Modelling Common Bubbles in Cryptocurrency Prices

Mauri Hall*, and Joann Jasiak †

This version: April 8, 2024

Abstract

The bubbles and spikes in cryptocurrency prices increase considerably the risk on investments in these assets. In the traditional time series literature bubbles are viewed as nonstationary and non-estimable components of a process. In this paper, we adopt a different approach and consider the bubbles as inherent features of a strictly stationary causal-noncausal (mixed) Vector Autoregressive (VAR) process. This approach allows us to model and estimate the common bubbles and spikes in cryptocurrency prices. It also provides us linear combinations of cryptocurrencies that eliminate common bubbles analogously to the cointegrating vectors eliminating common trends in unit root processes. They are used to build cryptocurrency portfolios immune to the risk of common bubbles that ensure stable investment strategies. The mixed VAR model is estimated from the US Dollar prices of Bitcoin, Ethereum, Ripple, and Stellar over the period 2017-2019. We document the common bubbles and illustrate the behavior of bubble-free portfolios.

Keywords: Noncausal Process, Bubble, Bitcoin, Ethereum, Cryptocurrency.

*York University, Canada, *e-mail*: hall1472@yorku.ca.

†York University, Canada, *e-mail*: jasiakj@yorku.ca.

The authors thank A. Hecq and C. Gourieroux for helpful comments. We acknowledge financial support of York University's Catalyzing Interdisciplinary Research Cluster "Digital Currencies" and Natural Sciences and Engineering Council of Canada (NSERC).

1 Introduction

The cryptocurrency prices display bubbles and spikes that considerably increase the risk associated with investments in these assets. The bubbles and spikes, referred to short-lived, i.e. local explosive patterns in Gourieroux & Zakoian (2017) are distinctive features of cryptocurrency prices due to their highly speculative character, in addition to time varying volatility, a nonlinear characteristic they share with fiat currency exchange rates and stock returns. Despite of high risks, the cryptocurrencies continue to attract investors, with the recently introduced Exchange Traded Funds (ETF) on Bitcoin and Ether contributing to further expansion of cryptocurrency trading. At the beginning of 2024, Bitcoin and Ethereum are the market leaders with respective capitalizations of over 1.3 Trillions and 300 Billions US Dollars. Ripple and Stellar are medium and small size cryptocurrencies with market capitalizations of over 30 and 4 Billions US Dollars, respectively.

In this paper, we study the comovements of US Dollar prices of these four cryptocurrencies and document their common bubbles and spikes. Our objective is to model and estimate the common bubbles in cryptocurrency, as well as to reveal and estimate linear combinations of cryptocurrencies that eliminate them, by analogy to the cointegrating vectors eliminating common trends in unit root processes, or common features of Engle, Kozicki (1993). The linear combinations of cryptocurrencies that eliminate common bubbles in cryptocurrency can be interpreted as portfolios hedging against the systemic risk due to common bubbles and ensuring the stability of investment strategy.

Under the traditional approach to time series analysis, bubbles are viewed as nonstationary patterns to be detected by unit root-type tests and separated from the stationary component of time series. There exists a variety of bubble models, such as the Watson bubble for example Blanchard & Watson (1982), and tests for bubbles such as those proposed in P. Phillips & Shi (2018), and P. C. B. Phillips et al. (2015a) and P. C. B. Phillips et al. (2015b). In this paper, an alternative approach is used. The cryptocurrency rates are modelled as a strictly stationary (mixed) causal-noncausal Vector Autoregressive (VAR) process. In this framework, bubbles and local explosive patterns, in general, arise as the inherent features of this strictly stationary multivariate process. This makes the common bubbles estimable and allows us to model and trace out their dynamics, which cannot be done when bubbles are considered nonstationary. In addition, linear combinations of cryptocurrencies that eliminate common bubbles follow directly from the state-space representation of the causal-noncausal VAR model and are estimable and predictable as well. Hence, a common bubble in cryptocurrency rates can be interpreted as a "common feature", in the sense of Engle & Kozicki (1993) and compared with bubble cointegration of Cubadda, Giancaterini, et al. (2023). The bubble-free linear combinations of cryptocurrency represent stable investment portfolios.

The comovements of cryptocurrency rates and the presence of common bubbles in particular, can be explained not only by the speculative character of these digital assets, but also by the integration of cryptocurrency market with traditional financial markets, changes in the regulatory environment and investor sentiment about blockchain technology. Investor sentiment with respect to blockchain technology in general and cryptocurrencies in particular plays a significant role. Youssef & Waked (2022) argues that cryptocurrencies exhibit evidence of herding effects, especially in response to media coverage. The cryptocurrency markets are also influenced by the global financial markets. When large institutions invest in cryptocurrencies, they reveal increased interest in this entire asset class. Then spot market investors follow by incorporating information from cryptocurrency futures markets which are typically dominated by institutional investors [Doan B. & Nguyen Thanh (2022)]. In addition, the regulatory environment for cryptocurrencies is subject to abrupt changes which impact the entire cryptocurrency market simultaneously. Bhatnagar et al. (2023) has shown that news shocks have persistent and simultaneous effects on volatility of large cap cryptocurrency returns.

The relationships between cryptocurrency returns have received a lot of attention in recent years [see e.g. Cross et al. (2021), Antonakakis et al. (2019), Bouri et al. (2021), Dunbar & Owusu-Amoako (2022) and Nyakurukwa & Seetham

(2023). We model the rates with respect to the US Dollar (USD), i.e. cryptocurrency prices rather than returns, because their variation is bounded and the explosive patterns are local, i.e. are not infinitely lasting. Hence the rates seem to satisfy the assumption of strict stationarity, whereas the assumption of second-order (weak) stationarity appears too stringent. We examine the bivariate series of cryptocurrency rates BTC/USD and ETH/USD, on the one hand, and XRP/USD and XLM/USD on the other. Pairing up the cryptocurrencies this way is motivated by the fact that Bitcoin and Ethereum are the market capitalization leaders while Ripple and Stellar serve similar purposes and share similar technological features. Next, to detect the common dynamic patterns in all four cryptocurrency rate series, we examine a single multivariate process of cryptocurrency rates of dimension four.

So far, the applied literature on cryptocurrencies has employed the causal i.e. past dependent models, often including nonlinear features. For example, Catania et al. (2019) find that combinations of parameter varying models can improve the inference. We accommodate the nonlinearities along with the explosive features in cryptocurrency rates through the presence of a noncausal component in the mixed VAR model.

A mixed causal-noncausal VAR process [see Gouriéroux & Jasiak (2017), Davis & Song (2020)] (henceforth referred to as the mixed VAR) has a representation similar to the traditional VAR model in the sense that the current value of a multivariate process is written as a linear function of the past values. However, unlike the traditional VAR, the mixed VAR has an autoregressive matrix of coefficients with eigenvalues of modulus either strictly smaller or greater than one. The eigenvalue(s) strictly smaller than one are associated with the traditional past dependent i.e. causal stationary behaviour of the series. The eigenvalue(s) strictly larger than one are associated with locally explosive behaviour, generating the bubbles and spikes. In addition, the errors of the mixed VAR have to be non-Gaussian, identically distributed (i.i.d.) and serially independent to ensure that the dynamics of the mixed process can be identified.

The assumption of Gaussianity, common in the traditional time series literature is not supported by the empirical evidence on the fat tails of sample densities of cryptocurrency prices. This assumption imposed on the errors of a traditional VAR model implies that the forward and backward dynamics of that multivariate process are not distinguishable and, as a consequence the forward-looking, or noncausal component(s) cannot be identified. As mentioned earlier, the noncausal component captures the explosive features, and hence the assumption of error Gaussianity is too restrictive for the modelling of cryptocurrency rates. Our approach relaxes the traditional assumption of error Gaussianity and second-order stationarity of the VAR process allowing for the presence of noncausal components to accommodate the explosive and nonlinear patterns, including bubbles and spikes. The mixed VAR model enables us to filter out and estimate the causal and noncausal components, the latter one capturing the common bubbles in strictly stationary multivariate time series.

The estimation of common bubble dynamics leads to the monitoring and forecasting of these explosive patterns. Bubble monitoring can be beneficial for the investors because large changes in cryptocurrency rates can either adversely affect the returns or provide investment opportunities. The common bubble can also be predicted [see Gouriéroux & Jasiak (2016) , Lanne et al. (2012)]. Moreover, as mentioned earlier, the estimated state-space representation of the mixed VAR model provides us the linear combination eliminating the common bubbles i.e. a cryptocurrency portfolio that is risk neutral and hedges the investors against the risk of common bubbles and explosive patterns in general, ensuring a stable investment strategy.

To estimate the mixed VAR model we apply the Generalized Covariance estimator Gouriéroux & Jasiak (2017) which is a one-step, consistent semiparametric estimator for mixed causal noncausal multivariate non-Gaussian processes. The advantage of this approach is the possibility to study the cryptocurrency rates in a semi-parametric setup, i.e. without imposing any distributional assumptions on the errors of the VAR model, except for non-normality, which is justified by non-normal sample distributions of cryptocurrency prices. In addition, we show that while the (mixed) causal-noncausal Vector Autoregressive (VAR) process provides a good fit to the cryptocurrency rates, the traditional, i.e. past-dependent causal VAR model is flawed and fails to detect the comovements of cryptocurrencies.

The paper is organized as follows: Section 2 discusses the causal-noncausal Vector Autoregressive (VAR) model and the GCov estimator. Section 3 introduces the time series of cryptocurrencies : Bitcoin (BTC), Ethereum (ETH), Ripple (XRP) and Stellar (XLM) and shows the results on the empirical analysis of their respective USD exchange rates. Section 4 concludes. The Appendix contains the VAR(1) estimation results for Bitcoin and Ethereum. The summary statistics and supplementary graphs are provided online in Supplementary Material file.

2 Methodology

A mixed causal-noncausal VAR process resembles the traditional VAR model in that the current value of a multivariate process is written as a linear function of the past values. However, unlike the traditional VAR, the mixed VAR has an autoregressive matrix with eigenvalues of modulus either less or greater than 1.

The Vector Autoregressive of order p (VAR(p)) model representing the dynamics of a weakly stationary multivariate process $y_t, t = 1, 2, ...,$ of dimension n is:

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + ... + \Phi_p y_{t-p} + e_t, \tag{1}$$

where $\Phi_i, i = 1, ..., p$ are $n \times n$ matrices of autoregressive coefficients, e_t is an error vector of length n which follows a weak or strong white noise with mean zero and a positive definite variance matrix Σ. [1]

Under the classical approach y_t is also assumed causal, i.e. past dependent. This condition implies that the roots of $det(Id - \Phi_1 z - \Phi_2 z^2 - ... - \Phi_p z^p) = 0$ lie outside the unit circle and eliminates potential noncausal components in the dynamics. Even if the causality assumption were not imposed, the traditional normality-based estimation methods such as the normality-based Maximum Likelihood (ML) and Ordinary Least Squares (OLS) do not distinguish between the causal and noncausal dynamics due to the lack of identification issue. The are routinely implemented in practice under the standard Box-Jenkins approach for time series analysis, which is adequate only for causal time-series which are normally distributed, stationary and linear, admitting a moving average representation in weak white noise errors. The reason is that it is based on the identification and estimation of time series from the moments of order up to two only. Consequently, the normality-based methods are unable to accommodate bubbles, spikes and local trends which involve higher-order moments.

If a time-series is strictly stationary, noncausal and non-Gaussian, then we are able to distinguish its two-sided moving average representation (including the past present and future errors) written in terms of independent and identically distributed (i.i.d.) non-Gaussian errors from a one-sided moving average representation including only the current and lagged errors.

