On the predictability of high and low prices: The case of Bitcoin

(Previsibilidade em preços máximos e mínimos: O caso do Bitcoin)

Leandro Maciel[†]

Rosangela Ballini[‡]

Abstract

Bitcoin has attracted the attention of investors lately due to its significant market capitalization and high volatility. This work considers the modeling and forecasting of daily high and low Bitcoin prices using a fractionally cointegrated vector autoregressive (FCVAR) model. As a flexible framework, FCVAR is able to account for two fundamental patterns of high and low financial prices: their cointegrating relationship and the long memory of their difference (i.e., the range), which is a measure of realized volatility. The analysis comprises the period from January 2012 to February 2018. Empirical findings indicate a significant cointegration relationship between daily high and low Bitcoin prices, which are integrated on an order close to the unity, and the evidence of long memory for the range. Results also indicate that high and low Bitcoin prices are predictable, and the fractionally cointegrated approach appears as a potential forecasting tool for cryptocurrencies market practitioners.

Keywords: Bitcoin; high and low prices; fractional cointegration; cryptocurrencies; forecasting

JEL Code: C31, C58, G17.

1. Introduction

Bitcoin (BTC), the most popular cryptocurrency traded in the digital money markets, exhibited a capitalization of about \$40.5 billion by mid-2017, representing 89% of the capitalization of all cryptocurrencies¹. Launched in 2009, Bitcoin transactions are based on an information technology infrastructure and on the lack of a central authority. Instead of relying on central banks, a decentralized computer network validates the transactions and grows the money supply of Bitcoin (Yermack, 2017). Users and investors have perceived huge financial potential in the Bitcoin market, driving the Bitcoin price from US dollar parity in early 2011 to about 1,500 US\$/BTC in mid-2017. Further, the number of transactions using Bitcoin has increased considerably and, according to Polasik et al. (2015), Bitcoin transactions per month increased from 12,000 to 2.1 million from August 2010 to August 2014, and in December 2015, approximately 200,000 Bitcoin transactions were carried out per day.

Bitcoin's particular features, in the absence of a regulatory agency, make the digital money a volatile and speculative currency, resulting in a market that is quite sensitive to real (e.g., economic, social and political) and fake (e.g., rumors) news. As stated by Alvarez-Ramirez et al. (2018) and Baek & Elbeck (2015), the poorly-defined liquidity conditions of the market and the lack of certain rules for investment realization add fragility to transactions, which are reflected as large price jumps and excessive volatility when compared to traditional currencies and assets. Despite that, a literature on cryptocurrencies has emerged, mostly focused on the legal aspects and the underlying blockchain technology.

Some authors, for example, discuss the efficiency of virtual money markets. Bartos (2015) indicates that Bitcoin returns follow the hypothesis of efficient markets by showing fast responses to publicly announced information in 2013-2014. On the other hand, based on automatic variance tests, Urquhart (2016) reveals that Bitcoin returns are significantly inefficient over the period from August 1st, 2010 to July 31st, 2016. However, when the sample is divided into two subsample periods, Urquhart (2016) finds that the Bitcoin market is efficient in the latter period – August 1st, 2013 to July 31st, 2016.

Submitted on November 12, 2018. Revised on March 10, 2019. Accepted on July 17, 2019. Published online in December, 2019. Editor in charge: João Caldeira.

[†](Corresponding author) Universidade Federal de São Paulo, Brasil. E-mail: maciel.leandro@unifesp.br

[‡]Universidade Estadual de Campinas, Brasil. E-mail: ballini@unicamp.br

¹ Other cryptocurrencies, based on blockchain technology, are for example the Litecoin (LTC), Ethereum (ETH), and Ripple (XRP). On the website <https://coinmarketcap.com/currencies/> market data regarding up to 641 of such assets can be found.

By verifying empirically whether or not there exist weekly price anomalies, Kurihara & Fukushima (2017) state that Bitcoin transactions are becoming more efficient, but its returns do not fulfill the efficient market hypothesis. Additionally, Urquhart (2017) find significant evidence of price clustering at round numbers as a Bitcoin price anomaly, with over 10% of prices ending with 00 decimals compared to other variations.

Bariviera et al. (2017) use detrended fluctuation analysis (DFA) over a sliding window to report that the Hurst exponent changed significantly during the first years of existence of Bitcoin, tending to stabilize from early 2014 to 2017. Also using DFA, Alvarez-Ramirez et al. (2018) estimate long-range correlations for price returns of Bitcoin. They find that the Bitcoin market exhibits periods of efficiency alternating with periods where the price dynamics are driven by anti-persistence.

The efficiency of Bitcoin market compared to gold, stock and foreign exchange markets is evaluated in the work of A-Yahyaee et al. (2018). They find that the long-memory feature and multifractality of the Bitcoin market is stronger, making the cryptocurrency more inefficient when compared to gold, stock and traditional currency markets.

Researchers have also devoted attention to the analysis of Bitcoin volatility. For instance, Balcilar et al. (2017) analyze the causal relationship between trading volume and Bitcoin returns, and trading volume and Bitcoin volatility. The causality-in-quantiles test reveals that volume can predict returns – except in Bitcoin bear and bull market regimes. It highlights the importance of modeling nonlinearities and accounting for the tail behavior when analyzing causal relationships between Bitcoin returns and trading volume. The authors additionally show, however, that volume cannot help predict the volatility of Bitcoin returns at any point of the conditional distribution.

Katsiampa (2017) explores the optimal conditional heteroskedasticity model with regards to goodness-of-fit to Bitcoin price data. The author shows that the best model is the AR-CGARCH model, highlighting the significance of including both a short-run and a long-run component for Bitcoin conditional variance. More recently, Lahmiri et al. (2018) investigate volatility nonlinear patterns in seven Bitcoin markets. The empirical findings identify the existence of long-range memory in Bitcoin volatility, irrespective of distributional inference. The same finding applies to Bitcoin entropy measurement, which indicates a high degree of randomness in the price series.

In general, the recent literature has stated that volatility of Bitcoin prices is an outcome of market sentiments, which can be attributed to the presence of significant “memory” (Katsiampa, 2017; Cheah et al., 2018; Cheah & Fry, 2015; Lahmiri et al., 2018). This emerges from a key element in the determination of Bitcoin prices: the assumption of full confidence of its users. Indeed, the literature has shown that Bitcoin is ideal for risk-averse investors in anticipation of negative shocks to the market (Dyhrberg, 2016a) and could be used as a hedging asset against market-specific risk (Dyhrberg, 2016b).

Due to the evidence of long memory of Bitcoin volatility,² this work aims to investigate whether or not the high and low prices of Bitcoin are predictable and which approach is appropriate to model these prices. Although much research has been devoted to the analysis of the predictability of daily market closing prices, few studies based on econometric time series models examined the case of high and low prices, as for instance the works of Baruník & Dvořáková (2015), Caporin et al. (2013), Cheung et al. (2010), Cheung et al. (2009), He & Hu (2009), and Cheung (2007). Indeed, Caporin et al. (2013) argue that the lack of studies regarding daily high and low asset prices is surprising for at least three reasons: i) the long histories of high and low prices data are readily available; ii) many technical analysis strategies use high and low prices to construct resistance and support levels; and iii) these prices can measure market liquidity and transaction costs.

