Modelling Comovements of Selected Large Cap Cryptocurrencies

Mauri Hall*, and Joann Jasiak †

This version: June 9, 2023

This paper examines the comovements of Bitcoin, Ethereum, Ripple and Stellar exchange rates against the US Dollar in the framework of a causal-noncausal mixed Vector Autoregressive (VAR) models. The cryptocurrency rates are modelled as bivariate and multivariate mixed VAR processes of dimension four. We estimate the mixed VAR models and filter out the common latent components determining the explosive pattern of the series. We also show the the traditional causal VAR model is inadequate for applications to cryptocurrency rates.

Keywords: Noncausal Process, Bitcoin, Ether, Cryptocurrency.

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

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

The authors thank A. Hecq and C. Gourieroux for helpful comments and acknowledge financial support of the Catalyzing Interdisciplinary Research Clusters Initiative, York University.

1 Introduction

The cryptocurrency market has grown considerably over the last decade. Bitcoin and Ethereum are the leaders in terms of market capitalizations of over 500 and 200 Billions, respectively. Stellar and XRP have much smaller capitalizations and lower prices. These four cryptocurrencies: BTC, ETH, XRP and XLM are examined in this paper to determine and explain their common characteristics in the behaviour of their exchange rates against the US Dollar over time. In particular, the four series display similar dynamic patterns, which include short-lived trends interpreted as bubbles and referred to as local explosive patterns in Gourieroux, Zakoian (2017). Other common dynamic price patterns are sudden spikes and time varying volatility, the latter one being a nonlinear pattern found in exchange rates and stock returns. The similarities in cryptocurrency behavior over time, can be examined by considering the dynamics of pairs of cryptocurrency prices, such as BTC and ETH, on the one hand, and XRP and XLM on the other. Pairing up the cryptocurrencies this way is motivated by the fact that Bitcoin and Ethereum were market capitalization leaders while Ripple and Stellar serve similar purposes and share similar underlying technological features. To detect the common dynamic patterns, all four cryptocurrency rate series can be considered as a single multivariate process. The presence of common bubbles in all of them is due to the speculative character of these digital assets.

Under the traditional approach to time series analysis, the bubbles are viewed as nonstationary phenomena, which need to be detected and modelled separately from the stationary component of a time series. There exist a variety of bubble models, such as the Watson bubble for example [Blanchard and Watson (1982)], and tests of bubbles such as those proposed in Phillips and Shi (2018), and Phillips and Yu (2015a) and (2015b).

In this paper, an alternative approach is used. The cryptocurrency rates are modelled as a strictly stationary (mixed) causal-noncausal Vector Autoregressive process and the bubbles and other local explosive patterns are considered as inherent features of this process. The objective of this paper is to estimate the common latent components of the cryptocurrencies, and to display and compare the common latent explosive components. 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 is flawed and fails to detect the comovements between the cryptocurrencies.

Research on modelling the cryptocurrencies as multivariate processes has primarily been focused on the causal i.e. past dependent models. For example, [Catania et al., 2019] find that combinations of parameter varying multivariate causal models can improve the inference. This suggests the importance of accommodating the nonlinearities in cryptocurrency rates, which is done in this paper by including the noncausal components. In addition, we model the rates, i.e. prices rather than returns which separates our approach from the literature. The relationships between the cryptocurrency returns have received a lot of attention in recent years [see e.g. Antonakakis et al. (2019), Bouri et al. (2021), Dunbar and Owusu-Amoako (2022) and Nyakurukwa and Seetharam (2023). We model the cryptocurrency rates as a strictly stationary multivariate process because these assets are mean reverting rather than explosive, i.e. they do not display global trends and long-lasting explosions. Our approach relaxes the traditional assumptions of linearity of the VAR process by allowing for noncausal components that accommodate nonlinear patters in the calendar time, including the aforementioned bubbles and spikes.

A mixed causal-noncausal VAR process Gouriéroux & Jasiak (2017) Davis & Song (2020).] (henceforth referred to as the mixed VAR) has a representation similar to the traditional VAR model. Unlike the traditional VAR, the autoregressive matrix of coefficients of the mixed VAR can have eigenvalues 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 behavior of the series. The eigenvalue(s) strictly larger than one are associated with locally explosive behavior, generating the bubbles and spikes. In addition, the errors of the mixed VAR have to be non-Gaussian and serially independent, identically distributed (i.i.d.) to ensure that the dynamics of the mixed process can be identified.

The assumption of Gaussianity, common in the time series literature, implies the restriction of the parameter space of a given time series model to the causal region. It implies that forward and backward dynamics of a given stochastic process are not distinguishable and, as a consequence the forward looking dynamics cannot be identified. Our empirical results suggest that the assumption of causality is insufficient for capturing the comovements of cryptocurrency rates.

To estimate the mixed VAR model we apply the Generalized Covariance estimator Gouriéroux & Jasiak (2017) which is a one-step, consistent semi-parametric estimator for mixed causal noncausal multivariate non-Gaussian processes. This approach allows us to study the cryptocurrency rates in a semi-parametric setup, i.e. without imposing any distributional assumptions on the errors of the VAR model.

The mixed VAR modelling allows us also to filter out the latent casual and noncausal components, the latter one capturing the bubble phenomena and locally explosive behaviour in strictly stationary time series. The noncausal components in the VAR process represents a common bubble component of the multivariate series. When the series share a common noncausal component we can monitor and forecast that explosive component to notify investors about explosive patterns, such as spikes and bubbles. The monitoring can be beneficial for the investors. The noncausal component can also be predicted [see Gourieroux, Jasiak (2016), Lanne, Luoto,Saikkonen (2013)].

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. Appendix A contains summary statistics, Appendix B contains supplementary graphs.

2 Mixed VAR(p) Model

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

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

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 white noise (i.e. a sequence of uncorrelated random vectors) with mean zero and a positive definite variance matrix Σ. [1]

Under the classical approach y_t is assumed to be causal i.e. past dependent and stationary. 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.

This assumption is too stringent for practical applications as it eliminates stationary noncasual or mixed (i.e. causal and noncausal) dynamics in non-Gaussian processes. Moreover, the normality-based estimation methods such as the normality-based Maximum Likelihood (ML) and Ordinary Least Squares (OLS) applied to such processes do not distinguish between the causal and noncausal dynamics of the process due to the lack of identification issue, and therefore yield inconsistent estimates. The standard Box-Jenkins approach for the estimation of causal, i.e past-dependent VAR, involving normality-based ML or OLS estimation is only adequate for causal linear time-series which are normally distributed, stationary, linear and which can be represented by a moving average with weak white noise errors. The

[1] In equation (1) y_t is assumed to have zero mean.

reason is that the Box-Jenkins approach is based on the identification and estimation of time series from moments of order up to only two. Consequently, the normality-based methods are unable to accommodate jumps, bubbles and local trends which involve higher moments. If the observed time-series is strictly stationary, noncausal and non-Gaussian then we are able to distinguish between the two-sided (those including the past present and future errors) and one-sided (those including only the current and lagged errors) moving average representations written in terms of i.i.d. non-Gaussian errors.

To solve this difficulty, Lanne and Saikkonen (2013) proposed a multiplicative vector autoregressive VAR model for strictly stationary non-Gaussian time series using a multiplicative polynomial representation:

$$\Pi(L)\Phi(L^{-1})y_t = \varepsilon_t,$$

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 limitations of this approach is the fact that the multiplicative polynomial representation with autoregressive orders of r and s may not be unique, and may not always exist for a VAR(p). Moreover, the Approximate Maximum Likelihood (AML) estimator of this model requires the assumption of parametric error distribution. If the distribution is correctly specified then the estimates stemming from the approximate likelihood method are consistent. If, on the other hand, the distribution is not correctly specified then the approximate likelihood method is unreliable.

Gouriéroux & Jasiak (2017), Davis & Song (2020) and Swensen (2022) consider the classical representation (1) under modified assumptions. More specifically Gouriéroux & Jasiak (2017) assume that the errors e_t follow a sequence of non-

Gaussian independent and identically distributed (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 involving the risk of mis-specification. Gouriéroux & Jasiak (2017) introduce a semi-parametric estimator called the Generalized Covariance Estimator (the GCov hereafter) for mixed causal noncausal multivariate non-Gaussian processes. The 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 the 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 (ε_t) is a strong multivariate non-Gaussian white noise of dimension n, and Φ is an $n \times n$ matrix. Gouriéroux & Jasiak (2017) assume that (ε_t) is square integrable with zero mean $E(\varepsilon_t) = 0$, and variance-covariance matrix $V(\varepsilon_t) = \Sigma$. [2]

The eigenvalues of matrix Φ are assumed to be of modulus different from 1

[2]The assumption of square integrability facilitates the derivation of asymptotic properties of the estimators.

as this ensures that a unique, strictly stationary solution to recursive equation (2.1) exists.

Since (ε_t) is not assumed independent of the lagged values of the process $Y_{t-1}, Y_{t-2}...$, the process ϵ_t cannot be interpreted as an innovation.

2.2 Representation Theorem GJ (2017)

This Section reviews the representation theorem of Gouriéroux & Jasiak (2017) introduced for the causal-noncausal mixed processes that distinguishes their purely causal and noncausal latent components. Let us consider the bivariate 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 whose modulus is 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 with dimension $n_1 \times n_1$ and J_2 with dimension $n_2 \times n_2$. The eigenvalues of J_1 (resp. J_2) with their modulus strictly less than 1 (resp. larger than 1) are such that :

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

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

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

where A_1, A_2 (resp A^1, A^2) are the blocks in the decomposition of matrix A as :

$$A = (A_1, A_2) \text{ [resp. in the decomposition of } A^{-1} \text{ as } A^{-1} = \begin{pmatrix} A^1 \\ A^2 \end{pmatrix}].$$

The matrices J_1 and J_2 are derived from the real Jordan canonical form of Φ such that

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

where A contains the eigenvectors of Φ as its columns.

