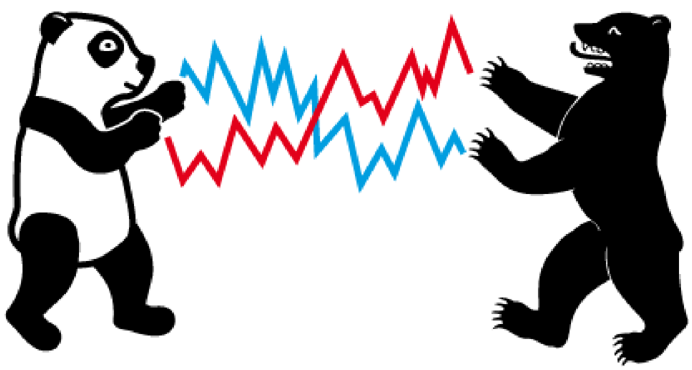
IRTG 1792 Discussion Paper 2020-012

On Cointegration and Cryptocurrency Dynamics

Georg Keilbar [*]
Yanfen Zhang [*2] [*]



[*] Humboldt-Universität zu Berlin, Germany
[*2] Xiamen University, China

This research was supported by the Deutsche Forschungsgesellschaft through the International Research Training Group 1792 "High Dimensional Nonstationary Time Series".

http://irtg1792.hu-berlin.de
ISSN 2568-5619

International Research Training Group 1792

Noname manuscript No.
(will be inserted by the editor)

On Cointegration and Cryptocurrency Dynamics

Georg Keilbar · Yanfen Zhang

Received: date / Accepted: date

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. A simple statistical arbitrage trading strategy is proposed showing a great in-sample performance.

Keywords Cointegration · VECM · Nonstationarity · Cryptocurrencies

Financial support from the Deutsche Forschungsgemeinschaft via the IRTG 1792 "High Dimensional Nonstationary Time Series", Humboldt-Universität zu Berlin, is gratefully acknowledged.

Georg Keilbar
IRTG 1792 "High Dimensional Nonstationary Time Series", School of Business and Economics, Humboldt-University of Berlin, Unter den Linden 6, 10099 Berlin, Germany
E-mail: georg.keilbar@hu-berlin.de

Yanfen Zhang
Wang Yanan Institute for Studies in Economics (WISE) and School of Economics, Xiamen University, Xiamen 361005, China.
IRTG 1792 "High Dimensional Nonstationary Time Series", School of Business and Economics, Humboldt-University of Berlin, Unter den Linden 6, 10099 Berlin, Germany
E-mail: zhanyanf@hu-berlin.de

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. (2019) 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 Yılmaz (2014). Sovbetov (2018) analyzesd 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. A crucial task is to select the number of those equlibria, 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. To analyze the cointegration of cryptocurrencies in a dynamic setting, we propose a novel nonlinear VECM specification, which we call *COINtensity* (cointegration intensity) VECM.

Our empirical results suggest that cointegration plays a crucial role for 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). A backtest of our trading strategy indicates that trading on large deviations from the long-run equlibria can be profitable.

The contributions of this paper are two-fold. 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 section 3. In section 4, we apply our methodology to a system of the largest ten cryptocurrencies. Section 5 introduces a simple cointegration-based trading strategy and section 6 concludes.

All codes of this paper are available on 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).

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

where D_t are deterministic variables and ε_t are zero-mean, independent error terms. We assume that each univariate time series is integrated of order one, $X_{it} \sim I(1), i = 1, \ldots, p$. 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.

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

where β is a $p \times r$ matrix of cointegration vectors and α is the $p \times r$ loading matrix. The order of cointegration is characterized by the rank r of β. Γ_i, $i = 1, \ldots, k-1$, are $p \times p$ parameter matrices associated with the impact of lagged values of ΔX_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.

