

Dissection of Bitcoin’s Multiscale Bubble History

from January 2012 to February 2018

J.C. Gerlach^{†*}, G. Demos[†], D. Sornette^{†‡}

[†] ETH Zürich, Dept. of Management, Technology and Economics, Zürich, Switzerland

[‡] Swiss Finance Institute, c/o University of Geneva, Geneva, Switzerland

Wednesday 18th April, 2018

Abstract

We present a detailed bubble analysis of the Bitcoin to US Dollar price dynamics from January 2012 to February 2018. We introduce a robust automatic peak detection method that classifies price time series into periods of uninterrupted market growth (drawups) and regimes of uninterrupted market decrease (drawdowns). In combination with the *Lagrange Regularisation Method* for detecting the beginning of a new market regime, we identify 3 major peaks and 10 additional smaller peaks, that have punctuated the dynamics of Bitcoin price during the analyzed time period. We explain this classification of long and short bubbles by a number of quantitative metrics and graphs to understand the main socio-economic drivers behind the ascent of Bitcoin over this period. Then, a detailed analysis of the growing risks associated with the three long bubbles using the *Log-Periodic Power Law Singularity* (LPPLS) model is based on the *LPPLS Confidence Indicators*, defined as the fraction of qualified fits of the LPPLS model over multiple time windows. Furthermore, for various fictitious present analysis times t_2 , positioned in advance to bubble crashes, we employ a clustering method to group LPPLS fits over different time scales and the predicted critical times t_c (the most probable time for the start of the crash ending the bubble). Each cluster is argued to provide a plausible scenario for the subsequent Bitcoin price evolution. We present these predictions for the three long bubbles and the four short bubbles that our time scale of analysis was able to resolve. Overall, our predictive scheme provides useful information to warn of an imminent crash risk.

Keywords: Cryptocurrency, Bitcoin, k-Means Clustering, Multiscale Bubble Indicator, Log-Periodic Power Law Singularity Analysis, Forecasting, Time Series Analysis, Market Crashes.

JEL Code: C2, C13, C32, C53, C55, C61, G1, G10.

*Corresponding author: jgerlach@ethz.ch

I. Introduction

From an investment point of view, in the past decade, Bitcoin has become known for two main reasons: its extraordinary return potential in phases of extreme price growth as well as regular massive crashes. For instance, as a consequence of the most recent crash in mid-December 2017, a book-to-market value of more than 200 Billion US Dollars of Bitcoin’s total market capitalization evaporated within only six weeks, resulting in a cumulative loss from the peak of about two thirds in market value of the digital currency. The massive crash was preceded by a no less impressive fiftyfold price boost over a time period of just two years. Given the turbulent market history that Bitcoin has undergone since its inception in 2008 ([Nakamoto, 2008](#)), the recent bubble is not exceptional. In fact, as we will demonstrate in this paper, multiple overlapping short- and long-term Bitcoin price bubbles have appeared between 2012 and 2018. It is our purpose here to document this sequence of bubbles and crashes.

At the time of writing, the combined capitalization of all existing cryptocurrencies still amounts to less than one percent of the world GDP ([Mossavar-Rahmani et al., 2018](#)), a fact illustrating the still low significance of this market in the global economic context. Nevertheless, cryptocurrencies, and especially Bitcoin as a precursor of this new asset class, have drawn increased attention of the scientific and investment communities, due to the strong growth of the sector over the past years as well as the technological and economic prospects that are promised.

There have already been a number of studies examining the statistical properties of Bitcoin returns. [Pichl and Kaizoji \(2017\)](#) model the time-varying realized volatility of Bitcoin and found it to be significantly larger compared to that of fiat currencies. [Urquhart and Zhang \(2018\)](#) study a variety of GARCH volatility models and test the hedging capability of the crypto-coin against other currencies. The hedging properties against other asset classes were investigated by [Bouri et al. \(2017\)](#). [Bariviera \(2017\)](#) provides evidence for volatility clustering through a long memory Hurst exponent analysis. Furthermore, [Osterrieder and Lorenz \(2017\)](#) find much larger magnitude in the heavy tail of the Bitcoin return distribution compared to conventional currencies. Additionally, [Begušić et al. \(2018\)](#) determine even larger tail risk than usually seen in stocks. [Donier and Bouchaud \(2015\)](#) investigate Bitcoin liquidity based on order book data and, from this, accurately predict the size of price crashes. Other approaches to econometric Bitcoin modeling are outlined in [Fantazzini et al. \(2016\)](#).

Besides the purely statistical properties of the Bitcoin financial time series, there has been growing interest in the social component shaping Bitcoin price dynamics, as well. [Kristoufek \(2013\)](#) firstly observed a bidirectional relationship between web search queries and Bitcoin prices. Moreover, [Garcia et al. \(2014\)](#) detect positive feedback loops between Bitcoin prices, user numbers of the Blockchain network and search queries. They successfully implement a profitable Bitcoin trading strategy exploiting these social dynamics ([Garcia and Schweitzer, 2015](#)). Likewise, [Glaser et al. \(2014\)](#) investigate the link between Bitcoin prices, Blockchain network and search query data. They observe that “Bitcoin users are limited in their level of professionalism and objectivity” because they primarily utilize Bitcoin as a speculative investment. An analysis of the Bitcoin user base in more depth is carried out by [Kim et al. \(2017\)](#) who apply topic modelling to Bitcoin forum posts and test the predictive power of a deep learning model that is trained on the used forum data. Further contributions to the characterization of the Bitcoin community are also summarized in [Fantazzini et al. \(2016\)](#).

These studies of the dynamics of Bitcoin’s price and of its associated social dynamics suggest that Bitcoin buyers have been attracted by the sky-rocketing price performance of the cryptocurrency and have been influenced by news and social media. This is typical of previous financial bubbles. Indeed, the media

as well as many pundits have drawn parallels between the Bitcoin phenomenon and former extraordinary financial bubbles such as the Tulip Mania or the Mississippi bubble. In the study presented below, we provide confirming evidence and quantitative analyses that strongly support the conclusion that Bitcoin has behaved as a highly speculative asset exhibiting strong bubble activity.

In order to test for financial bubbles, we use the *Log-Periodic Power Law Singularity* (LPPLS) model. First introduced as a calibration model by Sornette et al. (1995), it was elaborated into a rational expectation bubble model in (Johansen et al., 1999, 2000) to provide real-time diagnostics of bubbles and forecasts of crashes. A few studies have already applied the LPPLS model to analyse a few Bitcoin bubbles (see for instance (Macdonell, 2014; Bianchetti et al., 2018a,b)). Here, we present a comprehensive general methodology to identify and classify the complete set of Bitcoin bubbles (over time scales larger than one month) that occurred within the studied time period on the Bitcoin to US Dollar exchange rate series, both at short and long time scales. The methodology for bubble characterization is completely automatized and can be applied to any asset price time series, making it a robust measure for bubble detection.

We identify three massive long bubbles and ten short bubbles. The temporal coexistence of these bubbles underscores the multiscale nature of bubbles. It motivates the development of metrics that can be used to separately identify evolving short and long bubble dynamics. We explore two techniques for this: (i) the LPPLS Multi-Scale Confidence Indicator (Sornette et al., 2015; Zhang et al., 2016) and (ii) the characterization of bubble burst scenarios obtained by clustering the set of LPPLS calibrations over time scales and predicted critical times. The forecast ability of the metrics to diagnose bubbles in real time and predict crashes is evaluated. The present study complements and extends the one presented by Wheatley et al. (2018), which focused on the identification of a proxy for the fundamental value of Bitcoin and a preliminary assessment of predictability of a subset of the bubbles studied here. Here, we dwell much more in exploiting the information contained in the LPPLS model calibration using the two mentioned sophisticated indicators.

The paper is structured as follows. Section 2 describes firstly the generalised *Epsilon Drawdown Method* that serves to identify the end times of long and short bubbles from January 2012 to February 2018 in the Bitcoin price dynamics. The second part of the section explains how the *Lagrange Regularisation Approach* provides an efficient determination of the beginning of the bubble regimes. The section concludes with a summary of the main characteristics of the identified long and short bubbles. Section 3 analyses the main socio-economic drivers behind the three long bubbles, and presents a series of useful metrics. Section 4 simulates a real-time assessment of the risks of the on-going bubbles. Section 5 concludes.

II. Bitcoin Bubbles and Crashes

Before performing a predictive bubble analysis on Bitcoin, we need to construct a catalog of the bubbles that actually occurred. Identifying a bubble ex-post would seem to be a straightforward job, but it is not. A bubble could perhaps be defined as a large abnormal price increase, which then bursts in a crash. This intuitive description is however quite naive, hard to conceptualize and full of traps, as it requires the implicit definition of “abnormal price growth” and “crash”. In order to measure abnormal price increases, a reference frame or process against which deviations can be gauged, must be defined, with the obvious criticism that, when employing such a reference process, a bubble may be incorrectly diagnosed due to an untrue underlying benchmark model, an issue that makes the diagnostic of a bubble a joint-hypothesis problem (Fama, 1991). Moreover, a crash is also not easy to define, as it is a mixture of a large loss over some relatively short duration that seems exceptional compared to the regular asset price movements. In the literature, we find a plethora

of different definitions, giving an overall impression of unpleasant arbitrariness. Thus, in this section, we develop a methodology that combines several metrics to systematically and automatically identify the start and end of long and short bubbles, the size and duration of the subsequent crashes, as well as quantitative characteristics of the bubbles. The bubbles and following crashes, which are diagnosed in this way, will provide the targets for our subsequent predictive analyses.

A. Identification of the Peak Times of Long and Short Bubbles

Here, we present a straightforward method to identify the peaks of each bubble, where changing the value of a single parameter allows us to scan bubbles of different durations. Our bubble diagnostic algorithm is based on an extension of the *Epsilon Drawdown Method* developed by [Johansen and Sornette \(1998, 2001\)](#) and further used in ([Johansen and Sornette, 2010](#); [Filimonov and Sornette, 2015](#)). See Appendix A for the detailed description.

The key idea is to identify a bubble as a price run-up, the “drawup”, defined as a succession of positive returns that may be interrupted by negative returns no larger in amplitude than some pre-specified tolerance level ε . The bubble crash or correction phase corresponds to the subsequent “drawdown”, defined as a succession of negative returns that may be interrupted by positive returns no larger in amplitude than the pre-specified tolerance level ε . By construction, drawup and drawdown periods must be alternating and directly following each other. In this way, a price trajectory can be systematically decomposed into a sequence of alternating, consecutive price drawup and drawdown phases.

The tolerance parameter ε is defined as the product of the realized volatility σ , estimated over a running time window of duration w , and a constant, pre-set multiplier ε_0 . The larger ε_0 or σ , the bigger are the oppositely directed price movements that are tolerated in a given developing drawup or drawdown. A drawup (resp. drawdown) ends when a negative (resp. positive) (successive) return, whose amplitude exceeds ε , is observed.

In order to obtain robust results, we scan ε_0 from 0.1 to 5 in steps of 0.1 and w from 10 to 60 days in steps of 5 days, yielding $50 \times 11 = 550$ pairs (ε_0, w) . For each pair, we run the *Epsilon Drawdown Method* over the time series of the Bitcoin price expressed in US Dollars (*btc/usd*) from Jan. 2012 to Jan. 2018, and obtain a unique sequence of drawups and drawdowns. For each sequence, we record the peak times $\{t_p\}$, defined as the end dates of all drawup periods that are contained in the sequence. Scanning over the 550 pairs (ε_0, w) gives us 550 sets of peak times $\{t_p\}$. Next, for each daily date t in the observed time frame, we count the number of times that this date was identified as a peak time over the whole set of 550 pairs (ε_0, w) . Dividing the resulting number by 550 gives us the fraction f_t of pairs (ε_0, w) that have the date t qualified as a drawup peak. We define large peaks, interpreted at the end of long bubbles, as those dates such that $f_t > 0.95$. We define intermediate sized peaks, interpreted at the end of short bubbles, as those dates such that $f_t > 0.65$. Intuitively, we therefore assume that large peaks are recognized in most of the trials while small to intermediate peaks are identified fewer times.

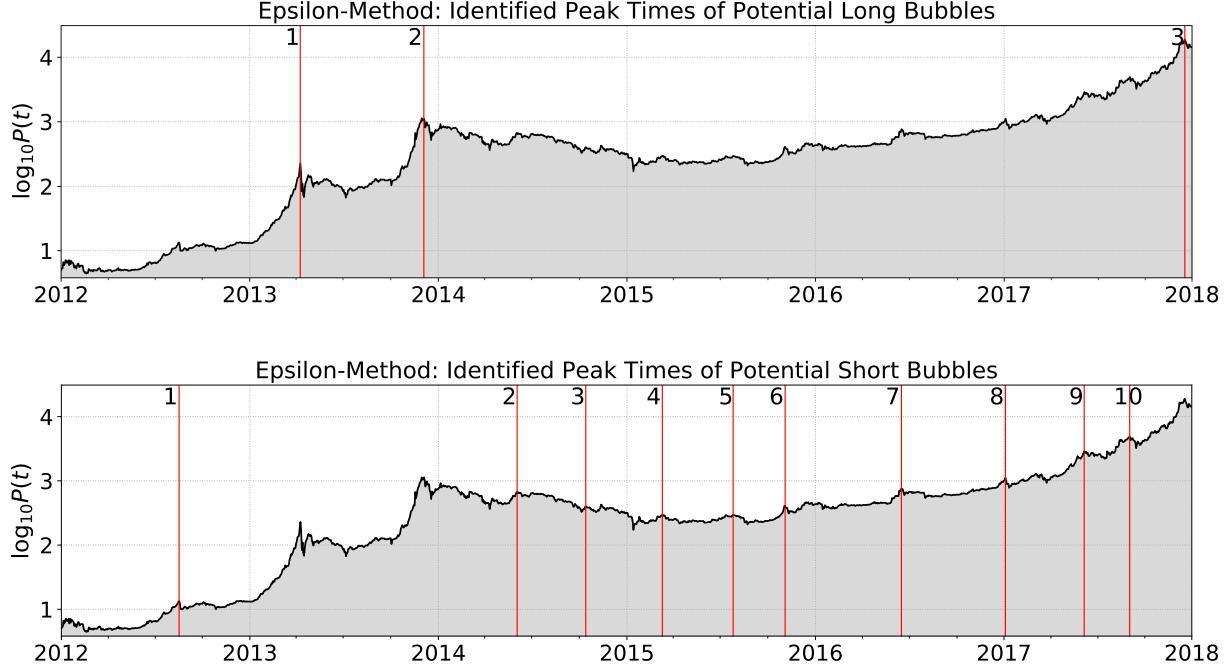


Figure 1. Bitcoin Bubble Peaks identified by Application of the Generalized *Epsilon Drawdown Procedure* between 2012 and 2018. Depicted in both panels is the logarithm of the Bitcoin to US Dollar exchange rate (black line, grey region) from Jan. 2012 to Jan. 2018. The vertical red lines indicate the bubble peaks that were identified according to the procedure outlined in the text. (Upper frame): the three vertical lines indicate the peaks obtained with the condition $f_t > 0.95$ (see text) and can be interpreted as the peak times of long bubbles. (Lower frame): the 10 vertical lines indicate 10 peaks, other than the three peaks of the upper frame, selected with the condition $f_t > 0.65$. They can be interpreted as the peaks of short bubbles.

