Decisions in Economics and Finance (2020) 43:187–228
https://doi.org/10.1007/s10203-019-00262-x

Market attention and Bitcoin price modeling: theory, estimation and option pricing

Alessandra Cretarola[1] · **Gianna Figà-Talamanca**[2] · **Marco Patacca**[3,4]

Received: 16 February 2018 / Accepted: 13 July 2019 / Published online: 26 July 2019

Abstract
The goal of this paper is to provide a novel quantitative framework to describe the Bitcoin price behavior, estimate model parameters and study the pricing problem for Bitcoin derivatives. To this end, we propose a continuous time model for Bitcoin price motivated by the findings in recent literature on Bitcoin, showing that price changes are affected by sentiment and attention of investors, see e.g., (Kristoufek in Sci Rep 3:3415, 2013, PLoS ONE 10(4):e0123923, 2015; Bukovina and Marticek in Sentiment and bitcoin volatility. Technical report, Mendel University in Brno, Faculty of Business and Economics 2016). Economic studies, such as Yermack (Handbook of Digital Currency, chapter second. Elsevier, Amsterdam, pp 31–43, 2015), have also classified Bitcoin as a speculative asset rather than a currency due to its high volatility. Building on these outcomes, the price dynamics in our suggestion is indeed affected by an exogenous factor which represents market attention in the Bitcoin system. We prove the model to be arbitrage-free under a mild condition and we fit the model to historical data for the Bitcoin price; after obtaining a approximate formula for the likelihood, parameter values are estimated by means of the profile likelihood method. In addition, we derive a closed pricing formula for European-style derivatives on Bitcoin, the performance of which is assessed on a panel of market prices for Plain Vanilla options quoted on www.deribit.com.

Keywords Bitcoin · Market attention · Stochastic models · Option pricing · Maximum likelihood estimation

Mathematics Subject Classification 91B70 · 91G20 · 91G70

JEL Classification C22 · C58 · G02

✉ Alessandra Cretarola
alessandra.cretarola@unipg.it

Extended author information available on the last page of the article

1 Introduction

Bitcoin was first introduced in 2009 by a computer scientist, known under the pseudonym Satoshi Nakamoto, as an electronic payment system between peers and is based on an open-source software which generates a peer-to-peer network, see (Nakamoto 2008). Opposite to traditional transactions, which are based on trust in financial intermediaries, this system relies on the network, on the fixed rules and on cryptography. Bitcoins can be purchased on online trading platforms (so-called *exchanges*) by using fiat currencies. Further, payments can be made in Bitcoins for several online services and goods and its use is increasing. At very low expenses it is also possible to send the cryptocurrency internationally. Concerns about the use of Bitcoin include reputation issues due anonymous transactions, such as money laundering or the possible financing of criminal activities , and cyber-security issues, since it can be only deposited in digital wallets which are vulnerable to hacking attacks. In spite of the above critics, Bitcoin has experienced a rapid growth both in value and in the number of transactions. From an economic viewpoint, one of the main concerns about Bitcoin is whether it should be considered a currency, a commodity or a stock. The conclusion in Yermack (2015) is that Bitcoin behaves as a high volatility stock and that most transactions on Bitcoins are aimed to speculative investments. Recently, literature about Bitcoin has paid much attention to Bitcoin price dynamics and, in particular, to the identification of possible price drivers. In Kristoufek (2013, 2015); Kim et al. (2015), Figà-Talamanca and Patacca (2019) it is shown that Bitcoin price and volatility are affected by the volume or number of transactions, by the number of Google searches on the topic, and by Wikipedia inquires on Bitcoin. Alternatively, in Bukovina and Martiček (2016) Bitcoin price is related to a sentiment measure obtained from the Web site[1] `Sentdex.com`.

It is also worth noticing that a market for derivatives on Bitcoin has recently raised on appropriate Web sites such as https://coinut.com and https://deribit.com trading European Calls and Puts and others are likely to come soon according to The Wall Street Journal (2017b), Binary options are also traded in some platforms endorsing the idea in Yermack (2015) that Bitcoins are traded for speculative purposes. Besides and more importantly, the Chicago Board Options Exchange (CBOE) has launched standardized Future contracts on the cryptocurrency in December 2017; this has given rise to a new era for Bitcoin trades and probably opened the way to other standardized derivatives, see (The Wall Street Journal 2017a). Motivated by the evidences in the above quoted papers and the increasing interest in Bitcoin derivatives, we propose a bivariate model in continuous time to describe the dynamics of both price and market attention for Bitcoin. As an additional novel feature, we allow for a possible delay between the attention factor and its delivered effect on Bitcoin price. We assume that the price and the attention factor are two observable variables, each described by a continuous time diffusion process; in addition, we assume that the attention factor influences directly the drift and the diffusion functions describing the price. In the

[1] This Web site collects data on sentiment through an algorithm, based on Natural Language Processing techniques, which is capable of identifying string of words conveying positive, neutral or negative sentiment on a topic (Bitcoin in this case).

numerical exercise, we use some proxies for the attention factor, such as the trading volume and the volume of internet searches. An alternative specification which takes into account market attention is suggested in Kou (2002) where the author assumes that the dynamics of the stock price is a jump diffusion and sudden jumps in returns may be induced by outside news. However, the jump process is not directly related to some attention factor or news indicator.

In this paper, we give several contributions: first, the market model is defined and it is proven that a strong solution exists; then, mild conditions are given for the model to be arbitrage-free and its statistical properties are investigated. Further, the likelihood for a discrete sample of the model is computed and an approximated closed formula is derived, so that maximum likelihood estimates are obtained for model parameters. Precisely, we apply the profile likelihood method described in Davison (2003), Pawitan (2001) to fit the model to Bitcoin market data from January 15, 2015 to March 31, 2017; market attention is measured either through the traditional *trading volume*, see (Barber and Odean 2007; Gervais et al. 2001; Hou et al. 2009), or by the *Google SVI search volume index*, as suggested in Da et al. (2011). Finally, based on risk-neutral evaluation, a quasi-closed formula is derived for any European-style derivative on the Bitcoin, which makes it possible to price derivatives when their values are not known or to calibrate model parameters when derivative prices are given by market trades or quotations.

The paper is organized as follows. In Sect. 2 we describe the model for the Bitcoin price dynamics, discuss its statistical properties and show that the market is arbitrage-free. In Sect. 3, we compute the joint distribution of the discretely sampled model as well as a closedform approximation which is useful to introduce a parameter estimation procedure. Section 4 is devoted to test possible proxies for the attention factor and to estimate model parameters on historical market data obtained from http://blockchain.info. In Sect. 5, we prove a quasi-closed formula for European-style derivatives with detailed computations for Plain Vanilla and Binary option prices, while in Sect. 6.2 we give some further insights on numerical applications of the pricing formula. Some concluding remarks are given in Sect. 7 as well as directions for interesting future investigations. For the sake of readability, side results as well as most technical proofs are collected in Appendices.

