with regards to $\{y_{t\text{-}1}, \dots, y_{t\text{-}p}, x_{t\text{-}1}, \dots, x_{t\text{-}p}\}$ if

$$Q_\theta\left(y_t \middle| y_{t\text{-}1}, \dots, y_{t\text{-}p}, x_{t\text{-}1}, \dots, x_{t\text{-}p}\right) \neq Q_\theta\left(y_t \middle| y_{t\text{-}1}, \dots, y_{t\text{-}p}\right) \quad (2)$$

Here, $Q_\theta(y_t | \cdot)$ is the θ-th quantile of y_t. The conditional quantiles of y_t, $Q_\theta(y_t | \cdot)$, depends on *t* and the quantiles are restricted between zero and one, i.e., $0 < \theta < 1$.

For a compact presentation of the causality-in-quantiles tests, we define the following vectors $Y_{t\text{-}1} \equiv (y_{t\text{-}1}, \dots, y_{t\text{-}p})$, $X_{t\text{-}1} \equiv (x_{t\text{-}1}, \dots, x_{t\text{-}p})$, and $Z_t = (X_t, Y_t)$. Let also define the conditional distribution functions $F_{y_t|Z_{t\text{-}1}}(y_t|Z_{t\text{-}1})$ and $F_{y_t|Y_{t\text{-}1}}(y_t|Y_{t\text{-}1})$, which signify the distribution functions of y_t conditioned on vectors Z_{t-1} and Y_{t-1}, respectively. Moreover, the conditional distribution $F_{y_t|Z_{t\text{-}1}}(y_t|Z_{t\text{-}1})$ is presumed to be completely continuous in y_t for nearly all Z_{t-1}. By defining $Q_\theta(Z_{t\text{-}1}) \equiv Q_\theta(y_t|Z_{t\text{-}1})$ and $Q_\theta(Y_{t\text{-}1}) \equiv Q_\theta(y_t|Y_{t\text{-}1})$, we can see that $F_{y_t|Z_{t\text{-}1}}\{Q_\theta(Z_{t\text{-}1})|Z_{t\text{-}1}\} = \theta$, which holds with a probability equal to one. As a result, the hypotheses to be evaluated for the causality-in-quantiles based on the equations (1) and (2) can be represented as:

$$H_0\colon\ P\left\{F_{y_t|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\} = \theta\right\} = 1 \quad (3)$$

$$H_1\colon\ P\left\{F_{y_t|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\} = \theta\right\} < 1 \quad (4)$$

In order to define a measurable metric for the practical implementation of the causality-in-quantiles tests, Jeong *et al.,* (2012) make use of the distance measure $J = \{\varepsilon_t E(\varepsilon_t|Z_{t\text{-}1}) f_Z(Z_{t\text{-}1})\}$, where ε_t denotes the regression error and $f_Z(Z_{t\text{-}1})$ denotes the marginal density function of Z_{t-1}. Consequently, the causality-in-quantiles test is based on the regression error ε_t. The

[3] The exposition in this section closely follows Nishiyama *et al.,* (2011) and Jeong *et al.,* (2012).

regression error ε_t arises based on the null hypothesis specified in equation (3), which would be true, if and only if $E[\mathbf{1}\{y_t \leq Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\}] = \theta$. In order to make the regression error explicit, we rewrite this last statement as $\mathbf{1}\{y_t \leq Q_\theta(Y_{t\text{-}1})\} = \theta + \varepsilon_t$, where $\mathbf{1}\{\cdot\}$ is an indicator function. Now, following Jeong *et al.,* (2012), based on the regression error, the distance metric can be defined as:

$$J = E\left[\left\{F_{y_t|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\}\text{-}\theta\right\}^2 f_Z(Z_{t\text{-}1})\right] \tag{5}$$

In relation to equations (3) and (4), it is crucial to understand that $J \geq 0$. The statement will hold with an equality, i.e., $J = 0$, if and only if the null H_0 in equation (3) is true, while $J > 0$ holds under the alternative H_1 in equation (4). The feasible counterpart of the distance measure J in equation (5) gives us a kernel-based causality-in-quantiles test statistics for the fixed quantile θ and defined as:

$$\hat{J}_T = \frac{1}{T(T\text{-}1)h^{2p}} \sum_{t=p+1}^{T} \sum_{s=p+1,s\neq t}^{T} K\left(\frac{Z_{t\text{-}1}\text{-}Z_{s\text{-}1}}{h}\right)\hat{\varepsilon}_t\hat{\varepsilon}_s \tag{6}$$