Figure 1 shows the results of the application of our procedure for the identification of bubble peak times to the *btc/usd* log-price trajectory from Jan. 2012 to Jan. 2018. The upper frame shows three vertical lines identifying the three peaks with the condition $f_t > 0.95$. These three peak times are pleasantly coincident with the ends of the three largest bubble regimes of the Bitcoin price in US Dollars from Jan. 2012 to Feb. 2018. The lower frame shows 10 vertical lines identifying the 10 peaks, other than the three previous peaks, selected with the condition $f_t > 0.65$. These 10 peak times will be interpreted below as the ends of shorter bubble regimes. Note that these sets of peak times are robust with respect to reasonable changes of the thresholds 0.95 and 0.65.

B. Identification of the Beginning Times of Long and Short Bubbles

Following the above systematic determination of the main bubble peak times, we next need an automatic, unbiased determination of the beginning times of the corresponding bubble regimes. For this, we use the *Log-Periodic Power Law Singularity* (LPPLS) Model first introduced by Sornette et al. (1995) and elaborated upon by Johansen et al. (1999, 2000), which combines a description of the faster-than-exponential transient price acceleration with an increasing frequency of volatility fluctuations that are deemed characteristic of a bubble regime. A summary of how the LPPLS model derives from the Johansen-Ledoit-Sornette model

(Johansen et al., 1999, 2000) is given in Appendix B. The LPPLS formula reads

$$LPPL(\phi, t) := E[\ln p(t)] = A + (t_c - t)^m (B + C_1 \cos(\omega \ln(t_c - t)) + C_2 \sin(\omega \ln(t_c - t))) \quad (1)$$

The model includes seven parameters $\phi = \phi_1 \cup \phi_2$, of which four are linear $\phi_1 = \{A, B, C_1, C_2\}$ and three are nonlinear $\phi_2 = \{m, \omega, t_c\}$. The calibration methodology of the LPPLS model to a given price time series is explained in Appendix C. In a nutshell, we apply a two-step procedure. For a fixed value of the triplet (t_c, m, ω) , we solve analytically for the four linear parameters, which are solution of a linear matrix equation deriving from the Ordinary Least Squares formulation. We perform a grid search in the (t_c, m, ω) -domain, followed by a nonlinear minimization, with the conditions that $t_c - t_2$ lies between 0 and $t_2 - t_1$, $0 < m < 1$ and $1 < \omega < 50$.

The fit of *bt/usd* log-price trajectory by the LPPLS formula requires the choice of a specific time window $[t_1, t_2]$. The end time t_2 can be chosen at the bubble peak in post-mortem analyses. In order to diagnose and eventually forecast the future price evolution, in an ex-ante predictive framework, t_2 would represent the “present time”, up to which the price history is available for analysis. Choosing the beginning time t_1 (or equivalently the window size $dt := t_2 - t_1 + 1$) is less straightforward. Ideally, one would like to have t_1 as the beginning of the current bubble or later if a bubble is developing, such that non-bubble phases, i.e. phases of regular price growth, are excluded from the data to analyse.

To determine the beginning of a bubble, we implement the *Lagrange regularisation Approach* introduced by Demos and Sornette (2017b). For a fixed t_2 , we scan t_1 from $t_2 - 29$ to $t_2 - 719$, corresponding to 691 windows ranging from 30 to 720 trading days. In principle, taking into account all LPPLS fits over the 691 time windows, the beginning of the bubble is defined as the time t_1^* that gives the “best” fit quality. Naively, one could use the square root of the normalized sum of squared errors, also known as the root-mean-square (*rms*), and choose t_1^* as the value that minimizes the *rms*. However, this procedure is incorrect as it does not account for the fact that smaller windows will be favoured, due to their smaller number of degrees of freedom. Demos and Sornette (2017b) observed that applying a simple correction to account for this bias, which is linear in the window size $(t_2 - t_1)$, behaves quite well and was efficient in a number of tests and real-life case studies. This *Lagrange Regularisation Approach* is recalled in Appendix D. We apply the method to the *btc/usd* rate in the same time frame as before. We impose the constraint that, for a given developing bubble, its start time t_1^* cannot be earlier than the previous peak, as determined in Figure 1.

C. List and Main Characteristics of the Long and Short Bubbles

Combining subsection II.A which provides the set of bubble end times with subsection II.B for the determination of their beginning, we obtain a set of bubbles encoded by their time interval $[t_1^*, t_p]$, i.e. the time frame ranging from start to peak of the bubble. By definition, after the peak time of a bubble, a drawdown starts, which initiates the crash or correction regime. We define the end of this correction regime as the time at which the price reaches its minimum value over the time interval starting from the beginning of the drawdown up to the start time of the next identified bubble. For the last analyzed bubble, we simply take the price minimum between the bubble peak and the last available data point. The intermediate regime separating the previous crash end and the next bubble start, if there exists one, is then defined as a phase of normal price growth.

From the knowledge of start, peak and crash end dates of a bubble, as well as its price trajectory, we calculate the bubble size in percent, defined as the cumulative return between bubble start and peak for

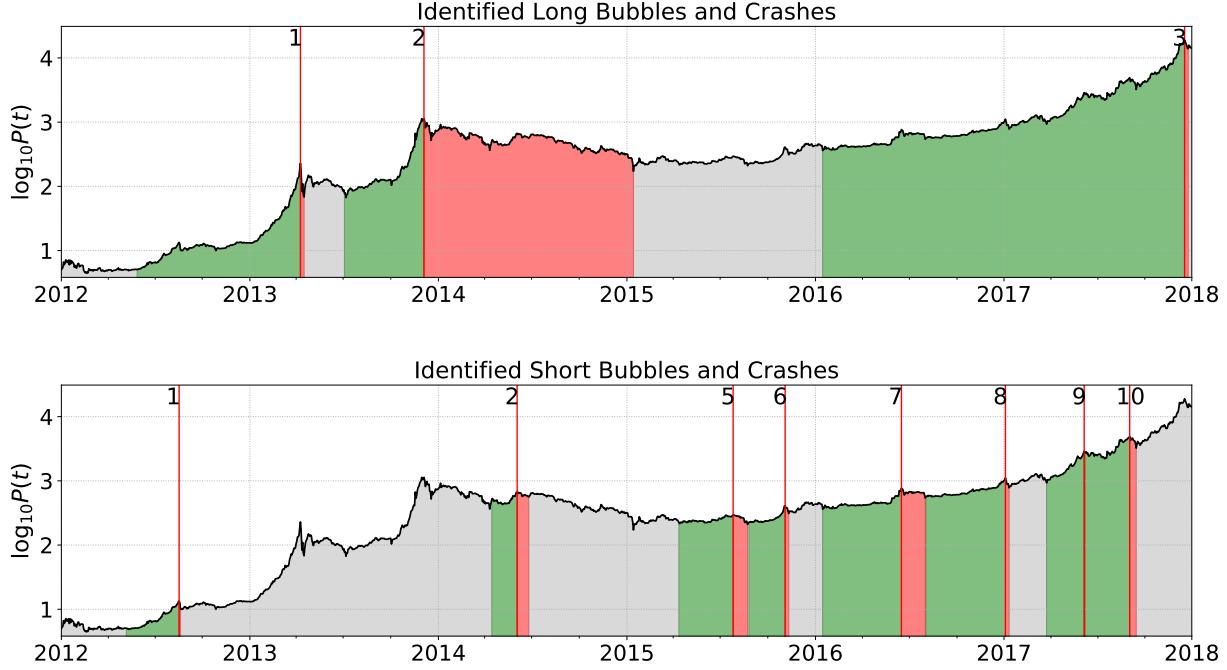


Figure 2. Bitcoin Bubbles from January 2012 to February 2018. The three resulting identified long (top panel) and eight short (bottom panel) bubbles, remaining after filtering the bubble characteristics. The green bands delineate the bubble growth phases, while the red shaded regions represent the associated crashes or correction regimes. The grey areas indicate phases of normal price growth, as defined in the text. The numbers to the left of the vertical red lines map to those given in Figure 1.

each bubble. The bubble duration is computed as the number of days from bubble start to peak. The crash size is calculated as the cumulative return between the peak time and crash end date.

Putting together the start, peak and crash end dates, the three peak times of the top frame of Figure 1 are thus the end time of three “long” bubbles, for which we now have the systematically determined beginning time and other characteristics, which are summarized in Table I.

For short bubbles to be determined among the 10 candidates associated with the peaks shown in the lower panel of figure 1, we want to avoid very short spells and thus impose two additional filters: a “short” bubble is such that it duration is at the minimum 30 days and its size larger than 25%, i.e. a short bubble has its price increase by at least 25% over at least 30 days. For the set of ten short bubble peak times shown in the bottom frame of Figure 1, we find that eight of them qualify as end times of short bubbles, whose characteristics are also summarized in Table I. Two of the ten peaks (Peaks 3 and 4) are excluded as not being preceded by a sufficiently large price increase (actually, they are local peaks within or just after a very low drawdown shown in Figure 2(a)).

Figure 2 shows the three identified long (top panel) and eight short (bottom panel) bubbles. At a first coarse-grained level of description, the top panel suggests that the Bitcoin history can be divided into three main regimes: (i) a pre-2014 phase of intense price growth, (ii) a 2014-2016 long drawdown and side-way regime, followed by (iii) the massive two-year long bubble that recently burst at the end of 2017. The bottom panel shows that the two large-scale bubbles from 2012 to end of 2013 and from 2016 to end of 2017 are themselves composed of shorter bubbles, some of them still with impressive amplitudes. We now turn to

describe them briefly from a socio-economic point of view.

Long Bubble Data							
Nr.	Bubble Start t_1^*	Bubble Peak t_{peak}	Crash End Date	Duration [days]	Bubble Size [%]	Crash Size [%]	
1	2012-05-28	2013-04-09	2013-04-16	316	4416	-70.27	NA
2	2013-07-03	2013-12-04	2015-01-14	154	1367	-84.83	NA
3	2016-01-15	2017-12-18	2017-12-25	703	5152	-26.55	NA
Short Bubble Data							
Nr.	Bubble Start t_1^*	Bubble Peak t_{peak}	Crash End Date	Duration [days]	Bubble Size [%]	Crash Size [%]	
1	2012-05-07	2012-08-16	2012-08-20	101	165	-25.11	Y
2	2014-04-15	2014-06-03	2014-06-25	49	28	-16.38	Y
3	2014-06-30	2014-10-14	2014-10-23	106	-37	-11.45	N
4	2014-10-23	2015-03-11	2015-04-13	139	-16	-24.75	N
5	2015-04-13	2015-07-27	2015-08-24	105	31	-28.74	Y
6	2015-08-26	2015-11-04	2015-11-11	70	80	-24.04	Y
7	2016-01-15	2016-06-16	2016-08-02	153	112	-29.56	Y
8	2016-08-03	2017-01-04	2017-01-11	154	97	-30.16	Y
9	2017-03-24	2017-06-06	2017-06-07	74	210	-6.86	Y
10	2017-06-07	2017-09-01	2017-09-14	86	83	-34.42	Y

Table I. Basic Information about all Qualified Bubbles. Dates and characteristics defining each bubble cycle. The last column indicates whether a potential bubble peak passed the bubble conditions defined in the main text.

III. Socioeconomic Drivers behind Bitcoin Bubble Dynamics

Having characterized the main bubble events, it is useful to put them into context and expose their key drivers, as well as the developments and events that promoted their nucleation or caused their sudden crashes. In particular, we analyze the possible causes of the nucleation and main drivers for the growth of the three major identified Bitcoin long bubbles that occurred between 2012 and 2018. The history of Bitcoin at its early stage was highly influenced by fiscal and monetary measures undertaken during the Eurozone crisis. Later on, the development of the cryptocurrency was to a great extent influenced by increasing demand from China, which evolved as a result of the emergence of major Chinese Bitcoin exchanges opening up the markets for investors. Although it had a strong temporary impact on Bitcoin in the short term, the shutdown of Chinese exchanges that occurred in early 2017 caused no persistent loss in its capitalization. Quite the contrary, the year of 2017 is characterized by remarkable growth dynamics due to the contagion of the Bitcoin bubble to a general cryptocurrency bubble. Finally, at the end of 2017, the most recent long bubble crashed with a shocking intensity, removing 60% (up to Feb. 2018) of the cryptocurrency market capitalization measured from its peak in Dec. 2018. The following sections examine this history in more details.

A. First Long Bubble: May 2012 - April 2013

The financial crisis of 2007-2008 put economies worldwide into a fragile state. It also helped reveal the unsustainable debt level and the longstanding financial mismanagement of some small and medium sized economies within the European Union such as Greece, Ireland, Portugal, Spain, Iceland and Cyprus. In particular, the crises in Cyprus and Greece stand out by their intensity and consequences. Notice that both countries reached their worst economic state for decades in 2012, as demonstrated in the bottom panel of

Figure 3. The economic decline was accompanied by ongoing discussions about further bailout programs for Greece, as well as requests of the Cypriot government for a bailout of its financial sector which the country's economic well-being largely relied on (Wilson et al., 2012).

The emerging mentality of distrust in governments and financial institutions triggered a wave of bank runs and hunts for monetary safe havens. Bitcoin, proposed as an alternative store of value that was primarily intended to be uncontrollable by governments and independent of monetary policies (Nakamoto, 2008), appeared to ideally meet these requirements. It is no coincidence that the nucleation of the first Bitcoin long bubble occurred at the exact time when the Greece and Cypriot indices reached their troughs.

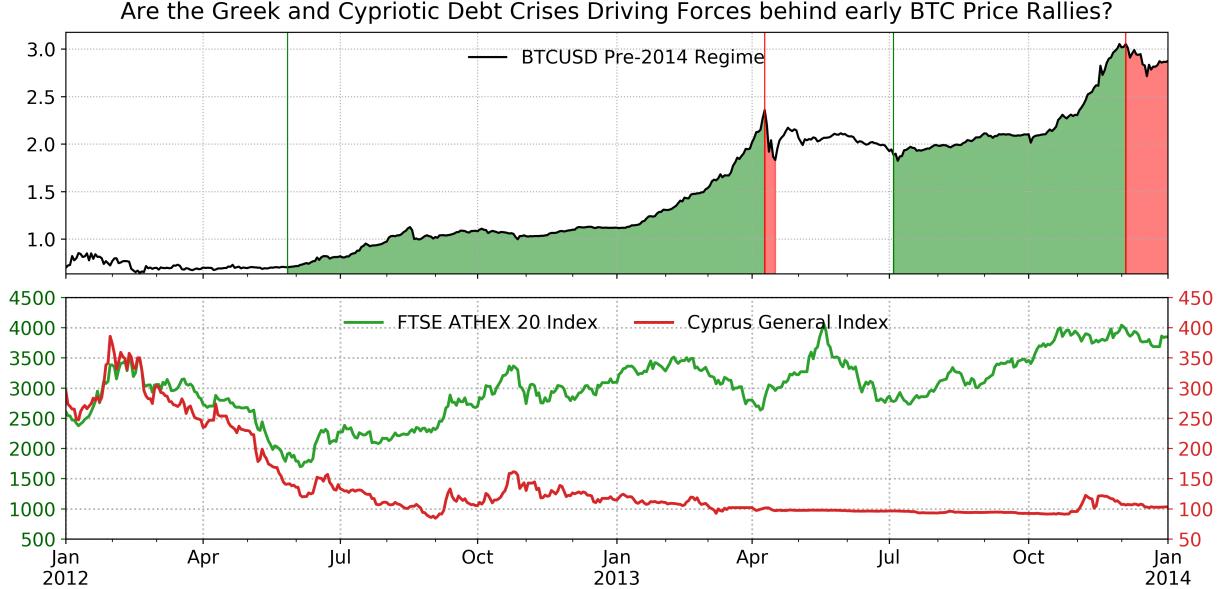


Figure 3. Comparison of the Evolution of Bitcoin Price and Greece as well as Cypriot Financial Market Indices. The top panel shows the Bitcoin log-price trajectory from Jan. 2012 to Jan. 2014. In the bottom panel, the Greek ATHEX 20 Index (left scale), as well as the Cyprus General Index (right scale) are depicted over the same period. The financial market indices of Spain, Portugal and Ireland follow paths highly correlated with that of the Greece index, as their indices bottomed out during approximately the same time. [Data Source: Datastream (2018)]

On March 16, 2013, a critical point was reached for the Cypriot economy, when a bail-in tax that largely affected the deposits of account holders at Cypriot banks was declared. The announcement created a massive wave of bank runs triggered by account holders trying to protect their personal savings (Thompson, 2013). This coincided with the last accelerating growth of Bitcoin's price, before its crash in April 2013 (Farrell, 2013).