By premultiplying both sides of equation (2.1) and (2.4) by matrix A^{-1} we can decompose Y_t into 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 :

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

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 of (2.2).

Given that the eigenvalues of J_1 have modulus strictly less than 1, the recursive equation below is causal :

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

and recursive backward substitutions can be used derive the causal one-sided moving average representation of $Y_{1,t}^*$ given by the expression below (where L is the lag operator) :

$$Y_{1,t}^* = \Sigma_{h=0}^{\infty} J_1^h \varepsilon_{1,t-h}^* = (Id - J_1 L)^{-1} \varepsilon_{1,t}^*, \tag{2.5}$$

and where

$$(Id - J_1 L)^{-1} \equiv \Sigma_{h=0}^{\infty} J_1^h L^h. \tag{2.6}$$

The second recursive equation is noncausal : $Y_{2,t}^* = J_2 Y_{2,t-1}^* + \varepsilon_{2,t}^*$ can, using recursive substitution, be written thus :

$$Y_{2,t}^* = J_2^{-1} Y_{2,t+1}^* - J_2^{-1} \varepsilon_{2,t+1}^* = (Id - J_2 L)^{-1} \varepsilon_{2,t}^*, \tag{2.7}$$

where :

$$(Id - J_2 L)^{-1} \equiv -\Sigma_{h=1}^{\infty} J_2^{-h} L^{-h}. \tag{2.8}$$

It follows that:

1. There exists a strong (i.i.d) two-sided moving average representation of the

solution of (2.1). Processes $(Y^*_{1,t})$ and $(Y^*_{2,t})$ are purely causal and noncausal processes, respectively. They can be interpreted as the causal and noncausal latent components of process (Y_t);

2. 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 explosive component that follows a strictly stationary noncausal (V)AR process. This component represents the common bubbles, and spikes. It can be univariate or multivariate, depending on the dimension of the time series Y_t.

2.3 Bivariate VAR(1) - Interpretation

To better understand the comovements of Y_t components and their contribution to the latent causal and non-causal components, let us consider again the bivariate VAR(1) process. When matrix Φ is triangular, then depending on the eigenvalues, all components of y_t do not always contribute to both the explosive (i.e. noncausal) and regular (causal) dynamics.

Suppose that matrix Φ has the following Jordan decomposition:

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

where J is the 2 by 2 matrix of eigenvalues A is the 2 by 2 matrix with columns containing the eigenvectors of Φ. Suppose also that $\phi_{12} = 0$ or $\phi_{21} = 0$, so that matrix Φ is 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 entries of the matrix

2) For any triangular matrix, a vector with all elements 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 latent 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 $y^*_{2t} = y_{2t}$:

$$y^*_{1,t} = a^{11}y_1 + a^{12}y_2 = \sum_{j=0}^{+\infty} \lambda_1 \varepsilon^*_{1,t-j}, \tag{2.9}$$

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

$$y^*_{2,t} = a^{22}y_{2,t} = -\sum_{j=0}^{+\infty} [\lambda_2^{-j-1}\, a^{22}\varepsilon_{2,t+j+1}]. \tag{2.10}$$

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 y^*_{2t}.

$$y^*_{1,t} = a^{11} y_1 = \sum_{j=0}^{+\infty} \lambda_1 \varepsilon^*_{1,t-j}, \tag{2.11}$$

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

$$y_{2,t}^{*} = a^{21}y_{1,t} + a^{22}y_{2,t} = -\sum_{j=0}^{+\infty}[\lambda_2^{-j-1}\,(a^{21}\varepsilon_{1,t+j+1} + a^{22}\varepsilon_{2,t+j+1})]. \tag{2.12}$$

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

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

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

Consider the VAR(p) process:

$$Y_t = \Phi_1 Y_{t-1} + ... + \Phi_p Y_{t-p} + \varepsilon_t \, , \tag{2.13}$$

where (ε_t) is a sequence of independent and identically distributed (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

$$X_t = (Y_t', Y_{t-1}', ..., Y_{t-p+1}')' \, , \tag{2.14}$$

to get

$$X_t = \Psi X_{t-1} + u_t. \tag{2.15}$$

The autoregressive polynomial Ψ can be written as

$$\Psi = \begin{bmatrix} \Phi_1 & ... & ... & \Phi_p \\ Id & 0 & ... & 0 \\ 0 & Id & ... & 0 \\ 0 & ... & Id & 0 \end{bmatrix} . \tag{2.16}$$

and

$$u_t = \begin{bmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{1,n} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.17}$$

By the representation theorem given in section (2.2) Ψ can also be written in Jordan canonical form:

$$\Psi = B \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} B^{-1}. \tag{2.18}$$

Similar to Y_t in the VAR(1) case, X_t can be written as the sum of causal and noncausal components as follows

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

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

$$X^*_{1,t} = B^1 X_t,$$

$$X^*_{2,t} = B^2 X_t.$$

Like in the VAR(1) model, the causal and noncausal errors are deterministic functions of the 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 and dim $u^*_{1,t}$ + dim $u^*_{2,t} = n_1 + n_2 = np$ and np is greater than dim $\epsilon_t = n$ whenever $p > 1$.

2.5 Estimation from Nonlinear Autocovariances

The semi-parametric estimation method applicable to mixed causal non-causal multivariate 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 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 $a_k, k = 1, ...K$.

Let us denote by $\Psi_l(Y_t, \phi)$, $l = 1, ..., KH$, function $a_k(Y_{t-h} - \Phi Y_{t-h-1})$, where $\phi = vec\Phi\prime$. For each covariance $c_{kl} = Cov[\Psi_k(Y_t, \phi), \Psi_l(Y_t, \phi)], k, l = 1, ..., KH$, we can write its sample counterpart:

$$\hat{\gamma}_{k,l,T} = \widehat{Cov}[\Psi_k(Y_t, \phi), \Psi_l(Y_t, \phi)],\ k, l = 1, ..., KH.$$

We can then define a covariance estimator as follows:

The Covariance estimator $\tilde{\phi_T}$ of $\phi = vec\Phi\prime$ 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.

The selection of the function a_j determines the semi-parametric efficiency bound of the covariance estimator and the asymptotic efficiency of a Covariance estimator based on a given set of covariances depends on the matrix of weights Ω the estimator, in general, asymptotically semi-parametrically efficient.

One can use the optimal weights Ω that ensure asymptotic semi-parametric efficiency and the associated estimator is then called the Generalized Covariance (GCov) estimator [see, Gourieroux, Jasiak (2022)].

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 and 2) reduced dimension of the objective function to be minimized.

2.5.1 Estimation of Φ_j

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. Note that the GMM estimator is not available for the models with noncausal components. As mentioned earlier in Section 2.1, the error term cannot be interpreted as an innovation and no instruments are available, see Lanne & Saikkonen (2011b) and Lanne & Saikkonen (2011a) for further discussion on the issue.

Let us consider nonlinear functions $a_k, k = 1, ..., K$, and sample autocorrelations $\hat{\rho}_{j,k}(h, \Phi) = Corr[a_j(Y_t - \Phi Y_{t-1}), a_k(Y_{t-h} - \Phi Y_{t-h-1})]$, $\Gamma(0)^{-1}\Gamma(1)$. Then the GCov estimator can be represented as a weighted covariance estimator:

$$\hat{\Phi}_T = \underset{\Phi}{\text{argmin}} \sum_{j=1}^{K} \sum_{k=1}^{K} [\sum_{h=1}^{H} \hat{\rho}_{j,k}^2(h, \Phi)] \tag{2.19}$$

where H is the highest selected lag and the theoretical auto correlations $\rho_{j,k}$ are replaced by their sample counterparts $\hat{\rho}_{j,k}$, seeGouriéroux & Jasiak (2017).

The efficiency of the GCov estimator depends on the choice of functions $a_k, k = 1, ..., K$ and the maximum lag H.

2.5.2 Asymptotic Properties of the GCov Estimator

Gouriéroux & Jasiak (2017) show that GCov estimator $\hat{\phi}$ of $vec(\Phi')$ is consistent and asymptotically normal with asymptotic variance given by: $(D'\Sigma^{-1}D)^{-1}$ $= V_{asy}[\sqrt{T}(\hat{\phi_T} - \phi)]$.

The rows of matrix D are: $D_{k,l} = -\frac{\partial}{\partial \phi'}\widehat{Cov}[\Psi_k(Y_t, \phi), \Psi_l(Y_t, \phi)]$.

The elements of matrix Σ are:

$$\sigma_{(k,l),(k',l')} = Cov_{asy}(\sqrt{T}\widehat{Cov}[\Psi_k, \Psi_l], \sqrt{T}\widehat{Cov}[\Psi_{k'}, \Psi_{l'}])$$

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

2.5.3 Estimation and Identification Steps

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. 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. Compute the residuals $\widehat{\epsilon_t} = Y_t - \widehat{\Phi_1} Y_{t-1} - ... - \widehat{\Phi_p} Y_{t-p}$ and compute their nonlinear autocorrelation function. If the residual autocorrelations are still significant at some lags, then re-estimate the model increasing p and repeat until the residual autocorrelations are no longer significant.

3 Empirical Analysis of Four Cryptocurrencies

3.1 Cyptocurrencies

Let us consider the US dollar exchange rates of the following cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP) and Stellar (XLM) observed at daily frequency obtained from the Bitfinex exchange (www.bitfinex.com). The data display short lived local trends which suggest the presence of noncausal dynamics.