To model the mixed dynamics in multivariate processes, Lanne & Saikkonen (2013) proposed a multiplicative model for strictly stationary non-Gaussian time series: $\Pi(L)\Phi(L^{-1})y_t = \varepsilon_t$, where $\Pi(L) = Id - \Pi_1 L - ... - \Pi_r L^r$, and, $\Phi(L^{-1}) = Id - \Phi_1 ... - \Phi_s L^{-s}$ are $n \times 1$ autoregressive causal and noncausal polynomials such that $det\Phi(z) \neq 0$ for $|z| \leq 1$ and $det\Pi(z) \neq 0$ for $|z| \leq 1$, and ε_t is a $n \times 1$ sequence of independent and identically distributed (i.i.d.) non-Gaussian random vectors with zero mean and finite positive definite variance-covariance matrix. A limitation of this approach is that $\Pi(L)$ and $\Pi(L)$ do not necessarily commute [Cubbada et al. (2023)]. The multiplicative model representation with autoregressive orders r and s may not

[1] In equation (1) we assume that y_t has zero mean.

always exist for a mixed VAR(p) process and it may not be unique [Gouriéroux & Jasiak (2017), Gouriéroux & Jasiak (2023), Davis & Song (2020), Swensen (2022)].

Gouriéroux & Jasiak (2017) and Davis & Song (2020) consider the classical representation (1) under modified assumptions. More specifically Gouriéroux & Jasiak (2017) assume that the errors e_t is a sequence of non-Gaussian i.i.d. vectors with positive definite variance-covariance matrix Σ, and the roots of the autoregressive polynomial lie either outside or inside the unit circle. Both articles discuss the identification and estimation of the causal-noncausal VAR(p) models. Davis & Song (2020) rely on the ML estimation which requires a distributional assumption on the error terms entailing the risk of misspecification. Gouriéroux & Jasiak (2017) introduce a semi-parametric estimator called the Generalized Covariance Estimator (GCov hereafter) for mixed causal noncausal multivariate non-Gaussian processes. The GCov estimator does not require an assumption of a specific parametric error distribution and uses the nonlinear autocovariances for identification of causal and noncausal components. Gouriéroux & Jasiak (2017) show that the GCov estimator is consistent and asymptotically normally distributed.

The next Section presents the causal-noncausal VAR model (referred to as the mixed VAR), recalls its representation in terms of purely causal and noncausal components and summarizes the results on the GCov estimator.

2.1 The Mixed VAR(1) Model

Let us consider a strictly stationary n-dimensional mixed VAR(1) process:

$$Y_t = \Phi Y_{t-1} + \varepsilon_t, \tag{2.1}$$

where Φ is an $n \times n$ matrix and (ε_t) is a i.i.d. multivariate non-Gaussian sequence of dimension n. We assume that (ε_t) is square integrable[2] with zero mean $E(\varepsilon_t) = 0$, and variance-covariance matrix $V(\varepsilon_t) = \Sigma$. Since (ε_t) is not assumed independent of the lagged values of the process $Y_{t-1}, Y_{t-2}...$, it cannot be interpreted as an innovation. The eigenvalues of matrix Φ are assumed to be of modulus different from 1 to ensure the existence of a unique, strictly stationary solution to equation (2.1) with a two-sided strong (i.i.d.) moving average representation.

2.2 State-Space Representation

This Section reviews the state-space representation of mixed VAR model based on the representation theorem of Gouriéroux & Jasiak (2017) for mixed processes that distinguishes their purely causal and noncausal latent components. Let us consider the VAR(1) model for ease of exposition. In the mixed VAR(1) model, if n_1 (resp. $n_2 = n - n_1$) represents the number of eigenvalues of Φ of modulus strictly less than 1 (resp. strictly larger than 1), then there exists an invertible $n \times n$ matrix A, and two square matrices: J_1 of dimension $n_1 \times n_1$ and J_2 of dimension $n_2 \times n_2$. The eigenvalues of J_1 (resp. J_2) with moduli strictly less than 1 (resp. larger than 1) are such that :

$$Y_t = A_1 Y_{1,t}^* + A_2 Y_{2,t}^*, \tag{2.2}$$

$$Y_{1,t}^* = J_1 Y_{1,t-1}^* + \varepsilon_{1,t}^*, \quad Y_{2,t}^* = J_2 Y_{2,t-1}^* + \varepsilon_{2,t}^*, \tag{2.3}$$

$$\varepsilon_{1,t}^* = A^1 \varepsilon_t, \quad \varepsilon_{2,t}^* = A^2 \varepsilon_t, \tag{2.4}$$

where A_1, A_2 are the blocks in the decomposition of matrix A as : $A = (A_1, A_2)$, and

[2] The assumption of square integrability can be satisfied by a nonlinear function of ε in the GCov estimator.

A^1, A^2 are the blocks in the decomposition of A^{-1} as $A^{-1} = \begin{pmatrix} A^1 \\ A^2 \end{pmatrix}$. The matrices J_1 and J_2 are derived from the real Jordan canonical form of Φ:

$$\Phi = A \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} A^{-1},$$

where the columns of matrix A correspond to an appropriate basis.

In a VAR(1) model, matrix Φ can be diagonalizable so that J_1 and J_2 are diagonal matrices containing distinct eigenvalues. Then, matrices A^1 and A^2 contain the eigenvectors. In general, however, this may not be the case and the eigenvalues may not be distinct, and the dimension of the eigenspace may be strictly smaller than the multiplicity order of the eigenvalue, especially when the number of component series n is large. Then, matrix Φ can be written in the (real) Jordan canonical form as $\Phi = AJA^{-1}$, where matrix J is block-diagonal and divided into sub-matrices J_1 and J_2 containing non-diagonal Jordan blocks associated to eigenvalues of modulus less and greater than 1, respectively.

By premultiplying both sides of equation (2.2) and (2.4) by matrix A^{-1} we can decompose Y_t into its latent causal and noncausal components as follows :

$$Y_t^* = \begin{pmatrix} Y_{1,t}^* \\ Y_{2,t}^* \end{pmatrix} \equiv A^{-1} Y_t, \quad \varepsilon_t^* = \begin{pmatrix} \varepsilon_{1,t}^* \\ \varepsilon_{2,t}^* \end{pmatrix} \equiv A^{-1} \varepsilon_t.$$

We get :

$$Y_t^* = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} Y_{t-1}^* + \varepsilon_t^*, \quad \text{and} \quad Y_{j,t}^* = J_j Y_{j,t-1}^* + \varepsilon_{j,t}^*, j = 1, 2,$$

In addition, equation $Y_t = AY_t^*$, is equivalent to $Y_t = A_1 Y_{1,t}^* + A_2 Y_{2,t}^*$, which is the decomposition (2.2).

Given that the eigenvalues of J_1 are of modulus strictly less than 1, the set of equations below is causal:

$$Y_{1,t}^* = J_1 Y_{1,t-1}^* + \varepsilon_{1,t}^*. \tag{2.5}$$

It can be used to derive the causal one-sided moving average representation of $Y_{1,t}^*$ where L denotes the lag operator: $Y_{1,t}^* = \Sigma_{h=0}^{\infty} J_1^h \varepsilon_{1,t-h}^* = (Id - J_1 L)^{-1} \varepsilon_{1,t}^*$ with $(Id - J_1 L)^{-1} \equiv \Sigma_{h=0}^{\infty} J_1^h L^h$.

The second set of equations is noncausal :

$$Y_{2,t}^* = J_2 Y_{2,t-1}^* + \varepsilon_{2,t}^*. \tag{2.6}$$

It can be written as $Y_{2,t}^* = J_2^{-1} Y_{2,t+1}^* - J_2^{-1} \varepsilon_{2,t+1}^* = (Id - J_2 L)^{-1} \varepsilon_{2,t}^*$, where $(Id - J_2 L)^{-1} \equiv -\Sigma_{h=1}^{\infty} J_2^{-h} L^{-h}$.

Processes $(Y_{1,t}^*)$ and $(Y_{2,t}^*)$ are purely causal and noncausal, respectively. They can be interpreted as the causal and noncausal latent components of process (Y_t) in a state-space representation (2.2)-(2.5)-(2.6). Moreover, these components are deterministic functions of (Y_t) since : $Y_{j,t}^* = A^j Y_t, j = 1, 2$.

The component $Y_{2,t}^*$ is the locally explosive component following a strictly stationary noncausal (V)AR process. It represents the common bubbles, local trends and spikes. It can be univariate or multivariate, depending on the dimension of the process $Y_{2,t}^*$.

The causal component $Y_{1,t}^*$ is the stationary linear combination $Y_{1,t}^* = A^1 Y_t$ of the observed multivariate process that eliminates its local explosive features, and $cov(Y_{1,t}^*, Y_{2,t}^*) = 0$.

In this respect, the above approach can be compared with the "common features" representation of Engle & Kozicki (1993), as it satisfies the following definition: "A feature that is present in each of a group of series is said to be common to those series if there exists a nonzero linear combination of the series that does not have the feature". This definition concerns the stationary time series, which is consistent with the assumption of strict stationarity of y_t. A common feature in Y_t is the locally explosive noncausal component along with its explosion rate determined by J_2.

The linear combination $Y_{1,t}^* = A^1 Y_t$ can also be compared to the cointegrating relation of Engle & Granger (1987), and interpreted as bubble "cointegration". An important difference is that Y_t is not a unit root process, but it is nevertheless (locally) explosive. The explosive patterns of Y_t are not global, i.e. do not involve unbounded growth (decline) or oscillations. The strictly stationary process Y_t is characterized by local, i.e. short-lasting explosions that end more or less suddenly.

If we were to impose an assumption analogous to the condition of equal integration order I(1), it would be the assumption of equal, multiple eigenvalues of modulus greater than 1 and a diagonalizable matrix Φ. Then, all local explosive patterns would have the same rate of explosion and would be "common" in the sense analogous to the equal explosion rates of global trends in the cointegrated, nonstationary processes, associated with and determined by the eigenvalues equal to 1. Technically, this would require imposing a constraint on the Jordan canonical form of matrix Φ with a fixed multiplicity n_2 of roots of equal values.

Alternatively, we could pre-test the series individually for noncausal roots by fitting a univariate noncausal model to each of the processes. Then, we could select the processes with one noncausal root and equal noncausal autoregressive coefficients, to ensure an equal explosion rate. As a consequence, the joint VAR(1) model of these series would have one noncausal component. Intuitively, we expect that if each component series has at most n_2 distinct noncausal, i.e. local explosive features, the joint process will have at most n_2 noncausal features.

An alternative approach to bubble cointegration is proposed in Cubadda, Hecq, & Voisin (2023) Their method is based on the multiplicative representation of the VAR(1) model of Lanne & Saikkonen (2013), written as a product of a lead and lag vector autoregressive polynomials and discussed in the introduction to Section 2. The linear combination introduced in Cubadda, Hecq, & Voisin (2023) eliminates the coefficients on the future values of Y_t in the factorized vector autoregressive representation.

2.3 Bivariate VAR(1) - Example

To better understand the comovements of Y_t components and their contribution to the causal and non-causal components, let us consider a bivariate mixed VAR(1) process with a diagonalizable matrix Φ. When matrix Φ is triangular, then depending on the eigenvalues, one component of y_t may not contribute to either the explosive (i.e. noncausal) or regular (causal) dynamics.

Suppose that matrix Φ has the following spectral decomposition:

$$\Phi = AJA^{-1}$$

where J is the 2 by 2 matrix of real eigenvalues, A is the 2 by 2 matrix with columns, which are the eigenvectors of Φ. Suppose also that either $\phi_{12} = 0$ or $\phi_{21} = 0$, so that matrix Φ is either upper or lower triangular. It is known that for any $n \times n$ triangular matrix the following properties hold:

1) The eigenvalues of an upper or lower triangular matrix are the diagonal elements of the matrix.

2) For any triangular matrix, a vector with all elements equal to zero, except the first one is an eigenvector. There is a second eigenvector with all elements zero, except the first two, etc.

Therefore, a triangular 2 by 2 matrix Φ has a triangular matrix A, with a triangular inverse A^{-1}. It follows that the past values of one component of y_t do not contribute to either the explosive dynamics y^*_{2t}, or the regular dynamics $y^*_{1,t}$.

Let the matrix J be written as $J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix}$, where $J_1 < 1 < J_2$. Then matrix A has entries $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and its inverse is $A^{-1} = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & a^{22} \end{pmatrix}$.

