

Keilbar, Georg; Zhang, Yanfen

**Article — Published Version**

## On cointegration and cryptocurrency dynamics

Digital Finance

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Keilbar, Georg; Zhang, Yanfen (2021) : On cointegration and cryptocurrency dynamics, Digital Finance, ISSN 2524-6186, Springer International Publishing, Cham, Vol. 3, Iss. 1, pp. 1-23,  
<https://doi.org/10.1007/s42521-021-00027-5>

This Version is available at:

<https://hdl.handle.net/10419/287694>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# On cointegration and cryptocurrency dynamics

Georg Keilbar<sup>1</sup> · Yanfen Zhang<sup>1,2</sup>

Received: 27 May 2020 / Accepted: 7 January 2021 / Published online: 17 February 2021  
© The Author(s) 2021

## Abstract

This paper aims to model the joint dynamics of cryptocurrencies in a nonstationary setting. In particular, we analyze the role of cointegration relationships within a large system of cryptocurrencies in a vector error correction model (VECM) framework. To enable analysis in a dynamic setting, we propose the *COINtensity* VECM, a nonlinear VECM specification accounting for a varying systemwide cointegration exposure. Our results show that cryptocurrencies are indeed cointegrated with a cointegration rank of four. We also find that all currencies are affected by these long term equilibrium relations. The nonlinearity in the error adjustment turned out to be stronger during the height of the cryptocurrency bubble. A simple statistical arbitrage trading strategy is proposed showing a great in-sample performance, whereas an out-of-sample analysis gives reason to treat the strategy with caution.

**Keywords** Cointegration · VECM · Nonstationarity · Cryptocurrencies

**JEL Classification** C32 · C58

---

We would like to thank the two anonymous referees and the editor for very valuable comments which helped to significantly improve the paper. Financial support from the Deutsche Forschungsgemeinschaft via the IRTG 1792 “High Dimensional Nonstationary Time Series”, Humboldt-Universität zu Berlin, is gratefully acknowledged. This research has also received funding from the European Union’s Horizon 2020 research and innovation program “FIN-TECH: A Financial supervision and Technology compliance training programme” under the grant agreement No 825215.

---

✉ Georg Keilbar  
georg.keilbar@hu-berlin.de

<sup>1</sup> IRTG 1792 “High Dimensional Nonstationary Time Series”, School of Business and Economics, Humboldt-University of Berlin, Unter den Linden 6, 10099 Berlin, Germany

<sup>2</sup> School of Economics, Xiamen University, Xiamen 361005, China

## 1 Introduction

Cryptocurrencies have emerged as a new asset class over recent years. As of 2020, the crypto universe includes almost 5000 currencies with a total market capitalization close to 200 bn USD (coinmarketcap.com). We refer to Härdle et al. (2020) for a general overview on cryptocurrencies. While the market is still dominated by Bitcoin (BTC), the analysis of the interdependence of cryptocurrencies received a lot of attention from researchers as well as practitioners. For instance, Guo et al. (2018) analyzed latent communities from a network perspective. A large strand of literature is concerned with the relation of cryptocurrencies to other more traditional classes of assets (Shahzad et al. 2019, Corbet et al. 2018). Yi et al. (2018) and Ji et al. (2019) analyzed directional volatility spillover effects using the variance decomposition method of Diebold and Yilmaz (2014). Sovbetov (2018) analyzed the cointegration of a VAR system of four cryptocurrencies. Leung and Nguyen (2019) proposed and discussed cointegration-based trading strategies.

While existing research contributions on cointegration restrict their focus to a small number of currencies, we argue that this only paints an incomplete picture. This paper aims to model the joint dynamics of cryptocurrencies in a nonstationary and high dimensional setting. In particular, we investigate the role of potential cointegration relationships among cryptocurrencies. In our empirical analysis, we consider the ten largest currencies in terms of market capitalization in the period from July 2017 to February 2020.

Our methodology is based on the vector error correction model (VECM), developed by Engle and Granger (1987), which augments the standard vector autoregressive (VAR) model with an additional role for deviations from long-run equilibria. To analyze the cointegration of cryptocurrencies in a dynamic setting, we propose a novel nonlinear VECM model, which we call *COINtensity* (cointegration intensity) VECM. The use of nonlinear specifications to model time series has a long tradition, see the monographs of Granger and Teräsvirta (1993) and Fan and Yao (2008). Examples for nonlinear time series models include the smooth transition autoregressive (STAR) model (Luukkonen et al. 1988; Teräsvirta 1994) and neural networks (Kuan and White 1994; Lee et al. 1993). Nonlinear error correction models are discussed in Dv et al. (2002) and extended to the vector case by Kristensen and Rahbek (2010). An advantage of nonlinear time series models is the increased flexibility compared to linear specifications. Usually, this flexibility comes at the expense of a large number of parameters to estimate. Our *COINtensity* VECM specification has the advantage that the number of additional parameters is equal to the cointegration rank, i.e. it is non increasing in the dimension of the VAR system. The nonlinear part of the model introduces a time-varying intensity effect for the error adjustment, which implies that the cryptocurrencies will return to the long-run equilibrium with varying speed. A crucial task is to select the number of those equilibria, also referred to as cointegration relations. Johansen (1988, 1991) proposed a likelihood ratio test, which is now commonly used. However, the testing procedure suffers from poor finite sample performance in systems of more than three variables (Johansen 2002; Liang and Schienle 2019). We, therefore, follow (Onatski and

Wang 2018), who proposed an alternative test for cointegration that is designed for a high-dimensional setting.

Our empirical results suggest that cointegration plays a crucial role in cryptocurrencies. In particular, we find four stationary long-run equilibria. We also find that all currencies are significantly affected by long-term stochastic trends, rejecting the hypothesis of weak exogeneity. The results of our dynamic *COINtensity* VECM show a time-varying dependence of cryptocurrencies on these stochastic trends. We find that the nonlinearity of error correction is stronger during the time of the cryptocurrency bubble, compared to a later time period. Based on our estimated cointegration vectors, we construct a simple trading rule, following and generalizing the strategy of Leung and Nguyen (2019). An in-sample analysis of our trading strategy indicates that trading on large deviations from the long-run equilibria can be profitable, while the out-of-sample analysis is more cautious. In particular, the success of such a statistical arbitrage strategy is dependent on the condition that the equilibrium relations will hold in the long-run.

The contributions of this paper are twofold. First, it is the first attempt to model a system of cryptocurrencies in a large vector autoregression while accounting for nonstationary effects. Second, we propose a novel, nonlinear VECM specification which increases the flexibility and also has a good interpretability even in large dimensions.

The remainder of the paper is organized as follows. Section 2 describes in detail the steps of our modelling and estimation procedure. To show the validity of our approach, we conduct a small simulation study in Sect. 3. In Sect. 4, we apply our methodology to a system of the largest ten cryptocurrencies. Section 5 introduces a simple cointegration-based trading strategy and Sect. 6 concludes.

All codes of this paper are available on [quantlet.de](https://quantlet.de).

## 2 Modelling framework

### 2.1 VECM and testing for cointegration

As a baseline model we consider the following  $p$ -dimensional vector autoregressive model with error correction term (VECM).

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^k \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t, \quad (1)$$

where  $D_t$  are deterministic variables and  $\varepsilon_t$  are zero-mean, independent error terms. We assume that each univariate time series is integrated of order one,  $X_{it} \sim I(1), i = 1, \dots, p$ . Under cointegration, there exists a linear combination which is stationary, i.e.  $\beta^\top X_t \sim I(0)$ . Thus, we can rewrite (1) in the following way,

$$\Delta X_t = \alpha \beta^\top X_{t-1} + \sum_{i=1}^k \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t, \quad (2)$$

where  $\beta$  is a  $p \times r$  matrix of cointegration vectors and  $\alpha$  is the  $p \times r$  loading matrix. The order of cointegration is characterized by the rank  $r$  of  $\beta$ .  $\Gamma_i$ ,  $i = 1, \dots, k$ , are  $p \times p$  parameter matrices associated with the impact of lagged values of  $\Delta X_t$ .