The described developments in Cyprus were anxiously observed in other Eurozone countries, for instance in Spain, where people feared that similar governmental interventions could lead to the loss of their own savings (Warner, 2013). Serving as a store of value that could not be seized by any institution, Bitcoin arrived on the scene, perceived as the perfect alternative investment. Figure 4 illustrates the consequential flow towards the cryptocurrency in form of soaring Blockchain transactions around the time of the bubble nucleation of Long Bubble 1. As a response to the increasing demand, in the course of bubble growth, the price of a Bitcoin was catapulted up by an incredible 4400% above the price mark of 100 USD per Bitcoin.

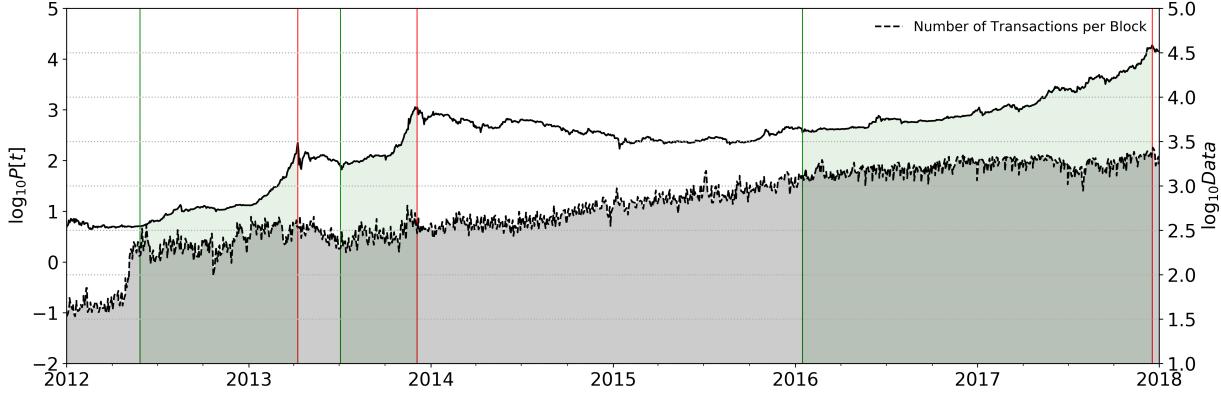


Figure 4. Evolution of the Number of Transactions on the Blockchain Network. The grey shaded region delineates the logarithmic number of Bitcoin Blockchain transactions (right scale). In 2012, a jump of the transaction number occurred in response to the large demand surge by investors seeking Bitcoin as a safe haven. [Data Source: [Blockchain.info \(2018\)](#)]

The bubble reached maturity in early April 2013. Specifically, after peaking on April 9th, it burst, with approximately 70% of Bitcoin's market capitalization disappearing within one week. The actual cause of the crash was a reaction to instability announcements on *MtGox*, the by then worldwide biggest Bitcoin Exchange in terms of Bitcoin exchange-traded volume, which was suffering the consequences of a hacker attack ([Buterin, 2013](#)).

Obviously, the event leading to the crash was not directly related to the political backgrounds of the Eurozone and Cypriot Crises that actually initiated the bubble. As is often the case in the bursting of a bubble, one should distinguish the proximate cause of the crash, which is in general unrelated to the source of the bubble, from the fundamental origin of the crash, which is that the Bitcoin market had progressively evolved towards a fragile, unstable, critical state, associated with a large susceptibility to adverse news.

B. Second Long Bubble: July 2013 - December 2013

The second long bubble matured at the end of 2013 after an approximate thirteen-fold increase from its post-crash recovery level after the first long bubble. As the European debt crisis was still well underway during this time, again, the attraction to the alternative decentralized Bitcoin contributed to its price surge. But, one should not underestimate a number of additional factors that contributed to the emergence of this bubble: the adoption of Bitcoin in China, the FBI shutdown of the Darknet drug market *Silk Road*, as well as the growth and increasing technical sophistication of the Bitcoin mining community.

Figure 5 presents the history of the birth and trading volumes of the main Bitcoin exchanges from 2012 onward. The blue inset plots show the births and absolute volumes handled by the different exchanges. The light grey area shows the logarithmic summed transaction volume of all of these exchanges that are still active today, while the dark area sums the volume of all listed exchanges. It can be seen that major Chinese exchanges were founded in the nucleation phase of the second long bubble, amongst them *Huobi* and *OKCoin*, which later on became the dominating Bitcoin exchanges in China ([Bovaird, 2016](#)).

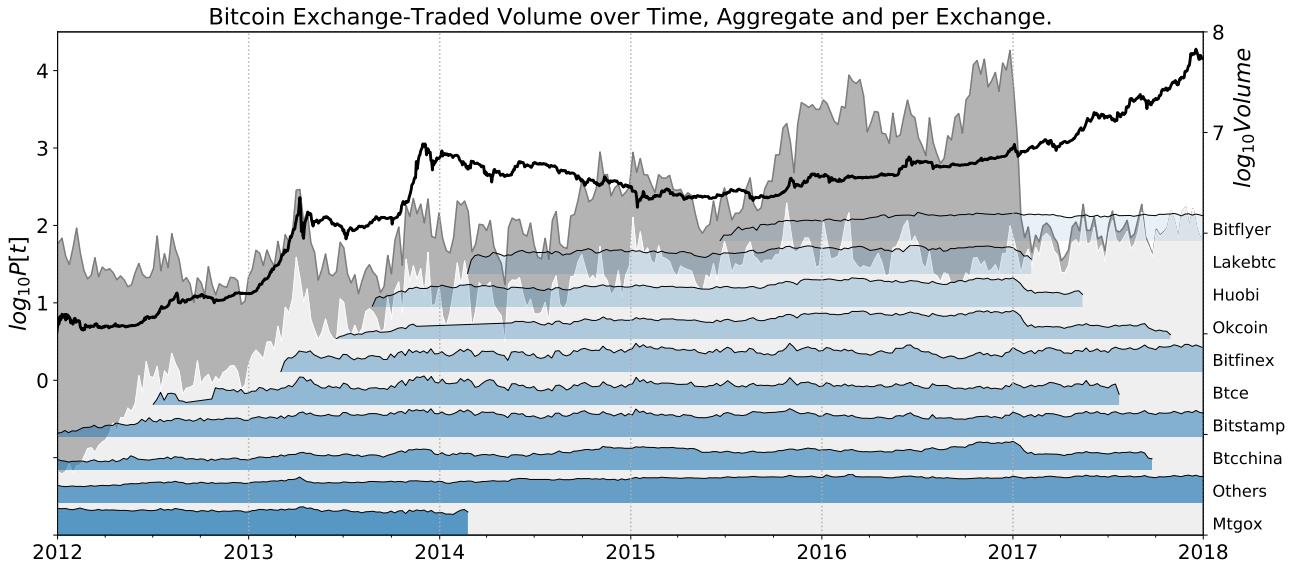


Figure 5. The Evolution of Bitcoin Exchanges in Terms of Trading Volume. The dark grey area (right scale) represents the total log-volume of all analyzed exchanges. The light area shows the volume of only the exchanges that are still in business as of end 2018. The blue inset plots schematically show the evolution of traded log-volume of the single exchanges that are listed next to the right plot axis. Note that the inset plots are ordered according to the dates of entry into service of the different exchanges. Hence, from these plots, the temporal origination of exchanges can be compared. [Data Source: [Bitcoinity.org \(2018\)](#)]

The emergence of new Bitcoin exchanges significantly facilitated market entrance for the numerous primarily China-based investors to the Bitcoin market. Ultimately, the wave of new investments drove prices above the 1000 USD level for the first time. Figure 6 shows that the share in total traded volume of Chinese exchanges started to really break through during the second long bubble, in the phase of most intense growth. Before that, *MtGox* held the majority of the market share. However, its market share plummeted from more than 50% at the beginning of the bubble down to a mere 10% at its peak. The corresponding lost fraction had been replaced by trading volume on the uprising Chinese exchanges.

As the bubble was developing at fast pace, on October 2nd, 2013, the online drug market *Silk Road*, which enabled customers to buy illegal substances anonymously, was shut down by the FBI and its alleged founder imprisoned ([Pagliery, 2013](#)). This signaled to the Bitcoin community that the legal authorities had their eye on the cryptocurrency and intended to prevent any illegal activities related to it by all means. Although *Silk Road* was the first, but not the only darknet drug market, its seizure had wide implications and hit the news headlines heavily. The closure of *Silk Road* symbolically set free Bitcoin as a proper investment for more cautious investors who, until then, were deterred by its illegal usage as drug money. The subsequent build-up of a clean image for Bitcoin was also recognized by the US senate, which as a consequence announced an official hearing about the utility and future prospects of the digital currency ([US Senate, 2015](#)). The event is seen as a further beneficial factor contributing to the immense price surge seen in the second long bubble ([Wile, 2013](#)).

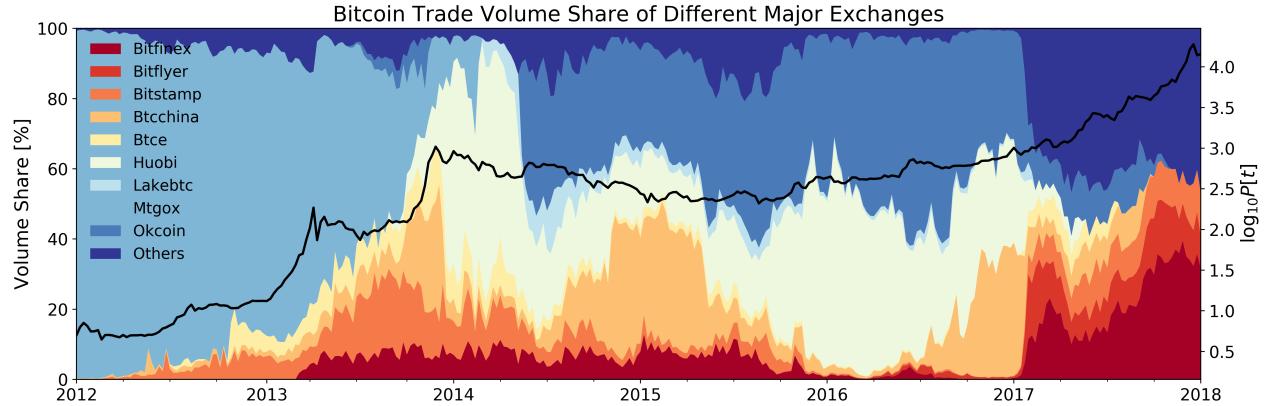


Figure 6. Volume Share of the Top Ten Bitcoin Exchanges. Shown on the left scale is the percentage share of the Bitcoin volume traded on each exchange with respect to total traded volume. Notice the surge of the two main Chinese exchanges, *OKcoin* and *Huobi* towards the end of the second long bubble, as well as the constant decrease of the market share of until then dominating Japan-based exchange *MtGox*. [Data Source: [Bitcoinity.org \(2018\)](#)]

Figure 7 hints at the third contributing factor for the price growth during the bubble. The logarithmic hash rate power of the miners computing transaction blocks on the Blockchain is shown, as well as the logarithmic number of registered wallets. Notice that the hash rate grew at a much faster rate compared to the number of wallets over the period starting with the first long bubble. The highest rate of increase was reached during the second half of 2013, coinciding with the second long bubble.

The number of registered wallets is a measure for the size of the network of users operating directly on the Blockchain. Hence, the hash rate increasing faster than the user base signals that on average miners enhanced their individual mining power during the period. This was most likely achieved through the usage of more efficient technical mining hardware. We conclude that, in addition to Bitcoin adoption effects in China and the image improvement of Bitcoin, we can see the effect of the ramp-up behavior of hash power as an indication of increased mining sophistication.

Concerning the improvement in hardware and mining technology, we note that the first long bubble played an important role as a precursor to the second one. As the Bitcoin price increased for the first time to a fairly high level, miners were incentivized to invest in mining hardware. At the relatively high price of a Bitcoin compared to the computational effort required to create it, it was profitable for them to enter the mining business. The increasing number of miners itself then triggered a feedback mechanism. It can be seen as a self-fulfilling, self-reinforcing bubble in the following sense: the larger the price, the larger the incentive to invest in hardware computing power; the more mining there is, the more activity there is on the blockchain that attracts more and more users and buyers, the more the price increases. But, the faster growing variable should quickly become the hashing difficulty. The loop is closed when the incentive of miners to invest in hardware rises again.

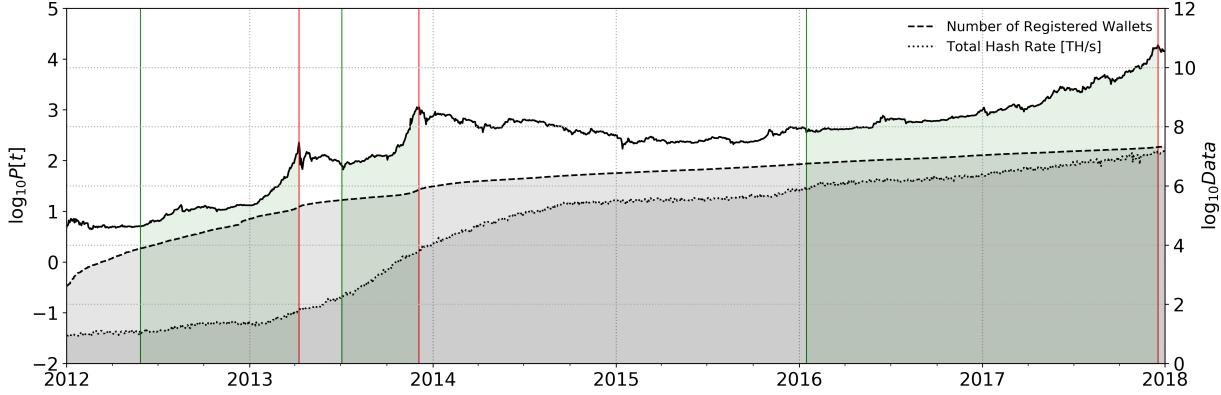


Figure 7. Growth of Bitcoin Blockchain Network User Base and Increase in Mining Power. Superimposed to the Bitcoin price in log-scale (left axis), the total hash rate and the number of registered wallets are shown (right axis). The largest rate of increase of the total hash rate can be observed over the course of the second long term bubble in 2013, a fact signalling an increasing level of miners' technique sophistication. A second, more subtle acceleration of the hash rate can be seen around the nucleation of the third long bubble. [Data Source: [Blockchain.info \(2018\)](#)]

There are two major triggers for the crash that started in Dec. 2013 and was followed by a long drawdown. Firstly, the Chinese Government suddenly prohibited financial institutions from using Bitcoin when the price of a Bitcoin was close to the peak of the bubble in December 2013 ([Hill, 2013](#)). The announcement destabilized the currency and sent it into free fall. Secondly, during February 2014, when there was still hope for the price to recover, *MtGox* suspended trading (see Fig.6) and filed for bankruptcy protection from creditors after a major account robbery of at least 700,000 missing or stolen Bitcoins of customers had been reported ([Zuegel, 2014](#)).