where $K(\cdot)$ denotes a known kernel function, h is the bandwidth for the kernel estimation, T denotes the sample size, and p represents the lag-order used for defining vector Z_t. Jeong *et al.* (2012) establish that the re-scaled statistics $Th^p\hat{J}_T/\hat{\sigma}_0$ is asymptotically distributed as standard normal, where $\hat{\sigma}_0 = \sqrt{2}\theta(1-\theta)\sqrt{1/(T(T\text{-}1)h^{2p})}\sqrt{\sum_{t\neq s} K^2((Z_{t\text{-}1}\text{-}Z_{s\text{-}1})/h)}$. The most crucial element of the test statistics $\hat{J}_T$ is the regression error $\hat{\varepsilon}_t$. In our particular case, the estimator of the unknown regression error is defined as:

$$\hat{\varepsilon}_t = \mathbf{1}\{y_t \leq \hat{Q}_\theta(Y_{t\text{-}1})\}\text{-}\theta \tag{7}$$

In equation (7), the quantile estimator $\hat{Q}_\theta(Y_{t\text{-}1})$ yields an estimate of the θ-th conditional quantile of y_t given Y_{t-1}. We estimate $\hat{Q}_\theta(Y_{t\text{-}1})$ by employing the nonparametric kernel approach as:

$$\hat{Q}_\theta(Y_{t\text{-}1}) = \hat{F}^{-1}_{y_t|Y_{t\text{-}1}}(\theta|Y_{t\text{-}1}) \tag{8}$$

where $\hat{F}_{y_t|Y_{t\text{-}1}}(y_t|Y_{t\text{-}1})$ denote the *Nadarya-Watson* kernel estimator given by:

$$\hat{F}_{y_t|Y_{t\text{-}1}}(y_t|Y_{t\text{-}1}) = \frac{\sum_{s=p+1,s\neq t}^{T} L\left(\frac{Y_{t\text{-}1}\text{-}Y_{s\text{-}1}}{h}\right)\mathbf{1}\{y_s \leq y_t\}}{\sum_{s=p+1,s\neq t}^{T} L\left(\frac{Y_{t\text{-}1}\text{-}Y_{s\text{-}1}}{h}\right)} \tag{9}$$

with $L(\cdot)$ denote a known kernel function and h is the bandwidth used in the kernel estimation. It is of interest for us also to test for Granger causality from the traded volume to variance (square) of Bitcoin returns. The causality in variance implies volatility transmission, which may exist even there is no causality in the mean (1st moment). Testing for Granger causality in the second or higher moments has some complications and the procedure for such tests should be carefully defined since rejection of causality in the moment m does not imply non-causality in the moment k for $m < k$. We begin by employing Nishiyama *et al.,* (2011) nonparametric Granger quantile causality method. In order to demonstrate the causality in higher order moments, first we examine the process below for y_t:

$$y_t = g(Y_{t\text{-}1}) + \sigma(X_{t\text{-}1})\varepsilon_t \tag{10}$$

where ε_t denote an independently and identically distributed (*iid*) process; the unknown functions $\sigma(\cdot)$ and $g(\cdot)$ satisfy some properties that are sufficient for the stationarity of y_t. Although, this representation does not permit linear or non-linear causalities from $X_{t\text{-}1}$ to y_t, it does allow $X_{t\text{-}1}$ to have predictive content for y_t^2 when $\sigma(\cdot)$ is an established nonlinear function. The representation in equation (10) illustrates that squares for X_{t-1} does not necessarily enter into the nonlinear function $\sigma(\cdot)$. Thus, we re-specify equations (3) and (4) into a null H_0 and alternative H_1 hypothesis for causality in variance as follows:

$$H_0\text{:}\;\; P\left\{F_{y_t^2|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\} = \theta\right\} = 1 \tag{11}$$

$$H_1\text{:}\;\; P\left\{F_{y_t^2|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\} = \theta\right\} < 1 \tag{12}$$

In order to get a feasible test statistic for testing the null hypothesis H_0 in equation (11), we substitute y_t in equations (6) to (9) with y_t^2. A problem may arise with the causality test based on the definition given in equation (10), since there may be causality in the second moment (variance) along with the causality in the conditional first moment (mean). We can illustrate this with the following model:

$$y_t = g(X_{t\text{-}1}, Y_{t\text{-}1}) + \varepsilon_t \tag{13}$$

Therefore, the higher order causality-in-quantiles can be stated as:

$$H_0:\ P\left\{F_{y_t^k|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\} = \theta\right\} = 1 \quad \text{for } k = 1,2,\dots,K \tag{14}$$

$$H_1:\ P\left\{F_{y_t^k|Z_{t\text{-}1}}\{Q_\theta(Y_{t\text{-}1})|Z_{t\text{-}1}\} = \theta\right\} < 1 \quad \text{for } k = 1,2,\dots,K \tag{15}$$