Johansen (1988, 1991) developed a sequential likelihood testing procedure to determine the cointegration rank  $r$ . Under the null hypothesis there are at most  $r$  cointegration relationships.

$$H_0 : \text{rank}(\Pi) \leq r \quad \text{vs} \quad H_1 : \text{rank}(\Pi) > r \quad (3)$$

In the special case of  $r = 0$ , there is no cointegration and we have to proceed with a stationary VAR model in first differences. On the other hand, if  $r = p$ , we can use a stationary VAR model in levels without any error correction terms. In all other cases,  $0 < r < p$ , the series are cointegrated.

The test statistic  $LR$  is based on the squared canonical correlations between the residuals obtained by regressing  $\Delta X_t$  and  $X_{t-1}$  on the lagged differences  $(\Delta X_{t-1}, \dots, \Delta X_{t-k})$  and the deterministic variables  $D_t$ , respectively. These correspond to the eigenvalues  $\lambda_1 \geq \dots \geq \lambda_p$  of the matrix  $S_{01} S_{11}^{-1} S_{01}^\top S_{00}^{-1}$ , with  $S_{00} = \frac{1}{T} R_{0t} R_{0t}^\top$ ,  $S_{01} = \frac{1}{T} R_{0t} R_{1t}^\top$  and  $S_{11} = \frac{1}{T} R_{1t} R_{1t}^\top$ .  $R_{0t}$  are the residuals of regressing  $\Delta X_t$  and  $R_{1t}$  are the residuals of regressing  $X_{t-1}$  on  $(\Delta X_{t-1}, \dots, \Delta X_{t-k})$  and  $D_t$ .

$$LR = -T \sum_{i=r+1}^p \log(1 - \lambda_i). \quad (4)$$

Under the null hypothesis, the test statistic converges in distribution to a function of Brownian motions. The limiting distribution is different according to the specific form of  $D_t$ , see Proposition 8.2 in Lütkepohl (2005). The critical values of the Johansen test are obtained by simulations.

The test has been proved to have issues in small samples, in particular if the dimension of the VAR model,  $p$ , becomes large. This issue is addressed in Johansen (2002). Onatski and Wang (2018), therefore, developed a different asymptotic setting. In particular, they consider the case where  $T$  and  $p$  go to infinity simultaneously such that  $p/T \rightarrow c \in (0, 1]$ . Consider a simplified representation of (1) without lagged differences.

$$\Delta X_t = \Pi X_{t-1} + \Phi D_t + \varepsilon_t \quad (5)$$

Under this asymptotic regime and under the null hypothesis of no cointegration, the empirical distribution function of the eigenvalues of the matrix  $S_{01} S_{11}^{-1} S_{01}^\top S_{00}^{-1}$  converges weakly to the Wachter distribution.

$$F_p(\lambda) \Rightarrow W_c(\lambda) \stackrel{\text{def}}{=} W(\lambda; c/(1+c), 2c/(1+c)) \quad (6)$$

where  $F_p(\lambda) = \frac{1}{p} \sum_{i=1}^p \mathbf{I}(\lambda_i \leq \lambda)$  and  $W(\lambda, \gamma_1, \gamma_2)$  denotes the Wachter distribution function with parameters  $\gamma_1, \gamma_2 \in (0, 1)$  and density  $f_W(\lambda, \gamma_1, \gamma_2) = \frac{1}{2\pi\gamma_1} \frac{\sqrt{(b_+-\lambda)(\lambda-b_-)}}{\lambda(1-\lambda)}$

on  $[b_-, b_+]$  with  $b_{\pm} = \left( \sqrt{\gamma_1(1-\gamma_2)} \pm \sqrt{\gamma_2(1-\gamma_1)} \right)^2$  and atoms of size  $\max(0, 1 - \gamma_2/\gamma_1)$  at zero and  $\max(0, 1 - \frac{1-\gamma_2}{\gamma_1})$  at unity. The rank of cointegration can be determined graphically by comparing the empirical quantiles of the calculated eigenvalues with the theoretical quantiles of the Wachter distribution. Under the null hypothesis of no cointegration the empirical quantiles of eigenvalues should lie close to the theoretical quantiles of the Wachter. Onatski and Wang (2018) suggest to select the cointegration rank by the number of eigenvalues which deviate from the 45 degree line. We show the validity of this approach in a simulation study in Sect. 3.

If the rank of the matrix of cointegration vectors is known, we can estimate cointegration vectors  $\beta$  by reduced rank maximum likelihood estimation, corresponding to the  $r$  largest eigenvalues of the matrix  $S_{01}S_{11}^{-1}S_{01}^{\top}S_{00}^{-1}$ , which we defined in the previous subsection. Without normalization, this estimator is not unique. Therefore, we set the  $j$ -th element in the  $j$ -th cointegration vector to one (Johansen 1995). Then, we can estimate the remaining parameters  $\alpha$  and  $\Gamma = (\Gamma_1 : \dots : \Gamma_k)$  with equation-wise OLS by plugging in the estimator for  $\beta$ , and give their asymptotically normal distribution using standard arguments for stationary processes.

## 2.2 COINtensity VECM

As an extension to the baseline setting, we consider a nonlinear VECM specification. Such models originate from (Granger and Teräsvirta 1993), who introduced the smooth transition error correction model (STECM). A vector version was proposed by Dv et al. (2002). Kristensen and Rahbek (2010) considered the general setting of likelihood-based estimation with nonlinear error correction. Corresponding linearity tests and inference-related issues are discussed in Kristensen and Rahbek (2013). The general setting can be formulated as follows.

$$\Delta X_t = g(\beta^{\top} X_{t-1}; \theta) + \sum_{i=1}^k \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t, \quad (7)$$

where  $g(\cdot)$  is a parametric error correction function with parameter vector  $\theta$ . The error correction function can be nonlinear in the long term stochastic trends as well as in  $\theta$ . In the baseline linear setting,  $g(z; \theta) = \alpha z$  and  $\theta = \text{vec}(\alpha)$ . In the vector version of the STECM we have  $g(z; \theta) = \{\alpha + \tilde{\alpha}\psi(z; \psi)\}$ , where  $\psi(z; \phi)$  is a fixed function satisfying  $|\psi(z; \phi)| = \mathcal{O}(1)$  as  $\|z\| \rightarrow \infty$ , and  $\theta = (\text{vec}(\alpha)^{\top}, \text{vec}(\tilde{\alpha})^{\top}, \text{vec}(\phi)^{\top})^{\top}$ , where  $\text{vec}$  is the vector operator that transforms matrix  $A_{m \times n}$  into an  $(mn \times 1)$  vector by stacking the columns.

The advantage of using nonlinear models is an increased degree of flexibility. However, often this flexibility comes at the expense of worse interpretability and of overfitting the data. We therefore introduce a new class of vector error correction models, which we call *COINtensity* (cointegration intensity) VECM.

$$\Delta X_t = \alpha \beta^\top X_{t-1} \{1 + G(s_t; \gamma)\} + \sum_{i=1}^k \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t, \quad (8)$$

where  $s_t$  is a  $d$ -dimensional vector of transition variables and  $G(\cdot) : \mathbb{R}^d \rightarrow (-1, 1)$  is a parametric function with parameter vector  $\gamma \in \mathbb{R}^d$ . We propose the following parameterisation,  $G(s_t; \gamma) = \tanh(s_t^\top \gamma)$  and  $s_t = \beta^\top X_{t-1}$ , where  $\tanh$  is the sigmoid tangent function. We denote  $G(\cdot)$  as the *COINtensity* (cointegration intensity) function. This function has a universal effect for all cryptocurrencies and measures the intensity of the impact of cointegration.  $G(\cdot)$  takes values in  $(-1, 1)$ . In this model specification, we still have a loading matrix  $\alpha$  which measures currency-specific marginal effects. Please note that our *COINtensity* VECM is a generalization of the baseline model, as model (8) reduces to model (2) if  $\gamma = 0$ .

