Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Econometric Analysis of Bitcoin and its 2013 Bubbles

Author: **Pavel Fišer**

Supervisor: **PhDr. Jiří Kukačka**

Academic Year: **2014/2015**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author hereby declares that this thesis has not been used to obtain another university degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 1, 2015

Signature

Acknowledgments

I am especially grateful to my supervisor PhDr. Jiří Kukačka for his guidance and devoted time during the creation of this thesis. Especially, for his willingness to answer all my questions and to discuss the related topic.

I would also like to thank my family and my friends for their support during my university studies. The usual caveat applies.

Abstract

This thesis examines Bitcoin in 2012-2015 period along with the two Bitcoin bubbles — April 2013 and November 2013 — using ARIMA, GARCH and LPPL models. First, we perform standard GARCH analysis along with GARCH rolling estimation and find that the volatility of Bitcoin differs substantially over time and that this relation is best captured by GARCH(1,1) in all studied periods. We also conclude that during the November bubble the number of irrational traders entering the market was much higher than in the April bubble which probably caused greater instability on the Bitcoin market. However, based on Ljung-box test we find these results to be questionable. For that reason, we present LPPL model and study its key parameters — power law growth rate β, frequency of log oscillation ω and its scaling ratio λ — in more detail using standard methodology and "loop analysis". We find that the November bubble experiences much faster oscillation and lower acceleration rate of power law in comparison with the April bubble. By the end we propose hypothesis that $\Delta\lambda$ serves as a better indicator of the upcoming bubble crash than simple scaling ratio which we concluded to be true in our analysis of the two Bitcoin bubbles. However, further examination of other financial bubbles is needed, in order to support this hypothesis.

JEL Classification	C22, C53, C58, E47, G17
Keywords	Bitcoin, bubble, LPPL model, scaling ratio, log-periodicity
Number of words	13532

Author's e-mail	pauli.fiser@seznam.cz
Supervisor's e-mail	jiri.kukacka@fsv.cuni.cz

Abstrakt

V bakalářské práci zkoumáme pomocí ARIMA, GARCH a LPPL modelů tři periody Bitcoinové řady mezi roky 2012-2015 — kompletní řadu, dubnovou bublinu v roce 2013 a listopadovou bublinu v roce 2013. Nejdříve provádíme standardní GARCH analýzu, na niž navazuje GARCH rolující estimace. Výsledky těchto analýz ukázaly, že volatilita Bitcoinu se v průběhu zkoumané periody výrazně liší, a že pro všechny tři časové periody se jako nejlepší specifikace zachycující volatilitu ukázala GARCH(1,1). Dále jsme vypozorovali, že v průběhu listopadové bubliny byl výrazně větší počet nově vstupujících iracionálních obchodníků na trh než v průběhu dubnové bubliny, což pravděpodobně způsobilo větší nestabilitu na trhu s Bitcoiny. Nicméně výsledky Ljung-box testu částečně zpochybňují konzistentnost odhadnutých parametrů GARCH analýzou. Z tohoto důvodu používáme LPPL model a zkoumáme jeho klíčové proměnné — rychlost růstu β, frekvence log oscialace ω spolu se “scaling ratio” λ — nejdříve standardní metodologií a poté “loop analýzou”. Z výsledků jsme vyvodili, že listopadová bublina měla výrazně rychlejší oscilaci a nižší akceleraci růstu ve srovnání s dubnovou bublinou. Na závěr přicházíme s hypotézou, že $\Delta\lambda$ slouží jako lepší indikátor nadcházejícího krachu bubliny než prosté “scaling ratio”. Výsledky naší analýzy dvou bublin tuto hypotézu potvrzují, nicméně jsme si vědomi toho, že k jejímu plnohodnotné potvrzení je potřeba zkoumání dalších finančních bublin.

Klasifikace JEL	C22, C53, C58, E47, G17
Klíčová slova	Bitcoin, bublina, LPPL model, scaling ratio, log-periodicita
Počet slov	13532

E-mail autora	`pauli.fiser@seznam.cz`
E-mail vedoucího práce	`jiri.kukacka@fsv.cuni.cz`

Contents

List of Tables

List of Figures

Acronyms

ACF	Autocorrelation Function
ADF	Augmented Dickey Fuller test
AIC	Akaike Information Criterion
APPL	Apple Inc.
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Schwarz Bayesian Information Criterion
BPI	Bitcoin Price Index
BRL	Brazilian Real
BTC	Bitcoin
CAD	Canadian Dollar
CHF	Swiss Franc
CNY	Chinese Yuan Renminbi
DEM	Deutsche Mark (former German currency)
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
EUR	Euro
FOREX	Foreign Exchange
FTSE100	Financial Times Stock Exchange index
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
JPY	Japanese Yen
KPSS	Kwiatkowski, Phillips, Schmidt and Shin test
LPPL	Log Periodic Power Law

MA	Moving Average
MXN	Mexican Peso
NASDAQ	National Association of Securities Dealers Automated Quotations Composite Index
NIKKEI	Nihon Keizai Shimbun (Japanese stock market index)
PACF	Partial Autocorrelation Function
RMSE	Root Mean Squared Error
RUB	Russian Rubble
S&P 500	Standard and Poor's 500 index
TGARCH	Threshold Generalized Autoregressive Conditional Heteroskedasticity
USD	United States Dollar
vol-of-vol	Volatility of Volatility

Bachelor Thesis Proposal

Author	Pavel Fišer
Supervisor	PhDr. Jiří Kukačka
Proposed topic	Econometric Analysis of Bitcoin and its 2013 Bubbles

Topic characteristics In this thesis I would like to focus on phenomena of Bitcoin and econometric analysis in time period of years 2012 and 2013 using standard econometric tools. I will try to analyze Bitcoin price development during this period in "standard development" time, but also in two crashes, occurred on April 2013 and on November 2013, during the period of excessive boom and subsequent crash, then compare them and examine, whether they have similar characteristics or not. Secondly I will examine whether statistical features of Bitcoin differ from statistical features of FOREX currencies stocks and indexes.

Hypotheses

1. The two crashes, occurring in 2013, have experienced similar progress during their period of boom and crash.
2. With certain probability, we are able to predict price development of Bitcoin in a horizon of days.
3. Bitcoin is experiencing similar development as other "standard" currencies in times of boom and crisis.
4. Price development of Bitcoin can be described with standard econometric tools.

Methodology ARIMA, GARCH, non-linear models

Outline

1. Introduction
2. Theoretical Background of Bitcoin
3. Description of data and methodology used in the thesis
4. Empirical analysis of crash and the bubble
5. Interpretation of results - comparison of methodologies
6. Conclusion

Core bibliography

1. Kristoufek, L. (2013), BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. Sci. Rep. 3, 3415; DOI:10.1038/srep03415.
2. Hamilton, J. D. (1994) Time Series Analysis, Princeton University Press.
3. Lutkepohl, H. & Kratzig, M. (2007) Applied Time Series Econometrics , Cambridge Uiversity Press.
4. Cipra, T. (1986) Analýza časových řad s aplikacemi v ekonomii, SNTL – Nakladatelství technické literatury, Praha.
5. Enders, W. (2003) Applied Econometric Time Series.
6. http://bitcoinmagazine.com, Bitcoin's New All-time High: Exploring the Latest Bubble. [online]. Availbale from: http://bitcoinmagazine.com/8104/bitcoins-new-all-time-high-exploring-the-latest-bubble/
7. Buterin V. The Bitcoin Crash: An Examination. [online]. Available from: http://bitcoinmagazine.com/4113/the-bitcoin-crash-an-examination/
8. Vondra A. (2013) Structure of the Invisible Hand: Hierarchy and Others [Bachelor thesis]. Prague: Institute of Economic Studies.
9. Šurda P. Economics of Bitcoin: is Bitcoin an alternative to fiat currencies and gold? [Diploma Thesis]. Vienna: WU Vienna University of Economics and Business; 2013.

Author   Supervisor

Chapter 1

Introduction

In quickly developing Internet era we can observe several new interesting phenomena in the financial markets. One of them is the emergence of digital currencies such as the most popular one — Bitcoin and others like Litecoin, Ripple, Sharkcoin etc.[1] In our thesis we focus on Bitcoin which was introduced in 2009. It is not under control of any central authority and serves as an alternative to standard fiat currencies. In the last two years, Bitcoin's publicity has substantially increased which attracted not only large number of investors but also many businessmen who now accept Bitcoin as mean of payment. During its history, Bitcoin price experienced several periods of explosive bubbles followed by crashes, where post-bubble price level never returned to its initial value.

The objective of this thesis is to study two bubbles occurred in 2013 — the April and November bubble — and compare their characteristics. For analysis of price volatility during the bubbles we use standard econometric time series tools — ARIMA and GARCH models. In addition, we use rolling estimation to study the changes of ARCH and GARCH coefficients over time. However, these standard methodologies are found to be insufficient when capturing the volatile behavior of price returns in periods of excessive price growth and subsequent crash. Therefore, we implement LPPL model, introduced by Johansen *et al.* (2000), which is designed to capture the price development during stock market bubbles and crashes. They study the crashes by examining the key variables β and ω which represent power-law growth and frequency of the fluctuation during the bubble, respectively. Moreover, they define the key criterion

[1] List of the most popular digital currencies and their values can be found on Coinmarketcap.com

— scaling ratio λ — which is used as a tool for empirical examination and prediction of the time of the crash. However, when analyzing the November bubble the λ seems to be unable to predict the crash. For that reason we propose a hypothesis that the change in the scaling ratio $\Delta\lambda$ could be a better indicator of the upcoming crash. Even thought this method predicts false alarm of crash in some periods, it also predicts the peaks occurred during the bubble period with very good precision.
The thesis is structured as follows: after the introduction, one chapter is dedicated to related literature to the thesis. In Chapter 3 we offer basic characterization of Bitcoin along with empirical study about price formation and description of price development during the bubbles. Two following chapters offer the characterization of the dataset and details on ARIMA and various GARCH models methodology used in econometrics study of Bitcoin bubbles and the Whole period sample. Chapter 6 summarizes the results of GARCH estimation. Next, Chapter 7 presents the estimation process of GARCH rolling estimation along with the outcomes. In the following Chapter 8 we present the LPPL model — its methodology and theoretical background, fitting procedure with the results and finally, the "loop analysis" in which we study the over time changes of key variables β, ω and the standard scaling ratio λ along with introduced $\Delta\lambda$ criterion. Chapter 9 summarizes our overall findings.

Chapter 2

Literature review

There is generally not a large amount of literature related to the Bitcoin phenomena as Bitcoin currency was created only a few years ago and presented in Nakamoto (N/A). Although, during its short history, and especially after 2013 when Bitcoin gained publicity, a lot of economists started to pursue this topic and more and more economic papers have been published.

Surda (2012) discusses the Bitcoin future evolution and whether the currency can become an alternative to the fiat currencies or gold. He argues that Bitcoin has the advantage over fiat currencies in low transaction costs and inelastic supply. Among other things, Surda (2012) also studies the connection between liquidity and volatility of Bitcoin during the 2009-2012 period. He finds negative correlation between these two parameters which supports his arguments about Bitcoin being a medium of exchange.

Kristoufek (2013) studies the dynamic relationship between the Bitcoin price and interest in the currency measured by search queries on Google Trends and Wikipedia. His empirical results show strong correlation between price level and searched terms. More importantly, he argues that this relationship is bidirectional, that is the search queries influence the prices and also the prices influence the search queries.

Ciaian *et al.* (2014) and Kristoufek (2014) focus on formation of Bitcoin price and its short-term and long-term relationship with supply-demand fundamentals of Bitcoin and global macro-financial indicators. Moreover, Kristoufek (2014) studies the connections between the Chinese and the USD market. Results of both papers confirm that the Bitcoin prices are driven by interest of investors in the crypto-currency and that presence of standard fundamental factors — velocity, money supply, price level, size of the Bitcoin economy —

has strong impact on price dynamics. Furthermore, Ciaian *et al.* (2014) concludes that in presence of standard supply-demand factors and attractiveness of Bitcoin for investors the macro-financial indicators are irrelevant to Bitcoin price formation.

Very useful are freely provided and very detailed statistical data about Bitcoin markets. Coindesk (www.coindesk.com/price/) provides Bitcoin price index (BPI) data about exchange rate between USD and the Bitcoin. Wide range of time series about Bitcoin market (total bitcoins in circulation, number of transactions, trade volume etc.) are reported on daily basis on Blockchain (www.blockchain.info). Prices along with currency and Bitcoin traded volume series are available on www.bitcoincharts.com with frequency up to one minute for several currencies and almost every exchange.

Safka (2014) examines the connections between Bitcoin and the real economy and also inspects the Bitcoin volatility, from August 2010 until February 2014, using several variations of Autoregressive conditional heteroskedasticity (ARCH) models with structural breaks. He finds that Bitcoin volatility differs significantly during the studied period and that this behavior is best captured by TGARCH(1,1) model.

For analysis of Bitcoin price volatility we apply several models from ARCH family — basic ARCH model introduced in Engle (1982), enlarged to GARCH by Bollerslev (1986) and TGARCH along with EGARCH model capturing the leverage effect, described in Zakoian (1994) and Nelson (1991), respectively. Summary and application of ARCH models can be found in Engle & Patton (2000), Engle (2001) and in applied econometric textbooks Luthkepohl & Kratzig (2007) and Franses & Dijk (2000).

Geraskin & Fantazzini (2013) and Bree & Joseph (2013) offer very nice and well arranged information about the theory and application of LPPL model studying the financial bubbles and crashes originally presented in Johansen & Sornette (2001), Johansen *et al.* (2000) and Johansen *et al.* (1999). For comparison of estimated LPPL to Bitcoin we use empirical results of fitted LPPL to various stock markets and FOREX rates all around the world presented in Johansen & Sornette (2001), Johansen *et al.* (1999) and Johansen (2003).