Accordingly, we have row vectors $A^1 = [a^{11}\ a^{12}]$ and $A^2 = [a^{21}\ a^{22}]$ corresponding to the components y^*_{1t} and y^*_{2t} with regular and explosive dynamics, respectively.

Example 1: Upper triangular Φ

Suppose the element $\phi_{21} = 0$ in matrix

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

which makes it an upper triangular matrix

$$\Phi_U = \begin{pmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{pmatrix}$$

If $J_1 = \phi_{11},\ J_2 = \phi_{22}$ so that $J_2 > J_1$, we get $a^{21} = 0$

$$A^{-1} = \begin{pmatrix} a^{11} & a^{12} \\ 0 & a^{22} \end{pmatrix}$$

Then, both y_{1t} and y_{2t} contribute to the regular component y^*_{1t}, but process y_{1t} does not contribute to the explosive component $y^*_{2t} = y_{2t}$:

$$y^*_{1,t} = a^{11} y_1 + a^{12} y_2 = \sum_{j=0}^{+\infty} \lambda_1 \varepsilon^*_{1,t-j}, \tag{2.7}$$

with $\varepsilon^*_{1,t}$ as the causal error $\varepsilon^*_{1,t} = a^{11}\varepsilon_{1,t} + a^{12}\varepsilon_{2,t}$, and

$$y^*_{2,t} = a^{22} y_{2,t} = -\sum_{j=0}^{+\infty} [\lambda_2^{-j-1}\, a^{22} \varepsilon_{2,t+j+1}].$$

The noncausal error $\varepsilon^*_{2,t} = a^{22}\varepsilon_{2,t}$ is a function of ε_2 only. We observe that $\underline{y_{1,T}}$ affects only the error term associated with $y_{1,T+1}$, i.e. the non-explosive error.

If $J_1 = \phi_{22},\ J_2 = \phi_{11}$ so that $J_2 > J_1$, we get $a^{11} = 0$

$$A^{-1} = \begin{pmatrix} 0 & a^{12} \\ a^{21} & a^{22} \end{pmatrix}$$

In this case, process y_{1t} is explosive and does not contribute to the regular component $y_{1t}^{*} = y_{2t}$, while both y_{1t} and y_{2t} contribute to the explosive component y_{2t}^{*}.

Example 2: Lower triangular Φ

Suppose the element $\phi_{12} = 0$ in matrix

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

which makes a lower triangular matrix

$$\Phi_L = \begin{pmatrix} \phi_{11} & 0 \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

Then, if $J_1 = \phi_{11}$, $J_2 = \phi_{22}$ so that $J_2 > J_1$, we get $a^{12} = 0$

$$A^{-1} = \begin{pmatrix} a^{11} & 0 \\ a^{21} & a^{22} \end{pmatrix}$$

Process y_{2t} does not contribute to the regular component $y_{1t}^{*} = y_{1t}$, but both processes contribute to the explosive component y_{2t}^{*}.

$$y^*_{1,t} = a^{11} y_1 = \sum_{j=0}^{+\infty} \lambda_1 \varepsilon^*_{1,t-j},$$

where $\epsilon^*_{1,t}$ is the causal error $\epsilon^*_{1,t} = a^{11}\epsilon_{1,t}$ and a function of $\epsilon_{1,t}$ only. The explosive component is

$$y^*_{2,t} = a^{21} y_{1,t} + a^{22} y_{2,t} = -\sum_{j=0}^{+\infty} [\lambda_2^{-j-1} \left(a^{21}\varepsilon_{1,t+j+1} + a^{22}\varepsilon_{2,t+j+1}\right)].$$

If $J_1 = \phi_{22}$, $J_2 = \phi_{11}$ so that $J_2 > J_1$, we get $a^{22} = 0$ and

$$A^{-1} = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & 0 \end{pmatrix}$$

In this case process y_{2t} does not contribute to the explosive component $y^*_{2t} = y_{1t}$, while both y_{1t} and y_{2t} contribute to the regular component y^*_{1t}.

Independence

The independence between y_1 and y_2 arises when $\phi_{12} = \phi_{21} = 0$ and the joint density of errors can be written as: $g(\varepsilon_{1,t}, \varepsilon_{2,t}) = g_1(\varepsilon_{1,t}) g_2(\varepsilon_{2,t}), \forall t$.

2.4 VAR(1) representation of the VAR(p) model

The mixed VAR(p) can be easily transformed into a mixed VAR(1) model for estimation and inference purposes. In that context, the causal and noncausal components can be easily determined too. Consider the mixed VAR(p) process:

$$Y_t = \Phi_1 Y_{t-1} + ... + \Phi_p Y_{t-p} + \varepsilon_t \, , \tag{2.8}$$

where (ε_t) is a sequence of i.i.d. random vectors of dimension n with variance-covariance matrix Σ. We can write this model as a VAR(1) model for X_t where X_t is obtained by stacking the current and lagged values of Y_t into $X_t = (Y_t', Y_{t-1}', ..., Y_{t-p+1}')'$, to get

$$X_t = \Psi X_{t-1} + u_t. \tag{2.9}$$

The autoregressive matrix Ψ can be written as the augmented matrix:

$$\Psi = \begin{bmatrix} \Phi_1 & ... & ... & \Phi_p \\ Id & 0 & ... & 0 \\ 0 & Id & ... & 0 \\ 0 & ... & Id & 0 \end{bmatrix}. \tag{2.10}$$

and the errors in (2.9) are:

$$u_t' = \begin{bmatrix} \epsilon_{1,t} & \cdots & \epsilon_{n,t} & 0 & \cdots & 0 \end{bmatrix}.$$

By the representation theorem given in Section 2.2, matrix Ψ can also be written in the Jordan canonical form:

$$\Psi = B \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} B^{-1}.$$

Similarly to Y_t in the VAR(1) case, X_t is the sum of causal and noncausal components

$$X_t = B_1 X^*_{1,t} + B_2 X^*_{2,t},$$

where

$$X^*_{1,t} = J_1 X^*_{1,t-1} + u^*_{1,t},$$

$$X^*_{2,t} = J_2 X^*_{2,t-1} + u^*_{2,t},$$

and the causal and noncausal errors are deterministic functions of the process u_t,

$$u^*_{1,t} = B^1 u_t, \quad u^*_{2,t} = B^2 u_t.$$

Errors $u^*_{1,t}$ and $u^*_{2,t}$ satisfy $n(p-1)$ linearly independent and deterministic relationships since they both depend on ε_t, dim $u^*_{1,t}$ + dim $u^*_{2,t} = n_1 + n_2 = np$ and np is greater than dim $\varepsilon_t = n$ whenever $p > 1$.

Moreover, it follows that:

$$X^*_{1,t} = B^1 X_t, \text{ and } X^*_{2,t} = B^2 X_t. \tag{2.11}$$

The expression $X^*_{1,t} = B^1 X_t$, is the linear combination that eliminates the local explosive patterns. Because X_t contains the present and past values of Y_t, the common bubble in a VAR(p) is not necessarily "common and contemporaneous".

As one can see in Figure 1, for example, the bubble dynamics are quite complex, and the bubbles are not always necessarily contemporaneous. The bubbles in the component processes do not start and end at the same time, which justifies the presence of lagged values in the non-explosive combination.

Suppose that the process of interest is VAR(2) for p=2, where $n_1 + n_2 = 2n$. Then we have n_2 common bubbles $X^*_{2,t} = B^2 X_t = B^2_1 Y_t + B^2_2 Y_{t-1}$, which are functions of the current and lagged Y_t, up to a pre-multiplication by an invertible matrix of dimension $n_2 \times n_2$. Among these combinations, we can distinguish n_{21} combinations that depend on Y_t only and depict the "common and contemporaneous" bubbles. Then, there exists $\gamma \in N(B^2_2) = \{\gamma' B^2_2 = 0\}$ such that $\gamma' X^*_{2,t} = \gamma' B^2_1 Y_t$, where $N(.)$ denotes the null space. In addition, there can be n_{22} combinations involving Y_{t-1} only. Then $\gamma \in N(B^2_1)$ so that $\gamma' X^*_{2,t} = \gamma' B^2_2 Y_{t-1}$. There are also $n_2 - n_{21} - n_{22}$ combinations $\gamma' X^*_{2,t}$ orthogonal to the previous ones that are mixed, and involve both the past and present values of the process. Then, $\gamma \in (N(B^2_1) + N(B^2_2))^\perp = N(B^2_1)^\perp \cap N(B^2_2)^\perp$.

The "contemporaneous" aspect of common bubbles can be easily tested by estimating the matrix B^2 and testing the null hypothesis of zero elements associated to the lags of Y_t.

There also exist n_1 locally non-explosive combinations $X^*_{1,t} = B^1 X_t = B^1_1 Y_t + B^1_2 Y_{t-1}$ which are functions of the current and lagged Y_t, up to a pre-multiplication by an invertible matrix of dimension $n_1 \times n_1$. By analogy to the cointegration, these combinations eliminate the local explosive patterns, instead of common global trends (nonstationarity). Among them, we can distinguish n_{11} combinations that depend on Y_t only and eliminate the "common and contemporaneous" bubbles. Then, there exists $\gamma \in N(B^1_2) = \{\gamma' B^1_2 = 0\}$ such that $\gamma' X^*_{1,t} = \gamma' B^1_1 Y_t$. In addition, there can be n_{12} non-explosive combinations involving Y_{t-1} only, such that $\gamma \in N(B^1_1) = \{\gamma' B^1_1 = 0\}$ yielding $\gamma' X^*_{1,t} = \gamma' B^1_2 Y_{t-1}$. Finally, there are $n_1 - n_{11} - n_{12}$ non-explosive combinations $\gamma' X^*_{1,t}$ orthogonal to the previous ones that are mixed, and involve both the past and present values of the process with $\gamma \in (N(B^1_1) + N(B^1_2))^\perp = N(B^1_1)^\perp \cap N(B^1_2)^\perp$. Note that in the context of cointegration, the lags of Y_t appear in the cointegrating equation that eliminates the common trends, unless the VAR(1) dynamics and equal order of integration on all nonstationary component series are assumed [e.g. Engle & Granger (1987)].

In this context, the alternative approach to bubble cointegration proposed in Cubadda, Hecq, & Voisin (2023) is also applicable. It is restricted to the "common and contemporaneous" bubbles only and does not involve the lags of Y_t. However, that method is based on the multiplicative representation of the VAR(p). The factorization of the vector autoregressive polynomial of Y_t into the causal and noncausal polynomials may not always be feasible, or unique Davis & Song (2020), Gouriéroux & Jasiak (2023).

2.5 Semi-Parametric Estimation

The semi-parametric estimation method for mixed causal noncausal processes, called the Generalized Covariance Estimator was introduced by Gouriéroux & Jasiak (2017). It follows from the nonlinear identification result in Ming-Chung & Kung-Sik (2007) that there exist nonlinear covariance based conditions that can be used to identify causal and noncausal components of a given series provided the error terms ε_t are serially independent. The nonlinear covariance based conditions for VAR(1) model (2.1), for example, could be the covariances between nonlinear transforms of the error terms defined for a given set of functions as:

$$c_{j,k}(h, \Phi) = Cov[a_j(Y_t - \Phi Y_{t-1}), a_k(Y_t - \Phi Y_{t-h-1})], \ j, k = 1, ..., K, \ h = 1, .., H,$$

for a given set of functions $a_k, k = 1, ...K$, satisfying the regularity conditions given in Gouriéroux & Jasiak (2017) and Gouriéroux & Jasiak (2023).

Let us denote by $\Theta_l(\underline{Y}_t, \phi)$, $l = 1, ..., KH$, the function $a_k(Y_{t-h} - \Phi Y_{t-h-1})$,

$k = 1, ..., K$, $h = 1, ..., H$ where $\phi = vec\Phi$. For each covariance $c_{kl} = Cov[\Theta_k(\underline{Y}_t, \phi)$ $, \Theta_l(\underline{Y}_t, \phi)]$, $k, l = 1, ..., KH$, we can write its sample counterpart: $\hat{\gamma}_{k,l,T} = \widehat{Cov}[\Theta_k(\underline{Y}_t, \phi),$ $\Theta_l(\underline{Y}_t, \phi)]$, $k, l = 1, ..., KH$, where $\underline{Y}_t$ contains the values of the process up to and including time t.