In particular, daily high and low prices provide valuable information regarding the dynamic process of an asset over time, which can be seen as reference values for investors in trading analysis, e.g. through candlestick charts (Xiong et al., 2017; Cheung & Chinn, 2001). He & Wan (2009) state that highs and lows are referred to prices at which excess demand changes its direction. Additionally, these prices are related to the concept

² The literature also presents substantial evidence of long memory in the volatility process of asset prices, interest rate differentials, inflation rates, forward premiums and exchange rates (Yalama & Celik, 2013; Garvey & Gallagher, 2012; Breidt et al., 1998; Andersen & Bollerslev, 1997).

of volatility. Alizadeh et al. (2002) show that the difference between the highest and lowest (log) prices of an asset over a fixed sample interval, also known as the (log) range, is a highly efficient volatility measure. Brandt & Diebold (2006) and Shu & Zhang (2006) point out that the range-based volatility estimator appears robust to microstructure noise such as bid-ask bounce. This overcomes the limitations of traditional volatility models based on closing prices that fail to use the information contents inside the reference period of the prices, resulting in inaccurate forecasts.

Daily highs and lows can be used as stop-loss bandwidths, providing information about liquidity provisioning and the price discovery process. Further, Caporin et al. (2013) state that high (low) prices are more likely to correspond to ask (bid) quotes; thus, transaction costs and other frictions, such as price discreteness, the tick size (i.e., the minimal increments) or stale prices, might represent disturbing factors. Finally, high and low prices are more likely to be affected by unanticipated public announcements or other unexpected shocks. Therefore, aspects such as market resiliency and quality of the market infrastructure can be determinant (Caporin et al., 2013).

Hence, this paper suggests a fractionally cointegrated vector autoregressive model (FCVAR), as proposed by Johansen (2008) and Johansen & Nielsen (2010, 2012), to model and predict the relationship between Bitcoin highs and lows. The motivation of this approach is twofold. First, FCVAR modeling captures the cointegrating relationship between high and low prices, i.e. in the short term they may diverge, but in the long term they have an embedded convergence path. Second, the range (the difference between high and low prices), as an efficient volatility measure, is assumed to display a long memory, which allows for greater flexibility.³ Baruník & Dvořáková (2015) give a more general fractional or long-memory framework, where the series are assumed to be integrated of order d and cointegrated of order less than d , i.e. $CI(d - b)$, where $d, b \in \mathfrak{R}$ and, $0 < b \leq d$, which is more useful for capturing the empirical properties of data, in accordance with the evidence of long memory in the volatility of Bitcoin returns (Katsiampa, 2017; Cheah et al., 2018; Cheah & Fry, 2015; Lahmiri et al., 2018). Therefore, the FCVAR framework is able to model both the cointegration between highs and lows, and the long-memory property of the range. The results of FCVAR are then compared against traditional benchmarks over different prediction horizons using traditional accuracy measures, statistical tests for competitive forecasters, and also considering an economic criterion by performing a simple trading strategy.

The usefulness of the FCAR framework for financial time series modeling has been addressed by the literature. Baruník & Dvořáková (2015) explore the fractional cointegration relationship between daily high and low of stock exchange indices. The authors find that the range of all indices display long memory and are mostly in the non-stationary region, supporting the idea that volatility might not be a stationary process. Considering highs and lows prices of equity shares traded on the Brazilian stock exchange, Maciel (2018) states that the FCVAR approach can improve forecasting, compared to traditional time series models. Alternatively, Dolatabadi et al. (2016) use the FCVAR model to analyze the relationship between spot and futures prices in commodity markets, concluding that both prices are cointegrated, and the cointegration is of the fractional type.

Empirical findings in this work indicate that high and low prices of Bitcoin are non-stationary with fractional integration parameters statically equal and both close to unit. The range displays long memory. Further, Bitcoin highs and lows are fractionally cointegrated with one cointegration relationship, indicating the adequacy of FCVAR to estimate the dynamic between these prices. Parameter estimates of the FCVAR show that a linear combination of highs and lows is integrated in a non-zero order, and the range is in the stationary region. In addition, forecasting analysis suggests that Bitcoin high and low prices are predictable and that FCVAR does provide more accurate forecasts than traditional benchmarks in terms of traditional error measures and also according to the payoffs of a simple trading strategy.

This paper is outlined as follows. Section 2 describes the data and provides a preliminary analysis of daily high and low Bitcoin prices and the range, focusing on their integration, cointegration, and long memory properties. A FCVAR model for high and low Bitcoin prices is presented in Section 3. The predictability

³ The literature considers asset prices to be integrated of order 1, i.e. $I(1)$. However, the choice between stationary, $I(0)$, and non-stationary, $I(1)$, processes can be too restrictive for the degree of integration of daily high and low prices (Baruník & Dvořáková, 2015). Since these prices can be considered as a possibly fractionally cointegrated relationship, it improves flexibility, mainly when the error correction term from the cointegrating relationship between high and low prices is the range (Cheung, 2007; Fiess & MacDonald, 2002).

analysis is discussed in Section 4. Finally, Section 5 concludes and suggests topics for future investigation.

2. Analysis of daily high and low Bitcoin prices

This section describes the database and provides an analysis regarding the integration, cointegration and long memory properties of daily high and low Bitcoin prices and their difference, the range. Further, tests for the possible fractional cointegration relationship between highs and lows are also presented.

2.1 Database

The dynamic properties and the predictability of daily high and low Bitcoin prices are investigated considering data for the period from January 1st, 2012 to February 28th, 2018 with a total of 2,251 observations.⁴ High and low prices of Bitcoin (BTC) to US dollar (USD) exchange rate data were collected in www.coindesk.com.

Let $p_t^H = \log(P_t^H)$ be the daily high log-price, $p_t^L = \log(P_t^L)$ the daily low log-price, and $R_t = p_t^H - p_t^L$ the daily range, where P_t^H and P_t^L are the high and low prices at t , respectively. Figure 1 shows the daily BTC/USD exchange rate high and low log-prices for the period from 2012 to 2018. Daily low log-prices in Figure 1 are the actual daily low log-prices minus 0.25 to improve visibility. The decrease of prices after May 2013 was derived from the failure of Mt. Gox to protect transaction details, which also provoked the suspension of trading.⁵ From early 2016 to date, the Bitcoin market experienced significant growth as a result of the subsequent implementation of cryptosystems to guarantee transaction privacy, stabilizing the Bitcoin exchange system. Clearly daily highs and lows dynamic suggests the presence of a common trend, indicating that the series are cointegrated and also non-stationary. Figure 2 shows alternatively the temporal evolution of the difference between high and low Bitcoin, i.e. the range. It is worth to note that higher values of the range are associated with periods of high price variability, confirming its property as a volatility measure, the volatility range.

2.2 Cointegration and memory properties of Bitcoin highs and lows

We first conducted the evaluation of the stationarity of the high and low prices. Table 1 shows the Augmented Dickey-Fuller (ADF) (Dickey & Fuller, 1979) test results for the daily high and low log-prices (p_t^H and

⁴ We begin the analysis in 2012 since in this period the prices of Bitcoin start to show some relevant volatility.

⁵ In the early stage (2010–2013), Bitcoin exchange was handled by Mt. Gox, an administration system based on Shibuya, Japan. By the last months of 2013, Mt. Gox handled about 70% of transactions worldwide.

Figure 1
High and low log-prices of Bitcoin to US dollar exchange rate

Figure 2
Daily range of Bitcoin

Table 1

The p -values of ADF test for unit root for high (H) and low (L) log-prices and range (R) of Bitcoin based on levels and first-differences, where c, t denotes the inclusion of both a constant and a trend for daily high and low log-prices in levels only, while the others include only a constant (drift). Lags denotes the number of lags included in the model, based on the Bayesian Information Criteria (Schwarz, 1978).

model	lags	ADF _H		ADF _L		ADF _R
		level	first-differences	level	first-differences	level
c, t	2	0.5847	0.0000	0.6431	0.0001	0.0001

p_t^L) as well as the range (R_t). Summarizing, daily high and low prices are non-stationary at a 0.05 significance level and the range is a stationary process.⁶ This finding indicates that daily high and low prices may be cointegrated. However, the ADF test has as null hypothesis the possibility of a unit root against the $I(0)$ alternative, i.e. with a very low power against fractional processes.