2 The Bitcoin market model

We fix a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ endowed with a filtration $\mathbb{F} = \{\mathcal{F}_t,\ t \geq 0\}$ that satisfies the usual conditions of right-continuity and completeness. On the given probability space, we consider a main market in which heterogeneous agents buy or sell Bitcoins and denote by $S = \{S_t,\ t \geq 0\}$ the price process of the cryptocurrency. Following an idea suggested in Cretarola et al. (2018), we assume that the Bitcoin price dynamics is described by the following equation:

$$\mathrm{d}S_t = \mu_S P_{t-\tau} S_t \mathrm{d}t + \sigma_S \sqrt{P_{t-\tau}} S_t \mathrm{d}W_t, \quad S_0 = s_0 \in \mathbb{R}_+, \tag{2.1}$$

where $\mu_S \in \mathbb{R}\backslash\{0\}, \sigma_S \in \mathbb{R}_+,\ \tau \in \mathbb{R}_+$ represent model parameters; $W = \{W_t,\ t \geq 0\}$ is a standard $\mathbb{F}$-Brownian motion on $(\Omega, \mathcal{F}, \mathbf{P})$ and $P = \{P_t,\ t \geq 0\}$ is a stochastic factor, representing the attention index in the Bitcoin market, and satisfying

$$\mathrm{d}P_t = \mu_P P_t \mathrm{d}t + \sigma_P P_t \mathrm{d}Z_t, \quad P_t = \phi(t),\ t \in [-L, 0]. \tag{2.2}$$

Here, $\mu_P \in \mathbb{R}\backslash\{0\}$, $\sigma_P \in \mathbb{R}_+$, $L \in \mathbb{R}_+$are constant parameters, $Z = \{Z_t,\ t \geq 0\}$ is a standard $\mathbb{F}$-Brownian motion on $(\Omega, \mathcal{F}, \mathbf{P})$, which is $\mathbf{P}$-independent of W, and $\phi : [-L, 0] \to [0, +\infty)$ is a continuous (deterministic) initial function. Note that, the nonnegative property of the function ϕ corresponds to requiring that the minimum level for attention is zero.

We assume that the reference filtration $\mathbb{F} = \{\mathcal{F}_t,\ t \geq 0\}$, describing the information on the Bitcoin market, is of the form

$$\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^Z, \quad t \geq 0,$$

where $\mathcal{F}_t^W$ and $\mathcal{F}_t^Z$ denote the σ-algebras generated by W and Z, respectively, up to time $t \geq 0$. Note that $\mathcal{F}_t^Z = \mathcal{F}_t^P$, for each $t \geq 0$, with $\mathcal{F}_t^P$ being the σ-algebra generated by P up to time $t \geq 0$. Since at any time t the Bitcoin price dynamics is affected by the attention index only up to time $t - \tau$, we consider the filtration $\widetilde{\mathbb{F}} = \{\widetilde{\mathcal{F}}_t,\ t \geq 0\}$, defined by

$$\widetilde{\mathcal{F}}_t = \mathcal{F}_t^W \vee \mathcal{F}_{t-\tau}^P, \quad t \geq 0.$$

to describe the traders information on the digital market. This assumption plays an important role in the derivation of the pricing formula in Sect. 5. We also remark that all filtrations satisfy the usual conditions of completeness and right-continuity (see e.g., Protter 2005).

In (2.1), the attention factor P affects explicitly both the drift and the diffusion of the Bitcoin price S_t, up to a certain preceding time $t - \tau$. We further assume that $\tau < L$ and that the process P is observed within the period $[-L, 0]$, to make the bivariate model jointly feasible. It is worth noticing that the instantaneous variance of the Bitcoin price process increases with the delayed process P; this may appear counter-intuitive if P is interpreted only as a *positive attention* indicator. However, in our perspective, the factor P is mathematically a nonnegative process but does not necessary represent a *positive attention*. Possible proxies for P are the volume or the number of daily transactions as well as the number of internet searches.

It is well known that the solution of (2.2) is available in closedform and that P_t has a log-normal distribution for each $t > 0$, see (Black and Scholes 1973). Further, we also prove that the system given by equations (2.1) and (2.2) admits a unique strong solution in $\mathbb{R}_+$ which is given in explicit form, see Theorem B.2 in "Appendix B".

In order to visualize the dynamics implied by the model in equations (2.1) and (2.2), we plot in Fig. 1 a possible simulated path of daily observations for the attention factor P and the corresponding Bitcoin prices S within one year horizon by letting τ vary; as expected, market reaction to attention is delayed when τ increases.

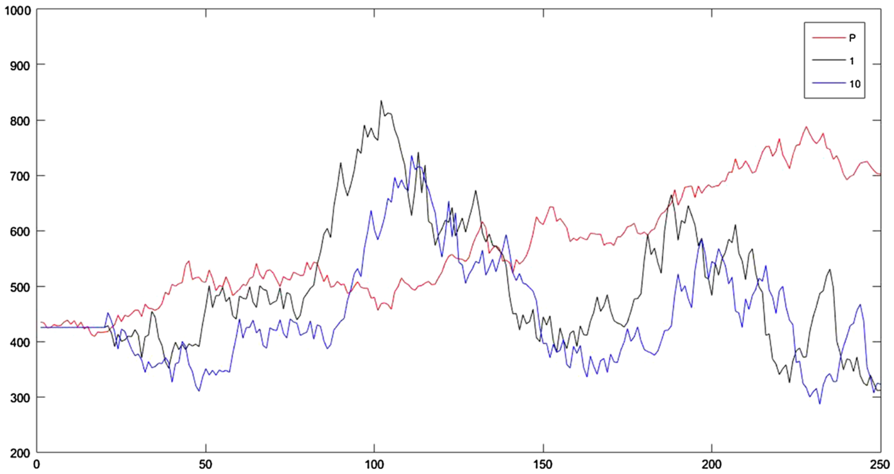

Fig. 1 An example of Bitcoin price dynamics given the evolution of the attention index (red): $\tau = 1$ day (black), $\tau = 10$ days (blue). Model parameters are set to $\mu_P = 0.03, \sigma_P = 0.35, \mu_S = 10^{-5}, \sigma_S = 0.04$

Let us introduce a key definition for the rest of the paper, the *integrated attention process* $X^\tau = \{X_t^\tau,\ t \geq 0\}$, associated to the factor P, given by:

$$X_t^\tau := \begin{cases} \int_0^t P_{u-\tau} \mathrm{d}u = \int_{-\tau}^0 \phi(u)\mathrm{d}u + \int_0^{t-\tau} P_u \mathrm{d}u = X_\tau^\tau + \int_0^{t-\tau} P_u \mathrm{d}u, & 0 \leq \tau \leq t, \\ \int_{-\tau}^{t-\tau} \phi(u)\mathrm{d}u, & 0 \leq t \leq \tau. \end{cases} \tag{2.3}$$

Note that, for $t \in [0, \tau]$, we have $X_t^\tau = \int_{-\tau}^{t-\tau} \phi(u)\mathrm{d}u$ which is deterministic. In addition, for a finite time horizon $T > 0$, the corresponding change over the interval $[t, T]$, for $t \leq T$, is defined as $X_{t,T}^\tau := X_T^\tau - X_t^\tau$. Obviously, $X_{T,T}^\tau = 0$; moreover, for $t < T$,

$$X_{t,T}^\tau := \begin{cases} \int_{t-\tau}^{T-\tau} P_u \mathrm{d}u & \text{if } 0 \leq \tau \leq t < T, \\ \int_{t-\tau}^{0} \phi(u)\mathrm{d}u + \int_0^{T-\tau} P_u \mathrm{d}u & \text{if } 0 \leq t \leq \tau < T, \\ \int_{t-\tau}^{T-\tau} \phi(u)\mathrm{d}u & \text{if } 0 \leq t < T \leq \tau. \end{cases}$$

Note that for $T \leq \tau$, we get $X_{t,T}^\tau = \int_{t-\tau}^{T-\tau} \phi(u)\mathrm{d}u$ which is deterministic. Basic statistical properties for the integrated attention process and for its changes are given in Lemma B.1 in "Appendix B".

