Madalena Oliveira Figueiredo

FORECASTING THE VOLATILITY OF BITCOIN USING MARKET AND BLOCKCHAIN INFORMATION

Dissertação no âmbito do Mestrado em Métodos Quantitativos em Finanças, orientada pelo Professor Doutor Helder Miguel Correia Virtuoso Sebastião e pelo Professor Doutor Pedro Miguel Avelino Bação, apresentada ao Departamento de Matemática da Faculdade de Ciências e Tecnologia e à Faculdade de Economia.

Março de 2021

Forecasting the Volatility of Bitcoin using Market and Blockchain Information

Madalena Oliveira Figueiredo

1 2 9 0

UNIVERSIDADE DE COIMBRA

Master in Quantitative Methods in Finance

Mestrado em Métodos Quantitativos em Finanças

MSc Dissertation | Dissertação de Mestrado

Março 2021

Acknowledgements

To my research supervisors, Professor Helder Sebastião and Professor Pedro Bação, who helped me solve all the setbacks that arose throughout this project, my deepest thanks not only for their valuable advice and guidance but also for their constant availability and support.

To the University of Coimbra and to everyone I came across in my academic path that, in one way or another, helped me to get here.

To the friends I made in this master's degree, for all the help, companionship, for all the outbursts and the laughter that allowed this whole course to become less tumultuous. I will never forget what it was like to experience Coimbra with you.

To my longtime friends, who have become family, for accompanying me and being present in all phases, both good and less good, for all your affection, advice, understanding and for knowing that I can count on you for everything, unconditionally. For all the last-minute dinners and getaways. Know that I will always be here to return it.

To my family, who celebrate all my conquests as if they were theirs, with special emphasis on my mother, without whom nothing would be possible. For being my base, for all the support and understanding, for all the conversations, for never judging, for aligning with everything to cheer me up, for being my number one fan and for her ability to cheer me up in my grayest days... for everything. I will never be able to repay. To my granny and to Gabriela, for your ability to make me laugh, you are very special.

And lastly, to all those who are no longer present, but that wherever they are, I have no doubt that they will be feeling the proudest.

Abstract

Since its creation in 2008, Bitcoin, the ground-breaking and most popular cryptocurrency until this day, has grown exponentially along several dimensions, attracting attention not only from investors but also from the general public. It has been the subject of extensive research, especially regarding its speculative nature, market efficiency, integration, and, a popular topic, volatility. The present work focuses on daily forecasting the volatility in the overall Bitcoin market, and aims to answer the question of whether other market information besides prices and Blockchain information are helpful for forecasting that volatility. Our research design is based on the strong assumptions that CoinMarketCap provides reliable price information on the overall Bitcoin market, and that its volatility is well measured by the Parkinson range-based estimator. Accordingly, an examination is conducted regarding the forecasting performance of Autoregressive and GARCH models with exogenous variables obtained both from the online market and from the Blockchain network. The results show that, as already documented in the literature, trading volume has no incremental information value, but realized volatility of liquid online exchanges and other carefully chosen Blockchain variables may improve the forecasting accuracy especially when using a GARCH-type model. The most striking result is that Autoregressive models are clearly superior to their GARCH models, independently of the predictor sets used, with the lag structure of the dependent variable corresponding to the main source of predictability.

Resumo

Desde a sua criação em 2008, Bitcoin, a primeira criptomoeda e a mais popular até ao momento, cresceu exponencialmente em diversas dimensões, atraindo não só a atenção dos investidores, mas também do público em geral. Esta tem sido objeto de extensas investigações, especialmente no que diz respeito à sua natureza especulativa, eficiência de mercado, integração e, um tópico particularmente recorrente, volatilidade. O presente trabalho foca a sua atenção na previsão diária da volatilidade no mercado da Bitcoin, como um todo, e visa concluir se outras informações de mercado, para além das informações de preços, e de Blockchain são úteis para prever essa volatilidade. A premissa desta investigação baseia-se em fortes suposições de que o CoinMarketCap fornece informações confiáveis acerca dos preços na generalidade do mercado da Bitcoin, e que sua volatilidade é bem medida pelo estimador de Parkinson. Nesse sentido, é estudado o desempenho de previsão de modelos Autoregressivo e GARCH que incluem variáveis exógenas obtidas tanto no mercado online quanto na rede Blockchain. Os resultados mostram que, em conformidade com o que já fora documentado na literatura, o volume de transação não possui valor de informação incremental, no entanto a volatilidade realizada proveniente de bolsas online líquidas e outras variáveis Blockchain que foram cuidadosamente escolhidas podem produzir melhorias de precisão da previsão, especialmente quando são utilizados modelos GARCH. O resultado mais surpreendente corresponde ao facto de que os modelos Autorregressivos são claramente superiores aos modelos GARCH, independentemente dos conjuntos de previsores utilizados, correspondendo o desfasamento da variável dependente à principal fonte de previsibilidade.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

CoinMarketCap[1] defines cryptocurrency as a digital medium of exchange, which assures financial transactions, oversees the creation of additional units, and verifies the transfer of assets using strong cryptography technology.

Since the creation of Bitcoin (BTC) in 2008, by Satoshi Nakamoto[2], cryptocurrencies' popularity has grown tremendously, attracting attention from the general public and investors (Bação et al., 2018). For instance, online commerce needs the presence of a financial institution that serves as intermediary, i.e. as a third party that guarantees the processing of transactions. However, there are some disadvantages associated with transactions processed by financial intermediaries, such as slowness, and high costs. With the introduction of Bitcoin, this reality may change, due to its open-source protocol, controlled by the users without the need of a reliable third party to ensure safeness.

This peer-to-peer system uses network nodes to verify the transactions, which are then recorded in a public distributed ledger called Blockchain. This ledger can record transactions between two parties in a way that is efficient, verifiable, and immutable (Reid and Harrigan, 2013). The security of Blockchain is based on the assumption that "miners", motivated by financial incentives (rewards in new units of Bitcoin), are working to maintain the integrity of the ledger. Bitcoin and some other cryptocurrencies are designed to gradually decrease the production of new units, placing an artificial cap on the total amount of circulating coins. This has been seen as a feature that has some affinity with the real world maximum offer of precious metals (Barber et al., 2012).

Lansky (2018) emphasizes the unique features of cryptocurrencies, namely decentralization, pseudo-anonymity, and ability to avoid double-spending attacks. The first feature concerns independence from central authorities or other third parties, which assures that changes in the rules are rare and can only happen in the case of a consensus among the majority of operators. The last two features are provided by Blockchain mechanisms, that give users the possibility of keeping their real life identity anonymous, and prevent the holder of the digital currency from using the same digital entity in multiple payments at the same time.

[1] https://coinmarketcap.com/alexandria/article/what-are-cryptocurrencies

[2] Satoshi Nakamoto is a pseudonym created by a person or group of persons, whose real identity is, to the present date, still unknown.

The increase in popularity of Bitcoin led to the emergence of other digital currencies, sometimes referred to as altcoins[3], offering more and more diversity to the world of cryptocurrencies. One of these cryptocurrencies is Ethereum, which is not in fact just another digital currency. Ethereum employs advanced Blockchain technology, when compared to other digital currencies, and therefore enables new applications, usually referred to as smart contracts.

Presently, Bitcoin holds the top place in the cryptocurrencies market, with its market capitalization situated at 954.3 billion USD (approximately 61.8% of the total cryptocurrency market capitalization). Ethereum is the second biggest cryptocurrency, with a market capitalization of 208.3 billion USD, representing 13.47% of the total market capitalization. At the present time, the cryptocurrency industry consists of more than 8500 cryptocurrencies and 35000 online exchanges, with different user bases and trade volumes[4].

Forecasting the volatility of Bitcoin returns is of great interest if investors are considering including this digital currency in their investment portfolios (Shen et al., 2020). Given the growing interest in Bitcoin, it is not surprising the attention that the issue of selecting reliable models for forecasting the risk associated with Bitcoin investments has been receiving in the literature (Ardia et al., 2019). This work intends to contribute to this strand of the literature, that is aimed at studying the volatility forecasting of Bitcoin, and more precisely to answer the question of whether market and Blockchain information have incremental value for forecasting volatility. With this purpose in mind the methodologies of Taylor and Xu (1997) and Pong et al. (2004) will be employed.

This study uses daily data for the period from January 3, 2014 to December 1, 2020. This period was then divided into an in-sample period, from January 3, 2014 to December 31, 2017, and an out-of-sample period, from January 1, 2018 to December 1, 2020. The data was collected from three distinct sources, which will later be described: CoinMarketCap (https://coinmarketcap.com/), Bitcoincharts website (https://bitcoincharts.com/) and Coinmetrics (https://coinmetrics.io/).

We start from the premise that comparison of volatility forecasts can provide more evidence about incremental information. Thus, in order to reach a conclusion about the incremental value from the inclusion of other market variables other than prices and Blockchain's variables, we computed forecasts using two classes of alternative models, with different sets of predictors. The first class consists of Autoregressive Models with exogenous variables (AR-X) where the dependent variable is the logarithm of the Parkinson measure of volatility, and the predictors include combinations of lags of the dependent variable with market and Blockchain variables. The second class are Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, ranging from a simple GARCH(1,1) model to more encompassing models with exogenous variables in the variance equation (i.e. GARCH(1,1)-X). Next, 1-step-ahead forecasts of volatility were computed and compared using several metrics. The goal is to assess whether the inclusion of additional variables has incremental value for forecasting Bitcoin volatility. As a final step in this research, and in order to clarify whether the observed differences between the forecasts are significant, we applied the test of Diebold and Mariano (1995).

The contribution of this research to the literature is twofold. First, it introduces Blockchain variables as a possible source of information regarding the volatility of Bitcoin. Second, it focus on the

[3]The term has several similar definitions, but basically means an alternative digital currency to Bitcoin (Chaum, 1983).
[4]Data obtained from the website https://coinmarketcap.com/, last accessed on February 17, 2021.

Parkinson metric, which is a range-based volatility estimator, as a good measure for the unobserved volatility of the Bitcoin market. The use of this estimator allows the measurement of volatility across the entire Bitcoin market, i.e., considers the information from various online exchanges. It is almost consensual the claim that realized volatility computed using 5-minute returns is the best measure of volatility, however this approach would require reliable information on all transaction data from several online exchanges, which besides being time consuming and a real computational challenge, creates the additional problem on how to combine the returns from different exchanges. That is why studies that use Bitcoin realized volatility, only gather information from one online exchange and assume it to be representative of the overall Bitcoin market.

The remainder of this dissertation is structured as follows. Chapter 2 presents a brief historical contextualization regarding cryptocurrencies. Chapter 3 presents a literature review, referring several articles about digital currencies, that are relevant in the context of this study. Chapter 4 lays out the methodologies on which this work is based. Chapter 5 describes the data employed in this study, as well as the forecasts obtained with the several models under analysis. Finally, Chapter 6 highlights the main conclusions.

Chapter 2

Brief Historical Contextualization

One of the first steps in the field of digital currencies was taken by Chaum (1983), who conceived the idea of an anonymous cryptographic electronic money, eCash. The eCash software could be installed in any user's local computer, providing storage of a digital form of money. The use of this digital currency would require a cryptographic signature by a bank, allowing users to interact with government-issued and regulated currencies, however it did not completely resolved the double-spending issue, and hence did not succeed in attracting a significant number of users.

In 2009 the first successful decentralized cryptocurrency was created, following the cryptography proposal of Nakamoto (2008), where it is stated that "a purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution". Satoshi Nakamoto, a person or group of people developed Bitcoin (BTC), as well as Blockchain, the underlying mechanism on which cryptocurrencies are based and legitimized (Lakhani and Iansiti, 2017; Sebastião et al., 2021).

The persistent cypherpunk movement in the eighties and nineties contributed to promoting the concept of cryptocurrency (Karlstrøm, 2014). With the primary focus on the premise that governments and large corporations are unable to guarantee individual privacy, this movement argued that these institutions monitor and censor the communications between individuals and therefore constrain individual freedom. Pursuing an ideological and philosophical agenda, the cypherpunk movement pushed for individual autonomy, freedom of choice and voluntarism. Thus, in the beginning, cryptocurrency enthusiasts contemplated cryptocurrencies as a close approximation to the ideal concept of money advocated by Friedrich von Hayek and right-libertarians of the Austrian school of economics (Hayek, 1990), who argued in favour of ending the monopoly of central banks in the production, distribution, and management of money (Sebastião et al., 2021).

The subprime mortgage crisis from 2007 to 2010, which had a worldwide impact, involved not only international financial markets but also the banking sector, and deeply affected the European sovereign debt. The crisis contributed to distrust of banks and financial institutions in general, and also affected negatively the views on monetary authorities, regulators, and politicians. This created favorable conditions for attention to be given to alternative monetary systems, namely to Bitcoin (Sebastião et al., 2021).