The demise of *MtGox* was perceived as a major setback for the Bitcoin world. Account holders of the exchange lost money that was not or hardly recoverable due to the difficulty of tracing it. Once more, this led to a huge loss of Bitcoin's trustworthiness, which already had been questioned over and over again in the past. These adverse circumstances set the currency off into a one-year long decline.

C. Third Long Bubble: January 2016 - December 2017

The winners of the long Bitcoin price drawdown that took place from 2014 until mid-2015 were clearly the China-based Bitcoin exchanges. At the time of the closure of *MtGox*, they had already occupied the majority (more than 90%) of all exchange-traded volume, as can be seen by looking back at Figure 6. Within the three years following 2014, Bitcoin volume formation on Chinese exchanges contributed to roughly an hundredfold increase in global volume (see Fig.5). It is this rising demand for Bitcoin from Chinese markets originating during that time, which can be seen as a major precipitating factor for the nucleation of the third long term bubble, whose start we identify around the beginning of 2016. This leaves open the question where this increased demand originated from? We identify the devaluation of the Chinese Yuan as a main promoting factor for the rising interest in cryptocurrencies in China and thereby the formation of the third long bubble.

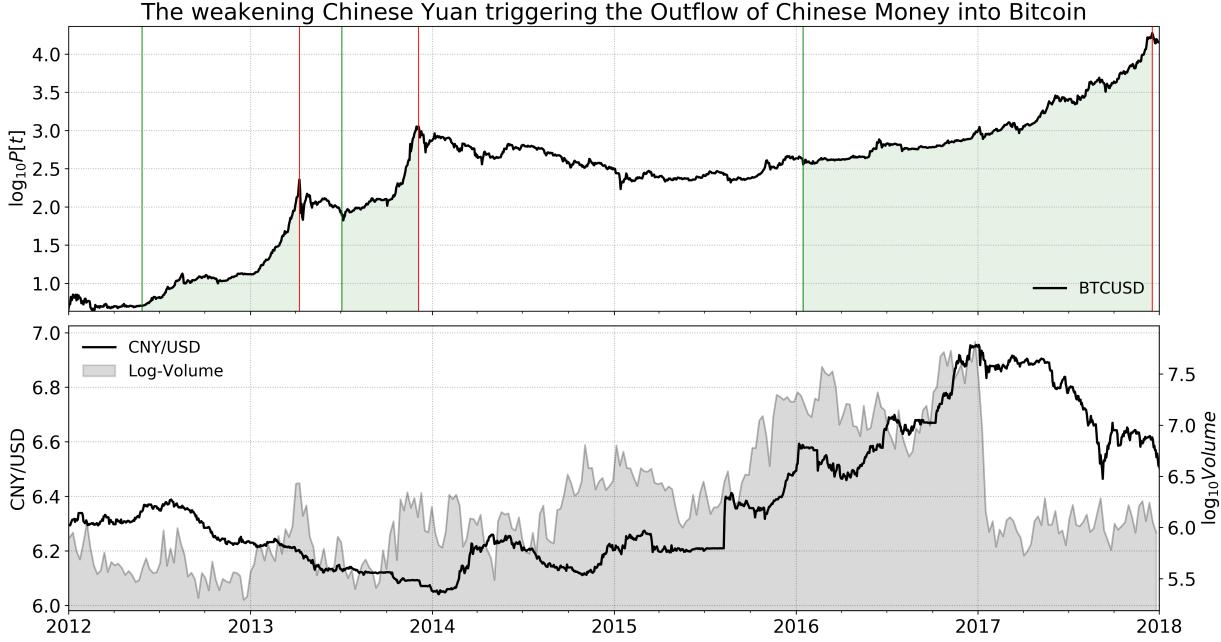


Figure 8. Development of Bitcoin in Parallel to the Chinese Yuan. The Bitcoin log-price (top panel) and the Chinese Yuan vs. US Dollar (CNY/USD) exchange rate (bottom panel, left axis). A change of regime occurred in the Chinese currency simultaneously to the decline of Chinese exchange-traded Bitcoin volume (bottom panel, right axis) at the beginning of 2017. [Data Source: [Datastream \(2018\)](#)]

Figure 8 shows a change of regime in the exchange rate of the Chinese Yuan (CNY) versus the US Dollar from 2014 onward. In August 2015, the depreciation of the Yuan was enforced through a devaluation of the currency by the People's Bank of China, *PBoC*, as seen by the sudden jump in the rate. This devaluation was motivated by the desire to raise the competitiveness of exporting firms. From there on, a continuous weakening of the Renminbi developed until January 2017.

As a reaction to the depreciation of their currency from 2014 onward, Chinese market participants tried to transfer their money to what they perceived as safer stores of value, causing an outflow of capital from China ([Gautham, 2016](#)). As for the average Chinese investor, limitations in terms of foreign-exchange investments were quite restrictive, Bitcoin was a straightforward solution to store value ([Suberg, 2016](#)). As a response to the devaluation of the Renminbi, one can observe indeed a large increase in demand for Bitcoin in form of rising trading volume and growing prices from mid 2015 onwards. Note specifically the period of devaluation of the Chinese Yuan (during the last quarter of 2015) preceding the start of the third long bubble.

In January 2017, the Central Bank of China instructed the - until then widely unregulated - Chinese Bitcoin exchanges to comply with the country's financial regulations ([Rizzo, 2017](#)), as it suspected illicit exchange activities such as money laundering, as well as volume manipulation (through wash trading) that was made possible by the zero trading fee policy of exchanges ([Olszewicz, 2017](#)). In a first intervention, the biggest Chinese Bitcoin exchanges *BTCC*, *Huobi* and *OKCoin* ([Higgins, 2017](#)) were forced to reintroduce a non-zero trading fee structure and to stop leveraged trading ([Bovaird, 2017](#)). The measure led to the huge slump in exchange trading volume that can be observed in Figure 5. Simultaneously, the Short Bubble 8 (see Figure 2 and Table I) burst in a 30% crash.

Chinese exchanges started with their zero trading fee policy in late 2013. It has been estimated that, due

to the ongoing manipulations of trading volume, the volume reported by exchanges had been overstated by up to forty times the true volume ([Woo, 2017](#)). Hence, the crash in trading volume partly seems to represent a normalization to realistic levels ([Wong, 2017](#)), as with the introduction of trading fees, volume could not be generated at no cost, any longer.

In a further regulatory move, in February 2017, the *PBoC* again exerted pressure on Chinese exchanges causing them to halt Bitcoin withdrawals ([Goh, 2017](#)), while still tolerating withdrawals executed in the domestic currency Yuan ([Solana, 2017](#)). Effectively, this measure was intended to cut off the outflow of Chinese money through Bitcoin.

In June 2017, the temporary withdrawal pause was ended, as exchanges had partly adapted to the regulatory standards demanded by the *PBoC* ([Arnold and Chen, 2017](#)). These news implied an overall positive outlook for Bitcoin's future in China, promoting further rise in the price of the digital coin. In September 2017, unexpectedly, Chinese regulators banned the so-called Initial Coin Offerings (ICOs) ([Chen and Lee, 2017](#)), a novel procedure for the emission of new digital coins, emulating the Initial Public Offerings of regular firms. Finally, in mid September, the Central Bank ordered Chinese exchanges to shut down all trading activities on the Chinese market ([Huang, 2017](#)). This quick series of unanticipated events officially put an end to cryptocurrency exchange business in China. The gradual decrease in trading volume and closure of exchanges during 2017 is shown again in Figure 5.

Although in late 2017, cryptocurrency trading was proclaimed "Officially Dead In China" ([Rapoza, 2017](#)), the actions undertaken by China's Central Bank led to (i) a shift of exchange-based trading to OTC trading, running on so-called peer-to-peer exchanges ([Barclay, 2017](#)), the most known of which was *LocalBitcoins*, and (ii) Chinese exchanges such as *Huobi* and *OKCoin* moving their business abroad ([Zhao, 2017](#)). In this way, although *PBoC* managed to prevent Yuan-based trading, the still present large if not increased demand was eventually just redirected to other markets ([Redman, 2017](#)). Thus, the imposed regulatory constraints did not have any permanent influence on the global evolution of the price of Bitcoin. This is also backed by the fact that, throughout 2017, the Bitcoin price recovered fast from drawdowns resulting from China-related negative news.

Besides the large contribution of Chinese-based Bitcoin investing to the formation of the third long bubble, there were other factors driving the growth of the bubble. The third long bubble was punctuated by a number of short bubbles that burst abruptly, but were followed by a fast recovery of the Bitcoin price, as seen in Figure 2. There are several reasons for this behavior.

Going back to the nucleation of the third bubble in early 2016, as in the case of the second long bubble, one witnesses an increase in the growth rate of the hash rate, with the total hash power then exceeding 1 ExaHash/second ($= 10^{18} H/s$!). Moreover, during 2016, several of the nowadays dominating BTC mining pools became active (*viabtc*, *btc.com*, *btc.top*) ([ViaBTC, 2018](#); [BTC.com, 2018](#); [BTC.TOP, 2018](#)). This development produced additional publicity for Bitcoin and pushed mining efficiency for individuals once more. The fact that mining gains of pools are proportionally shared amongst miners according to their power contribution provided a large incentive especially for weak miners to resume their mining activity, as their chance of being rewarded for mining (within their lifetime) suddenly was ensured. The growth of the mining network and associated hash power is a factor that may well have contributed to higher resilience and fast recovery of the Bitcoin price during the massive, third long bubble.

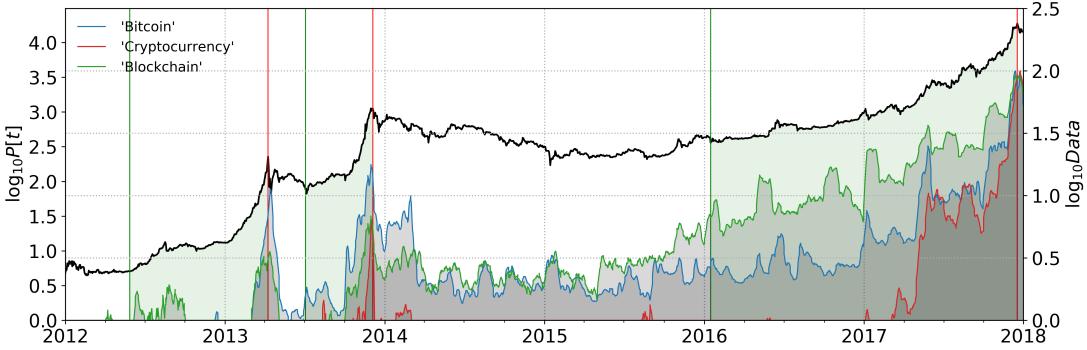


Figure 9. Google Trends Search Queries related to Bitcoin. The three grey shaded areas depict the development of search requests for the terms *Bitcoin*, *Cryptocurrency* and *Blockchain*. For studies demonstrating the relationship between such search queries and the Bitcoin price trajectory, we refer the reader to (Kristoufek, 2013; Garcia et al., 2014). [Data Source: [Google Trends \(2018\)](#)]

The most distinct feature driving the acceleration of the Bitcoin price after the China exchange shutdown was the growth of the cryptocurrency market as a whole. Figure 9 demonstrates that, besides skyrocketing Google search queries for the term *Bitcoin*, from 2017 onward, a sharp increase in search queries for the term *Cryptocurrency* occurred. Already from 2015 onward, rising requests for the term *Blockchain* can be seen. This signals interest in Blockchain technology on its own that has been growing over the years, as well as strongly enhanced interest in alternatives to Bitcoin from 2017 onward.

As a response to investor's growing demand for alternative investments in the cryptocurrency market, the emergence of a multitude of new digital coins ensued. Throughout 2017, the cryptocurrency market changed its structure from being dominated by Bitcoin to a more diversified market offering numerous technologies and variants of cryptocurrencies (Wu et al., 2018).

Figure 10 shows the dominance of Bitcoin at the beginning of 2017 over the complete cryptocurrency market with its market share, measured with respect to the total market capitalization of the top 1000 cryptocurrencies, being as high as 90%. Within three months after February 2017, its relative market share dropped by 50%, while its total capitalization value and that of other crypto-currencies continued to grow at accelerating pace until Dec. 2017.

During the last two months of 2017, the overall capitalization of the crypto-market multiplied by a factor four. As the idiom claims, “the rising tide lifts all boats”, and thus, the large inflow of fresh money to the cryptocurrency market impacted most of the Altcoins as well as the still dominating Bitcoin. However, with the crash following Dec. 18th, 2017, the value of Bitcoin and that of many cryptocurrencies has been dropping, with Bitcoin losing sixty percent of its total value (as of February 2018), putting the Bitcoin market share to an all-time low.

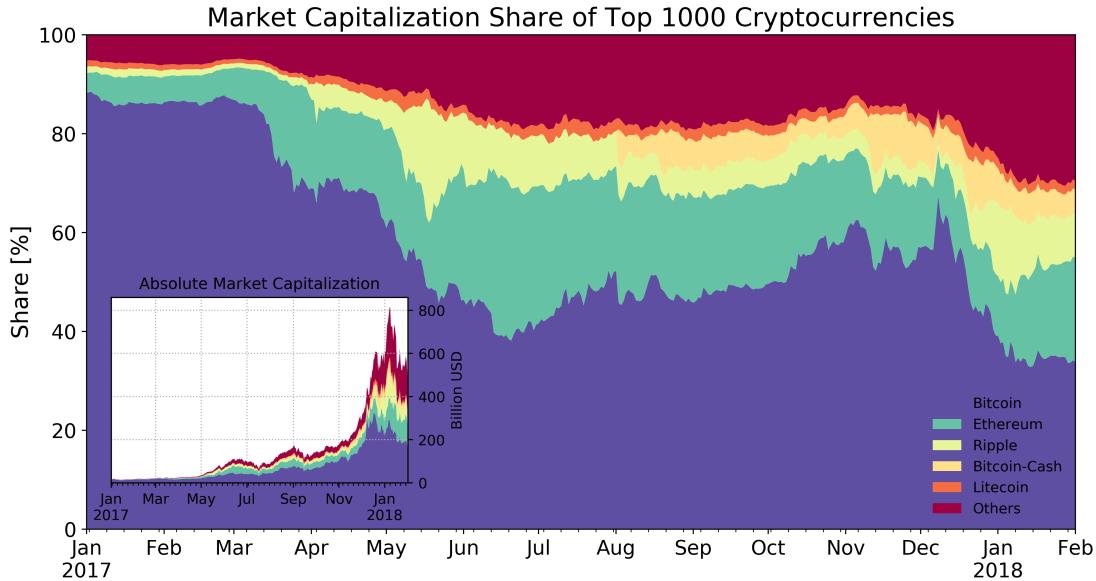


Figure 10. Progressive Maturation of the Cryptocurrency Market during the Year of 2017. Shown is the market share of the five largest capitalized cryptocurrencies *Bitcoin*, *Ethereum*, *Ripple*, *Bitcoin Cash* and *Litecoin*, as well as the aggregate capitalization of the top thousand cryptocurrencies excluding the five largest ones in the year 2017 (shown as “Others”). Similarly, the inset plot depicts the evolution of the absolute market capitalization. We can see that, within one year, Bitcoin lost a majority of its share in the total market capitalization. This signals its possibly decreasing significance and, at the same time, the growing number of competitive cryptocurrencies in the market. [Data Source: [CoinMarketCap \(2018\)](#)]

One can say that 2017 was the year of the cryptocurrencies. New competitors were constantly pushing their coins as well as new tokens into the market. These coins and tokens were associated with new proposed ideas of the application of Blockchain technology. As the market is becoming ever more complex every day, the possibility for the release of a new coin making Bitcoin obsolete is perceived more likely. In contrast, others note that Bitcoin may survive as a kind of relatively scarce digital “commodity” or asset, used by investors as a storage of value and diversification vehicle.