3.2 Bitcoin (BTC) and Ethereum (ETH)

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

Figure 1a displays the daily BTC/USD and ETH/USD exchange 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 but had, as of 2018, lost a large proportion of their respective US dollar values since their peak in late 2017. In addition, both exchange rates show evidence of bubble phenomena, i.e. explosive trends characterized by periods of explosive increases in level followed by rapid decreases in level or vice versa.

Figure 1b displays BTC/USD and ETH/USD exchange rates with medians subtracted. The BTC/USD exchange rate is divided by a factor of ten for comparison and further modelling and, along with the original ETH/USD exchange rate, it is hereafter referred to as the adjusted series.

In Figure 1c we show in the grey region 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.

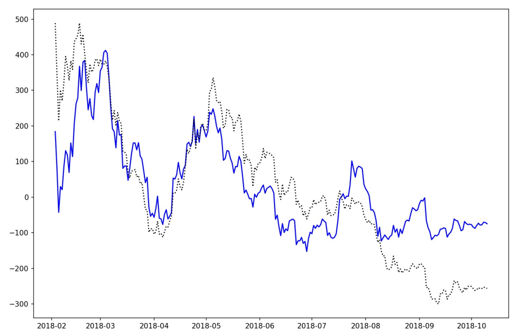
We chose this sub-sample with $T = 250$ because it displayed many spikes in the late 2017 when the large bubble was bursting and the price of cryptocurrency was decreasing. This sub-sample shown in Figures 1b and 1d is then detrended with Python package Obspy [3] using a spline of order 2 with a knot every 30 observations.

[3] Specifically, from Obspy we import the spline package from obspy.signal.detrend

(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 Exchange Rates. BTC/USD solid line, ETH/USD dotted line.

We detrend the subsample using splines instead of the Hodrick–Prescott filter because the Hodrick–Prescott filter is parametric and can destroy the serial dependence pattern Hecq & Voisin (2022). Moreover, one cannot predict it or extrapolate from data detrended using this method. As such we use splines because they are non-parametric. The original and detrended data for BTC and ETH exchange rates can be seen in Figures 2. Figure 2c displays the detrended, adjusted BTC/USD exchange rate as a solid line and the detrended ETH/USD exchange rate as a dotted line.

The autocorrelation function (ACF) in Figure 3 of the detrended data shows finite range of serial dependence. The shaded region in Figure 3 is the asymptotically valid confidence interval at 95%.

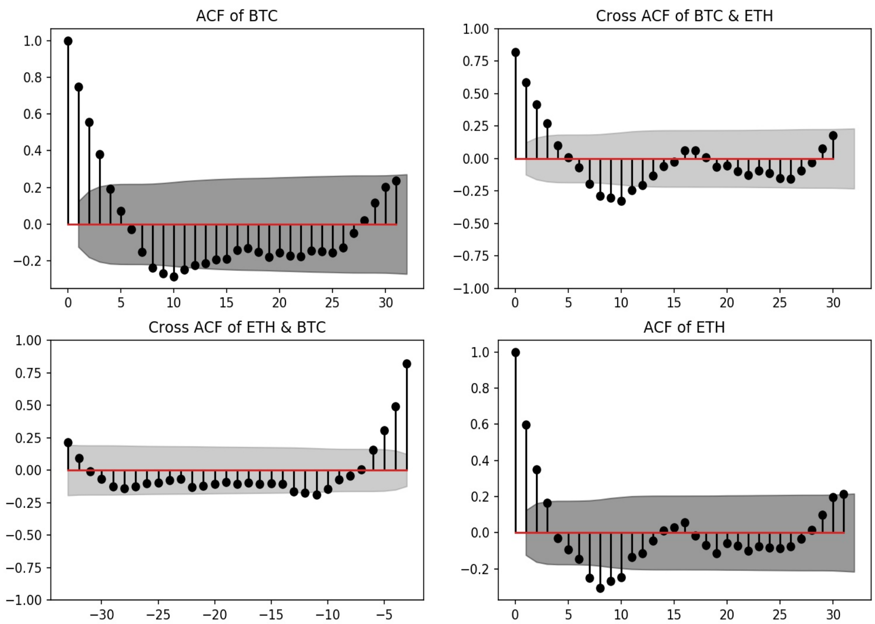
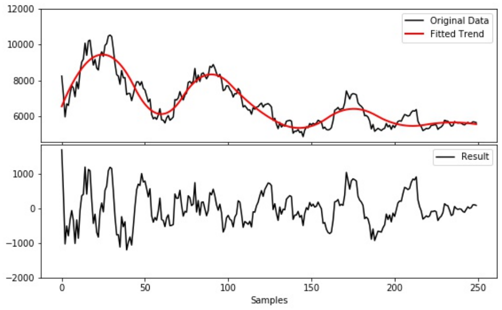
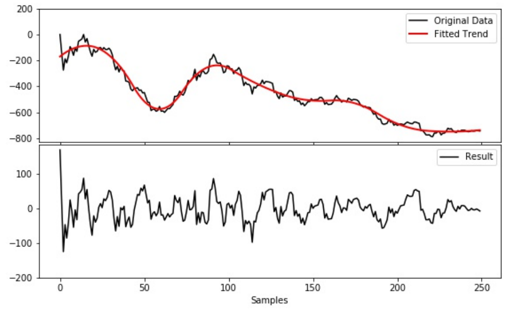


Figure 3: ACF of Detrended BTC and ETH VAR(1)

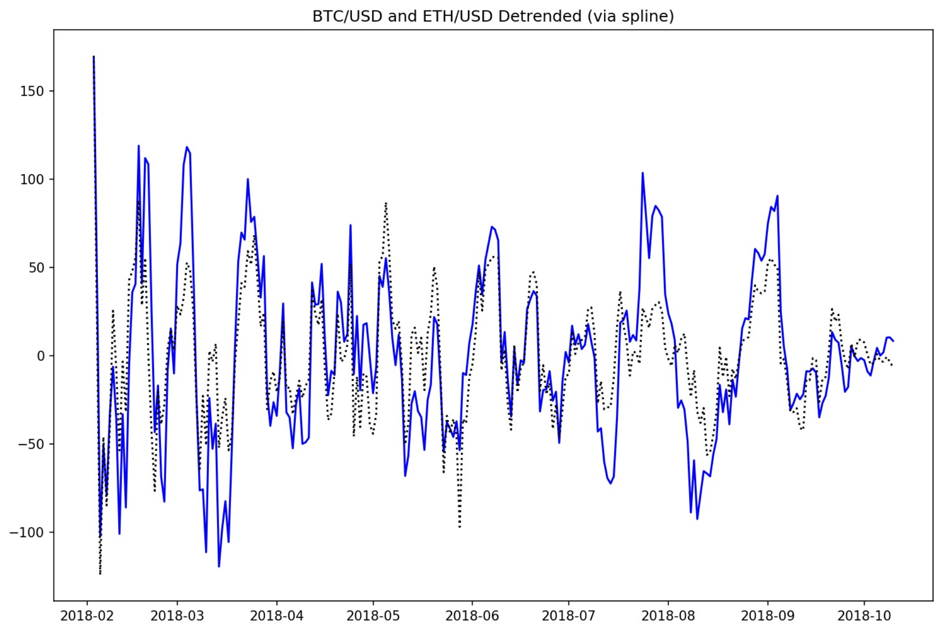
By applying the Augmented Dickey-Fuller test to the detrended data we



(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

find that the resulting process is stationary as all p-values are near zero (for models with and without a constant, constant, with a linear trend and with both linear and quadratic trends - results available on request). Moreover, the detrended data is not normally distributed, with excess kurtosis and skewness of 0.245 and 0.31, respectively in BTC, and of 2.15 and 0.338 in ETH.

3.3 VAR(1) Model of BTC and ETH

The VAR(1) model is estimated minimizing the objective function (2.19) 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 of the VAR(1) model is shown in Appendix B in Figure B.1 and of the squared residuals in Figure B.2 . We find that most serial correlation has been removed but there still exists evidence of slight correlation at lags 1 and 2, especially in the squared residuals.

The histograms and QQ plots shown in Figures 4 and 5 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. Their densities are difficult to specify parametrically, and are characterized by departures from normality.

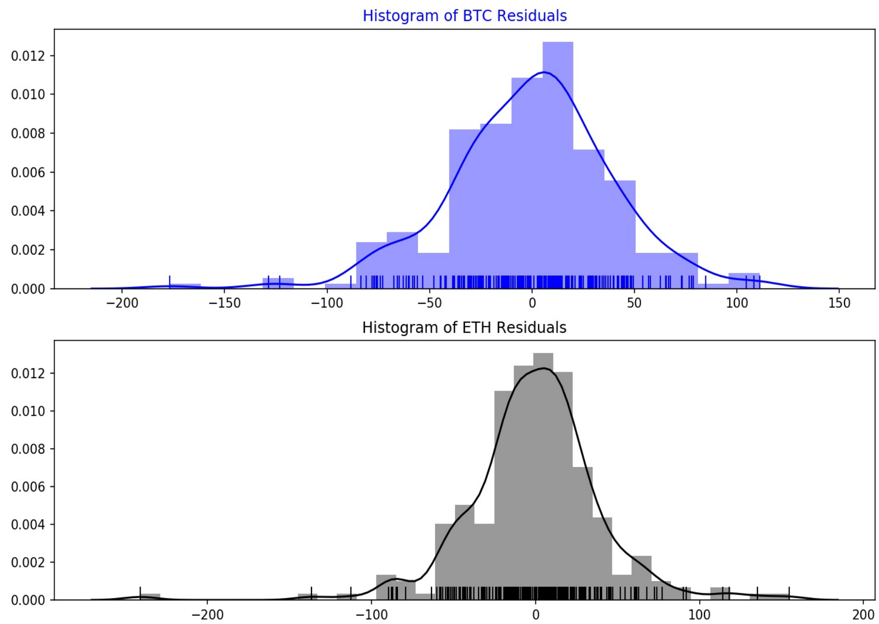
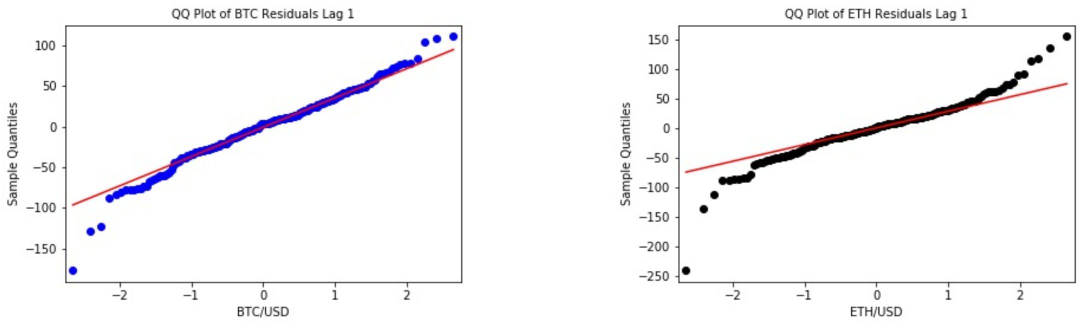