We also performed the KPSS test of Kwiatkowski et al. (1992), appropriate in situations when the tested series are close to being a unit root (Maciel, 2018). The KPSS test results, reported in Table 2, confirm the non-stationarity of the high and low log-prices. However, regarding the range, the KPSS test suggests the presence of a unit root for the range, opposite to the ADF test. This conflicting result may be caused by the possible long memory property of the range. Table 2 provides results concerning short lags and long lags in the KPSS test, i.e. reported p -values are the average of the p -values from model specifications using from 1 to 10 lags (short lags) and also from 11 to 30 lags (long lags). Results for high and low log-prices for both short and long lags confirm the non-stationarity of the series. On the other hand, when long lags are concerned, the KPSS test results suggest that the range is stationary at a 0.05 significance level. This finding provides further evidence on the long memory of the range of Bitcoin prices.⁷

Due to the high degree of persistence of the range, cointegration analysis may not be satisfactory in explain-

⁶ Results are similar despite the model structure used in the ADF test, i.e. including a drift, a drift and a trend, and with no drift nor trend.

⁷ A high degree of persistence is also verified in the autocorrelation function (ACF) of the Bitcoin range, as significant autocorrelations were achieved even after 30 lags, confirming the results of the KPSS test and the evidence of long memory of the Bitcoin range - see Appendix A, Figure 4. The same results are also verified by analyzing Bitcoin squared range ACF (Appendix A, Figure 5).

Table 2

The p -values of KPSS test for unit root for high (H) and low (L) log-prices and range (R) of Bitcoin based on levels and two lag specifications, short lag and long lag, where c,t denotes the inclusion of both a constant and a trend for daily high and low log-prices in levels only, while the rest include a only a constant (drift). * indicates stationarity at a 0.05 significance level.

model	KPSS _H		KPSS _L		KPSS _R	
	short lag	long lag	short lag	long lag	short lag	long lag
c,t	0.0000	0.0001	0.0001	0.0001	0.0001	0.0892*

ing the relationship between high and low prices, which requires the use of a more appropriate framework such as the idea of fractional cointegration.

2.3 Testing the fractional cointegration order

Since the “error correction” term in the cointegrated relationship between Bitcoin high and low prices, i.e. the range, may contain long memory, this paper considers the use of the fractional cointegration framework. Let $X_t \equiv (p_t^H, p_t^L)'$ be a vector composed of the high and low Bitcoin prices, p_t^H and p_t^L , respectively. If the elements of X_t are $I(1)$ and there exists a linear combination $\beta'X_t$ that is an $I(0)$ process, X_t is called a cointegrated vector (Maciel, 2018). To relax the restriction on the choice between stationary $I(0)$ and non-stationary $I(1)$ processes, the series can be considered an $I(d)$ process with $d \in \mathfrak{R}$, where d is the fractional differencing parameter, fractional degree of persistence or fractional order of integration.

The series X_t is an $I(d)$ process if $u_t = (1 - L)^d X_t$ is $I(0)$, with L standing for the lag operator and $d < 0.5$ (Robinson & Yajima, 2002). If $d \geq 0.5$, X_t is defined as a non-stationary $I(d)$ series with $X_t = (1 - L)^{-d} u_t I\{t \geq 1\}$, where $t = 0, \pm 1, \pm 2, \dots$, and $I\{\cdot\}$ is an indicator function. For $d > 0$ ($d < 0$) the process has long-memory (anti-persistence). If $d = 0$, the process collapses to the random walk, i.e. a stationary process (Maciel, 2018).

To test the fractional order of integration of high and low log-prices and the range of Bitcoin, we employed the univariate exact local Whittle (ELW) estimator, as a semi-parametric approach, proposed by Nielsen & Shimotsu (2007). The ELW estimator is consistent in the presence or absence of cointegration, as to both stationary and non-stationary cases. Estimates of the fractional integration order do not imply the presence nor absence of cointegration. In addition, to test the equality of integration orders, $H_0 : d^H = d^L = d$, we also employed the test suggested by Nielsen & Shimotsu (2007), which is robust to the presence of fractional cointegration.⁸

The first six columns of Table 3 display the ELW estimates of $\hat{d}^H$, $\hat{d}^L$ and $\hat{d}^R$ for the Bitcoin, where the exponent denotes daily high (H), daily low (L) and daily range (R). The estimates of integration orders were calculated base on two specifications of bandwidth, $m_d = T^{0.5}$ and $m_d = T^{0.6}$, as in the works of Nielsen & Shimotsu (2007), Caporin et al. (2013), and Baruník & Dvořáková (2015). For both bandwidths, the order of integration of daily highs and lows are generally high and close to 1, indicating that the series are non-stationary for both bandwidth parameter. The difference between high and low prices (the range) displays long memory ($\hat{d}^R > 0$) with parameter $\hat{d}^R < 0.5$. Summarizing, the daily high and low Bitcoin prices are not stationary and the range displays long memory, in line with the results of Caporin et al. (2013) and Baruník & Dvořáková (2015).

Concerning the test for the equality of integration orders, the last two columns of Table 3 present the test statistics estimated with $m_d = T^{0.5}$ and $m_d = T^{0.6}$ as bandwidth parameters. Since the critical value of χ_1^2 is 2.71 in a 90% confidence interval, the null hypothesis of equality of the integration orders cannot be rejected for both bandwidth parameters (Maciel, 2018). These results suggest the use of a FCVAR model with the same degree of integration orders $d^H = d^L$. The generalization to the presence of fractional cointegration between highs and lows is novel in modeling of Bitcoin prices.

⁸ For details regarding the ELW estimator and the Nielsen & Shimotsu (2007) test for the equality of integration order, refer to Appendix B.

Table 3

Estimates of the fractional order of integration parameter d of high ($\hat{d}^H$) and low ($\hat{d}^L$) log-prices and the range ($\hat{d}^R$) of Bitcoin using the exact local Whittle (ELW) estimator, and test statistics for the equality of integration orders ($\hat{T}_0$). All estimates use both $m_d = T^{0.5}$ and $m_d = T^{0.6}$ as bandwidths.

ELW _{$m_d=T^{0.5}$}			ELW _{$m_d=T^{0.6}$}			$\hat{T}_0$	
$\hat{d}^H$	$\hat{d}^L$	$\hat{d}^R$	$\hat{d}^H$	$\hat{d}^L$	$\hat{d}^R$	$m_d = T^{0.5}$	$m_d = T^{0.6}$
0.9524	0.9361	0.3982	0.8973	0.8819	0.4113	0.1207	0.2238

3. FCVAR modeling for daily high and low Bitcoin prices

The fractionally cointegrated vector autoregression (FCVAR), suggested by Johansen (2008) and Johansen & Nielsen (2010, 2012), generalizes the classical cointegration analysis by allowing X_t to be fractional of order d and cofractional of order $d - b$, which suggests that $\beta'X_t$ should be fractional of order $d - b \geq 0$. This framework allows for the existence of a common stochastic trend, integrated with order d , and the short-term divergences from the long-run equilibrium integrated of order $d - b$. The parameter b is the strength of the cointegrating relationships, called the cointegration gap, in which a higher b means less persistence in the cointegrating relationships (Maciel, 2018).