The Covariance estimator $\tilde{\phi}_T$ of $\phi = vec\Phi$ minimizes the following objective function:

$$\tilde{\phi}_T = \hat{\gamma}_T\prime(\phi)\Omega\hat{\gamma}_T(\phi),$$

with respect to ϕ where $\hat{\gamma_T}(\phi)$ denotes the vector obtained by stacking $\hat{\gamma}_{k,l,T}(\phi)$ and Ω is a $(KH \times KH)$ positive definite weighting matrix.

Under the assumption that the model is well-specified, and the identification condition given in Gouriéroux & Jasiak (2017), Gouriéroux & Jasiak (2023) there exists a unique solution to the limiting objective function. Under mild regularity conditions, the GCov estimator exists and is consistent and asymptotically normally distributed Gouriéroux & Jasiak (2017), Gouriéroux & Jasiak (2023). The asymptotic efficiency of a Covariance estimator based on a given set of nonlinear autocovariances depends on the matrix of weights Ω. The estimator is asymptotically semi-parametrically efficient when the optimal weighting matrix is used. The optimal weights Ω that ensure asymptotic semi-parametric efficiency are based on the inverse sample variance matrices of $a(Y_t)$ [see, Gouriéroux & Jasiak (2023)].

For our implementation, we consider the sample autocorrelations $\hat{\rho}_{j,k}(h, \Phi) = Corr[a_j(Y_t - \Phi Y_{t-1}), a_k(Y_{t-h} - \Phi Y_{t-h-1})]$. Then the GCov estimator can be represented as a weighted covariance estimator that minimizes a portmanteau-type criterion [see e.g. Cubadda & A. Hecq (2011)]:

$$\hat{\phi}_T = \underset{\phi}{\text{argmin}} \sum_{j=1}^{K} \sum_{k=1}^{K} [\sum_{h=1}^{H} \hat{\rho}^2_{j,k,T}, (h, \phi)] \tag{2.12}$$

where H is the highest selected lag and the theoretical autocorrelations $\rho_{j,k}$ are replaced by their sample counterparts $\hat{\rho}_{j,k,T}$, see Gouriéroux & Jasiak (2017). This estimator is easy to implement, but is not optimally weighted.

The definition of the Generalized Covariance estimator is similar to the definition of a Generalized Method of Moments (GMM) estimator since by analogy, we can obtain a consistent covariance estimator with a simple weighting scheme such as an identity matrix (although that first step estimator may not be fully semi-parametrically efficient). The differences lie in 1) the use of the central moments only in the GCov approach, 2) the reduced dimension of the objective function to be minimized, and 3) the portmanteau statistic interpretation of the objective function [see e.g. Cubadda & A. Hecq (2011)].

The choice of nonlinear covariances is a problem similar to choosing the instruments in a GMM setting. One can choose a combination of quadratic and linear transformations to capture the absence of leverage effect at lag $h, h \leq 0$ for example, or other nonlinear functions, such as higher powers or logarithms. The choice of tuning parameters H and K has a limited impact on the variance of the estimator [3].

The GCov estimator has no closed-form and is obtained numerically from a minimization procedure. In general, the objective function of the GCov is not convex and an algorithm, such as the commonly used BFGS or BHHH can converge to a local minimum. In practice, the choice of initial conditions is important. For example, using the OLS estimator as the initial condition may cause the algorithm to converge to a local minimum associated to the inverse roots. The practical implementation problems and their solutions are discussed in Cubadda, Giancaterini, et al. (2023) who propose a Simulated Annealing (SA) procedure to eliminate potential computational difficulties.

Gouriéroux & Jasiak (2017) show that GCov estimator $\hat{\phi}$ of $\phi = vec(\Phi')$ is asymptotically normal with the asymptotic variance given by: $V_{asy}[\sqrt{T}(\hat{\phi}_T -$

[3] See Gouriéroux & Jasiak (2017) and Gouriéroux & Jasiak (2023) for the discussion of choices of H and K.

$\phi)]$=$(D'\Sigma^{-1}D)^{-1}$. The rows of matrix D are: $D_{k,l} = -\frac{\partial}{\partial\phi'}\widehat{Cov}[\Theta_k(Y_t,\phi),\Theta_l(Y_t,\phi)]$. The elements of matrix Σ are:

$$\sigma_{(k,l),(k',l')} = Cov_{asy}(\sqrt{T}\widehat{Cov}[\Theta_k,\Theta_l],\sqrt{T}\widehat{Cov}[\Theta_{k'},\Theta_{l'}])$$

where $\Theta_i = \Theta_i(Y_t,\phi) = a_k(Y_{t-h} - \Phi Y_{t-h-1})$ for $i = (k,l,k',l')$.

In practice, the estimation of a (bivariate) VAR(p) from the GCov estimator can be accomplished along the following steps:

1. Estimate $\Phi_1,...,\Phi_p$ for a given autoregressive order p using the GCov estimator. This can be done using linear and nonlinear functions of $\epsilon_t(\phi) = Y_t - \Phi_1 Y_{t-1} - ... - \Phi_p Y_{t-p}$.
2. Compute the residuals $\widehat{\epsilon_t} = Y_t - \widehat{\Phi_1}Y_{t-1} - ... - \widehat{\Phi_p}Y_{t-p}$ and their nonlinear autocorrelation functions. If the residual autocorrelations are significant at some lags, then re-estimate the model by increasing the autoregressive order p and repeat until the residual autocorrelations are no longer significant.
3. Using the p estimated autoregressive coefficients $\widehat{\Phi_1},...,\widehat{\Phi_p}$ compute $\widehat{\Psi}$ and derive the Jordan canonical form of $\widehat{\Psi}$ The decomposition will yield $\widehat{n_1}$ and $\widehat{J_i}$ and $\widehat{B_i}$ for i = {1,2}.

3 Empirical Analysis of Cryptocurrencies

3.1 Cryptocurrencies

We consider the US Dollar prices of the following cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP) and Stellar (XLM) over the period 2017-2019, obtained from the Bitfinex exchange (www.bitfinex.com). The presence of common bubbles in cryptocurrency markets during that period is attributable to several factors, such as herding behaviour among traders, regulatory changes and news shocks.

Devenow & Welch (1996) defines herding as correlated patterns of behaviour among traders. Herding behaviour occurs in financial markets when investors imitate the investment decisions of others without reference to fundamentals [see Hwang & Salmon (2004)]. Then, traders mimic the trading decisions of others instead of relying on their own private information Yuan Zhao (2022). Since herding leads to the adoption of similar investment strategies among traders, it can help explain local trends and common bubbles in asset prices [see Corbet et al. (2019)]. Bouri et al. (2019) investigate herding in cryptocurrency markets and find evidence of herding over the period 2016 - 2018. Building on the work of Bouri et al. (2019), da Gama Silva et al. (2019) examines 50 high liquidity cryptocurrencies between March 2015 and November 2018 and confirms the evidence of herding behaviour, which is particularly strong in the down-market period of 2018. Esra Alp Coskun (2020) study 14 leading cryptocurrencies between 2013 and 2018 and find that cryptocurrencies with smaller market capitalization generally herd with cryptocurrencies that have a larger market capitalization.

The regulatory environment for cryptocurrencies also plays a role. The global cryptocurrency regulation is not homogeneous and is subject to abrupt changes in specific jurisdictions which can impact the entire cryptocurrency market. For example, in February 2018 the People's Bank of China blocked access to all domestic and foreign cryptocurrency exchanges Perper (2018).

News events also contribute to the comovements of cryptocurrencies. For example, Djogbenou et al. (2023) show that the appointment of Peter Warrack, a veteran anti-money laundering specialist at Royal Bank of Canada, to the position of CEO of Bitfinex on May 7th 2018 had an impact on the exchange rates of the stable coin Tether. [4] More specifically, Tether hovered above its one dollar peg until

[4] A stable coin is a cryptocurrency that is pegged to another currency, commodity, or financial instrument Hayes (2023). Tether aims to maintain a 1:1 peg to the US dollar.

between May and September of 2018. During this same period there was a common spike in the prices of Bitcoin, Ethereum, Ripple and Stellar in the US dollar which lasted approximately until September 2018 as well.

3.2 Bitcoin (BTC) and Ethereum (ETH)

The sample of US Dollar (USD) prices of Bitcoin and Ethereum (BTC and ETH hereafter) consists of $T = 885$ daily observations collected between January 01, 2017 and June 04, 2019.

Figure 1a displays the daily BTC/USD and ETH/USD rates over the entire sampling period of 885 days. Both Bitcoin and Ethereum experienced a large increase in value relative to the US dollar since early 2017. In 2018 they lost a large proportion of that increase, compared to the peak in late 2017. In addition, both rates show evidence of bubbles, i.e. local explosive trends with periods of explosive increases and sudden declines. Figure 1b displays BTC/USD and ETH/USD rates with medians subtracted. The BTC/USD rate is divided by a factor of ten for comparison and further modelling. These rates are hereafter referred to as the adjusted series. In Figure 1c the grey region indicates a sub-sample of $T = 250$ observations over the period February 02, 2018 and October 10, 2018 selected for further analysis of the series. This sub-sample is shown in Figure 1d again to document the comovements between the series.

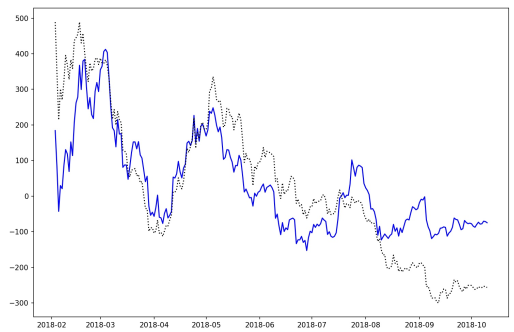
We chose this sub-sample with $T = 250$ because it displayed many spikes in the late 2017 while the large bubble was bursting and the cryptocurrency prices were decreasing. This sub-sample shown in Figures 1b and 1d is detrended using Python package Scipy [5] using a spline of order 2 with a knot every 30 observations. Alternatively, the Hodrick-Prescott (HP) filter could be used [see Paige & Trindade (2010)]. Hecq & Voisin (2023) find that the Hodrick-Prescott filter does not introduce significant distortions to the mixed causal-noncausal dynamics when applied to their

[5] More specifically we use LSQUnivariateSpline from the Scipy.interpolate package

(a) 2017-01-01 to 2019-06-04

(b) Adjusted 2017-01-01 to 2019-06-04

(c) Sub-sample 2018-02-03 to 2018-10-10 in grey



(d) Sub-sample 2018-02-03 to 2018-10-10

Figure 1: BTC/USD and ETH/USD Rates. BTC/USD solid line, ETH/USD dotted line.

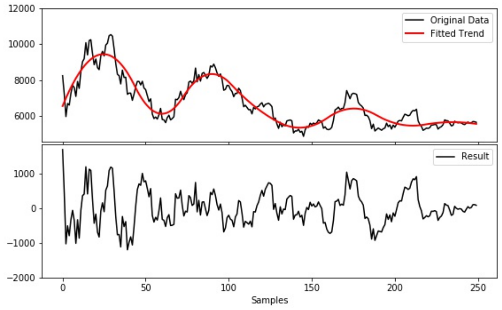
(monthly) data on oil. However, the HP filter requires the selection of a value for the parameter lambda that is conventionally an increasing function of the sampling frequency of the data. In our analysis of daily closing rates the value of lambda is very large leading to computational errors, which motivates the use of a cubic spline to detrend our data.

The original and detrended BTC and ETH rates are shown in Figures 2. Figure 2c displays the detrended, adjusted BTC/USD rate as a solid line and the detrended ETH/USD rate as a dotted line.