Figure 4: Histograms of Residuals from VAR(1) for BTC and ETH



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

Figure 5: 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 can be seen in Table 1 and excess kurtosis and skewness 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

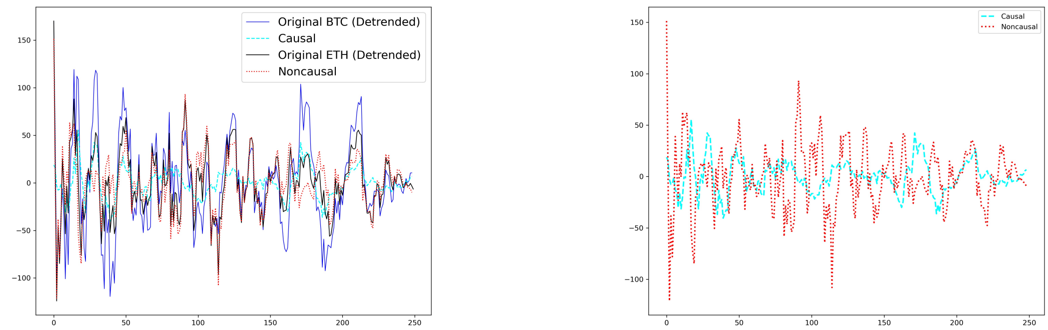
The non-normality is also evidenced by the skewness and the excess kurtosis of 1.63 and -0.42, respectively for the BTC residuals and of 5.38 and -0.53, respectively for the ETH residuals.

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 noncausal components of the process and causal noncausal components of the residuals of both processes.

In order to calculate the causal and noncausal components of the process (which are deterministic functions of the estimated errors) we use blocks of of A^{-1},

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

Below, we plot the two series of exchange rates along with their causal and noncausal components representing the 'regular' and 'explosive' common dynamics in Figure**??**. Next we show only the causal and noncausal components in Figure 6b.



(a) Detrended Series with Causal and Noncausal Latent Components

(b) Causal and Noncausal Latent Components

Figure 6: BTC and ETH Series with Causal and Noncausal Latent Components, VAR(1)

Figure 6a shows the graph 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. Figure 6b shows that the noncausal component displays more volatility than the common causal component. This is expected since 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 pre-determined 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 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.

3.4 VAR(3) Model of BTC and ETH

The VAR(3) model is estimated by setting H in the objective function (2.19) equal to 11 and minimizing it with respect to Φ. We obtain the estimated autoregressive coefficient matrices at lags one, two and three, which are:

$$\hat{\Phi}_{GCOV_{BTC/ETH}i} = \begin{bmatrix} -0.792 & 2.059 \\ -1.268 & 2.06 \end{bmatrix}, \begin{bmatrix} 1.717 & -1.439 \\ -1.268 & 2.06 \end{bmatrix}, \begin{bmatrix} -0.497 & 0.242 \\ 0.087 & -0.099 \end{bmatrix}$$

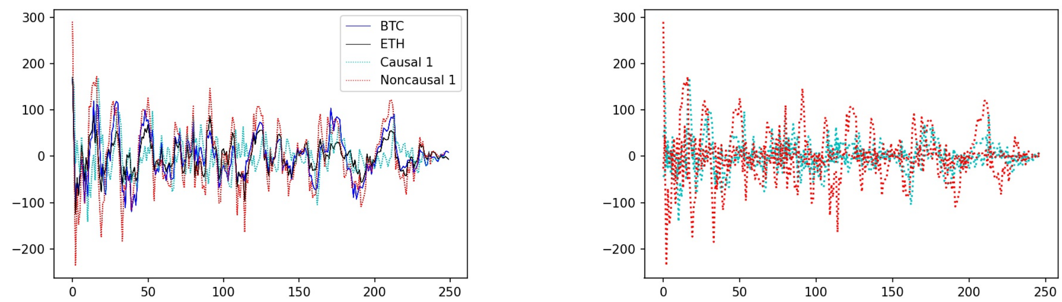
with the following augmented VAR(1) representation:

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

The VAR(3) has two eigenvalues outside the unit circle equal to 1.42 and -1.079. There are two real valued eigenvalues 0.4 and -0.09 and a pair of two complex eigenvalues 0.576+0.4i and 0.576-0.4i, both of modulus 0.7.

Figure B.5 in Appendix B shows that there remains no statistically significant serial correlation in the residuals. Figure B.6 also shows that there does not remain much statistically significant serial correlation in the squared residuals. The VAR(3) models provides a good fit to the data.

The histograms of VAR(3) residuals for BTC and ETH are given in Figures B.3 Appendix B. Both series display large tails indicating non-normality.



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

(b) Causal and Noncausal Latent Components

Figure 7: BTC and ETH Series with Causal and Noncausal Latent Components, VAR(3)

There are two noncausal components in the VAR(3) model of BTC and ETH. In Figure 7 the data and the highest variance latent causal and noncausal components are graphed. Panel (a) contains both the original series and the causal and noncausal components while Panel (b) contains only the causal and noncausal latent components. The causal components are graphed as solid blue lines while the noncausal components are graphed as dotted red lines.

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

We observe that the noncausal components of VAR (1) displayed in Figure

6 (b) has very similar dynamics to the Noncausal 1 component of VAR(3) shown in Figure 7.

A linear regression of the noncausal component of the BTC/ETH VAR(1) on the two noncausal components of the BTC/ETH VAR(3) show close relationship with and R-squared of 0.92.

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

Let us compare the fit of the mixed VAR(3) with a pure causal VAR(3) for BTC and ETH USD exchange rates.

The OLS estimated VAR(3) coefficients are as follows:

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

with the augmented VAR(1) representation:

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

The eigenvalues for this augmented matrix are, 0.943, 0.456+0.43i, 0.456-0.43i (modulus 0.627), 0.463, -0.55, -0.23. The coefficient values 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, Appendix B, shows the presence of serial dependence in the squared residuals of a linear and 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 that a standard linear causal VAR model is unable to accommodate. 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 correlation matrices for the mixed and causal VAR models are shown below. We can see that these two models have similar contemporaneous correlations between their respective residuals.

$$Corr_{[OLS_{BTC/ETH}]} = \begin{bmatrix} 1 & 0.857 \\ 0.857 & 1 \end{bmatrix}, \; Corr_{GCov_{BTC/ETH}} = \begin{bmatrix} 1 & 0.889 \\ 0.889 & 1 \end{bmatrix}$$

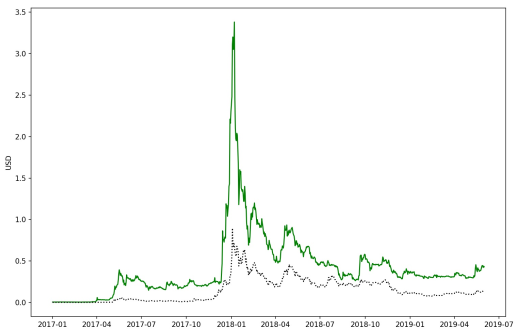
Although the correlations between residuals are similar, the OLS model fails to capture the comovements and feedback effects that the mixed causal noncausal model captures. This is because the OLS model assumes all eigenvalues lie within the unit circle and is therefore misspecified.

3.6 XRP (Ripple) and XLM (Stellar)

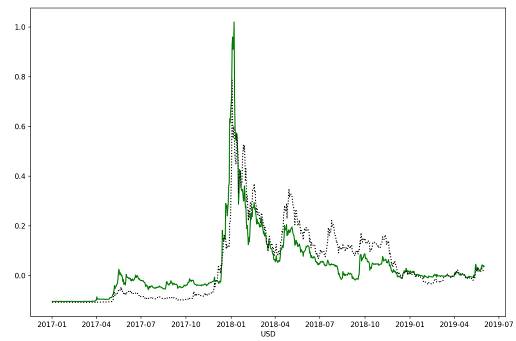
Figure 8a shows the USD exchange rates for Ripple and Stellar for the full sample of 882 daily exchange rate observations between 2017 01 and 2019 06, referred to as XRP and XLM hereafter. Figure 8b displays the same two time series with their medians subtracted (referred to as the adjusted series henceforth) and Figure 8c shows the sub-sample of $T = 250$ observations between 2018 03 25 and 2018 11 29 used for analysis in the context of the full sample in grey and Figure 8d shows the adjusted sub-sample.

The summary statistics for the two series are given in Summary Sample Statistics, Table A.1. We find that the series are not normally distributed. In addition the series display features indicating the presence of bubbles and spikes.