Incorporating the whole concept, we specify that x_t Granger causes y_t in quantile θ up to K-th moment using equation (14) to formulate the test statistic of equation (6) for each k. Nishiyama *et al.* (2011) construct nonparametric Granger causality tests using the density-weighted approach as in Jeong *et al.* (2011) and show that density-weighted nonparametric tests in higher moments have the same asymptotic normal distribution as the test for causality in first moment, although some stronger moment conditions might be necessary. Nevertheless, it is not an easy task to test for all $k = 1,2,\dots,K$ jointly, since the statistics are jointly correlated (Nishiyama *et al.*, 2011). In order to systematically overcome this issue, we follow the sequential testing approach in Nishiyama *et al.* (2011) to test for causality in both models defined in equations (10) and (13). In this approach, we first test for nonparametric Granger causality in the first moment ($k = 1$), but still continue for testing causality in variance even if the non-causality is not rejected. That is, if the null for $k = 1$ is not rejected, then there might still be causality in the second moment and, thus, we construct the tests for $k = 2$. This approach allows us to test the existence of causality only in variance as well as the causality in the mean and variance successively. Conclusively, we can investigate the existence of causality-in-mean and causality-in-variance sequentially. The empirical application of causality testing through quantiles require identifying three crucial choices: the lag order p, the bandwidth h, and the kernel type for $K(\cdot)$and $L(\cdot)$ in equations (6) and (9), respectively. In this study, we make use of lag order of 7 based on the Schwarz Information Criterion (SIC) under a VAR involving Bitcoin returns and traded volume. Moreover, when it comes to choosing lags, the SIC is considered being parsimonious compared to other lag-length selection criteria. The SIC helps overcome the issue of overparametrization usually arising with nonparametric frameworks.[4] The bandwidth value is chosen by employing the least squares cross-validation techniques.[5] Finally, for $K(\cdot)$ and $L(\cdot)$ Gaussian-type kernels was employed.

4. Data and Empirical Findings

4.1 Data

Our analysis is based on two variables, namely, the Bitcoin index and the trading volume as a measure the level of trading activity. The Bitcoin price index is denominated in USD, the currency against which Bitcoin is the most traded on Bitstamp – the largest European Bitcoin exchange (Brandvold et al., 2015). The data is sourced from: www.bitcoincharts.com.

Both the Bitcoin index and volume are non-stationary in log-levels as indicated by standard unit root tests.[6] Since our methodology requires stationary data, we work with Bitcoin returns, obtained as the first-differences of the natural logarithmic values of the index expressed in percentage. The squared values of returns measure the volatility of the Bitcoin returns. On the other hand, following Gebka and Wohar (2013), we use a detrended measure of volume. Specifically, we consider the natural log of the volume series and remove its trend by regressing it on a constant, (t/T) and $(t/T)^2$, where T is the total sample size. Our period of analysis covers the

[4] Hurvich and Tsai (1989) examine the Akaike Information Criterion (AIC) and show that it is biased towards selecting an overparameterized model, while the SIC is asymptotically consistent.

[5] For each quantile, we determine the bandwidth h using the leave-one-out least-squares cross validation method of Racine and Li (2004) and Li and Racine (2004).

[6] Complete details of the unit root tests are available upon request from the authors.

daily period of 19th December, 2011 to 25th April, 2016 (i.e., 1587 observations), with the start and end date being purely driven by data availability. Interestingly, the sample period covers the Bitcoin crash of December 2013 (Cheah and Fry, 2015) and the recovery that started in the fourth quarter of 2014, thus, it allows us to examine how the return-volume relationship in the Bitcoin market was affected as a result.

Table 1 presents the summary statistics for the Bitcoin returns and traded volume. We observe that volume is more volatile than returns in the Bitcoin market. More importantly for our context of causality-in-quantiles, both the variables are skewed to the left, with excess kurtosis, resulting in non-normal distributions, as indicated by the strong rejection of Jarque-Bera statistic at 1 percent level of significance. The heavy-tails of the distributions of both returns and volume provide a preliminary justification for the causality-in-quantiles test used in this paper. The natural logarithm of the data for the Bitcoin index and traded volume, and their respective transformations to returns and detrended volume, are presented in Figure 1. We clearly notice the long bull market that lasted almost three years before it ended in December 2013 - the month during which we captured a major structural break in Bitcoin prices (see the next sub-section for more details), with such an observation also made by Bouri et al., (2016).