Our model specification has two advantages. First, it has only a few additional parameters compared to the baseline specification. The overfitting problem of nonlinear error correction models can therefore be contained. Second, the modified model enables us to analyze cointegration and the exposure of cryptocurrencies to long-term equilibrium relationships in a dynamic context.

If the cointegration vectors  $\beta$  are estimated a priori, model parameters can be estimated by quasi maximum likelihood estimation (QMLE). For convenience, we write  $\theta \stackrel{\text{def}}{=} (\text{vec}(\alpha)^\top, \text{vec}(\Gamma)^\top, \gamma^\top)^\top$ . The QMLE,  $\hat{\theta}$  of  $\theta$ , is defined as the minimizer of the following negative log-likelihood criterion,

$$L_T(\theta) = \sum_{t=1}^T \varepsilon_t^\top(\theta) \varepsilon_t(\theta). \quad (9)$$

We split the parameters into two parts and write  $\theta = (\text{vec}(\theta_1)^\top, \theta_2^\top)^\top$ , with  $\theta_1 = (\alpha, \Gamma)^\top$  and  $\theta_2 = \gamma$ . Note that  $\theta_1$  is a  $(r + pk) \times p$  parameter matrix. Further, we define

$$W_t(\theta_2) \stackrel{\text{def}}{=} \left( [\beta^\top X_{t-1} \{1 + \tanh(\theta_2^\top \beta^\top X_{t-1})\}]^\top, \Delta X_{t-1}^\top, \dots, \Delta X_{t-k}^\top \right)^\top, \quad (10)$$

where  $W_t(\theta_2) \in \mathbb{R}^{r+pk}$ . Now, we can rewrite model (8) as follows.

$$\Delta X_t = \theta_1^\top W_t(\theta_2) + \varepsilon_t \quad (11)$$

The profile estimator for  $\theta_1(\theta_2)$  can be obtained by standard OLS.

$$\hat{\theta}_1(\theta_2) = \left\{ \sum_{t=1}^T W_t(\theta_2) W_t^\top(\theta_2) \right\}^{-1} \sum_{t=1}^T W_t(\theta_2) \Delta X_t^\top \quad (12)$$

We proceed by obtaining the corresponding vector of residuals.

$$\hat{\varepsilon}_t(\theta_2) = \Delta X_{t-1} - \hat{\theta}_1^\top W_t(\theta_2) \quad (13)$$

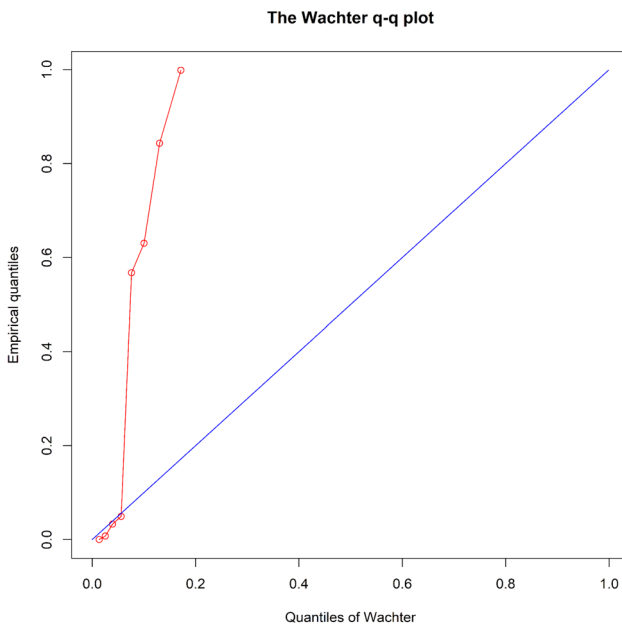
Given the profile estimator, we can estimate  $\theta_2$  by

$$\hat{\theta}_2 = \arg \min_{\theta_2 \in \Theta_2} L_T(\hat{\theta}_1(\theta_2), \theta_2), \quad (14)$$

where  $\Theta_2$  is the parameter space of  $\theta_2$ . The final estimator for  $\theta_1$  can be obtained by plugging (14) into (12). The interpretation of the parameters  $\alpha$ ,  $\beta$  and  $\Gamma$  is almost completely analogous to the linear VECM. In particular, the cointegration vector has the same function as before, governing the long run equilibrium relations. The only difference in the interpretation of parameters is that the loading intensity is now time-varying. Regarding the selection of the cointegration rank  $r$  we cannot make a definitive statement whether the asymptotics of Onatski and Wang (2018) also hold in the *COINtensity* model. However, the procedure seems to work well in practice, as shown in our simulation study.

### 3 Simulation study

In the first part of this simulation study, we examine the validity of the procedure of Onatski and Wang (2018) to test for cointegration. They suggest to determine the cointegration rank graphically by comparing the empirical quantiles of the eigenvalues with the theoretical quantiles of the Wachter distribution. The cointegration rank is chosen according to the number of eigenvalues deviating from the 45° line. Here, we calibrate the numerical example in Liang and Schienle (2019), which is an 8-dimensional VAR(2) process with four unit roots, i.e.  $p = 8$ ,  $r = 4$ ,  $k = 1$ .



**Fig. 1** The Wachter  $Q$ - $Q$  plot for the linear data generating process. The plot shows that the number of eigenvalues deviating from the 45° line is equal to the true cointegration rank,  $r = 4$



$$\Delta X_t = \alpha \beta^\top X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t, \quad (15)$$

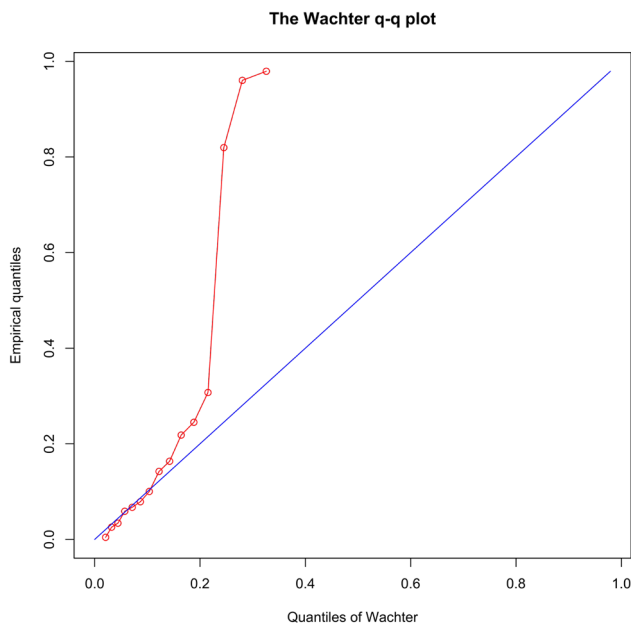
with full-rank matrices  $\alpha$ ,  $\beta$  of dimension  $p \times r$  and iid-distributed  $\varepsilon_t$  generated from  $N(0, I_8)$ . We consider  $T = 200$ , matrices  $\alpha$ ,  $\beta$  and  $\Gamma_1$  are listed in the appendix (setting 1.1). Figure 1 shows that there are exactly four eigenvalues that deviate from the 45 degree line, which also supports the simulation result of Onatski and Wang (2018), while the Johansen test rejects the null hypothesis of a cointegration rank smaller than or equal to four at 5% significance level, implying five cointegration relationship.

As a second setting, we consider the *COINtensity* VECM as our data generating process,

$$\Delta X_t = \alpha \beta^\top X_{t-1} \{1 + \tanh(\gamma^\top \beta^\top X_{t-1})\} + \varepsilon_t. \quad (16)$$

We consider the high dimensional case of  $p = 15$  and  $r = 3$ . The error term is iid and generated from  $N(0, 0.05 \cdot I_{15})$ . We choose a sample size of  $T = 200$ . The parameter matrices are listed in the appendix (setting 1.2). Figure 2 shows exactly three eigenvalues deviating from the 45° line. The testing procedure seems to also work well in the high dimensional and nonlinear setting. So, we apply the Wachter  $Q$ – $Q$  plot to decide the number of cointegration in our large dimensional model.