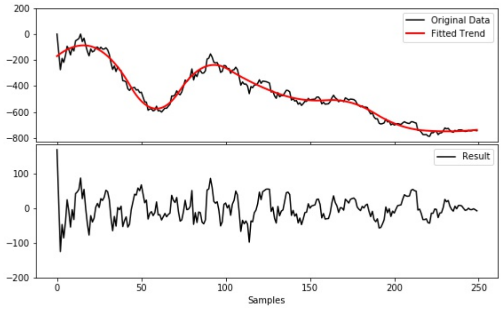
The autocorrelation function (ACF) of the detrended data in Figure B.1 in Supplementary Material shows a finite range of serial dependence. The shaded region in Figure B.1 marks the asymptotically valid confidence interval at 95%. Moreover, the detrended data is not normally distributed, with excess kurtosis and skewness of 0.25 and 0.31, respectively in BTC, and of 2.15 and 0.34 in ETH.

3.3 VAR(3) Model of BTC and ETH

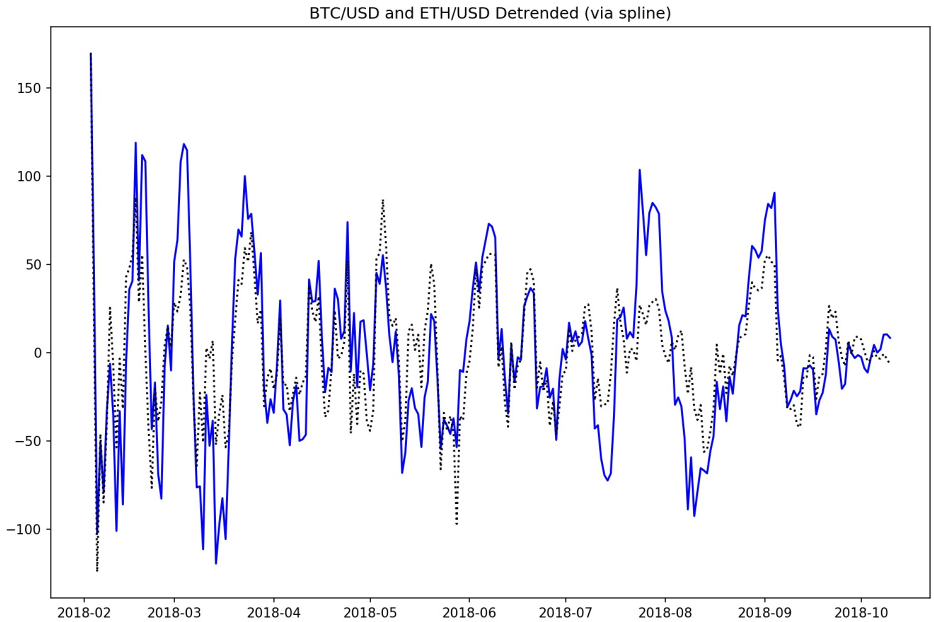
We estimate a mixed VAR(3) model for the BTC and ETH rates to improve upon the fit of the mixed VAR(1) model summarized in the Appendix. We increase the autoregressive order to eliminate the remaining serial correlation in the squared residuals. The VAR(3) model is estimated by setting H in the objective function (2.12) equal to 11 and minimizing it with respect to Φ. We obtain the following augmented matrix of estimated autoregressive coefficients:



(a) BTC (Adjusted) Detrended by Spline



(b) ETH (Adjusted) Detrended by Spline



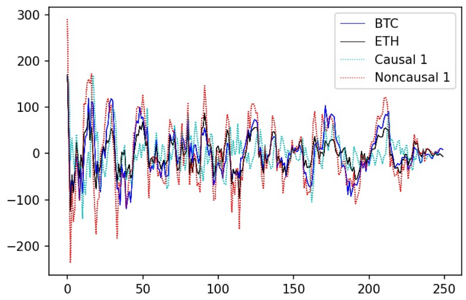
(c) BTC/USD: solid line, ETH/USD: dotted line

Figure 2: BTC and ETH (Adjusted) Detrended by Spline

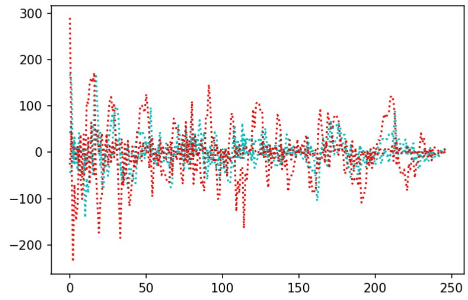
$$\hat{\Psi}_{GCov_{BTC/ETH}} = \begin{bmatrix} -0.792 & 2.059 & 1.717 & -1.439 & -0.497 & 0.242 \\ -1.268 & 2.06 & -1.268 & 2.06 & 0.087 & -0.099 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

This autoregressive augmented matrix has two eigenvalues outside the unit circle equal to 1.42 and -1.079. Inside the unit circle, there are two real valued eigenvalues 0.4 and -0.09 and a pair of complex conjugate eigenvalues 0.576+0.4i and 0.576-0.4i, both of modulus 0.7.

Figures B.7 and B.8 in Supplementary Material shows that there remains no statistically significant serial correlation in the residuals and in the squared residuals, and the VAR(3) model provides a good fit to the data. The histograms of VAR(3) residuals for BTC and ETH are given in Figure B.5 of Supplementary Material. The sample densities of both residual series display large tails indicating non-normality.



(a) BTC and ETH Series with Causal and Noncausal Components



(b) Causal and Noncausal Components

Figure 3: BTC and ETH Series with Causal and Noncausal Components, VAR(3)

There are two noncausal components in the VAR(3) model of BTC and ETH. Figure 3 presents the data and the highest variance causal and noncausal components. Panel 3a displays the original series along with the causal and noncausal components while Panel 3b contains only the causal and noncausal components. The causal components are graphed as solid blue lines while the noncausal components are graphed as dotted red lines.

The dynamics of the causal components represents the portfolios of cryptocurrency that are immune to local explosions and ensure stable investment paths.

The Noncausal 1 component is the most explosive stationary combination of the two processes while the Causal 1 is the highest variance non-explosive combination. The dynamics of the Noncausal 1 match the bubbles and spikes, such as those between observations 200 and 230.

We observe that the dynamics of noncausal component of the VAR(1) model of BTC and ETH rates displayed in Figure 11b in the Appendix is similar to the dynamics of Noncausal 1 component of VAR(3) shown in Panel 3b above. A linear regression of the noncausal component of the VAR(1) model of BTC/ETH on the two noncausal components of the VAR(3) model of BTC/ETH reveals a strong linear relationship with an R-squared of 0.92.

3.4 Comparison of Mixed VAR(3) and Causal VAR(3) for BTC and ETH

Let us compare the fit of the mixed VAR(3) and pure causal VAR(3) models for BTC and ETH rates. The augmented matrix of OLS estimated VAR(3) coefficients is shown below:

$$\hat{\Psi}_{OLS_{BTC/ETH}} = \begin{bmatrix} 0.867 & -0.166 & 0.108 & 0.0574 & -0.172 & 0.006 \\ -0.066 & 0.677 & 0.143 & 0.029 & -0.057 & -0.104 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

The eigenvalues for the augmented matrix are, 0.943, 0.456+0.43i, 0.456-0.43i (modulus 0.627), 0.463, -0.55, -0.23. The values of autoregressive coefficients of the causal and mixed VAR models are different, as well as their statistical significance.

In the equation of BTC in the OLS estimated causal VAR(3) model, the only statistically significant coefficient is on BTC at time $t-1$. In the ETH equation the only statistically significant coefficient is on ETH at time $t-1$. No other coefficients are statistically significant. The results show no evidence of a feedback effect or comovements.

The ACF in Figure B.7 in Supplementary Material, reveals the presence of serial dependence in the squared residuals of the causal VAR(3) model estimated by the OLS estimator from the same sample. The mixed causal noncausal VAR is able to capture nonlinear serial dependence in the data whereas a standard linear causal VAR model is unable to accommodate it. We observe that the autocorrelation of the squared residuals at lags one to three are statistically significant for the causal VAR(3) while they were not for the mixed VAR(3).

The OLS-estimated causal VAR model fails to capture the comovements and feedback effects accommodated by the mixed causal-noncausal model. This is because the causal VAR model assumes that all eigenvalues lie within the unit circle and is therefore misspecified.

3.5 XRP (Ripple) and XLM (Stellar)

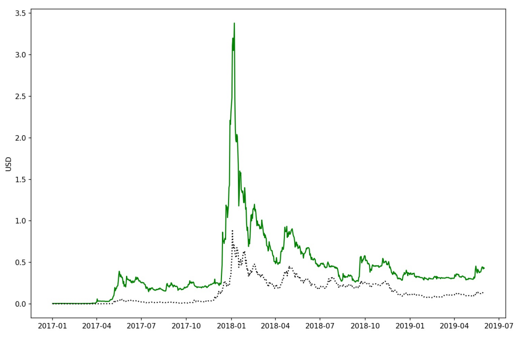
Figure 4a shows the full sample of 882 daily observations on US Dollar prices of Ripple and Stellar, referred to as XRP and XLM hereafter, between 2017 01 and 2019 06. Figure 4b displays the same two time series with their medians subtracted (referred to as the adjusted series henceforth). The grey region in Figure 4c indicates the sub-sample of $T = 250$ observations between 2018 03 25 and 2018 11 29 used for analysis and Figure 4d shows the adjusted sub-sample.

We observe that the series display explosive features with the presence of common bubbles and spikes. The summary statistics for the two series are given in Supplementary Material, Table A.1. and provide evidence that the series are not normally distributed.

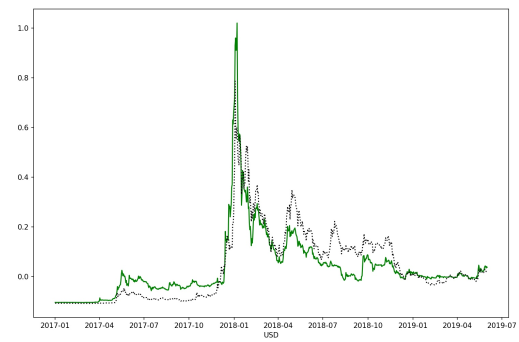
The XRP and XLM rates are detrended by using a spline of order three and with a knot at every 25 observations using Python package Scipy [6]. Figures 5a and 5b show the original and detrended series of XRP and XLM rates, respectively.