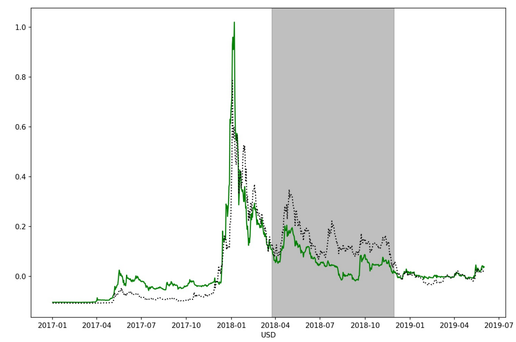
The XRP and XLM exchange rate data is detrended by using a spline of order three and with a knot at every 25 observations using Python package Obspy



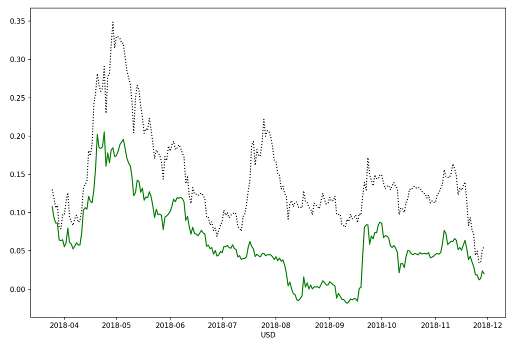
(a) 2017-01-01 to 2019-06-04



(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 8: XRP/USD and XML/USD Exchange Rates. XRP/USD solid line, XML/USD dotted line.

[4]. Figures 9a and 9b show the original and detrended series for XRP and XLM respectively.

Figure 9c shows the adjusted and detrended sub-sample for XLM and XRP with XRP as the solid line and XLM as the dotted line.

By applying the Augmented Dickey-Fuller test to the detrended data we find that the resulting process is stationary as all p-values are near zero (for models with and without a constant, constant, with a linear trend and with both linear and quadratic trends - results available on request). 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 15 shows the ACF of the detrended series.

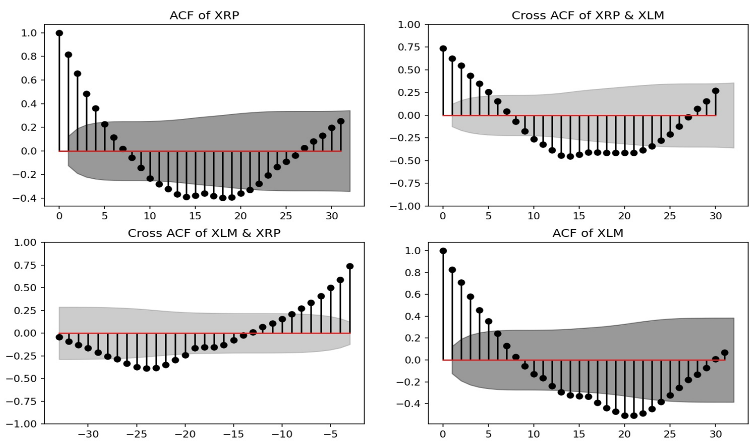
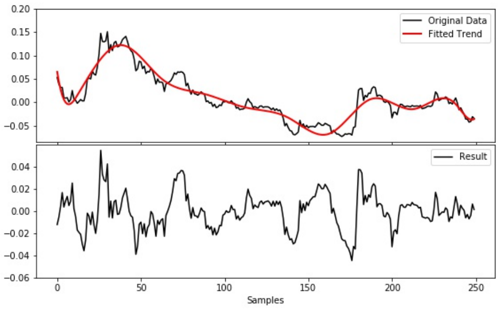


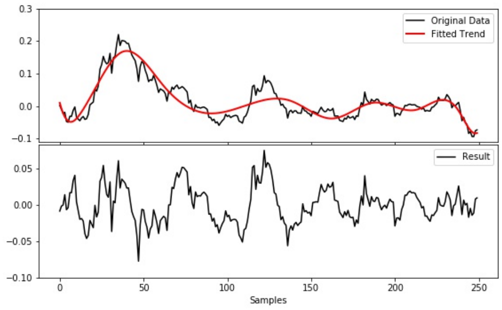
Figure 10: ACF of Adjusted and Detrended XRP and XLM

We observe that the autocorrelations of the detrended XRP and XLM series are decaying gradually to 0.

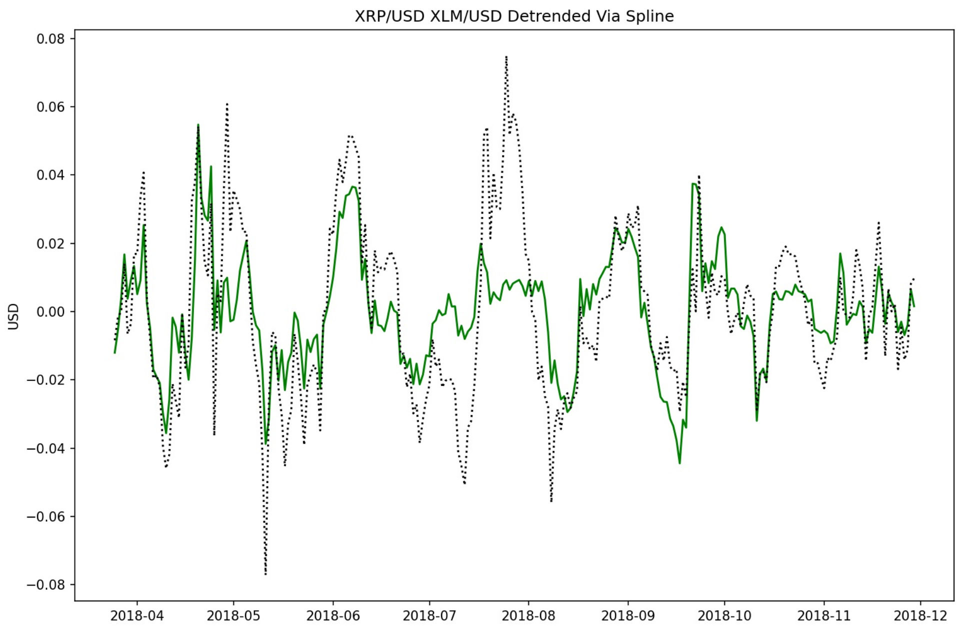
[4]Specifically, from Obspy we import the spline package from obspy.signal.detrend



(a) XRP (Adjusted) Detrended by Spline



(b) XLM (Adjusted) Detrended by Spline



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

Figure 9: XLM and XRP (Adjusted) Detrended by Spline

3.7 VAR(3) Model of XRP and XLM

The mixed VAR(1) model estimated from the XRP and XLM rates does not remove completely the serial correlation in the residuals. Hence, to account for the serial correlation in the squared residuals, we increase the order of the model to VAR(3), as it was done in Section 3.4 for the BTC and ETH series. By setting H in the objective function (2.19) equal to 6 and minimizing it with respect to Φ we obtain the following estimated autoregressive coefficient matrices at lags one, two and three given below.

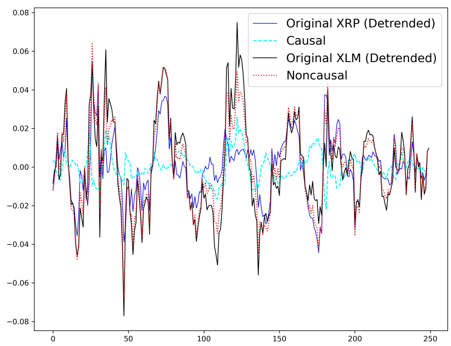
$$\hat{\Phi}_{GCOV_{XRP/XLM}-} = \begin{bmatrix} 1.52 & 0.04 \\ 1.70 & 0.67 \end{bmatrix}, \begin{bmatrix} -2.19 & 1.61 \\ -4 & 2.97 \end{bmatrix}, \begin{bmatrix} 1.66 & -1.35 \\ 3.33 & -2.53 \end{bmatrix}$$

All coefficients are statistically significant according to the standard Wald test.

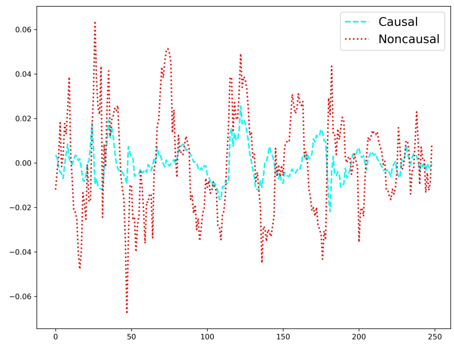
The augmented matrix $\hat{\Psi}_{GCOV_{XRP/XLM}}$ has the following eigenvalues, with one above the unit circle: 1.8, 0.75, -0.64+0.2i and -0.64-0.2i (of modulus 0.67) 0.46+0.52i 0.46-0.52i (of modulus 0.69) which is consistent with a mixed causal noncausal dynamics.

The ACFs for the residuals and squared residuals from the VAR(3) model for XRP and XLM an are shown in Figures B.8 and B.9 respectively in APpendix B. These plots indicate that the noncausal VAR(3) has captured the linear and nonlinear serial dependence in the residuals and is thus an improvement over the mixed bivariate VAR(1) model.

The histograms of VAR(3) residuals for XRP and XLM are given in Figures B.4 Appendix B. Both series display large tails and non-normality of their sample distributions.



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



(b) Causal and Noncausal Components

Figure 11: XRP and XLM Series with Causal and Noncausal Latent Components, VAR(3)

Figure 11 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 mimics closely the bubbles and spikes of the series.