In the second part of the simulation study, we investigate the finite-sample properties of our estimator for the *COINtensity* VECM. We follow the study design of Kristensen and Rahbek (2010), focusing on the case where  $p = 2$  and



**Fig. 2** The Wachter  $Q$ – $Q$  plot for the nonlinear data generating process. The plot shows that the number of eigenvalues deviating from the 45° line is equal to the true cointegration rank,  $r = 3$

**Table 1** RMSE for QMLE of individual parameters in our *COINtensity* VECM [Qhttps://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

	$T = 250$	$T = 500$	$T = 1000$	$T = 2000$
$\alpha_1$	0.1136	0.0910	0.0736	0.0581
$\alpha_2$	0.0448	0.0322	0.0246	0.0187
$\gamma$	0.3584	0.2887	0.2218	0.1696

the number of cointegration relation is  $r = 1$ . Further, the number of lagged differences entering our model is  $k = 1$ . We consider four different sample sizes,  $T \in \{250, 500, 1000, 2000\}$ . The cointegration vector is assumed to be estimated in advance,  $\beta = (1, -1)^\top$ . The loading parameters are set to  $\alpha_1 = 0.5$  and  $\alpha_2 = -0.5$ . The elements matrix of parameters associated with the lagged first differences are set to  $\Gamma_{jk} = 0.05$  for  $j, k = 1, 2$ . Finally, we set  $\gamma = 0.2$ . For each sample size, we simulate 1000 sample paths of our VECM specification. We evaluate the performance of the estimator by the root mean square error (RMSE). The simulation results can be found in Table 1.

For the individual-specific parameters,  $\alpha_1$  and  $\alpha_2$ , we can observe a good estimation accuracy already in small and moderate samples. As expected, the estimates become more precise with increasing sample size  $T$ . This is also the case for  $\gamma$ , which governs the intensity by which the individual series are affected by deviations from the long-term equilibrium. However, the estimates for  $\gamma$  are not as precise as for the former parameters.

We consider a second setting for the evaluation of our estimation procedure. In this setting, we have the relative large dimensional case of  $p = 8$  and  $r = 3$ . The true values for  $\alpha$ ,  $\beta$  and  $\gamma$  are again listed in the appendix (setting 2.2). The simulation results, based on 1000 Monte Carlo iterations, can be found in Table 2. We report the average Frobenius error of estimating the loading matrix  $\alpha$ ,  $\|\hat{\alpha} - \alpha\|_F$ , as well as the RMSE of  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$ . The result confirms that the estimation error can be effectively reduced with increasing sample size, even if the dimensionality is comparably high.

**Table 2** The first row shows the average Frobenius error,  $\|\hat{\alpha} - \alpha\|_F$ , for the estimated parameter matrix  $\hat{\alpha}$  [Qhttps://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

	$T = 250$	$T = 500$	$T = 1000$	$T = 2000$
$\alpha$	0.3965	0.3440	0.2948	0.2669
$\gamma_1$	0.0634	0.0457	0.0327	0.0298
$\gamma_2$	0.0627	0.0416	0.0466	0.0363
$\gamma_3$	0.0752	0.0714	0.0638	0.0447

The remaining rows show the RMSE for QMLE of  $\hat{\gamma}$  in our *COINtensity* VECM

## 4 Dynamics of cryptocurrencies

### 4.1 Data and descriptive statistics

In the empirical part of the paper, we analyze the joint dynamics of the largest cryptocurrencies. In particular, we are interested in the following set of questions.

- I. Do cointegration relations exist among cryptocurrencies?
- II. Which cryptocurrencies affect and which are affected by long-term equilibrium effects?
- III. How does the impact of the cointegration relationships change in a dynamic setting?

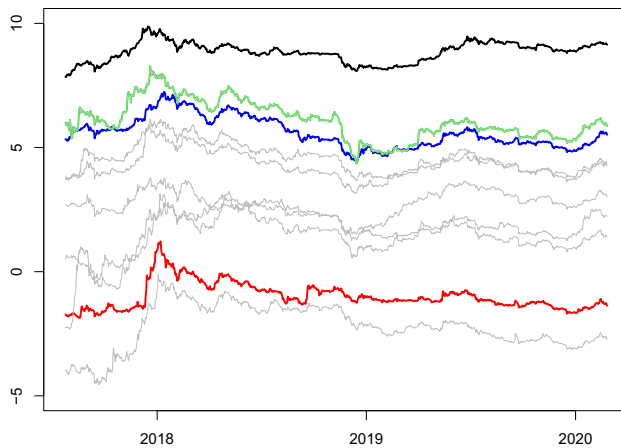
We use daily time-series data of the largest ten cryptocurrencies, which we obtained from Coinmarketcap.com. Since some of the currencies have a very short trading history, we restrict our analysis to those with a time series dating back to at least July 2017. The reason for this decision is to include the boom and the bust of the crypto-bubble at the end of 2017 and start of 2018. To avoid pathological cases, we also remove stable coins such as Tether (USDT). Stable coins are characterized by a fixed exchange rate with the USD and are therefore expected to be stationary in levels. The list of currencies included in our analysis can be found in Table 3. In total, we have 945 daily price observations from July 25, 2017 until February 25, 2020.

The aggregated market capitalization of our sample is around 230 bn USD and captures more than 95% of the total market capitalization of cryptocurrencies. Our analysis, therefore, has a high degree of external validity. By looking at Table 3, it becomes apparent that the crypto market is still dominated by Bitcoin. However, also ETH and XRP occupy a dominant position in the market.

Figure 3 shows the development of the log prices over time. The multivariate time series reveals a strong co-movement of cryptocurrencies. For instance, we can observe a sharp rise in prices for all currencies at the end of 2017, followed by a

**Table 3** List of cryptocurrencies and descriptive statistics. Market capitalization as of February 25, 2020, obtained from Coinmarketcap.com [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

Currency	Symbol	Market Cap (10 <sup>6</sup> USD)	Avg Return (%)	$\sigma$
Bitcoin	BTC	170,370	0.181	0.019
Ethereum	ETH	27,223	0.077	0.020
XRP	XRP	11,087	0.028	0.022
Bitcoin Cash	BCH	6477	0.133	0.047
Litecoin	LTC	4567	0.092	0.025
EOS	EOS	3764	0.107	0.034
Binance Coin	BNB	3164	0.338	0.053
Monero	XMR	1978	0.204	0.031
Stellar	XLM	1320	0.222	0.045
Ethereum Classic	ETC	1076	0.066	0.032



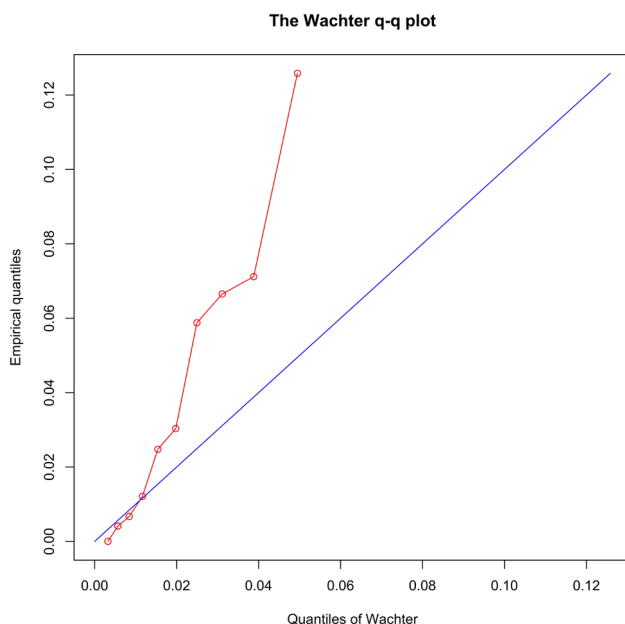
**Fig. 3** Time Series of log prices from July 2017–February 2020. BTC, ETH, XRP, BCH and all others  
[https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

sharp decrease at the beginning of 2018 during burst of the cryptocurrency bubble. This empirical observation suggests a dependence of currencies in levels, not only in first differences. It is thus an essential task to account for cointegration, when analyzing the joint dynamics of cryptocurrencies. Failing to do so would only paint an incomplete picture.