3 Statistical properties of discretely observed quantities and parameter estimation

In this section, we derive statistical properties for a sample of discretely observed prices and suggest a possible closedform approximation for the joint probability density of the discrete sample. Let us fix a discrete observation step Δ and consider the discrete time process $\{S_i,\ i \in \mathbb{N}\}$, where $S_i := S_{i\Delta}$. Define the corresponding logarithmic returns process $\{R_i,\ i \in \mathbb{N}\}$ as

$$R_i = \log(S_i) - \log(S_{i-1}),$$

where

$$\log(S_i) = \log(S_0) + \left(\mu_S - \frac{\sigma_S^2}{2}\right) \int_0^{i\Delta} P_{u-\tau} du + \sigma_S \int_{(0}^{i\Delta} \sqrt{P_{u-\tau}} \mathrm{d}W_u,$$

see Theorem B.2 in "Appendix B". Define

$$M_i := \int_{(i-1)\Delta}^{i\Delta} \sqrt{P_{u-\tau}} \mathrm{d}W_u, \quad \forall i \in \mathbb{N},$$

so that we can write,

$$R_i = \left(\mu_S - \frac{\sigma_S^2}{2}\right) A_i^\tau + \sigma_S M_i, \quad i \in \mathbb{N},$$

where $A_i^\tau := X^\tau_{(i-1)\Delta,\, i\Delta}$, with $X^\tau_{t,T}$ being the variation of the integrated attention process introduced in (2.3); since τ is fixed we omit hereafter the dependence on it and, without loss of generality we assume $\tau < \Delta$ so that $A_1 = X_\tau^\tau + \int_0^{\Delta-\tau} P_u \mathrm{d}u$.

Note that if $j\Delta \le \tau < (j+1)\Delta$ the quantities $A_1, \ldots, A_j$ are deterministic and the outcomes in what follows still hold if A_1 is replaced by the first non-deterministic value A_{j+1}.

Let us consider a finite time horizon $T = n\Delta$; under model assumptions the conditional probability distribution of the vector $\mathbf{M} = (M_1, M_2, \ldots, M_n)$, given the vector $\mathbf{A} = (A_1, A_2, \ldots, A_n)$, is a multivariate normal with covariance matrix $Diag(\mathbf{A})$.

Hence, the vector of discretely observed logarithmic returns $\mathbf{R} = (R_1, R_2, \ldots, R_n)$, conditionally on $\mathbf{A}$, is jointly normal with mean $\left(\mu_S - \frac{\sigma_S^2}{2}\right)\mathbf{A}$ and covariance matrix $\Sigma = \sigma_S^2 Diag(\mathbf{A})$ where we have omitted the superscript τ for the ease of reading.

The application of Bayes' rule allows to write the unconditional joint probability distribution of $(\mathbf{R}, \mathbf{A})$, i.e., the density function $f_{(\mathbf{R},\mathbf{A})} : \mathbb{R}^n \times \mathbb{R}_+^n \times N^n \longrightarrow \mathbb{R}$ as

$$f_{(\mathbf{R},\mathbf{A})}(\mathbf{r}, \mathbf{a}) = f_{A_1}(a_1) \prod_{i=2}^{n} f_{(A_i|A_{i-1})}(a_i) \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_S^2 a_i}} e^{-\frac{1}{2} \frac{\left(r_i - \left(\mu_S - \frac{\sigma_S^2}{2}\right) a_i\right)^2}{\sigma_S^2 a_i}}. \quad (3.1)$$

with $\mathbf{r} = (r_1, r_2, \ldots, r_n) \in \mathbb{R}^n$ and $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}_+^n$.

The probability distribution functions $f_{A_1}(.)$ and $f_{A_i|A_{i-1}}(.)$ for $i = 2, 3, \ldots, n$ are not available in closedform; though, several approximations exist among which those introduced in Levy (1992) and Milevsky and Posner (1998). Of course, any approximation available for such densities can be applied in order to find a closed formula approximating the joint density $f_{(\mathbf{R},\mathbf{A})}(\mathbf{r},\mathbf{a})$; in what follows we adopt the one suggested in Levy (1992), see "Appendix A" for further details. Note that the inverse gamma approach suggested in Milevsky and Posner (1998) holds in the limit when T tends to infinity, a condition which is not at all consistent with the applications we have in mind; further discussion on the approximating distribution is beyond the scope of our paper.

3.1 The approximated likelihood

One of the pillar in statistical inference is the maximum likelihood (in short ML) estimation approach where model parameters are estimated so as to maximize the probability of the realized sample to be extracted randomly; the likelihood function shares the same mathematical expression of the probability density function but it is computed "ex-post" when a realization of involved random variables is available and assuming the underlying model parameters to be unknown. It is well known that ML estimates are consistent and asymptotically normal and they achieve efficiency, i.e., they have the lowest variance among estimators sharing the same asymptotic properties (see Davison 2003).

By applying the approximation of Levy (1992), we prove the following Lemma.

Lemma 3.1 *Let $\phi(t) > 0$, for each $t \in [-L, 0]$, in (2.2) and $\tau < \Delta$. Then, in the market model outlined in Sect. 2, we have*

(i) *the distribution of $A_1 - X_\tau^\tau$ is approximated by a log-normal with mean α_1 and variance ν_1^2 given by*

$$\begin{aligned}\alpha_1 =& \log\phi(0) + 2\log\frac{e^{\mu_P(\Delta-\tau)}-1}{\mu_P}\\ &-\frac{1}{2}\log\left(\frac{2}{\mu_P+\sigma_P^2}\left[\frac{e^{(2\mu_P+\sigma_P^2)(\Delta-\tau)}-1}{2\mu_P+\sigma_P^2}-\frac{e^{\mu_P(\Delta-\tau)}-1}{\mu_P}\right]\right)\\ \nu_1^2 =& \log\left(\frac{2}{\mu_P+\sigma_P^2}\left[\frac{e^{(2\mu_P+\sigma_P^2)(\Delta-\tau)}-1}{2\mu_P+\sigma_P^2}-\frac{e^{\mu_P(\Delta-\tau)}-1}{\mu_P}\right]\right)\\ &-2\log\left(\frac{e^{\mu_P(\Delta-\tau)}-1}{\mu_P}\right)\end{aligned}$$

(ii) *the distribution of A_i given A_{i-1} (shortly $A_i|A_{i-1}$), for $i = 1, \dots, n$, is approximated by a log-normal with means α_i and variances ν_i^2 given by*

$$\begin{aligned}\alpha_i &= \log(A_{i-1}) + \left(\mu_P - \frac{\sigma_P^2}{2}\right)\Delta, \quad \text{for } i = 1, \dots, n,\\ \nu_i^2 &= \sigma_P^2\Delta, \quad \text{for } i = 1, \dots, n.\end{aligned}$$

The proof is postponed to "Appendix B".

Now, we are in the position to state the following theorem.

Theorem 3.2 *Under the same assumptions of Lemma 3.1, given the realized sample* $(\bar{\mathbf{r}}, \bar{\mathbf{a}})$*, the log-likelihood function* $\log \mathcal{L}_{\mathbf{R},\mathbf{A}}(\mu_P, \mu_S, \sigma_P, \sigma_S) : \mathbb{R}^2 \times \mathbb{R}_+^2 \longrightarrow \mathbb{R}$ *can be approximated by*

$$\log \mathcal{L}_{\mathbf{R},\mathbf{A}}(\mu_P, \mu_S, \sigma_P, \sigma_S) = \sum_{i=1}^{n} \left[\log \left(\frac{1}{\sqrt{2\pi\sigma_S^2 a_i}} \right) - \frac{1}{2} \frac{\left(r_i - \left(\mu_S - \frac{\sigma_S^2}{2} \right) a_i \right)^2}{\sigma_S^2 a_i} \right] + \sum_{i=1}^{n} \left[\log \left(\frac{1}{a_i \nu_i \sqrt{2\pi}} \right) - \frac{(\log(a_i) - \alpha_i)^2}{2\nu_i^2} \right], \tag{3.2}$$

where upper case letters are used for random variables and lowercase for the corresponding realizations.