A linear regression of the noncausal components of the XRP/XLM VAR(1) Noncausal Component on the noncausal components of the BTC/ETH VAR(1) and VAR(3) show a linear relationship as the regression has an R-squared of 0.3.

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

Let us compare the mixed VAR(3) estimated by using the GCov estimator with the results obtained from a causal VAR(3) estimated by OLS on the XRP and XLM data.

The OLS estimated VAR(3) coefficients are as follows:

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

In the equation of XRP_t of the OLS estimated VAR(3) 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.10 shows that the causal VAR(3) 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 the causal VAR(3) model. In contrast, the mixed causal noncausal VAR(3) model is able to remove the nonlinear serial dependence.

The correlation matrix is shown below which shows that both the mixed and causal VAR models have similar correlations in their respective residuals. $Corr_{[OLS_{XRP/XLM}]} = \begin{bmatrix} 1 & 0.74 \\ 0.74 & 1 \end{bmatrix}$, $Corr_{GCov_{XRP/XLM}} = \begin{bmatrix} 1 & 0.82 \\ 0.82 & 1 \end{bmatrix}$

As was the case with the BTC/ETH pair, the causal OLS model shows strong correlation in the residuals but fails to capture the feedback effects because of the misspecification of the model due to the assumption of causality.

3.9 VAR(1) For Bitcoin, Ethereum, Ripple and Stellar

We now consider a noncausal VAR(1) of all four cryptocurrencies using 200 observations recorded between March 5th 2018 and October 10th 2018 in which the values of BTC and ETH have been divided by a factor of 1000 in order to adjust the data to a common range of values. The data has been demeaned and scaled in order to perform the estimation and can be seen in Figure 12.

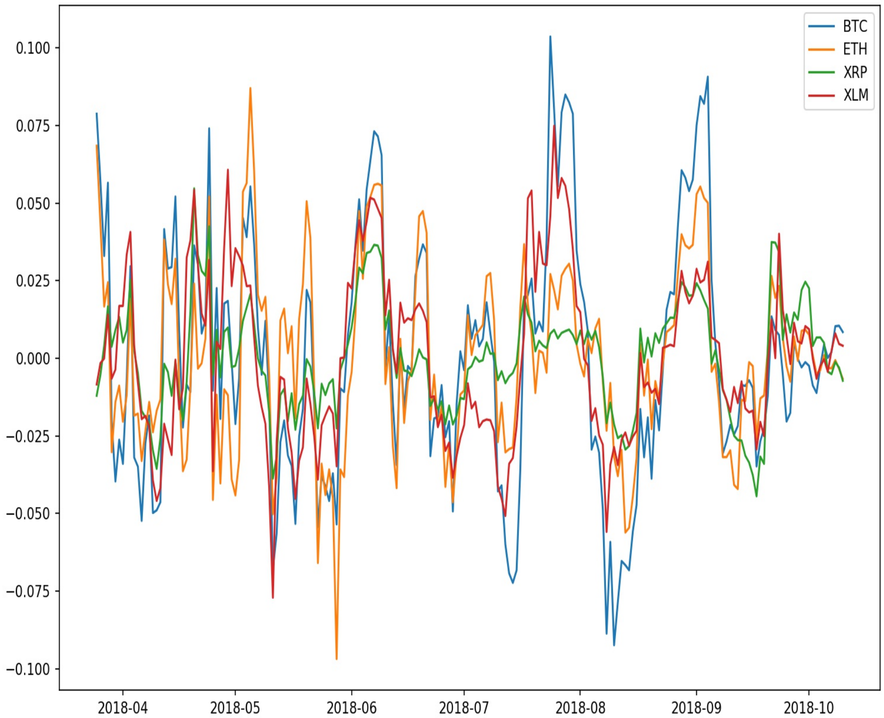


Figure 12: BTC, ETH, XRP, and XLM Exchange Rates Spline Detrended (adjusted)

Figure 12 suggests that the series display comovements. The explosive patterns of the series resemble one another, in particular. This suggests modelling 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 [5], we expect the VAR(1) model to provide a satisfactory fit to the data.

By setting H=14 in the objective function (2.19)) and minimizing it with respect to Φ with powers two as the nonlinear functions, we obtain the following estimated autoregressive matrix:

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

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 (i.e. a process containing both causal and noncausal components).

The histograms of the residuals and QQ plots of the residuals, both in Supplementary Graphs, Figures B.11 and B.12 respectively show large tails consistent with a 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.

Located in Supplementary Graphs Figures B.13a and B.13c display the ACFs for the residuals and squared residuals for BTC and ETH respectively, while Figures B.13b and B.13d display the ACFs for the residuals and squared residuals

[5]The mixed VAR models are Markov in both the calendar and reverse time.

for XRP and XLM.

We observe here that the model removes the serial correlation in the residuals and the squared residuals. This implies that the model provides a good fit to the data.

There is one common noncausal component determining the common explosive patterns of the four cryptocurrency series. The noncausal component is displayed in Figure 13 below.

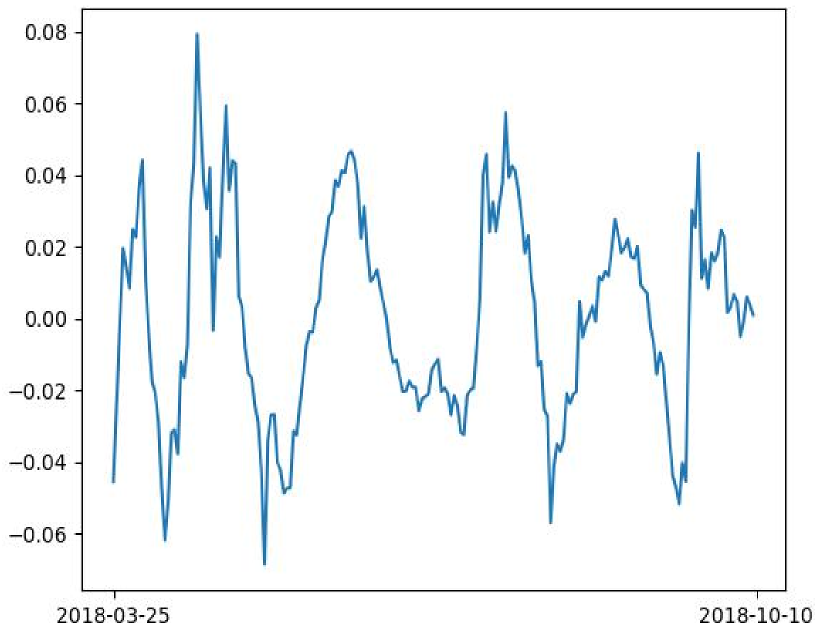


Figure 13: Common Noncausal Component of VAR(1) with four Cryptocurrencies

The noncausal component of the VAR(1) 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 series for both lags one and three 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 US dollar exchange rates for the following cryptocurrencies: Bitcoin, Ethereum, Ripple and Stellar. We modelled these cryptocurrency exchange rates as bivariate VAR(1) and VAR(3) mixed processes for the pairs Bitcoin/Ethereum and Ripple/Stellar and as a VAR(1) mixed process for the four cryptocurrency exchange rates (i.e. a VAR(1) of dimension four).

The mixed causal noncausal modelling has allowed us to decompose the processes into their causal (i.e. 'regular') and noncausal (i.e. 'explosive') latent components. The noncausal component can be monitored over time to provide inference on the local explosive patterns and bubbles. It can be also predicted by using the prediction methods for noncausal processes given in Gourieuroux, Jasiak (2016).

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

Bibliography

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

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.

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

Elie Bouri, X. V. V. D. R., Tareq Saeed. (2021). Quantile Connectedness in the Cryptocurrency Market. *International Review of Financial Analysis*(71), 101302.

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)

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., & 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.

Hecq, A., & Voisin, E. (2022). Predicting Crashes in Oil Prices During the COVID-19 Pandemic with Mixed Causal-Noncausal Models. *Advances in Econometrics in honour of Joon y. Park*, *forthcoming*.

Kwamie Dunbar, J. O.-A. (2022). Cryptocurrency Returns under Empirical Asset Pricing. *International Review of Financial Analysis*(82), 102216.

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.

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

Nikolaos Antonakakis, D. G., Ioannis Chatziantoniou. (2019). Cryptocurrency Market Contagion: Market Uncertainty, Market Complexity, and Dynamic Portfolios,. *Journal of International Financial Markets, Institutions and Money*(61), 37-51.