We now turn to a more quantitative analysis of the various bubbles from 2012 to 2017 to recount how well these bubbles (could) have been diagnosed in real time with metrics derived from the *Log-Periodic Power Law Singularity Model* presented above in section II.B.

IV. Real-Time Bubble Identification

A. LPPLS Multiscale Confidence Indicators as Bubble Diagnostics

Following the methodology of [Sornette et al. \(2015\)](#) and [Zhang et al. \(2016\)](#), we use the *LPPLS Confidence Indicator* as a diagnostic tool for the recognition of bubbles. In a nutshell, the LPPLS Confidence Indicator at a given time t_2 is the fraction of windows $[t_1, t_2]$, obtained by scanning $t_2 - t_1$ over a certain window size interval, for which the calibration of the log-price of Bitcoin by the LPPLS formula (1) passes the criteria shown in Table II. A large LPPLS Confidence Indicator value indicates that the LPPLS patterns with model parameters passing the filtering conditions are found in a large fraction of time scales at this particular

time t_2 . This is then translated into a bubble diagnostic analysis, based on the hypothesis that the LPPLS pattern is a characteristic feature of bubbles.

Parameter	Filter Bounds	Description	Pre-Condition
Power Law Amplitude B	$(-\infty, 0)$	-	-
Power Law Exponent m	$(0, 1)$	-	-
Log-periodic Frequency ω	$[4, 25]$	-	-
Critical Time t_c	$(0, dt_i)$	-	-
Damping D	$[0.5, \infty)$	$D = \frac{m B }{\omega C }$	-
Number of Oscillations O	$[2.5, \infty)$	$O = \frac{\omega}{2\pi} \ln \frac{t_c - t_1}{t_c - t_2}$	if $ C/B \geq 0.05$

Table II. Filter Conditions for Qualified LPPLS Fits. The table gives the parameter bounds that were used to filter for qualified LPPL fits. The constraint on B ensures the existence of a positive bubble. For m and t_c , the boundaries are excluded to avoid singular behaviors in the search algorithm. The damping factor quantifies the allowed downwards movements of bubble fits. The O -parameter measures the number of oscillations that occur within the fit window $[t_1, t_2]$. The filter on O is applied only when the amplitude of the oscillations, as quantified by C , is sufficiently large relative to the power law amplitude B .

In order to capture the main relevant time scales of investment horizons, we propose to partition the window sizes $dt := t_2 - t_1 + 1$ over which the LPPLS Confidence Indicator is calculated into three classes: $dt \in [30, 120]$ (short time scale), $dt \in [100, 240]$ (medium time scale) and $dt \in [200, 720]$ (long time scale). The short time scale goes approximately from one month to five months. The medium time scale ranges from 6 months to a year (when there are approximately 250 trading days in a calendar year). The large time scale goes from 10 months to approximately 3 years. Thus, for each t_2 , we construct three LPPLS Confidence Indicators, one for each set of window sizes. There are respectively 91, 141 and 521 differently sized windows for the short, intermediate and long time scales. We refer to the ensemble of the three indicators calculated in this way as *LPPL Multiscale Confidence Indicators*.

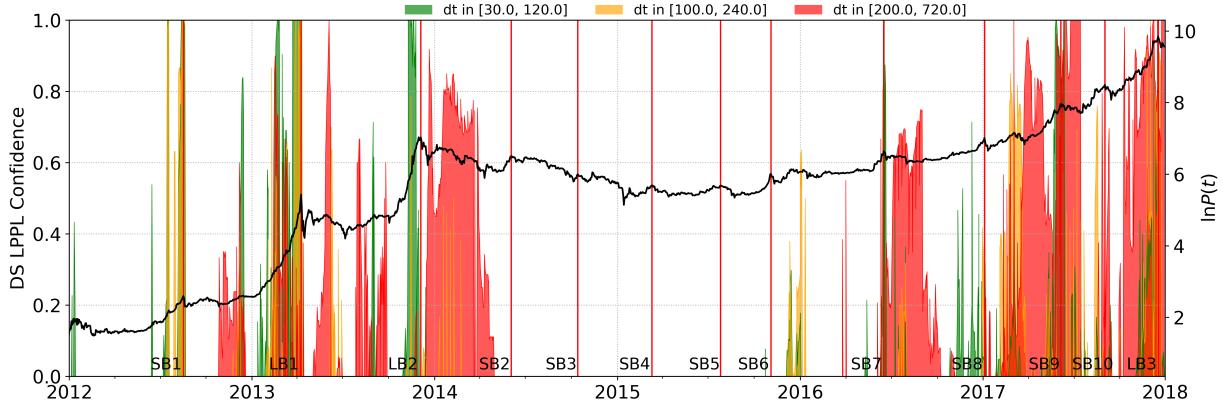


Figure 11. Diagnosing Bubbles over Multiple Scales using the LPPLS Confidence Indicator. The LPPLS Confidence Indicator diagnoses bubbles nucleating over short (green), intermediate (yellow) and long time scales (red). The larger the value of the metric at a given point in time, the larger the portion of qualified LPPLS fits, indicating stronger bubble activity. The peaks of the bubbles, as documented in Figure 2 and Table I, are depicted in red vertical lines.

Figure 11 depicts the LPPLS Multiscale Confidence Indicators, as described above, constructed on the Bitcoin/USD price. One can observe that the short and intermediate indicators peak at or close to the

corresponding bubble tops, providing useful diagnostics, in particular for the first three pre-2014 bubbles. For the shortest bubble episodes such as Short Bubbles 5 and 6, there are no signals. This can be understood from the choice of time scale: even the short time scale $dt \in [30, 120]$ is too long, compared with the duration of these two bubbles. This shows the importance of having a suitable set of indicators scanning the time scales of interest. Although not all of the minor short bubbles are identified due to the mentioned limitation of time scales, in all cases for which the LPPLS Confidence Indicator reaches its maximum value 1, a crash follows.

After 2016, the LPPLS Confidence Indicator at the long time scale exhibits an increasing amplitude, associated with the development of the third long bubble. Close to the peak in December 2017, the large-scale LPPLS Confidence Indicator is close to its maximum level, providing a warning for an imminent massive crash, following bubble price dynamics that evolved over the long term. Since the end of 2016, we have been implementing this and other indicators to watch the development of asset price bubbles, amongst them the many bubbles in the Bitcoin price: for instance, in our monthly report of the *Financial Crisis Observatory* on the 1st December 2017, we pointed out the alarm generated in advance to the burst of the third long bubble (see 1st December 2017 Synthesis report at www.er.ethz.ch/financial-crisis-observatory.html).

Overall, the LPPLS Multiscale Confidence Indicators over the three monitored scales provide a valuable risk management metric for early recognition of emerging bubbles at short and long time scales as well as their approaching bursts. In particular, the long bubbles are clearly identified in advance by the fact that the three LPPLS Confidence Indicators peaked jointly for these long bubbles. The other non-synchronized indicator peaks testify of the multiscale nature of bubbles in Bitcoin.

B. Predictions of Bubble Burst Times Using Ensemble Forecasting

B.1. Methodology

We now present systematic mock-up forecasts of the burst times of the three long and eight short bubbles identified in Figure 2 and Table I. For each bubble, we position the mock-up “present” time of analysis t_2 ten business days prior to the corresponding identified bubble peak date given in Table I. Next, for each analysis time t_2 , we identify the associated bubble start time t_1^* by means of the Lagrange Regularisation Methodology, as explained in section II.A.

Based on the complete “frame” of each bubble $[t_1^*, t_2]$, we perform LPPLS calibrations over all windows by scanning t_1 from t_1^* to $t_2 - 29$. Out of these $t_2 - t_1^* - 28$ LPPLS fits, we again only keep those that pass the filter conditions listed in Table II. For a given t_2 , each of the remaining qualified fits is then characterized by (i) its window size $dt := t_2 - t_1 + 1$ corresponding to the observed time scale and (ii) a forecasted horizon $t_c - t_2$, which is the estimated remaining time from the present time t_2 up to the predicted end t_c of the bubble.

Because analyses performed over different window sizes may result in different forecasts, corresponding to multiple price dynamics, we use the k-mean clustering method to identify clusters in the $(dt, t_c - t_2)$ space whose components yield common predictions for the critical time t_c . Then, if a unique, well-defined cluster is identified, this can be interpreted as a consensus forecast obtained over all the time windows included in the cluster. If there are two well-defined clusters, similarly, one should conclude that the analysis suggests two different scenarios for the future development of the bubble, and so on. We augment the k-mean clustering method by using the Silhouette metric to determine the optimal number \hat{k} of clusters and thus of forecasted scenarios. The details of the k-mean clustering implementation and application of the Silhouette metric are

given in Appendix E.

Long Bubble Data				
Nr.	Analysis Date t_2	Qualified Fits [%]	Clusters \hat{k}	Silhouette $\bar{s}_{\hat{k}}$
1	2013-03-28	43	2	0.85
2	2013-11-22	40	2	0.84
3	2017-12-06	68	9	0.67
Short Bubble Data				
Nr.	Analysis Date t_2	Qualified Fits [%]	Clusters \hat{k}	Silhouette $\bar{s}_{\hat{k}}$
1	2012-08-06	59	2	0.58
2	2014-05-22	0	-	-
3	2014-10-02	0	-	-
4	2015-02-27	0	-	-
5	2015-07-15	0	-	-
6	2015-10-23	0	-	-
7	2016-06-06	12	2	0.91
8	2016-12-23	33	3	0.75
9	2017-05-25	93	2	0.57
10	2017-08-22	0	-	-

Table III. Summary of Cluster Data for Long and Short Bubbles. The first column enumerates the three long bubbles and the ten additional peaks, which include the eight short bubbles, as qualified by the method of Section II.A. The second column gives the “present” time t_2 that was chosen ten business days prior to the peak of each bubble. The third column gives the percentage of fits that are qualified. The fourth column gives the optimal number of clusters. Finally, the fifth column lists the values of the Silhouette coefficient corresponding to the optimal cluster configuration. For further information on the details of the calculation, see Appendix E.

B.2. Long Bubbles

Table III shows that, for the long bubbles, between 40 and 68 percent of the fits are qualified according to the filtering constraints from Table II. For the first and second long bubbles, two main scenarios are identified while for the third longest bubble, we find up to nine scenarios, most of them being far to the future and with few elements, thus unreliable and discarded. Figure 12 contains three panels, one for each of the three long bubbles. Each panel shows the log-price of Bitcoin over the time interval of the full development of the bubble together with, for each of the identified clusters, the 15 LPPS fits (when they exist) having the lowest sum of squared errors (SSE).

The top-left panel shows the first long bubble that peaked on April 9, 2013. The scenario corresponding to the first largest cluster predicts the mean of t_c to be 14 days in the future (after t_2) with a standard deviation of 9 days. The second, smaller cluster predicts t_c to be 262 days in the future with a standard deviation of 34 days. Hence, the first scenario is essentially telling us that the bubble peak is imminent, with a quite precise bracketing of the true peak. Interpreting the fraction of fits belonging to this cluster as a proxy for the probability for this scenario, we can attribute a probability of 80% (101 fits) for this scenario.

The second less probable (20% probability, 25 fits) scenario considers as possible that the bubble may continue much longer, up to another year, before bursting. Given the interplay between the endogenous herding dynamics captured by these scenarios and the exogenous shocks that continuously punctuated the bubble development, in retrospect, these two scenarios appear reasonable possibilities when placing oneself at the time t_2 of the analysis. In hindsight, the highest probability scenario was the one that unfolded.

Turning to the second long bubble that peaked on December 4, 2013 and is shown in the top-right panel

of Figure 12, we again identify two clusters. The first cluster predicts t_c to be 4 days in the future (after t_2) with a standard deviation of 2 days and probability 82% (41 fits). The second cluster predicts t_c to be 76 days in the future with a standard deviation of 17 days, with a probability of 18% (9 fits). In this case, the bubble peaked 10 days after t_2 , i.e. later than predicted by the first cluster, but this discrepancy can be reconciled when noticing the double peak structure and flatness of this bubble end.

The third longest bubble that ended in December 2017 has 9 clusters, but only the two most probable scenarios in terms of the number of fits contained in the associated clusters are depicted in the lower left panel of Figure 12. The first cluster predicts t_c to be at the time t_2 of analysis with all the 90 qualified fits yielding the same critical time t_c , hence a vanishing standard deviation for the predicted t_c 's. This scenario has a probability of 20%. The second cluster predicts t_c to be 113 days in the future with a standard deviation of 7 days, with a probability of 15% (69 fits). The lower quality of these predictions resonates with their much lower probabilities, compared to the two first bubbles.

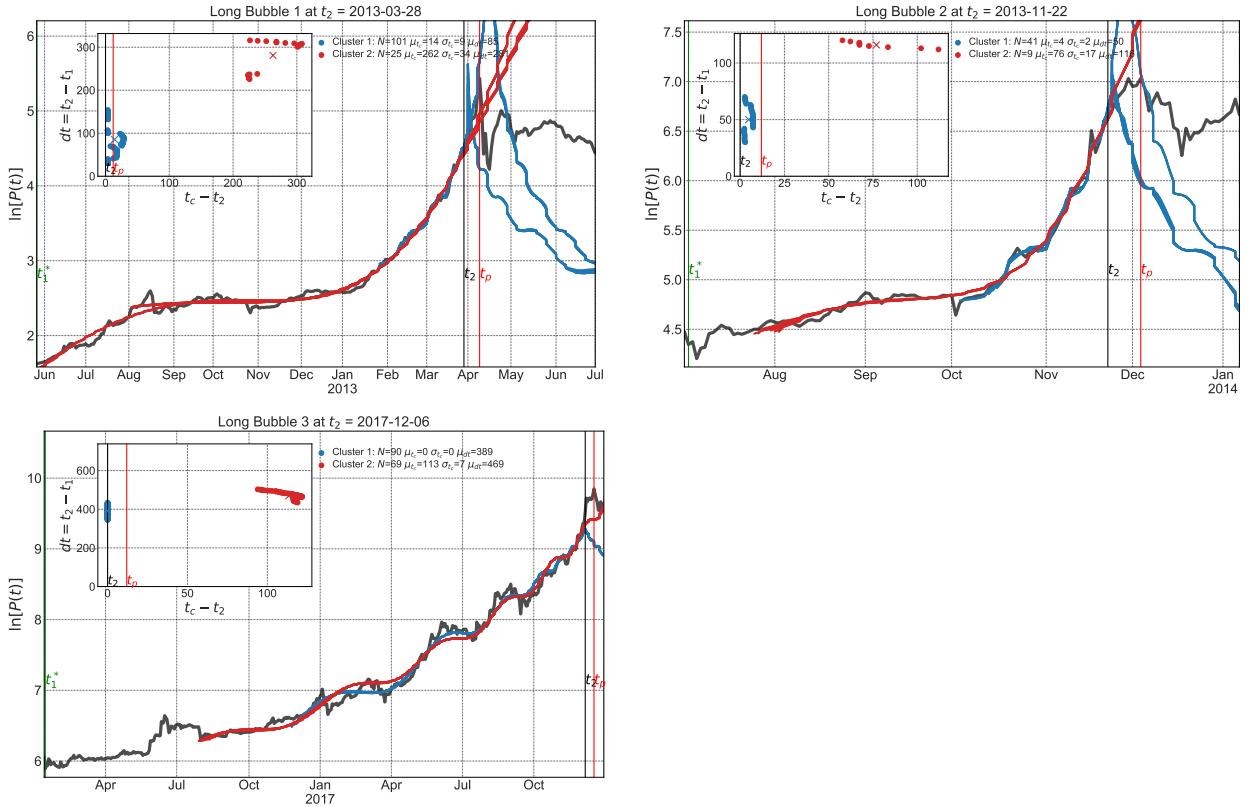


Figure 12. Scenario Projections for the Three Long Bubbles. Together with the log-price of Bitcoin over the time interval including the full development of each bubble, two sets of 15 LPPLS fits with Formula (1) having the lowest sums of square errors (SSE) are shown for each of the identified clusters (blue for Cluster 1 and red for Cluster 2). The “present” time t_2 of analysis is shown by the black vertical line, while the time t_p at which the bubble peaked is shown by a red vertical line. N represents the number of qualified fits, according to the filtering conditions in Table II associated to each cluster. For information about the construction of the clusters, we refer to Appendix E. In the legend, the pair (μ_{t_c}, μ_{dt}) represents the centroid value for the predicted $t_c - t_2$ and window sizes of Clusters 1 and 2. The corresponding standard deviations σ_{t_c} and σ_{dt} for the predicted t_c and associated windows of each cluster are also given. The inset plots show the two largest clusters in the $(dt := t_2 - t_1 + 1, t_c - t_2)$ space. Thereby, t_2 is placed at the origin on the horizontal axis and the vertical red line for t_p is placed relative to t_2 . The centroid of each cluster is indicated by a cross with the corresponding color (blue for Cluster 1 and red for Cluster 2).