Proof First, recall that the likelihood function of a parameter corresponds to the probability density function where random variables are replaced by they realizations and parameters are unknown. Then, by simply applying the logarithmic function to (3.1) we get

$$\log f_{(\mathbf{R},\mathbf{A})}(\mathbf{r}, \mathbf{a}) = -\frac{1}{2} \sum_{i=1}^{n} \left[\log \left(2\pi\sigma_S^2 a_i \right) + \frac{\left(r_i - \left(\mu_S - \frac{\sigma_S^2}{2} \right) a_i \right)^2}{\sigma_S^2 a_i} \right] + \log \left(f_{A_1}(a_1) \right) + \sum_{i=2}^{n} \log \left(f_{A_i|A_{i-1}}(a_i) \right). \tag{3.3}$$

Replacing the unknown densities in (3.3) according to Lemma 3.1 gives the desired result. □

Maximum likelihood estimates for the model can be obtained by maximizing the log-likelihood approximation in (3.2), i.e.,

$$(\widehat{\mu}_P, \widehat{\mu}_S, \widehat{\sigma}_P, \widehat{\sigma}_S) = \arg \max_{\substack{\mu_P, \mu_S \\ \sigma_P, \sigma_S}} \log \mathcal{L}_{\mathbf{R},\mathbf{A}}(\mu_P, \mu_S, \sigma_P, \sigma_S).$$

In this case, the methodology is referred to as Quasi-Maximum Likelihood since the exact expression of the likelihood is not available; under suitable conditions, quasi-maximum likelihood estimates are asymptotically equivalent to the Maximum Likelihood estimates, see e.g., (White 1982; Gourieroux et al. 1984). We also performed a simulation study to assess the finite sample behavior of the estimates which is summed up in "Appendix C".

It is worth to stress that the above estimation method does not assume the process P to be observed, as far as X_τ^τ and $A_i, i \geq 1$ are known (note that A_i is the cumulative of P along the time interval $[(i-1)\Delta - \tau, i\Delta - \tau]$).

3.2 Estimation of the delay parameter

The delay parameter τ directly affects the definition of the discrete process A_i. Hence, in order to proceed with its estimation we need to observe the process P at a finer observation step δ with respect to the log-returns. In what follows we set $\Delta = \delta r$ and we adopt a two step estimation procedure known as Profile Likelihood in order to estimate the delay. We briefly describe the Profile Likelihood approach to estimation and its application in our specific case; interested readers are referred to Davison (2003), Pawitan (2001) for details on the profile likelihood. The basic idea of this approach is to split the parameter vector which has to be estimated, say θ, in two sub-vectors, one representing the parameter of interest and the other the so-called nuisance parameter, i.e., $\theta = (\beta, \lambda)$; to estimate β and λ jointly we should maximize at once the likelihood, i.e.,

$$\max_{\beta,\lambda} \log \mathcal{L}(\beta, \lambda).$$

When this is not feasible and provided the likelihood computed with respect to the nuisance parameter vector λ is available and easy to maximize, we can apply a two step procedure by maximizing, for each β in its parametric space,

$$\mathcal{L}_p(\beta) = \max_{\lambda} \mathcal{L}(\beta, \lambda) = \mathcal{L}(\beta, \widehat{\lambda}_\beta),$$

where p is the length of parameter β and $\widehat{\lambda}_\beta$ is the maximum likelihood estimate of λ for fixed β. Then, the best estimate for β is obtained as

$$\widehat{\beta} = \arg\max_{\beta} \log \mathcal{L}_p(\beta).$$

Classical confidence intervals cannot be defined in this setting; indeed, it is possible to obtain a confidence region for β using the likelihood ratio statistics (see Davison 2003), defined as

$$W_p(\beta_0) = 2\left\{\mathcal{L}(\widehat{\beta}, \widehat{\lambda}) - \mathcal{L}(\beta_0, \widehat{\lambda}_{\beta_0})\right\},$$

where

$$W_p(\beta_0) \xrightarrow{D} \chi_p^2$$

and β_0 is an assigned value for β. These results imply that the confidence region for β is the set

$$\left\{\beta : \mathcal{L}_p(\beta) \geq \mathcal{L}_p(\widehat{\beta}) - \frac{1}{2}c_p(1-2\alpha)\right\}, \qquad (3.4)$$

with $c_p(\alpha)$ is the α-quantile of the χ_p^2 distribution.

In our exercise we split $\theta := (\mu_P, \mu_S, \sigma_P, \sigma_S, \tau)$ in $\theta = (\tau, \lambda)$ where τ is the parameter on which we are focusing and $\lambda = (\mu_P, \mu_S, \sigma_P, \sigma_S)$ is the nuisance parameter vector. The Profile Likelihood approach is feasible in our case since a closed approximating expression for the likelihood with respect to the nuisance parameter is indeed available. The parametric space for $\beta := \tau$ is the interval $[0, L]$ in this case but, for practical purposes, τ is chosen on a grid, i.e., $\tau \in \{\tau_0, \tau_1, \tau_2, \dots, \tau_k\}$; the maximization of the likelihood $\log \mathcal{L}_{\mathbf{R},\mathbf{A}}(\mathbf{r}, \mathbf{a})$ is then performed with respect to λ for each value τ_j in the grid, obtaining $\mathcal{L}_p(\tau_j)$ for $j = 0, 1, \dots, k$. An estimate for τ is then obtained as $\widehat{\tau} = \arg\max_j \mathcal{L}_p(\tau_j)$. Finally we get $\widehat{\theta} = \left(\widehat{\tau}, \widehat{\lambda}_{\widehat{\tau}}\right)$. Of course, the estimation error decreases with the mesh of the grid so that it sufficiently spans the parametric set for τ.

4 Model fitting on historical data

In order to fit the model described in (2.1)–(2.2) on real data we need samples for both the Bitcoin price and the attention indicator. Recently, in Da et al. (2011) the authors suggest a new and direct measure of investor attention using cumulative internet search frequency. This measure is particularly consistent with our approach since Bitcoin is an internet-based digital currency and internet users commonly collect information through a search engine such as Google[2]. Besides, "the search volume is likely to be representative of the internet search behavior of the general population and more critically, search is a revealed attention measure: if you search for a stock in Google, you are undoubtedly paying attention to it. Therefore, aggregate search frequency in Google is a direct and unambiguous measure of attention", quoting (Da et al. 2011). The authors in Da et al. (2011) also find strong evidence that search volume index (SVI), provided by Google, captures the attention of individual/retail investors, i.e., of non informed investors, that we named as followers in the Introduction.

Further, the number and volume of transactions, which are examples of traditional measures for investor attention, as well as the SVI volume of Google searches or Wikipedia requests have been proven to affect Bitcoin prices in Kristoufek (2015), Kristoufek (2013).

Hence, once consistency is investigated, for the above measures of attention, with the dynamics assumed in (2.2), we fit the model to both the attention factor and Bitcoin prices data. Daily observations for Bitcoin prices, volume and number of transactions are obtained through the Web site http://blockchain.info which provides an average price among main exchanges trading on Bitcoin and the total exchanged volume. Daily data for Wikipedia requests on the term "bitcoin" are obtained through the Web site http://tools.wmflabs.org/pageviews. Google computes the search volume index (SVI) for a search term as the number of searches for that term scaled by its time series average and make this measure available by the Web site Google-trends http://www.google.com/trends from which we obtained weekly data for the number of Google searches on the term "bitcoin" (daily data are not available at the time of writing).

[2] It is well documented that Google is the most popular search engine.

4.1 Proxies for attention

The univariate process P is a geometric Brownian motion; it is well known that the corresponding discrete process of logarithmic changes, with time step δ, is a sequence of independent and identically distributed normal random variables with mean $\left(\mu_P - \frac{\sigma_P^2}{2}\right)\delta$ and variance $\sigma_P^2\delta$. Simple tests, such as the augmented Dickey–Fuller test, see (Tsay 2005), and the one-sample Kolmogorov-Smirnov test, see (Massey 1951), are applied in what follows to investigate such property on candidate proxies for the process P.