In the FCVAR modeling approach, the usual lag operator and the difference operator are replaced by the fractional lag operator and the fractional difference operator, $L_b = 1 - \Delta^b$ and $\Delta^b = (1 - L)^b$, respectively (Johansen & Nielsen, 2012; Nielsen & Morin, 2016). The fractional difference operator is defined by the binomial expansion $\Delta^b Z_t = \sum_{n=1}^{\infty} (-1)^n \binom{b}{n} Z_{t-n}$ (Baruník & Dvořáková, 2015). Thus, the model is applied to $Z_t = \Delta^{d-b} X_t$. A fractionally cointegrated vector autoregressive FCVAR _{d,b} (p) model for $X_t \equiv (p_t^H, p_t^L)'$ as the vector of high and low prices can be defined as

$$\Delta^d X_t = \Delta^{d-b} L_b \alpha \beta' X_t + \sum_{i=1}^p \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \quad t = 1, \dots, T, \quad (1)$$

where α and β are $2 \times r$ matrices comprised by the long-run parameters, $0 \leq r \leq 2$, the rank r is termed the cointegration, or cofractional, rank, $d \geq b > 0$, $\Gamma = (\Gamma_1, \dots, \Gamma_p)$ are the autoregressive augmentation parameters related to the short-run dynamics, and ε_t is an p -dimensional *i.i.d* $(0, \Omega)$, with positive-definite variance matrix Ω .

The columns of β constitute the r cointegration (cofractional) vectors such that $\beta'X_t$ are the cointegrating combinations of the variables in the system, i.e. the long-run equilibrium relations. The parameters in α are the adjustment or loading coefficients. They represent the speed of adjustment towards equilibrium for each of the variables (Nielsen & Morin, 2016). If $d - b < 0.5$, $\beta'X_t$ is asymptotically a zero-mean stationary process. Denoting $\Pi = \alpha\beta'$, where the $2 \times r$ matrices α and β with $r \leq 2$ are assumed to have full column rank r , the columns of β are then the r cointegrating (cofractional) relationship determining the long-run equilibrium (Maciel, 2018).

The model parameters are estimated by maximum likelihood as described in Nielsen & Morin (2016). Before estimating the FCVAR models for daily high and low prices of Bitcoin, we need to test and determine the cointegration rank in the model, as discussed next.

3.1 Cointegration rank

A time series X_t is fractionally cointegrated $CI(d, b)$ if X_t has $I(d)$ elements and for some $b > 0$, there exists a vector β such that $\beta'X_t$ is integrated of order $(d - b)$. The cointegration rank test proposed by Nielsen & Shimotsu (2007) is considered in this work. It allows for both stationary and non-stationary fractionally integrated processes and is based on the exact local Whittle estimate of d , used to examine the rank of the spectral density matrix G and its eigenvalues. The test is described in Appendix C.

Table 4

Estimates of the fractional cointegration rank test statistics and their respective eigenvalues by the approach of Nielsen & Shimotsu (2007) using $\bar{d}_*$, the average of the estimated integration orders of daily high and low Bitcoin prices from the ELW estimator with $m_d = T^{0.6}$ as bandwidth parameter, in the fractional cointegration analysis for both $v(T) = m_L^{-0.45}$ and $v(T) = m_L^{-0.05}$, with $m_L = T^{0.5}$.

$\bar{d}_*$	eigenvalues		rank estimates					
			$v(T) = m_L^{-0.45}$			$v(T) = m_L^{-0.05}$		
	$\hat{\delta}_1$	$\hat{\delta}_2$	$L(0)$	$L(1)$	$\hat{r}$	$L(0)$	$L(1)$	$\hat{r}$
0.9442	0.2982	0.0004	-1.3201	-1.7822	1	-0.4297	-1.2033	1

Table 5

Likelihood ratio (LR) statistics and p -values from the cointegration test by Johansen & Nielsen (2012) for each rank $r = 0, 1, 2$, as well as the corresponding estimates of the parameter of the fractional order of integration ($\hat{d}$) and the parameter of the cointegration gap ($\hat{b}$) for Bitcoin high and low prices.

$r = 0$				$r = 1$				$r = 2$	
$\hat{d}$	$\hat{b}$	LR	p -value	$\hat{d}$	$\hat{b}$	LR	p -value	$\hat{d}$	$\hat{b}$
0.6892	0.6101	29.9827	0.000	0.9446	0.5987	0.1455	0.7255	0.9972	0.6212

Table 4 reports the results of the cointegration rank test of Nielsen & Shimotsu (2007) using $v(T) = m_L^{-0.45}$ and $v(T) = m_L^{-0.05}$. Estimates suggest one cointegration relationship for Bitcoin highs and lows. When $L(1) < L(0)$, this can be taken as strong evidence in favor of fractional cointegration between p_t^H and p_t^L (see Appendix C).

Additionally, the cointegration rank test proposed by Johansen & Nielsen (2012) is also considered. It tests the hypothesis $H_r : \text{rank}(\Pi) = r$ against the alternative $H_n : \text{rank}(\Pi) = n$. The test is described in Appendix D. Table 5 shows the a significant cointegration relationship between high and low Bitcoin prices according to the test of Johansen & Nielsen (2012). For $r = 0$, larger values of the likelihood ratio (LR) statistics indicate rejection of the null hypothesis of zero cointegrating relationship. Otherwise, when $r = 1$, the LR statistics are smaller and the corresponding p -values indicate that we cannot reject the null of one cointegrating relationship.

3.2 Empirical FCVAR model for Bitcoin highs and lows

Based on the previous evidence of one significant cointegrating vector for the high and low prices of Bitcoin, a fractionally cointegrating VAR (FCVAR) model was estimated with $p = 1$ for the short-term deviations, which is sufficient to capture the autocorrelation of the residuals.⁹

Table 6 reports the FCVAR estimates for the high and low log-prices of Bitcoin. Parameter estimates of the fractional integration order and the cointegration gap, $\hat{d}$ and $\hat{b}$ respectively, are significantly different from zero and different from each other. Estimate of $\hat{d}$ indicates that daily high and low prices are integrated of an order close to the unity. Regarding the cointegrating vector, $\hat{\beta}$, the estimates are very close to the vector $[1, -1]$, which is expected.

The results also suggest that a linear combination of the daily high and low prices (the range) is integrated of a non-zero order, and the range is in the stationary region ($d - b < 0.5$) – see Table 6. This finding is in line with work of Baruník & Dvořáková (2015) and Caporin et al. (2013) that considers several US stocks and different stock exchange indices, respectively. However, these authors indicate that the ranges of equities and indices fall mostly in the non-stationary region. Further, Baruník & Dvořáková (2015) also show that estimates of range

⁹ MacKinnon & Nielsen (2014) suggested that a single lag is usually sufficient in the fractional model, in contrast with the standard cointegrated VAR where more lags are required to account for the serial correlation in the residuals.

Table 6
FCVAR model estimates results

	$\hat{d}$	$\hat{b}$	$\hat{\beta}$	$\hat{\alpha}_H$	$\hat{\alpha}_L$	$\hat{\gamma}_{11}$	$\hat{\gamma}_{12}$	$\hat{\gamma}_{21}$	$\hat{\gamma}_{22}$
estimate	0.9649	0.5841	[1,-0.9980]	-0.1682	0.0936	-0.0167	0.2562	0.1918	0.1360
standard error	0.0290	0.0546		0.0200	0.0178	0.0322	0.0382	0.0271	0.0360
p-value	0.0000	0.0000		0.0000	0.0000	0.3020	0.0000	0.0000	0.0001

integration order have thus changed from the stationary region into the non-stationary region, suggesting that range long memory changes over time.

The estimates of the adjustment coefficients, $\hat{\alpha}^H$ and $\hat{\alpha}^L$, which describe the speed of adjustment of p_t^H and p_t^L toward equilibrium, are significantly different from zero (Table 6). Parameter $\hat{\alpha}^H$ is negative and $\hat{\alpha}^L$ is positive, indicating that they move in opposite directions to restore equilibrium after a shock to the system. Considering the absolute value of these parameter estimates, $\hat{\alpha}^H$ is greater than $\hat{\alpha}^L$, implying that the correction in the equation for daily highs overshoots the long-run equilibrium. These results are also verified by Baruník & Dvořáková (2015) and Caporin et al. (2013) in an analysis of equity prices. However, in more than 50% of the cases $\hat{\alpha}^H$ estimates were smaller than $\hat{\alpha}^L$.