Chapter 3

Literature Review

The initial research on cryptocurrencies was produced in the fields of Computer Science, Cryptography and Law, with primary focus on Bitcoin and its technical and methodological characteristics (Bação et al., 2018). Subsequently, the literature approaching the issue from a Finance perspective quickly became very large, mainly due to the rise in interest from the public, together with the surprising explosive price behaviour of cryptocurrencies. In fact, the speculative nature of Bitcoin is a major topic in the literature. For example, Cheah and Fry (2015) provide evidence of the existence of speculative bubbles in the prices of Bitcoin, while Fry and Cheah (2016) resort to the close relationship between Mathematical Finance and Statistical Physics to develop an Econophysics model which suggests that Bitcoin and Ripple had episodes of negative bubbles.

Despite those concerns, cryptocurrencies, and especially Bitcoin, have been establishing themselves as an highly speculative investment asset, often being referred to as New Gold. For instance, Matkovskyy et al. (2020) investigate the relationship — regarding returns and volatility — between Bitcoin and traditional financial markets, represented by five stock market indices, namely NASDAQ100, S&P500, Euronext100, FTSE100 and NIKKEI225, with a focus on the effects of Economic Policy Uncertainty (EPU). The conclusion of that analysis is that Bitcoin may provide a hedge against uncertainty in traditional markets, especially in the USA, as it seems to be the case with gold. Conversely, the analysis of this topic presented in Klein et al. (2018) lead to the conclusion that Bitcoin is not the new gold, as its properties appear to differ markedly from those of gold, especially in periods of market distress.

An important part of the literature on cryptocurrencies focuses on cryptocurrency volatility modeling and forecasting, the issue to which the present study aims to contribute. A wide range of models have been employed, a significant part of which include exogenous variables in the variance equation. An exception is provided by Chaim and Laurini (2019), who study the return and volatility dynamics of the nine major cryptocurrencies: Bitcoin, Ethereum, Ripple, Litecoin, Stellar, Dash, Monero, NEM, and Verge. The authors rely on the estimation of a multivariate stochastic volatility model with common discontinuous jumps to returns and volatility. The analysis emphasizes two high volatility periods in 2017 and at the beginning of 2018. The permanent volatility component appears to be influenced by developments in the market and by the interest of the public in these markets. The transitory mean jumps become more important in the second part of the sample, indicating a shift in the dynamics of these cryptocurrencies. Another important result is that stationary models with jumps seem to account for the long-memory characteristics of cryptocurrencies.

Nevertheless, the most popular models for volatility appear to be in the GARCH family. One part of this literature has used GARCH-MIDAS models to add information to the standard GARCH model. For example, Conrad et al. (2018) examine the economic determinants of long-term Bitcoin volatility. The variables considered are macroeconomic and financial variables, such as the Baltic dry index and VIX (the Chicago Board Options Exchange Volatility Index), as well as Bitcoin variables, such as trading volume. The results show that Bitcoin volatility is pro-cyclical, i.e., it increases in periods when economic activity is higher. Walther et al. (2019) apply the same approach to five major cryptocurrencies: Bitcoin, Ethereum, Litecoin, Ripple, and Stellar. The authors find Global Real Economic Activity to be the most important exogenous driver of volatility, outperforming all the others that were examined. Fang et al. (2020) also include five cryptocurrencies (Bitcoin, Ethereum, Ripple, Litecoin and New Economy Movement) in the analysis, but the focus is limited to the News-based Implied Volatility (NVIX), a text-based index that captures investors' perceptions of future uncertainty. The results show that NVIX is an important determinant of the long-term volatility of the five cryptocurrencies. The authors conclude that investor's perception of uncertainty is more important than uncertainty regarding economic fundamentals.

The GARCH family of models is very large. Therefore it is possible to find papers using many different GARCH-type models. To allow for asymmetry in the response to shocks to the volatility of twenty major cryptocurrencies, Baur and Dimpfl (2018) use both a TGARCH and a quantile autoregressive model of order 1. They find that a positive shock produces larger volatility increases than a negative shock, i.e., that there is asymmetry, which the authors relate to the behaviour of informed and uninformed investors. Bouri et al. (2016) also find evidence of asymmetric response to volatility shocks in Bitcoin. However, they include VIX as an additional variable in their asymmetric GARCH model. The results show that Bitcoin volatility responds negatively to the VIX. Nevertheless, the results also show that the dynamics of volatility has changed over time. A model in the GARCH class that addresses the issue of time-varying behaviour is the Markov–switching GARCH model (MSGARCH). Ardia et al. (2019) used an MSGARCH to test for regime changes in the volatility dynamics of Bitcoin. They find strong evidence of regime changes in the volatility of Bitcoin.

Instead of focusing on a single GARCH model, Trucíos (2019) compares the forecasting performance of alternative GARCH models. The forecasts concern one-step-ahead volatility and Value-at-Risk in the Bitcoin market. The results highlight the importance of employing robust procedures to deal with outliers. Using a similar approach, Köchling et al. (2020) evaluate the forecasting ability of a wide range of GARCH-type models. Again the focus is on one-day-ahead forecasts of the volatility of Bitcoin returns. The main conclusion is that it is difficult to select one model as clearly superior to the rest, highlighting the difficulty of forecasting volatility in a changing environment.

An alternative to GARCH models are then heterogeneous autoregressive (HAR) models. Bouri et al. (2021) use a HAR model, as well as the machine-learning technique of random forests, to study the importance of the US-China trade war in forecasting out-of-sample daily realized volatility of Bitcoin. The basic HAR model is extended to include controls, namely for jumps, while the measure of US-China trade tensions is based on Google Trends data. The authors conclude that this measure improves the accuracy of volatility forecasts. Another study of the volatility of Bitcoin where HAR models are used is Shen et al. (2020). This paper compares the forecasting accuracy of several HAR models. The results show that the HAR models that allow for jumps or structural breaks perform better. A further example of HAR modeling of Bitcoin volatility is provided by Aalborg et al. (2019). Their HAR models include several additional variables: returns, trading volume, change in the number of unique Bitcoin addresses, the VIX index and Google searches for "Bitcoin". Of these variables, at the 5% significance level, only the trading volume improves the basic HAR model for daily volatility. Aalborg et al. (2019) also analyze similar models for Bitcoin returns, but the variables are not consistently significant.

The common procedure in HAR models is to use realized volatility (RV) as the dependent variable, i.e., as a proxy for the latent volatility, following Andersen and Bollerslev (1998). In the present dissertation we use both RV and a range-based estimator of volatility, proposed by Parkinson (1980). Range-based estimators are based on the evidence that daily extreme prices, i.e. the lowest and highest prices of each day, reveal information about the entire price process. In fact, Christensen and Podolskij (2007) proposed an alternative volatility measure, called the realized range-based volatility (RRV), which is given by the difference between the maximum and minimum prices during a certain period, while the RV is defined as the sum of non-overlapping squared returns within a fixed period. For further information regarding realized volatility measures see Liu et al. (2015).

Another approach that has become more popular in recent years makes use of quantile regression techniques combined with Granger causality tests. This approach is used by Balcilar et al. (2017) to study the causal relationships between trading volume and Bitcoin's volatility and returns. Specifically, they apply a non-parametric causality-in-quantiles test to the analysis of the causal relationship between trading volume and Bitcoin returns and volatility, testing at different quantiles across the conditional distributions. The results show that volume can predict Bitcoin returns in the middle part of the distribution, but not in the tails. However, the volume has no information regarding the volatility of Bitcoin returns in any part of the conditional distribution. Although the result regarding the ability to forecast returns is contradicted by Bouri et al. (2019), this latter study also concludes that volume does not predict volatility. Relative to Balcilar et al. (2017), the innovation in Bouri et al. (2019) is the addition of copulas to the quantile causality approach, i.e. they use a copula-quantile causality approach to study Granger causality from trading volume to returns and volatility. Additionally, they apply the approach not just to Bitcoin, but to seven leading cryptocurrency markets (Bitcoin, Ripple, Ethereum, Litecoin, Nem, Dash, and Stellar). Bouri et al. (2019) find evidence of Granger causality from trading volume to the returns of the cryptocurrencies, at both left and right tails, whereas Balcilar et al. (2017) only find that in the middle of the distribution. But, as mentioned above, Bouri et al. (2019), like Balcilar et al. (2017), do not find Granger causality from volume to volatility. Granger causality is also used in Yang and Kim (2015) in the context of network theory, i.e., they view the Bitcoin market as a network of agents that establish connections by trading. Therefore, they compute complexity measures that quantify the activity in the network, and test the Granger causality from those complexity measures to returns and volatility in the Bitcoin market. The authors conclude that one of the complexity measures that they compute is useful for forecasting both returns and volatility in the Bitcoin market.

Chapter 4

Methodology

The essential elements of the methodology used in this dissertation are based on two papers: Taylor and Xu (1997), concerning the testing of incremental information about volatility, and Pong et al. (2004), which compares the forecasting ability of volatility models. The first paper, Taylor and Xu (1997), extends a previous study by the same authors (Xu and Taylor, 1995), by moving from daily data (Forex exchange rates) to high frequency data, and thus incorporating more information in the computation of volatility measures. Taylor and Xu (1997) use the intraday data to compute a measure of observed volatility, realized volatility. Each day t in the sample is divided into n periods, with corresponding returns r_{ti}. The realized volatility at day t is then computed as

$$RV_t = \sqrt{\sum_{i=1}^{n} r_{ti}^2} \tag{4.1}$$

where r_{ti}^2 represents the squared logarithm return at the i^{th} interval. The value of n is chosen so that the returns are computed on a time interval with a certain length, usually chosen to be 5 minutes.

Taylor and Xu (1997) then proceed to estimate GARCH-type models of log-return conditional volatility. They estimate several versions of the baseline model, excluding or including, realized volatility and another measure of volatility, given by the volatility implicit in the price of options (implied volatility). To test the existence of additional information on those two measures of volatility, the authors test the significance of the corresponding parameters and compare the log-likelihood of the alternative models. In addition, the authors also evaluate the out-of-sample forecasting performance of the alternative models. Differently, in Pong et al. (2004) the variable to be forecasted is realized volatility, as in Equation (4.1). Pong et al. (2004) compare the forecasting performance of several models: ARMA, ARFIMA, GARCH, and a regression model with option implied volatility as an explanatory variable. The forecasts are compared on the basis of the mean squared error, the Diebold-Mariano test (Diebold and Mariano, 1995), and of the R-squared in a regression of the actual values on the forecasts.

The empirical strategy used in this dissertation takes inspiration from the procedures employed in both those papers, while moving the focus of attention from the foreign exchange rate market to the Bitcoin market.

As volatility is not an observable variable, it is necessary to begin by computing a measure of it. In this dissertation we assume that the volatility range-based estimator of Parkinson measure

(Parkinson, 1980) provides a good measure of that latent variable. Hence, the variable to be forecasted is this measure but we also assume that this task can be accomplished using the conditional variance forecasts provided by GARCH models. In addition, as in Taylor and Xu (1997), realized volatility in a particular online exchange is also used as a predictor. Realized volatility has been shown to be a highly accurate volatility measure (Andersen and Bollerslev, 1998). The measure of realized volatility used was calculated with 5-minute returns. The returns were computed with the transaction prices obtained from the Bitstamp exchange, one of the most liquid, and hence representative, Bitcoin exchanges. Because online exchanges are open for trading 24/7, the choice of 5-minute returns implies that, in Equation (4.1), $n = 288$, i.e.

$$RV_t = \sqrt{\sum_{i=1}^{288} r_{ti}^2} \tag{4.2}$$

where r_{ti}^2 represents the squared logarithm return at the i^{th} 5-minute interval, counting from 00:00:00 UTC.

The Parkinson realized-range-based volatility is here applied in a daily basis (Christensen and Podolskij, 2007). This estimator requires the maximum and minimum Bitcoin prices observed in each day t, also called the high and low prices, respectively, as shown in the following formula:

$$\sigma_t = \ln(H_t/L_t)/2\sqrt{\ln(2)} \tag{4.3}$$

where H_t and L_t are the high and low Bitcoin prices recorded during day t.

The reason why the Parkinson estimator is used as a measure of the volatility to be forecasted, while realized volatility is only used as a predictor, is that the interest is on measuring volatility in the overall Bitcoin market instead of volatility in one particular Bitcoin exchange. In fact, the Parkinson estimate is computed by aggregating the information from many Bitcoin exchanges, as the high and low prices refer to a price index weighted by trading volume of several relevant exchanges constructed by CoinMarketCap. In contrast, the measure of realized volatility, which requires intraday data, is computed using data from just one Bitcoin exchange. For further details on the data used in this dissertation, see Section 5.1.1.

As predictors in the models used for forecasting Bitcoin volatility, besides realized volatility, this dissertation will use the transaction volume (another market variable), as well as daily Blockchain information.