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

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

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

Phillips, S. S., P., & 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

A Summary Sample Statistics

	BTC	ETH	XRP	XLM
T	250	250	250	250
Mean	5613	317.1	0.41	0.15
Std	3460	263.71	0.4	0.145
Excess Kurtosis	1.52	-0.81	14.74	1.89
Skew	0.96	1.33	3.16	1.27
Min	778.6	8.2	0.005	0.0017
Max	19187	1380	3.38	0.9

Table A.1: Summary Statistics

B Supplementary Graphs

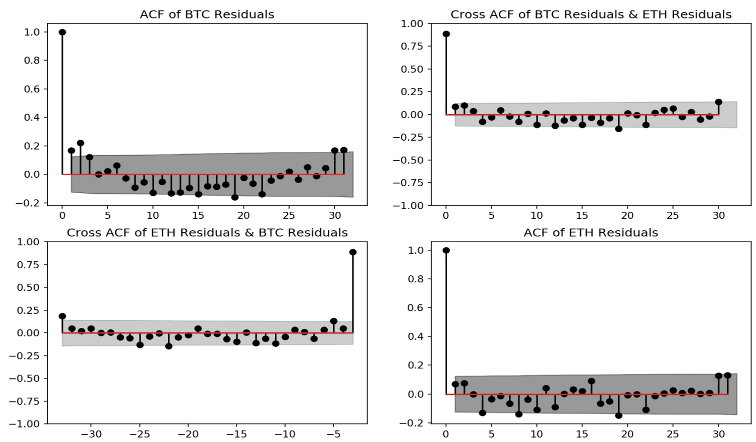


Figure B.1: ACF of Residuals from VAR(1) for BTC and ETH

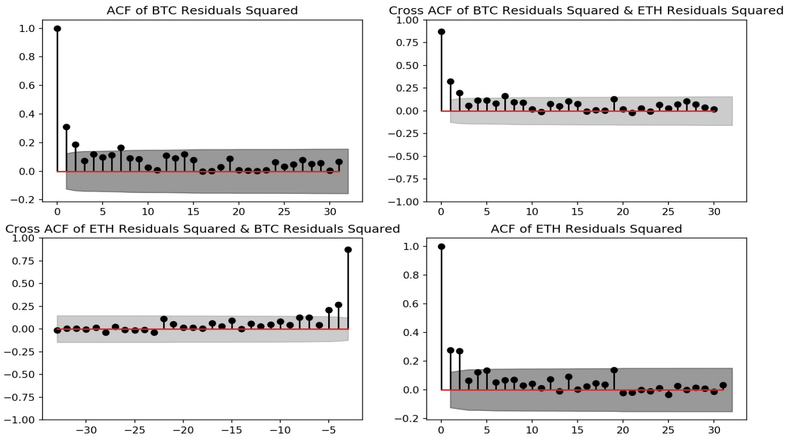


Figure B.2: ACF of Squared Residuals from VAR(1) for BTC and ETH

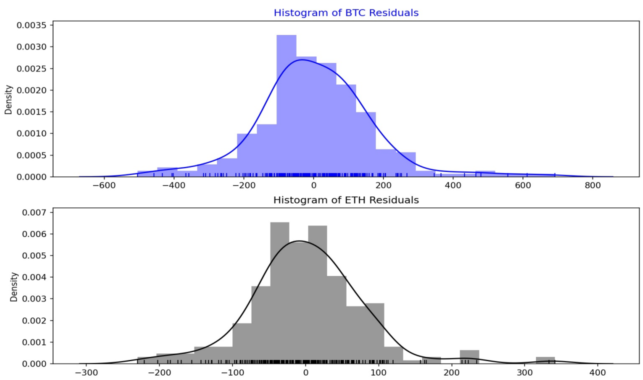


Figure B.3: BTC/ETH Histograms of VAR(3) Residuals

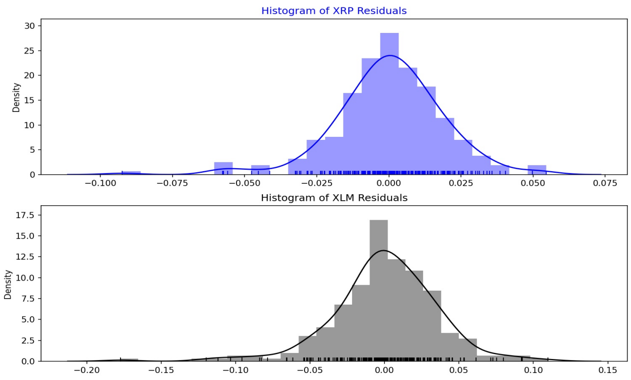


Figure B.4: XRP/XLM Histograms of VAR(3) Residuals

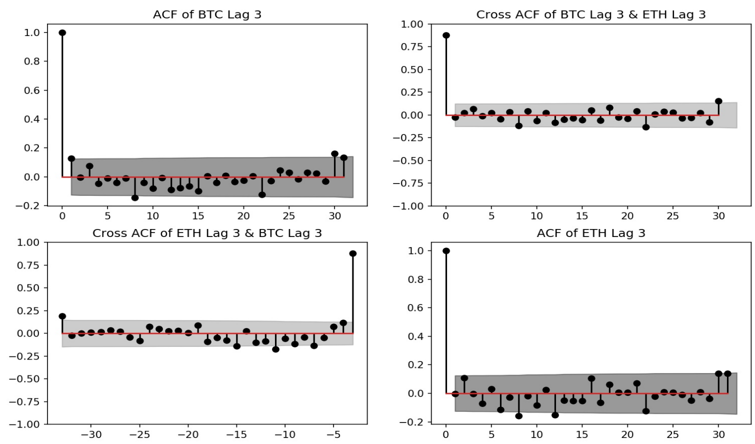


Figure B.5: ACF of Residuals from VAR(3) for BTC and ETH

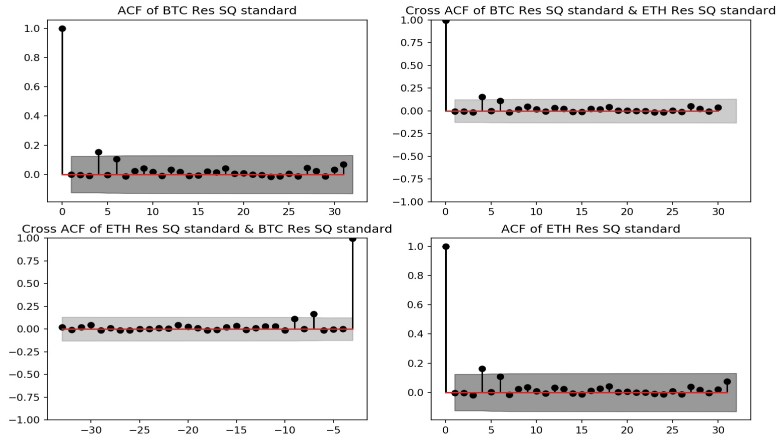


Figure B.6: ACF of Squared Residuals from VAR(3) for BTC and ETH

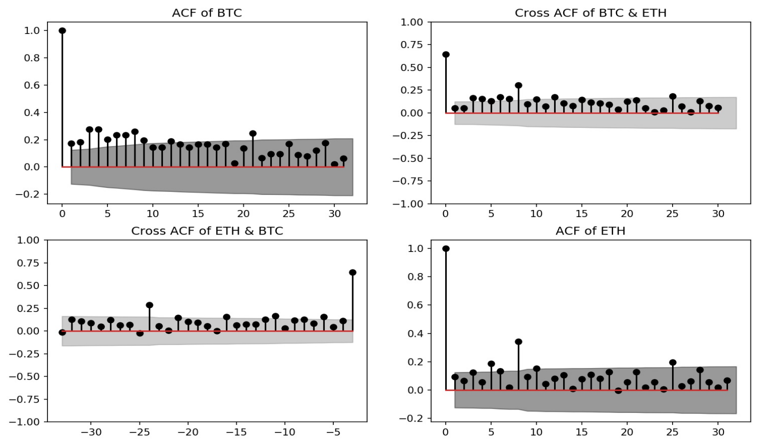


Figure B.7: ACF of Squared Residuals from Causal VAR(3) for BTC and ETH