[INSERT TABLE 1]

4.2 Empirical Findings

Though our objective is to analyse the causality-in-quantiles running from volume to the Bitcoin returns and its volatility, for the sake of completeness and comparability, we also conducted the standard linear Granger causality test based on a VAR(7) model. The resulting $\chi^2(7)$ statistic for the null that volume does not *Granger* cause returns is 9.4305 with a p-value of 0.2232. In other words, there is no evidence of predictability emanating from volume to returns based on the linear causality test even at the ten percent level of significance.[7]

Next, to motivate the use of the nonparametric quantile-in-causality approach, we statistically investigate the possibility of nonlinearity in the relationship between the Bitcoin returns and volume. To this end, we apply the Brock *et al.,* (1996, BDS) test on the residuals of an AR(7) model for returns, and the returns equation in the VAR(7) model involving traded volume. As can be seen from Table 2, we find strong evidence, at highest level of significance, for the rejection of the null of *i.i.d.* residuals at various embedded dimensions (*m*).[8] These results provide strong evidence of nonlinearity in not only the returns, but also in its relationship with volume. This means that, the result of causality based on the linear Granger causality test, cannot be deemed robust and reliable.

[INSERT TABLE 2]

Next, we turn to the powerful *UDmax* and *WDmax* tests of Bai and Perron (2003), which determines 1 to *M* (multiple) globally determined breaks, applied again to the AR(7) model for Bitcoin returns, and the returns equation in the VAR(7) model involving the volume. There is only one break (19th December, 2013) in the AR(7) model of returns, that corresponds to the Bitcoin price crash of December 2013 identified in Cheah and Fry (2015). Two breaks (18th April, 2013 and 18th December, 2013) are detected for the returns equation in the VAR(7) model involving volume;[9] the first corresponds to the Bitcoin bubble of April 2013 following the technical glitch in the Bitcoin software as identified in Fry and Cheah (2016), whereas the second break of December 2013 was aforementioned identified in the return series. So, as under the

---

7 The $\chi^2(7)$ statistic for the null that returns does not *Granger* cause volume is 9.9290 with a *p*-value of 0.1926, i.e., returns does not cause volume based on the linear causality test at conventional levels of significance.

8 Similar results were obtained for the AR(7) model of volume, and for the volume equation in the VAR(7) model involving the returns series. Hence, not only is volume nonlinear but also it has a nonlinear relationship with returns, implying non-reliability of the linear test of causality running from returns to volume. Complete details of these results are available upon request from the authors.

9 The AR(7) model for volume, as well as, the volume equation in the VAR(7) model with Bitcoin returns was found to have a regime change at 22nd October, 2013. Complete details of these results are available upon request from the authors. So, based on the BDS and the structural break tests, we can conclude that traded volume evolves in a nonlinear fashion and is also related nonlinearly with returns.

BDS test which detected nonlinearity, existence of structural breaks in the returns, and in its relationship with the traded volume, imply that the Granger causality tests based on a linear framework is likely to suffer from misspecification. Given the strong evidence of nonlinearity and regime changes in the relationship between returns and volume, we now turn our attention to the causality-in-quantiles test, which is robust to linear misspecification, being a nonparametric (i.e., data-driven) approach.

In Figure 2, we present the results obtained from the quantile causality test for Bitcoin returns and squared returns (i.e., volatility) due to the traded volume. As can be seen, the null that volume does not *Granger* cause returns is rejected at the five percent level of significance (critical value of 1.96) over the quantile range (τ) of 0.25 to 0.75 of the conditional distribution of returns. However, we fail to reject the null of volume does not *Granger* cause volatility over the entire conditional distribution.[10] In other words, volume can predict returns, but not volatility, with the evidence of causality for returns holding over the entire conditional distribution barring the two ends. Specifically speaking, volume has no predictive content for Bitcoin returns when the market is in bearish (lower quantiles) and bullish (upper quantiles) phases. But, when the market is functioning around the normal (median) mode, volume can indeed predict returns.

So in sum, evidence that volume predicts Bitcoin returns is non-existent in the linear model. However, as we show, the standard Granger causality result cannot be relied upon due to the existence of nonlinearity and structural breaks. Given this, when we look into the nonparametric causality-in-quantiles test, which is robust to misspecifications, we find evidence that volume can predict returns around the median of the conditional distribution, but not volatility. This result seems to suggest that investors in the Bitcoin market can obtain valuable predictive information for returns from traded volume, when the market is in normal mode. This result confirms that

[10] We also conducted the causality-in-quantiles test with detrended volume and its squared value as dependent variables. However, we could not detect any evidence of causality for either volume or squared volume resulting from returns at any point of the respective conditional distributions of the dependent variables. Complete details of these results are available upon request from the authors.

the Bitcoin market is generally weak-form efficient. But in bear and bull regimes, volume plays no role in predicting returns, with past (seven) lags of returns providing relatively more valuable information for future returns than volume. This finding partially implies that in stress periods, and unlike the case of equities, trading volume is not associated with extremely high/low returns in the Bitcoin market.