After collecting all the data and organizing the database, the next step was to estimate in-sample (see Section 5.1.1) two different classes of models. The first class consists of AR models where the dependent variable is the time series of the Parkinson estimate of volatility in the Bitcoin market. Several versions of the AR model will be estimated, differing in the combination of exogenous variables added to the autoregressive terms (AR-X). The second class of models is Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. Again, we estimate both a baseline model, GARCH(1,1), and more more complex GARCH(1,1) models that include exogenous variables in the variance equation (GARCH(1,1)-X models). As remarked in the Literature Review (Section 2), this type of model is very popular, not only among academics, but also among practitioners.

The AR-X models estimated in this dissertation can be written as

$$\ln(\sigma_t) = \sum_{i=1}^{p} a_i \ln(\sigma_{t-i}) + \sum_{j=1}^{q} \mathbf{B}_j \mathbf{X}_{t-j} + \varepsilon_t \tag{4.4}$$

where σ_t is the Parkinson estimator from Equation (4.3), a_i is the coefficient of lag i of the dependent variable, p is the order of the autoregressive part, $\mathbf{X}$ is the column vector of exogenous variables, $\mathbf{B}_j$ is the row vector of coefficients of the lag j of the exogenous variables, and ε_t is a white-noise error term. In practice we may use a more flexible model, allowing for the exogenous variables to have different lag structures.

As for the GARCH(1,1)-X models, Equations (4.5) and (4.6) are, respectively, the mean equation and the variation equation:

$$r_t = \mu + \varepsilon_t \tag{4.5}$$

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 + \mathbf{B}\mathbf{X}_{t-1} \tag{4.6}$$

In Equations (4.5) and (4.6), $r_t = \ln(P_t) - \ln(P_{t-1})$ are the log-returns, computed as the difference between two consecutive close prices, μ is a constant, and ε_t is the error term, assumed to follow a normal distribution with mean zero and conditional standard deviation h_t. $\mathbf{B}$ is the row vector of coefficients of the lagged exogenous variables, and ω is a positive constant, while α and β are the (non-negative) parameters of the ARCH and GARCH components, respectively. Notice also that in the conditional variance equation we only consider the first lag of the exogenous variables.

Given the estimated models, 1-step-ahead volatility forecasts are produced for the out-of-sample period. The procedure for a given model is the following: the model is initially estimated using the in-sample data, that is, until T, the last observation in-sample, using the estimated model we compute the forecast for $T+1$, then we re-estimate the model using data until $T+1$ and compute the forecast for $T+2$ and so forth until we obtain all forecasts for the out-of-sample period. In sum, the models are reiterated on an expanded window. These forecasts are then compared with the realized values of σ_t in order to assess the forecasting performance of the several models that, within each class, distinguish themselves by the different predictors space. The statistics used in this study are the Mean Error (ME), the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), the Mean Percent Error (MPE), the Mean Absolute Percent Error (MAPE) and Theil's U. Letting n be the number of forecasts, $t = 1, \ldots, n$ be the forecasted periods, y_t be the actual value of the forecasted variable at time t, and $\hat{y}_t$ be the forecast of y_t, the formulas for these statistics are the following:

$$ME = \frac{1}{n} \sum_{i=1}^{n} (y_t - \hat{y}_t) \tag{4.7}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_t - y_t| \tag{4.8}$$

$$MPE = \frac{100\%}{n} \sum_{i=1}^{n} \left(\frac{y_t - \hat{y}_t}{y_t} \right) \tag{4.9}$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{4.10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_t - \hat{y}_t)^2}{n}} \tag{4.11}$$

$$U = \sqrt{\frac{\sum_{i=1}^{n} \left(\frac{\hat{y}_t - y_t}{y_{t-1}}\right)^2}{\sum_{i=1}^{n} \left(\frac{y_t - y_{t-1}}{y_{t-1}}\right)^2}} \tag{4.12}$$

It should be noted that in Equations (4.7) to (4.12), y_t is given by $ln(\sigma_t)$ and $\hat{y}_t$ is given by $ln(\hat{\sigma}_t)$ in the case of AR-X models and $ln(\hat{h}_t)$ in the case of GARCH(1,1)-X models.

The Mean Error (ME) measures the average difference between the estimated values and the actual observed values. Its purpose is to detect biases. The Mean Absolute Error (MAE) gives a measure of the average magnitude of the errors. The usual alternative to the MAE is the Root Mean Square Error (RMSE), which also gives a measure of the average magnitude of the errors, but giving a larger weight to larger errors.

For positive variables, it is also common to evaluate the size of the errors as a proportion of the actual value of variable. This leads to the Mean Percent Error (MPE) and Mean Absolute Percent Error (MAPE), i.e., the mean of the percent errors and the mean of the absolute percent errors, respectively.

Finally, and also for use with positive variables, Theil's U is useful to evaluate whether a forecasting model is superior to naïve forecasts, i.e., to using the actual value at time t as the forecast of the value at time $t+1$. The forecasting model will be superior to naïve forecasts when Theil's U is less than one. When it equals one, then the forecast performance is identical. If Theil's U is larger than one, then the model performs worse than naïve forecasts.

To complement the comparison of the forecast performance based on the previous statistics, this study will also employ the test proposed by Diebold and Mariano (1995) on the null hypothesis that two forecast sets are equal. This test provides information about whether the difference in forecast performances is statistically significant or not.

Chapter 5

Empirical Results

5.1 Data and Preliminary Analysis

5.1.1 Data

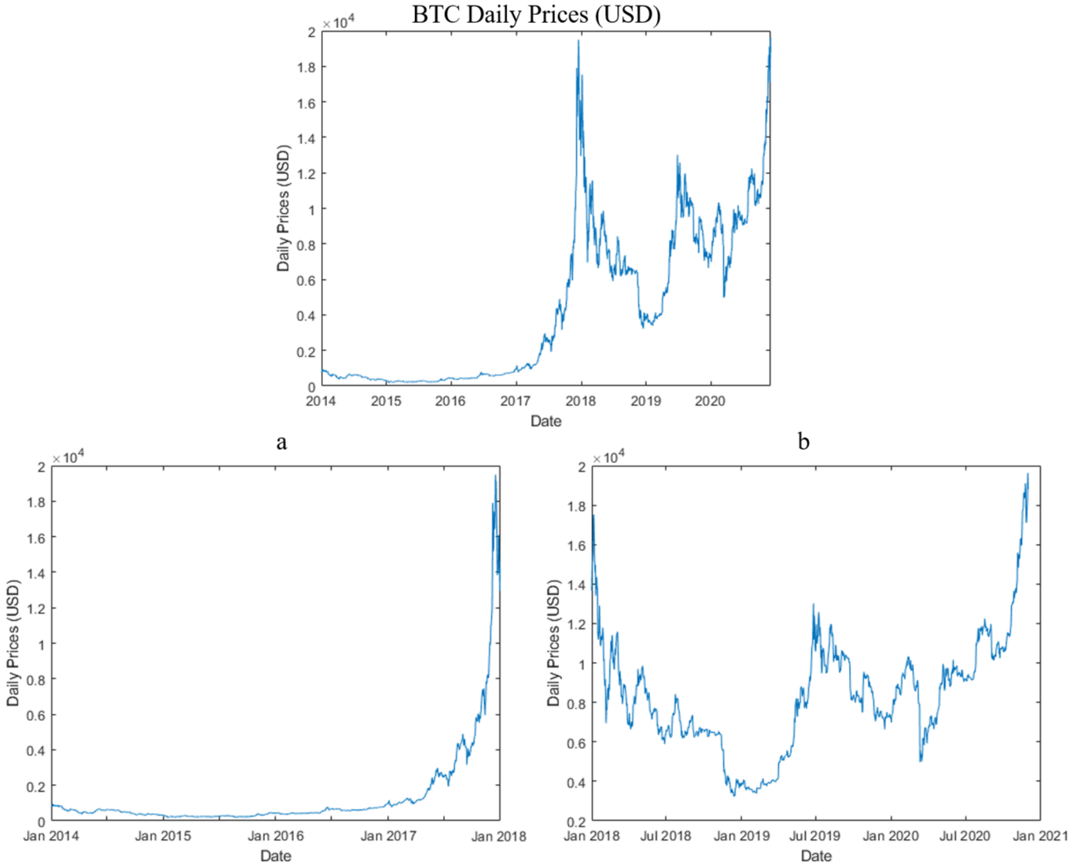
In this dissertation the object of study is the digital currency Bitcoin (BTC), in the period from January 3, 2014 to December 1, 2020 (which gives a total of 2525 daily observations). The period covered by the sample was divided into two sub-periods. The first is the "in-sample", and corresponds to the period from January 3, 2014 to December 31, 2017. The second sub-period is the "out-of-sample" period, which starts in January 1, 2018 and ends in December 1, 2020.

The data was collected from three sources: CoinMarketCap (https://coinmarketcap.com/), the bitcoincharts website (https://bitcoincharts.com/) and Coinmetrics (https://coinmetrics.io/).

CoinMarketCap provided daily data on close prices (i.e., the last reported prices before 00:00:00 UTC of the next day), daily high and low prices per day, and the daily trading volume, all denominated in USD. CoinMarketCap computes these series by aggregating the information from several exchanges, intending to give a comprehensive portrait of the overall Bitcoin market. For instance, the close prices are not trading prices but instead weighted average prices, with the weights given by the previous 24h market shares of the exchanges and are eligible. In this dissertation, the close prices are used to compute the daily log-returns, $r_t = \ln(P_t) - \ln(P_{t-1})$. The high and low daily prices are used to compute the Parkinson daily range-based volatility estimator, using the formula in Equation (4.3). Trading volume information is used as a predictor of volatility.

Figure 5.1 presents the evolution of Bitcoin prices. It displays an erratic behaviour marked by a long run positive trend. From the beginning of the period, January 2014, to the end of November 2017, at which time Bitcoin reached 8000USD, Bitcoin presents an approximately linear growth. From then on, an explosive growth behaviour is perceptible, with the Bitcoin price reaching almost 20000USD in mid December 2017. This could be related to the recognition granted in Japan to 11 firms as exchange houses of this digital currency in September of that year. However, the end of December was characterized by a sharp decline in the Bitcoin price, to around 14000USD, which was mainly attributed to the creation at that time of the Bitcoin futures by the the Chicago Board Options Exchange (CBOE) and the Chicago Mercantile Exchange (CME) (Sebastião and Godinho, 2020). The oscillations continued in the beginning of 2018, with accentuated growth followed by rapid decline. In 2020 the global pandemic caused by the COVID-19 virus had worldwide economic effects, including the markets for digital currencies. A large fall, to close to 5000USD, in the Bitcoin price was observed during the first half of March. In the rest of 2020, despite the fluctuations, the overall trend was clearly positive, leading to the high values registered at the end of 2020.

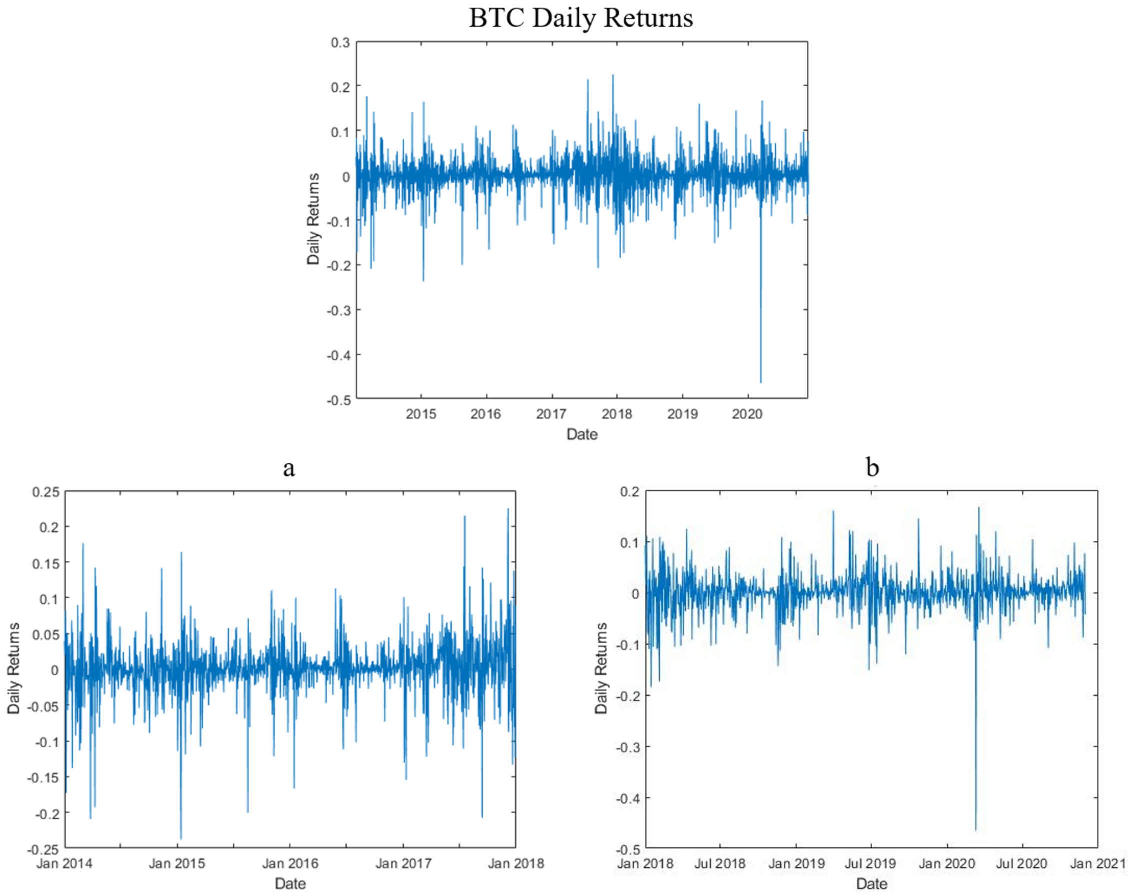
The evolution of returns is in Figure 5.2. An extreme negative drop around the time when the pandemic spread out worldwide is quite visible in Figure 5.2.