Figure 5c shows the adjusted and detrended sub-sample for XLM and XRP

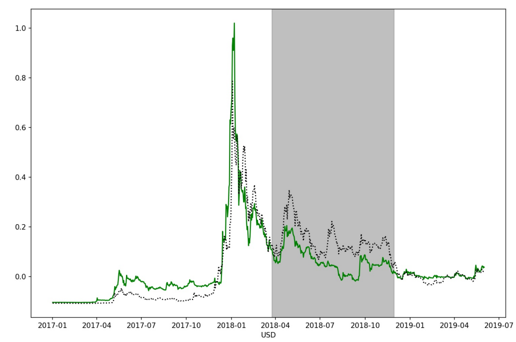
[6] More specifically we use LSQUnivariateSpline from the Scipy.interpolate package



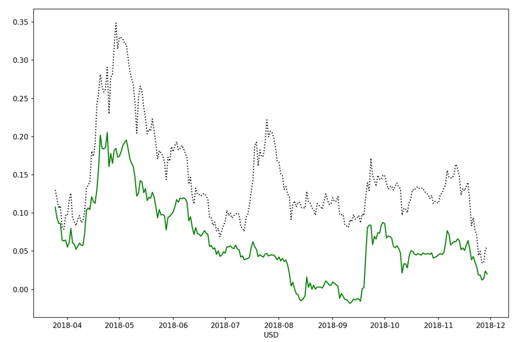
(a) 2017-01-01 to 2019-06-04



(b) Adjusted 2017-01-01 to 2019-06-04

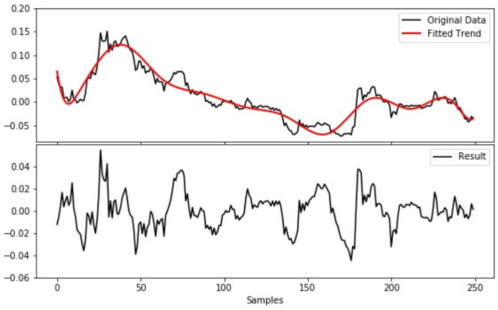


(c) Sub-sample 2018-03-25 to 2018-11-29 in grey

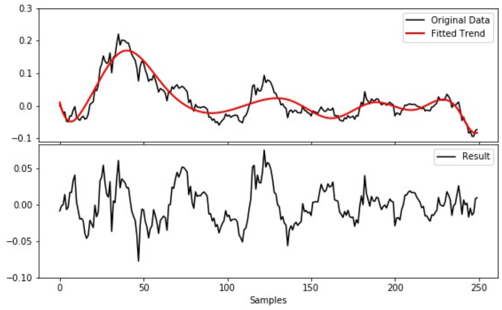


(d) Sub-sample 2018-03-25 to 2018-11-29

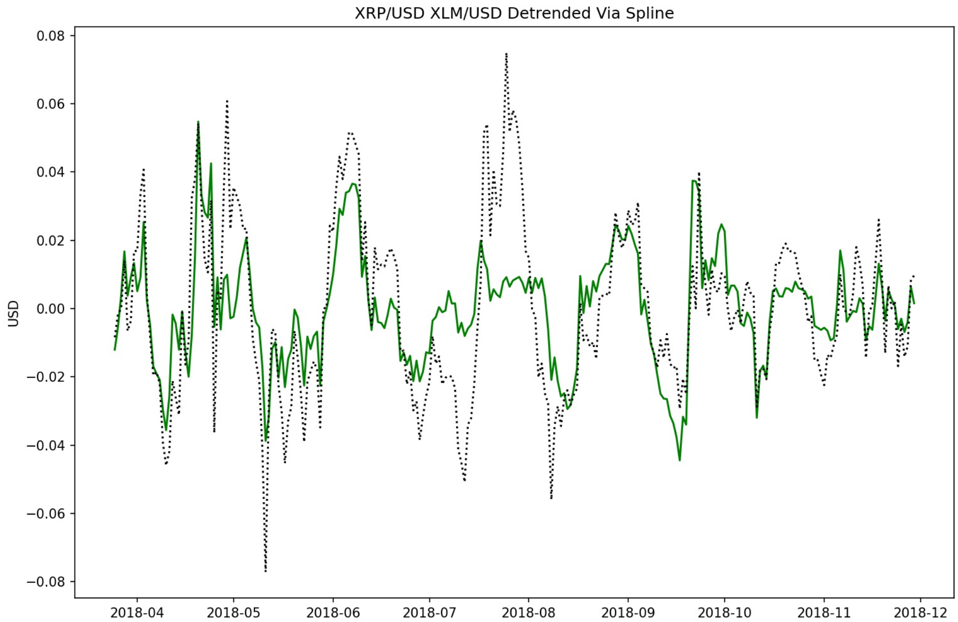
Figure 4: XRP/USD and XML/USD Rates. XRP/USD solid line, XML/USD dotted line.



(a) XRP (Adjusted) Detrended by Spline



(b) XLM (Adjusted) Detrended by Spline



(c) XRP/USD: solid line, XLM/USD: dotted line

Figure 5: XLM and XRP (Adjusted) Detrended by Spline

with XRP as the solid line and XLM as the dotted line. The detrended data is not normally distributed, with non-zero excess kurtosis and skewness equal to 0.48 and 0.102, respectively in XRP and equal to 0.017 and 0.28 in XLM. Figure B.2 in Supplementary Material shows the ACF of the detrended series for XRP and XLM. We observe that the autocorrelations of the detrended XRP and XLM series are decaying gradually to 0 at a slow rate.

3.6 VAR(3) Model of XRP and XLM

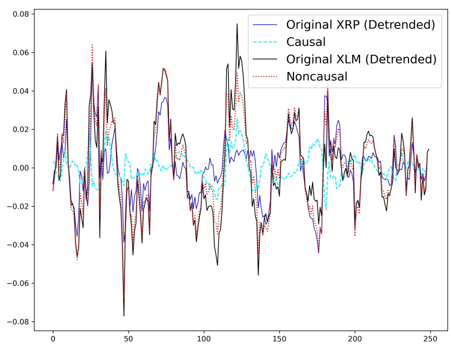
The mixed VAR(1) model estimated from the XRP and XLM rates does not completely remove the serial correlation in the residuals [7]. Hence, to account for the serial correlation in the squared residuals, we increase the autoregressive order of the model as it was done in Section 3.3 for the BTC and ETH series. We set H in the objective function (2.12) equal to 6 and minimize it with respect to Φ. We obtain the augmented autoregressive matrix of coefficients given below.

$$\hat{\Psi}_{GCOV_{XRP/XLM}} = \begin{bmatrix} 1.52 & 0.04 & -2.19 & 1.61 & 1.66 & -1.35 \\ 1.7 & 0.67 & -4.01 & 2.97 & 3.33 & -2.53 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

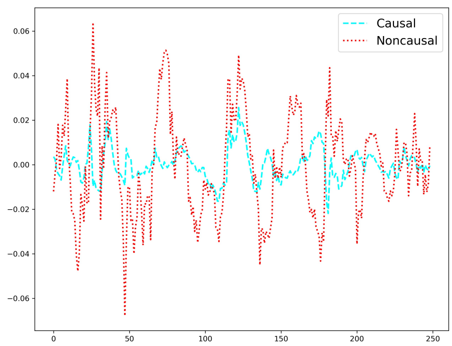
[7]We estimate a mixed VAR(1) model for the XRP and XLM rates by setting the starting values for the minimization procedure to zero and using the BFGS algorithm.

All coefficients are statistically significant according to the standard Wald test. The augmented matrix $\hat{\Psi}_{GCOV_{XRP/XLM}}$ has one eigenvalue of 1.8 outside the unit circle and the following eigenvalues inside the unit circle 0.75, -0.64+0.2i and -0.64-0.2i (of modulus 0.67) 0.46+0.52i 0.46-0.52i (of modulus 0.69) which is consistent with mixed causal noncausal dynamics.

The autocorrelation functions of the residuals and squared residuals from the VAR(3) model for XRP and XLM rates are shown in Figures B.10 and B.11 of Supplementary Material. These plots indicate that the noncausal VAR(3) has captured the linear and nonlinear serial dependence in the residuals. The histograms of VAR(3) residuals for XRP and XLM are given in Figure B.6 of Supplementary Material. The sample distributions of both residual series have large tails indicating non-normality.



(a) XRP and XLM Series with Causal and Noncausal Components



(b) Causal and Noncausal Components

Figure 6: XRP and XLM Series with Causal and Noncausal Components, VAR(3)

Figure 6 above displays the real causal and noncausal components of the XRP/XLM pair of cryptocurrencies. There is only one noncausal component of the VAR(3) model which closely mimics the bubbles and spikes of the series.

The dynamics of the causal component depicts the behavior of a portfolio of cryptocurrency that is free of locally explosive patterns.

Linear regressions of the noncausal component of the XRP/XLM VAR(1) on either the noncausal component of the BTC/ETH VAR(1), or those of VAR(3) reveal strong linear relationships with high R-squared. This motivates our joint analysis of all four series in Section 3.8.

3.7 Comparison of Mixed VAR(3) and Causal VAR(3) for XRP and XLM

Let us compare the mixed VAR(3) model estimated by the GCov estimator with the results obtained from a causal VAR(3) model estimated by the OLS on the XRP and XLM data. The OLS estimated coefficients of a causal VAR(3) model are as follows:

$$\hat{\Psi}_{OLS_{XRP/XLM}} = \begin{bmatrix} 0.867 & -0.166 & 0.108 & 0.0574 & -0.172 & 0.006 \\ -0.066 & 0.677 & 0.143 & 0.029 & -0.057 & -0.104 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

In the OLS estimated equation of XRP_t there are two statistically significant coefficients on XRP_{t-1}, and XLM_{t-2}. In the equation of XLM_t there is only one statistically significant coefficient on XLM_{t-1} at time $t-1$. No other coefficients are statistically significant.

The ACF of the squared residuals in Figure B.12 of Supplementary Material shows that the causal VAR(3) model estimated by OLS, fails to remove serial correlation from the squared residuals. We observe that the autocorrelation of the squared residuals at lags one to three are statistically significant. In contrast, the mixed causal noncausal VAR(3) model is able to remove the nonlinear serial dependence.

As it was the case for the BTC/ETH pair, the causal VAR model fails to capture the feedback effects because of the misspecification due to the assumption of causality.

3.8 VAR(1) of Bitcoin, Ethereum, Ripple and Stellar

We now consider a noncausal VAR(1) model of all four cryptocurrencies estimated from 200 observations recorded between March 5 2018 and October 10 2018, with the values of BTC and ETH divided by a factor of 1000 in order to adjust the data to a closer range of values. The data adjusted by subtracting the median and rescaling are displayed in Figure 7.

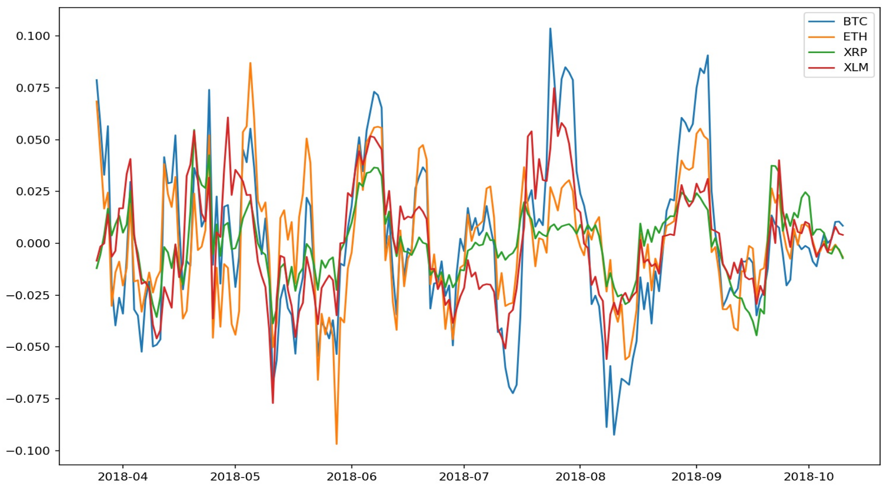


Figure 7: BTC, ETH, XRP, and XLM Rates Spline Detrended (adjusted)

Figure 7 suggests that the cryptocurrency series display comovements. The local explosive patterns of the series resemble one another, indicating the presence of bubbles common to all of them, which motivates the modelling of the series jointly as a mixed model. Because there is a trade-off between the lag order and the dimension of the series in Markov processes [8], we expect the VAR(1) model to provide a satisfactory fit to the data.

By setting H=14 in the objective function (2.12) and minimizing it with respect to Φ with powers two as the nonlinear functions, we obtain the following estimated autoregressive matrix:

[8]The mixed VAR(1) models are Markov of order 1 in both the calendar and reverse time.

$$\hat{\Phi}_{GCOV_{BTC/ETH/XRP/XLM}} = \begin{bmatrix} 0.69 & 0.075 & 0.099 & 0.56 \\ -0.094 & 0.918 & -0.15 & 0.588 \\ -0.2 & 0.0979 & 0.995 & 0.295 \\ 0.306 & -0.326 & -0.0115 & 1.068 \end{bmatrix}.$$

The eigenvalues of the autoregressive matrix given above are as follows: 1.16, 0.79, 0.79+0.23i, 0.79-0.23i with one eigenvalue outside the unit circle and three eigenvalues inside the unit circle (the complex eigenvalues have modulus 0.832). This result implies a mixed VAR(1) process containing three causal and one noncausal components.

The histograms and QQ plots of the residuals given in Supplementary Material in Figures B.13 and B.14 respectively, show large tails of sample densities, consistently with non-normal distribution of the VAR(1) residuals. The Jarque-Bera and Shapiro Wilk test statistics both indicate that the residuals for BTC, ETH, XRP and XLM are not normally distributed.