These findings on returns predictability based on volume add to information documented in prior studies trying to define the underlying factors that determine the price of Bitcoin (see for example, Bouoiyour and Selmi, 2014; Bouoiyour et al., 2014; Polasik et al., 2014; Kristoufek, 2014; Ciaian et al., 2016). In fact, our findings provide a broader picture of the dependence between volume and returns in the Bitcoin market and uncover the insignificant role of trading volume in predicting Bitcoin price volatility. In terms of the finding of returns predictability around the median, our result is different from what is observed with equity markets, where the predictability of returns from volume holds mostly at the tails (see for example, Chuang et al., (2009), Gebka and Wohar (2013)). This could possibly be due to the short-selling constraint in the Bitcoin market in the lower-end of the market. While, when the market is booming there is more herding into the market and less interest in searching out for relevant information (in this case volume) on behalf of the market agents. The lack of predictability of volatility due to traded volume provides support to the so-called Mixture of Distribution Hypothesis (MDH) developed by Clark (1973). MDH assumes that the volume-volatility relation is dependent on the rate of information flow into the market. The basic intuition is that, because all traders simultaneously receive new information, there will be no intermediate equilibrium. Since, the variables contemporaneously change in response to new information, it is impossible to use past volume data to predict volatility. The possibility of MDH holding in the Bitcoin market rather than the Sequential Information Arrival Hypothesis (SIAH; Copeland, 1976) is likely to be higher due to possible easy dissemination of information across traders, given that Bitcoin involves an open source software based online payment system.

5. Conclusion

There exit a large literature which has analyzed the role of traded volume in predicting movements in stock returns and volatility (see Gebka and Wohar (2013) for a detailed literature review in this regard). We build on this literature by analyzing the predictability of traded volume for the returns and volatility of the Bitcoin market. In this regard, we use daily data covering the period of 19th December, 2011 to 25th April, 2016. To statistically evaluate predictability, we first apply standard linear Granger causality test, but fail to detect any evidence of volume causing returns. Test of nonlinearity indicate that returns and its relationship with volume evolves in a nonlinear manner. In addition, evidence of regime changes is also detected for returns and in the equation relating it to volume, based on tests of multiple structural breaks. The evidence of nonlinearity and structural breaks in turn, implies that the linear set-up to test causality is in fact misspecified, with the results obtained from the same being unreliable. Given this, we resort to a nonparametric causality-in-quantiles test, which combine elements of the test for nonlinear causality of k-th order developed by Nishiyama *et al.* (2011) with the causality-in-quantiles test developed by Jeong *et al.* (2012). The causality-in-quantile approach allows us to test for not only causality-in-mean, but also causality that may exist in the tails of the joint distribution of the variables. In addition, we are also able to investigate causality-in-variance (volatility spillovers), when causality in the conditional-mean may not exist, yet higher order interdependencies might emerge. The test reveals that the null that volume does not *Granger* cause returns is rejected at the conventional levels of significance over the quantile range of 0.25 to 0.75 of the conditional distribution of returns. However, we fail to reject the null of volume does not *Granger* cause volatility over the entire conditional distribution. In other words, volume can predict returns, but not volatility, with causality for returns non-existent in bearish (lower quantiles) and bullish (upper quantiles) phases. But, when the market is functioning around the normal (median) mode, volume can indeed predict returns, thus providing investors in the Bitcoin market with valuable predictive information for returns. This result implies that, when the market is performing well or poorly, all that matters for predicting future returns is its past values, and information from volume is irrelevant. In general, our results, via the volume-returns causality in the Bitcoin market, highlight the importance of detecting and modeling nonlinearity when analyzing predictability via causal relationships.

References

Bai, J. and Perron, P. 2003. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18, 1-22.

Balcilar, M., Gupta, R., Kyei, C., and Wohar, M.E. 2016a. Does Economic Policy Uncertainty Predict Exchange Rate Returns and Volatility? Evidence from a Nonparametric Causality-in-Quantiles Test. Open Economies Review, 27 (2), 229-250.

Balcilar, M., Gupta, R., Pierdzioch, C. 2016b. Does uncertainty move the gold price? New evidence from a nonparametric causality-in-quantiles test. Resources Policy, 49(1), 74-80.

Balduzzi, P., Elton, E. and Green, T. (2001). Economic News and Bond Prices: Evidence from the U.S. Treasury Market. Journal of Financial and Quantitative Analysis, 36(4), 523–43.

Bekiros, S., Gupta, R., and Majumdar, A., (Forthcoming). Incorporating Economic Policy Uncertainty in US Equity Premium Models: A Nonlinear Predictability Analysis. Finance Research Letters.

Bouoiyour, J. and Selmi, R. (2014). What Bitcoin Looks Like? MPRA Paper No. 58091.