Notes: The top figure represents Bitcoin daily prices in USD in the full sample period, while (a) concerns the in-sample period and (b) the out-of-sample period.

Fig. 5.1 Bitcoin daily prices (USD)

The growth of Bitcoin trading volume in the final part of the in-sample accompanied that of the Bitcoin price — see Figure A.1 in Appendix A. In the out-of-sample period there also seems to be some association between the evolution of prices and the evolution of volume, for example, in March 2020. Thus, it is possible that trading volume contains information about the evolution of prices.



Notes: The top figure represents Bitcoin daily returns in USD in the full sample period, while (a) concerns the in-sample Period and (b) the out-of-sample period.

Fig. 5.2 Bitcoin daily returns

Figure 5.3 presents the evolution of the Parkinson estimator. The volatility displays strong oscillating movements throughout the full sample. Nevertheless, it is clearly visible the extreme peak in mid-March 2020, coinciding with the quick spread of the COVID-19 pandemic across the world. Besides that peak, special attention should also be given to the period centered around 2018, which is characterized by persistently high levels of volatility.

The second database used in this dissertation contains transaction data from the Bitstamp exchange. Bitstamp was chosen because of its reputation as one of the most important exchanges in the transmission of information on the USD/Bitcoin hourly prices since the bankruptcy of MtGox (Sebastião et al., 2018), because of its importance in terms of trading volume (Hileman and Rauchs, 2017), and because it is one of the exchanges from which the Chicago Mercantile Exchange computes the Bitcoin Reference Rate (BRR). The transaction prices were sampled at 5-minute intervals (considering the last known price at each sampling point) and used to compute the daily realized volatility, RV_t, according to Equation (4.2). Figure A.2 in Appendix A shows the evolution of this variable. As was the case for the Parkinson estimator, realized volatility also exhibits large fluctuations throughout the full sample.

There are two major peaks: the first in January 2015, and the second one in March 2020, similarly to the Parkinson measure.

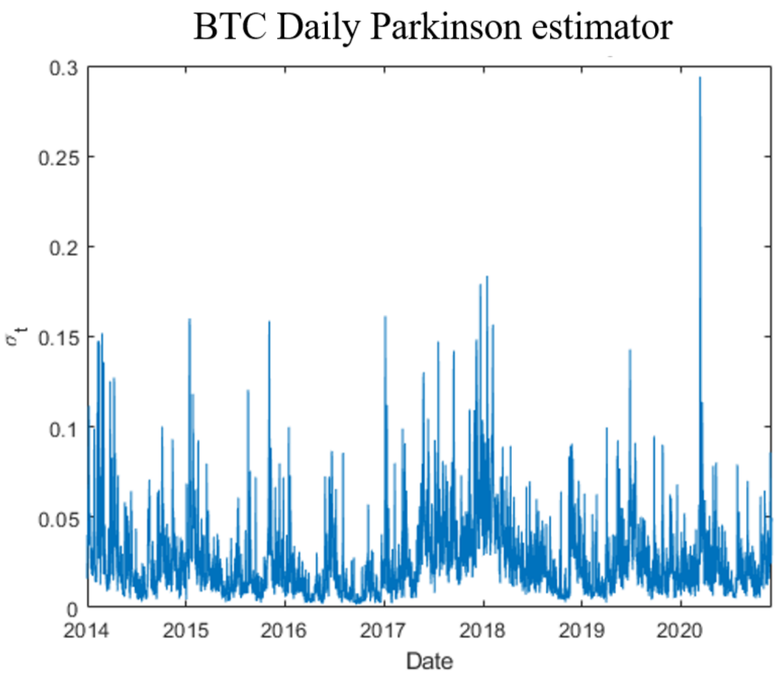


Fig. 5.3 Parkinson estimator

Finally, daily Blockchain information was collected from Coinmetrics. This dataset contains 42 variables (see Table A.1 in Appendix A, which describes all the 42 variables). These variables were subjected to visualization, and 10 were dropped out because they did not present almost any variability. Hence, the analysis proceeds with 32 Blockchain variables.

5.1.2 Stationarity Analysis and Selection of Blockchain's Variables

Before continuing with the analysis of the stationarity of trading volume and of the 32 blockchain variables, the variables that only assume positive values and are clearly skewed were transformed using logarithms. The stationarity analysis was conducted using the ADF test. The results are presented in Table 5.1. The variables are only tested in levels when at least one variant of the test suggests that they do not have a unit root at the 1% significance level; otherwise the first difference of the variables was also tested. The latter was the case of the trading volume; thus henceforth the first difference of the log of the trading volume, *d_log_Vol*, is used in the analysis.

Table 5.1 ADF unit root tests

		ADF				
	Log	Level (L)		Fist Difference (D)		L or D
		T	C	C	NC	
Trading Volume	Y	-3.06	-0.39	-13.06***	-12.93***	D
AdrActCnt	Y	-2.18	-1.94	-14.65***	-14.41***	D
BlkCnt	N	-6.89***	-6.37***	–	–	L
BlkSizeByte	Y	-2.43	-2.32	-16.16***	-16.00***	D
BlkSizeMeanByte	Y	-2.71	-2.61**	-15.73***	-15.57***	D
DiffMean	N	-0.83	1.80	-32.95***	-32.80***	D
FeeMeanNtv	N	-3.70**	-3.71***	–	–	L
FeeMeanUSD	N	-5.42***	-5.25***	–	–	L
FeeMedNtv	N	-3.85**	-3.85**	–	–	L
FeeMedUSD	N	-5.18***	-5.04***	–	–	L
FeeTotNtv	N	-3.64**	-3.60***	–	–	L
FeeTotUSD	N	-5.33***	-5.16***	–	–	L
HashRate	Y	-2.89	-4.66***	–	–	L
IssContNtv	N	-2.66	-1.39	-14.90***	-14.83***	D
IssContPctAnn	N	-2.44	-1.80	-15.37***	-15.23***	D
IssContUSD	N	-3.02	-1.89	-12.83***	-12.83***	D
IssTotNtv	N	-2.66	-1.39	-14.90***	-14.83***	D
IssTotUSD	N	-3.02	-1.89	-12.83***	-12.83***	D
NVTAdj	N	-2.69	-2.51	-15.98***	-15.98***	D
NVTAdj90	N	-3.21*	-2.68*	-50.76***	-50.77**	D
SplyFF	N	-1.12	-1.28	-8.50***	-7.47***	D
TxCnt	Y	-1.99	-2.04	-13.92***	-13.80***	D
TxTfrCnt	Y	-2.68	-2.24	-16.52***	-16.41***	D
TxTfrValAdjNtv	Y	-2.70	-2.71*	-15.62***	-15.62***	D
TxTfrValAdjUSD	Y	-2.38	-0.56	-13.49***	-13.41***	D
TxTfrValMeanNtv	N	-12.90***	-11.53***	–	–	L
TxTfrValMeanUSD	Y	-2.66	-1.29	-17.82***	-17.82***	D
TxTfrValMedNtv	N	-6.90***	-3.65***	–	–	L
TxTfrValMedUSD	N	-4.16***	-3.08**	–	–	L
TxTfrValNtv	N	-9.84***	-9.65***	–	–	L
TxTfrValUSD	Y	-2.38	-1.06	-15.39***	-15.35***	D

Notes: The number of lags included in the test regressions was chosen according to the AIC criterion. “T” identifies tests ran with a constant and a trend. “C” identifies tests ran with only a constant. “NC” identifies tests ran without a deterministic term. The null hypothesis of the ADF test is the existence of a unit root. Significance at the 1%, 5% and 10% levels is denoted by “***”, “**” and “*”, respectively.

Having decided whether to use the level or the first difference of the variables, the next step was to reduce the number of variables under analysis. The first selection criterion was based on the in-sample correlation between the lagged variable and the volatility measure $\ln(\sigma_t)$. Thus, only those variables for which the null hypothesis of no lagged correlation with log-volatility — i.e. $Corr(\ln(\sigma_t), X_{t-1}) = 0$ — was rejected at the 5% significance level were considered for inclusion in the volatility forecasting models. The correlations are presented in Table 5.2. According to the previous criterion, there are 11 candidates to be included in the volatility forecasting models. However, some of these variables are highly cross-correlated, which may result in a multicollinearity issue. At a first glance, this seems to be the case of the variables measuring transaction fees. To solve this problem, we use a second criterion: if two Blockchain variables have a cross-correlation equal or higher than 0.85, the one with the lowest correlation with $\ln(\sigma_t)$ is discarded. This pairwise analysis results in the following information set to be included in the models, ranked by descending order of correlation with $\ln(\sigma_t)$: TxTfrValMedUSD, FeeMeanNtv, FeeMeanUSD, NVTAdj90, BlkCnt, HashRate, and TxTfrValAdjNtv. The cross-correlation matrix is presented in Table 5.3 and the evolution of the seven selected Blockchain variables is in Appendix A, Figure A.3.

Table 5.2 Correlation between the Parkinson volatility series and lagged Blockchain variables

AdrActCnt	BlkCnt	BlkSizeByte	BlkSizeMeanByte	DiffMean
0.0161	0.0703***	-0.0005	0.0098	0.0072
FeeMeanNtv	FeeMeanUSD	FeeMedNtv*	FeeMedUSD	FeeTotNtv
0.2990***	0.2922***	0.2842***	0.2834***	0.2896***
FeeTotUSD	HashRate	IssContNtv	IssContPctAnn	IssContUSD
0.2751***	0.0488**	-0.0054	-0.0036	-0.0302
IssTotNtv	IssTotUSD	NVTAdj	NVTAdj90	SplyFF
-0.0054	-0.0302	-0.0354*	-0.0704***	0.0178
TxCnt	TxTfrCnt	TxTfrValAdjNtv*	TxTfrValAdjUSD	TxTfrValMeanNtv
0.0017	0.0074	0.0461**	0.0383*	0.0089
TxTfrValMeanUSD	TxTfrValMedNtv	TxTfrValMedUSD	TxTfrValNtv	TxTfrValUSD
0.0359*	0.0104	0.3142***	0.0254*	0.0382*

Table 5.3 Blockchain's variables crosscorrelation matrix

	TxTfrValMedUSD	FeeMeanNtv	FeeMeanUSD	NVTAdj90	BlkCnt	HashRate	TxTfrValAdjNtv
TxTfrValMedUSD	1						
FeeMeanNtv	0.5256	1					
FeeMeanUSD	0.8106	0.7856	1				
NVTAdj90	-0.0137	-0.0640	-0.0592	1			
BlkCnt	-0.0199	0.0033	0.0060	-0.0003	1		
HashRate	0.5954	0.0338	0.2208	0.0103	-0.2019	1	
TxTfrValAdjNtv	0.0455	0.0133	0.0042	-0.03208	-0.0426	-0.0030	1

5.1.3 Data Descriptive Analysis

The descriptive statistics for Bitcoin returns are presented in Table 5.4, for the full sample and the two subsamples. The most striking difference is in the mean: in the second sub-sample (the out-of-sample period) the mean is substantially smaller than in either the full sample or the first subsample. However, the variance does not change visibly across samples. The consequence is that the coefficient of variation does change markedly. The skewness and kurtosis statistics indicate that the distribution of returns is negatively skewed and leptokurtic. The non-normality of these distributions is corroborated by the large values of the Bera-Jarque test of the null hypothesis of normality, which is rejected at the 5% significance level.

Table 5.4 also shows the autocorrelation coefficients for lags 1 to 7, both for the returns and for the squared returns. The Ljung-Box test was used to test for joint residual autocorrelation up to lag 7. For the returns, there may be autocorrelation in the first subsample, if one uses the 1% significance level. For the squared returns, as expected, the test shows that there are significant ARCH effects in the residuals of the returns, in all samples.

The same analysis was applied to the natural logarithm of the Parkinson estimator — see Table 5.5. In this case, the descriptive statistics are very similar across subsamples. Again, the Jarque-Bera test rejects of the null hypothesis of normality at the 5% significance level for all samples. Finally, also as expected, there is clear evidence of strong positive autocorrelation in this series.