The Figures B.15 Panel (a) and B.15 Panel (c) in Supplementary Material display the autocorrelation functions of the residuals and squared residuals for BTC and ETH respectively, while Figures B.13 Panel (b) and B.13 Panel (d) display the autocorrelation functions of the residuals and squared residuals for XRP and XLM. We observe that the model removes the serial correlation in the residuals and squared residuals, and provides a good fit to the data.

There is one common noncausal component representing the common bubble and local explosive pattern of the four cryptocurrency series. The noncausal component is displayed in Figure 8 below.

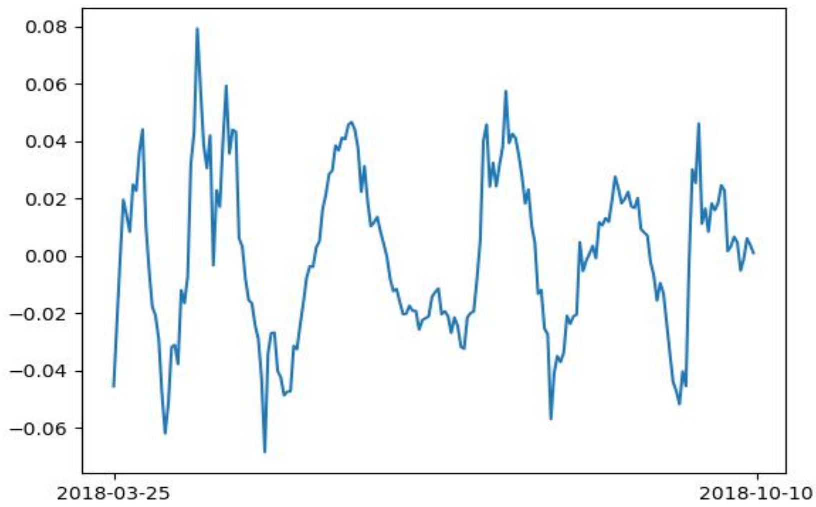


Figure 8: Common Noncausal Component of VAR(1) with four Cryptocurrencies

The noncausal component of the VAR(1) model of the four cryptocurrencies is closely related to the noncausal components of the bivariate processes. A linear regression of the noncausal components of the VAR(1) of dimension four on the noncausal components of the two bivariate VAR(1) and VAR(3) models shows a close relationship between the noncausal components of all series with an R-squared of 0.88.

4 Conclusion

In this paper we examined the prices of the following cryptocurrencies: Bitcoin, Ethereum, Ripple and Stellar in US dollars, documented their comovements and displayed the common bubbles. We modelled the pairs Bitcoin/Ethereum and Ripple/Stellar as bivariate mixed VAR(1) and VAR(3) processes and all four cryptocurrency rates as a mixed VAR(1) process of dimension four.

The mixed causal-noncausal modelling allowed us to decompose the processes into their causal (i.e. 'regular') and noncausal (i.e. 'explosive') components. The noncausal component can be estimated and monitored over time to provide inference on the common bubbles and local explosive patterns, in general. It can be also predicted by using the prediction methods for noncausal processes given in Gouriéroux & Jasiak (2016). The causal component eliminates the common local explosive dynamics of the series depicted by the common bubbles and spikes. Therefore, the common bubble can be interpreted as a common feature in the sense of Engle & Kozicki (1993). The causal component itself is a portfolio of cryptocurrencies, which is immune to the risk of common bubbles and provides a stable investment strategy.

We compared the results from the OLS estimation of causal VAR models with the semi-parametrically estimated mixed causal noncausal models. We found that modelling the cryptocurrency rates as processes with causal and noncausal components enables us to detect nonlinear dependencies within and between these series as well as their comovements, which cannot be captured by the standard linear causal VAR models.

Bibliography

(n.d.).

Antonakakis, N., Chatziantoniou, I., & Gabauer, D. (2019). Cryptocurrency Market Contagion: Market Uncertainty, Market Complexity, and Dynamic Portfolios. *Journal of International Financial Markets, Institutions and Money*, *61*(C), 37-51. Retrieved from https://ideas.repec.org/a/eee/intfin/v61y2019icp37-51.html doi: 10.1016/j.intfin.2019.02.

Bhatnagar, M., Taneja, S., & Rupeika-Apoga, R. (2023). Demystifying the Effect of the News (Shocks) on Crypto Market Volatility. *Journal of Risk and Financial Management.*, *16*(2).

Blanchard, O., & Watson, M. (1982). *Bubble, Rational Expectations and Financial Market.* Lexington, MA: Lexington Books.

Bouri, E., Gupta, R., & Roubaud, D. (2019). Herding Behaviour in Cryptocurrencies. *Finance Research Letters*, *29*(C), 216–221. Retrieved from https://ideas.repec.org/a/eee/finlet/v29y2019icp216-221.html doi: 10.1016/j.frl.2018.07.008

Bouri, E., Saeed, T., Vo, X. V., & Roubaud, D. (2021). Quantile Connectedness In The Cryptocurrency Market. *Journal of International Financial Markets, Institutions and Money*, *71*, 101302. Retrieved from https://www.sciencedirect

.com/science/article/pii/S1042443121000214 doi: https://doi.org/10.1016/j.intfin.2021.101302

Breid, F., Davis, R. A., Lh, K.-S., & Rosenblatt, M. (1991). Maximum Likelihood Estimation for Noncausal Autoregressive Processes. *Journal of Multivariate Analysis*, *36*(2), 175-198. Retrieved from https://www.sciencedirect.com/science/article/pii/0047259X91900568 doi: https://doi.org/10.1016/0047-259X(91)90056-8

Breidt, F., & Davis, R. (1992). Time-reversibility, Identifiability, and Independence of Innovations for Stationary Time Series. *J. of Time Series Analysis*(13), 377–390.

Breidt, F., Davis, R., & Dunsmuir, W. (1992). *On Backcasting in Linear Time Series Models, New Directions in Time Series Analysis, Part I* (G. P. R. Brillinger Caines & Taqqu, Eds.). Springer-Verlag.

Breidt, F., Davis, R., & Lii, K. (1990). Nonminimum Phase Non-Gaussian Autoregressive Processes. *Proc. Natl. Acad. Sci.*(87), 179–181.

Breidt, J., & Davis, A. (1992). Time-Reversibility, Identifiability and Independence of Innovations for Stationary Time-Series. *Journal of Time Series Analysis*, *13*, 377–390.

Brockwell, P., & Davis, R. (1987). *Time Series: Theory and Methods.* Springer.

Catania, L., Grassi, S., & Ravazzolo, F. (2019). Forecasting cryptocurrencies under model and parameter instability. *International Journal of Forecasting*, *35*(2), 485–501.

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, *62*, 182-199. Retrieved from https://www.sciencedirect.com/science/article/pii/S1057521918305271 doi: https://doi.org/10.1016/j.irfa.2018.09.003

Cross, J. L., Hou, C., & Trinh, K. (2021). Returns, volatility and the cryptocurrency bubble of 2017–18. *Economic Modelling*, *104*, 105643. Retrieved from https://www.sciencedirect.com/science/article/pii/S0264999321002327 doi: https://doi.org/10.1016/j.econmod.2021.105643

Cubadda, G., & A. Hecq, A. (2011). Testing for Common Autocorrelation in Data-Rich Environments. *Journal of Forecasting*, *30*, 325-–335.

Cubadda, G., Giancaterini, F., Hecq, A., & Jasiak, J. (2023). Optimization of the Generalized Covariance Estimator in Noncausal Processes. *arXiv preprint arXiv:2306.14653*.

Cubadda, G., Hecq, A., & Voisin, E. (2023). Detecting Common Bubbles in Multivariate Mixed Causal- Noncausal Models. *Econometrics*, *11*, 1–16.

da Gama Silva, P. V. J., Klotzle, M. C., Pinto, A. C. F., & Gomes, L. L. (2019). Herding behavior and contagion in the cryptocurrency market. *Journal of Behavioral and Experimental Finance*, *22*, 41–50. Retrieved from https://www.sciencedirect.com/science/article/pii/S2214635018302090 doi: https://doi.org/10.1016/j.jbef.2019.01.006

Davis, R. A., & Song, L. (2020). Noncausal Vector AR Processes with Application to Economic Time Series. *Journal of Econometrics*, *216*(1), 246–267. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407620300221 (Annals Issue in Honor of George Tiao: Statistical Learning for Dependent Data) doi: https://doi.org/10.1016/j.jeconom.2020.01.017

Devenow, A., & Welch, I. (1996). Rational Herding in Financial Economics. *European Economic Review*, *40*((3-5)), 603–615. Retrieved from https://www.sciencedirect.com/science/article/pii/0014292195000739 doi: https://doi.org/10.1016/0014-2921(95)00073-9

Djogbenou, A., Inan, E., & Jasiak, J. (2023). Time-Varying Coefficient DAR Model and Stability Measures for Stablecoin Prices: An Application to Tether. *Journal of International Money and Finance*, *139*.

Doan B., H., Pham, & Nguyen Thanh, B. (2022). Price Discovery in the Cryptocurrency Market: Evidence from Institutional Activity. *Journal of Industrial and Business Economics*, *49*(1), 111–131.

Dunbar, K., & Owusu-Amoako, J. (2022). Cryptocurrency Returns Under Empirical Asset Pricing. *International Review of Financial Analysis*, *82*, 102216. Retrieved from https://www.sciencedirect.com/science/article/pii/S1057521922001776 doi: https://doi.org/10.1016/j.irfa.2022.102216

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation. *Econometrica*, *55*, 251–276.

Engle, R. F., & Kozicki, S. (1993). Testing for Common Features. *Journal of Business and Economic Statistics*, *11*, 369–380.

Esra Alp Coskun, H. K., Chi Keung Marco Lau. (2020). Uncertainty and Herding Behavior: Evidence from Cryptocurrencies. *Research in International Business and Finance*, *54*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0275531920300957 doi: https://doi.org/10.1016/j.ribaf.2020.101284

Fries, S. (2019). Conditional Moments of Noncausal α-Stable Markov Processes and the Prediction of Bubble Burst Odd. *arXiv*. (preprint: 1805.05397, revised December 2019)

Gourieroux, C., & Zakoian, J.-M. (2017). Local Explosion Modelling by Non-causal Process. *Journal of the Royal Statistical Society Series B*, *79*(3), 737-

756. Retrieved from https://EconPapers.repec.org/RePEc:bla:jorssb:v:79:y:2017:i:3:p:737-756

Gouriéroux, C., & Jasiak, J. (2016). Filtering, Prediction and Simulation Methods for Noncausal Processes. *Journal of Time Series Analysis*, *37*, 405–430.

Gouriéroux, C., & Jasiak, J. (2017). Noncausal Vector Autoregressive Process: Representation, Identification and Semi-parametric Estimation. *Journal of Econometrics*, *200*, 118–134.

Gouriéroux, C., & Jasiak, J. (2018). Misspecification of Noncausal Order in Autoregressive Processes. *Journal of Econometrics*, *205*, 226–248.

Gouriéroux, C., & Jasiak, J. (2023). Generalized Covariance Estimator. *Journal of Business and Economic Statistics*, *41*.

Gouriéroux, C., & Monfort, A. (2014). Revisiting Identification and Estimation in Structural VARMA Models. *CREST*, 2014-30.

Hannen, J. (1973). The Asymptotic Theory of Linear Time-Series Models. *Journal of Applied Probability*, *10*(1), 130–145.

Hayes, A. (2023). *Stablecoins: Definition, how they work, and types.* Retrieved 2023-12-29, from https://www.investopedia.com/terms/s/stablecoin.asp

Hecq, A., & Voisin, E. (2023). Predicting Crashes in Oil Prices During The Covid-19 Pandemic with Mixed Causal-Noncausal Models. In Y. Chang, S. Lee, & J. I. Miller (Eds.), *Advances In Econometrics* (Vol. 45B, pp. 209–233). Emerald Publishing Limited.