Chapter 3

Theoretical background of Bitcoin

3.1 General description of Bitcoin

Bitcoin is peer-to-peer payment network introduced by Satoshi Nakamoto in 2009. Bitcoin, as a digital currency, is purely electronic with no physical form and serves as an alternative currency to the standard fiat currencies, e.g. dollar, euro, Chinese yuan. One of the key distinguishing features is that Bitcoin is not issued or controlled by any authority such as central bank or government, but it is managed by an open source software algorithm that uses global internet network to create the currency and also to record and to verify the transactions (Kristoufek 2013; Voorhees 2015; Ciaian *et al.* 2014).
Bitcoins are created in “mining” process, where miners use their computers to provide computing power for verifying and recording transactions into a public ledger called “blockchain”. As a reward for this service, miners receive transaction fees from validated transactions and more importantly certain number of Bitcoins, which provide inflow of new Bitcoins into circulation. In order to ensure that supply of Bitcoins evolves according to publicly known algorithm, the difficulty of solving the computational problem, measured in hashes, increases with every “block” of transaction verified (Kristoufek 2014; Ciaian *et al.* 2014).
Bitcoin enables any two parties anywhere on the earth to make transaction freely with each other with low or no fees. This is one of the main reasons why increasing number of companies accept Bitcoins as a mean of payment in exchange for goods and services.[1] Bitcoin can be also traded for other currencies in many exchanges, operating 24/7, with easy access for anyone with

[1]List of companies accepting bitcoins available at Coindesk.com

a computer and internet connection, therefore the entry costs are one of the lowest on financial market (Ciaian *et al.* 2014; Pieters & Vivanco 2015).
As mentioned before, one of the main advantages of Bitcoin in contrast with standard fiat currencies is that person is able to make a transaction anywhere in the world at any time with very low fess. Furthermore, all finalized transactions are completely public and can be verified, however personal information is hidden thus the information transparency is secured while personal identities are saved.
On the other hand Bitcoin is not flawless. One of the biggest issues is that the relative anonymity of Bitcoin represents perfect environment for organized crime, money laundering and other illegal activities. Also with increasing popularity and public attention some Bitcoin exchanges have been targeted by hackers and their attacks have caused significant losses for Bitcoin owners and represented serious problems to attacked exchanges which often led to shutting them down. As in the case of one of the largest and most popular exchanges — MtGox — in the beginning of 2014. Furthermore, the public awareness about digital currencies and Bitcoin is very low as it is still a developing currency, plus there is no central authority and the price of Bitcoin is highly volatile and does not serve well as a store of value which represents a potential threat for its future growth (CoinReport 2014; Kristoufek 2014; Bambani & Beer 2013).

3.2 Bitcoin price formation

According to Kristoufek (2013; 2014) the price formation of Bitcoin cannot be explained by standard economic and financial theories because supply-demand fundamentals, normally forming the basis of standard currency price formation, are missing on Bitcoin market. Firstly, supply function evolves according to publicly known algorithm and the currency is not issued by any central bank or other entity. Secondly, the demand is mainly driven by speculative behavior of investors, as there is no interest rate for holding the digital currencies and thus profits can be obtained only from price changes. So investors' behavior and sentiment becomes a key variable in price forming.
Moreover considering that Bitcoins' price dynamics has been changing significantly during its evolution in recent years it would be naive to think that the driving forces of the price have remained unchanged during its existence. Therefore when examining Bitcoin price dynamics both frequency (scale of the

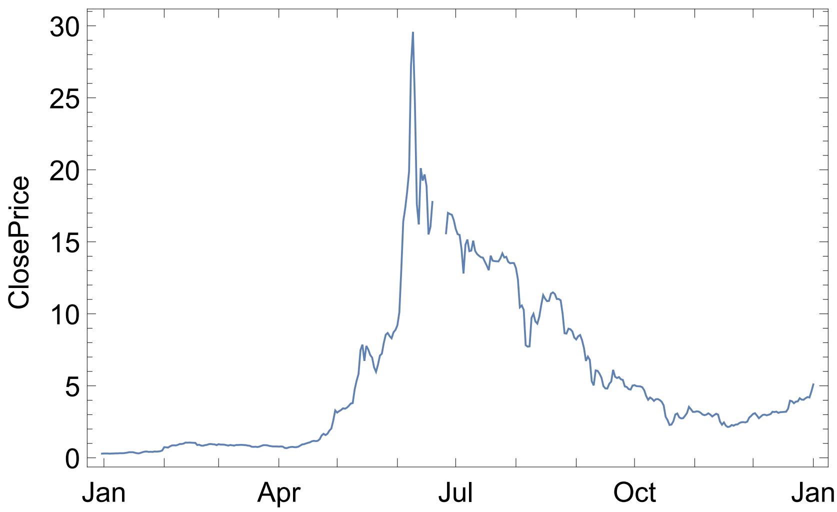
interconnections between variables) and time have to be focused on.
Based on the findings of Ciaian *et al.* (2014), the supply-demand fundamentals have a strong impact on Bitcoin price, moreover the demand side (size of the Bitcoin economy and velocity of Bitcoin circulation) appears to have stronger influence on price then the supply side (stock of Bitcoin in the circulation). However Bitcoin attractiveness for investors — measured by Wikipedia views, number of new members and new posts[2] — has the strongest and statistically most significant impact on the currency price. This is consistent with Kristoufek (2013; 2014) results, that Google and Wikipedia searches for the word "Bitcoin", quantifying the interest in Bitcoin, are highly positively correlated with Bitcoin price. Moreover, he found out that this relationship is most evident in the long run and that interest in Bitcoin have an asymmetric effect in bubbles periods. During bubble formation (periods of explosive prices) interest drives prices further up, and during the bubble bursting (rapid prices declines) interest pushes prices further down. Ciaian *et al.* (2014) also argue that if supply-demand variables and Bitcoin's attractiveness for investors are included in analysis, the macro-financial indicators captured by stock exchange indices, exchange rates and oil prices do not significantly affect Bitcoin price in the long run.
Furthermore, Kristoufek (2014) discovered several other interesting findings about drivers of the Bitcoin price. First, he found out that with increasing usage of Bitcoin in real trade transactions the currency appreciates in the long run and that increasing prices boost demand for the currency at the exchanges. Second, the increasing price attract more miners into the system so there is a positive relationship between mining difficulty and price in the long run, however this relationship becomes weaker in time as the price of Bitcoin is slowly decreasing and difficulty is too high. Third, there is no sign that Bitcoin is a safe haven investment as gold is (or at least once was). Finally, Kristoufek shows that there is no clear evidence of interconnection between the CNY volume and USD price and that there is no causal relationship between the Chinese and USD Bitcoin markets, even though these two markets are closely connected.

[2]For detailed description of variables see Ciaian *et al.* (2014), chapter 4 — Data.

3.3 Historical price development

To examine the Bitcoin price development we use data from formerly one of the largest and most popular exchange on Bitcoin market — MtGox — which went bankrupt on February 2014 as a result of a loophole in their security system (BBC 2014). However MtGox data from period 2010-2013 are not affected by the bankruptcy so they appear to be the best choice for the analysis in this period. For period 2014-2015 we are using data from currently the largest exchange trading with US dollars — Bitfinex.
If we look at Bitcoin price development from the beginning we can observe several exponential increases in prices throughout the history. The first one occurred in the beginning of 2011. The price increased from \$ 0.1 to \$1 (900% increase) and later that year, the price grew from 1\$ to a maximum of 30\$ within two months thus creating the first bubble with absurd profit of approximately 3000%. But only after few days the price dropped to half its value and continued decreasing until it reached the level below \$5. Evolution of Bitcoin price during 2011 is shown in Figure 3.1.

Figure 3.1: **Bitcoin price**, 2011

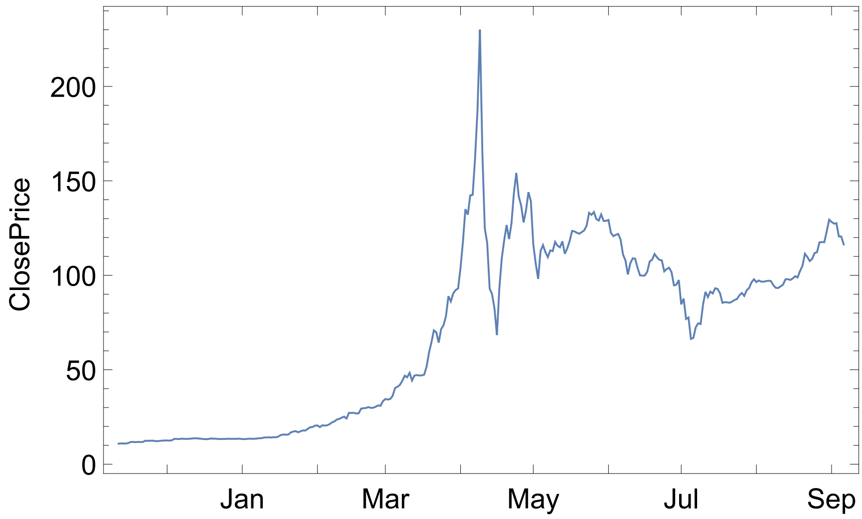


Source: The Authors' own computations via *Wolfram Mathematica* using BPI data.

3.3.1 April bubble

During 2012 the price experienced slow increasing trend but nothing worth mentioning happen. But in the beginning of 2013 so called April bubble appeared. On 1 January 2013 the price of Bitcoin was at the level of \$13 and then started growing exponentially until it reached maximum of \$265 on 10 April 2013. In the last ten days Bitcoin more than doubled its value and the potential profit was almost 2000% in less than four months. Next day closing price dropped about 50% compared to the peak and continued decreasing for a few days until it hit the bottom at level of \$50. The Bitcoin value has stabilized around the range of \$100-120 after the subsequent corrections and stayed there in upcoming months. April bubble price development is depicted in Figure 3.2.

Figure 3.2: **Bitcoin price**, April bubble 2013



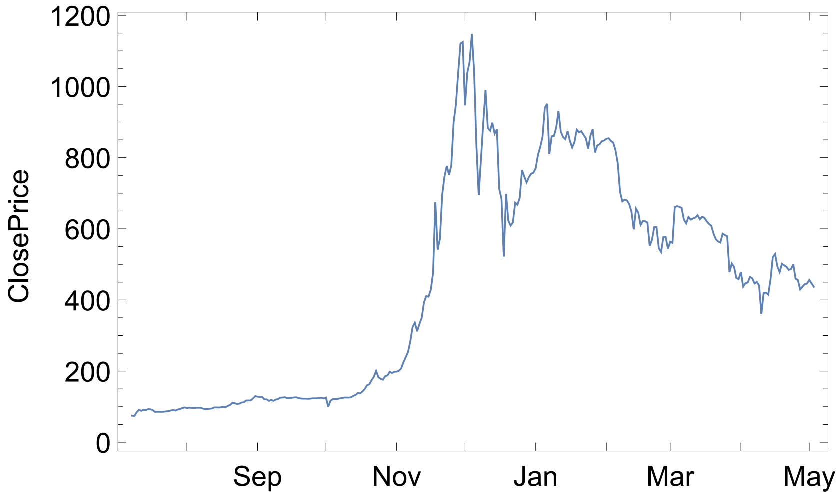
Source: The Authors' own computations via *Wolfram Mathematica* using BPI data.

3.3.2 November bubble

The second bubble which occurred later that year in November is depicted in Figure 3.3. It can be observed that the Bitcoin price started to rise steadily in October and exponentially in the beginning of November. Later that month Bitcoin broke the level of \$1000 and peaked at \$1242 by the end of November. This represents absolutely unimaginable potential profit of approximately 9200% for buy-and-hold strategy from the beginning of the year, i.e. in 11 months (Kristoufek 2014). In the beginning of December Bitcoin lost more than 50% of its value with price around \$600 and after following corrections its exchange rate with US dollar has stabilized around \$900 per Bitcoin in the beginning of 2014.

But, as already mentioned in this chapter, the largest exchange — MtGox — went bankrupt and Bitcoin suffered enormous strike to its credibility on February 2014. Since then the price started to have a decreasing trend with

Figure 3.3: **Bitcoin price**, November bubble 2013



Source: The Authors' own computations via *Wolfram Mathematica* using BPI data.

few jumps and falls but not on such a large scale as in the cases of April or November bubble. The current value (March 2015) of Bitcoin is slightly below $300.

Chapter 4

Dataset

4.1 Data Description

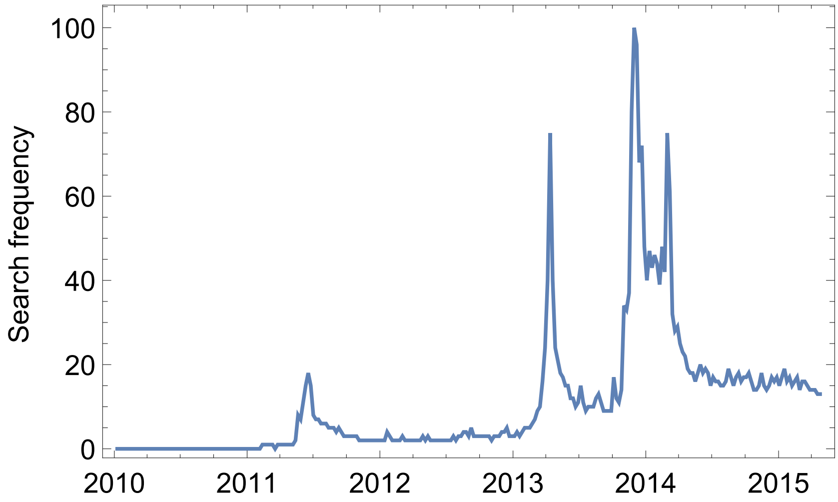
Since the analysis of a specific exchange is not feasible because of the Mt-Gox bankruptcy we use the Bitcoin price index (BPI) which was introduced by CoinDesk (2015) in September 2013. The BPI is an index of the exchange rate between US dollar (USD) and Bitcoin (BTC) and it is constructed as a simple average of the Bitcoin prices across the most liquid global exchanges. In order to be included in the BPI exchanges have to meet specific criteria, which are currently met by five exchanges — Bitfinex, Bitstamp, BTC-e, LakeBTC and OKCoin. The BPI historical data commence on 1 July 2013. Prior to this date MtGox closing price data are used as BPI and they are available from 18 July 2010. The series is freely available at Coindesk.com. In our analysis we use daily closing price data.