Before any cointegration analysis can be done, one has to assure that all the currencies series are non-stationary and integrated of the same order. Performing the Augmented Dickey-Fuller (ADF) test with a constant and a time trend, the null hypothesis of a unit root cannot be rejected for the individual logged prices at 90% level. The lag length  $k$  for the ADF test has been selected by the Ng and Perron (1995) downtesting procedure starting with a maximum lag of 12. However, the

**Table 4**  $p$  Values of the stationary tests for the level and first difference data

	$X_t$		$\Delta X_t$	
	ADF	KPSS	ADF	KPSS
BTC	0.76	< 0.01	< 0.01	> 0.1
ETH	0.56	< 0.01	< 0.01	> 0.1
XRP	0.21	< 0.01	< 0.01	> 0.1
BCH	0.59	< 0.01	< 0.01	> 0.1
LTC	0.60	< 0.01	< 0.01	> 0.1
EOS	0.41	< 0.01	< 0.01	> 0.1
BNB	0.40	< 0.01	< 0.01	0.04
XMR	0.62	< 0.01	< 0.01	> 0.1
XLM	0.28	< 0.01	< 0.01	0.07
ETC	0.39	< 0.01	< 0.01	> 0.1



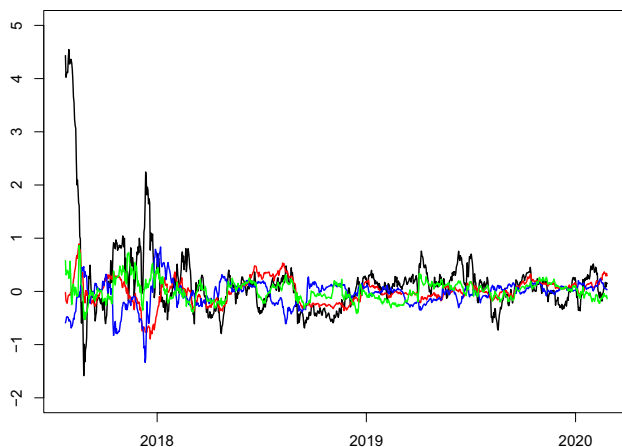
**Fig. 4** Wachter  $Q$ – $Q$  plot to determine the cointegration rank  $r$

**Table 5** Estimated cointegration vectors  $\hat{\beta}$  [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

	BTC	ETH	XRP	BCH	LTC	EOS	BNB	XMR	XLM	ETC
$\beta_1$	1.00	0.00	0.00	0.00	1.98	0.13	− 0.94	− 3.42	0.57	0.70
$\beta_2$	0.00	1.00	0.00	0.00	− 0.28	− 0.27	0.24	− 1.09	0.11	0.31
$\beta_3$	0.00	0.00	1.00	0.00	− 0.97	0.39	0.20	0.54	− 0.76	0.00
$\beta_4$	0.00	0.00	0.00	1.00	0.53	− 0.43	− 0.06	− 1.27	0.37	− 0.42

results of the ADF test are not sensitive to the choice of  $k$  and the null cannot be rejected for any number of lagged terms in each of the series.

In the next step, we apply differences of the time series and compute the ADF test statistic on the differenced data. This time, the null of non-stationarity is rejected for all indices at the 99% level. This suggests that daily returns follow a stationary process. Since the original series must be differenced one time to achieve stationarity, we conclude that the cryptocurrency prices are integrated of order one, such that the vector  $X_t$  is  $I(1)$ . The results of the tests are summarized in Table 4. Having confirmed that all the series are integrated of the same order, this allows to test for cointegration.



**Fig. 5** Time series of the long-term stochastic trends,  $\hat{\beta}_1^\top X_{t-1}$ ,  $\hat{\beta}_2^\top X_{t-1}$ ,  $\hat{\beta}_3^\top X_{t-1}$  and  $\hat{\beta}_4^\top X_{t-1}$  [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

**Table 6** Estimated loading matrix  $\hat{\alpha}$  [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

	$ECT_1$	$ECT_2$	$ECT_3$	$ECT_4$
BTC	0.0045	− 0.0017	− 0.0119	0.0005
ETH	<i>0.0084</i>	− <b>0.0228</b>	0.0098	0.0085
XRP	− 0.0010	− <b>0.0385</b>	− <b>0.0287</b>	0.0164
BCH	0.0044	0.0024	− <b>0.0348</b>	<b>−0.0403</b>
LTC	0.0019	− 0.0144	− 0.0185	− 0.0138
EOS	− 0.0068	− <b>0.0293</b>	− <b>0.0382</b>	<i>0.0438</i>
BNB	<i>0.0308</i>	− 0.0199	0.0142	− 0.0047
XMR	0.0024	0.0026	− 0.0123	<i>0.0268</i>
XTM	0.0067	− <b>0.0276</b>	0.0205	0.0205
ETC	− 0.0009	− 0.0096	0.0007	<i>0.0282</i>

Bold indicates significance of negative coefficients, italic indicates significance of positive coefficients, with significance at 5%, 1% and 0.1% level

## 4.2 Estimation results for linear VECM

In the first step, we determine the cointegration rank graphically using the Wachter  $Q$ – $Q$  plot proposed by Onatski and Wang (2018). As explained in the last section, large deviations of the empirical quantiles of eigenvalues from the theoretical quantiles of the Wachter distribution indicate that the present matrix does not have full rank. We conclude from Fig. 4 that there are four cointegration relations since we can observe four eigenvalues deviating from the 45° line.

Having fixed the cointegration rank, we can proceed with estimating the cointegration vectors. The estimated coefficients can be found in Table 5. To make the estimator unique, we normalize the  $j$ -th entry of the  $j$ -th cointegration vector to 1. Due

to this normalization, we have one vector associated with each of the four largest currencies. For instance, we can observe for  $\beta_1$  that the entry for BTC is one whereas the entries for ETH, XRP and BCH are all close to zero. Based on these estimation results, we plot the time series of our four stochastic trends in Fig. 5. Apart from the beginning of our observation period and apart from the crypto bubble of 2017/2018, we can observe steady and mean-reverting stochastic trends. These observations can be confirmed statistically. Results from the ADF test reject the hypothesis that these trends have a unit root. We can continue to estimate the short-run parameters  $\alpha$  and  $\Gamma$ . In the following, we select the lag order,  $k = 1$ , using the Bayesian information criterion (BIC).

The estimation results of our baseline VECM indicate that cointegration plays an important role for cryptocurrencies. See Table 6 for the estimation of the loading matrix  $\alpha$ . The  $(j,i)$ -th entry of the table shows how currency  $j$  is affected by error correction term  $i$ , where  $ECT_{j,t-1} \stackrel{\text{def}}{=} \hat{\beta}_j^\top X_{t-1}$ . Almost all currencies are significantly affected by at least one stochastic trend, with BTC and LTC being the only exceptions. We additionally test the hypothesis of weak exogeneity to examine whether a given currency is unaffected by all stochastic trends. The null and alternative hypotheses are:

$$H_0 : \alpha_{j,1} = \dots = \alpha_{j,r} = 0 \quad \text{vs.} \quad H_1 : \exists \alpha_{j,k \leq r} \neq 0 \quad (17)$$

The test statistic is constructed as a classical Wald statistic. We reject the null hypothesis for all currencies at a significance level of 0.1%. Cointegration, therefore, has universal effects. The long-run linkages between the indices suggest that cryptocurrency prices are not independent, but predictable using information of others. The results also suggest that investors who seek to diversify their portfolios internationally should be aware that the ten cryptocurrency prices in the system follow a common stochastic trend. This means that these markets generate similar returns in the long-run. Therefore, diversification across the markets is limited and investors should include other markets with lower correlation to hedge their risk.