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

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

The test statistic is based on the squared canonical correlations between the residuals obtained by regressing ΔX_t and X_{t-1} on the lagged differences and the exogenous variables, respectively. These correspond to the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_p$ 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 $S_{11} = \frac{1}{T}R_{1t}R_{01}^\top$. R_{0t} are the residuals of regressing ΔX_t and R_{1t} are the residuals of regressing X_{t-1}.

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

Under the null hypothesis, the test statistic converges in distribution to a function of Brownian motions.

$$LR \xrightarrow{\mathcal{L}} \operatorname{tr}\left\{ \left(\int_0^1 W dW^\top \right)^\top \left(\int_0^1 W W^\top ds \right)^{-1} \left(\int_0^1 W dW^\top \right) \right\}, \tag{5}$$

where W is a $(p-r)$-dimensional Brownian motion.

The Johansen test of cointegration has proved to have issues in small samples, in particular if the dimension of the VAR model, p, becomes large. This issue is adressed 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 at a constant rate $c \stackrel{\text{def}}{=} \frac{p}{T} \in [0, 1]$. Consider a simplified representation of (1) without lagged differences.

$$\Delta X_t = \Pi X_{t-1} + \Phi D_t + \varepsilon_t \tag{6}$$

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.

$$F_p(\lambda) \Rightarrow W_c(\lambda) \stackrel{\text{def}}{=} W\left(\lambda; c/(1+c), 2c/(1+c)\right) \tag{7}$$

where $F_p(\lambda) = \frac{1}{p}\mathbf{I}(\lambda_i \leq \lambda)$ 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 section 3.

If the rank of the matrix of cointegration vectors is known, we can estimate the parameters of the VECM by reduced rank maximum likelihood estimation (Johansen (1995)). We differentiate between the long-run parameters, β, and the short-run parameters, α and Γ. In order to derive estimators for these parameters, it is convenient to rewrite model (2) in matrix notation.

$$\Delta X = \alpha\beta^\top X_{-1} + \Gamma Z + \varepsilon \tag{8}$$

with $\Delta X = (\Delta X_1, \dots, \Delta X_T)$, $X_{-1} = (X_0, \dots, X_{T-1})$ and $Z = (Z_0, \dots, Z_{T-1})$ with $Z_{t-1} = (\Delta X_{t-1}, \dots, \Delta X_{t-k})^\top$. First, we estimate $\widehat{\beta}$ by the eigenvectors 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. We can now estimate the remaining parameters with equation-wise OLS by plugging in the estimator for β.

$$[\hat{\alpha} : \hat{\Gamma}] = [\Delta X X_{-1}^\top \beta : \Delta X Z^\top] \begin{bmatrix} \beta^\top X_{-1} X_{-1}^\top \beta & \beta^\top X_{-1} Z^\top \\ Z X_{-1}^\top \beta & Z Z^\top \end{bmatrix}^\top \tag{9}$$

Using standard arguments for stationary processes, the estimator's distribution is asymptotically normal.

$$\sqrt{T}\mathrm{vec}([\hat{\alpha} : \hat{\Gamma}] - [\alpha : \Gamma]) \xrightarrow{\mathcal{L}} N(0, \Sigma_{\alpha,\Gamma}), \tag{10}$$

where

$$\Sigma_{\alpha,\Gamma} = \begin{bmatrix} \beta^\top X_{-1} X_{-1}^\top \beta & \beta^\top X_{-1} Z^\top \\ Z X_{-1}^\top \beta & Z Z^\top \end{bmatrix}^\top \otimes \Sigma_\varepsilon \tag{11}$$

This enables us to set up and interpret t-tests in the usual way because they have a standard normal limiting distribution under our assumptions. Also, Wald tests and the corresponding F-tests of linear restrictions on the parameters have the usual asymptotic χ^2- or approximate F-distributions that are obtained 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 Dijk 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.