For whole sample analysis we filtered the available data according to three criteria. Two of them were taken from Kristoufek (2013) but the last one was created by us.

Based on Kristoufek (2013), we first examine the Google trends data about frequency of searching for term "Bitcoin"[1] which represents the investors interest and attention to the Bitcoin currency. From Figure 4.1 it can be seen that before the bubble on May 2011 the interest about Bitcoin currency was nearly at zero level, which questions market efficiency before May 2011.

[1]Data are available with weekly frequency from https://www.google.com/trends and they are normalized, so the maximum value of the series is equal 100. For our analysis we chose time period from 3 January 2010 to 4 April 2015

Figure 4.1: **Google Trends search query**, weekly series



Source: The Authors' own computations via *Wolfram Mathematica* using data from https://www.google.com/trends.

Second, we apply the same criterion as introduced in Kristoufek (2013) where he examines the liquidity of the Bitcoin market at MtGox exchange and finds out that the market has been liquid since the beginning of May 2011. This fact confirms our conclusions resulting from Google trends data inspection.

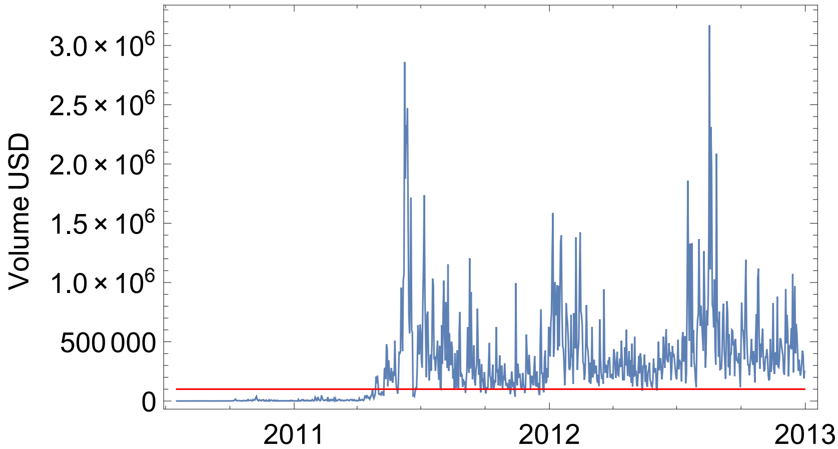
Finally, we implement our own restriction where the starting date has to satisfy the condition about the USD volume traded on seven major USD exchanges operating between years 2010-2013.[2] We choose the first day when the sum of USD volume traded in one day exceeds 100000 USD and stays above this level for at least 98 days from 100 upcoming days, without violation of this condition for the rest of the time period, as a starting date for our sample. From Figure 4.2 we can see that before May 2011 the volume was clearly under 100000 USD level which supports our claim about market inefficiency in this period. By further examination of the data we can notice that at the end of 2011 there was a period where the volume declines again below the 100000 USD level.[3]

Therefore, we analyze the whole sample series starting on 19 December 2011 with an ending date of 31 March 2015 as this period satisfies all defined conditions. As a result, our data consist of 1199 observation in total.

[2]Exchanges with highest volume in USD currency between 2010-2013 based on data available from bitcoincharts.com/charts — Mt.Gox, btc.e, btcex.com, BitStamp, Camp BX, Intersango and TradeHill.

[3]We also take into consideration the data about the number of transaction per day, available from blockchain.info. We use similar approach as in USD volume analysis and set the bar for number of transaction per day to 5000. Obtained results for sample starting date are consistent with the conclusion of USD volume examination.

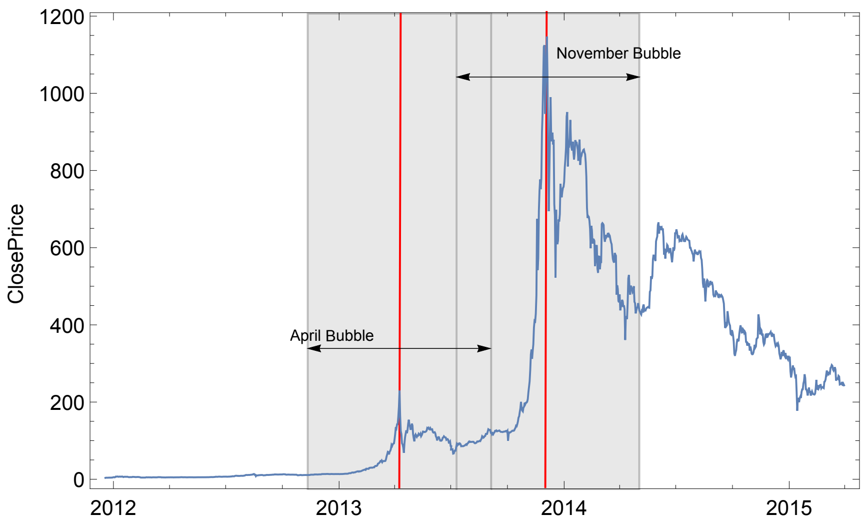
Figure 4.2: **USD Volume**, seven major USD exchanges, daily series



Source: The Authors' own computations via *Wolfram Mathematica* using data from bitcoincharts.com.

For the analyses of the bubble periods, we take 150 days before and after the peak day so we obtain time series with 300 observation in total. The starting and ending dates of the analysis of April bubble are therefore set to 11 November 2012 and 6 September 2013. Similarly, for the November bubble the dates are 8 July 2013 and 3 May 2014. Graphical representation of these three periods is depicted in Figure 4.3.

Figure 4.3: **Raw Bitcoin price series** — ClosePrice



Source: The Authors' own computations via *Wolfram Mathematica* using BPI data.

4.1.1 Stationarity

We test our data for stationarity in order to avoid possible spurious regression problem and potential long memory of shocks which can be present in non-stationary time series data. This is done by studying the original series (referred as ClosePrice) and its first logarithmic differences transformation (referred as ld_ClosePrice) in our three sample periods of interest — Whole sample, April bubble and November bubble. To test for stationarity, we first examine Autocorrelation (ACF) and Partial Autocorrelation (PACF) functions described in detail in Luthkepohl & Kratzig (2007).
From Figure A.1 (see Appendix) it can be observed that for original series the ACF is declining extremely slowly towards zero with increasing lags in all periods of interest which is specific for non-stationary series. The PACFs first lag is close to one and then drops close to zero for second and higher lag which suggests that we should use differentiated data of order one. Looking at the transformed series, the Whole sample and November bubble series seem to be white noise since in ACF almost all autocorrelations are close to zero and the ones above the significance level are not consistent with economic theory. They seem to be coincidence and do not have any significant effect on the stationarity of the data. Regarding the April bubble the series seems to be stationary as ACFs first lag is significant but small and for second and higher lag autocorrelations drop to zero level again with some exceptions without affecting the stationarity.
To verify these preliminary conclusions formally we perform Augmented Dickey-Fuller (ADF) and KPSS tests for stationarity, described in detail in Luthkepohl & Kratzig (2007). According to Kristoufek (2013), ADF test, with non-stationarity of the series as null hypothesis, and KPSS test, with stationarity of the series as null hypothesis, form ideal pair of tests for stationarity as they have the opposite null and alternative hypotheses. Results of these tests are summarized in Table 4.1.

We find the original series in all sample periods to be non-stationary and to contain the unit-root. On the contrary for log-difference transformation we reject the null hypothesis of ADF test that series contains unit-root at 5% significance level for all sample periods. Performing the KPSS test we failed to reject the null hypothesis of time series to be integrated of order zero at 10% significance level for April bubble, at 5% significance level for Whole sample but only at 1% significance level for November bubble. Kwiatkowski *et al.*

Table 4.1: ADF and KPSS tests results

	ADF	p-value	KPSS	p-value
Whole sample				
ClosePrice	-0.8919	0.3299	8.7168	$< .01$
log-difference	-5.7196	0.000	0.4578	0.052
April bubble				
ClosePrice	0.2388	0.7556	3.6805	$< .01$
log-difference	-2.2352	0.0245	0.2245	$> .10$
November bubble				
ClosePrice	-0.3079	0.575	2.9276	$< .01$
log-difference	-2.4238	0.0148	0.5191	0.042

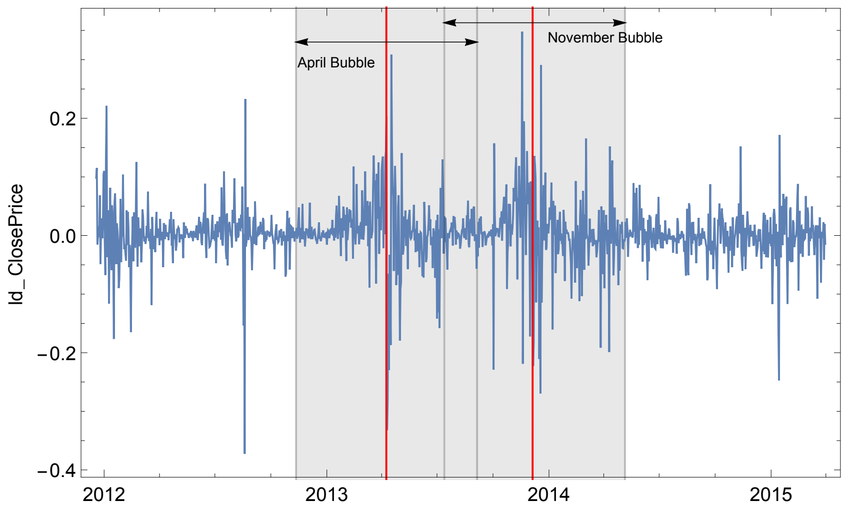
Source: The Authors' own computations via *Gretl* using BPI data.

(1992) propose following KPSS test statistic

$$KPSS = \frac{1}{T^2} \sum_{t=1}^{T} \frac{S_t^2}{\hat{\sigma}_\infty^2} \tag{4.1}$$

where $S_t = \sum_{j=1}^{T} \hat{\omega}_t$ with $\hat{\omega}_t = y_t - \bar{y}$ and $\hat{\sigma}_\infty^2$ is an estimator of the long-run variance of the stationary process z_t ($y_t = x_t + z_t$, where x_t is a random walk). Based on the properties of this test statistic we can see that with the larger data sample (with the same distribution of ω_t) we get lower test statistic and higher p-value which corresponds to the higher significance of the test result of not rejecting the H_0.[4]

Figure 4.4: **Bitsoins' daily returns** — *ld_ClosePrice*



Source: The Authors' own computations via *Wolfram Mathematica* using BPI data.

[4]For empirically testing this property we simulate the data with 10000, 20000 and 30000 observations using *Wolfram Mathematica* software and perform the KPSS test using *Gretl* and obtain the consistent results with the theory.

Therefore, we can see that the KPSS test results largely depend on number of observation. For that reason we decided to take into consideration the results of ADF test in case of the November bubble (as we have only 300 observations). Based on the tests results we conclude that the log-difference transformation of the series is stationary. Therefore, we use continuous daily returns (ld_ClosePrice) for our econometric analysis, defined as

$$r_t = ln(p_t) - ln(p_{t-1}) = ln\left(\frac{p_t}{p_{t-1}}\right) \tag{4.2}$$

and shown in Figure 4.4. For descriptive statistics see Table A.1 in Appendix A.

Chapter 5

Methodology

The econometric models used for analyses of the Bitcoin price volatility are described in the following section in order to give a theoretical background.

5.1 ARIMA Process

Luthkepohl & Kratzig (2007) define general auto-regressive moving average model of orders p,q (ARMA(p,q)) process as combination of AR(p) and MA(q) models as follows

$$y_t \;=\; a_1 y_{t-1} + \cdots + a_p y_{t-p} + u_t + m_1 u_{t-1} + \cdots + m_q u_{t-q} \tag{5.1}$$

where u_t is white noise process with zero mean $E(u_t) = 0$ and time invariant variance $E(u_t^2) = \sigma_u^2$. For an efficient estimation, ARMA methodology needs to be applied to stationary series. The ARMA(p,q) model is found to be stationary when both the AR and MA parts of the process are stationary. As MA process always satisfies this condition, that is achieved when absolute value of sum of alpha coefficients is lower than one — $|\sum_{i=1}^{p} a_i| < 1$.

Since we analyze daily returns series ($y_t = r_t$), which is already differentiated series of order one and it is stationary, we do not need further differentiation. Therefore, the ARMA(p,q) process of daily returns represents the ARIMAp,1,q) process where the 1 stands for differentiated data of order one.

5.2 Univariate GARCH models

Univariate autoregressive conditional heteroskedasticity (ARCH) models are widely used in financial empirical studies for modeling volatility clustering. The basic ARCH model was first introduced by Engle (1982) and later Bollerslev (1986) came up with more complex model - the generalized ARCH (GARCH).

5.2.1 ARCH process

According to Luthkepohl & Kratzig (2007) and Engle (1982) the ARCH(q) process can be defined as

$$\begin{aligned} u_t &= \sigma_t \varepsilon_t;\ \varepsilon_t \sim i.i.d\ N(0,1) \\ h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-q}^2 \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 \end{aligned} \tag{5.2}$$

where u_t is residual from ARMA model defined in Equation 5.1, following an autoregressive conditionally heteroskedastic process of order q (ARCH(q)) given past information represented by $\Omega_{t-1} = \{u_{t-1}, u_{t-2}, \ldots\}$, so that $u_t|\Omega_{t-1} \sim (0, h_t)$ and conditional variance is $E(u^2) = h_t$. To have a stationary process the following condition $\sum_{i=1}^{q} \alpha_i < 1$ needs to be satisfied. According to Luthkepohl & Kratzig (2007) it was observed that the ARCH model usually needs a large number of squared lagged residuals to have a correct specification of the model which can be considered as the main weakness of this model.