Concerning the short-run dynamic parameter estimates $\Gamma_1 = (\hat{\gamma}_{11}, \dots, \hat{\gamma}_{22})$, according to Table 6, the coefficients of the lagged daily highs and lows are mostly positive, which suggests spillover effects.¹⁰ Finally, the residuals were also tested for the remaining autocorrelation and heteroskedasticity. In most cases, the null of no autocorrelation was rejected according to the Ljung-Box Q-test, but based on the visualization of the autocorrelation functions, the dependency is weak, and it disappears after the second lag. Some heteroskedasticity was also detected by the autocorrelation function of squared residuals, however, it is very weak.

4. Predictability of daily high and low Bitcoin prices

Next, the forecasting ability of the FCVAR modeling framework will be examined. Forecasts were performed using the FCVAR in an out-of-sample set comprised by the last two years of data. As competing models, we consider the VECM model of Cheung (2007); the random walk, RW; the ARIMA model; the 5-day moving average, MA₅; and the 22-day moving average, MA₂₂; the latter two of which correspond to weekly and monthly averages, respectively, and are often employed by technical analysts.

The Diebold & Mariano (1995) test is carried out to measure the forecasting superiority of the FCVAR, focusing on the mean squared error (MSE) of the forecasts. The error of the model i for the h -step ahead forecasting horizon is defined by

$$\epsilon_{t+h,i}^H = p_{t+h}^H - \hat{p}_{t+h,i}^H, \quad (2)$$

for the daily high, and

$$\epsilon_{t+h,i}^L = p_{t+h}^L - \hat{p}_{t+h,i}^L, \quad (3)$$

for the daily low, with $i = \text{FCVAR, VECM, RW, ARIMA, MA}_5, \text{MA}_{22}$, where p_t^H (p_t^L) and $\hat{p}_t^H$ ($\hat{p}_t^L$) are the actual and predicted high (low) Bitcoin prices at t .

Forecasting to assess the prediction performance of fractionally cointegration models for high and low Bitcoin prices is performed not only one-step-ahead, as done by Caporin et al. (2013) concerning asset prices, but also five- and ten-step-ahead to examine the medium- and long-term forecasting ability of the empirical FCVAR and selected competitors.

Models' ability to predict daily high and low Bitcoin log-prices is measured using traditional accuracy measures, such as the root mean square error (RMSE) and the mean absolute percentage error (MAPE) metrics,

¹⁰ Cheung (2007) states that negative coefficients imply a regressive behavior, whereas positive coefficients are an indication of spillover effects. In this case, higher daily highs tend to fall to a lower level, lower daily highs tend to drift up to a higher level, and higher daily lows lead to higher daily highs (Baruník & Dvořáková, 2015).

calculated as

$$\text{RMSE}_i^B = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(\frac{p_{t+h}^B - \hat{p}_{t+h}^B}{p_{t+h}^B} \right)^2}, \quad (4)$$

$$\text{MAPE}_i^B = \frac{100}{T} \sum_{t=1}^T \frac{|p_{t+h}^B - \hat{p}_{t+h}^B|}{p_{t+h}^B}, \quad (5)$$

where $B = \{L, H\}$ represents the low and high log-prices (i.e., the interval bounds of the prices), T is the sample size, and RMSE_i^L (MAPE_i^L) and RMSE_i^H (MAPE_i^H) are the RMSE (MAPE) for the low and high Bitcoin prices, respectively, of model i , where $i = \text{FCVAR}, \text{VECM}, \text{RW}, \text{ARIMA}, \text{MA}_5, \text{MA}_{22}$.

In some practical situations, a correct prediction of the direction of change can be more important than the magnitude of the error. Thus, the results are also compared in terms of a conventional measure of direction accuracy:

$$\text{DA}_i^B = \frac{1}{T} \sum_{t=1}^T Z_{i,t+h}^B, \quad (6)$$

where

$$Z_{i,t+h}^B = \begin{cases} 1, & \text{if } (\hat{p}_{t+h}^B - p_{t+h}^B)(p_{t+h}^B - p_{t+h-1}^B) > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (7)$$

again, $B = \{L, H\}$ represents the low and high log-prices (i.e., the interval bounds of the prices).

Table 7 shows models' prediction performance in terms of RMSE, MAPE, and DA, computed individually for both low (L) and high (H) Bitcoin log-prices. Better results are highlighted in bold. Concerning RMSE and MAPE measures, lower values correspond to better forecasters. For both low and high Bitcoin prices, the FCVAR model outperforms all of other competitors besides the forecasting horizon, except for two cases where VECM provides better but similar error measures than FCVAR. In general, FCVAR and VECM provide similar results, but better than the remaining models, indicating that accounting for the relationship between lows and highs can provide improvements in forecasting. The RW performs similarly to ARIMA, and they are both slightly better than the moving average techniques. Overall, the rankings for Bitcoin lows and highs from the best to worst in terms of RMSE and MAPE are: FCVAR, VECM, ARIMA, RW, MA₅, and MA₂₂.

When forecasters are compared in terms of direction accuracy, results from Table 7 highlight the superiority of the VECM and FCVAR over the other methods, as they provide higher and similar hit rates. Here, higher DA values indicate better accuracy. It is conceivable that the reason for the inferiority of the alternative models is that they ignore the mutual dependency between the daily highs and lows of the Bitcoin log-prices. On the other hand, the performance of moving average methods are in general 50%, which is similar to a 50-50 situation. Regarding the ARIMA method, besides presenting a performance inferior to VECM and FCVAR, in general the direction accuracy is higher than 50%.

Figure 3 illustrates the performance of FCVAR modeling framework using actual Bitcoin candlesticks with the corresponding predicted high-low bands by FCVAR for the last three months of data considering one-step-ahead predictions. FCVAR provides a good fit of the high-low dispersion, indicating the potential of the proposed method which can enhance chart analysis, a tool often used by technical traders. In general, the accuracy of the FCVAR model is lower when predicting Bitcoin drops than rises (Figure 1), mainly for the low prices. This evidence is also confirmed by the results of Table 7, where in general, models perform better when predicting highs than low prices, indicating some sort of asymmetry in Bitcoin highs and lows. This issue has been discussed in the literature, suggesting that financial prices tend to respond differently to good and bad news (Medovikov, 2016).

Table 8 shows ranking results of the Diebold& Mariano (1995) test for the out-of-sample forecasts of daily high and low Bitcoin log-prices obtained using the FCVAR against the benchmark models. For each forecasting horizon and for both highs and lows, models are ranked in terms of the RMSE values, from 1 (better - lower RMSE) to 5 (worst - higher RMSE). Therefore, pairwise comparisons are made across the ranking such that $A > B$ indicates that models A and B can be considered equally accurate, while $A >^* B$ means that model

Table 7

Performance comparison of competitive approaches in terms of accuracy measures and direction accuracy for low and high Bitcoin prices.