Table 5.6 shows the descriptive statistics for the first difference of the log of trading volume (*d_log_Vol*) and for the realized volatility in the Bitstamp exchange (RV_t). In both cases, the mean and the variance seem to decrease in the second subsample (the coefficient of variation does not change much in the case of realized volatility). Similarly to the previous variables, the Jarque-Bera test on again rejects of the null hypothesis of normality at the 5% significance level for these variables in all samples.

Table 5.4 Descriptive statistics of Bitcoin return series

	Returns (full data sample)	in-sample Returns	out-of-sample Returns
Observations	2525	1459	1066
Mean	0.0012	0.0020	0.0003
Median	0.0015	0.0018	0.0012
Std. Dev.	0.0392	0.0388	0.0398
Var. Coeff.	31.373	19.697	149.40
Minimum	-0.4647	-0.2376	-0.4647
Maximum	0.2251	0.2251	0.1671
Skewness	-0.9149	-0.3859	-1.5816
Kurtosis	15.491	9.7777	22.416
Jarque-Bera	16767	2828.8	17189
Autocorrelations Returns			
lag(1)	-0.0241	0.0090	-0.0661
lag(2)	-0.0154	-0.0804	0.0606
lag(3)	0.0179	0.0351	0.0006
lag(4)	0.0042	-0.0081	0.0169
lag(5)	0.0128	0.0030	0.0244
lag(6)	0.0458	0.0612	0.0371
lag(7)	-0.0293	-0.0265	-0.0364
$Q(7)$	10.802	18.018	12.458
$Q(7)$ p-value	0.1475	0.0119	0.0865
Autocorrelations Squared Returns			
lag(1)	0.1295	0.2673	0.0603
lag(2)	0.0429	0.0763	0.0254
lag(3)	0.0339	0.0669	0.0157
lag(4)	0.0713	0.0810	0.0669
lag(5)	0.0471	0.0837	0.0269
lag(6)	0.0406	0.0820	0.0175
lag(7)	0.0943	0.0547	0.1124
$Q^2(7)$	95.153	153.67	24.332
$Q^2(7)$ p-value	0.0000	0.0000	0.0010

Table 5.5 Descriptive statistics of the natural logarithm of the Parkinson estimator series

	full sample	in-sample	out-of-sample
Observations	2525	1459	1066
Mean	-3.9451	-3.9938	-3.8784
Median	-3.9226	-3.9660	-3.8710
Std. Dev.	0.8194	0.8789	0.7253
Var. Coeff.	-0.2077	-0.2201	-0.1870
Minimum	-6.4575	-6.4575	-5.8893
Maximum	-1.2244	-1.7200	-1.2244
Skewness	-0.0483	-0.0735	0.1431
Kurtosis	2.7595	2.6233	2.7077
Jarque-Bera	7.0691	9.9408	7.4303
Autocorrelations			
lag(1)	0.6289	0.6687	0.5416
lag(2)	0.5395	0.5838	0.4376
lag(3)	0.5345	0.5681	0.4532
lag(4)	0.5191	0.5323	0.4780
lag(5)	0.4725	0.4920	0.4129
lag(6)	0.4630	0.4834	0.4001
lag(7)	0.4670	0.4836	0.4106
$Q(7)$	4801.4	3081.1	1519.5
$Q(7)$ p-value	0.0000	0.0000	0.0000

Table 5.6 Descriptive statistics of Bitcoin trading volume and realized volatility in the full sample, in-sample, and out-of-sample periods

	d_log_Vol FS	*d_log_Vol* IS	*d_log_Vol* OS	RV_t FS	RV_t IS	RV_t OS
Observations	2525	1459	1066	2525	1459	1066
Mean	0.0028	0.0039	0.0013	0.0404	0.0434	0.0363
Median	-0.0140	-0.0188	-0.0109	0.0332	0.0360	0.0295
St. Dev.	0.3463	0.4266	0.1871	0.0271	0.0284	0.0246
Var. Coeff	122.07	108.18	141.58	0.6706	0.6544	0.6784
Minimum	-1.6634	-1.6634	-1.0214	0.0000	0.0000	0.0073
Maximum	2.4917	2.4917	0.8709	0.3460	0.3460	0.2972
Skewness	0.5206	0.4584	0.2813	3.0905	3.0752	3.1361
Kurtosis	6.5160	4.7454	5.4493	20.869	20.390	21.738
Jarque-Bera	1414.7	236.29	280.51	37612	20684	17342

Notes: FS, IS and OS refer to full sample;in-sample out-of-sample, respectively.

As for the 7 previously selected Blockchain variables, the descriptive statistics are in Tables A.2) (full sample), A.3) (in-sample) and A.4) (out-of-sample) in Appendix A. Some of the statistics show considerable variation across the subsamples. Again, the Jarque-Bera tests reject the null hypothesis of normality at the default 5% significance level for all variable in all subsamples.

5.2 Forecasting Models

The next step in this study is to compare the forecasting performance of the competing models. In the first sections the comparison is made within each class of models — The autorregressive models on the natural logarithm of the Parkinson estimator (AR-X) and the GARCH(1,1)-X models, with different sets of exogenous variables. The third section performs a comparison between selected models using the teste of Diebold and Mariano (1995).

5.2.1 AR-X Models

Firstly, the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the dependent variable, $\ln(\sigma_t)$, until lag 21 (3 weeks) were analyzed to assess the order of the lag structure (Figure B.1). According to the PACF there are 7 significant lags, which coincides with a calendar week. Having this into consideration, the same number of lags are initially considered for the other market variables, RV_t e d_log_Vol. When regressing the models on the 7 lags of the three variables, it was noticed that only the first 2 lags of d_log_Vol were significant. Hence we proceed running the several models with just 2 lag of d_log_Vol.

We estimated 32 AR-X (see Table B.2). All these models use 7 lags of the dependent variable and then use different combinations of the exogenous variables. The two last models deserve further explanation. Model ALL includes all the regressors (7 lags lags of the dependent variable and of realized volatility, 2 lags of trading volume and 1 lag of the 7 Blockchain variables). The Model SLCT is obtained from Model ALL by removing in a sequential order the regressors with higher p-value, until all the regressors in the final model have a p-value lower than 10%. That is, the variable that presented a higher p-value was removed, and the model was re-estimated, the process was iterated until all the remaining variables presented a significance level of, at least, 10%. This Model SLCT includes the 7 lags of the dependent variable, some of the lags of RV_t and 1 lag of four of the seven previously selected Blockchain variables.

Before we proceeded with the interpretation of the results on the metrics of the 1-step-ahead forecasts, we compared the adjusted coefficient of determination, $\bar{R}^2$, of the estimations in-sample. Model A, which is the most basic model achieved an $\bar{R}^2 = 0.5155$ meaning that just the 7 lags of the dependent variable explain more than 50% of the variation of the response variable around its mean. Model ALL and Model SLCT have an $\bar{R}^2 = 0.5310$ and an $\bar{R}^2 = 0.5313$, respectively. These are slight increases in relation to Model A, indicating that most probably the trading volume, realized volatility and Blockchain variables do not add significant information to forecast volatility, at least measured by the Parkinson estimator.

This claim is supported by the results presented in Table B.2. The AR-X models produce very similar results in terms of forecasting performance, with very small differences in the metrics (with a magnitude of only 10^{-3}), and without a clear winner. For instance the RMSE ranges from 0.5645 (Model R) to 0.5678 (Model P), the MAE ranges from 0.4380 (Model U) to 0.4459 (Model P), the MAPE ranges from 12.331 (Model R) to 12.477 (Model SLCT) and the Theil's U ranges from 0.7977 (Model R) to 0.8123 (Model SLCT). In a nutshell, the lag structure of the dependent variable captures almost all the forecasting potential, and the differences produced by introducing other explanatory variables are not statistically or economically significant.

5.2.2 GARCH(1,1)-X Models

Here the focus shifts into the GARCH(1,1)-X models, that is, on the conditional volatility and their one-step-ahead forecasts. Table B.3 presents the forecasting statistics for 28 GARCH models. The first model (Model A) is the GARCH(1,1), and the following models consider several combinations of the first lag of 1, 2 and 3 exogenous variables. The last one, Model ALL, includes all the relevant exogenous variables in the variance equation.

One should notice that too many exogenous variables in the variance equation increases non-trivially the potential for non-converge problems in the GARCH estimations. Hence in order to reduce the dimensionality of the predictors space we proceed in the following manner: We run the GARCH(1,1)-X using *d_log_Vol* as the only exogenous variable (Model B) and compute the RMSE statistic for one-step-ahead forecasts. Then we use this value as a threshold to select the other exogenous variables, i.e. if the RMSE of the GARCH(1,1)-X with just one of the other exogenous variables (Model C to Model I) is below that threshold, the variable is used afterward in other GARCH(1,1)-X models with more than one exogenous variable, otherwise that variable is discarded. The reason behind the choice of this threshold is the evidence provided in the literature that volume may have some predictive ability on volatility. After this additional filtering, besides RV_t, 4 Blockchain variables remain: FeeMeanNtv, NVTAdj90, BlkCnt and HashRate. All these models produce very similar results. In terms of RMSE the best one is Model C that considers RV_t as the unique exogenous variable and the worst one is Model B that considers *d_log_Vol*. Basically, the RV_t stands out as the best exogenous variable as the models that include it are the ones with the best forecasting performance. This is the case of models C, K, L and M. The Model ALL even has worst results than those models suggesting that other variables, especially *d_log_Vol*, introduce more noise than information on the estimations.

5.2.3 Statistical tests on the forecasting performance of different models

This subsection summarizes and provides a comparative analysis between the several models within and between the two different models' classes, namely by applying the test of Diebold and Mariano (1995) on the null hypothesis that one-step-ahead forecasts are equal.

The full set of AR-X and GARCH(1,1)-X models are presented in Appendix, Tables B.2 and B.3, respectively. The most important performance metrics for the base model, and the best and second-best models, according to the RMSE, in each class are presented in Table 5.7.

Table 5.7 Summary of the forecasting performance of selected models

	AR-X			GARCH(1,1)-X		
	Base Model	Best (Model R)	Second-Best (Model Q)	Base Model	Best (Model L)	Second-Best (Model K)
RMSE	0.5668	0.5645	0.5647	0.7990	0.7633	0.7649
MAE	0.4419	0.4397	0.4385	0.6659	0.6329	0.6340
MAPE	12.41%	12.33%	12.37%	16.47%	15.74%	15.76%
Theil's U	0.8086	0.7977	0.8090	1.0958	1.0456	1.0495

Notes: In the AR-X models, the base model corresponds to the model including 7 lags of the dependent variable, $\ln(\sigma_t)$, and in the GARCH(1,1)-X models to the GARCH(1,1) model. In the AR-X, the best model (Model R) correspond to a model with a predictors set formed by 7 lags of $\ln(\sigma_t)$, 2 lags of *d_log_Vol* and 1 lag of NVTAdj90, the second-best model (Model Q) uses 7 lags of $\ln(\sigma_t)$, 2 lags of *d_log_Vol*, and 1 lag of FeeMeanNtv. In the GARCH(1,1)-X, the best model (Model L) considers as exogenous variables RV_t and NVTAdj90, the second-best model (Model K) uses RV_t and FeeMeanNtv.

The results presented in Table 5.7 clearly show that AR-X models are superior to the GARCH(1,1)-X models, independently of the predictor sets used. This is quite striking even when the AR-X base model is compared with the Best GARCH(1,1)-X model: all the performance measures of former model are lower than the correspondent ones last model. It remains to know if these differences are statistically significant and if the smaller differences between models within each class are also statistically significant. This more refined analysis is presented in Table 5.8, wich presents the results of the Diebold-Mariano test.

Table 5.8 Diebold-Mariano Tests between Selected Models

	$D-M$	p-value
AR-X Base model vs Best AR-X model	1.2445	0.2133
AR-X Best model vs AR-X Second Best Model	-0.0776	0.9381
AR-X Base model vs GARCH(1,1)-X Base model	-14.103	0.0000
GARCH(1,1)-X Base model vs GARCH(1,1)-X Best model	13.663	0.0000
GARCH(1,1)-X Best model vs GARCH(1,1)-X Second-Best model	-1.0344	0.3009
Best AR-X model vs Best GARCH(1,1)-X model	-12.875	0.0000

Notes: See notes of the previous table on the predictor sets of the different models. The column D-M is the statistic of the Diebold and Mariano (1995) test on the null that two forecast sets are equal, and the last column refers to its p-value.

Several inferences can be made from the Diebold-Mariano statistics ($D-M$).

First, as it was already mentioned, the D-M test is not able to reject the null that the forecasts of the AR-X model and the Best AR-X model are equal at any usual significance level (the p-value is 0.2133) implying in this class of models the predictive potential is almost all captured by the lag structure of the dependent variable, $ln(\sigma_t)$, and hence the exogenous variable only contribute marginally to increase the forecasting performance of AR-X models.