5.2.2 GARCH process

To overcome this weakness, Bollerslev (1986) introduced GARCH model allowing for a more flexible lag structure. Unlike the ARCH process, in which the conditional variance is only a linear function of past sample variances u_t, the lagged conditional variances of GARCH are allowed to enter the model. In his paper, he also argues that even simple GARCH model provides marginally better fit and more plausible learning mechanism than the ARCH model with large number of lags. The GARCH(p,q) process is then given by

$$\begin{aligned} u_t &= \sigma_t \varepsilon_t;\ \varepsilon_t \sim i.i.d\ N(0,1) \\ h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_q u_{t-q}^2 + \beta_1 h_{t-1} + \cdots + \beta_p h_{t-p} \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \end{aligned} \tag{5.3}$$

The starionarity of the process is achieved when $(\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j) < 1$. It is important to note that for $p = 0$ the process is the same as ARCH(q).

5.3 Modifications of GARCH models

Both basic models assume that positive as well as negative innovations have the same impact on the conditional variance h_t. However, empirical literature examining the returns of assets shows that negative news affect the future volatility more than positive news. The existence of such a leverage effect supports the need to improve these models (Luthkepohl & Kratzig 2007).

5.3.1 TGARCH Process

For capturing the leverage effect Zakoian (1994) introduced threshold GARCH (TGARCH) model. The basic and mostly used TGARCH(1,1) process is defined as

$$\begin{aligned} u_t &= \sigma_t \varepsilon_t;\ \varepsilon_t \sim \textit{i.i.d N(0,1)} \\ h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + +\gamma_1 I_{t-1} u_{t-1}^2 \end{aligned} \tag{5.4}$$

where I_{t-1} is an indicator function equal to 1 if $u_{t-1} < 0$ and equal to 0 otherwise. If $\gamma_1 = 0$ then there is no leverage effect in the estimated model and TGARCH process becomes GARCH process.

5.3.2 EGARCH Process

Last modification of the basic model is exponential GARCH (EGARCH) first described by Nelson (1991). SAS Online Documentation (2015) defines basic EGARCH(1,1) model as

$$\begin{aligned} u_t &= \sigma_t \varepsilon_t;\ \varepsilon_t \sim \textit{i.i.d N(0,1)} \\ log(h_t) &= \alpha_0 + \beta_1 log(h_{t-1}) + \theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\frac{|u_{t-1}|}{\sqrt{h_{t-1}}} - E\left(\frac{|u_{t-1}|}{\sqrt{h_{t-1}}} \right) \right] \\ \varepsilon_t &= \frac{u_t}{\sqrt{h_t}} \end{aligned} \tag{5.5}$$

and if $\varepsilon_t \sim N(0,1)$ then $E(|\varepsilon_{t-1}|) = \sqrt{\frac{2}{\pi}}$. It is possible to account for the leverage effect if coefficient $\theta < 0$. Since we model $log(h_t)$ the main advantage of the model is that independent on the sign direction of the parameters the conditional variance h_t remains positive.

5.4 Information criteria

According to Franses & Dijk (2000) we choose the best model selection based on information criteria which measure the balance between goodness of fit — representing quality of the estimation by log-likelihood — and parsimony — number of parameters used for estimation. We use two criteria, Akaike information criterion (AIC) and Schwarz Bayesian information criterion (BIC) formulated as

$$\begin{aligned} AIC &= -2L + 2k \\ BIC &= -2L + 2k\, log(n) \end{aligned} \tag{5.6}$$

and L representing the maximum log likelihood function defined as

$$L = -\frac{n}{2}\left(1 + log(2\pi)\right) - \frac{n}{2}\, log\left(\frac{SSR}{n}\right), \tag{5.7}$$

where n is the number of observations in the sample, k denotes the number of estimated parameters and SSR denotes sum of squared residuals.

In order to get the best selection of a model the chosen criterion needs to be minimized. We are using BIC as it penalizes additional parameters of the model as the sample size grows more heavily then AIC. This ensures selection of correctly specified parsimonious model over larger model with more parameters.

Chapter 6

Results

6.1 ARIMA analysis

Choosing the models with the lowest values of the BIC, we end up with ARIMA(1,1,0) for the April bubble with statistically significant AR coefficient $\alpha_1 = 0.216$ which suggests very small possibility to predict Bitcoin returns. For the Whole sample and the November bubble samples we choose ARIMA(0,1,0) model. Therefore, we can conclude that we are not able to predict future development of Bitcoin returns using ARIMA process for these two periods. Further results of the ARIMA analysis are summarized in Table A.2 in Appendix A.
In order to find out if the ARMA processes suffer from conditional heteroskedasticity in error terms we implement an ARCH-LM test, described in detail in Luthkepohl & Kratzig (2007). Executing the test for the Whole sample as well for the bubbles the p-values are all zero. Thus, the null hypothesis can be rejected which indicates the presence of conditional heteroskedasticity. To overcome this issue an ARCH and GARCH framework is implemented.

6.2 GARCH analysis

Based on BIC, described in Equation 5.6, we estimate the GARCH(1,1) model with robust standard errors, defined by Equation 5.3 with p=1 and q=1, for all three periods of interest.

Looking at the results reported in Table 6.1 it can be seen that the conditional volatility significantly depends on lagged squared error term u_{t-1}^2 (ARCH term)

Table 6.1: GARCH (1,1) model — results

	Coefficient	**z-statistic**
Whole sample		
α_0	0.000156*	1.788
α_1	0.274***	3.224
β_1	0.703***	8.945
$\alpha_1 + \beta_1$	0.977	
	Q-statistic	**p-value**
L-B (5)	7.2214	0.205
L-B (10)	21.1074	0.020
April bubble		
α_0	0.000034	1.346
α_1	0.216***	4.926
β_1	0.784***	23.700
$\alpha_1 + \beta_1$	1.000	
	Q-statistic	**p-value**
L-B (5)	9.3991	0.094
L-B (10)	31.2635	0.001
November bubble		
α_0	0.000498	1.137
α_1	0.267	1.401
β_1	0.643***	3.190
$\alpha_1 + \beta_1$	0.909	
	Q-statistic	**p-value**
L-B (5)	12.6832	0.027
L-B (10)	18.8807	0.042

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Source: The Authors' own computations via *Gretl* using BPI data.

and on lagged conditional variance of the error term h_{t-1} (GARCH term) in all periods except November bubble where only GARCH term is statistically significant. This could be caused by the characterization of this bubble where large number of new traders traders enter the Bitcoin market. Most of these people do not have any experience with trading and they formed expectations about the profit solely based on the last bubble (April bubble) price development rather than on the actual market condition. This idea is described further in the Section 6.5.

Furthermore, in Whole sample and November bubble periods the sum of coefficients of ARCH and GARCH term $\alpha_1 + \beta_1$ is relatively high which indicates persistent volatility shocks. In April bubble sample the sum of coefficients is equal to 1 so it can be said that in this period the GARCH process is not station-

ary and the conditional variance does not converge to a constant unconditional variance in the long run i.e. the estimated (forecasted) variance grows linearly over the forecast horizon (Durnel 2012).
We also report Ljung-Box Q-statistics and p-value of the 5th and 10th lag for standardized squared residuals of fitted GARCH model. It can be seen that for the Whole sample we reject the null hypothesis of no serial correlation in residuals but only for 5 lags. For 10 lags we cannot reject the null hypothesis i.e. we cannot say that the residuals are serially uncorrelated for 10 lags. With respect to April bubble we can again reject the null hypothesis for 5 lags but only at 5% significance level, for 10 lags the situation is similar to Whole sample period. In the case of November bubble we cannot reject the null hypothesis neither for 5 lags nor 10 lags. These results of Ljung-Box test raise the question whether the models we are using are adequately capturing all of the persistence in the variance of returns (Engle & Patton 2000).

6.3 TGARCH and EGARCH analysis

We also estimate the TGARCH(1,1) and EGARCH(1,1) models described by Equation 5.4 and Equation 5.5, respectively. Results are summarized in Table 6.2.

Table 6.2: TGARCH(1,1) and EGARCH(1,1) models — results

	Coefficient	**z-statistic**		**Coefficient**	**z-statistic**
Whole sample	**TGARCH**			**EGARCH**	
α_0	0.000237**	2.145	α_0	-1.123**	-2.316
α_1	0.286***	4.344	γ	0.495***	3.913
β_1	0.714***	9.697	β_1	0.874***	12.950
γ_1	-0.002	-0.023	θ	0.010	0.231
April bubble					
α_0	0.000067*	1.947	α_0	-0.588***	-4.909
α_1	0.256***	5.132	γ	0.466***	5.803
β_1	0.791***	22.760	β_1	0.962***	63.840
γ_1	-0.036	-0.265	θ	-0.005	-0.096
November bubble					
α_0	0.000641	1.245	α_0	-1.977*	-1.921
α_1	0.317**	2.058	γ	0.621**	2.432
β_1	0.633***	3.360	β_1	0.725***	4.393
γ_1	-0.044	-0.250	θ	0.078	0.605

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Source: The Authors' own computations via *Gretl* using BPI data.

We can observe that the asymmetric ARCH terms γ for TGARCH and θ for EGARCH are both statistically insignificant for all periods. Therefore, we do not find any significant leverage effect in conditional variance. These conclusions are inconsistent with Safka (2014) which is probably caused by the fact that he performed an analysis of a different time period (2010 - 2014) and especially that he included the period before 2012 in which we claim that the Bitcoin market is not efficient (see Section 4.1).

6.4 Comparison with stock market indexes and FOREX rates

In this section we compare the results of estimated GARCH(1,1) model for Bitcoin Whole sample period and five global stock market indexes — NASDAQ, S&P 500, FTSE100, NIKKEI and HANG SENG — reported in Jiang (2012) along with one single stock — APPL[1] — and FOREX rates of USD with eight other currencies — EUR, CHF, JPY, CNY, RUB, CAD, MXN, BRL.
From the results of stock market indexes, summarized in Table 6.3, it can be observed that the ARCH coefficient α_1 of Bitcoin has much higher value than coefficients of market indexes and the value of Bitcoin GARCH coefficient β_1 is much lower than coefficients of market indexes. According to Alexander (2008) higher α_1 which is often associated with lower β_1 produces GARCH volatility with high volatility of volatility (vol-of-vol). In other words GARCH volatility is more "spiky". Based on that we can conclude that the stock markets are much more stable than Bitcoin market as in Bitcoin market the GARCH volatility has larger volatility.

Table 6.3: GARCH (1,1) model — Bitcoin comparison with stock market indexes

	Bitcoin	NASDAQ	S&P 500	FTSE100	NIKKEI	HANG SENG	AAPL
α_1	0.274	0.0924	0.0994	0.1292	0.1339	0.1062	0.0865
β_1	0.703	0.8964	0.8863	0.8603	0.8496	0.8867	0.8813

Source: Based on results presented in Jiang (2012) except for APPL.

[1] For APPL analysis the data has been taken from http://finance.yahoo.com/q/hp?s=AAPL and estimated in the period 30. 4. 2008 - 25. 6. 2015

Comparing Bitcoin with FOREX rates[2] GARCH(1,1) results from Table 6.4, we observe that even now the Bitcoin GARCH coefficient is lower and ARCH coefficient higher. Therefore the conclusion is the same as in comparison with stock market indexes — FOREX rates are more stable than Bitcoin returns and have lower vol-of-vol. We get similar GARCH coefficient only when comparing with exchange rate of USD and Chinese Yuan but the ARCH coefficient still differs significantly.

Table 6.4: GARCH (1,1) model — Bitcoin comparison with FOREX rates

	Bitcoin	USD/EUR	USD/CHF	USD/JPY	USD/CNY
α_1	0.274	0.029	0.057	0.057	0.176
β_1	0.703	0.970	0.929	0.928	0.740
		USD/RUB	USD/CAD	USD/MXN	USD/BRL
α_1		0.088	0.047	0.118	0.114
β_1		0.909	0.950	0.873	0.885

Source: The Authors' own computations via *Gretl* using data from http://www.oanda.com/currency/historical-rates/.

6.5 Comparison of April and November bubble

In the end, we compare the two bubbles which occurred in 2013.
Firstly, we compare the two bubbles by the same logic as in previous section where we compare Bitcoin with market stock indices. Based on the magnitude of ARCH and GARCH coefficients from Table 6.1, we observe that the November bubble period GARCH volatility has higher vol-of-vol than April bubble as November bubble has higher α_1 and lower β_1 coefficients.
Johansen & Sornette (2001) summarize the development of financial bubbles and crashes in five stages.

1. Smooth start of the bubble with some increasing demand for asset.

2. Price appreciation due to increased interest of international investors who see good potential gains from the asset.

3. This attracts less sophisticated investors, the so called "noise traders" which leads to demand for faster stock rising.

[2]Data about FOREX rates are taken from http://www.oanda.com/currency/historical-rates/. We choose the period 30. 4. 2008 - 25. 6. 2015, which give us 2613 observations, and apply the same methodology as in GARCH analysis.

4. At this stage, market behavior is driven mostly by the irrational trading of noise investors.

5. As the price skyrockets, the number of new investors entering the speculative market decrease and the market becomes very nervous until the instability is revealed and the market collapses.

Based on this characterization, we claim that during the later bubble the number of irrational investors entering the market was much higher[3] as the public awareness about Bitcoin significantly increase (as seen in Google Trends Figure 4.1) and as investors expectation about the profit was mainly based on previous bubble development more than on actual market condition. This cause larger appreciation of Bitcoin price during the bubble and consequently much larger instability in the market which could be the main reason for relatively small GARCH and insignificant ARCH coefficient in November bubble analysis. Secondly, the GARCH process of April bubble appears not to be stationary which corresponds with more persistent volatility shocks than in the November bubble GARCH process, as mentioned before in GARCH analysis (see Section 6.2). Thirdly, the most appropriate model based on BIC is AR(1)-GARCH(1,1) for April bubble, while for the November bubble it is GARCH(1,1) model.

[3]Several charts about the trade volume, number of transactions per day, number of unique Bitcoin Addresses used, My Wallet number of users etc. available on blockchain.info/charts shows that during the November bubble all of these indicators experienced significant increase suggesting more people trading in the Bitcoin market. The trade volume had increase to extreme levels and represented overall maximum, number of transaction and number of unique Bitcoin addresses used substantially rose and My Wallet had experience exponential growth in number of users.