$$\Delta X_t = g\left(\beta^\top X_{t-1}; \theta\right) + \sum_{i=1}^{k} \Gamma_i \Delta X_{t-i} + D_t + \varepsilon_t \tag{12}$$

where $g(\cdot)$ is a parametric error correction function with parameter vector θ. The error correction function can be nonlinear in the long term stochastic trends as well as in θ. In the baseline linear setting, $g(z;\theta) = \alpha z$ and $\theta = \mathrm{vec}(\alpha)$. In the vector version of the STECM we have $g(z;\theta) = \{\alpha + \widetilde{\alpha}\psi(z;\psi)\}$, where $\psi(z;\phi)$ is a fixed function satisfying $|\psi(z;\phi)| = \mathcal{O}(1)$ as $\|z\| \to \infty$, and $\theta = (\mathrm{vec}(\alpha)^\top, \mathrm{vec}(\widetilde{\alpha})^\top, \mathrm{vec}(\phi)^\top)^\top$.

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.

$$\Delta X_t = \alpha\beta^\top X_{t-1}\left\{1 + G\left(\gamma, s_t\right)\right\} + \sum_{i=1}^{k} \Gamma_i \Delta X_{t-i} + D_t + \varepsilon_t \tag{13}$$

where s_t is a d-dimensional vector of transition variables and $G(\cdot) : \mathbb{R}^d \to (-1,1)$ is a parametric function with parameter vector $\gamma \in \mathbb{R}^d$. We propose the following parameterisation, $G(\cdot) = \tanh(s_t^\top \gamma)$ and $s_t = \beta^\top X_{t-1}$. 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 α which measures currency-specific marginal effects. Please note that our *COINtensity* VECM is a generalization of the baseline model, as model (13) 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, addressing question III. of this research.

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

$$L_T(\theta) = \sum_{t=1}^{T} \varepsilon_t^\top(\theta)\varepsilon_t(\theta) \tag{14}$$

We split the parameter vector into two parts and write $\theta = (\theta_1^\top, \theta_2^\top)^\top$, with $\theta_1 = (\mathrm{vec}(\alpha), \mathrm{vec}\,\Gamma)^\top$ and $\theta_2 = \mathrm{vec}(\gamma)$. Further, we define

$$W_t(\theta_2) \stackrel{\text{def}}{=} \left(\left[\beta^\top X_{t-1}\left\{1 + \tanh\left(\theta_2\beta^\top X_{t-1}\right)\right\}\right]^\top, \Delta X_{t-1}^\top, \ldots, \Delta X_{t-k}\right)^\top, \tag{15}$$

where $W_t(\theta_2) \in \mathbb{R}^{r+pk}$. Now, we can rewrite model (13) as follows.

$$\Delta X_t = \theta_1^\top W_t(\theta_2) + \varepsilon_t \tag{16}$$

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

$$\widehat{\theta}_1(\theta_2) = \left\{ \sum_{t=1}^{T} W_t(\theta_2) W_t^\top(\theta_2) \right\}^{-1} \sum_{t=1}^{T} W_t(\theta_2) \varDelta X_{t-1}^\top \tag{17}$$

We proceed by obtaining the corresponding vector of residuals.

$$\widehat{\varepsilon}_t(\theta_2) = \varDelta X_{t-1} - \widehat{\theta}_1 W_t(\theta) \tag{18}$$

Given the profile estimator, we can estimate θ_2 by

$$\widehat{\theta}_2 = \arg \min_{\theta_2 \in \varTheta_2} L_T(\theta_1(\theta_2), \theta_2), \tag{19}$$

where $\varTheta_2$ is the parameter space of θ_2. The final estimator for θ_1 can be obtained by plugging (19) into (17).

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 eigenvalues of the Wachter distribution. The cointegration rank is chosen according to the number of eigenvalues deviating from the 45 degree 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 $p = 8$, $r = 4$, $k = 1$.

$$\varDelta X_t = \alpha \beta^\top X_{t-1} + \varGamma_1 \varDelta X_{t-1} + \varepsilon_t, \tag{20}$$

with full-rank matrices α, β of dimension $p \times r$ and iid-distributed ϵ_t generated from $N(0, I_8)$. We consider $T = 200$, matrices α, β and $\varGamma_1$ are listed in the appendix.