Second, the contrary happens with the GARCH(1,1)-X models. The statistic on the test between the GARCH(1,1)-X Base model vs GARCH(1,1)-X Best model is 13.663, statistically significant at the 1% level. Hence within this class of models some exogenous variables have incremental information that can be used to forecast volatility, namely the realized volatility, RV_t, and NVTAdj90. This last variable corresponds to the "ratio of the network value (or market capitalization, current supply) to the 90-day moving average of the adjusted transfer value".

Third, the selection of Blockchain information to include in the GARCH(1,1)-X models should be made carefully, as different variables will have a non-trivial impact on the accuracy of the forecasts.

Fourth, the claim that the AR-X models are generally better than the GARCH(1,1)-X models is supported by the D-M statistic for the comparison between the AR-X Base model and GARCH(1,1)-X Base model ($D-M=-14.103$, significant at the 1% level) and by the D-M statistic for the comparison between the Best AR-X model vs Best GARCH(1,1)-X model ($D-M=-12.875$, significant at the 1% level).

Chapter 6

Conclusion

Launched in 2008, Bitcoin was the first cryptocurrency to solve the double-spending problem. It thrived and remains nowadays the most popular and influential digital currency.

This dissertation examines the several alternatives to forecast the volatility of the overall Bitcoin market, using daily data that spans from January 3, 2014 to December 1, 2020. We should reinforce that our goal is different from previous studies on Bitcoin volatility forecasting. These studies aim at forecasting the volatility of a particular online exchange, and for that matter the latent volatility is proxied by the realized volatility at that exchange. Here, we intend to forecast the overall Bitcoin volatility, which has the price input from several exchanges. Our research design is in fact based on two strong assumptions. First, it is assumed that CoinMarketCap provides reliable price information on the overall Bitcoin market, and second that its volatility is well measured by the Parkinson range-based estimator. With this in mind, we proceed by examine the forecasting performance of Autoregressive and GARCH models with exogenous variables obtained from the online market, namely the realized volatility of a liquid online exchange — Bitstamp, overall trading volume, and other variables obtained from the Blockchain network. The analysis is conducted on 1-step ahead forecasts, i.e. in a daily basis, using some procedures already documented in the literature for other more traditional markets, such the Forex (Pong et al., 2004; Taylor and Xu, 1997).

Our results show that while realized volatility may help explaining the overall volatility, trading volume has no incremental information value, which support the claims already presented elswere (see, for instance, Balcilar et al., 2017; Bouri et al., 2019). The incremental value of Realized Volatility and other carefully chosen Blockchain variables is especially important when using GARCH-type models. The most compelling result is that Autoregressive models are clearly superior to their GARCH models, and that the lag structure of the dependent variable is the main source of predictability. However, one should relative this result as the methodology used here is biased in this direction, that is while the Autoregressive models tackle directly the problem of forecasting the Parkinson estimator, the GARCH models intend to forecast the conditional volatility and hence only indirectly forecast that estimator. There are other analyses that could be made to reinforce or contradict our main conclusions. Use different forecasting horizons or improving on the GARCH forecasts by introducing a second step in which the forecasts of the conditional volatility are regressed on the realized Parkinson estimates, are among some of these additional analyses.

References

Aalborg, H. A., Molnár, P., and de Vries, J. E. (2019). What can explain the price, volatility and trading volume of bitcoin? *Finance Research Letters*, 29:255–265.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 39(4):885–905.

Ardia, D., Bluteau, K., and Rüede, M. (2019). Regime changes in bitcoin garch volatility dynamics. *Finance Research Letters*, 29:266–271.

Bação, P., Duarte, A. P., Sebastião, H., and Redzepagic, S. (2018). Information transmission between cryptocurrencies: does bitcoin rule the cryptocurrency world? *Scientific Annals of Economics and Business*, 65(2):97–117.

Balcilar, M., Bouri, E., Gupta, R., and Roubaud, D. (2017). Can volume predict bitcoin returns and volatility? a quantiles-based approach. *Economic Modelling*, 64:74–81.

Barber, S., Boyen, X., Shi, E., and Uzun, E. (2012). Bitter to better — how to make bitcoin a better currency. In *International conference on financial cryptography and data security*, volume 7397 of *Lecture Notes in Computer Science*, pages 399–414. Springer.

Baur, D. G. and Dimpfl, T. (2018). Asymmetric volatility in cryptocurrencies. *Economics Letters*, 173:148–151.

Bouri, E., Azzi, G., and Dyhrberg, A. H. (2016). On the return-volatility relationship in the bitcoin market around the price crash of 2013. *Available at SSRN 2869855*.

Bouri, E., Gkillas, K., Gupta, R., and Pierdzioch, C. (2021). Forecasting realized volatility of bitcoin: The role of the trade war. *Computational Economics*, 57:29–53.

Bouri, E., Lau, C. K. M., Lucey, B., and Roubaud, D. (2019). Trading volume and the predictability of return and volatility in the cryptocurrency market. *Finance Research Letters*, 29:340–346.

Chaim, P. and Laurini, M. P. (2019). Nonlinear dependence in cryptocurrency markets. *The North American Journal of Economics and Finance*, 48:32–47.

Chaum, D. (1983). Blind signatures for untraceable payments. In *Advances in cryptology*, pages 199–203. Springer.

Cheah, E.-T. and Fry, J. (2015). Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin. *Economics Letters*, 130:32–36.

Christensen, K. and Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349.

Conrad, C., Custovic, A., and Ghysels, E. (2018). Long-and short-term cryptocurrency volatility components: A garch-midas analysis. *Journal of Risk and Financial Management*, 11(2):23.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.

Fang, T., Su, Z., and Yin, L. (2020). Economic fundamentals or investor perceptions? the role of uncertainty in predicting long-term cryptocurrency volatility. *International Review of Financial Analysis*, 71:101566.

Fry, J. and Cheah, E.-T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47:343–352.

Hayek, F. A. (1990). *Denationalisation of money: the argument refined: an analysis of the theory and practice of concurrent currencies*, volume 70. Institute of economic affairs London.

Hileman, G. and Rauchs, M. (2017). Global cryptocurrency benchmarking study. *Cambridge Centre for Alternative Finance*, 33:33–113.

Karlstrøm, H. (2014). Do libertarians dream of electric coins? the material embeddedness of bitcoin. *Distinktion: Scandinavian Journal of Social Theory*, 15(1):23–36.

Klein, T., Thu, H. P., and Walther, T. (2018). Bitcoin is not the new gold–a comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, 59:105–116.

Köchling, G., Schmidtke, P., and Posch, P. N. (2020). Volatility forecasting accuracy for bitcoin. *Economics Letters*, 191:108836.

Lakhani, K. and Iansiti, M. (2017). The truth about blockchain. *Harvard Business Review*, 95(1):119–127.

Lansky, J. (2018). Possible state approaches to cryptocurrencies. *Journal of Systems Integration*, 9(1):19–31.

Liu, L. Y., Patton, A. J., and Sheppard, K. (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187(1):293–311.

Matkovskyy, R., Jalan, A., and Dowling, M. (2020). Effects of economic policy uncertainty shocks on the interdependence between bitcoin and traditional financial markets. *The Quarterly Review of Economics and Finance*, 77:150–155.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. online. *Available in https://bitcoin. org/bitcoin. pdf*.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business*, 53(1):61–65.

Pong, S., Shackleton, M. B., Taylor, S. J., and Xu, X. (2004). Forecasting currency volatility: A comparison of implied volatilities and ar (fi) ma models. *Journal of Banking & Finance*, 28(10):2541–2563.

Reid, F. and Harrigan, M. (2013). An analysis of anonymity in the bitcoin system. In *Security and privacy in social networks*, pages 197–223. Springer.

Sebastião, H., Duarte, A. P., and Guerreiro, G. (2018). Where is the information on usd/bitcoin hourly prices? *Notas Económicas*, 45:7–25.

Sebastião, H., Cunha, P., and Godinho, P. (2021). Cryptocurrencies and blockchain. overview and future perspectives. *International Journal of Economics and Business Research*, 21(3):305–342.

Sebastião, H. and Godinho, P. (2020). Bitcoin futures: An effective tool for hedging cryptocurrencies. *Finance Research Letters*, 33:101230.

Shen, D., Urquhart, A., and Wang, P. (2020). Forecasting the volatility of bitcoin: The importance of jumps and structural breaks. *European Financial Management*, 26(5):1294–1323.

Taylor, S. J. and Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, 4(4):317–340.

Trucíos, C. (2019). Forecasting bitcoin risk measures: A robust approach. *International Journal of Forecasting*, 35(3):836–847.

Walther, T., Klein, T., and Bouri, E. (2019). Exogenous drivers of bitcoin and cryptocurrency volatility–a mixed data sampling approach to forecasting. *Journal of International Financial Markets, Institutions and Money*, 63:101133.

Xu, X. and Taylor, S. J. (1995). Conditional volatility and the informational efficiency of the PHLX currency options market. *Journal of Banking & Finance*, 19(5):803–821.

Yang, S. Y. and Kim, J. (2015). Bitcoin market return and volatility forecasting using transaction network flow properties. In *2015 IEEE Symposium Series on Computational Intelligence*, pages 1778–1785. IEEE.

Appendix A

More on trading volume, realized volatility and Blockchain variables

Table A.1 Description of the 42 collected variables from Coinmetrics site (https://coinmarketcap.com/).

ID	Name	Description
AdrActCnt	Addresses, active, count	The sum count of unique addresses that were active in the network (either as a recipient or originator of a ledger change) that interval. All parties in a ledger change action (recipients and originators) are counted. Individual addresses are not double-counted if previously active.
BlkCnt	Block, count	The sum count of blocks created that interval that were included in the main (base) chain.
BlkSizeByte	Block, size, bytes	The sum of the size (in bytes) of all blocks created that interval.
BlkSizeMeanByte	Block, size, mean, bytes	The mean size (in bytes) of all blocks created that day.
BlkSizeMeanByte	Block, size, mean, bytes	The mean size (in bytes) of all blocks created that day.
CapMVRVCur	Capitalization, MVRV, current supply	The ratio of the sum USD value of the current supply to the sum "realized" USD value of the current supply.
CapMrktCurUSD	Capitalization, market, current supply, USD	The sum USD value of the current supply. Also referred to as network value or market capitalization.

Continues on the next page

ID	Name	Description
CapRealUSD	Capitalization, realized, USD	The sum USD value based on the USD closing price on the day that a native unit last moved (i.e., last transacted) for all native units.
DiffMean	Difficulty, mean	The mean difficulty of finding a hash that meets the protocol-designated requirement (i.e., the difficulty of finding a new block) that interval. The requirement is unique to each applicable cryptocurrency protocol. Difficulty is adjusted periodically by the protocol as a function of how much hashing power is being deployed by miners.
FeeMeanNtv	Fees, transaction, mean, native units	The mean fee per transaction in native units that interval.
FeeMeanUSD	Fees, transaction, mean, USD	The USD value of the mean fee per transaction that interval.
FeeMedNtv	Fees, transaction, median, native units	The median fee per transaction in native units that interval.
FeeMedUSD	Fees, transaction, median, USD	The USD value of the median fee per transaction that interval.
FeeTotNtv	Fees, total, native units	The sum of all fees paid to miners that interval. Fees do not include new issuance.
FeeTotUSD	Fees, total, USD	The sum USD value of all fees paid to miners that interval. Fees do not include new issuance.
HashRate	Hash rate, mean	The mean rate at which miners are solving hashes that interval. Hash rate is the speed at which computations are being completed across all miners in the network. The unit of measurement varies depending on the protocol.
IssContNtv	Issuance, continuous, native units	The sum of new native units issued that interval. Only those native units that are issued by a protocol-mandated continuous emission schedule are included (i.e., units manually released from escrow or otherwise disbursed are not included).
IssContPctAnn	Issuance, continuous, percent, annualized	The percentage of new native units (continuous) issued over that interval, extrapolated to one year (i.e., multiplied by 365), and divided by the current supply at the end of that interval. Also referred to as the annual inflation rate.