Chapter 7

Rolling Estimation

In this chapter we would like to further examine the changes of ARCH and GARCH coefficients over time using rolling estimation. With almost 1200 daily observation in our dataset we estimate GARCH(1,1) process on 250 days rolling samples with step of ten days.[1]
The 250 days window is reasonably large as such a sample is still able to provide solid statistical results ans it is sufficiently small to capture the structural breaks in the data, but also large enough to provide required statistical properties (eq. stationarity). Moreover, the step of ten days is chosen as it is able to capture the frequently changing scale of returns during the bubbles.[2]

7.1 Stationarity

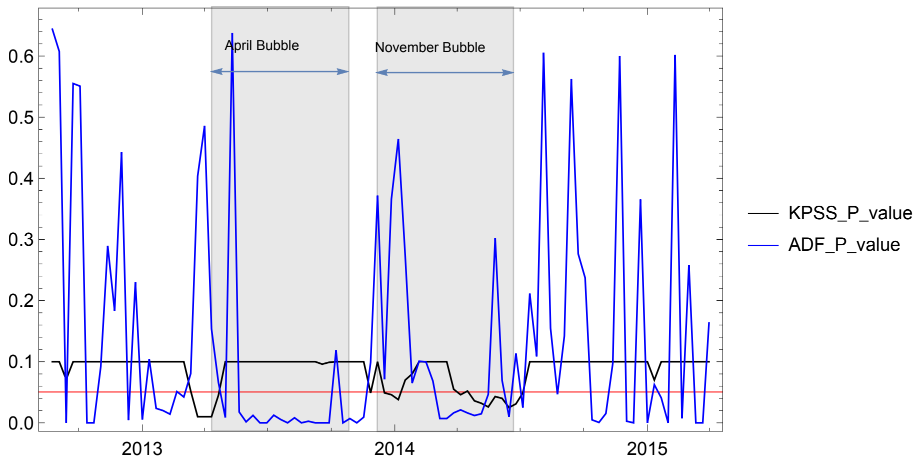
Before interpreting the actual results we examine the stationarity of individual rolling estimation samples. We again use the combination of ADF and KPSS tests, as described above (Subsection 4.1.1), to test the stationarity. We want to obtain p-value of KPSS test higher then 0.05 (for not rejecting H_0 at 5% significance level) and p-value of ADF test lower than 0.05 (for rejecting H_0 at 5% significance level) to confirm the stationarity assumption of the series. Figure 7.1 shows the p-values of both tests. While the ADF test results are very unstable, the KPSS p-values seem to have much smoother development

[1] All figures in this chapter are constructed in a way that the last observation of the sample period represents the whole period.

[2] Various combinations of rolling sample lengths and steps had been used in the analysis. The results of these analyses are available upon request.

over time. Therefore, we decided to follow KPSS tests' results when identifying the stationarity of the series.

Figure 7.1: **KPSS and ADF tests results**, p-values



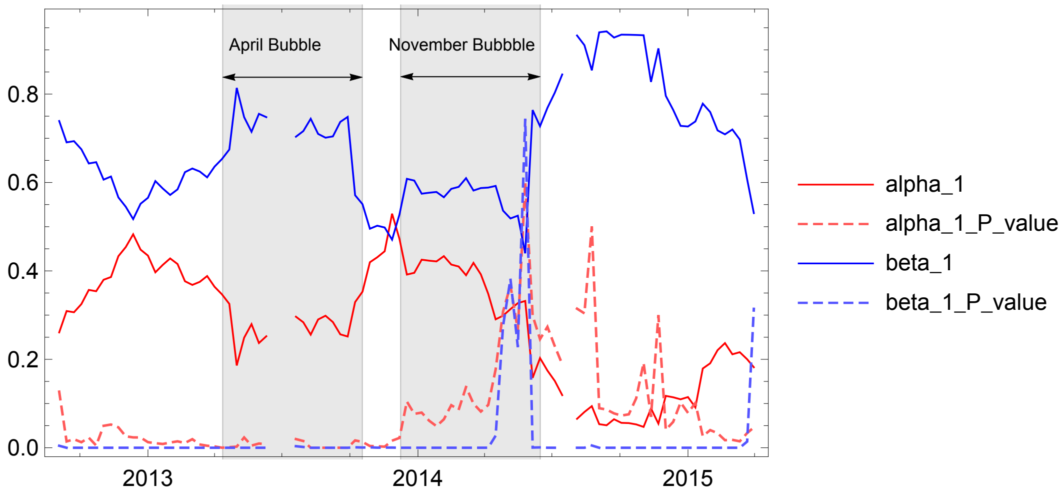
The red line represents the 5% significance level. When performing KPSS test via *Gretl*, we obtain specific number of p-value only when the value is lower than 0.1. Otherwise we only get the information that the p-value is higher than 0.1. For that reason, the p-value higher than 0.1 is depicted as equal to 0.1.
Source: The Authors' own computations via *Gretl* using BPI rolling estimation data.

When the period of formation of April bubble gets into the tested sample period it is observed that the series become non-stationary for a few periods (this conclusion is also supported by ADF tests' results). Whereas for the rest of the bubble period the assumption of stationarity of the series is valid. Unfortunately, this can't be claimed about the November bubble — the figure shows that when the peak of the bubble enters the sample periods the series become non-stationary. This state occurs once more in the end when the post-bubble period starts to influence the tests' result. The series are stationary for most of the time in samples where none of the two bubble peaks are present in the tested period.

7.2 GARCH rolling analysis

Moving to interpretation of the actual rolling results, Figure 7.2 shows the rolling coefficient estimates along with their significances.[3] First, we can notice that in the end of November bubble period both ARCH (α_1) and GARCH (β_1) estimated coefficients are statistically very insignificant. This is probably caused by the violation of the stationarity assumption. For that reason these estimates have been excluded from further interpretation of results.

Figure 7.2: **GARCH rolling estimation** — α_1 and β_1



Continuos line represent magnitude of the coefficients and dashed line their p-values. α_1 is in red and β_1 in blue.
Source: The Authors' own computations via *Gretl* using BPI rolling estimation data.

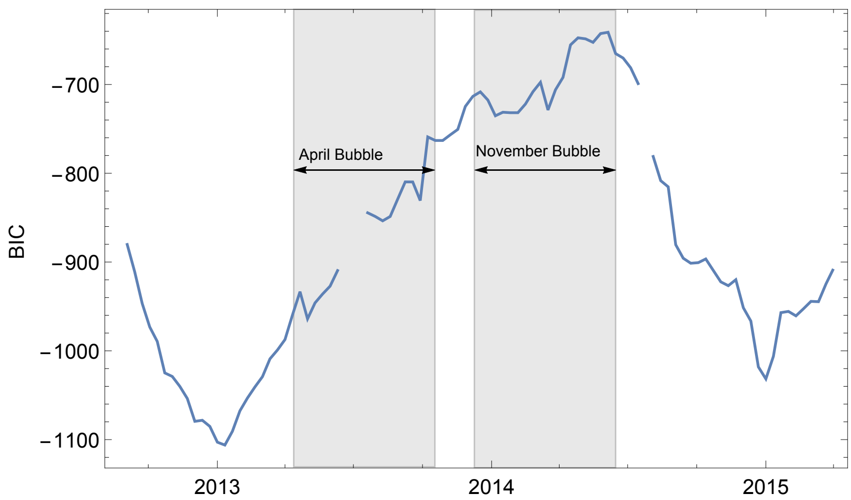
The ARCH coefficients are almost all statistically significant for the first half of the dataset, including the April bubble. But the coefficients become statistically insignificant for the rest of the estimated samples including the November bubble with the exception of the last few sample periods. On the other hand, the GARCH coefficients are all statistically significant except the last period and the period already excluded because of the non-stationarity.

It is obvious that the larger β_1 is often associated with the lower α_1 from looking at the magnitude of these two coefficients. Therefore, we observe that in April bubble the GARCH coefficients are larger and that the ARCH coefficients are lower in contrast with the November bubble when comparing their values

[3]The blank space in figures represents four sample periods which were not estimated as *Gretl* was unable to estimate these periods because of the problem with meeting the convergence criterion.

during the bubbles. This means that the predicted variance for current period has much larger influence over the new information entering the market in this period during the first bubble. While in the second bubble the influence of predicted variance is lower and the new information gain more importance, but become statistically insignificant so its importance is questionable. The simple average of estimated coefficients for the April bubble is α_1=0.277 and β_1=0.720 while for the November bubble (the coefficients from the second part of the period excluded) $\alpha_1 = 0.402$ and $\beta_1 = 0.584$.
We are not able to study the gap between the two bubbles in more detail because the period is not very large, but we are able to take further look at the pre-bubbles and post-bubbles periods. In pre-bubbles period the ARCH and GARCH coefficients get closer to each other and start to diverge again when the process of formation of the April bubble begins. As oppose to the post-bubbles period the GARCH coefficients become even higher and stay on relatively high level of values for the rest of the analyzed period, except the last sample period. This increase in β_1 suggests larger stability in volatility of Bitcoin returns as the predicted variance gain more influence and the values of this coefficient get closer to the estimated coefficients for stock market indexes (see Table 6.3) which can be considered as series with a low vol-of-vol of returns. These conclusions are consistent with the results of GARCH analysis presented in Section 6.5.

Figure 7.3: **GARCH rolling estimation** — BIC



Source: The Authors' own computations via *Gretl* using BPI rolling estimation data.

Lastly, we check the BIC information criteria for the rolling estimation samples. The results in Figure 7.3 show that during the bubbles, the GARCH process quality of estimation decreases due to higher volatility in data during these periods. Which again supports our conclusion from GARCH analysis (Section 6.2) that during the bubbles the GARCH model is not that efficient in capturing the development of volatility of Bitcoin returns.

Chapter 8

Log-Periodic Power Law model

Johansen & Sornette (2001) argues that GARCH(1,1) model does a reasonable job of reproducing fluctuations in "normal trading", but it is unable to capture the fluctuations connected with large crashes. Therefore, another type of model is necessary to estimate these fluctuations. Since the results of GARCH analysis, especially during the bubbles, supported Johansen & Sornette (2001) argument we implement the original Log-Periodic Power Law (LPPL) model for financial bubble modeling proposed by Johansen & Sornette (2001) and Johansen *et al.* (2000).

8.1 Methodology

Key assumption of the LPPL model is presence of two types of agents in the market. First, group of traders with rational expectations and second, irrational traders with herding behavior — the so called "noise" traders (Johansen & Sornette 2001).

According to Johansen & Sornette (2001) a crash is not certain, but it is characterized by its hazard rate $h(t)$ — the probability per unit of time that the crash will take place, given that it has not yet occurred. In practice it means that large group of agents places sell orders simultaneously and doing that creates the imbalance in the order book which market is unable to absorb without substantial decrease in prices. The key question here is to determine by what mechanism the traders decided for sudden coordinated sell off.

Johansen & Sornette (2001) and Johansen *et al.* (1999) also proposed the answer. They assume that agents are organized into a network (of friends, colleagues etc.) and that through this network they influence each other with the information exchange. According to their theory, market traders form their opinion based on two factors — opinion of other participants in the network (external factor) and idiosyncratic signals generated alone by the trader (internal factor). In "normal" times, buyers and sellers disagree with one another and submit approximately as many buy orders as sell orders — they balance each other out (internal factor overweights the external one) but when crash happens everybody agrees on the same opinion — selling (external factor overweights the internal one).

8.1.1 Price and hazard rate

Following Johansen & Sornette (2001) and Johansen *et al.* (1999), the simplest way to describe imitation model between traders is to assume that hazard rate evolves according to the following equation

$$\frac{dh(t)}{dt} = Ch^{\delta}, \qquad with \;\; \delta > 1, \tag{8.1}$$

where C is a positive constant and δ is the number of interactions between traders. Meaning that h^{δ} determines whether the hazard rate increases or decreases based on presence of interactions between traders. The condition $\delta > 1$ ensures that the critical point (crash) occurs in finite time.
Integrating Equation 8.1 we get power law dependence of the hazard rate

$$h(t) = B(t_c - t)^{-\alpha}, \qquad with \;\; \alpha = \frac{1}{\delta - 1}, \tag{8.2}$$

where B is a positive constant and t_c is the critical point or the most probable time of the crash. The exponent α must lie between 0 and 1 otherwise the hazard rate would go to zero and price to infinity when approaching to t_c.
The evolution of the price before the critical date, using hazard rate, is given by

$$p(t) \approx p_c - \frac{\kappa B}{\beta}(t_c - t)^{\beta} \tag{8.3}$$

where $z = 1 - \alpha \in (0, 1)$, p_c is the price at the critical time (conditioned the crash has not occurred yet) and $\kappa \in (0, 1)$ presents fixed percentage price drop during a crash.

8.1.2 Log-periodicity

Adding the log-periodicity, Johansen & Sornette (2001) and Johansen *et al.* (2000) generalize the hazard rate Equation 8.2 as follows

$$h(t) = B(t_c - t)^{-\alpha} + C(t_c - t)^{-\alpha} cos\left[\omega log(t_c - t) + \psi\right] \tag{8.4}$$

where $\frac{\omega}{2\pi}$ determines "log-frequency" of the oscillation term and ψ is a phase constant shifting the oscillation. The crash hazard rate still explodes near the critical date and in addition, it now displays the log-periodic oscillations. The evolution of the price before the critical date is then given by

$$p(t) \approx p_c - \frac{\kappa}{\beta}\left\{B(t_c - t)^{\beta} + C(t_c - t)^{\beta} cos\left[\omega\, log(t_c - t) + \phi\right]\right\} \tag{8.5}$$

where ϕ is another phase constant.
The key feature is that oscillations appear in the price just before the critical time. The ratio of successive time between local maxima of the function intervals tends to zero and is constant

$$\lambda = e^{\frac{2\pi}{\omega}}. \tag{8.6}$$

This scaling ratio is very useful for empirical examination and can be used for prediction of critical time t_c as it contains information about the frequency acceleration. Therefore, we use it as a key criterion in the analysis of critical time in the two bubble periods.