Bouoiyour, J., Selmi, R. and Tiwari, A. (2014). Is BitCoin business income or speculative bubble? Unconditional vs. conditional frequency domain analysis. MPRA Paper No. 59595.

Bouri, E., Gil-Alana, L.A., Gupta, R., Roubaud, D. (2016). Modelling Long Memory Volatility in the Bitcoin Market: Evidence of Persistence and Structural Breaks. Department of Economics, University of Pretoria, Working Paper No. 2016-54.

Brandvold, M., Molnár, P., Vagstad, K., and Valstad, O. C. A. (2015). Price discovery on Bitcoin exchanges. Journal of International Financial Markets, Institutions and Money 36: 18–35.

Brock, W., Dechert, D., Scheinkman, J., LeBaron, B., 1996. A test for independence based on the correlation dimension. Econometric Reviews*, 15* 197–235.

Cheah, E-T. and Fry, J. M. (2015). Speculative bubbles in Bitcoin markets? An empiricalinvestigation into the fundamental value of Bitcoin. Economic Letters, 130, 32–36.

Chen, G.M., Firth, M. and Rui, O. (2001). The Dynamic Relation between Stock Returns, Trading Volume and Volatility. The Financial Review, 36(3), 153–174.

Chen, (Cathy) W.S., So, (Mike) K.P., and Chiang, (Thomas) C. (2016). Evidence of Stock Returns and Abnormal Trading Volume: A Threshold Quantile Regression Approach, 67(1), 96–124.

Chiang, T. C. and J. Li (2012) Stock Returns and Risk: Evidence from Quantile Regression Analysis, Journal of Risk and Financial Management, 5, 20–58.

Chiarella, C., Kang, B., Nikitopoulos, C. S., and To, T-D. (2016). The return-volatility relation in commodity futures markets. Journal of Futures Markets, 36, 127–152.

Chuang, C. C., C. M. Kuan and H. Y. Lin (2009) Causality in Quantiles and Dynamic Stock Return-Volume Relations, Journal of Banking and Finance, 33, 1351–1360.

Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices, Econometrica, 41, 135-155.

Copeland, T. (1976). A model of asset trading under the assumption of sequential information arrival, Journal of Finance, 31, 1149-1168.

Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. Journal of Financial Stability, 17, 81-91.

Fry, J. and Cheah, E-T. (2016). Negative bubbles and shocks in cryptocurrency markets. International Review of Financial Analysis, doi:10.1016/j.irfa.2016.02.008.

Garcia, D., Tessone, C., Mavrodiev, P. and Perony, N. (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. Journal of the Royal Society Interface, 11(99), 1-8.

Gebka, B., and Wohar, M.E. (2013). Causality between trading volume and returns: Evidence from quantile regressions. International Review of Economics and Finance, 27, 144–159.

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

Hanley, B. P. (2013). The false premises and promises of Bitcoin. arXiv preprint arXiv:1312.2048.

Hayes, A.S. (2016). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. Telematics and Informatics, doi:10.1016/j.tele.2016.05.005.

Hurvich, C.M. and Tsai, C.-L. (1989). Regression and Time Series Model Selection in Small Samples. Biometrika, 76, 297-307.

Jeong, K., Härdle, W. K. and Song, S., 2012. A consistent nonparametric test for causality in quantile. Econometric Theory 28, 861-887.

Karpoff, J.M. (1987). The relation between price changes and trading volume: a survey. Journal of Financial and Quantitative Analysis, 22, 109–126.

Kristoufek, L. (2014). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. http://arxiv.org/pdf/1406.0268.pdf.

Li, Q. and Racine, J.S. (2004). Cross-validated local linear nonparametric regression. Statistica Sinica, 14, 485-512.

Lin, H-Y. (2013). Dynamic Stock Return- Volume Relation: Evidence from Emerging Asian Markets. Bulletin of Economic Research, 65(2), 178-193.

Marsh, T. A. and Wagner, N. (2000). Return-Volume Dependence and Extremes in International Markets. Working Paper RPF-293, Research Program in Finance, University of California, Berkeley.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System.

Nishiyama, Y., Hitomi, K., Kawasaki, Y., and Jeong, K. 2011. A consistent nonparametric Test for nonlinear causality - specification in time series regression. Journal of Econometrics 165, 112-127.

Polasik, M., Piotrowska, A., Wisniewski, T. P., Kotkowski, R. and Lightfoot, G. (2014). Price Fluctuations and the Use of Bitcoin: An Empirical Inquiry. Available at SSRN 2516754.

Popper, N. (2015). Digital gold: The untold story of Bitcoin. London: Penguin.

Puri, T.N. and Philippatos, G.C. (2008). Asymmetric Volume-Return Relation and Concentrated Trading in LIFFE Futures. European Financial Management, 14, 528 - 563.