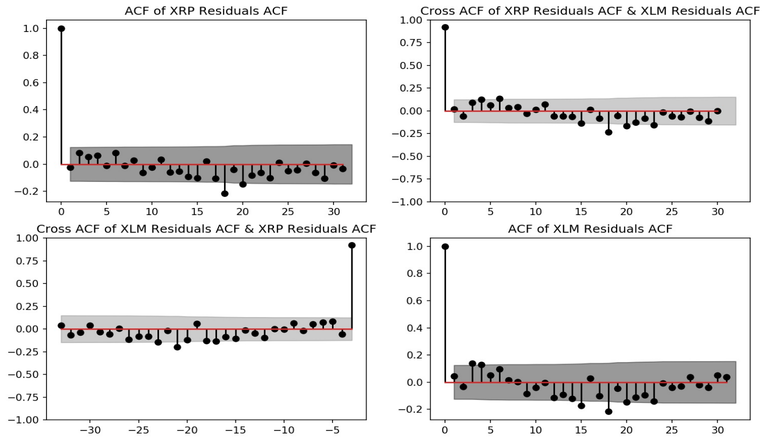


Figure B.8: ACF of VAR(3) Residuals for XRP and XLM

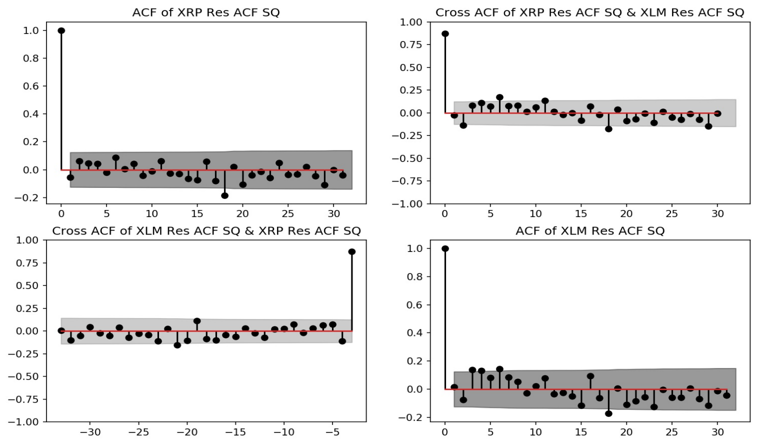


Figure B.9: ACF of VAR(3) Squared Residuals for XRP and XLM

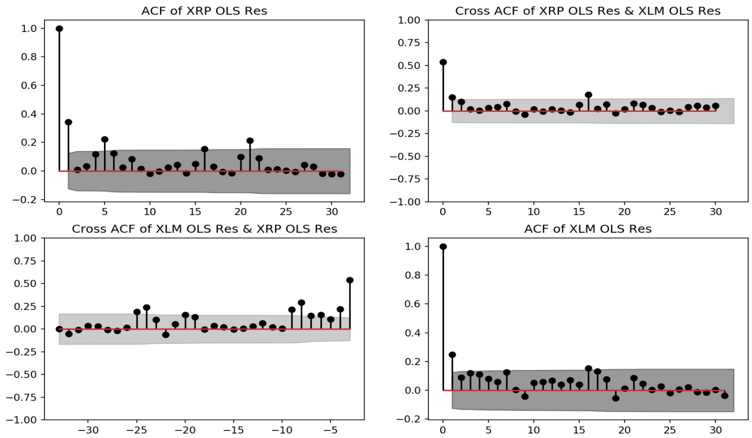
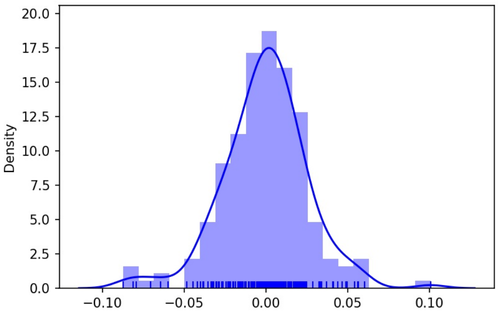
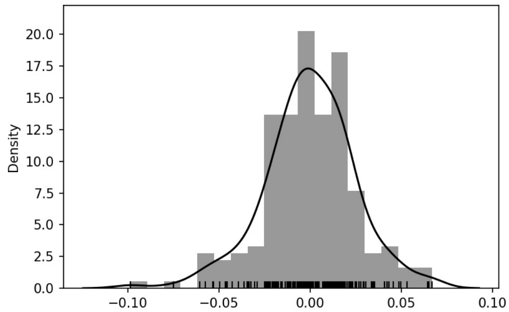


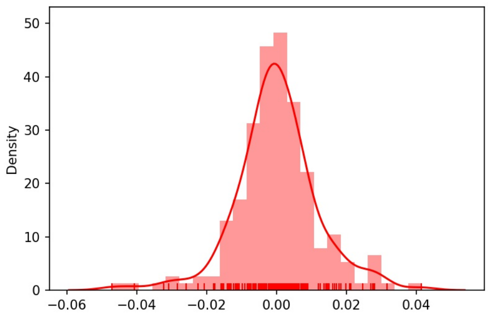
Figure B.10: ACF of Causal VAR(3) Squared Residuals for XRP and XLM



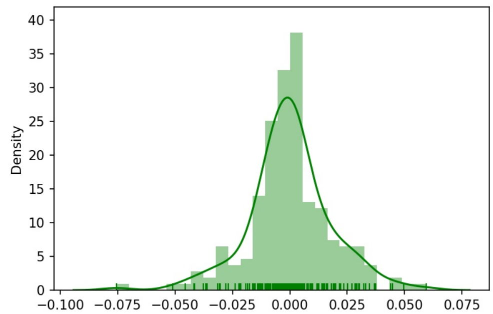
(a) BTC Residuals



(b) ETH Residuals

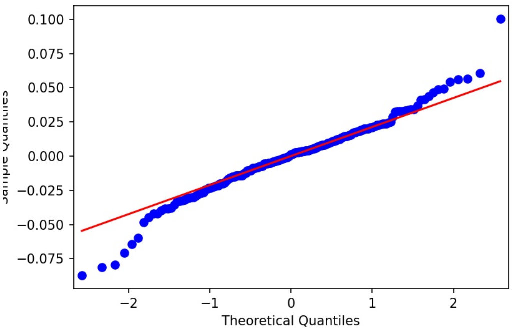


(c) XRP Residuals

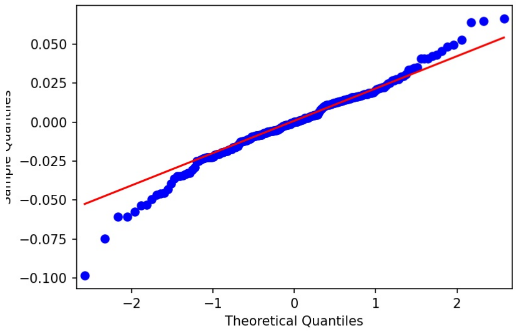


(d) XLM Residuals

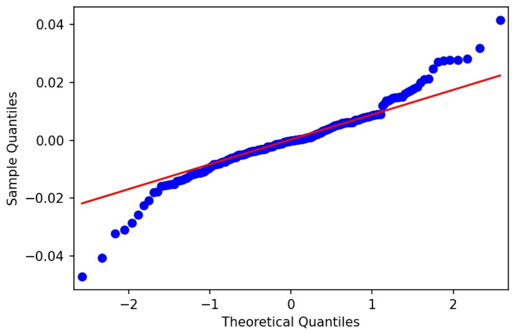
Figure B.11: BTC/ETH/XRP/XLM Histograms of Residuals from VAR(1)



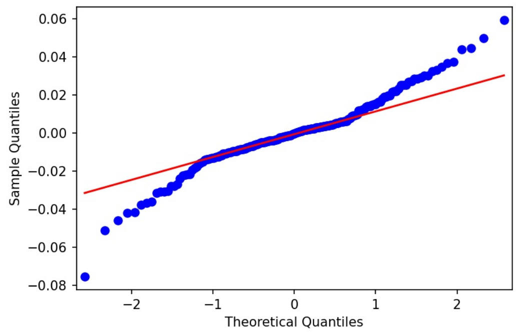
(a) BTC QQ Plot



(b) ETH QQ Plot

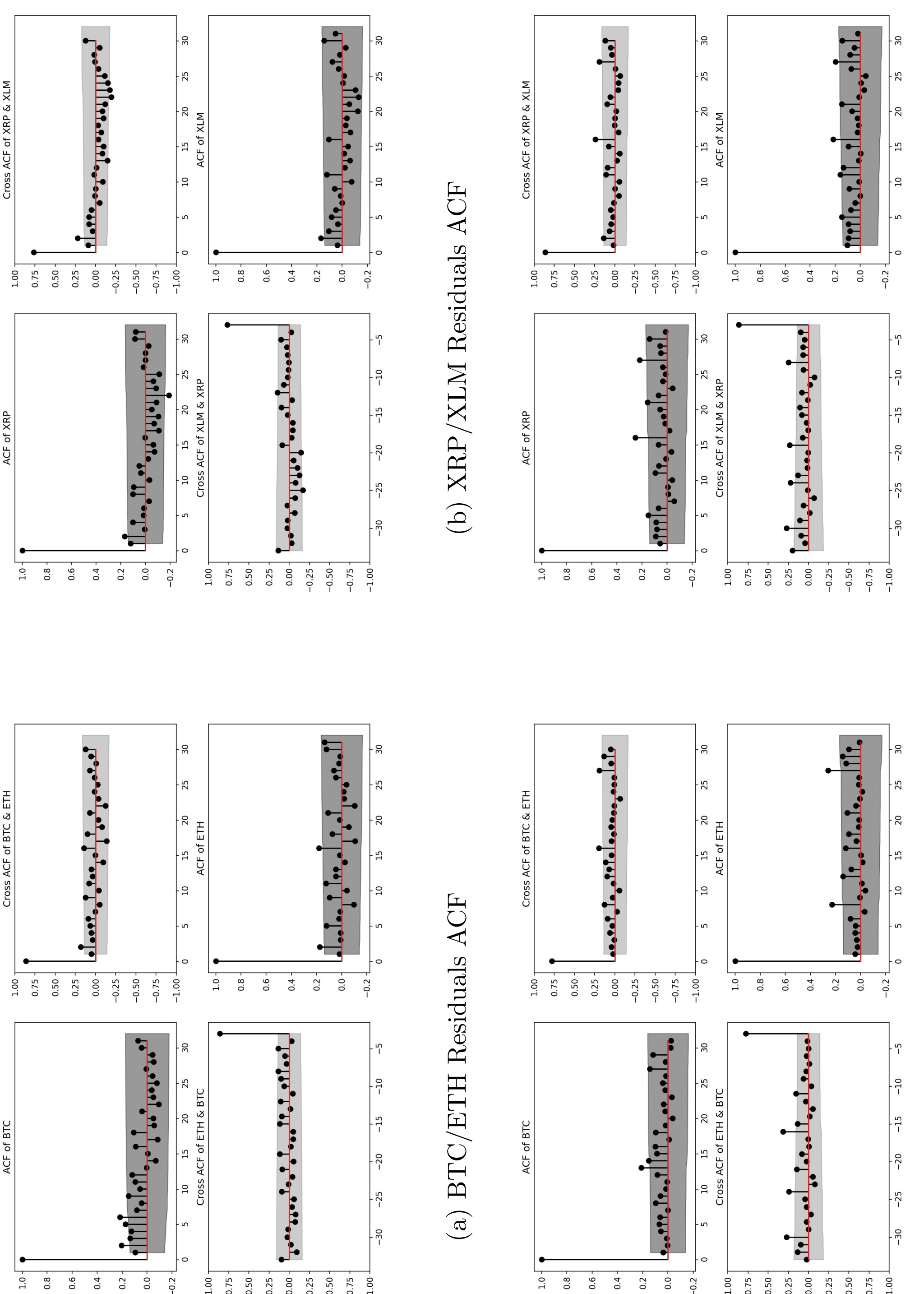


(c) XRP QQ Plot



(d) XLM QQ Plot

Figure B.12: BTC/ETH/XRP/XLM QQ Plots of Residuals from VAR(1)



(a) BTC/ETH Residuals ACF

(b) XRP/XLM Residuals ACF

(c) BTC/ETH Squared Residuals ACF

(d) XRP/XLM Squared Residuals ACF

Figure B.13: ACF of Demeaned VAR(1) Residuals and Squared Residuals for BTC, ETH, XRP, XLM pairs (n=4)