8.1.3 LPPL model

Based on Johansen & Sornette (2001) and Geraskin & Fantazzini (2013), the final LPPL model is obtained by rewriting the Equation 8.5 to a more suitable form for fitting financial time series. So the price evolution before the critical time is then defined as follows

$$p(t) \approx A + B(t_c - t)^{\beta} + C(t_c - t)^{\beta} cos\left[\omega\, log(t_c - t) + \phi\right]. \tag{8.7}$$

To summarize the estimated parameters, $A > 0$ is the price at the critical time $p(t_c)$, $B < 0$ embodies the scale of power law — the increase in p_t as the price approach closer to critical time, $C \neq 0$ is the magnitude of the oscillation around the price growth, β determines the power law acceleration of price and should satisfy the condition $\beta \in (0, 1)$ to ensure a finite price at the critical time, ω express the frequency of the oscillation term and ϕ is a fixed phase constant satisfying $\phi \in (0, 2\pi)$. Geraskin & Fantazzini (2013) also remarks that A, B, C and ϕ do not carry any structural information and are just units of distribution of β and ω parameters.

8.1.4 Fitting the LPPL parameters

The numerical procedure of fitting Equation 8.7 is minimization of the variance

$$\begin{aligned} Var &= \frac{1}{n}\sum_{t=t_1}^{t_n}(y_t - f(t))^2 \\ &= \frac{1}{n}\sum_{t=t_1}^{t_n}\left\{y_t - A - B(t_c - t)^{\beta} - C(t_c - t)^{\beta}cos\left[\omega log(t_c - t) + \phi\right]\right\} \end{aligned} \quad (8.8)$$

between n data points y_t and the fit function $f(t)$ i.e. sum of the squared differences between the actual and fitted prices. Further details of the procedure are given in Johansen *et al.* (2000) and Bree & Joseph (2013).
Seven parameters of the LPPL model need to be estimated and the chosen values of these parameters should be the ones that give the lowest root mean squared error (RMSE) between the actual and fitted price. RMSE is defined as

$$RMSE = \sqrt{Var} = \sqrt{\frac{1}{n}\sum_{t=t_1}^{t_n}(y_t - f(t))^2}. \quad (8.9)$$

Starting date

Other issue is to choose the right time interval prior to the crush to fit. According to Johansen & Sornette (2001) and Johansen *et al.* (1999), the last point should be the one with highest value of the price before the crash and the first point with the lowest value of the price when the bubble started. Given this indefinite description of starting date, we decided to fit the model for several local minima of Bitcoin price and choose the one with the lowest RMSE.

Raw price versus log price

Lastly, it needs to be decided whether to use raw price or logarithmic transformation of price. Johansen & Sornette (2001) introduce the assumption that one should use the raw price rather than the log price data if the following condition is satisfied. The increase in price from the beginning of the bubble is much lower than difference between the price at the beginning of the bubble and the fundamental value i.e.

$$p(t) - p(t_0) \ll p(t_0) - p_1, \tag{8.10}$$

where t_0 is the time of the beginning of the bubble and p_1 is the fundamental value. Bree & Joseph (2013) argues that even thought we usually do not know the fundamental price of the asset, the condition of Equation 8.10 is fulfilled only when $p(t) < 2p(t_0)$ i.e. the price rise during the bubble is lower than the price at the beginning of the bubble unless the fundamental value is negative. Therefore, we should use log price data rather than the raw data only if $p(t) > 2p(t_0)$.

8.2 Application to Bitcoin

In this section we apply the LPPL model to the two 2013 bubbles. In order to get enough observations and since the actual price growth during bubbles was rather quick (in terms of days) we are forced to use more frequent data then in original LPPL papers. However, the different time unit does not affect the coefficient estimation. Therefore, we use six hour closing price data from MtGox exchange for the analysis.

8.2.1 Choosing the starting and critical time

The time variable t in a model is constructed in a way that one unit of t corresponds to one year. For our purposes it is sufficient to match the initial position $t = 0$ to date 1.1.2013 (t=1 represents 1.1.2014 and we assume 365 days in a year).

April bubble

In case of April bubble the critical time t_c is nicely visible as there is only one peak in 6 AM 10.4.2013 which correspond to $t_c = 0.27192$ (see Figure 3.2). As a starting date we choose two different dates. First, 12 PM 23.3.2013 which matches to $t_0 = 0.22329$ leaving us 72 observations. This date was picked because it has the lowest RMSE among all other fitted dates. But since this day is practically in the middle of the bubble and it represents relatively small sample we decided to pick a second date which would better describe the definition of the starting date from Johansen & Sornette (2001). The second starting date is 12 PM 23.2.2013 corresponding to $t_0 = 0.14658$ and this sample contains 184 observations. Even though, it does not have the lowest RMSE from all fitted periods, it has the lowest RMSE among other local minima from period before the bubble.[1]

November bubble

The situation is rather different when analyzing the critical time for the November bubble as this period have two peaks with very close price values (see Figure 3.3). We decided to use the one occurred in 12 AM 30.11.2013 ($t_c = 0.91233$) as only this critical time give us consistent results with Johansen & Sornette (2001) assumptions about β coefficient. The most convenient starting date, among all others, appears to be 6 PM 2.10.2013 ($t_0 = 0.75274$), so the estimated period consists of 234 observations.[2]

8.2.2 Fitting to the log prices

For fitting the raw prices, the condition from Equation 8.10 needs to be satisfied. In Section 8.1.4 we show that this condition is fulfilled only when $p(t) < 2p(t_0)$. But from Table 8.1 it can be observed that the $p(t_c)/p(t_0)$ ratio is much larger than two therefore the condition from Equation 8.10 is violated for all bubble periods and we should use log prices. However, as Bree & Joseph (2013) argues, the Johansen & Sornette (2001) and Johansen *et al.* (1999) fit the LPPL with the raw indexes rather then their logs, even though they should not. We decided to estimate both raw and log Bitcoin prices in order to be able to compare the results.

Table 8.1: Ratio of Bitcoin prices on the critical time t_c to the prices on the starting date t_0

t_0	t_c	$p(t_0)$	$p(t_c)$	Ratio $p(t_c)/p(t_0)$
0.14658	0.27192	28.89	255.01	8.83
0.22329	0.27192	63.50	255.01	4.02
0.75274	0.91233	123.00	1228.66	9.99

Source: The Authors' own computations using MtGox six hour data.

[1] Various combinations of samples with different starting date had been used in the analysis. The results of these analyses are available upon request.

[2] Various combinations of samples with two different critical times — one presented in the paper and other occurring on 4.12.2013 — and several starting dates had been used in the analysis. The results of these analyses are available upon request.

8.3 Best model fits

We fit the LPPL model, defined in Equation 8.7, for the two sample periods in April bubble and one in the later one, each time for both raw and log prices.

Table 8.2: Fit LPPL parameters of Bitcoin bubbles

Bubbble L/R	t_0	t_c	β	ω	λ	A	B	C	ϕ	RMSE
A(23.2.-10.4.) L	0.147	0.272	0.52	7.04	2.44	5.64	-6.80	-0.26	6.94	0.04004
A(23.2.-10.4.) R	0.147	0.272	0.07	6.67	2.57	786.1	-885.7	-6.67	5.65	4.915
A(23.3.-10.4.) L	0.223	0.272	0.63	7.34	2.35	5.54	-9.12	-0.65	8.30	0.05542
A(23.3.-10.4.) R	0.223	0.272	0.25	6.65	2.57	336.2	-592.3	18.36	2.59	5.234
N(2.10.-30.11) L	0.753	0.912	0.66	3.08	7.69	7.25	-9.19	0.96	-0.86	0.07208
N(2.10.-30.11) R	0.753	0.912	0.23	3.05	7.87	1899.0	-2909.3	157.3	-0.73	38.396

L/R in the name of the bubble refer to whether we use raw or log prices in the estimation. For description of parameters see Equation 8.7. It is worth mentioning that all parameters presented in the table are statistically significant.
Source: The Authors' own computations via *Gretl* using MtGox six hour data.

From results of the fitted LPPL parameters, summarized in Table 8.2, we are mainly interested in parameters capturing the power law growth rate β, frequency of oscillation term ω and the scaling ratio λ, defined in Equation 8.6, as they carry the structural information. We also report graphical representation of all three periods where the actual and LPPL fitted prices are depicted. For the April bubble starting from 23. 2. and 23. 3. see Figure 8.1 and Figure 8.2 respectively. For November bubble see Figure 8.3.

First, it can be noticed that the results of the frequency of the fluctuations ω have approximately the same values regardless of using the raw or log prices. This does not apply for the growth parameter β where the coefficients differ significantly.
Johansen (2003) summarizes results of over 30 crashes on the major financial markets and he found that the distributions of fitted log frequencies and fitted growth exponents are $\omega \approx 6.36 \pm 1.56$ and $\beta \approx 0.33 \pm 0.18$. It is necessary to mention that these conclusions are based on fitting LPPL to raw indexes therefore when comparing the β coefficients (as log prices seem to have no significant effect on ω parameter) we need to look only at the fitted values of raw prices. On the other hand, Johansen & Sornette (2001) found that when examining emerging markets, which can be considered as less stable, the values of β and ω experience larger fluctuations (β coefficients seems to be slightly lower and on the contrary, ω higher). Moreover, Johansen & Sornette (2001) analyze stock

Figure 8.1: **LPPL fit** — April bubble (23. 2. - 10. 4.), raw and log Bitcoin prices

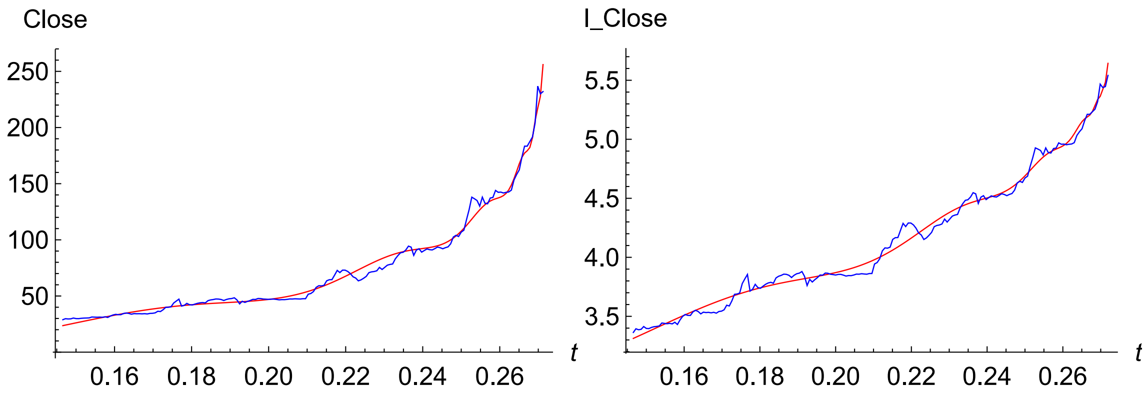


Figure on the left side represents raw price fit and figure on the right side log price fit. LPPL fitted values are captured by red line while the actual values by blue line. For parameter values of the fits see Table 8.2.
Source: The Authors' own computations via *Mathematica* using MtGox six hour data.

Figure 8.2: **LPPL fit** — April bubble (23. 3. - 10. 4.), raw and log Bitcoin prices

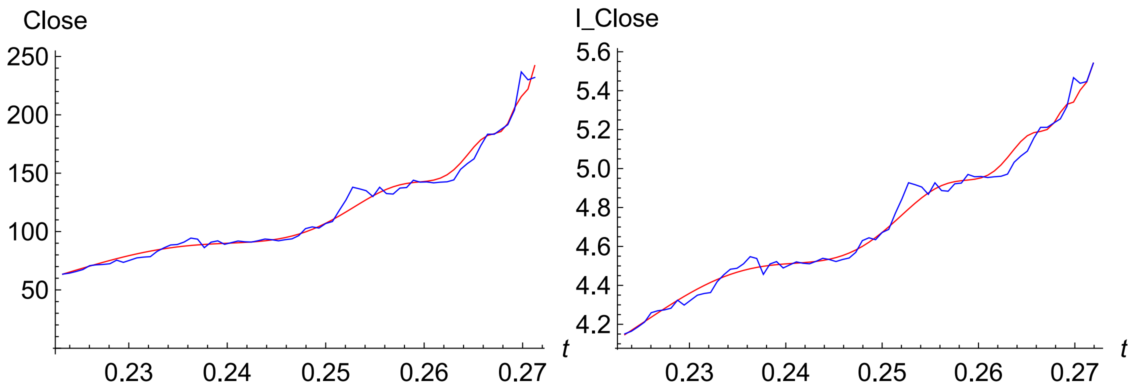


Figure on the left side represents raw price fit and figure on the right side log price fit. LPPL fitted values are captured by red line while the actual values by blue line. For parameter values of the fits see Table 8.2.
Source: The Authors' own computations via *Mathematica* using MtGox six hour data.