Racine, J.S. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. Journal of Econometrics, 119, 99-130.

Woo, D., Gordon, I. and Iarolov, V. (2013) Bitcoin: A First Assessment. FX and Rates, Global. Bank of America. Merrill Lynch. December.

Yermack, D. (2013). Is Bitcoin a real currency? An economic appraisal (No. w19747). National Bureau of Economic Research. URL: http://www.nber.org/papers/w19747.

Table 1. Summary Statistics

	Variable	
Statistic	**Returns**	**(Detrended) Volume**
Mean	0.0031	0.0013
Median	0.0020	0.0554
Maximum	0.3375	2.7966
Minimum	-0.6639	-6.6199
Std. Dev.	0.0516	0.9022
Skewness	-1.6387	-0.7632
Kurtosis	28.8730	6.7148
Jarque-Bera	44975.1100	1066.5960
Probability	0.0000	0.0000

Note: Std. Dev. stands for standard deviation, while probability corresponds to the Jarque-Bera test of normality.

Table 2. BDS Test Statistic

Model of Returns Equation	m				
	2	3	4	5	6
AR(7)	12.9500***	15.5327***	17.4181***	19.1123***	20.9129***
VAR(7)	13.0976***	15.6403***	17.5329***	19.2779***	21.1136***

Note: m stands for the number of (embedded) dimension which embed the time series into m-dimensional vectors, by taking each m successive points in the series. Values in the cells represent BDS z-statistic; *** indicates rejection of the null of *i.i.d.* residuals at 1 percent level of significance.