Figure 1 shows that there are exactly 4 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. So, we apply the Wachter Q-Q plot to decide the number of cointegration in our large dimensional model.

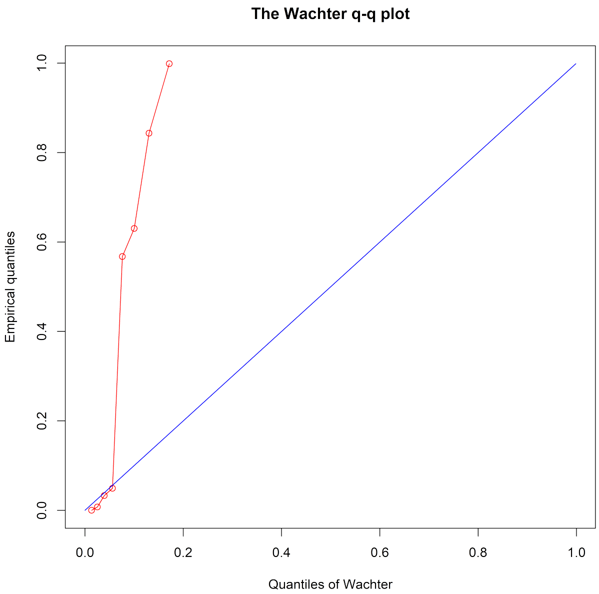


Fig. 1: The Wachter Q-Q plot shows that the number of eigenvalues deviating from the 45 degree line is equal to the true cointegration rank, $r = 4$.

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). In particular, we focus on the case where $p = 2$ and 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 known, $\beta = (1, -1)^\top$. The loading parameters are set to $\alpha_1 = 0.2$ and $\alpha_2 = -0.2$. The elements matrix of parameters associated with the lagged first differences are set to $\Gamma_{jk} = 0.05$ for $j, k = 1, 2$. For each sample size, we simulate 1000 sample paths of our VECM specification. Finally, we set $\gamma = 0.2$. We evaluate the performance of the estimator by the root mean square error (RMSE). The simulation results can be found in Table 1.

	$T = 250$	$T = 500$	$T = 1000$	$T = 2000$
α_1	0.0624	0.0498	0.0403	0.0314
α_2	0.0622	0.0501	0.0400	0.0314
γ	0.4034	0.3476	0.2838	0.2202

Table 1: RMSE for QMLE of individual parameters in our *COINtensity* VECM.

CryptoDynamics_Simulation

For the individual-specific parameters, α_1 and α_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 γ, which governs the intensity by which the individual series are affected by deviations from the long-term equlibrium. However, the estimates for γ are not as precise as for the former parameters.

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 2. In total, we have 945 daily price observations from July 25, 2017 until February 25, 2020.

Currency	Symbol	Market Cap (10^6 USD)	Avg Return (%)	σ
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	6,477	0.133	0.047
Litecoin	LTC	4,567	0.092	0.025
EOS	EOS	3,764	0.107	0.034
Binance Coin	BNB	3,164	0.338	0.053
Tezos	XTZ	1,978	0.204	0.031
Stellar	XLM	1,320	0.222	0.045
Ethereum Classic	ETC	1,076	0.066	0.032

Table 2: List of cryptocurrencies and descriptive statistics. Market capitalization as of February 25, 2020, obtained from Coinmarketcap.com.
CryptoDynamics_Scraping

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 2, it becomes apparent that the crypto market is still dominated by Bitcoin. However, also ETH and XRP occupy a dominant position in the market.