metric	method					
	MA ₂₂	MA ₅	RW	ARIMA	VECM	FCVAR
<i>panel A: one-step-ahead prediction horizon</i>						
RMSE ^L	0.0064	0.0063	0.0057	0.0055	0.0048	0.0040
RMSE ^H	0.0059	0.0055	0.0046	0.0044	0.0039	0.0032
MAPE ^L	0.3726	0.3701	0.3669	0.3653	0.3576	0.3417
MAPE ^H	0.3254	0.3187	0.2967	0.2817	0.2715	0.2645
DA ^L	0.5241	0.5199	-	0.5365	0.5454	0.5542
DA ^H	0.5098	0.5111	-	0.5243	0.5310	0.5472
<i>panel B: five-steps-ahead prediction horizon</i>						
RMSE ^L	0.0103	0.0109	0.0154	0.0079	0.0063	0.0054
RMSE ^H	0.0098	0.0102	0.0113	0.0068	0.0058	0.0050
MAPE ^L	0.4988	0.4765	0.4893	0.4437	0.3987	0.3644
MAPE ^H	0.4777	0.4811	0.4425	0.3798	0.3566	0.3588
DA ^L	0.5019	0.5187	-	0.5309	0.5398	0.5462
DA ^H	0.5102	0.4993	-	0.5210	0.5287	0.5409
<i>panel C: ten-steps-ahead prediction horizon</i>						
RMSE ^L	0.0287	0.0299	0.0236	0.0229	0.0156	0.0149
RMSE ^H	0.0247	0.0265	0.0210	0.0199	0.0146	0.0150
MAPE ^L	0.5676	0.5565	0.5454	0.5216	0.4988	0.4765
MAPE ^H	0.5465	0.5654	0.5391	0.5300	0.4876	0.4487
DA ^L	0.4987	0.4875	-	0.5186	0.5322	0.5289
DA ^H	0.4882	0.5087	-	0.5087	0.5240	0.5277

Figure 3
Bitcoin candlesticks and FCVAR predicted high-low bands

Table 8
Forecasting models ranking from Diebold-Mariano test for high and low Bitcoin prices

* indicates the mean difference between the two competing methods is significant at the 5% level.

price	rank of methods											
	1		2		3		4		5		6	
Panel A: one-step-ahead prediction horizon												
high	FCVAR	>*	VECM	>*	ARIMA	>	RW	>*	MA ₅	>	MA ₂₂	
low	FCVAR	>*	VECM	>*	ARIMA	>	RW	>*	MA ₅	>	MA ₂₂	
Panel B: five-steps-ahead prediction horizon												
high	FCVAR	>*	VECM	>*	ARIMA	>	RW	>*	MA ₅	>	MA ₂₂	
low	FCVAR	>	VECM	>*	ARIMA	>	RW	>*	MA ₅	>	MA ₂₂	
Panel C: ten-steps-ahead prediction horizon												
high	VECM	>	FCVAR	>*	ARIMA	>	RW	>*	MA ₅	>	MA ₂₂	
low	FCVAR	>	VECM	>*	ARIMA	>	RW	>*	MA ₅	>	MA ₂₂	

A is statistically more accurate than method B. From Table 8, when forecasts are performed for one-step-ahead, FCVAR statistically outperforms the remaining models. VECM is statistically better than ARIMA, which is equally as accurate as the RW model. Also, the RW do provide more accurate forecasts in statistical terms than the moving average approaches. Results are the same for the five-steps-ahead forecasting horizon, except for Bitcoin lows, where forecasts from FCVAR and VECM can be considered equally accurate. Moving across long horizons, as ten-steps-ahead, results from Table 8 indicate that FCVAR and VECM are equally accurate, but both techniques show statistically superior forecasts than the alternative competitors, indicating that when the relationship between highs and lows is taken into account by the model, accuracy is improved. Summing up, the results indicate the predictability of daily high and low prices of Bitcoin. Moreover, use of a long memory framework such as the FCVAR does improve forecasting performance in short- and long-term prediction horizons.

In addition to the Diebold-Mariano test, this paper also conducted a multiple forecasting comparison test through the Model Confidence Set (MCS), suggested by Hansen et al. (2011). The MCS test constructs a set $\mathcal{M}_{1-\alpha}^*$ containing the best forecasting model(s) with probability greater than or equal to $(1 - \alpha)$ in a given set of competing models $\mathcal{M}_0$. The test comprises an elimination rule that identifies the object (competitor) to be removed from the set of objects (competitors) where the algorithm stops when the $\mathcal{M}_{1-\alpha}^*$ set contains only the “surviving” objects (competitors).¹¹

Table 9 shows the results from the Model Confidence Set (MCS) test. In general, conclusions regarding the MCS test are similar for high and low Bitcoin prices considering a 5% significance level. For one-step-ahead, the $\mathcal{M}_{75\%}^*$ includes FCVAR, VECM, RW, and ARIMA as the best forecasting models. In a more restrictive set, $\mathcal{M}_{90\%}^*$ is comprised of only FCVAR and VECM. Moving across the forecasting horizon, for five- and ten-steps-ahead only FCVAR and VECM remain in the $\mathcal{M}_{75\%}^*$ set, except at a five-steps-ahead forecasting horizon where the ARIMA and RW are still in the $\mathcal{M}_{75\%}^*$ set only for Bitcoin highs. The $\mathcal{M}_{90\%}^*$ set in these cases are empty, indicating that models’ forecasting accuracy deteriorates as the forecasting horizon rises. The moving average approaches, MA₅ and MA₂₂, are consistently kicked out of the $\mathcal{M}_{75\%}^*$ and $\mathcal{M}_{90\%}^*$ sets, confirming the results from the Diebold-Mariano test (Table 8).

Finally, results are also compared using an economic criterion. Considering the predicted highs and lows, the performance of a simple trading strategy is calculated, as in the work of Xiong et al. (2017). Let p_t^O and p_t^C be the opening and closing Bitcoin prices at t , respectively, and $\hat{p}_{t+1}^H$ and $\hat{p}_{t+1}^L$ be the forecasted high and low prices for day $t + 1$ after the market closes on day t . The trading strategy is comprised of four steps as follows (Xiong et al., 2017): i) on a given day t , a “buy” signal for the asset is generated if $\hat{p}_{t+1}^H - p_t^O > p_t^O - \hat{p}_{t+1}^L$; ii) if

¹¹ Details on the MCS test are found in Hansen et al. (2011).

Table 9

Model Confidence Set (MCS) p -values for high and low Bitcoin log-prices forecasts considering one- ($h = 1$), five- ($h = 5$) and ten-steps-ahead ($h = 10$) prediction horizons. Forecasts in $\mathcal{M}_{90\%}^*$ and $\mathcal{M}_{75\%}^*$ are identified by one and two asterisks, respectively.

model	p -value – highs			p -value – lows		
	$h = 1$	$h = 5$	$h = 10$	$h = 1$	$h = 5$	$h = 10$
FCVAR	0.127*	0.326**	0.248**	0.140*	0.517**	0.318**
VECM	0.188*	0.298**	0.291**	0.128*	0.453**	0.229**
RW	0.540**	0.254**	0.032	0.426**	0.011	0.017
ARIMA	0.441**	0.281**	0.010	0.452**	0.040	0.020
MA ₅	0.047	0.012	0.003	0.040	0.037	0.001
MA ₂₂	0.032	0.044	0.001	0.038	0.013	0.001

Table 10

Performance comparison of competitive approaches in terms of economic criteria (annualized returns and percentage of trades with positive annualized returns) of predicted high and low Bitcoin prices.

metric	method					
	MA ₂₂	MA ₅	RW	ARIMA	VECM	FCVAR
annualized returns	14.35%	15.27%	15.84%	20.09%	29.44%	32.22%
positive trades	38.45%	40.19%	41.88%	49.13%	55.29%	58.90%

the “buy” signal is observed for k consecutive days beginning with day t , buy the asset on day $t + k - 1$ using the closing value p_{t+k-1}^C ; otherwise, hold the capital; iii) on another day s subsequent to buying the asset, a “sell” signal is generated if $\hat{p}_{s+1}^H - p_s^O < p_s^O - \hat{p}_{s+1}^L$; iv) sell the asset on day $s + k - 1$ using the closing value p_{s+k-1}^C of that day if a “sell” signal has been observed for k consecutive trading days beginning with day s ; otherwise, hold the asset (Maciel, 2018). Based on one-step-ahead forecasts, we take $k = 2$ as an example.