Hwang, S., & Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance*, *11*(4), 585-616. Retrieved from https://www.sciencedirect.com/science/article/pii/S0927539804000507 (Special Issue on Behavioral Finance) doi: https://doi.org/10.1016/j.jempfin.2004.04.003

Lanne, M., Luoto, J., & Saikkonen, P. (2012). Optimal forecasting of noncausal autoregressive time series. *International Journal of Forecasting*, *28*(3), 623-631. Retrieved from https://www.sciencedirect.com/science/article/pii/S016920701100135X doi: https://doi.org/10.1016/j.ijforecast.2011.08.003

Lanne, M., & Saikkonen, P. (2011a). GMM Estimators with Non-Causal Instruments. *Oxford Bulletin on Economics and Statistics*, *71*, 581-591.

Lanne, M., & Saikkonen, P. (2011b). Noncausal Autoregressions for Economic Time Series. *Journal of Time Series Econometrics*, *3*(3), 1941–1928.

Lanne, M., & Saikkonen, P. (2013). Noncausal Vector Autoregression. *Econometric Theory*, *3*(3), 447–481.

Lux, T. (1995, July). Herd Behaviour, Bubbles and Crashes. *Economic Journal*, *105*(431), 881–896. Retrieved from https://ideas.repec.org/a/ecj/econjl/v105y1995i431p881-96.html

Ming-Chung, L., & Kung-Sik, C. (2007). Multivariate Reduced-Rank Nonlinear Time Series Modelling. *Statistica Sinica*, *71*(1), 139–159.

Nyakurukwa, K., & Seetham, Y. (2023). Higher Moment Connectedness of Cryptocurrencies: a Time-Frequency Approach. *Journal of Economics and Finance*, *forthcoming*.

of the Ontario Securities Commission, S. (2020). *Quadrigacx: A review by staff of the ontario securities commission.* OSC Review. Washington, DC, USA. Retrieved from https://www.osc.gov.on.ca/quadrigacxreport/web/files/QuadrigaCX-A-Review-by-Staff-of-the-Ontario-Securities-Commission.pdf

Paige, P. R. L., & Trindade, A. A. (2010). The Hodrick-Prescott Filter: A special case of penalized spline smoothing. *Electronic Journal of Statistics*, *4*, 856—874.

Perper, R. (2018, February). China is Moving to Eliminate All Cryptocurrency Trading with a Ban on Foreign Exchanges. *Business Insider*. Business Insider. Retrieved 2017-03-12, from https://www.businessinsider.com/china-eliminates-all-cryptocurrency-trading-2018-2

Phillips, P., & Shi, S. (2018). Financial Bubble Implosion and Reverse Regression. *Econometric Theory*(34), 705–753.

Phillips, P. C. B., Shi, S., & Yu, J. (2015a). Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P500. *International Economic Review*(56), 1043–1075.

Phillips, P. C. B., Shi, S., & Yu, J. (2015b). Testing for Multiple Bubbles: Limit Theory of Real Time Detectors. *International Economic Review*(56), 1079–1134.

Swensen, A. (2022). On Causal and Non-Causal Cointegrated Vector Autoregressive Time Series. *Journal of Time Series Analysis*(43), 178–196. doi: 10.1111/jtsa.12607

Youssef, M., & Waked, S. S. (2022). Herding behavior in the cryptocurrency market during COVID-19 pandemic: The role of media coverage. *The North American Journal of Economics and Finance*, *62*.

Yuan Zhao, W. L., Nan Liu. (2022). Industry Herding in Crypto Assets. *International Review of Financial Analysis*, *84*. Retrieved from https://www.sciencedirect.com/science/article/pii/S1057521922002848 doi: https://doi.org/10.1016/j.irfa.2022.102335

Appendix: Mixed VAR(1) Model of BTC and ETH

This Appendix summarizes the estimation results for the VAR(1) model of Bitcoin and Ethereum [9]. The VAR(1) model is estimated by minimizing the objective function (2.12) with respect to Φ with H equal to 11 and power two as the nonlinear function. We obtain the following estimates of the autoregressive matrix:

$$\hat{\Phi}_{GCOV_{BTC/ETH}} = \begin{bmatrix} 0.12 & 1.18 \\ -0.56 & 2.08 \end{bmatrix}.$$

The eigenvalues for this matrix are 0.55 and 1.6 respectively, which is consistent with a mixed causal-noncausal process. The standard errors for the first row are 0.059, 0.093, respectively and the standard errors for the second row are 0.064 and 0.1, respectively. The coefficients are statistically significant based on the standard Wald test.

The residual variance covariance matrix estimated for the BTC/ETH VAR(1) model is

$$\hat{\Sigma}_{BTC/ETH} = \begin{bmatrix} 1002.37 & 676.4 \\ 676.4 & 650.1 \end{bmatrix}.$$

The ACF of the residuals and squared residuals of the VAR(1) model are shown in Supplementary Material in Figures B.3 and B.4 respectively. .

We find that most serial correlation has been removed, but there still exists evidence of slight autocorrelations at lags 1 and 2, especially in the squared residuals.

[9]Setting the starting values for the minimization procedure to zero and using the BFGS algorithm.

The histograms and QQ plots shown in Figures 9 and 10 respectively, display the sample distributions of the residuals for the VAR(1) BTC and ETH providing evidence of their non-Gaussian distributions. The residual densities have long left tails indicating departures from normality.

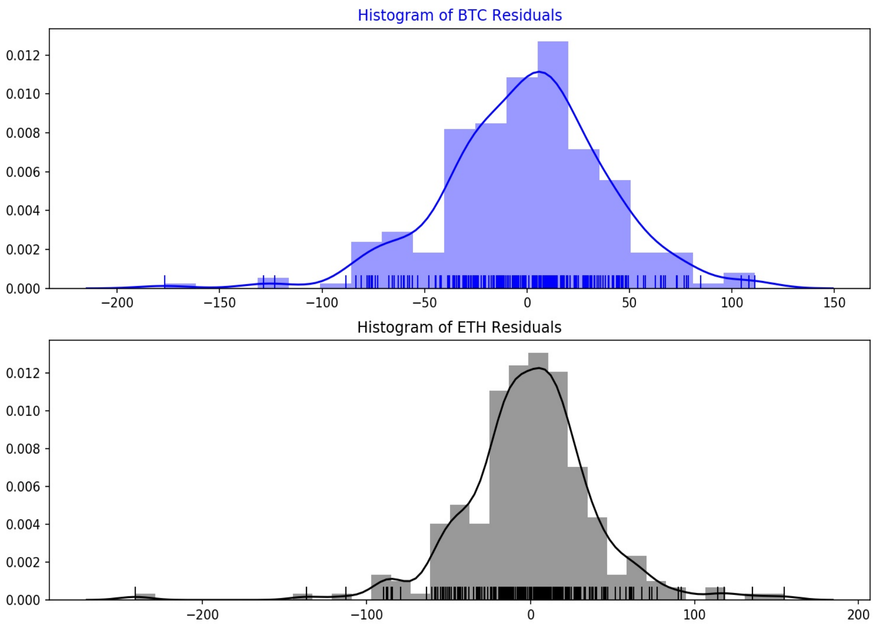
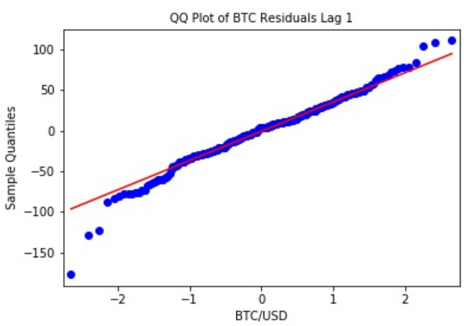
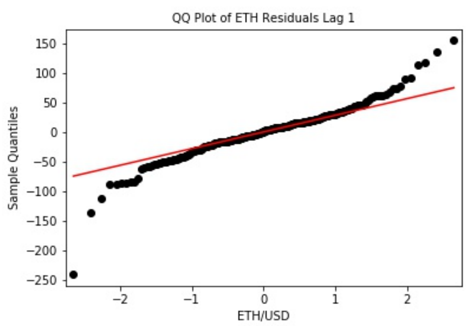


Figure 9: Histograms of Residuals from VAR(1) for BTC and ETH



(a) BTC QQ Plot of Residuals VAR(1)



(b) ETH QQ Plot of Residuals VAR(1)

Figure 10: BTC and ETH Plot of Residuals VAR(1)

In order to further investigate the normality of the residuals we employ a battery of statistical tests: JB - Jarque-Bera, KS - Komolgorov and Smirnoff, DP - D'Agostino and Pearson, Sh - Shapiro whose test statistics and p-values are given in Table 1, where 'p' stands for 'p-value'.

	JB	JB-p	KS	KS-p	DP	DP-p	Sh	Sh-p
BTC	35.03	0.0	0.514	0.0	0.97	0.0	19.2	0.0
ETH	312.5	0.0	0.485	0.0	48.04	0.0	0.93	0.0

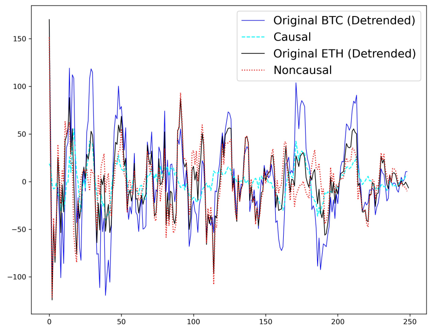
Table 1: BTC and ETH Normality Tests for VAR(1) Residuals

The non-normality is also evidenced by the skewness and the excess kurtosis of 1.63 and -0.42, respectively for the BTC residuals, and of 5.38 and -0.53, respectively for the ETH residuals.

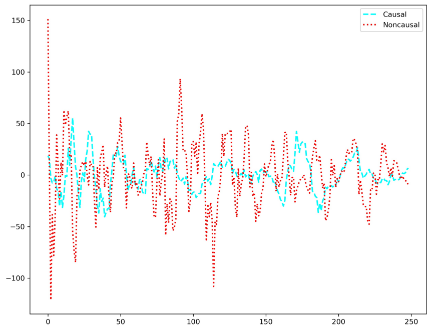
The estimated autoregressive matrix has distinct eigenvalues and is diagonalizable. Having decomposed the autoregressive coefficient matrix $\hat{\Phi} = \hat{A}\hat{J}\hat{A}^{-1}$ (i.e. into Jordan normal form) we can use the blocks of matrix $\hat{A}^{-1}$ to obtain the causal and noncausal components of the process:

$$\hat{Y}^*_{1,t} = \hat{A}^1 Y_t, \qquad \hat{Y}^*_{2,t} = \hat{A}^2 Y_t, \text{ where } \hat{A}^{-1} = \begin{pmatrix} \hat{A}^1 \\ \\ \hat{A}^2 \end{pmatrix}.$$

Below, we plot the two series of cryptocurency rates along with their causal and noncausal components representing the regular and explosive common dynamics in Figure**??**. The causal and noncausal components are shown separately in Figure 11b.



(a) Detrended Series with Causal and Noncausal Components



(b) Causal and Noncausal Components

Figure 11: BTC and ETH Series with Causal and Noncausal Components, VAR(1)

Figure 11a shows the dynamics of the causal and noncausal components of the multivariate process for the sub-sample of BTC and ETH without the original series. We see that the causal component of the model is more smooth compared to the noncausal component. It is the combination that eliminates the common bubbles from the cryptocurrency and provides the investor with a portfolio that is immune to the local explosive patterns, ensuring a stable investment strategy.

Figure 11b shows that the noncausal component displays more volatility than the common causal component. This is because the noncausal component represents the common bubble or explosive local trend in the series. It can be monitored in practice to provide insights to investors, for example when the explosive component exceeds in absolute value a predetermined threshold.

Since the process shows autocorrelation in the squared residuals at lag 1 we increase the number of lags in the VAR model to remove the remaining serial correlation in the squared residuals. This autocorrelation appears to be removed by lag 3, i.e. when the VAR(3) model is fitted to the time series.