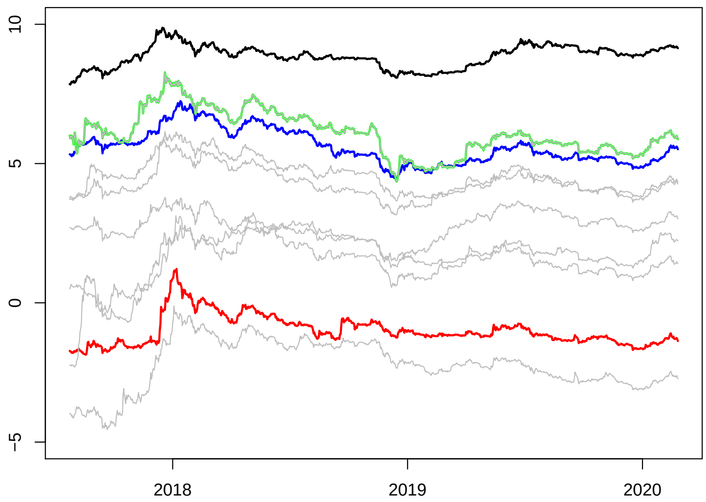


Fig. 2: Time Series of log prices from July 2017 - February 2020. BTC, ETH, XRP, BCH and all others.

CryptoDynamics_Series

Figure 2 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 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 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, which corresponds to a time span of about three months. However, the 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 in order 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 3.

	X_t		ΔX_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
XTZ	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

Table 3: p-values of the stationary tests for the level and first difference data.

Having confirmed that all the series are integrated of the same order, this allows to test for cointegration.

4.2 Estimation Results

	BTC	ETH	XRP	BCH	LTC	EOS	BNB	XMR	XLM	ETC
β_1	1.00	0.00	0.00	0.00	1.98	0.13	-0.94	-3.42	0.57	0.70
β_2	0.00	1.00	0.00	0.00	-0.28	-0.27	0.24	-1.09	0.11	0.31
β_3	0.00	0.00	1.00	0.00	-0.97	0.39	0.20	0.54	-0.76	0.00
β_4	0.00	0.00	0.00	1.00	0.53	-0.43	-0.06	-1.27	0.37	-0.42

Table 4: Estimated cointegration vectors $\widehat{\beta}$.

CryptoDynamics_Estimation

In the first step, we determine the cointegration rank graphically by using the Wachter QQ 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 Figure 3 that there are four cointegration relations since we can observe four eigenvalues deviating from the 45 degree line.

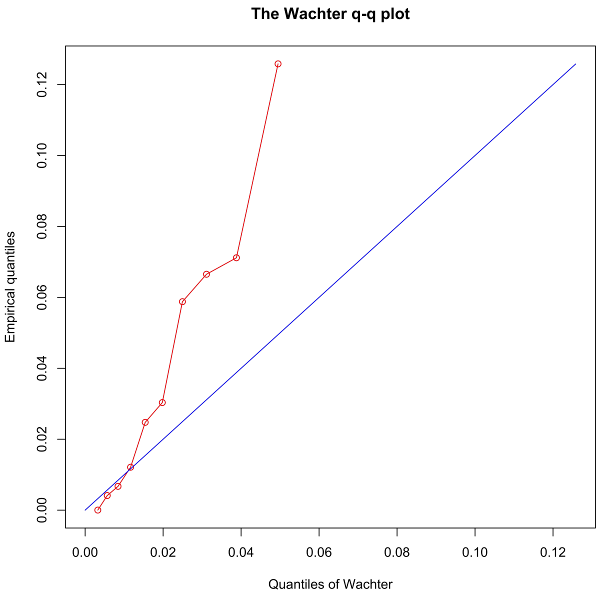


Fig. 3: Wachter QQ plot to determine the cointegration rank r.
CryptoDynamics_Wachter

Having fixed the cointegration rank, we can proceed with estimating the cointegration vectors. The estimated coefficients can be found in Table 4. 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 β_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 Figure 4. 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 α and Γ. In the following, we assume that the lag order $k = 1$.