We consider the number and the volume of transactions as possible traditional attention measures and the number of Google searches and Wikipedia requests as examples of web-based attention indicators. The first three series were investigated from 01/01/2015 to 31/03/2017 while Wikipedia requests are considered from 01/07/2015 to 31/03/2017[3]

Non-stationarity is rejected for all proxies, whereas log-normality is not rejected only for the trading volume and the SVI index. For this reason, we used the these two proxies for our empirical estimation.

4.2 Estimation results

According to the outcomes in previous subsection, we consider the daily time series of the volume of transactions and the weekly time series of the Google searches from 01/01/2015 to 31/03/2017 as suitable proxies for process P. Hence, we fit the model described in (2.1) and (2.2) by applying the procedure described in Sect. 3, that we may briefly refer as *Profile Quasi-Maximum Likelihood* (PQML). Bitcoin prices are considered from 15/01/2015 in order to account for a maximum of two weeks for the time delay with respect to attention measures. Note that Google-trends provides a scaled time series for the number of searches so the maximum value is 100; in order to compare outcomes we do the same for the trading volume time series. In what follows we assume that $\Delta = 1$ week is the observation step for Bitcoin log-returns.

4.2.1 Attention measured by volume

Given daily observations $\{P_i\}_i$ of the volume of transactions we are able to compute the cumulative weekly attention $\{A_i\}_i$; for $\tau = 0$ A_i is simply the mean volume during the preceding week, i.e.,

$$A_i = \int_{(i-1)\Delta}^{i\Delta} P_u \mathrm{d}u = \sum_{j=1}^{7} \int_{(i-1)\Delta+(j-1)\delta}^{(i-1)\Delta+j\delta} P_u \mathrm{d}u = \sum_{j=1}^{7} P_{(i-1)\Delta+j\delta}\delta$$
$$= \frac{\sum_{j=1}^{7} P_{(i-1)\Delta+j\delta}}{7}\Delta.$$

[3] Data available only from 01/07/2015.

Table 1 Parameter fit with $\tau = 5$

Variable	Fitted value	SE	t value	$P(>\|t\|)$
μ_P	1.0404	0.7373	1.4110	0.1610
σ_P	1.1092	0.0725	15.2924	0.0000
μ_S	0.0153	0.0083	1.8434	0.0679
σ_S	0.0830	0.0054	15.2403	0.0000

The generalization to $\tau > 0$ is straightforward as soon as we assume $\tau = r\delta$ for some positive integer r; in which case A_i would be the mean volume of the 7 days preceding time $i\Delta - r\delta$.

By applying the PQML method we obtain $\tau = 5$ days; the estimated value of other parameters is summed up in Table 1.

In order to assess significance of parameters, the usual t-statistics is also computed and shown in the table as well as the p-value of the t test. It is clear from Table 1 that μ_P is not statistically significant and that μ_S is weakly significant. Finally, we evaluate the confidence region for τ using (3.4) and $\tau \in \{0, 1, 2, \dots, 10\}$ days. We find that the confidence region is given by $\mathcal{T} = \{2, 3, 4\}$ days. Estimates of other parameters are indeed very similar for any $\tau \in \mathcal{T}$ and analogous comments apply. We stress that the value 0 is not within the confidence region, meaning that the introduction of the delay parameter is relevant in the model specification.

4.2.2 Attention measured by Google searches

Assume now that Google searches are representative of attention on Bitcoin. Since we have weekly data, we should aggregate both Google searches and Bitcoin returns to a coarser observation step; however, this would reduce the time series length dramatically and corresponding estimates might be unreliable. Hence, we assume that the available observations correspond to the cumulative attention time series A_i and we assume there exists a nonnegative integer c such that $\tau = c\Delta$; in this case, τ is on a weekly scale and not on a daily scale as in the previous case.

By applying the PQML procedure we obtain $\tau = 1$ week; the estimated value of other parameters is displayed in Table 2.

The t- statistics and its p-value are also reported for all parameters, as in Sect. 4.2.1. Again, μ_P is not statistically significant and μ_S is weakly significant. Note that, the confidence region for τ only includes $\tau = 1$ week.

Table 2 Parameter fit with $\tau = 1$

Variable	Fitted value	SE	t value	$P(>\|t\|)$
μ_P	0.9573	0.7315	1.3087	0.1935
σ_P	1.0818	0.0714	15.1611	0.0000
μ_S	0.0181	0.0092	1.9534	0.0535
σ_S	0.0867	0.0057	15.1774	0.0000

5 Risk-neutral evaluation of European-type contingent claims

In this section results concerning arbitrage-free conditions and pricing for European-style derivatives are discussed. Precisely, we derive a closedform pricing formula for European-style derivatives and, once parameter are estimated via the PQML method described above, we compare model prices with real quotations reported on the trading platform www.deribit.com, where standard and binary option on Bitcoin are issued regularly. Though Bitcoin is traded on different exchanges with different prices, the derivative contracts in the platform assume that the underlying price is given by the value of the Bitcoin Index available in www.blockchain.info , computed as a weighted average of prices on major exchanges. Here, we use the same approach and consider a single price for one Bitcoin.

Let us fix a finite time horizon $T > 0$ and assume the existence of a riskless asset, say the money market account, whose value process $B = \{B_t,\ t \in [0, T]\}$ is given by

$$B_t = e^{\int_0^t r(s)\mathrm{d}s}, \quad t \in [0, T],$$

where $r : [0, T] \to \mathbb{R}$ is a bounded, deterministic function representing the instantaneous risk-free interest rate. To exclude arbitrage opportunities, we need to check that the set of all equivalent martingale measures for the Bitcoin price process S is non-empty. More precisely, it will contain more than a single element, since P does not represent the price of any tradable asset, and therefore the underlying market model is incomplete, as shown in Lemma B.3 in "Appendix B".

Simple examples of candidate equivalent martingale measures are obtained under the assumption of a constant price γ of attention risk. Specifically these are defined as probability measures $\mathbf{Q}^\gamma$ with the following density

$$\left.\frac{\mathrm{d}\mathbf{Q}^\gamma}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{F}_T} =: L_T^{\mathbf{Q}^\gamma}, \quad \mathbf{P}-\text{a.s.},$$

where $L_T^{\mathbf{Q}^\gamma}$ is the terminal value of the $(\mathbb{F}, \mathbf{P})$-martingale $L^{\mathbf{Q}^\gamma} = \{L_t^{\mathbf{Q}^\gamma},\ t \in [0, T]\}$ given by

$$L_t^{\mathbf{Q}^\gamma} := \mathcal{E}\left(-\int_0^{\cdot} \frac{\mu_S P_{s-\tau} - r(s)}{\sigma_S\sqrt{P_{s-\tau}}}\mathrm{d}W_s - \int_0^{\cdot} \gamma \mathrm{d}Z_s\right), \quad t \in [0, T],$$

for a constant parameter γ.