Continues on the next page

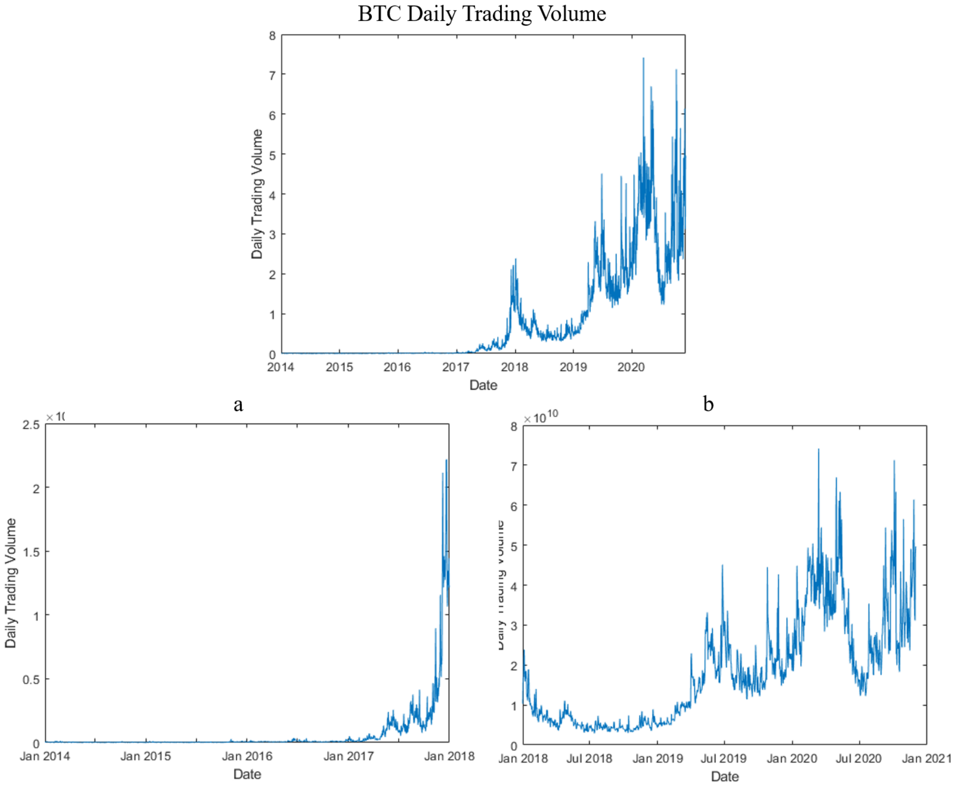
ID	Name	Description
IssContUSD	Issuance, continuous, USD	The sum USD value of new native units issued that interval. Only those native units that are issued by a protocol-mandated continuous emission schedule are included (i.e., units manually released from escrow or otherwise disbursed are not included).
IssTotNtv	Issuance, total, native units	The sum of all new native units issued that interval.
IssTotUSD	Issuance, total, USD	The sum USD value of all new native units issued that interval.
NVTAdj	NVT, adjusted	The ratio of the network value (or market capitalization, current supply) divided by the adjusted transfer value. Also referred to as NVT.
NVTAdj90	NVT, adjusted, 90d MA	The ratio of the network value (or market capitalization, current supply) to the 90-day moving average of the adjusted transfer value. Also referred to as NVT.
PriceBTC	Price, BTC	The fixed closing price of the asset as of 00:00 UTC the following day (i.e., midnight UTC of the current day) for end-of-day data or the closest prior hour (nearest to that block) for block-by-block data, denominated in BTC.
PriceUSD	Price, USD	The fixed closing price of the asset as of 00:00 UTC the following day (i.e., midnight UTC of the current day) denominated in USD. This price is generated by Coin Metrics' fixing/reference rate service. Real-time PriceUSD is the fixed closing price of the asset as of the timestamp set by the block's miner.
ROI1yr	ROI, percent, 1yr	The return on investment for the asset assuming a purchase 12 months prior.
ROI30d	ROI, percent, 30d	The return on investment for the asset assuming a purchase 30 days prior.
SplyCur	Supply, current	The sum of all native units ever created and visible on the ledger (i.e., issued) at the end of that interval. For account-based protocols, only accounts with positive balances are counted.

Continues on the next page

ID	Name	Description
SplyExpFut10yrCMBI	Supply, future expected, next 10yr (CMBI)	The sum of all native units counting current supply and including all those expected to be issued over the next 10 years from that interval if the current known continuous issuance schedule is followed. Future expected hard-forks that will change the continuous issuance are not considered until the day they are activated/enforced.
SplyFF	Supply, free float	Free float supply refers to the number of native units of a crypto asset that are readily available to trade in open markets (i.e. not restricted) at the end of the time interval. It includes all native units visible on the ledger minus company, foundation and founding team native units.
TxCnt	Transactions, count	The sum count of transactions that interval. Transactions represent a bundle of intended actions to alter the ledger initiated by a user (human or machine). Transactions are counted whether they execute or not and whether they result in the transfer of native units or not (a transaction can result in no, one, or many transfers). Changes to the ledger mandated by the protocol (and not by a user) or post-launch new issuance issued by a founder or controlling entity are not included here.
TxTfrValAdjNtv	Transactions, transfers, value, adjusted, native units	The sum of native units transferred between distinct addresses that interval removing noise and certain artifacts.
TxTfrValAdjUSD	Transactions, transfers, value, adjusted, USD	The USD value of the sum of native units transferred between distinct addresses that interval removing noise and certain artifacts.
TxTfrValMeanNtv	Transactions, transfers, value, mean, native units	The mean count of native units transferred per transaction (i.e., the mean "size" of a transaction) between distinct addresses that interval.
TxTfrValMeanUSD	Transactions, transfers, value, mean, USD	The sum USD value of native units transferred divided by the count of transfers (i.e., the mean "size" in USD of a transfer) between distinct addresses that interval.

Continues on the next page

ID	Name	Description
TxTfrValMedNtv	Transactions, transfers, value, median, native units	The median count of native units transferred per transfer (i.e., the median "size" of a transfer) between distinct addresses that interval.
TxTfrValMedUSD	Transactions, transfers, value, median, USD	The median USD value transferred per transfer (i.e., the median "size" in USD of a transfer) between distinct addresses that interval.
TxTfrValUSD	Transactions, transfers, value, USD	The sum USD value of all native units transferred (i.e., the aggregate size in USD of all transfers) between distinct addresses that interval.
VtyDayRet180d	Volatility, daily returns, 180d	The 180D volatility, measured as the standard deviation of the natural log of daily returns over the past 180 days.
VtyDayRet30d	Volatility, daily returns, 30d	The 30D volatility, measured as the standard deviation of the natural log of daily returns over the past 30 days.
VtyDayRet60d	Volatility, daily returns, 60d	The 60D volatility, measured as the standard deviation of the natural log of daily returns over the past 60 days.
TxTfrCnt	Transactions, transfers, count	The sum count of transfers that interval. Transfers represent movements of native units from one ledger entity to another distinct ledger entity. Only transfers that are the result of a transaction and that have a positive (non-zero) value are counted.



Notes: The top figure represents BTC Trading Volume in the full sample period, while (a) concerns the in-sample Period and (b) the out-of-sample period.

Fig. A.1 BTC Trading Volume

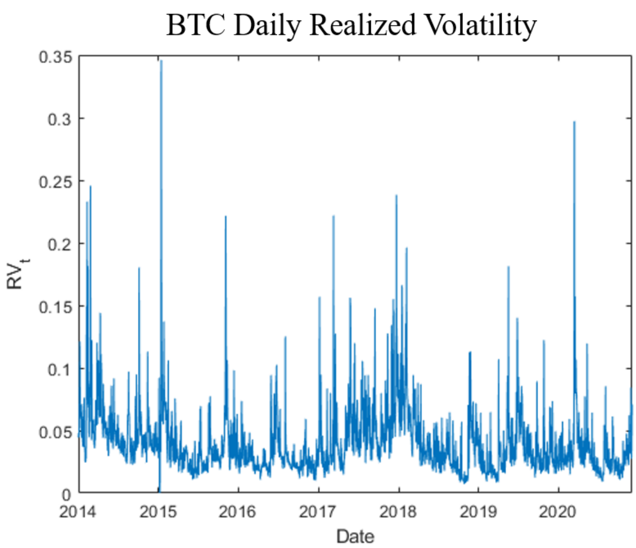


Fig. A.2 BTC Realized Volatility of the Bitstamp exchange

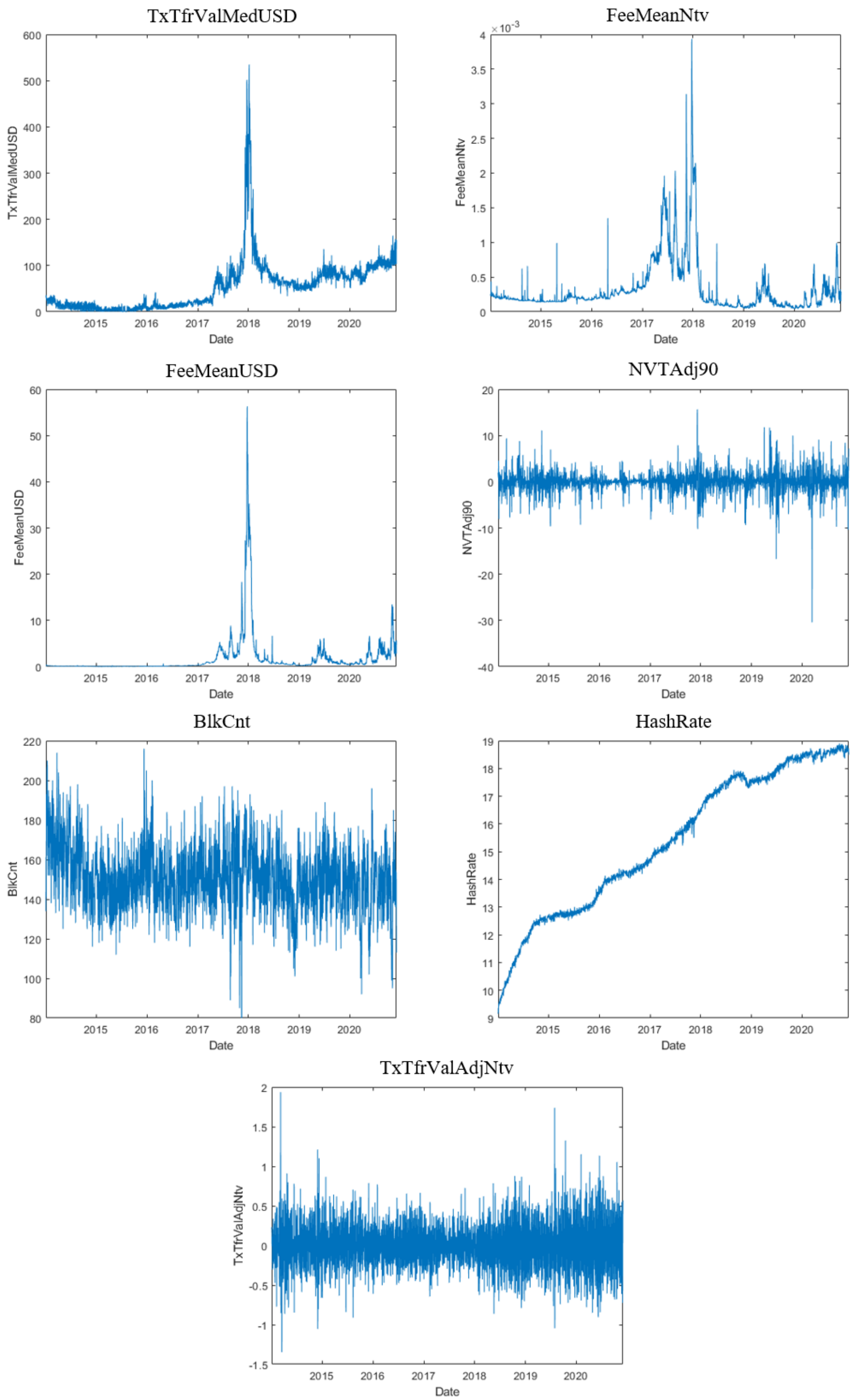


Fig. A.3 Evolution of the 7 previously selected Blockchain's variables

Table A.2 Descriptive statistics of the 7 previously selected Blockchain's variables

	a)	b)	c)	d)	e)	f)	g)
Observations	2525	2525	2525	2525	2525	2525	2525
Mean	54.433	0.0003	1.6806	0.0212	150.97	15.389	0.0003
Median	48.609	0.0002	0.4235	0.0850	151.00	15.420	-0.0142
Std. Dev.	56.609	0.0004	4.4718	2.3932	16.609	2.5329	0.2972
Var. Coeff.	1.0400	1.2049	2.6608	112.89	0.1100	0.1646	1030.9
Minimum	0.0023	0.0265	0.0243	-30.419	80.000	9.1511	-1.3464
Maximum	534.85	0.0039	56.308	15.660	216.00	18.905	1.9368
Skewness	3.1421	3.4790	6.4458	-0.9104	0.0788	-0.2990	0.3651
Kurtosis	20.236	18.829	53.757	18.146	3.6997	1.8884	4.9923
Jarque-Bera	35410	31455	288530	24482	54.1258	167.62	473.71

Notes: Blockchain variables are here denoted as a) TxTfrValMedUSD; b) FeeMeanNtv; c) FeeMeanUSD; d) NVTAdj90; e) BlkCnt; f) HashRate; g) TxTfrValAdjNtv.