The estimation results of our baseline VECM indicate that cointegration plays an important role for cryptocurrencies. See Table 5 for the estimation of the loading matrix α. The (j,i)-th entry of the table shows how currency j is affected by error correction term i, where $ECT_{j,t-1} \stackrel{\text{def}}{=} \widehat{\beta}_j^\top X_{t-1}$. 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

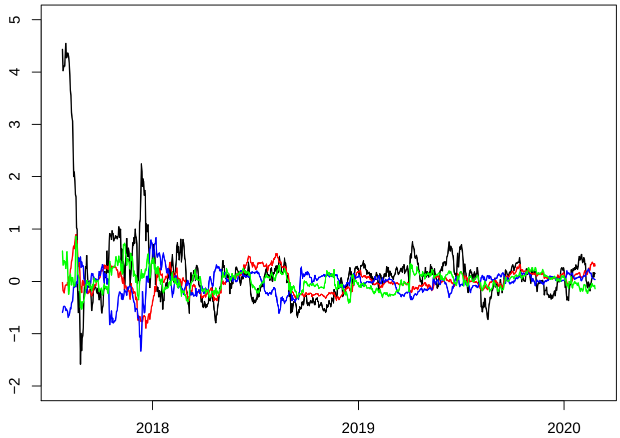


Fig. 4: Time series of the long-term stochastic trends. $\beta_1^\top X_{t-1}$, $\beta_2^\top X_{t-1}$, $\beta_3^\top X_{t-1}$ and $\beta_4^\top X_{t-1}$.

CryptoDynamics_Estimation

stochastic trends. The null and alternative hypotheses are:

$$H_0 : \alpha_{j,1} = \ldots = \alpha_{j,r} = 0 \quad \text{vs.} \quad H_1 : \exists\, \alpha_{j,k \leq r} \neq 0 \tag{21}$$

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 LTC, 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 short-term adjustment parameters. These results suggest that distortions in the long-run equilibrium will be corrected slowly and unevenly among the 10 cryptocurrencies. In the third one, XLM is the leader in the system, XRP, BCH and EOS carry the burden of adjustment to return to the long-run relationship. In the fourth one, EOS, XRM, ETC are the leaders in the system and that BCH carries the burden of adjustment to return to the long-run relationship.

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

Table 5: Estimated loading matrix $\widehat{\alpha}$. Red color indicates significance of negative coefficients, blue color indicates significance of positive coefficients, with significance at 5%, 1% and 0.1% level.

CryptoDynamics_Estimation

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

Table 6: Estimated coefficient matrix $\widehat{\Gamma}$. Red color indicates significance of negative coefficients, blue color indicates significance of positive coefficients, with significance at 5%, 1% and 0.1% level.

CryptoDynamics_Estimation

The estimation results for the lagged differences can be found in Table 6. 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.

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 section 2.3. In the first step, we estimate the cointegration vectors β 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 10000.

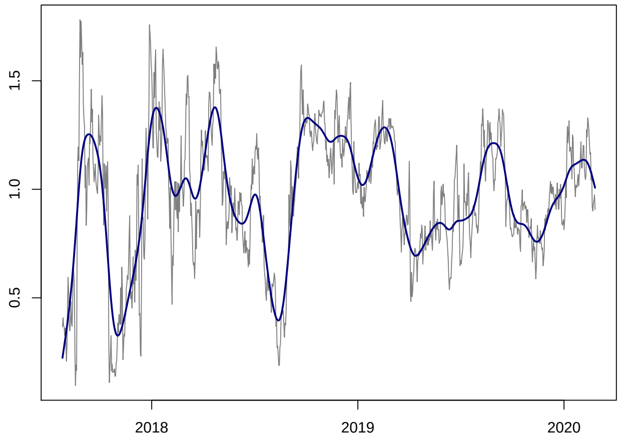


Fig. 5: Time series of cointegration intensity (grey) and spline interpolation (blue).