Figure 8.3: **LPPL fit** — November bubble (2. 10. - 30. 11.), raw and log Bitcoin prices

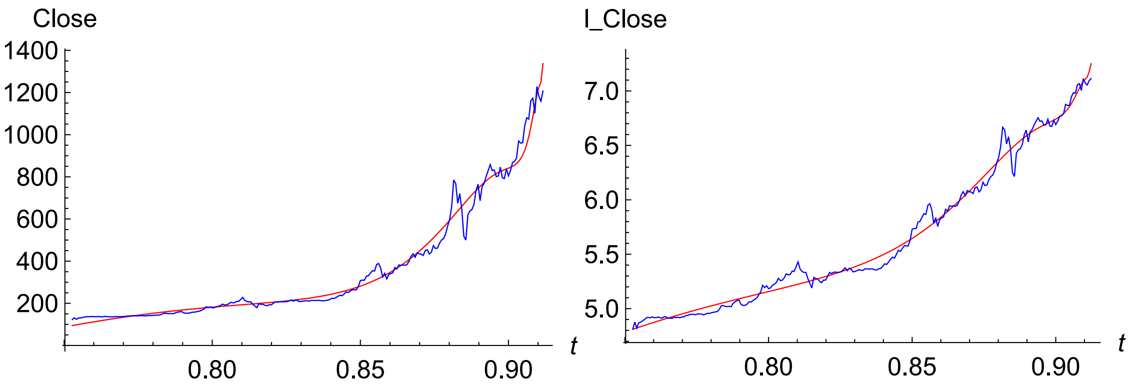


Figure on the left side represents raw price fit and figure on the right side log price fit. LPPL fitted values are captured by red line while the actual values by blue line. For parameter values of the fits see Table 8.2.
Source: The Authors' own computations via *Mathematica* using MtGox six hour data.

markets from Europe, Asia and Pacific and observe that the λ is very consistent $\lambda \approx 2.0 \pm 0.3$. Johansen *et al.* (1999) study the behavior of market indexes Wall Street and Hong-Kong (four crashes) before crashes as well as collapses of the USD against the DEM, CHF, CAD and JPY currencies. They notice the small fluctuations of $\lambda \approx 2.5 \pm 0.3$ for all data except the CHF. By further examination of ω and corresponding λ we notice that its value is also relatively stable and independent of the starting date. Based on these findings, it appears that frequency of log-oscillation ω can quite precisely indicate whether the bubble is about to crash while the growth coefficient β seems to be rather informative parameter about the speed of the price growth before the bubble crash.
Looking at the magnitude of β coefficients we observe that the November bubble acceleration rate of power law is lower as β is higher for both log and raw prices.[3] Moreover, the β of April bubble is outside the distribution of fitted growth exponent, found in Johansen (2003), as it is very low suggesting very fast acceleration of price before the crash. Contrarily, the November bubble growth coefficient is inside the range and indicating the upcoming crash.
The situation is opposite to β coefficient results when examining the ω and λ. In November bubble the λ suggests faster log-periodic oscillations but much faster (larger λ) than the one found in Johansen & Sornette (2001) and Johansen *et al.* (1999) while in April Bubble period the scaling ratio coincides

[3] We exclude the April bubble (23. 3. - 10. 4.) from comparison of two bubbles as it does not match properly the definition of starting date, as mentioned in Section 8.2.1.

with these findings and successfully indicate the crash. But we will study the scaling ratio of these two bubbles in more detail in the following section.

8.4 Loop analysis

In this section we try to capture the development of the LPPL parameters β, ω and λ before and after the crash to see the coefficients' changes as we proceed through the bubble. For that we use "loop analysis" where we fix the starting date of the bubbles and fit the LPPL model from Equation 8.7 while moving with the critical time with step of six hours i.e. one observation. We choose the starting date based on our LPPL analysis in previous section, therefore for April bubble we choose the date 22. 3. 2013 ($t_0 = 0.147$) and for November bubble 2. 10. 2013 ($t_0 = 0.753$). Critical time period for the earlier bubble is from 6 PM 26. 3. 2013 (t_{c_1}=0.232) to 6 PM 15. 4. 2013 ($t_{c_{81}}$=0.287), in the later bubble we study the period from 6 AM 6. 11. 2013 (t_{c_1}=0.847) to 12 AM 10. 12. 2013 ($t_{c_{136}}$=0.940).
In Figure 8.4 and Figure 8.6 we report development of log prices and β and ω coefficients of April and November bubble, respectively. Results of the "loop analysis" for scaling ratio of April and November bubble are depicted in Figure 8.5 and Figure 8.7, respectively.

Figure 8.4: **Loop** — *l_close*, β, ω — April (26. 3. - 15. 4.)

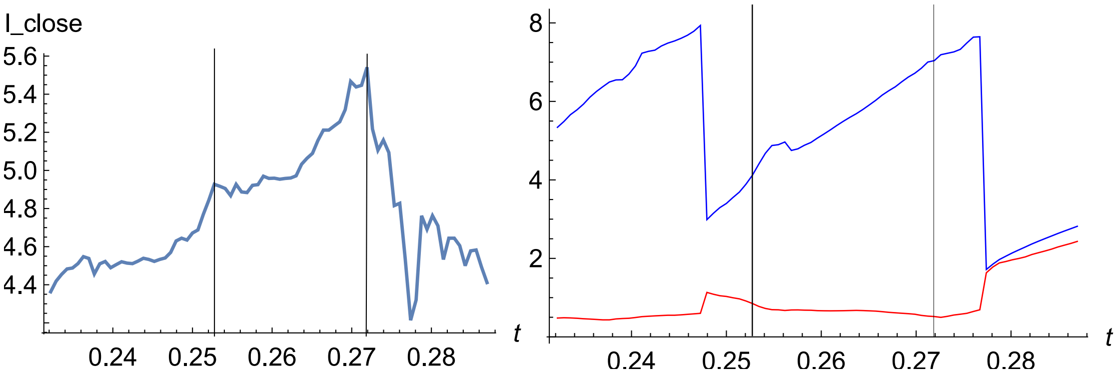


Figure on the left side represents log price data. Figure on the right side shows β coefficient development (red line) and ω coefficient development (blue line) over time period. The black vertical lines represent peaks occurred during the bubble (second peak is the actual crash of the bubble).
Source: The Authors' own computations via *Gretl* using MtGox six hour data.

Figure 8.5: **Loop** — λ, $\Delta\lambda$ — April (26. 3. - 15. 4.)

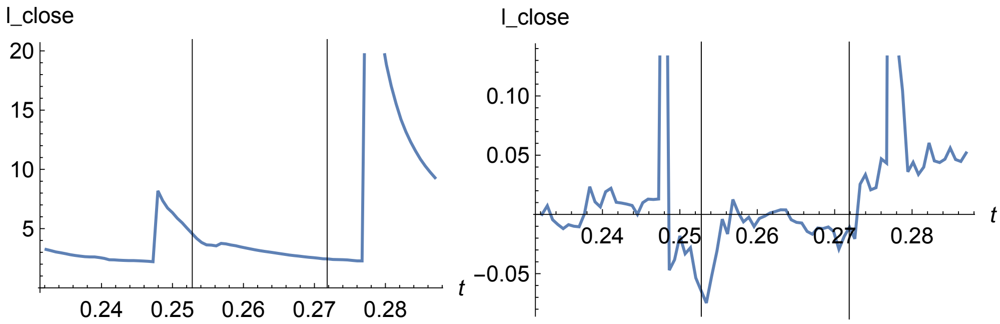


Figure on the left side represents loop for log scaling ratio λ. Figure on the right side shows development of first differences of λ i.e. $\Delta\lambda = \lambda_t - \lambda_{t-1}$. The black vertical lines represent peaks occurred during the bubble (second peak is the actual crash of the bubble).

Source: The Authors' own computations via *Gretl* using MtGox six hour data.

Figure 8.6: **Loop** — *l_close*, β, ω — November (6. 11. - 10. 12.)

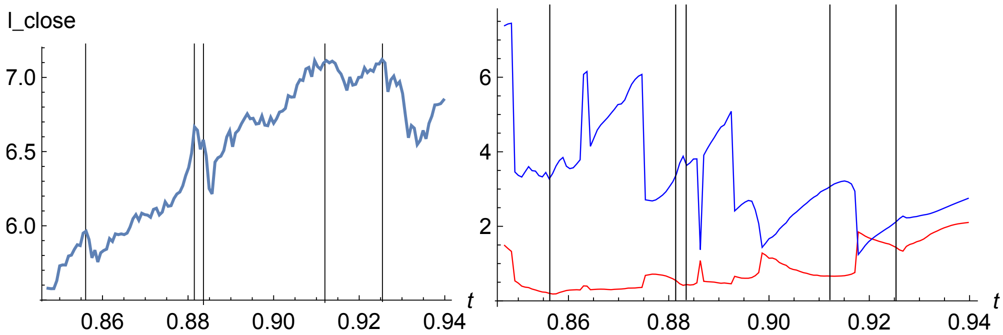


Figure on the left side represents log price data. Figure on the right side shows β coefficient development (red line) and ω coefficient development (blue line) over time period. The black vertical lines represent several potential peaks occurred during the bubble (last peak is the actual crash of the bubble).

Source: The Authors' own computations via *Gretl* using MtGox six hour data.

Figure 8.7: **Loop** — λ, $\Delta\lambda$ — November (6. 11. - 10. 12.)

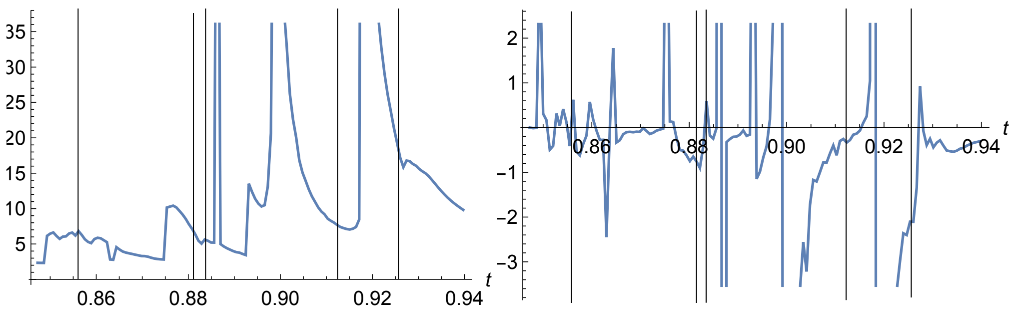


Figure on the left side represents loop for log scaling ratio λ. Figure on the right side shows development of first differences of λ i.e. $\Delta\lambda = \lambda_t - \lambda_{t-1}$. The black vertical lines represent several potential peaks occurred during the bubble (last peak is the actual crash of the bubble).

Source: The Authors' own computations via *Gretl* using MtGox six hour data.

8.4.1 Power law acceleration results

From the results of the "loop analysis" for April bubble (see Figure 8.4) we notice that before the actual crash the β experiences slow decreases, meaning the growth is getting faster. In $t_c = 0.255$ the parameter significantly decrease suggesting the crash could occur. However this state where $\beta \approx 0.6 \pm 0.1$ lasts until the actual crash so the predicted power of β does not seem to be so significant. The behavior before the first potential peak (first black line) does not (correctly) indicate any crash.

Examining the November bubble (see Figure 8.6), the development of β before the first potential peak (first black line) suggests extremely fast growth of the power law as the coefficient relatively quickly and significantly declines. The same situation occurs before the second and third peaks. But the crash becomes more clear with the third peak ($t_c = 0.884$) as β losses almost half of its value in just sixty hours. The fourth peak is the one used in LPPL fit in Section 8.3. Drop in β precedes to this peak as well but in this case the parameter value was not that small indicating that the growth was not that fast. Even though the β declines before the actual crash as well, the value of beta does not show any sign of crash.

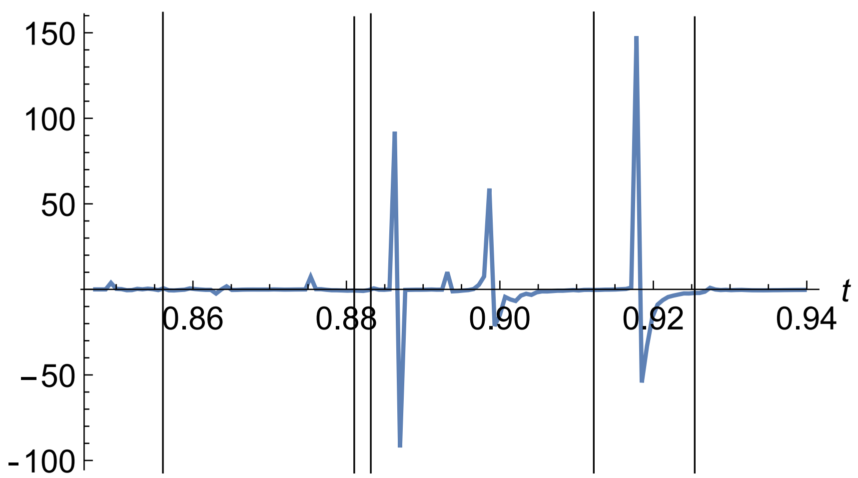
In conclusion, we can say that based solely on β the April bubble should end much earlier than it actually ends. On the other hand, in the November bubble period it suggests correctly the small drop in prices before the peak but the actual crash was not foreseen at all. However, this is not so surprising as the β expresses only the growth of power law and the log price development right before the crash does not suggest any extreme growth acceleration.

8.4.2 Scaling ratio results

In this section, we further examine relationship between crashes and the scaling ratio λ but also its first difference $\Delta\lambda$. The value of λ suggest that before the bubble, the coefficient is close to range $2.0 \leq \lambda \leq 3.0$, as noted in Section 8.3. In the April bubble loop results (see Figure 8.5) we can see that λ approaches to this range in $t_c = 0.234$ and stays there until $t_c = 0.247$. But no crash occurs during or after this period suggesting false alarm of a bubble. The scaling ratio again enters this range in $t_c = 0.264$ indicating the end of the bubble which indeed happens 66 hours later. By looking at the $\Delta\lambda$ we notice that before the crash, the λ is experiencing decreasing trend in its value. Also after this trend

changes i.e. the $\Delta\lambda$ is increasing, and the $\Delta\lambda$ is getting closer towards zero the bubble usually ends. Following this logic, in April bubble we can notice three of these situations. First, in period before $t_c = 0.237$ where indeed small drop in log price appears. Second, $t_c \in \langle 0.249, 0.256 \rangle$ period during which the market potential peak appears, however no crash of a bubble occurs. Therefore this can be considered again as a false alarm. By the end of the last period $t_c \in \langle 0.264, 0.273 \rangle$ and after the change of the trend of $\Delta\lambda$ the actual peak of the bubble appears thus correctly predicting the crash.
We try to verify the hypothesis about connection between $\Delta\lambda$ and the crashes of the bubble by studying the November bubble (see Figure 8.7). The λ is entering the range $2.0 \leq \lambda \leq 3.0$ only in two subperiods during the bubble. The first ends in $t_c = 0.849$ and is not followed by the crash. In the second subperiod, $t_c \in \langle 0.872, 0.875 \rangle$ λ indicates the second and third peak, but in advance of 78 hours. In further development λ always decreases before the potential peak but never enters this range again. For that reason we decided to take a closer look at $\Delta\lambda$.

Figure 8.8: **Loop** — $\Delta\lambda$ — November (6. 11. - 10. 12.), full scale range



The black vertical lines represent several potential peaks occurred during the bubble (Last peak is the actual crash of the bubble)
Source: The Authors' own computations via *Gretl* using MtGox six hour data.