Table 10 shows the annualized return and the percentage of trades resulting in positive returns from a trading strategy concerning Bitcoin high and low prices forecasts.¹² In general, one must notice that the models perform well in terms of annualized returns and percentage of positive trades. The annualized returns range from 14.35% (MA₂₂) to 32.22% (FCVAR). FCVAR and VECM provide similar results, but they are superior to the remaining approaches in terms of annualized returns. RW performs slightly better than the moving average techniques. Concerning the percentage of trades with positive returns, FCVAR and VECM are the only methodologies that achieve profitable trades larger than 50%. The highest percentage of profitable trades is 58.90% when the FCVAR forecasts are used. The remaining methods provide similar percentages of profitable trades.

5. Conclusion

This work evaluated the predictability and dynamic properties of daily high and low Bitcoin prices. The modeling of daily high and low prices considered a fractionally cointegrated VAR model (FCVAR), which accounts for two fundamental patterns of these prices: their cointegrating relationship and the long-memory of their difference (i.e., the range), as the error correction term is allowed to fall into a non-stationary region. The empirical analysis examined daily high and low prices of Bitcoin (BTC) to US dollar (USD) rate exchange during the period from January 2012 to February 2018. The findings indicate that daily high and low Bitcoin prices are integrated of an order close to the unity. The range displays long memory and is in the stationary region. A significant cointegration relationship was found between daily high and low prices. The empirical FCVAR model shows that high and low prices move in opposite directions to restore equilibrium after a shock

¹² A one-time 0.1% deduction was considered in order to mimic the transaction cost. Additionally, it is supposed that investors can enter the market at any time during the evaluation period.

to the system occurs. Also, the results indicate predictability of daily highs and lows of Bitcoin for different forecasting horizons, in which the fractional approach leads to better predictions than competitive methods, providing higher payoffs in a simple trading strategy. Future work should include estimation of the FCVAR with the restriction on the cointegrating vector β to be $(1, -1)$, which allows the interpretation of the difference $(d - b)$ as the order of integration of the range, as well as the analysis of asymmetries between high and low price forecasting. The evaluation of the results for intraday trading also appears as a topic for further research.

APPENDIX A - ACF of daily and squared daily range of Bitcoin

Figure 4
ACF of daily range of Bitcoin

Figure 5
ACF of daily squared range of Bitcoin

APPENDIX B - Univariate exact local Whittle (ELW) estimation

The univariate local exact Whittle estimators for highs, lows, and the range ($\hat{d}^H$, $\hat{d}^L$ and $\hat{d}^R$, respectively) are found by minimizing the following contrast function (Maciel, 2018):

$$Q_{m_d}(d^i, G_{ii}) = \frac{1}{m_d} \sum_{j=1}^{m_d} \left[\log \left(G_{ii} \lambda_j^{-2d^i} \right) + \frac{1}{G_{ii}} I_j \right], \quad i = H, L, R, \quad (\text{A.1})$$

which is concentrated with respect to the diagonal element of the 2×2 matrix G , a finite and nonzero matrix with strictly positive diagonal elements. Under the hypothesis that the spectral density of $U_t = [\Delta^{d^H} p_t^H, \Delta^{d^L} p_t^L, \Delta^{d^R} R_t]$, G satisfies

$$f_U(\lambda) \sim G \text{ as } \lambda \rightarrow 0, \quad (\text{A.2})$$

where $f_U(\lambda)$ is the spectral density matrix, I_j the coperiodogram at the Fourier frequency $\lambda_j = \frac{2\pi j}{T}$ of the fractionally differenced series U_t , m_d is the number of frequencies used in the estimation, and T is the sample size (Caporin et al., 2013). The matrix G is estimated as

$$\hat{G} = \frac{1}{m_d} \sum_{j=1}^{m_d} \text{Re}(I_j), \quad (\text{A.3})$$

with $\text{Re}(I_j)$ standing for the real part of the coperiodogram.

To test the equality of integration orders, $H_0 : d^H = d^L = d$ one may use the methodology suggested by Nielsen & Shimotsu (2007), which is robust to the presence of fractional cointegration. In the bivariate case under study, the test statistic is (Nielsen & Shimotsu, 2007):

$$\hat{T}_0 = m_d (S\hat{d})' \left(S \frac{1}{4} \hat{D}^{-1} (\hat{G} \odot \hat{G}) \hat{D}^{-1} S' + h(T)^2 \right)^{-1} (S\hat{d}), \quad (\text{A.4})$$

where $\odot$ is the Hadamard product, $\hat{d} = [\hat{d}^H, \hat{d}^L]$, $S = [1, -1]'$, $h(T) = \log(T)^{-k}$ for $k > 0$, $D = \text{diag}(G_{11}, G_{22})$.

According to Nielsen & Shimotsu (2007), if the variables are not cointegrated, i.e. the cointegration rank is $r = 0$, $\hat{T}_0 \rightarrow \chi_1^2$, while if $r \geq 1$, the variables are cointegrated and $T_0 \rightarrow 0$. For significant large values of the test statistic $\hat{T}_0$ with respect to the null density χ_1^2 , it provides evidence against the null hypothesis of the equality of integration orders (Maciel, 2018).

APPENDIX C - Cointegration rank test of Nielsen & Shimotsu (2007)

For the bivariate case, the cointegration rank test of Nielsen & Shimotsu (2007) estimates the rank r by

$$\hat{r} = \arg \min_{u=0,1} L(u), \quad (\text{C.1})$$

where

$$L(u) = v(T)(2-u) - \sum_{i=1}^{2-u} \hat{\delta}_i, \quad (\text{C.2})$$

for some $v(T) > 0$ which satisfies

$$v(T) + \frac{1}{m_L^{1/2} v(T)} \rightarrow 0, \quad (\text{C.3})$$

with $\hat{\delta}_i$ as the i -th eigenvalue of $\hat{G}$, and m_L a new bandwidth parameter.

The estimation of matrix G involves two steps. First, $\hat{d}^H$ and $\hat{d}^L$ are obtained using (A.1) with m_d as the bandwidth parameter. Given $\bar{d}_* = (\hat{d}^H + \hat{d}^L)/2$, the matrix G is estimated as follows:

$$\hat{G} = \frac{1}{m_L} \sum_{j=1}^{m_L} \text{Re}(I_j), \quad (\text{C.4})$$

such that $m_L/m_d \rightarrow 0$. The estimates of G are robust to all different choices of m_d and m_L (Nielsen & Shimotsu, 2007).

APPENDIX D - Johansen & Nielsen (2012) cointegration test

Let $L(d, b, r)$ be the profile likelihood function given rank r , where (α, β, Γ) have been concentrated out by regression and reduced rank regression (Nielsen & Morin, 2016). For the model with a constant, the test concerns the hypothesis $H_r : \text{rank}(\Pi, \mu) = r$ against $H_n : \text{rank}(\Pi, \mu) = n$, with $L(d, r)$ as profile likelihood function given rank r , where the parameters $(\alpha, \beta, \rho, \Gamma)$ have been concentrated out by regression and reduced rank regression (Maciel, 2018).

The profile likelihood function is maximized both under the hypothesis H_r and under H_n considering the LR test statistic is computed as follows:

$$LR(q) = 2 \log (L(\hat{d}_n, \hat{b}_n, n) / L(\hat{d}_r, \hat{b}_r, r)), \quad (\text{D.1})$$

where $q = n - r$ and

$$L(\hat{d}_n, \hat{b}_n, n) = \max_{d, b} L(d, b, n), \text{ and } L(\hat{d}_r, \hat{b}_r, r) = \max_{d, b} L(d, b, r). \quad (\text{D.2})$$

The asymptotic distribution of $LR(q)$ depends qualitatively (and quantitatively) on the parameter b . In the case of “weak integration”, $0 < b < 0.5$, $LR(q)$ has a standard asymptotic distribution (Nielsen & Morin, 2016):

$$LR(q) \xrightarrow{D} \chi^2(q^2), \quad 0 < b < 0.5. \quad (\text{D.3})$$

Otherwise, in the case of “strong cointegration”, when $0.5 < b \leq d$, asymptotic theory is nonstandard and

$$LR(q) \xrightarrow{D} \text{Tr} \left\{ \int_0^1 dW(s) F(s)' \left(\int_0^1 F(s) F(s)' ds \right)^{-1} \int_0^1 F(s) dW(s)' \right\}, \quad b \geq 1/2, \quad (\text{D.4})$$

where the vector process dW is the increment of ordinary (non-fractional) vector standard Brownian motion of dimension $q = p - r$ (Nielsen & Morin, 2016).