CryptoDynamics_Nonlinear

The time series of the estimated *COINtensity* function, $G(\widehat{\gamma}, \widehat{\beta}^\top X_{t-1})$, is visualized in Figure 5. 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 close to 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 no values exceeding the 0.5 and 1.5 thresholds. We conclude that nonlinearity was more prevalent in the turbulent period of the cryptocurrency bubble.

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, $j = 1, \ldots, r$, we have

$$\begin{aligned} S_{j,t} &= \beta_j^\top X_t \\ &= \beta_{j,1} X_{1,t} + \ldots + \beta_{j,p} X_{p,t} \end{aligned} \tag{22}$$

If the spread exceeds an upper threshold, we enter a long position, if the spread goes below the lower threshold, we enter a short 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. We choose three different threshold levels, $\tau \in (\pm 0.5\sigma_j, \pm\sigma_j, \pm 1.5\sigma_j)$, which are chosen to be symmetric around the long term mean of the stochastic trend, σ_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 position. 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 (2018) 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).

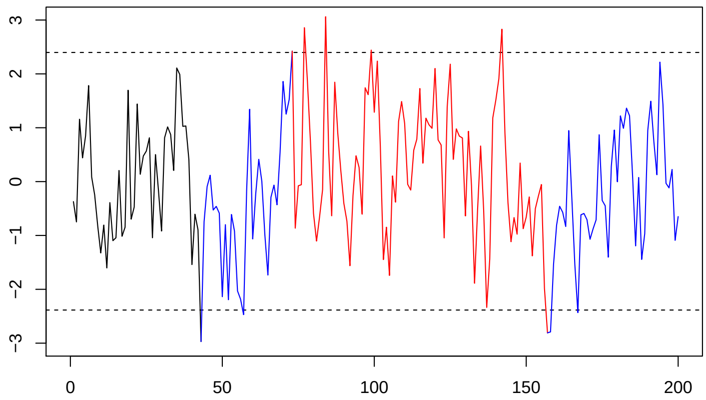


Fig. 6: Visualization of the statistical arbitrage trading strategy for simulated data. Neutral position, short position and long position.

Threshold	$\pm 0.5\sigma_j$	$\pm\sigma_j$	$\pm 1.5\sigma_j$	CRIX
Number of Trades	20	13	7	-
Net Profits	14,625	23,480	16,853	15,330
Maximal Drawdown	4,549	4,511	5,910	55,297
Annual Sharpe Ratio	0.59	0.93	0.66	0.22

Table 7: Performance statistics for different threshold levels.

CryptoDynamics_Trading

Table 7 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

Fig. 7: Performance of the trading strategy with thresholds $\tau = \pm\sigma$ (black) vs. CRIX (yellow).

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 the lowest maximal drawdown. While the net profits of the benchmark index portfolio (CRIX) are comparable to those of the strategies with thresholds $\pm 0.5\sigma_j$ and $\pm 1.5\sigma_j$, 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 7 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 largest 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.

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 performance in a backtesting study, beating the industry benchmark CRIX in terms of net profits, Sharpe ratio and maximal drawdown.

Appendix: Simulation Design

Baseline VECM specification:

$$\mathit{\Delta} X_t = \alpha\beta^\top X_{t-1} + \mathit{\Gamma}_1 \mathit{\Delta} X_{t-1} + \varepsilon_t,$$

with parameter matrices

$$\alpha = \begin{pmatrix} -1.47 & -1.3 & 0 & -1.26 \\ 0 & 0.97 & 0 & 0 \\ 0 & 0 & -0.74 & 0 \\ -1.19 & 0.85 & 0 & 0 \\ 0.55 & 0.78 & -1 & -1.37 \\ 0.8 & 0.75 & 0 & 0 \\ 0 & -0.74 & -1.26 & 0.78 \\ 0 & -1.4 & 0 & 0 \end{pmatrix}, \; \beta^\top = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -0.87 & 1.45 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1.48 \\ 0 & 0 & 1 & 0 & 0 & -1.29 & -0.53 & 0.9 \\ 0 & 0 & 0 & 1 & 0.8 & 1.49 & -0.82 & -0.69 \end{pmatrix},$$

$$\mathit{\Gamma}_1 = \mathrm{diag}\{0, 0.797929, 0, 0.793248, 0, 0.537687, 0, 0.722737\}.$$