Table A.3 Descriptive statistics of the 7 previously selected Blockchain's variables for the in-sample period

	a)	b)	c)	d)	e)	f)	g)
Observations	1459	1459	1459	1459	1459	1459	1459
Mean	28.296	0.0004	1.2787	-0.0016	153.25	13.522	0.1513
Median	15.210	0.0394	0.1081	0.0643	152.00	13.531	0.0008
Std. Dev.	47.142	0.0004	4.6613	2.0629	16.908	1.6139	0.2666
Var. Coeff.	1.6660	1.0791	3.6452	25.487	0.1103	0.1194	322.75
Minimum	0.0023	0.0001	0.0243	-10.209	80.00	9.1511	-1.3464
Maximum	501.84	0.0039	56.308	15.660	216	16.564	1.9368
Skewness	5.3616	3.1699	7.3368	0.0219	0.1962	-0.3139	0.3801
Kurtosis	40.606	16.493	65.826	10.155	3.8115	2.6100	20.236
Jarque-Bera	92963	13511	253040	3111.8	49.4021	33.2016	596.83

Notes: Blockchain variables are here denoted as: a) TxTfrValMedUSD; b) FeeMeanNtv; c) FeeMeanUSD; d) NVTAdj90; e) BlkCnt; f) HashRate; g) TxTfrValAdjNtv.

Table A.4 Descriptive statistics of the 7 previously selected Blockchain's variables for the out-of-sample period

	a)	b)	c)	d)	e)	f)	g)
Observations	1066	1066	1066	1066	1066	1066	1066
Mean	90.207	0.0002	2.2307	0.0524	147.84	17.944	-0.0004
Median	82.544	0.0001	0.9882	0.1591	149	17.879	-0.0135
Std. Dev.	48.326	0.0003	4.1380	2.7833	15.668	0.5758	0.3346
Var. Coeff.	0.5357	1.3561	1.8550	53.075	0.1060	0.0321	-746.82
Minimum	33.683	0.0000	0.1613	-30.419	92.000	16.435	-1.0441
Maximum	534.85	0.0021	35.268	11.801	196.00	18.9047	1.7386
Skewness	4.8616	4.3441	5.0328	-1.3994	-0.2102	-0.3175	0.3454
Kurtosis	34.347	24.512	31.420	19.387	3.1906	2.0935	4.0314
Jarque-Bera	47844	23908	40376	12276	9.4634	54.408	68.449

Notes: Blockchain variables are here denoted as: a) TxTfrValMedUSD; b) FeeMeanNtv; c) FeeMeanUSD; d) NVTAdj90; e) BlkCnt; f) HashRate; g) TxTfrValAdjNtv.

Appendix B

More on forecasting models

Table B.1 Autocorrelation function for the Log Parkinson

LAG	ACF	PACF	Q-stat.
1	0.6687***	0.6687 ***	653.7 ***
2	0.5838 ***	0.2473 ***	1152 ***
3	0.5681 ***	0.2099 ***	1625 ***
4	0.5323 ***	0.1064 ***	2040 ***
5	0.4920 ***	0.0532 **	2395 ***
6	0.4834 ***	0.0812 ***	2738 ***
7	0.4836 ***	0.0888 ***	3081 ***
8	0.4552 ***	0.0277	3385 ***
9	0.4298 ***	0.0132	3657 ***
10	0.4323 ***	0.0530 **	3932 ***
11	0.4149 ***	0.0188	4185 ***
12	0.3879 ***	-0.0037	4407 ***
13	0.3959 ***	0.0500 *	4638 ***
14	0.3988 ***	0.0446 *	4873 ***
15	0.4118 ***	0.0768 ***	5123 ***
16	0.3605 ***	-0.0602 **	5315 ***
17	0.3559 ***	0.0066	5502 ***
18	0.3383 ***	-0.0199	5671 ***
19	0.3159 ***	-0.0158	5819 ***
20	0.3207 ***	0.0292	5971 ***
21	0.3105 ***	-0.0036	6114 ***

Notes: "***", "**" and "*" indicate significance at the 1%, 5% and 10% levels, respectively. These p-values were computed using standard error computed as $1/T^{0.5}$, where T is the number of observations.

Table B.2 Forecasting Performance of AR-X models

Model	ME	RMSE	MAE	MPE	MAPE	Theil's U
A	0.0092	0.5668	0.4419	-2.6361	12.4090	0.8086
B	0.0095	0.5655	0.4412	-2.6501	12.395	0.8097
C	0.0098	0.5677	0.4439	-2.6220	12.460	0.8068
D	0.0099	0.5668	0.4423	-2.5786	12.375	0.8029
E	-0.0349	0.5667	0.4455	-1.4195	12.346	0.8112
F	0.0277	0.5660	0.4392	-3.1197	12.388	0.8077
G	0.0056	0.5655	0.4405	-2.5228	12.352	0.8076
H	0.0103	0.5662	0.4406	-2.6404	12.354	0.7987
I	0.0065	0.5669	0.4422	-2.5595	12.408	0.8086
J	-0.0113	0.5672	0.4442	-2.0789	12.404	0.8093
K	0.0092	0.5667	0.4428	-2.6283	12.427	0.8074
L	0.0101	0.5666	0.4432	-2.6395	12.446	0.8073
M	0.0236	0.5669	0.4414	-2.9877	12.429	0.8060
N	0.0098	0.5675	0.4437	-2.6102	12.443	0.8030
O	0.0076	0.5677	0.4442	-2.5574	12.460	0.8069
P	-0.0104	0.5678	0.4459	-2.0705	12.448	0.8079
Q	0.0279	0.5647	0.4385	-3.1291	12.374	0.8090
R	0.0107	0.5645	0.4397	-2.6556	12.331	0.7977
S	0.0067	0.5655	0.4416	-2.5699	12.392	0.8096
T	-0.0100	0.5658	0.4436	-2.1199	12.394	0.8101
U	0.0277	0.5655	0.4380	-3.0966	12.332	0.7983
V	0.0257	0.5661	0.4394	-3.0598	12.384	0.8077
W	0.0259	0.5673	0.4404	-3.0941	12.419	0.8102
X	0.0078	0.5663	0.4409	-2.5691	12.354	0.7988
Y	-0.0109	0.5666	0.4430	-2.0669	12.350	0.7996
Z	-0.0117	0.5673	0.4443	-2.0647	12.405	0.8093
AA	0.0239	0.5659	0.4410	-3.0038	12.421	0.8067
AB	0.0101	0.5662	0.4426	-2.6263	12.415	0.8018
AC	0.0079	0.5667	0.4436	-2.5732	12.446	0.8073
AD	-0.0095	0.5668	0.4453	-2.1010	12.437	0.8080
ALL	0.0111	0.5674	0.4439	-2.6172	12.441	0.8082
SLCT	0.0132	0.5680	0.4446	-2.6842	12.477	0.8123

Notes: All AR-X models bellow include 7 lags of the dependent variable, the Parkinson estimator, $ln(\sigma_t)$. The exogenous variables considered in each model are listed bellow. Model A is the baseline model that includes only the lags of the dependent variable.

Model A: n.a.;

Model B: 2 lags of *d_log_Vol*;

Model C: 7 lags of the RV_t ;

Model D: 1 lag of each of the 7 Blockchain variables;

Model E: 1 lag of TxTfrValMedUSD;

Model F: 1 lag of FeeMeanNtv;

Model G: 1 lag of FeeMeanUSD;

Model H: 1 lag of NVTAdj90;

Model I: 1 lag of BlkCnt;

Model J: 1 lag of HashRate;

Model K: 1 lag of TxTfrValAdjNtv;

Model L: 7 lags of the RV_t and 2 lags of *d_log_Vol*;

Model M: 7 lags RV_t and 1 lag of FeeMeanNtv;

Model N: 7 lags of the RV_t and 1 lag of NVTAdj90;

Model O: 7 lags of the RV_t and 1 lag of BlkCnt;

Model P: 7 lags of the RV_t and 1 lag of HashRate;

Model Q: 2 lags of *d_log_Vol* and 1 lag of FeeMeanNtv;

Model R: 2 lags of *d_log_Vol* and 1 lag of NVTAdj90;

Model S: 2 lags of *d_log_Vol* and 1 lag of the BlkCnt;

Model T: 2 lags of *d_log_Vol* and 1 lag of HashRate;

Model U: 1 lag of FeeMeanNtv and NVTAdj90;

Model V: 1 lag of FeeMeanNtv and BlkCnt;

Model W: 1 lag of FeeMeanNtv and HashRate;

Model X: 1 lag of NVTAdj90 and BlkCnt;

Model Y: 1 lag of NVTAdj90 and HashRate;

Model Z: 1 lag of BlkCnt and HashRate;

Model AA: 7 lags of the RV_t, 2 lags of *d_log_Vol* and 1 lag of FeeMeanNtv;

Model AB: 7 lags of the RV_t, 2 lags of *d_log_Vol* and 1 lag of NVTAdj90;

Model AC: 7 lags of the RV_t, 2 lags of *d_log_Vol* and 1 lag of BlkCnt;

Model AD: 7 lags of the RV_t, 2 lags of *d_log_Vol* and 1 lag of HashRate;

Model ALL: 7 lags of the RV_t variable, 2 lags of *d_log_Vol* and 1 lag of each of the 7 Blockchain variables;;

Model SLCT: Model selected from Model ALL by applying sequentially the rule of omitting one-by-one the variables with a p-value higher than 10%.

Table B.3 Forecasting Performance of GARCH(1,1)-X models

Model	ME	RMSE	MAE	MPE	MAPE	Theil's U
A	-0.5506	0.7990	0.6659	0.1197	0.1647	1.0958
B	-0.5554	0.8035	0.6699	0.1209	0.1659	1.1024
C	-0.5114	0.7716	0.6409	0.1091	0.1591	1.0580
D	-0.6344	0.8631	0.7273	0.1409	0.1782	1.1715
E	-0.5368	0.7888	0.6564	0.1162	0.1627	1.0829
F	-0.5819	0.8192	0.6863	0.1279	0.1690	1.1188
G	-0.5447	0.7959	0.6633	0.1182	0.1642	1.0914
H	-0.5373	0.7916	0.6585	0.1160	0.1631	1.0848
I	-0.5401	0.7953	0.6607	0.1170	0.1639	1.0915
J	-0.5380	0.7896	0.6573	0.1164	0.1630	1.0838
K	-0.5010	0.7649	0.6340	0.1065	0.1576	1.0495
L	-0.4962	0.7633	0.6329	0.1051	0.1574	1.0456
M	-0.5094	0.7691	0.6397	0.10867	0.1589	1.0540
N	-0.5384	0.7943	0.6607	0.1157	0.1634	1.0828
O	-0.5536	0.8024	0.6688	0.1205	0.1657	1.1018
P	-0.5410	0.8005	0.6658	0.1168	0.1652	1.0967
Q	-0.5546	0.8029	0.6697	0.1208	0.1658	1.1015
R	-0.5446	0.8005	0.6643	0.1181	0.1649	1.0989
S	-0.5316	0.7861	0.6541	0.1149	0.1622	1.0788
T	-0.5209	0.7807	0.6477	0.1118	0.1608	1.0709
U	-0.5030	0.7817	0.6447	0.1070	0.1606	1.0722
V	-0.5260	0.7856	0.6532	0.1130	0.1620	1.0762
W	-0.5454	0.8003	0.6657	0.1183	0.16501	1.0967
X	-0.5491	0.8047	0.6692	0.1188	0.1656	1.0998
Y	-0.5346	0.7878	0.6552	0.1156	0.1626	1.0820
Z	-0.5083	0.7737	0.6414	0.1082	0.1595	1.0600
AA	-0.5337	0.7852	0.6544	0.1153	0.1623	1.0768
ALL	-0.5348	0.8018	0.6637	0.1145	0.1644	1.0921

Notes: This table only presents successful estimations. Although other combinations of exogenous variables are possible, some of these combinations are not presented because the corresponding GARCH(1,1) did not converge at some point.The exogenous variables considered in each successful model are listed bellow. Notice that the variance equation only consider the first lag of these variables. Model A is the baseline model corresponding to a GARCH(1,1).

Model A: n.a.;

Model B: *d_log_Vol*;

Model C: RV_t;

Model D:TxTfrValMedUSD;

Model E: FeeMeanNtv;
Model F: FeeMeanUSD;
Model G: NVTAdj90;
Model H: BlkCnt;
Model I: HashRate;
Model J: *d_log_Vol* and RV_t;
Model K: RV_t and FeeMeanNtv;
Model L: RV_t and NVTAdj90;
Model M: RV_t and BlkCnt;
Model N: RV_t and HashRate;
Model O: *d_log_Vol* and FeeMeanNtv;
Model P: *d_log_Vol* and NVTAdj90;
Model Q: *d_log_Vol* and BlkCnt;
Model R: *d_log_Vol* and HashRate;
Model S: FeeMeanNtv and NVTAdj90;
Model T: FeeMeanNtv and BlkCnt;
Model U: FeeMeanNtv and HashRate;
Model V: NVTAdj90 and BlkCnt;
Model W: NVTAdj90 and HashRate;
Model X: BlkCnt and HashRate;
Model Y: *d_log_Vol* and RV_t and FeeMeanNtv as;
Model Z: *d_log_Vol*, RV_t and NVTAdj90 ;
Model AA: *d_log_Vol*, RV_t and BlkCnt ;
Model ALL: *d_log_Vol*, RV_t and 4 Blockchain variables (FeeMeanNtv, NVTAdj90, BlkCnt and HashRate).