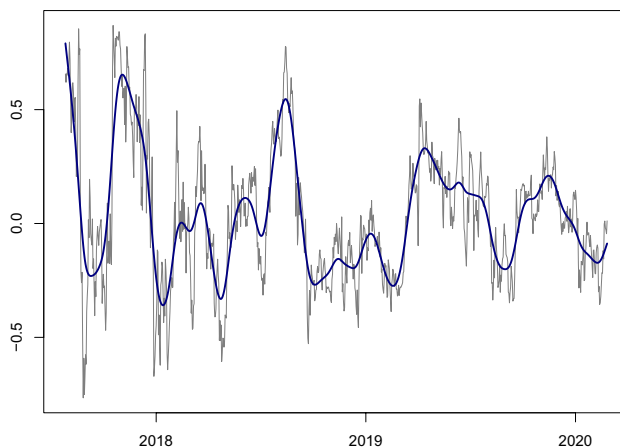
In the first error correction term, ETH and BNB do not tend to return to the long-run equilibrium as the coefficient on the error term is positive. In the second one, ETH, XRP, EOS and XLM all have the predicted negative sign, which indicates that the disequilibrium given in the error correction term will be reduced period by period. However, the size of the estimates differs widely and is quite small compared to the other short-term adjustment parameters. These results suggest that distortions in the long-run equilibrium will be corrected slowly and unevenly among the ten cryptocurrencies. In the third one, XRP, BCH and EOS carry the burden of adjustment to return to the long-run relationship. In the fourth one, EOS, XMR, ETC are the leaders in the system and that BCH carries the burden of adjustment to return to the long-run relationship.

The estimation results for the lagged differences can be found in Table 7. Compared to the estimated coefficients for the error correction terms, the lagged differences seem to be less important. Some currencies, such as BCH and BNB, have highly significant coefficients associated with their own lagged value. Another interesting observation is that BTC and BCH both depend on each other negatively.

**Table 7** Estimated coefficient matrix  $\hat{\Gamma}$  [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

	BTC	ETH	XRP	BCH	LTC	EOS	BNB	XMR	XLM	ETC
BTC	0.08	− 0.08	− 0.03	<b>− 0.05</b>	− 0.06	0.07	0.00	0.06	0.02	− 0.04
ETH	− 0.07	0.05	− 0.07	0.00	0.07	− 0.02	0.01	0.02	0.02	− 0.08
XRP	<b>− 0.17</b>	0.06	<i>0.11</i>	− 0.03	0.03	0.01	0.05	0.06	− 0.08	<b>− 0.12</b>
BCH	<b>− 0.28</b>	0.13	− 0.09	<i>0.19</i>	− 0.05	− 0.00	0.08	0.06	0.01	<b>− 0.14</b>
LTC	0.01	− 0.11	− 0.03	0.02	0.09	− 0.05	0.02	0.03	− 0.01	− 0.02
EOS	− 0.07	− 0.06	− 0.07	− 0.03	0.11	0.00	<i>0.08</i>	0.02	0.01	− 0.01
BNB	0.15	0.01	0.02	0.03	<b>− 0.18</b>	0.01	<i>0.18</i>	− 0.04	<b>− 0.13</b>	− 0.06
XMR	− 0.05	− 0.01	− 0.07	− 0.01	0.02	0.03	<i>0.07</i>	− 0.04	− 0.01	− 0.05
XLM	0.04	− 0.08	− 0.04	− 0.07	0.09	0.03	0.00	− 0.03	<i>0.13</i>	− 0.11
ETC	0.05	− 0.01	− 0.09	0.05	− 0.00	0.05	0.01	− 0.08	0.02	− 0.07

Bold indicates significance of negative coefficients, italic indicates significance of positive coefficients, with significance at 5%, 1% and 0.1% level

**Fig. 6** Time series of cointegration intensity  $G(\hat{\beta}^T X_{t-1}; \hat{\gamma})$  (grey) and spline interpolation (blue) [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation) (colour figure online)

### 4.3 Estimation results for *COINtensity* VECM

All the previous results are obtained in the baseline linear VECM setting. For a dynamic analysis, we henceforth rely on our *COINtensity* VECM. We estimate the model by the profile likelihood estimation framework introduced in Sect. 2.2. In the first step, we estimate the cointegration vectors  $\beta$  as before. In practice, we then estimate the nonlinear part of the model by random parameter search. We assume that the parameter vector  $\theta_2 = \gamma$  lies in  $\Theta_2 = [-1, 1]^r$ . The candidate parameters are generated from the  $r$ -dimensional uniform distribution in the same range. Our number of simulations is 10,000.



The time series of the estimated *COINtensity* function,  $G(\hat{\beta}^\top X_{t-1}; \hat{\gamma})$ , is visualized in Fig. 6. We can observe a time-varying pattern of the intensity by which cryptocurrencies are affected by long run equilibrium effects. Prior to the building of the bubble at the end of 2017, cointegration intensity was low with values below zero. The following increase goes along with the strong increase in prices across all cryptocurrencies in the last quarter of the same year. The subsequent months can be characterized by a highly volatile cointegration intensity. Recently, from the second half of 2018, we can observe a period of stabilization with only a few values exceeding the  $-0.5$  and  $0.5$  thresholds. We conclude that nonlinearity was more prevalent in the turbulent period of the cryptocurrency bubble.

We also evaluate the out-of-sample predictive power of the *COINtensity* VECM compared to the linear baseline model. Even if prediction is not the main purpose of this research, it can still provide insight into the usefulness of the nonlinear specification. For the out-of-sample analysis, we consider the period from February 26 to October 13, 2020. The results can be found in Table 8. We report the root mean square error (RMSE) of prediction for both models and for each cryptocurrency separately. It becomes evident that the *COINtensity* specification outperforms the linear model. For nine out of ten currencies the RMSE is lower. We apply the test of Diebold and Mariano (2002) to test whether this outperformance is significant. We find that only for one currency (BNB) the forecast is significantly better.

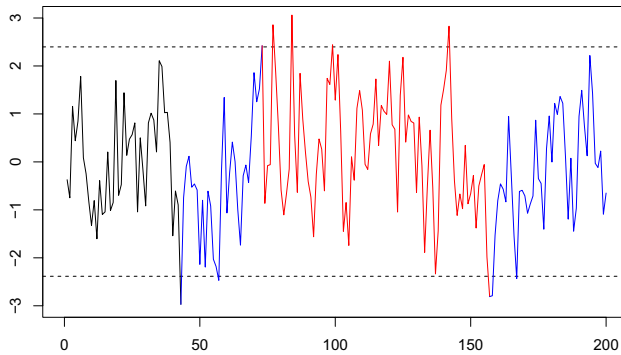
## 5 A simple statistical arbitrage trading strategy

In this section, we apply a simple cointegration-based trading strategy for cryptocurrencies. We use the same data as in the previous section. Under the assumption of mean reversion of the long term stochastic trends, a large deviation from the equilibrium relationships should lead to profitable investment opportunities. In the following, we define the cointegration spreads. For each cointegration relationship,  $j = 1, \dots, r$ , we have

**Table 8** Out-of-sample predictive performance in terms of root mean squared error (RMSE) for the linear VECM and *COINtensity* VECM specification

	Linear	<i>COINtensity</i>	DM-test
BTC	0.0451	<b>0.0450</b>	0.1883
ETH	0.0614	<b>0.0579</b>	0.8132
XRP	0.0488	<b>0.0446</b>	0.8226
BCH	0.0580	<b>0.0579</b>	0.8991
LTC	<b>0.0504</b>	0.0506	0.8904
EOS	0.0575	<b>0.0541</b>	0.8004
BNB	0.0675	<b>0.0589</b>	<b>0.0413</b>
XMR	0.0538	<b>0.0527</b>	0.1571
XLM	0.0559	<b>0.0524</b>	0.7931
ETC	0.0539	<b>0.0539</b>	0.6541