Figure 1a. Natural Logarithm of Closing Price of Bitcoin

Figure 1b. Natural Logarithm of Volume Traded

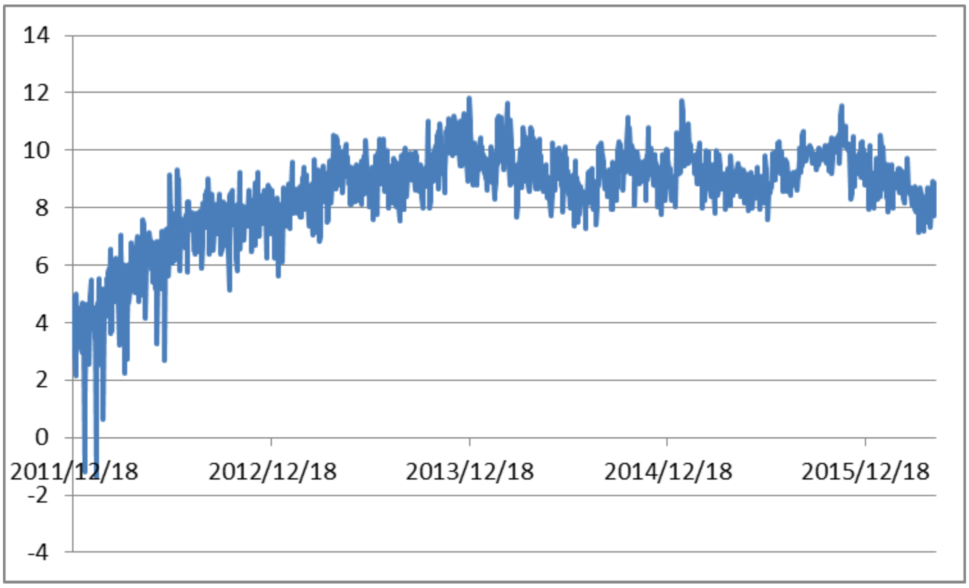


Figure 1c. Bitcoin Returns

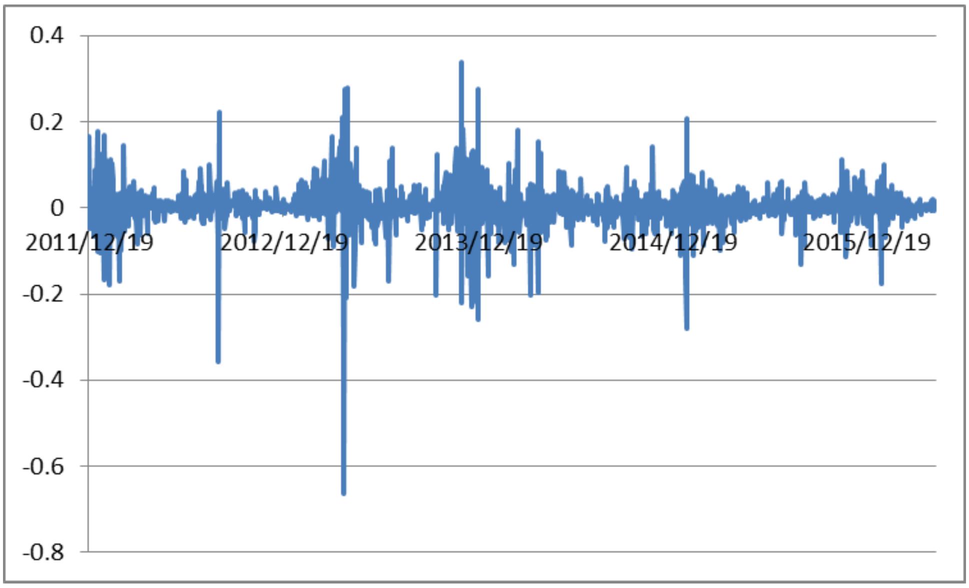


Figure 1d. Detrended Volume

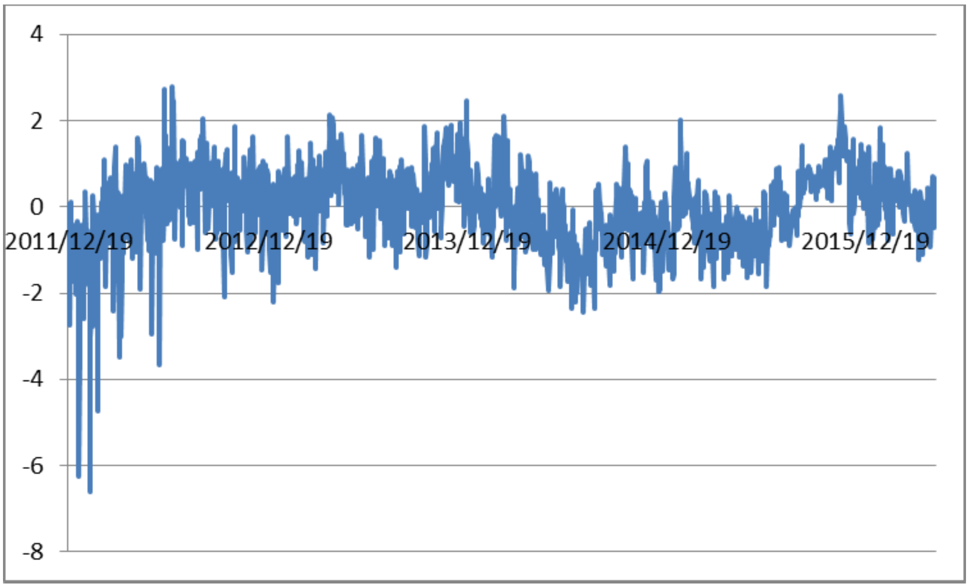
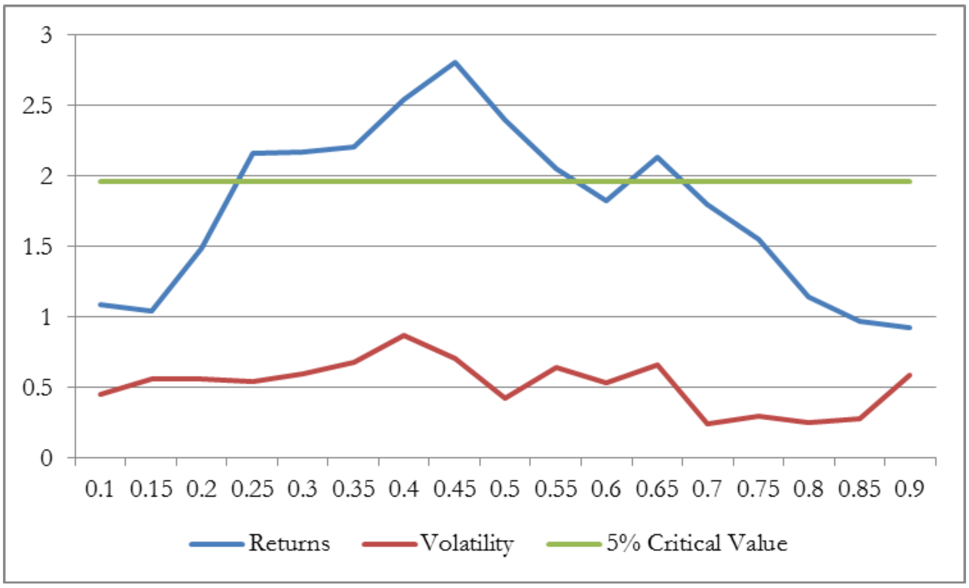


Figure 2. Causality-in-Quantiles: Volume does not Granger cause Bitcoin Returns and Volatility



Note: Vertical axis presents the test statistics corresponding to the null the volume does not *Granger* cause returns, and volume does not *Granger* cause volatility; Horizontal axis measures the quantiles; 5 percent critical value is 1.96.