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 FX, Yılmaz K (2014) On the network topology of variance decompositions: Measuring the connectedness of financial firms. Journal of Econometrics 182(1):119–134

Dijk Dv, Teräsvirta T, Franses PH (2002) Smooth transition autoregressive models—a survey of recent developments. Econometric reviews 21(1):1–47

Engle RF, Granger CW (1987) Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society pp 251–276

Granger C, Teräsvirta T (1993) Modelling nonlinear economic relationships oxford university press. New York

Guo L, Tao Y, Härdle WK (2018) A dynamic network perspective on the latent group structure of cryptocurrencies. arXiv preprint arXiv:180203708

Härdle WK, Harvey CR, Reule RC (2019) Understanding cryptocurrencies. Available at SSRN 3360304

Ji Q, Bouri E, Lau CKM, 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 pp 1551–1580

Johansen S (1995) Likelihood-based inference in cointegrated vector autoregressive models. Oxford University Press on Demand

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

Leung T, Nguyen H (2019) Constructing cointegrated cryptocurrency portfolios for statistical arbitrage. Studies in Economics and Finance

Liang C, Schienle M (2019) Determination of vector error correction models in high dimensions. Journal of econometrics 208(2):418–441

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 SJH, 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

Trimborn S, Härdle WK (2018) Crix an index for cryptocurrencies. Journal of Empirical Finance 49:107–122

Yi S, Xu Z, Wang GJ (2018) Volatility connectedness in the cryptocurrency market: Is bitcoin a dominant cryptocurrency? International Review of Financial Analysis 60:98–114

IRTG 1792 Discussion Paper Series 2020

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

001 "Estimation and Determinants of Chinese Banks' Total Factor Efficiency: A New Vision Based on Unbalanced Development of Chinese Banks and Their Overall Risk" by Shiyi Chen, Wolfgang K. Härdle, Li Wang, January 2020.

002 "Service Data Analytics and Business Intelligence" by Desheng Dang Wu, Wolfgang Karl Härdle, January 2020.

003 "Structured climate financing: valuation of CDOs on inhomogeneous asset pools" by Natalie Packham, February 2020.

004 "Factorisable Multitask Quantile Regression" by Shih-Kang Chao, Wolfgang K. Härdle, Ming Yuan, February 2020.

005 "Targeting Cutsomers Under Response-Dependent Costs" by Johannes Haupt, Stefan Lessmann, March 2020.

006 "Forex exchange rate forecasting using deep recurrent neural networks" by Alexander Jakob Dautel, Wolfgang Karl Härdle, Stefan Lessmann, Hsin-Vonn Seow, March 2020.

007 "Deep Learning application for fraud detection in financial statements" by Patricia Craja, Alisa Kim, Stefan Lessmann, May 2020.

008 "Simultaneous Inference of the Partially Linear Model with a Multivariate Unknown Function" by Kun Ho Kim, Shih-Kang Chao, Wolfgang K. Härdle, May 2020.

009 "CRIX an Index for cryptocurrencies" by Simon Trimborn, Wolfgang Karl Härdle, May 2020.

010 "Kernel Estimation: the Equivalent Spline Smoothing Method" by Wolfgang K. Härdle, Michael Nussbaum, May 2020.

011 "The Effect of Control Measures on COVID-19 Transmission and Work Resumption: International Evidence" by Lina Meng, Yinggang Zhou, Ruige Zhang, Zhen Ye, Senmao Xia, Giovanni Cerulli, Carter Casady, Wolfgang K. Härdle, May 2020.

012 "On Cointegration and Cryptocurrency Dynamics" by Georg Keilbar, Yanfen Zhang, May 2020.

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.