Additionally,  $p$  values from the Diebold-Mariano test are reported



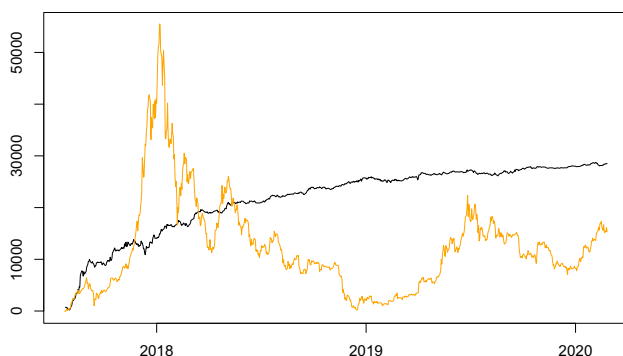
**Fig. 7** Visualization of the statistical arbitrage trading strategy for simulated data. Neutral position, short position and long position

**Table 9** In-sample performance statistics for different threshold levels [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

Threshold	$\pm\sigma_j$	$\pm 1.5\sigma_j$	$\pm 2\sigma_j$	CRIX
Number of Trades	40	21	12	–
Net Profits	28,474	24,876	21,247	15,330
Maximal Drawdown	3078	2989	2893	55,297
Annual Sharpe Ratio	2.24	1.99	1.77	0.22

$$\begin{aligned}
 S_{j,t} &= \beta_j^\top X_t \\
 &= \beta_{j,1}X_{1,t} + \dots + \beta_{j,p}X_{p,t}.
 \end{aligned}
 \tag{18}$$

These spreads are nothing more than weighted averages of log prices of cryptocurrencies, where the weighting is done by the cointegration vectors. If the spread exceeds an upper threshold, we enter a short position, if the spread goes below the lower threshold, we enter a long position. The reasoning behind the strategy is very intuitive. A large positive spread is a signal that the portfolio is overpriced and it is profitable to sell it. On the other hand, if we encounter a large negative spread, the portfolio is underpriced and we should buy it. The logic of the trading strategy is visualized in Fig. 7. We choose three different threshold levels,  $\tau \in (\pm\sigma_j, \pm 1.5\sigma_j, \pm 2\sigma_j)$ , which are chosen to be symmetric around the long term mean of the stochastic trend and  $\sigma_j$  is the estimated standard deviation. This investment decision is repeated for each estimated cointegration relationship and for each trading day. So each day, we have to make a decision to either buy, sell or hold our positions. The trading strategy follows (Leung and Nguyen 2019), who consider a similar statistical arbitrage strategy. However, our strategy differs in two aspects. First, Leung and Nguyen (2019) use the approach of Engle and Granger (1987) to estimate the cointegration vector and second, our paper utilizes  $r$  cointegration relations while their paper is restricted to a single one. We backtest our strategy and compare the performance to the cryptocurrency index CRIX (Trimborn and Härdle 2018).



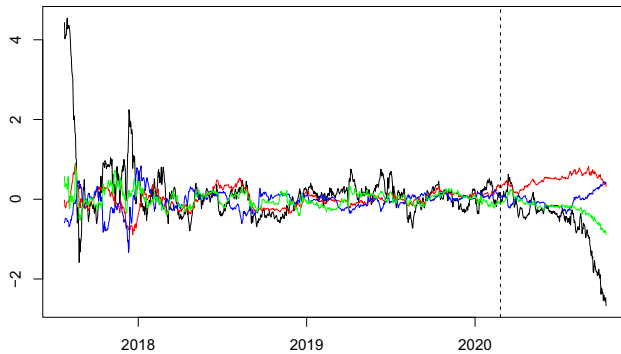
**Fig. 8** In-sample performance of the trading strategy with thresholds  $\tau = \pm 1.5\sigma$  (black) vs. CRIX (yellow) (colour figure online)

Table 9 summarizes the performance of our trading strategy for different threshold levels and compares it to the performance of the CRIX. The number of trades is decreasing with an increasing threshold level. For each of the candidate thresholds, we can make substantial profits. The optimal threshold in our analysis is  $\tau = \pm\sigma_j$ . It has the highest net profits, the largest Sharpe ratio and a similar maximal drawdown to the other threshold levels. While the net profits of the benchmark index portfolio (CRIX) are comparable to our arbitrage strategy, the risk is significantly higher. The maximal drawdown is more than ten times as large as for the optimal strategy. Also the Sharpe ratio, which relates expected returns to the standard deviation, is clearly smaller. Figure 8 visualizes the time series of the cumulative returns of our trading strategy and of the CRIX. As expected of an arbitrage strategy, there is almost no dependence of the cumulative returns to the market. An interesting observation is that the only substantial losses are made during the height of the crypto bubble at the end of 2017. The gains and losses are very volatile in this period. From the middle of 2018 until the beginning of 2020 we can observe small but steady profits.

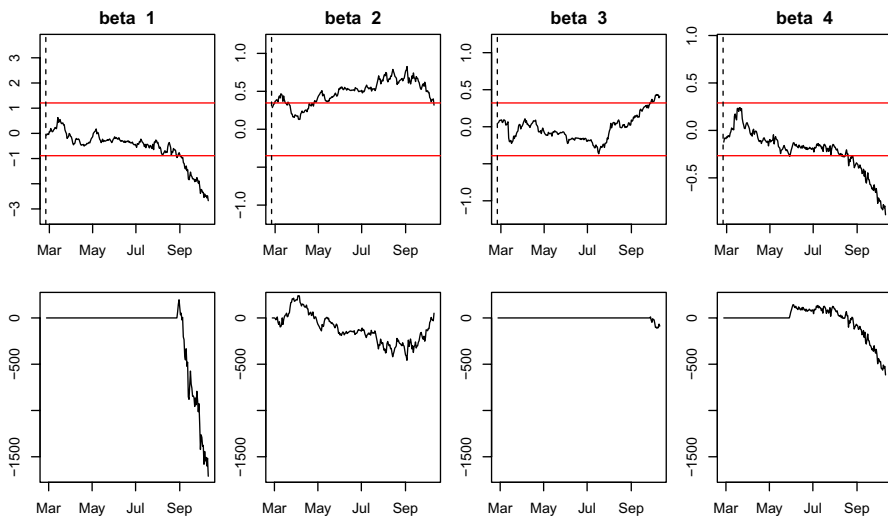
While the backtesting results show a great performance of our trading strategy, a word of caution is needed. First, backtesting is an in-sample evaluation with limited external validity. There is no guarantee that long-term relationship will hold in the future, which is an implicit assumption in our cointegration analysis. This problem is particularly severe in the case of cryptocurrencies due to their very short history. Another caveat is that we assume perfect markets. In reality, investors face short selling restrictions and transaction costs, even if some exchanges as Bitfinex allow for short selling.

**Table 10** Out-of-sample performance statistics for different threshold levels [https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics\\_Simulation](https://github.com/QuantLet/CryptoDynamics/tree/master/CryptoDynamics_Simulation)

Threshold	$\pm\sigma_j$	$\pm 1.5\sigma_j$	$\pm 2\sigma_j$	CRIX
Number of Trades	5	4	3	–
Net Profits	– 2134	– 2363	– 1660	10,692
Maximal Drawdown	675	571	543	10,186
Annual Sharpe Ratio	– 1.56	– 2.20	– 1.73	0.87



**Fig. 9** Time series of long-run stochastic trends. Dashed vertical line indicates the begin of the out-of-sample period



**Fig. 10** Out-of-sample analysis of long run equilibrium relationships and profits from corresponding trading strategies. The red horizontal lines visualize the thresholds  $\tau = \pm 1.5\sigma$

To further evaluate our trading strategy, we take a look at its out-of-sample performance. We consider the same set of cryptocurrencies in a time period from February 26 to October 13, 2020. The results are reported in Table 10. It becomes evident that for none of the threshold levels we can make a positive profit. The reason for this is the divergence of the cointegration relation in the out-of-sample period, as shown in Fig. 9. To analyze the trading performance in more detail, we consider each cointegration relation separately in Fig. 10 for a threshold of  $\pm 1.5\sigma$ . Most of the losses originate from the first long-run relationship, which begins to deviate from its mean at the end of August. A similar phenomenon can be observed for the fourth cointegration relation. The remaining two stochastic trends do not diverge significantly, leading to a close-to-zero profit. It will be interesting to observe in the

future whether our estimated cointegration relations are indeed mean-reversing, i.e. whether they will return to their equilibria, as predicted by our model. This would also provide more information on the profitability of our trading strategy. While the out-of-sample analysis can be seen as evidence against the possibility of statistical arbitrage, it is too soon to tell whether the long-run relations disappeared completely.