References

- A-Yahyaee, K. H., Mensi, W. & Yoon, S. (2018). Efficiency, multifractality, and the long-memory property of the Bitcoin market: A comparative analysis with stock, currency, and gold markets. *Finance Research Letters* **27**: 228-234.
- Alizadeh, S., Brandt, M. W. & Diebold, F. X. (2002). Range-based estimation of stochastic volatility. *Journal of Finance* **57**(3): 1047-1091.
- Alvarez-Ramirez, J., Rodriguez, E. & Ibarra-Valdez, C. (2018). Long-range correlations and asymmetry in the Bitcoin market. *Physica A* **492**: 948-955.
- Andersen, T. G. & Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance* **52**(3): 975-1005.
- Baek, C. & Elbeck, M. (2015). Bitcoins as an investment or speculative vehicle? A first look. *Applied Economic Letters* **22**(1): 30-34.
- Balcilar, M., Bouri, E., Gupta, R. & Roubaud, D. (2017). Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling* **64**: 74-81.
- Bariviera, A. F., Basgall, M. J., Hasperué, W. & Naiouf, M. (2017). Some stylized facts of the Bitcoin market. *Physica A* **484**: 82-90.
- Bartos, J. (2015). Does Bitcoin follow the hypothesis of efficient market? *International Journal of Economic Science* **4**(2): 10-23.
- Baruník, J. & Dvořáková, S. (2015). An empirical model of fractionally cointegrated daily high and low stock market prices. *Economic Modelling* **45**: 193-206.
- Brandt, M. W. & Diebold, F. X. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business* **79**(1): 61-74.
- Breidr, F. J., Crato, N. & Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, **83**(1-2): 325-348.
- Caporin, M., Ranaldo, A. & Magistris, P. C. (2013). On the predictability of stock prices: A case for high and low prices. *Journal of Banking & Finance* **37**: 5132-5146.
- Cheah, E. T. & Fry, J. M. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters* **30**: 32-36.
- Cheah, E. T., Mishra, T., Parhi, M. & Zhang, Z. (2018). Long memory interdependency and inefficiency in Bitcoin markets. *Economics Letters* **167**: 18-25.
- Cheung, Y. L., Cheung, Y. W., He, A. W. & Wan, A. T. (2009). A high-low model of daily stock price ranges. *Journal of Forecasting* **28**: 103-119.
- Cheung, Y. L., Cheung, Y. W. & Wan, A. T. (2010). A trading strategy based on callable bull/bear contracts. *Pacific-Basin Finance Journal* **18**: 186-198.
- Cheung, Y. W. (2007). An empirical model of daily highs and lows. *International Journal of Finance & Economics* **12**(1): 1-20.
- Cheung, Y. W. & Chinn, M. D. (2001). Currency traders and exchange rate dynamics: A survey of the US market. *Journal of International Money and Finance* **20**(4): 439-471.

- Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**(266): 427-431.
- Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economics Statistics* **13**(3): 253-265.
- Dolatabadi, S., Nielsen, M. O. & Xu, K. (2016). A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets. *Journal of Empirical Finance* **38**: 623-639.
- Dyhrberg, A. H. (2016a). Bitcoin, gold and the dollar - A GARCH volatility analysis. *Finance Research Letters* **16**: 85-92.
- Dyhrberg, A. H. (2016b). Hedging capabilities of Bitcoin. Is it the virtual gold? *Finance Research Letters* **16**: 139-144.
- Fiess, N. M. & MacDonald, R. (2002). Towards the fundamentals of technical analysis: Analysing the information content of high, low and close prices. *Economic Modelling* **19**(3): 353-374.
- Garvey, J. F. & Gallagher, L. A. (2012). The realised-implied volatility relationship: Recent empirical evidence from FTSE-100 stocks. *Journal of Forecasting* **31**(7): 639-660.
- Hansen, P. R., Lunde, A. & Nason, J. M. (2011). The model confidence set. *Econometrica* **79**(2): 453-497.
- He, A. W. & Wan, A. T. (2009). Predicting daily highs and lows of exchange rates: A cointegration analysis. *Journal of Applied Statistics* **36**(11): 1191-1204.
- He, L. T. & Hu, C. (2009). Impacts of interval computing on stock market variability forecasting. *Computational Economics* **33**(3): 263-276.
- Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press, New York.
- Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory* **24**(3): 651-676.
- Johansen, S. & Nielsen, M. O. (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics* **158**(1): 51-66.
- Johansen, S. & Nielsen, M. O. (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica* **80**(6): 2667-2732.
- Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters* **158**: 3-6.
- Kurihara, Y. & Fukushima, A. (2017). The market efficiency of Bitcoin: A weekly anomaly perspective. *Journal of Applied Finance and Banking* **7**: 57-64.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* **54**(1): 159-178.
- Lahmiri, S., Bekiros, S. & Salvi, A. (2018). Long-range memory, distributional variation and randomness of Bitcoin volatility. *Chaos, Solitons and Fractals* **107**: 43-48.
- Maciel, L. (2018). Technical analysis based on high and low stock prices forecasts: Evidence for Brazil using a fractionally cointegrated VAR model. *Empirical Economics*. <https://doi.org/10.1007/s00181-018-1603-8>

- MacKinnon, J. G. & Nielsen, M. O. (2014). Numerical distribution functions of fractional unit root and cointegration tests. *Journal of Applied Econometrics* **29**(1): 161-171.
- Medovikov, I. (2016). When does the stock market listen to economic news? New evidence from copulas and news wires. *Journal of Banking & Finance* **65**: 27-40.
- Nielsen, M. & Morin, L. (2016). A Matlab program and user's guide for the fractionally cointegrated VAR model. *Technical report*, QED Working Paper 1330. Queen's University.
- Nielsen, M. O. & Shimotsu, K. (2007). Determining the cointegrating rank in nonstationary fractional systems by the exact local Whittle approach. *Journal of Econometrics* **141**(2): 574-596.
- Polasik, M., Piotrowska, A. I., Wisniewski, T. P., Kotokowski, R. & Lightfoot, G. (2015). Price fluctuations and the use of Bitcoin: An empirical inquiry. *International Journal of Electronic Commerce* **20**(1): 9-49.
- Robinson, P. M. & Yajima, Y. (2002). Determination of cointegrating rank in fractional systems. *Journal of Econometrics* **106**(2): 217-241.
- Shu, J. H. & Zhang, J. E. (2006). Testing range estimators of historical volatility. *Journal of Future Markets* **26**(3): 297-313.
- Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters* **148**: 80-82.
- Urquhart, A. (2017). Price clustering in Bitcoin. *Economics Letters* **159**: 145-148.
- Xiong, T., Li, C. & Bao, Y. (2017). Interval-valued time series forecasting using a novel hybrid Holt and MSVR model. *Economic Modelling* **60**: 11-23.
- Yalama, A. & Celik, S. (2013). Real or spurious long memory characteristics of volatility: Empirical evidence from an emerging market. *Economic Modelling* **30**: 67-72.
- Yermack, D. (2017). Corporate governance and blockchains. *Review of Finance* **21**: 7-31.