B.3. Short Bubbles

Table III shows that, for the short bubbles, between 12 and 93 percent of the fits are qualified for the bubbles numbered 1 and 7-9, according to the filtering conditions listed in Table II. However, the Short Bubbles 2-6 and 10 do not have any qualified fits. This is in agreement with the previously calculated Multiscale Confidence Indicators, which show low or zero amplitude in advance of these bubbles, due to their short timespan. For the bubbles numbered 1 and 7-9, an optimal cluster configuration of two to three clusters are obtained according to the Silhouette procedure, described in Appendix E.

Figure 13 contains four panels, one for each of the four short bubbles having a non-zero percentage of qualified fits. This figure is constructed in the same way as Figure 12, with the condition that clusters are drawn in the plots only when they contain more than five fits.

For the Short Bubble 1 that peaked on August 16, 2012 and is shown in the top-left panel of Figure 13, there are two clusters. The first cluster predicts t_c to be 27 days in the future (after t_2) with a standard deviation of 2 days, with probability 51% (22 fits). The second cluster predicts t_c to be 26 days in the future with a standard deviation of 2 days, with a probability of 49% (21 fits).

Next, for the Short Bubble 7 that peaked on June 16, 2016 and is shown in the top-right panel of Figure 13, there are again two clusters, but only the first and largest one is represented as the second one has only three fits. The first cluster predicts t_c to be 1 days after t_2 with a standard deviation of 0 days, with probability 81% (13 fits). The small value of σ_{t_c} means that the fits in all time windows give the same t_c within one day of precision.

For the Short Bubble 8 that peaked on January 4, 2017 and is shown in the lower-left panel of Figure 13, there are three clusters. The first cluster predicts t_c to be 114 days after t_2 with a standard deviation of 9 days, with probability 51% (22 fits). The second cluster predicts t_c to be 52 days in the future with a standard deviation of 10 days, with a probability of 40% (17 fits). This poor result comes from the sudden acceleration of the price during the last 15 days before the peak, of which only the beginning phase was present in the windows of analysis. Before that sudden acceleration, the trend was informing the LPPLS calibration of a longer bubble duration. This poor result illustrates that a full operational prediction method should incorporate more time scales in order to be more reactive to such burgeoning acceleration. Again, it demonstrates the need to resolve also smaller time scales for complete short bubble detection, which may however become problematic when using data with a daily frequency.

For the Short Bubble 9 that peaked on June 6, 2017 and is shown in the lower-right panel of Figure 13, there are once more two clusters. The first cluster predicts t_c to be 10 days after t_2 with a standard deviation of 1 days, with probability 50% (21 fits). The second cluster predicts t_c to be also 10 days in the future with a standard deviation of 4 days and a probability of 50% (21 fits). This is a prediction very close to the true bubble peak.

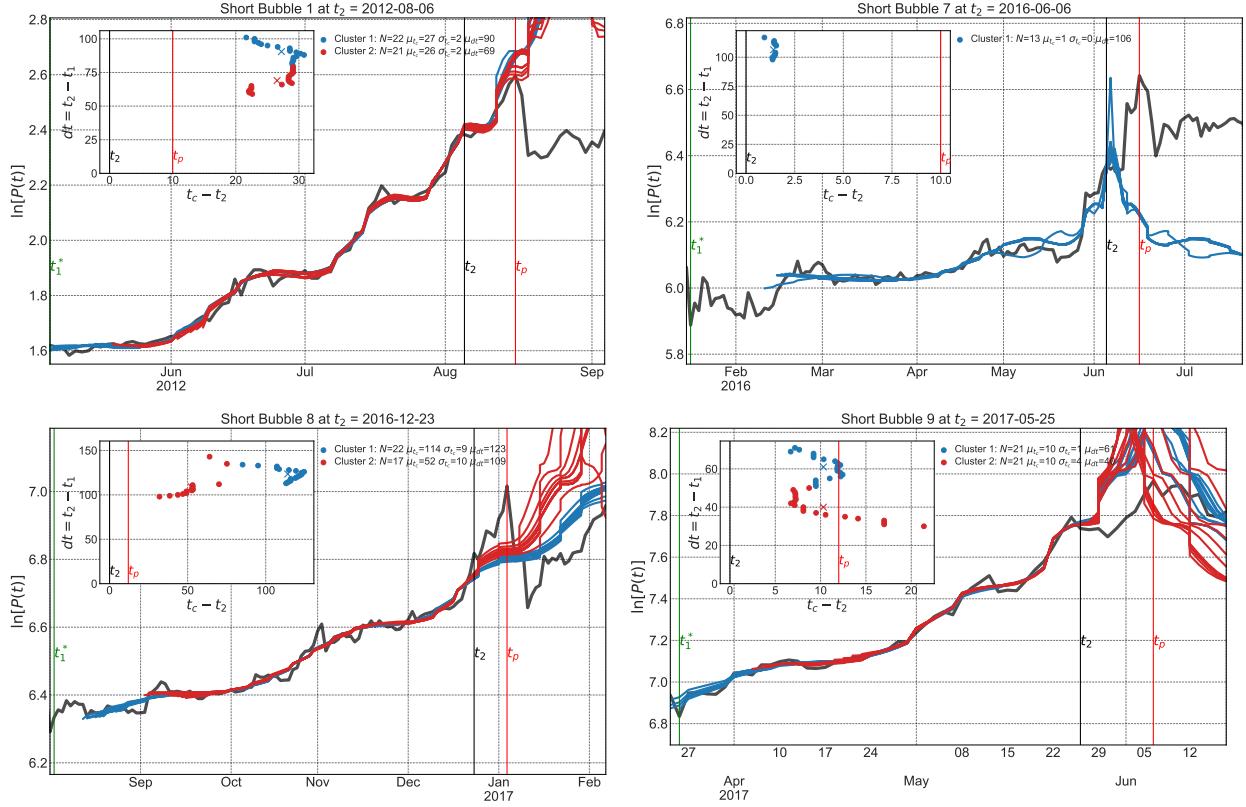


Figure 13. Scenario Projections for the Four Short Bubbles. Same presentation as in Figure 12.

V. Concluding Remarks

We have presented a detailed analysis of the dynamics of the price of bitcoin expressed in US dollars from January 2012 to February 2018. Given the impressive ascent of Bitcoin over this period, it is reasonable to ask whether bubbles have occurred, to characterize them if they exist and investigate the potential for their predictability. In order to robustly define bubble regimes and crashes in hindsight, as a reference for the predictive analysis, we have first presented a generalization of the *Epsilon Drawdown Method* to determine both periods of uninterrupted market growth (drawups) and regimes of uninterrupted market decrease (drawdowns). Together with the *Lagrange Regularisation Method* for detecting the beginning new bubble regimes, characterized by the transition from normal price growth to super-exponential price dynamics, we have been able to identify three major bubble peaks and ten additional smaller bubble peaks, that have punctuated the dynamics of Bitcoin price in the analyzed time period, as well as their associated bubble start times. The statistics of the found periods of abnormal price growth suggest that the identified peaks can indeed be attributed to genuine bubbles.

The three major peaks, associated to long bubbles, developed over a time scale from 6 months to almost 3 years and the price was multiplied from a factor almost 14 (1367% in 154 business days) to a factor 51 (5152% in 703 business days). If these numbers do not qualify these development as bubbles, then no other market regime could ever historically be called a bubble. For the 10 smaller peaks, the amplitude of growth was about 30% at the minimum over 50 days to a factor of 2 increase over 74 days. These numbers are also staggering in a general financial context.

We have then presented a number of quantitative metrics and graphs to understand at least a part of the socio-economic drivers behind the ascent of Bitcoin over the period of analysis, focusing primarily on the long bubbles. We stress in particular the attraction to Bitcoin as a novel asset type that was promoted and understood by many as a decentralized, independent medium of exchange and storage of value at the time of heavy interventions of central banks and growing skepticism about the reliability of the standard banking system. We highlight also the influence of mining and of miners' technical sophistication on Bitcoin price. The role of increasing demand from China, significantly contributing to the formation of the third long bubble, is emphasized.

Given the list of bubbles and a general understanding of their background, we proceeded by presenting a detailed quantitative, predictive analysis of these bubbles using metrics derived from the *Log-Periodic Power Law Singularity Model* (LPPLS), which embodies the super-exponential nature of price dynamics during a bubble, complemented by an accelerating spiral of volatility oscillations up to the bubble peak. We have presented two approaches.

The first one is based on the *LPPLS Multiscale Confidence Indicators*, defined as the fraction of qualified fits of the LPPLS model over multiple time windows. The time evolution of the LPPLS Confidence Indicators at three time scales has provided a global insight of the growing risks during bubble development.

The second approach uses a clustering method to group LPPLS fits over different time scales and predicted critical times t_c (the most probable time for the start of the crash ending the bubble). For each “present time” t_2 of the analysis, positioned in advance to the burst of the identified bubbles, clusters provided different plausible scenarios for the subsequent Bitcoin price evolution. We have documented these predictions for the three long bubbles and for the four short bubbles that our time scale of analysis was able to resolve. Overall, our predictive scheme seems significantly informative and useful for the three long bubbles. For three out of four short bubbles, the predictions are useful to warn of an imminent risk of crash.

In summary, we conclude from the bubble analysis that Bitcoin price, although highly volatile and frequently dropping sharply, has had the potential to recover from drawdowns in no time, enabling the price to grow to ever higher levels again and again during the past years. This leaves open the question whether the most recent change of regime has introduced a new era to Bitcoin's market behavior or whether the digital currency will continue its accelerating growth to new heights in the future. Technically, this academic study has gone one step closer to building a set of analytical tools and testing environment towards the real-time build-up of a early warning system of bubbles. We hope to report further on developments going on in the broader universe of cryptocurrencies in the future.

Appendix A: The Epsilon Drawup/Drawdown Methodology for Determination of Potential Bubble Peaks

We start out by calculating the daily discrete log-returns

$$r_i = \ln P[t_i] - \ln P[t_{i-1}], \quad i = 1, 2, \dots \quad (2)$$

with $t_i = t_0 + i\Delta_t$ and $\Delta_t = 1\text{day}$.

The first date t_{i_0} in our timeframe, corresponding to the discrete time $i_0 = 1$, is defined as the beginning of a drawup (drawdown) if $r_1 > 0$ ($r_1 < 0$). Then, for each subsequent $t_i > t_{i_0}$, we calculate the cumulative return up to this t_i as

$$p_{i_0, i} = \sum_{k=i_0}^i r_k = \ln P[t_i] - \ln P[t_{i_0}] \quad (3)$$

At each time, we need to check whether the current drawup (drawdown) phase is still active. We test this by calculating the largest deviation $\delta_{i_0, i}$ of the price trajectory from a previous maximum (minimum).

$$\delta_{i_0, i} = \begin{cases} \max_{i_0 \leq k \leq i} \{p_{i_0, k}\} - p_{i_0, i} & \text{for drawups} \\ p_{i_0, i} - \min_{i_0 \leq k \leq i} \{p_{i_0, k}\} & \text{for drawdowns} \end{cases} \quad (4)$$

The procedure is stopped at time i when the deviation exceeds a predefined tolerance ε .

$$\delta_{i_0, i} > \varepsilon \quad (5)$$

The stopping tolerance quantifies how much the price is allowed to move in the direction opposite to the drawup/drawdown trend.

When the procedure has been stopped, the end of the current drawup (drawdown) phase is determined as the time of the highest (lowest) price seen in the tested interval:

$$i_1 = \begin{cases} \arg \max_{i_0 \leq k \leq i} \{p_{i_0, k}\} & \text{for drawups} \\ \arg \min_{i_0 \leq k \leq i} \{p_{i_0, k}\} & \text{for drawdowns} \end{cases} \quad (6)$$

The Epsilon drawup/drawdown Procedure is restarted at time $i_1 + 1$. The start of the next drawup/drawdown period will then be located at this time. By construction of δ in Eq.(4) and the stopping condition, a drawup (resp. drawdown) is always followed by a drawdown (resp. drawup). The procedure is repeated until the full length of the analyzed time series is represented as a sequence of drawups and drawdowns. From the sequence identified by the Epsilon drawup/drawdown procedure, a set of peak times $\{t_{p,0}, t_{p,1}, \dots\}$, defined as the set containing the end times of all drawup periods in the sequence, is obtained. These times can be regarded as the peaks of candidate bubbles.

A reasonable way to find suitable values for the stopping tolerance ε is to incorporate the dynamics of realized return volatility and define

$$\varepsilon(\varepsilon_0, w) = \varepsilon_0 \sigma(w) \quad (7)$$

with realized volatility $\sigma(w)$ estimated over a moving window of the past w days and ε_0 being a constant multiplier.

The partitioning of the drawup-drawdown-sequence is strongly dependent on the choice of the pair (ε_0, w) . Selections resulting in large values of the tolerance ε will tend to yield a coarse long-term sequence while small ε -values result in more frequent interruption of a drawup or drawdown, yielding finer sequences. In order to account for this dependence, instead of selecting a single pair of values, we execute the Epsilon drawup/drawdown procedure for many different pairs $\{(\varepsilon_0, w)_j | j = 1, \dots, N_\varepsilon\}$. More specifically, we scan ε_0 in $[0.1, 5.0]$ in steps of $\Delta\varepsilon_0 = 0.1$ and w in $[10, 60]$ in steps of $\Delta w = 5$ days. This amounts to a total number of $N_\varepsilon = 50 \cdot 11 = 550$ pairs (ε_0, w) .