The dynamics of the (non-discounted) Bitcoin price with respect to the minimal martingale measure $\mathbf{Q}^\gamma$ is given by

$$\begin{cases} \mathrm{d}S_t = r(t)S_t\mathrm{d}t + \sigma_S\sqrt{P_{t-\tau}}S_t\mathrm{d}\widehat{W}_t, & S_0 = s_0 \in \mathbb{R}_+, \\ \mathrm{d}P_t = (\mu_P - \gamma\,\sigma_P)P_t\mathrm{d}t + \sigma_P P_t\mathrm{d}\widehat{Z}_t, & P_t = \phi(t),\ t \in [-L, 0], \end{cases}$$

where $\widehat{W} = \{\widehat{W}_t,\ t \in [0,T]\}$ and $\widehat{Z} = \{\widehat{Z}_t,\ t \in [0,T]\}$ are $(\mathbb{F}, \mathbf{Q}^\gamma)$-Brownian motions defined, respectively, by

$$\widehat{W}_t := W_t + \int_0^t \frac{\mu_S P_{s-\tau} - r(s)}{\sigma_S \sqrt{P_{s-\tau}}} \mathrm{d}s, \quad t \in [0,T],$$
$$\widehat{Z}_t := Z_t + \gamma\, t, \quad t \in [0,T].$$

A very special case within this family is the so-called *minimal martingale measure* (see e.g., Föllmer and Schweizer 1991, 2010), obtained by setting $\gamma = 0$, and denoted by $\widehat{\mathbf{P}}$. Note that, in this case, the Brownian motion driving the attention factor as well as its dynamics is not affected by the change of measures, i.e., $\widehat{Z}_t := Z_t$. The interested reader may find further details in "Appendix B".

Let us consider in what follows a European-type contingent claim, expiring at time T, traded on the underlying market and we assume that its final payoff is described by a $\widetilde{\mathfrak{F}}_T$-measurable random variable $H = \varphi(S_T)$, with $\varphi : \mathbb{R} \to \mathbb{R}$ being a Borel-measurable function[4] such that H is integrable under $\mathbf{Q}^\gamma$.

Recall that $X^\tau_{t,T} = X^\tau_T - X^\tau_t$, for each $t \in [0,T)$, refers to the variation of the process X^τ defined in (2.3), over the interval $[t,T]$. Then, denote by $\mathbb{E}^{\mathbf{Q}^\gamma}\left[\cdot \middle| \widetilde{\mathfrak{F}}_t\right]$ the conditional expectation with respect to $\widetilde{\mathfrak{F}}_t$ under the probability measure $\mathbf{Q}^\gamma$ and so on.

Theorem 5.1 *Let $H = \varphi(S_T)$ be the payoff a European-type contingent claim with date of maturity T. Then, the risk-neutral price $\Phi_t(H)$ at time t of H is given by*

$$\Phi_t(H) = \mathbb{E}^{\mathbf{Q}^\gamma}\left[\psi(t, S_t, X^\tau_{t,T}) \middle| S_t\right], \quad t \in [0,T),$$

where $\psi : [0,T) \times \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ is a Borel-measurable function such that

$$\psi(t, S_t, X^\tau_{t,T}) = B_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\frac{1}{B_T} G\left(t, S_t, X^\tau_{t,T}, Y_{t,T}\right) \middle| \mathfrak{F}^W_t \vee \mathfrak{F}^P_{T-\tau}\right], \tag{5.1}$$

for a suitable function G depending on the contract such that $G\left(t, S_t, X^\tau_{t,T}, Y_{t,T}\right)$ is $\mathbf{Q}^\gamma$-integrable.

Proof For the sake of simplicity suppose that $\tau < T$ and set $Y_{t,T} := \int_t^T \sqrt{P_{u-\tau}}\,\mathrm{d}\widehat{W}_u$, for each $t \in [0,T)$. Then, the risk-neutral price $\Phi_t(H)$ at time t of a European-type contingent claim with payoff $H = \varphi(S_T)$ is given by

[4] The function φ is usually referred to as the *contract function*.

$$\Phi_t(H) = B_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\frac{\varphi(S_T)}{B_T}\middle|\widetilde{\mathfrak{F}}_t\right]$$
$$= B_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathbb{E}^{\mathbf{Q}^\gamma}\left[\frac{\varphi\left(S_t e^{\int_t^T r(u)\mathrm{d}u - \frac{\sigma_S^2}{2}X^\tau_{t,T} + \sigma_S Y_{t,T}}\right)}{B_T}\middle|\mathfrak{F}_t^W \vee \mathfrak{F}_{T-\tau}^P\right]\middle|\widetilde{\mathfrak{F}}_t\right], \tag{5.2}$$

where $\mathbb{E}^{\mathbf{Q}^\gamma}\left[\cdot\middle|\widetilde{\mathfrak{F}}_t\right]$ denotes the conditional expectation with respect to $\widetilde{\mathfrak{F}}_t$ under the equivalent martingale measure $\mathbf{Q}^\gamma$. More generally, (5.2) can be written as

$$\Phi_t(H) = B_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathbb{E}^{\mathbf{Q}^\gamma}\left[\frac{G(t, S_t, X^\tau_{t,T}, Y_{t,T})}{B_T}\middle|\mathfrak{F}_t^W \vee \mathfrak{F}_{T-\tau}^P\right]\middle|\widetilde{\mathfrak{F}}_t\right], \tag{5.3}$$

for a suitable function G depending on the contract function φ. We can apply the same arguments used in point (ii) of the proof of Theorem B.2, to get that, for each $t \in [0, T)$, the random variable $Y_{t,T}$ conditioned on $\mathfrak{F}_{T-\tau}^P$ is Normally distributed with mean 0 and variance $X^\tau_{t,T}$. Then, we can write (in law) that $Y_{t,T} = \sqrt{X^\tau_{t,T}}\epsilon$, where ϵ is a standard Normal random variable and this allows to find a function ψ such that (5.1) holds, which means that the conditional expectation with respect to $\mathfrak{F}_t^W \vee \mathfrak{F}_{T-\tau}^P$ in (5.3) only depends on S_t and $X^\tau_{t,T}$, for every $t \in [0, T)$. Consequently, the risk-neutral price $\Phi_t(H)$ can be written as

$$\Phi_t(H) = \mathbb{E}^{\mathbf{Q}^\gamma}\left[\psi(t, S_t, X^\tau_{t,T})\middle|\widetilde{\mathfrak{F}}_t\right] = \mathbb{E}^{\mathbf{Q}^\gamma}\left[\psi(t, S_t, X^\tau_{t,T})\middle|S_t\right], \tag{5.4}$$

where the last equality holds since S is $\widetilde{\mathbb{F}}$-adapted and $X^\tau_{t,T}$ is independent of $\widetilde{\mathfrak{F}}_t$, for each $t \in [0, T)$, see e.g., (Pascucci 2011, Lemma A.108). More precisely, we have

$$\mathbb{E}^{\mathbf{Q}^\gamma}\left[\psi(t, S_t, X^\tau_{t,T})\middle|\widetilde{\mathfrak{F}}_t\right] = \mathbb{E}^{\mathbf{Q}^\gamma}\left[\psi(t, S_t, X^\tau_{t,T})\middle|S_t\right] = g(S_t),$$

where

$$g(s) = \mathbb{E}^{\mathbf{Q}^\gamma}\left[\psi(t, s, X^\tau_{t,T})\middle|S_t = s\right], \quad s \in \mathbb{R}_+.$$

□

Since the martingale measure is fixed, the risk-neutral price obtained above agrees with the arbitrage-free price for those payoffs which can be replicated by investing on the underlying market. Forward prices at settlement can also be obtained, by imposing an initial zero cash-flow between counterparts in the pricing formula.

Remark 5.2 It is worth to remark that $\psi(t, S_t, x)$, with $x \in \mathbb{R}_+$, represents the risk-neutral price at time $t \in [0, T)$ of the contract $H = \varphi(S_T)$ in a Black & Scholes framework, where the constant volatility parameter σ^{BS} is defined by

$$\sigma^{BS} := \sigma_S \sqrt{\frac{x}{T-t}}.$$

This is proved explicitly in Corollary 5.4 for the special case of a *Plain Vanilla* European Call option.

Remark 5.3 A pricing formula analogous to (5.4) is conjectured in Hull and White (1987) for a special example of the model suggested here (corresponding to $\tau = 0$ and $\sigma_S = 1$). However, the authors assumed from the very beginning to be within a risk-neutral setting, without defining the dynamics under the physical measure and without the need for a proof of the existence of any equivalent martingale measure. Theorem 5.1 extends their results to the more general case and gives a rigorous proof.