It can be seen that in November bubble the trend of scaling ratio growth is quite unstable. However, following the Figure 8.7 and Figure 8.8 we detect several prolonged periods in which $\Delta\lambda$ substantially decreases and then returns close to zero level. Initially, in the period $t_c \in \langle 0.861, 0.864 \rangle$ $\Delta\lambda$ quickly changes its trend, but no crash occurs. In the following subperiod $t_c \in \langle 0.865, 0.875\rangle$ the crash does not occur either. In the next noticeable period $t_c \in \langle 0.877, 0.883\rangle$ the second and third potential peak appear and $\Delta\lambda$ correctly predicts a drop in log prices. In the next subperiod $t_c \in \langle 0.887, 0.892\rangle$ $\Delta\lambda$ experiences an extreme drop after which log price slightly decreases. Now we are getting to the $t_c \in \langle 0.899, 0.915\rangle$ in which our fitted peak by LPPL is situated. It can be seen that the $\Delta\lambda$ is increasing and is approaching towards zero when the peak occurs. The situation is very similar in the last subperiod $t_c \in \langle 0.918, 0.926\rangle$ where the bubble ends.

Therefore, we can claim that the actual crash and the potential peaks occurred during the bubble were predicted with a good precision. However, we should add that in same cases the $\Delta\lambda$ indicator predicts a false alarm of the crash.

Chapter 9

Conclusion

Aim of this thesis is to examine the Bitcoin Price Index and volatility of its returns. In the first part we define the Bitcoin term and summarize Kristoufek (2014) and Ciaian *et al.* (2014) conclusions that Bitcoin price formation is mainly driven by investors' interest in the currency and by standard supply-demand factors. In the following chapters we study the volatility using standard time series econometric tools — ARIMA and GARCH approach — with focus on three periods — Whole sample, the April 2013 and the November 2013 bubble. Concerning the ARIMA analysis, the BIC suggested ARIMA(0,1,0) model for the Whole sample as well as for the November bubble period and ARIMA(1,1,0) model for April bubble sample. Performing ARCH-LM test, we found that ARIMA processes provide weak information about the real behavior of the Bitcoin price returns.

Therefore, we estimated GARCH, TGARCH and EGARCH processes as they are more suitable for modeling of Bitcoin returns in presence of heteroskedastic squared residuals. Based on the final estimates of the GARCH-analysis and rolling GARCH estimation, the best model for all sample periods is GARCH(1,1). Comparing the GARCH(1,1) models of Bitcoin with global stock market indices and FOREX rates, we observe that the Bitcoins' GARCH volatility is more volatile i.e. more "spiky" which suggests lower stability of the Bitcoin market. Furthermore, we find several differences in behavior of the April and the November bubble. The November bubble seems to have higher volatility of GARCH volatility and therefore it is less stable. Moreover, based on Johansen & Sornette (2001) summarized theory of financial bubbles and crashes we claim that during the November bubble the number of irrational investors entering the market was much higher and that investors formed their expectation about the profit based on April bubble experience rather than on current market condition. This probably caused larger appreciation of Bitcoin price and consequently much higher market instability during the November bubble. However, the results of the Ljung-Box test suggest serially correlated squared residuals of fitted GARCH model for all of the three periods which question the estimated results.

Therefore, we implement another model designed to capture the price development during the financial bubbles and crashes — the LPPL introduced by Johansen *et al.* (2000). We study the key variables capturing the power law growth rate β, frequency of oscillation term ω and corresponding scaling ratio λ defined in Equation 8.6. We observe that the November bubble has lower acceleration rate of power law as the β is higher for this period and its value is in a range suggesting the crash. This can not be said about the April bubble where the coefficient is outside of the range. Looking at the scaling ratio λ, the April bubble values are within the range suggesting the crash but in the case of the November bubble, λ suggests much faster oscillation and does not indicate a crash.

Since the results of standard LPPL model were not as we expected, we decided to perform "loop analysis" where we fixed the starting date and we were moving with the critical time as we went through the bubble. We examined the development of β, ω and especially λ over time. The LPPL model appears to be able to predict the April bubble crash better than the November bubble crash, but the results are still not very convincing. Therefore, we come up with an idea to study changes of λ as time gets closer to the crash and we find out that when $\Delta\lambda$ is increasing and is getting closer to zero the crash is about to happen. The indicator correctly detects the crashes over the bubble period and in spite of an occasional prediction of a false alarm, it appears to be quite precise. Therefore, our hypothesis that $\Delta\lambda$ serves as a better indicator of upcoming crash seems be true. However, to verify our hypothesis, analysis of more financial bubbles should be carried out.

Bibliography

ALEXANDER, C. (2008): *Market risk analysis*. John Wiley & Sons. ISBN 978-0-470-99789-5.

BAMBANI, R. & F. A.-S. BEER (2013): "Bubbles and Bitcoins." *Technical report*, Erasmus Univesiteit Rotterdam, Erasmus School of Economics.

BBC (2014): "MtGox bitcoin exchange files for bankruptcy." *BBC journal,* Retrieved 21 March 2015, from http://www.bbc.com/news/technology-25233230.

BOLLERSLEV, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **31(3)**: pp. 307 – 327.

BREE, D. S. & N. L. JOSEPH (2013): "Testing for financial crashes using the Log Periodic Power Law model." *International Review of Financial Analysis* **30**: pp. 287–297.

CIAIAN, P., M. RAJCANIOVA, & D'ARTIS KANCS (2014): "The Economics of BitCoin Price Formation." *Papers 1405.4498*, arXiv.org.

COINDESK (2015): "About the Bitcoin Price Index." Retrieved 2 May 2015, from http://www.coindesk.com/price/bitcoin-price-index/.

COINREPORT (2014): "What are the Advantages and Disadvantages of Bitcoin?" Retrieved 21 March 2015, from https://coinreport.net/coin-101/advantages-and-disadvantages-of-bitcoin/.

DURNEL, J. (2012): "Applied Econometric Time Series, ARCH GARCH Modelling." *Technical report*, Lund University, Sweden.

ENGLE, R. (2001): "GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives* **15(4)**: pp. 157–168.

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* **50(4)**: pp. pp. 987–1007.

ENGLE, R. F. & A. J. PATTON (2000): "What Good is a Volatility Model." *Quantitative Finance* pp. 237–245.

FRANSES, P. H. & D. v. DIJK (2000): *Nonlinear time series models in empirical finance*. Cambridge University Press. ISBN 0 511 01100 8.

GERASKIN, P. & D. FANTAZZINI (2013): "Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask." *The European Journal of Finance* **19(5)**: pp. 366–391.

JIANG, W. (2012): *Using the GARCH model to analyze and predict the different stock markets*. Master's thesis, Uppsala Universitet, Department of Statistics, Sweden.

JOHANSEN, A. (2003): "Characterization of large price variations in financial markets." *Physica A Statistical Mechanics and its Applications* **324**: pp. 157–166.

JOHANSEN, A., O. LEDOIT, & D. SORNETTE (2000): "Crashes as Critical Points." *International Journal of Theoretical and Applied Finance* pp. 219–255.

JOHANSEN, A. & D. SORNETTE (2001): "Bubbles and anti-bubbles in Latin-American, Asian adn Western stock markets: An empirical study." *International Journal of Theoretical and Applied Finance* **4(6)**: pp. 853–920.

JOHANSEN, A., D. SORNETTE, & L. OLIVIER (1999): "Predicting Financial Crashes Using Discrete Scale Invariance." *Journal of Risk* **1(4)**: pp. 5–32.

KRISTOUFEK, L. (2013): "BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era." *Sci. Rep. 3, 3415;* DOI: 10.1038/srep03415.

KRISTOUFEK, L. (2014): "What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis." *Papers 1406.0268*, ArXiv.org.

KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT, & Y. SHIN (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root

: How sure are we that economic time series have a unit root?" *Journal of Econometrics* **54(1-3)**: pp. 159–178.

Luthkepohl, H. & M. Kratzig (2007): *Applied Time Series Econometrics.* Cambridge University Press. ISBN 9780521839198.

Nakamoto, S. (N/A): "Bitcoin: A peer-to-peer electronic cash system,? http://bitcoin.org/bitcoin.pdf."

Nelson, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* **59(2)**: pp. pp. 347–370.

Pieters, G. & S. Vivanco (2015): "Bitcoin arbitrage and unofficial exchange rates." *Technical report*, Trinity University, San Antonio, Texas.

Safka, J. (2014): *Virtual currencies in real economy: Bitcoin.* Master's thesis, Charles Universitt in Prague, Faculty of Social Sciences, Institute of Economic Studies.

SAS Online Documentation (2015): "GARCH, IGARCH, EGARCH, and GARCH-M Models." Retrieved 9 May 2015, from http://v8doc.sas.com/sashtml/ets/chap8/sect22.htm.

Surda, P. (2012): *Economics of Bitcoin: is Bitcoin an alternative to fiat currencies and gold?* Master's thesis, University of Economics and Business, Vienna.

Voorhees, E. (2015): "What is Bitcoin?" *Bitcoin Magazine,* Retrieved 20 March 2015, from https://bitcoinmagazine.com/19020/bitcoin/.

Zakoian, J.-M. (1994): "Threshold heteroskedastic models." *Journal of Economic Dynamics and Control* **18(5)**: pp. 931 – 955.

Appendix A

Supplementary figures and tables

Table A.1: Summary Statistics for the variable ld_ClosePrice (1199 valid observations)

Mean	Median	Minimum	Maximum
0.00361522	0.00250672	−0.372425	0.347767
Std. Dev.	C.V.	Skewness	Ex. kurtosis
0.0535070	14.8005	−0.507002	10.0723
5% perc.	95% perc.	IQ Range	Missing obs.
−0.0684330	0.0858007	0.0353056	0

Source: The Authors' own computations via *gretl* using BPI data.

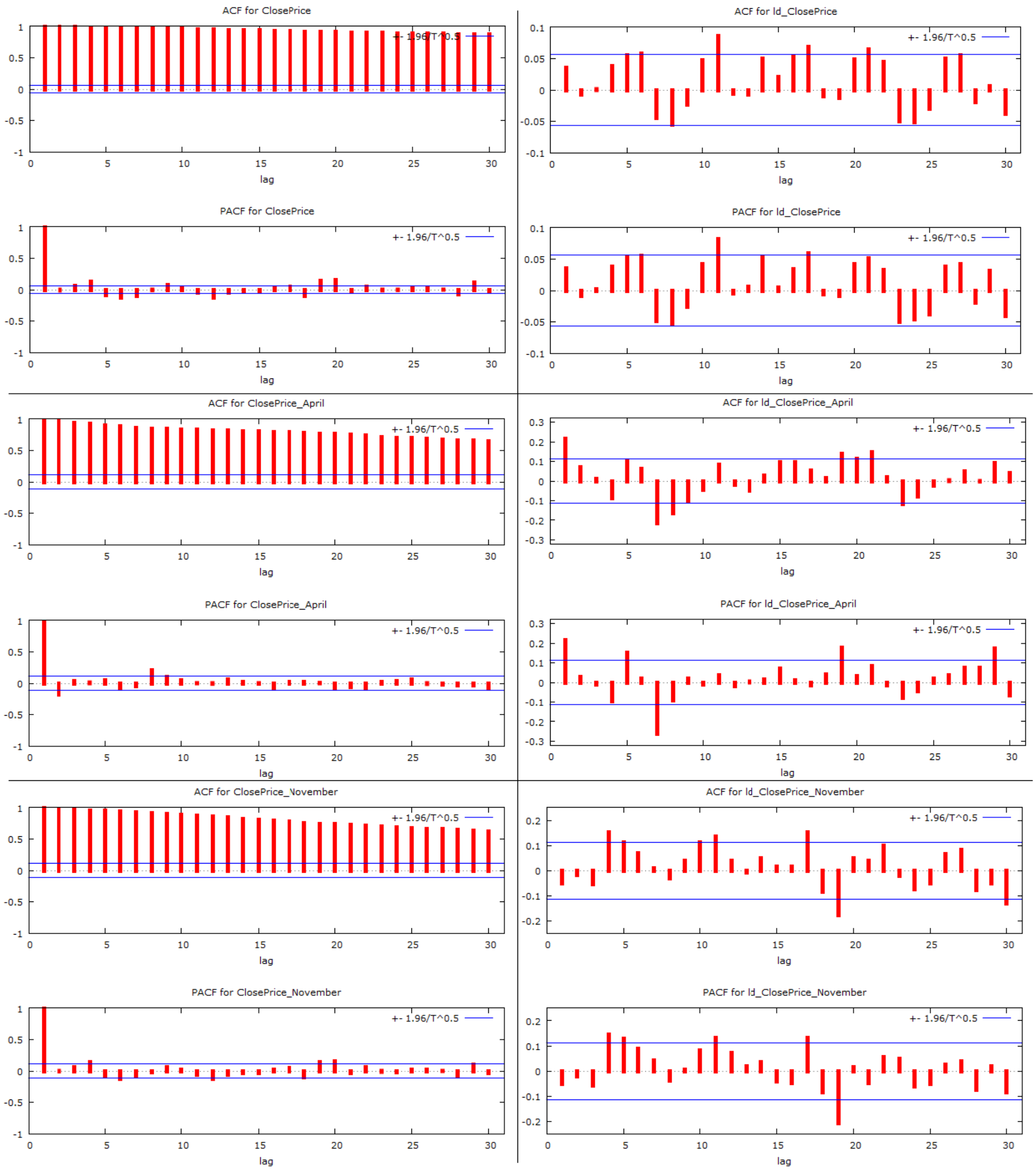
Table A.2: ARMA model — results

	Coefficient	**z-statistic**
Whole sample		
ARMA (0,0)		
α_0	0.004**	2.340
April bubble		
ARMA (1,0)		
α_0	0.008*	1.872
α_1	0.216***	3.828
November bubble		
ARMA (0,0)		
α_0	0.006	1.558

Significance levels : * : 10% ** : 5% *** : 1%

Source: The Authors' own computations via *Gretl* using BPI data.

Figure A.1: **ACF and PACF**, Whole sample, April bubble, November bubble



Source: The Authors' own computations via *Gretl* using BPI data.