For each pair $(\varepsilon_0, w)_j$, we obtain an associated set of peak times

$$\mathcal{T}_j = \{t_{p,0}, t_{p,1}, \dots\}_j, \quad j = 1, \dots, N_\varepsilon, \quad (8)$$

which may or may not overlap with the sets of other pairs. The collection of all peak times that have been identified at least once across all pairs is then the union of the single sets \mathcal{T}_j :

$$\mathcal{T} = \bigcup_{j=1}^{N_\varepsilon} \mathcal{T}_j \quad (9)$$

We define N_{t_p} as the number of unique peak times, i.e. the cardinal of the union set \mathcal{T} . Next, for each of the N_{t_p} peak times in \mathcal{T} , we count the number of times $N_{t_{p,k}}$ that they occurred over all trials,

$$N_{t_{p,k}} = \sum_{j=1}^{N_\varepsilon} I_j(t_{p,k}) \quad (10)$$

where I_j is given by the indicator function

$$I_j(t_{p,k}) = \begin{cases} 1 & \text{if } t_{p,k} \text{ in } \mathcal{T}_j \\ 0 & \text{else.} \end{cases} \quad (11)$$

The index k in $t_{p,k}$ registers the k^{th} unique peak time in \mathcal{T} . We further divide the number of counts $N_{t_{p,k}}$ by the fixed total number of tested pairs N_ε to obtain a series of values between $[0, 1]$, each indicating the fraction of occurrence of the corresponding peak time with respect to the total number of trials. We group these in the set:

$$\mathcal{N} = \{n_{t_{p,k}} = \frac{N_{t_{p,k}}}{N_\varepsilon} | k = 1, \dots, N_{t_p}\}. \quad (12)$$

In this way, the information covered by N_ε pairs is condensed down into a single figure for each peak time.

Based on the series \mathcal{N} , we can now search for potential bubble peak times. We introduce the thresholds $D_{LT} = 0.95$ for long bubbles and $D_{ST} = 0.65$ for short bubbles. We look for peak times that occur more

frequently than D_{LT} or D_{ST} , and obtain the sets of potential long and short bubble peaks as:

$$\begin{aligned}\mathcal{T}_{LT} &= \{t_{p,k} \mid n_{t_{p,k}} \geq D_{LT} \text{ for } k = 1, \dots, N_{t_p}\} \\ \mathcal{T}_{ST} &= \{t_{p,k} \mid n_{t_{p,k}} \geq D_{ST} \text{ for } k = 1, \dots, N_{t_p}\} \setminus \mathcal{T}_{LT}\end{aligned}\quad (13)$$

Thereby, note that we exclude identified long bubble peaks from the short bubble set in order to avoid taking the same peak times into account twice.

Appendix B: The Log-Periodic Power Law Singularity Model

In a bubble regime, the observed price trajectory of a given asset decouples from its intrinsic fundamental value (Kindleberger, 1978; Sornette, 2003). For a given fundamental value, the JLS model (Johansen et al., 1999, 2000) assumes that the logarithm of the observable asset price $p(t)$ follows

$$\frac{dp}{p} = \mu(t)dt + \sigma(t)dW - \kappa dj, \quad (14)$$

where $\mu(t)$ is the expected return, $\sigma(t)$ is the volatility, dW is the infinitesimal increment of a standard Wiener process and dj represents a discontinuous jump such that $j = n$ before and $j = n + 1$ after a crash occurs (where n is an integer). The parameter κ quantifies the amplitude of a possible crash.

In the network structure that is underlying the LPPL, two types of agents are considered: the first group consists of traders with rational expectations (Blanchard and Watson, 1982), while the second one is formed by noise traders. Noise traders are susceptible to show imitation and herding behaviour as a group. Their collective behaviour may destabilise asset prices. Johansen et al. (2000) propose that the behaviour of the agent network can be incorporated by writing the *crash hazard rate* $h(t)$ in the following form:

$$h(t) = \alpha(t_c - t)^{m-1}(1 + \beta \cos(\omega \ln(t_c - t) - \phi')), \quad (15)$$

where α, β, ω and t_c are parameters. Eq. (15) tells us that the risk of a crash resulting from herding behaviour is a sum of a power law singularity ($\alpha(t_c - t)^{m-1}$) and accelerating large scale amplitude oscillations that are periodic in the logarithm of the time to the singularity (or critical time) t_c . The power law singularity embodies the positive feedback mechanism associated with the herding behaviour of noise traders. At time $t = t_c$, the power law reaches the singularity. Seyrich and Sornette (2016) have recently presented a percolation-based model providing a micro-foundation for this singular behavior. The log-periodic oscillations represent the tension and competition between the two types of agents that tend to create deviations around the faster-than-exponential price growth as the market approaches a finite-time-singularity at t_c .

The no-arbitrage condition imposes that the excess return $\mu(t)$ during a bubble phase is proportional to the crash hazard rate given by Eq.(15). Indeed, setting $E[dp] = 0$, and assuming that no-crash has yet occurred ($dj = 0$), this yields $\mu = \kappa h(t)$, since $E[dj] = h(t)dt$ by definition of $h(t)$. By integration, we obtain the trajectory of the expected log-price during a bubble phase, conditional on the crash not having happened yet, as

$$E[\ln p(t)] = A + B|t_c - t|^m + C|t_c - t|^m \cos(\omega \ln |t_c - t| - \phi), \quad (16)$$

with $B = -\kappa\alpha/m$ and $C = -k\alpha\beta/\sqrt{m^2 + \omega^2}$. Note that the formula extends the price dynamics beyond t_c

by replacing $t_c - t$ by $|t_c - t|$, which corresponds to assuming a symmetric behavior of the average of the log-price around the singularity at t_c .

Bubble regimes are in general characterized by $0 < m < 1$ and $B < 0$. The first condition $m < 1$ writes that a singularity exists (the momentum of the expected log-price diverges at t_c for $m < 1$), while $m > 0$ ensures that the price remains finite at the critical time t_c . The second condition $B < 0$ expresses that the price is indeed growing super-exponentially towards t_c (for $0 < m < 1$).

Appendix C: Estimation of the The Log-Periodic Power-Law Singularity Model

The log-price of a given instrument can be described via

$$LPPL(\phi, t) = A + B(f) + C_1(g) + C_2(h), \quad (17)$$

where $\phi = \{A, B, C_1, C_2, m, \omega, t_c\}$ is a (1×7) vector of parameters that we would like to determine and

$$f \equiv (t_c - t)^m, \quad (18)$$

$$g \equiv (t_c - t)^m \cos(\omega \ln(t_c - t)), \quad (19)$$

$$h \equiv (t_c - t)^m \sin(\omega \ln(t_c - t)). \quad (20)$$

Fitting Eq.(17) to the log-price time-series amounts to search for the parameter set ϕ^* that yields the smallest N -dimensional distance between realisation and theory. Mathematically, using the L^2 norm, we form the following sum of squares of residuals

$$F(t_c, m, \omega, A, B, C_1, C_2) = \sum_{i=1}^N \left[\ln[P(t_i)] - A - B(f_i) - C_1(g_i) - C_2(h_i) \right]^2, \quad (21)$$

for $i = 1, \dots, N$. We proceed in two steps. First, slaving the linear parameters $\{A, B, C_1, C_2\}$ to the remaining nonlinear parameters $\phi = \{t_c, m, \omega\}$, yields the cost function $\chi^2(\phi)$

$$\chi^2(\phi) := F_1(t_c, m, \omega) = \min_{\{A, B, C_1, C_2\}} F(t_c, m, \omega, A, B, C_1, C_2) = F(t_c, m, \omega, \hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2), \quad (22)$$

where the hat symbol $\hat{\cdot}$ indicates estimated parameters. This is obtained by solving the optimization problem

$$\{\hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2\} = \arg \min_{\{A, B, C_1, C_2\}} F(t_c, m, \omega, A, B, C_1, C_2), \quad (23)$$

which can be obtained analytically by solving the following matrix equations

$$\begin{bmatrix} N & \sum f_i & \sum g_i & \sum h_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i & \sum f_i h_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 & \sum g_i h_i \\ \sum h_i & \sum f_i h_i & \sum g_i h_i & \sum h_i^2 \end{bmatrix} \begin{bmatrix} \hat{A} \\ \hat{B} \\ \hat{C}_1 \\ \hat{C}_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i f_i \\ \sum y_i g_i \\ \sum y_i h_i \end{bmatrix} \quad (24)$$

Second, we solve the nonlinear optimisation problem involving the remaining nonlinear parameters m, ω, t_c :

$$\{\hat{t}_c, \hat{m}, \hat{\omega}\} = \arg \min_{\{t_c, m, \omega\}} F_1(t_c, m, \omega). \quad (25)$$

The model is calibrated on the data using the Ordinary Least Squares method, providing estimations of all parameters $t_c, \omega, m, A, B, C_1, C_2$ in a given time window of analysis. For each fixed data point t_2 (corresponding to a fictitious ‘‘present’’ up to which the data is recorded), we fit the price time series in shrinking windows (t_1, t_2) of length $dt := t_2 - t_1 + 1$ decreasing from 720 trading days to 30 trading days. We shift the start date t_1 in steps of 1 trading day, thus giving us 691 windows to analyse for each t_2 . In order to minimise calibration problems and address the sloppiness of the model (16) with respect to some of its parameters (and in particular t_c), we use a number of filters to select the viable solutions, which are summarised in Table IV. For further information about the sloppiness of the LPPLS model, we refer to (Sornette et al., 2015; Demos and Sornette, 2017a; Filimonov et al., 2016). These filters derive from the empirical evidence gathered in investigations of previous bubbles (Zhou and Sornette, 2003; Zhang et al., 2016; Sornette et al., 2015). Only those calibrations that meet the conditions given in Table (IV) are considered valid and the others are discarded.

Previous calibrations of the JLS model have further shown the value of additional constraints imposed on the nonlinear parameters in order to remove spurious calibrations (false positive identification of bubbles) (Demos and Sornette, 2017a; Bree et al., 2013; Geraskin and Fantazzini, 2011).

Parameter		Filter Bounds	Description	Pre-Condition
Power Law Amplitude	B	$(-\infty, 0)$	-	-
Power Law Exponent	m	$(0, 1)$	-	-
Log-periodic Frequency	ω	$[4, 25]$	-	-
Critical Time	t_c	$(0, dt_i)$	-	-
Damping	D	$[0.5, \infty)$	$D = \frac{m B }{\omega C }$	-
Number of Oscillations	O	$[2.5, \infty)$	$O = \frac{\omega}{2\pi} \ln \frac{t_c - t_1}{t_c - t_2}$	if $ C/B \geq 0.05$

Table IV. Filter Conditions for Qualified LPPLS Fits. The table gives the parameter bounds that were used to filter for qualified LPPL fits. The constraint on B ensures the existence of a positive bubble. For m and t_c , the boundaries are excluded to avoid singular behaviors in the search algorithm. The damping factor quantifies the allowed downwards movements of bubble fits. The O -parameter measures the number of oscillations that occur within the fit window $[t_1, t_2]$. The filter on O is applied only when the amplitude of the oscillations, as quantified by C , is sufficiently large relative to the power law amplitude B .

Appendix D: The Lagrange Regularisation Approach for the Determination of Bubble Start Times

The LPPL Model is assumed to be a valid description for the log-price trajectory only if the underlying asset is in a bubble phase. Hence, if the LPPL is fit to time periods corresponding to phases of normal price growth, invalid / spurious fit results might be the consequence. Therefore, it is important to determine the start of a bubble at first, and then apply the LPPL for time windows with start points later than the found bubble start date.

A solution for the bubble identification problem has been introduced recently by Demos and Sornette (2017b), which proposed the Lagrange Regularisation Approach. The bubble start time t_1^* is found to be quite rigid over different points in time where the method is applied (Demos and Sornette, 2017a). The

idea of the method is based on the observation that, for fits conducted at the same fit window end time t_2 but at increasing fit window sizes dt_i (i.e. longer fit window horizons), the average sum of squared errors (SSE/N) exhibits an approximate linear behavior as a function of the fit window length, at least over the span of interest from a few tens to a few hundred data points. Indeed, smaller windows tend to give a smaller residual error per data point (SSE/N), since they present a smaller number of degrees of freedom. Therefore, in order to compare fits performed for different window sizes and find the true single best window in terms of the cost function, the slope of the dependence of SSE/N vs dt is calculated by linear regression and this dependence is detrended by subtracting the resulting straight regression line. This yields a set of detrended SSE/N values over all window sizes. The fit with the smallest detrended SSE/N is then selected as the “best” fit. The left-end point of the corresponding time window of the defines the start time t_1^* of the bubble. Based on this results, as no data earlier than the bubble start shall be used for fitting, all fits from the collection that were conducted at window sizes dt_i with $t_{1,i} = t_2 - dt_i < t_1^*$, i.e. had their starting times before the bubble start, are rejected and only the remaining fits are considered valid.

Appendix E: k-means Clustering and the Silhouette Metric

The k -means algorithm aims at grouping a set of N measured data points $X = \{x_1, x_2, x_3, \dots, x_n\}$ into a predefined number $k \geq 2$ of clusters $\mathcal{C}_k = \{C_1, C_2, \dots, C_k\}$, $C_i \subset X$, so as to minimize the variance of each cluster. The clusters are non-overlapping, $C_i \cap C_j = \emptyset$, i.e. no data point can be assigned to more than one cluster. Furthermore, any data point is contained in one of the clusters, $\cup_{i=1}^k C_i = X$.

Let $\mu = \{\mu_1, \mu_2, \dots, \mu_k\}$ be the centers of the clusters. They are calculated as the empirical mean of the data points in each cluster,

$$\mu_i = \frac{1}{N_i} \sum_{x_j \in C_i} x_j \quad (26)$$

with $\{N_1, N_2, \dots, N_k\}$ representing the number of members in each cluster, summing up in total to $N = \sum N_i$.

Based on the minimum variance condition, the optimal cluster configuration $\hat{\mathcal{C}}_k$ is obtained by solving the following minimization problem

$$\hat{\mathcal{C}}_k = \arg \min_C \sum_{i=1}^k \sum_{x_j \in C_i} \|x_j - \mu_i\|^2 \quad (27)$$

where $\|\cdot\|$ denotes the Euclidean distance measure. Indeed, the objective function is minimized when each point is assigned to the cluster to whose center μ_i its distance is minimal.

At the beginning, no points are assigned to any cluster. Thus, the algorithm must be initialized by randomly selecting or guessing the positions of the k cluster centers. The optimal positions of the cluster centers are then iteratively calculated. In each iteration, Equation (27) is solved. As long as this still changes the resulting composition of clusters, subsequently, the cluster centers are recalculated based on the new data point clusters according to Equation (26). As soon as no more data points are reassigned to new clusters, the algorithm stops.

The choice of the number of clusters k largely affects the capability of the results to accurately reflect the underlying structure of the data. If it is chosen too small, the variance of clusters might blow up, while an overly large number of clusters might cause unnecessarily fine data separation and excessive usage of

computational resources. A tool for the “right” selection of the number of clusters is the *Silhouette Metric*.

In order to optimize the number of clusters, we firstly calculate cluster configurations for various k ranging from 2 to 10. We then employ the Silhouette Metric to determine the optimal number of clusters \hat{k} .

1. The optimal cluster configuration $\hat{\mathcal{C}}_k$ for a fixed value of k is determined.
2. For each data point x_i in the resulting k clusters, its average distance a_i to all other points in its own cluster is calculated.
3. In the same manner, the average distance of each data point x_i to the data points in each of the other clusters is calculated. The minimum such value with respect to all clusters, which is the critical one, is denoted by b_i .
4. The Silhouette coefficient is defined as

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (28)$$

for each individual data point x_i .

5. From the individual Silhouette scores, the mean Silhouette score \bar{s}_k of all data points is computed for the respective cluster configuration.