## 6 Conclusion

This paper examined the joint behavior of cryptocurrencies in a non-stationary setting. We were in particular interested in three questions.

- I. Do cointegration relations exist among cryptocurrencies?
- II. Which cryptocurrencies affect and which are affected by long-term equilibrium effects?
- III. How does the impact of the cointegration relationships change in a dynamic setting?

To address problem (I) and (II), we tested for cointegration using the approach of Onatski and Wang (2018) and estimated a linear VECM. We found that our sample of currencies are indeed cointegrated with rank four. By testing for weak exogeneity, we were able to show that all cryptocurrencies are significantly affected by long term stochastic trends. To address problem (III), we proposed a new nonlinear VECM specification, which we call *COINtensity* VECM. The model has a good interpretability without the need of having to estimate many new parameters. The results of our dynamic VECM show a time-varying dependence of cryptocurrencies on deviations from long run equilibria. We find that the nonlinearity of error correction is stronger during the time of the cryptocurrency bubble, compared to a later time period.

Finally, we utilized the estimated cointegration relationships to construct a simple statistical arbitrage trading strategy, extending the one proposed in Leung and Nguyen (2019). Our strategy shows a great in-sample performance, beating the industry benchmark CRIX in terms of net profits, Sharpe ratio and maximal draw-down. A look at the out-of-sample performance takes a more cautious perspective. In particular, the trading strategy can only be successful if the cointegration equilibrium relations hold in the long-run.

## Appendix: simulation design

Setting 1.1: Baseline VECM specification:

$$\Delta X_t = \alpha \beta^\top X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t,$$

with parameter matrices

$$\alpha = \begin{pmatrix} -1.47 & -1.3 & 0 & -1.26 \\ 0 & 0.97 & 0 & 0 \\ 0 & 0 & -0.74 & 0 \\ -1.19 & 0.85 & 0 & 0 \\ 0.55 & 0.78 & -1 & -1.37 \\ 0.8 & 0.75 & 0 & 0 \\ 0 & -0.74 & -1.26 & 0.78 \\ 0 & -1.4 & 0 & 0 \end{pmatrix},$$

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.8 \\ 0 & 0 & -1.29 & 1.49 \\ -0.87 & 0 & -0.53 & -0.82 \\ 1.45 & 1.48 & 0.9 & -0.69 \end{pmatrix},$$

$$\Gamma_1 = \text{diag}\{0, 0.7979, 0, 0.7932, 0, 0.5377, 0, 0.7227\}.$$

Setting 1.2: *COINtensity* VECM specification:

$$\Delta X_t = \alpha \beta^\top X_{t-1} \{1 + \tanh(\gamma^\top \beta^\top X_{t-1})\} + \varepsilon_t.$$

$$\alpha = \begin{pmatrix} -0.10 & 0.47 & -0.12 \\ 0.10 & -0.20 & -0.12 \\ 0.67 & 0.20 & -0.30 \\ 0.37 & 0.34 & 0.30 \\ -0.71 & 0.49 & 0.28 \\ -0.62 & -0.45 & 0.10 \\ -1.00 & 0.03 & 0.32 \\ -0.82 & 0.72 & 0.88 \\ -0.77 & 0.50 & -0.70 \\ -0.08 & 0.76 & 0.24 \\ -0.01 & -0.82 & -0.05 \\ -0.62 & 0.58 & 0.68 \\ 0.27 & 0.91 & -0.28 \\ 0.95 & -0.44 & 0.03 \\ -0.58 & 0.91 & 0.11 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\gamma = (0.2, 0.3, -0.4)^\top.$$

Setting 2.2:

$$\alpha = \begin{pmatrix} -0.10 & 0.27 & 0.37 \\ 0.10 & -0.20 & -0.23 \\ -0.14 & 0.20 & -0.30 \\ -0.08 & 0.10 & 0.30 \\ 0.04 & 0.08 & 0.16 \\ 0.24 & -0.26 & -0.00 \\ -0.18 & -0.18 & 0.13 \\ 0.24 & -0.19 & 0.30 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\gamma = (0.2, 0.3, -0.4)^T.$$

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Corbet, S., Meegan, A., Larkin, C., Lucey, B., & Yarovaya, L. (2018). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, 165, 28–34.
- Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 20(1), 134–144.
- Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1), 119–134.
- Dv, Dijk, Teräsvirta, T., & Franses, P. H. (2002). Smooth transition autoregressive models-a survey of recent developments. *Econometric Reviews*, 21(1), 1–47.
- Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 55, 251–276.
- Fan, J., & Yao, Q. (2008). *Nonlinear time series: Nonparametric and parametric methods*. Berlin: Springer.
- Granger, C., & Teräsvirta, T. (1993). *Modelling nonlinear economic relationships*. New York: Oxford University Press.
- Guo, L., Tao, Y., Härdle, W. K. (2018) A dynamic network perspective on the latent group structure of cryptocurrencies. arXiv preprint. [arXiv:1802.03708](https://arxiv.org/abs/1802.03708).
- Härdle, W. K., Harvey, C. R., & Reule, R. C. G. (2020). Understanding Cryptocurrencies\*. *Journal of Financial Econometrics*, 18(2), 181–208. <https://doi.org/10.1093/jfinec/nbz033>.
- Ji, Q., Bouri, E., Lau, C. K. M., & Roubaud, D. (2019). Dynamic connectedness and integration in cryptocurrency markets. *International Review of Financial Analysis*, 63, 257–272.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2–3), 231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society*, 59, 1551–1580.
- Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press.

- Johansen, S. (2002). A small sample correction for the test of cointegrating rank in the vector autoregressive model. *Econometrica*, 70(5), 1929–1961.
- Kristensen, D., & Rahbek, A. (2010). Likelihood-based inference for cointegration with nonlinear error-correction. *Journal of Econometrics*, 158(1), 78–94.
- Kristensen, D., & Rahbek, A. (2013). Testing and inference in nonlinear cointegrating vector error correction models. *Econometric Theory*, 29(6), 1238–1288.
- Kuan, C. M., & White, H. (1994). Artificial neural networks: An econometric perspective. *Econometric Reviews*, 13(1), 1–91.
- Lee, T. H., White, H., & Granger, C. W. (1993). Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests. *Journal of Econometrics*, 56(3), 269–290.
- Leung, T., & Nguyen, H. (2019). Constructing cointegrated cryptocurrency portfolios for statistical arbitrage. *Studies in Economics and Finance*, 36(4), 581–599.
- Liang, C., & Schienle, M. (2019). Determination of vector error correction models in high dimensions. *Journal of Econometrics*, 208(2), 418–441.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin: Springer.
- Luukkonen, R., Saikkonen, P., & Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, 75(3), 491–499.
- Ng, S., & Perron, P. (1995). Unit root tests in arma models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90(429), 268–281.
- Onatski, A., & Wang, C. (2018). Alternative asymptotics for cointegration tests in large vars. *Econometrica*, 86(4), 1465–1478.
- Shahzad, S. J. H., Bouri, E., Roubaud, D., Kristoufek, L., & Lucey, B. (2019). Is bitcoin a better safe-haven investment than gold and commodities? *International Review of Financial Analysis*, 63, 322–330.
- Sovbetov, Y. (2018). Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero. *Journal of Economics and Financial Analysis*, 2(2), 1–27.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425), 208–218.
- Trimborn, S., & Härdle, W. K. (2018). Crix an index for cryptocurrencies. *Journal of Empirical Finance*, 49, 107–122.
- Yi, S., Xu, Z., & Wang, G. J. (2018). Volatility connectedness in the cryptocurrency market: Is bitcoin a dominant cryptocurrency? *International Review of Financial Analysis*, 60, 98–114.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.