5.1 A Black & Scholes-type option pricing formula

We consider here the special case of a European Call option with strike price K and maturity T. Define the function $C^{BS} : [0, T) \times \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ as follows

$$C^{BS}(t, s, x) := s\mathcal{N}(d_1(t, s, x)) - Ke^{-\int_0^t r(u)\mathrm{d}u}\mathcal{N}(d_2(t, s, x)), \tag{5.5}$$

where

$$d_1(t, s, x) = \frac{\log\left(\frac{s}{K}\right) + \int_0^t r(u)\mathrm{d}u + \frac{\sigma_S^2}{2}x}{\sigma_S\sqrt{x}}$$

and $d_2(t, s, x) = d_1(t, s, x) - \sigma_S\sqrt{x}$, or more explicitly

$$d_2(t, s, x) = \frac{\log\left(\frac{s}{K}\right) + \int_0^t r(u)\mathrm{d}u - \frac{\sigma_S^2}{2}x}{\sigma_S\sqrt{x}}.$$

Here, $\mathcal{N}$ stands for the standard Gaussian cumulative distribution function

$$\mathcal{N}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{z^2}{2}}\mathrm{d}z, \quad \forall\, y \in \mathbb{R}.$$

Corollary 5.4 *The risk-neutral price C_t at time t of a European Call option written on the Bitcoin with price S expiring in T and with strike price K is given by the formula*

$$C_t = \mathbb{E}^{\mathbf{Q}^\gamma}\left[C^{BS}(t, S_t, X_{t,T}^\tau)\middle| S_t\right], \quad t \in [0, T), \tag{5.6}$$

where the function C^{BS} is defined in (5.5).

Proof As in the proof of Theorem 5.1, let us assume that $\tau < T$. Under the martingale measure $\mathbf{Q}^\gamma$, the risk-neutral price C_t at time $t \in [0, T)$ of a European Call option written on the Bitcoin with price S expiring in T and with strike price K, is given by

$$\begin{aligned} C_t &= B_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\frac{\max(S_T - K, 0)}{B_T}\middle|\widetilde{\mathcal{F}}_t\right] \\ &= B_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\widetilde{S}_T \mathbf{1}_{\{S_T > K\}}\middle|\widetilde{\mathcal{F}}_t\right] - K e^{-\int_t^T r(u)\mathrm{d}u} \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathbf{1}_{\{S_T > K\}}\middle|\widetilde{\mathcal{F}}_t\right] \\ &= B_t J_1 - K e^{-\int_t^T r(u)\mathrm{d}u} J_2, \end{aligned}$$

where we have set $J_1 := \mathbb{E}^{\mathbf{Q}^\gamma}\left[\widetilde{S}_T \mathbf{1}_{\{S_T > K\}}\middle|\widetilde{\mathcal{F}}_t\right]$ and $J_2 := \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathbf{1}_{\{S_T > K\}}\middle|\widetilde{\mathcal{F}}_t\right]$. Setting $Y_{t,T} := \int_t^T \sqrt{P_{u-\tau}}\mathrm{d}\widehat{W}_u$, for every $t \in [0, T)$, the term J_2 can be written as

$$\begin{aligned} J_2 &= \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathbf{1}_{\{S_T > K\}}|\mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right]\middle|\widetilde{\mathcal{F}}_t\right] \\ &= \mathbb{E}^{\mathbf{Q}^\gamma}\left[\widehat{\mathbf{P}}\left(S_t e^{\int_t^T r(u)\mathrm{d}u - \frac{\sigma_S^2}{2} X_{t,T}^\tau + \sigma_S Y_{t,T}} > K \middle| \mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right)\middle|\widetilde{\mathcal{F}}_t\right] \\ &= \mathbb{E}^{\mathbf{Q}^\gamma}\left[\widehat{\mathbf{P}}\left(-\frac{Y_{t,T}}{\sqrt{X_{t,T}^\tau}} < \frac{\log\left(\frac{S_t}{K}\right) + \int_t^T r(u)\mathrm{d}u - \frac{\sigma_S^2}{2} X_{t,T}^\tau}{\sigma_S \sqrt{X_{t,T}^\tau}} \middle| \mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right)\middle|\widetilde{\mathcal{F}}_t\right] \\ &= \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathcal{N}\left(d_2(t, S_t, X_{t,T}^\tau)\right)\middle|\widetilde{\mathcal{F}}_t\right], \end{aligned}$$

as for each $t \in [0, T)$, the random variable $-\frac{Y_{t,T}}{\sqrt{X_{t,T}^\tau}}$, given $\mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P$, has a standard Gaussian law $\mathcal{N}(0, 1)$ under $\mathbf{Q}^\gamma$. Concerning J_1, let us introduce the auxiliary probability measure $\bar{\mathbf{Q}}$ on $(\Omega, \mathcal{F}_T)$ defined as follows:

$$\frac{\mathrm{d}\bar{\mathbf{Q}}}{\mathrm{d}\mathbf{Q}^\gamma} := e^{-\frac{\sigma_S^2}{2}\int_0^T P_{u-\tau}\mathrm{d}u + \sigma_S \int_0^T \sqrt{P_{u-\tau}}\mathrm{d}\widehat{W}_u}, \quad \widehat{\mathbf{P}} - \text{a.s.}.$$

By Girsanov's Theorem, we get that the process $\bar{W} = \{\bar{W}_t,\ t \in [0, T]\}$, given by

$$\bar{W}_t := \widehat{W}_t - \sigma_S \int_0^t \sqrt{P_{u-\tau}}\mathrm{d}u, \quad t \in [0, T],$$

follows a standard $(\mathbb{F}, \bar{\mathbf{Q}})$-Brownian motion. In addition, using (B.16) in "Appendix B", we obtain

$$\widetilde{S}_T = \widetilde{S}_t e^{\sigma_S \int_t^T \sqrt{P_{u-\tau}}\mathrm{d}\bar{W}_u + \frac{\sigma_S^2}{2}\int_t^T P_{u-\tau}\mathrm{d}u},$$

for every $t \in [0, T]$. Since S is $\widetilde{\mathbb{F}}$-adapted, by (B.16) and the Bayes formula on the change of probability measure for conditional expectation, for every $t \in [0, T)$ we get

$$\begin{aligned}
J_1 &= \mathbb{E}^{\mathbf{Q}^\gamma}\left[\widetilde{S}_T \mathbf{1}_{\{S_T > K\}}\middle|\widetilde{\mathcal{F}}_t\right] \\
&= \widetilde{S}_t \frac{\mathbb{E}^{\mathbf{Q}^\gamma}\left[e^{-\frac{\sigma_S^2}{2}X_T^\tau + \sigma_S Y_{0,T}}\mathbf{1}_{\{S_T > K\}}\middle|\widetilde{\mathcal{F}}_t\right]}{e^{-\frac{\sigma_S^2}{2}X_t^\tau + \sigma_S Y_{0,t}}} \\
&= \widetilde{S}_t \mathbb{E}^{\bar{\mathbf{Q}}}\left[\mathbf{1}_{\left\{\widetilde{S}_T > K B_T^{-1}\right\}}\middle|\widetilde{\mathcal{F}}_t\right] \\
&= \widetilde{S}_t \mathbb{E}^{\bar{\mathbf{Q}}}\left[\mathbb{E}^{\bar{\mathbf{Q}}}\left[\mathbf{1}_{\left\{\sigma_S \bar{Y}_{t,T} > \log\left(\frac{K}{S_t}\right) - \int_t^T r(u)\mathrm{d}u - \frac{\sigma_S^2}{2}X_{t,T}^\tau\right\}}\middle|\mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right]\middle|\widetilde{\mathcal{F}}_t\right] \\
&= \widetilde{S}_t \mathbb{E}^{\bar{\mathbf{Q}}}\left[\bar{\mathbf{Q}}\left(-\frac{\bar{Y}_{t,T}}{\sqrt{X_{t,T}^\tau}} < \frac{\log\left(\frac{S_t}{K}\right) + \int_t^T r(u)\mathrm{d}u + \frac{\sigma_S^2}{2}X_{t,T}^\tau}{\sigma_S\sqrt{X_{t,T}^\tau}}\middle|\mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right)\middle|\widetilde{\mathcal{F}}_t\right] \\
&= \widetilde{S}_t \mathbb{E}^{\bar{\mathbf{Q}}}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^\tau)\right)\middle|\widetilde{\mathcal{F}}_t\right],
\end{aligned} \tag{5.7}$$

with

$$d_1(t, S_t, X_{t,T}^\tau) = d_2(t, S_t, X_{t,T}^\tau) + \sigma_S\sqrt{X_{t,T}^\tau}.$$