The Silhouette coefficient of a data point is a measure of how closely it is related to data within its own or another cluster. By construction, it is bounded between -1 and 1. We wish to find the cluster configuration with minimal dissimilarity of each data point to its own cluster. This translates into the condition that the average Silhouette score should be as close to one as possible. The optimal number of clusters is then given by:

$$\hat{k} = \arg \min_k \{1 - \bar{s}_k\} \quad (29)$$

References

- Arnold, A. and Chen, L. Y. (2017). China's Largest Bitcoin Exchanges to Allow Withdrawals Again. *Bloomberg* (<https://www.bloomberg.com/news/articles/2017-06-01/china-s-largest-bitcoin-exchanges-to-again-allow-withdrawals>).
- Barclay, A. (2017). China's bitcoin crackdown raises cryptocurrency demand as they move to peer-to-peer platforms. *South China Morning Post* (<http://www.scmp.com/tech/innovation/article/2125282/chinas-bitcoin-crackdown-raises-cryptocurrency-demand-they-move-peer>).
- Bariviera, A. F. (2017). The inefficiency of Bitcoin revisited: A dynamic approach. *Economics Letters*, 161:1–4.
- Begušić, S., Kostanjčar, Z., Stanley, H. E., and Podobnik, B. (2018). Scaling properties of extreme price fluctuations in Bitcoin markets. (<http://arxiv.org/abs/1803.08405>).
- Bianchetti, M., Ricci, C., and Scaringi, M. (2018a). Are Cryptocurrencies Real Financial Bubbles? Evidence from Quantitative Analyses. Available at SSRN: <https://www.ssrn.com/abstract=3092427>.
- Bianchetti, M., Ricci, C., and Scaringi, M. (2018b). Forecasting Crypto Crashes with Bubble Analysis. *Risk.net* (<https://www.risk.net/comment/5388551/forecasting-crypto-crashes-with-bubble-analysis>).
- Bitcoinity.org (2018). [Online] Available at: <http://data.bitcoinity.org/markets/volume/> (Accessed February 2018).
- Blanchard, O. J. and Watson, M. W. (1982). Bubbles, rational expectations and financial markets. In Wachtel, P., editor, *Crisis in the Economic and Financial Structure*, pages 295–315. Lexington Books, Lexington, MA.
- Blockchain.info (2018). *Blockchain Luxembourg*. [Online] Available at: www.blockchain.info/charts (Accessed February 2018).
- Bouri, E., Molnár, P., Azzi, G., Roubaud, D., and Hagfors, L. I. (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20:192–198.
- Bovaird, C. (2016). Bitcoin Price Climbs Over 50% in First Half of 2016. *Coindesk* (<https://www.coindesk.com/bitcoin-price-h1-report-2016/>).
- Bovaird, C. (2017). Despite Drop in Volume, Traders Stick With China's Bitcoin Exchanges. *Coindesk* (<https://www.coindesk.com/volume-traders-china-bitcoin-exchanges/>).
- Bree, D., Challet, D., and Peirano, P. (2013). Prediction accuracy and sloppiness of log-periodic functions. *Quantitative Finance*, 3:275–280.
- BTC.com (2018). *btc.com Mining Pool Statistics*. [Online] Available at: <https://pool btc.com/pool-stats> (Accessed February 2018).
- BTC.TOP (2018). *btc.com Mining Pool Statistics*. [Online] Available at: <https://btc.com/stats/pool/BTC.TOP> (Accessed February 2018).

- Buterin, V. (2013). The Bitcoin Crash: An Examination. *Bitcoin Magazine* (<https://bitcoinmagazine.com/articles/the-bitcoin-crash-an-examination-1365911041/>).
- Chen, L. Y. and Lee, J. (2017). Bitcoin Tumbles as PBOC Declares Initial Coin Offerings Illegal. *Bloomberg* (<https://www.bloomberg.com/news/articles/2017-09-04/china-central-bank-says-initial-coin-offerings-are-illegal>).
- CoinMarketCap (2018). [Online] Available at: <https://coinmarketcap.com/> (Accessed February 2018).
- Datastream (2018). *Thomson Reuters Datastream*. [Online] Available at: Subscription Service (Accessed February 2018).
- Demos, G. and Sornette, D. (2017a). Birth or Burst of Financial Bubbles: Which One is Easier to Diagnose? *Quantitative Finance*, 17(5):657–675.
- Demos, G. and Sornette, D. (2017b). Lagrange Regularisation Approach to Compare Nested Data Sets and Determine Objectively Financial Bubbles' Inceptions. *Swiss Finance Institute Research Paper No. 18-20*. Available at SSRN: <https://ssrn.com/abstract=3007070>.
- Donier, J. and Bouchaud, J. P. (2015). Why do markets crash? Bitcoin data offers unprecedented insights. *PLoS ONE*, 10(10):23–25.
- Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5):1575–1617.
- Fantazzini, D., Nigmatullin, E., Sukhanovskaya, V., and Ivliev, S. (2016). Everything You Always Wanted to Know About Bitcoin Modelling But Were Afraid to Ask. *Applied Econometrics, Forthcoming*. Available at SSRN: <https://ssrn.com/abstract=2794622>.
- Farrell, M. (2013). Bitcoin prices surge post-Cyprus bailout. *CNNMoney* (<http://money.cnn.com/2013/03/28/investing/bitcoin-cyprus/index.html>).
- Filimonov, V., Demos, G., and Sornette, D. (2016). Modified profile likelihood inference and interval forecast of the burst of financial bubbles. *Quantitative Finance*, 19:1–20.
- Filimonov, V. and Sornette, D. (2015). Power law scaling and “dragon-kings” in distributions of intraday financial drawdowns. *Chaos, Solitons and Fractals*, 74:27–45.
- Garcia, D. and Schweitzer, F. (2015). Social signals and algorithmic trading of bitcoin. *Open Science*, 2(9). DOI: <https://doi.org/10.1098/rsos.150288>.
- Garcia, D., Tessone, C. J., Mavrodiev, P., and Perony, N. (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *Journal of The Royal Society Interface*, 11(99). DOI: <https://doi.org/10.1098/rsif.2014.0623>.
- Gautham (2016). Bitcoin Price \$800 Imminent; Boosted by China Capital Outflows. *NewsBTC* (<https://www.newsbtc.com/2016/11/01/bitcoin-price-800-imminent-boosted-china-capital-outflows/>).
- Geraskin, P. and Fantazzini, D. (2011). 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):366–391.

- Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M., and Siering, M. (2014). Bitcoin - Asset or currency? Revealing users' hidden intentions. *ECIS 2014 Proceedings - 22nd European Conference on Information Systems*.
- Goh, B. (2017). China's OkCoin, Huobi exchanges to stop bitcoin withdrawals. *Reuters* (<https://www.reuters.com/article/us-china-bitcoin/chinas-okcoin-huobi-exchanges-to-stop-bitcoin-withdrawals-idUSKBN15P0HE>).
- Google Trends (2018). [Online] Available at: <https://trends.google.de/trends/> (Accessed February 2018).
- Higgins, S. (2017). China's Central Bank to Continue Bitcoin Exchange Inspections. *CoinDesk* (<https://www.coindesk.com/chinas-central-bank-continue-bitcoin-exchange-inspections/>).
- Hill, K. (2013). Bitcoin in China : The Fall-out From Chinese Government Banning Real World Use. *Forbes* (<https://www.forbes.com/sites/kashmirhill/2013/12/06/bitcoin-in-china-the-fall-out-from-chinese-government-banning-real-world-use/#4ee4bcc8481a>).
- Huang, Z. (2017). It's over : China's biggest bitcoin exchanges will halt all trading services for local customers. *Quartz* (<https://qz.com/1079908/huobi-and-okcoin-chinas-two-biggest-bitcoin-exchanges-will-halt-all-trading-services-for-local-customers/>).
- Johansen, A., Ledoit, O., and Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance*, 2:219–255.
- Johansen, A. and Sornette, D. (1998). Stock market crashes are outliers. *European Physical Journal B*, 1:141–143.
- Johansen, A. and Sornette, D. (2001). Large Stock Market Price Drawdowns Are Outliers. *Journal of Risk*, 4(2):69–110.
- Johansen, A. and Sornette, D. (2010). Shocks, crashes and bubbles in financial markets. *Brussels Economic Review (Cahiers économiques de Bruxelles)*, 53(2):201–253.
- Johansen, A., Sornette, D., and Ledoit, O. (1999). Predicting Financial Crashes Using Discrete Scale Invariance. *Journal of Risk*, 1(4):5–32.
- Kim, Y. B., Lee, J., Park, N., Choo, J., Kim, J. H., and Kim, C. H. (2017). When Bitcoin encounters information in an online forum: Using text mining to analyse user opinions and predict value fluctuation. *PLoS ONE*, 12(5):1–14.
- Kindleberger, C. P. (1978). *Manias, Panics and Crashes: A History of Financial Crises*. Macmillan, London, 3 edition.
- Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, 3:1–7.
- Macdonell, A. (2014). *Popping the Bitcoin Bubble: An application of log-periodic power law modeling to digital currency*. PhD thesis.

- Mossavar-Rahmani, S., Nelson, B., Weir, M., Minovi, M., Ubide, A., Asl, F., Dibo, M., and Rich, M. C. (2018). (Un) Steady as She Goes. Technical Report January, Goldman Sachs Investment Management Division.
- Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *bitcoin.org* (<https://bitcoin.org/bitcoin.pdf>).
- Olszewicz, J. (2017). Bitcoin Price Analysis - PBoC crushes volume. *BraveNewCoin* (<https://bravenewcoin.com/news/bitcoin-price-analysis-pboc-crushes-volume/>).
- Osterrieder, J. and Lorenz, J. (2017). a Statistical Risk Assessment of Bitcoin and Its Extreme Tail Behavior. *Annals of Financial Economics*, 12(01):1750003.
- Pagliery, J. (2013). FBI shuts down online drug market Silk Road. *CNN Money* (<http://money.cnn.com/2013/10/02/technology/silk-road-shut-down/index.html>).
- Pichl, L. and Kaizoji, T. (2017). Volatility Analysis of Bitcoin Price Time Series. *Quantitative Finance and Economics*, 1(4):474–485.
- Rapoza, K. (2017). Cryptocurrency Exchanges Officially Dead In China. *Forbes* (<https://www.forbes.com/sites/kenrapoza/2017/11/02/cryptocurrency-exchanges-officially-dead-in-china/>).
- Redman, J. (2017). It Came from the East: How Asia Bolstered Cryptocurrencies in 2017. *Bitcoin.com* (<https://news.bitcoin.com/it-came-from-the-east-how-asia-bolstered-cryptocurrencies-in-2017/>).
- Rizzo, P. (2017). China's Central Bank Issues Warnings to Major Bitcoin Exchanges. *CoinDesk* (<https://www.coindesk.com/chinas-central-bank-issues-warnings-major-bitcoin-exchanges/>).
- Seyrich, M. and Sornette, D. (2016). Micro-foundation using percolation theory of the finite-time singular behavior of the crash hazard rate in a class of rational expectation bubbles. *International Journal of Modern Physics C*, 27(10):1650113 (20 pages).
- Solana, J. (2017). BTC China extends bitcoin withdrawal freeze into March. *CalvinAyre* (<https://calvinayre.com/2017/02/17/bitcoin/btc-china-extends-bitcoin-withdrawal-freeze-march/>).
- Sornette, D. (2003). *Why stock markets crash: Critical events in complex financial systems*. Princeton University Press, New Jersey.
- Sornette, D., Demos, G., Zhang, Q., Cauwels, P., Filimonov, V., and Zhang, Q. (2015). Real-Time Prediction and Post-Mortem Analysis of the Shanghai 2015 Stock Market Bubble and Crash. *Journal of Investment Strategies*, 4(4):77–95.
- Sornette, D., Johansen, A., and Bouchaud, J. P. (1995). Stock market crashes, precursors and replicas. *Journal of Physics*, 6:167–175.
- Suberg, W. (2016). China's Bitcoin Capital Flight Hits Mainstream As Analysts Fear Crackdown. *Cointelegraph* (<https://cointelegraph.com/news/chinas-bitcoin-capital-flight-hits-mainstream-as-analysts-fear-crackdown>).

- Thompson, M. (2013). Rush to ATMs in Cyprus on EU bailout tax. *CNN Money* (<http://money.cnn.com/2013/03/16/news/economy/europe-cyprus-bailout/index.html>).
- Urquhart, A. and Zhang, H. (2018). Is Bitcoin a Hedge or Safe-Haven for Currencies? An Intraday Analysis. pages 1–25. Available at SSRN: <https://ssrn.com/abstract=3114108>.
- US Senate (2015). Beyond Silk Road: Potential Risks, Threats, and Promises of Virtual Currencies. *Committee on Homeland Security and Governmental Affairs United States Senate*.
- ViaBTC (2018). *btc.com Mining Pool Statistics*. [Online] Available at: <https://btc.com/stats/pool/ViaBTC> (Accessed February 2018).
- Warner, B. (2013). Jittery Spaniards Seek Safety in Bitcoins. *Bloomberg* (<https://www.bloomberg.com/news/articles/2013-03-20/jittery-spaniards-seek-safety-in-bitcoins>).
- Wheatley, S., Sornette, D., Reppen, M., Huber, T., and Gantner, R. N. (2018). Are Bitcoin Bubbles Predictable? Combining a Generalised Metcalfe's Law and the LPPLS Model. *Swiss Finance Institute Research Paper No. 18-22*. Available at SSRN: <https://ssrn.com/abstract=3141050>.
- Wile, R. (2013). Bitcoin Crosses \$ 700 As Senate Hearing Wraps. *Business Insider* (<http://www.businessinsider.com/senate-bitcoin-hearing-2013-11>).
- Wilson, J., Dombey, D., and Spiegel, P. (2012). Cyprus requests eurozone bailout. *Financial Times* (<https://www.ft.com/content/80320e0e-bed0-11e1-b24b-00144feabdc0>).
- Wong, J. I. (2017). China's bitcoin traders are finding new ways to trade after an official clampdown. *Quartz* (<https://qz.com/915316/chinas-bitcoin-traders-are-finding-new-ways-to-trade-after-an-official-clampdown/>).
- Woo, W. (2017). Estimating China's Real Bitcoin Trading Volumes. *Coindesk* (<https://www.coindesk.com/estimating-data-china-real-bitcoin-trading-volumes/>).
- Wu, K., Wheatley, S., and Sornette, D. (2018). Classification of crypto-coins and tokens from the dynamics of their power law capitalisation distributions. (<http://arxiv.org/abs/1803.03088>).
- Zhang, Q., Zhang, Q., and Sornette, D. (2016). Early warning signals of financial crises with multi-scale quantile regressions of Log-Periodic Power Law Singularities. *PLoS ONE*, 11(11):e0165819, pp. 1–43.
- Zhao, W. (2017). China's Bitcoin Exchanges Are Shifting Business Models. *Coindesk* (<https://www.coindesk.com/chinas-bitcoin-exchanges-shift-p2p-model-domestic-crackdown/>).
- Zhou, W.-X. and Sornette, D. (2003). Evidence of a worldwide stock market log-periodic anti-bubble since mid-2000. *Physica A: Statistical Mechanics and its Applications*, 330(3–4):543–583.
- Zuegel, D. K. (2014). What happened at MtGox ? The Collapse of the World's Largest Bitcoin Exchange. *The Stanford Review* (<https://medium.com/stanfordreview/what-happened-at-mtgox-the-collapse-of-the-worlds-largest-bitcoin-exchange-1f4b18c4e207>).