In the above computations, analogously to before, we have set $\bar{Y}_{t,T} := \int_t^T \sqrt{P_{u-\tau}}\mathrm{d}\bar{W}_u$, for each $t \in [0, T)$. Consequently, we have that $\bar{Y}_{t,T}$ conditional on $\mathcal{F}_{T-\tau}^P$, is a Normally distributed random variable with mean 0 and variance $X_{t,T}^\tau$, for each $t \in [0, T)$, since Z is not affected by the change of measure from $\mathbf{Q}^\gamma$ to $\bar{\mathbf{Q}}$. Indeed, by the change of numéraire theorem, we have that the probability measure $\bar{\mathbf{Q}}$ turns out to be the minimal martingale measure corresponding to the choice of the Bitcoin price process as benchmark. Further, by applying again the Bayes formula on the change of probability measure for conditional expectation, we get

$$\begin{aligned}
J_1 &= \widetilde{S}_t \mathbb{E}^{\bar{\mathbf{Q}}}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^\tau)\right)\middle|\widetilde{\mathcal{F}}_t\right] \\
&= \widetilde{S}_t \frac{\mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^\tau)\right)e^{-\frac{\sigma_S^2}{2}\int_0^T P_{u-\tau}\mathrm{d}u + \sigma_S\int_0^T\sqrt{P_{u-\tau}}\mathrm{d}\widehat{W}_u}\middle|\widetilde{\mathcal{F}}_t\right]}{e^{-\frac{\sigma_S^2}{2}\int_0^t P_{u-\tau}\mathrm{d}u + \sigma_S\int_0^t\sqrt{P_{u-\tau}}\mathrm{d}\widehat{W}_u}} \\
&= \widetilde{S}_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^\tau)\right)e^{-\frac{\sigma_S^2}{2}X_{t,T}^\tau}\mathbb{E}^{\mathbf{Q}^\gamma}\left[e^{\sigma_S Y_{t,T}}\middle|\mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right]\middle|\widetilde{\mathcal{F}}_t\right] \\
&= \widetilde{S}_t \mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^\tau)\right)\middle|\widetilde{\mathcal{F}}_t\right],
\end{aligned} \tag{5.8}$$

since the conditional Gaussian distribution of $Y_{t,T}$ gives

$$\mathbb{E}^{\mathbf{Q}^{\gamma}}\left[e^{\sigma_S Y_{t,T}}\middle|\mathcal{F}_t^W \vee \mathcal{F}_{T-\tau}^P\right] = e^{\frac{\sigma_S^2}{2}X_{t,T}^{\tau}}.$$

Finally, gathering the two terms (5.8) and (5.7), for every $t \in [0, T)$ we obtain

$$\begin{aligned} C_t &= S_t \mathbb{E}^{\mathbf{Q}^{\gamma}}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^{\tau})\right)\middle|\widetilde{\mathcal{F}}_t\right] - K e^{-\int_t^T r(u)\mathrm{d}u}\mathbb{E}^{\mathbf{Q}^{\gamma}}\left[\mathcal{N}\left(d_2(t, S_t, X_{t,T}^{\tau})\right)\middle|\widetilde{\mathcal{F}}_t\right] \\ &= \mathbb{E}^{\mathbf{Q}^{\gamma}}\left[C^{BS}(t, S_t, X_{t,T}^{\tau})\middle|\widetilde{\mathcal{F}}_t\right] = \mathbb{E}^{\mathbf{Q}^{\gamma}}\left[C^{BS}(t, S_t, X_{t,T}^{\tau})\middle|S_t\right], \end{aligned}$$

where the last equality follows again from (Pascucci 2011, Lemma A.108), since for each $t \in [0, T)$, $X_{t,T}^{\tau}$ is independent of $\widetilde{\mathcal{F}}_t$ and S_t is $\widetilde{\mathcal{F}}_t$-measurable. □

It is worth noticing that the option pricing formula (5.6) only depends on the distribution of $X_{t,T}^{\tau}$ which is the same both under measure $\mathbf{Q}^{\gamma}$ and $\bar{\mathbf{Q}}$. As observed in Remark 5.2, formula (5.6) evaluated in S_t corresponds to the Black & Scholes price at time $t \in [0, T)$ of a European Call option written on S, with strike price K and maturity T, in a market where the volatility parameter is given by $\sigma_S\sqrt{\frac{x}{T-t}}$. Then, for every $t \in [0, T)$ it may be written as:

$$C_t = \int_0^{+\infty} C^{BS}(t, S_t, x) f_{X_{t,T}^{\tau}}(x)\mathrm{d}x,$$

where $f_{X_{t,T}^{\tau}}$ denotes the density function of $X_{t,T}^{\tau}$, for each $t \in [0, T)$, under measure $\mathbf{Q}^{\gamma}$. The price at time t for a *Plain Vanilla* European option may also be written as a Black & Scholes style price:

$$C_t = S_t Q_1 - K e^{-\int_t^T r(u)\mathrm{d}s} Q_2,$$

where

$$Q_1 := \mathbb{E}^{\mathbf{Q}^{\gamma}}\left[\mathcal{N}\left(d_1(t, S_t, X_{t,T}^{\tau})\right)\middle|S_t\right] = \int_0^{+\infty} \mathcal{N}\left(d_1(t, S_t, x)\right) f_{X_{t,T}^{\tau}}(x)\mathrm{d}x,$$

and

$$Q_2 := \mathbb{E}^{\mathbf{Q}^{\gamma}}\left[\mathcal{N}\left(d_2(t, S_t, X_{t,T}^{\tau})\right)\middle|S_t\right] = \int_0^{+\infty} \mathcal{N}\left(d_2(t, S_t, x)\right) f_{X_{t,T}^{\tau}}(x)\mathrm{d}x.$$

Similar formulas can be computed for other European-style derivatives as for binary options which, indeed, are quoted in Bitcoin markets. For the case of a Cash or Nothing

Call, which is essentially a bet of A on the exercise event, the risk-neutral pricing formula is given by

$$\begin{aligned} C_t^{Bin} &= Ae^{-\int_t^T r(u)\mathrm{d}s}\mathbb{E}^{\mathbf{Q}^\gamma}\left[\mathcal{N}\left(d_2(t,S_t,X_{t,T}^\tau)\right)\Big| S_t\right] \\ &= Ae^{-\int_t^T r(u)\mathrm{d}s}\int_0^{+\infty}\mathcal{N}\left(d_2(t,S_t,x)\right) f_{X_{t,T}^\tau}(x)\mathrm{d}x, \quad t\in[0,T). \end{aligned}$$

To compute derivative prices by applying the above formulas, we need an explicit form for the distribution function of $X_{t,T}^\tau$. One possible approach is to approximate this distribution by applying the outcomes in Levy (1992) where the author suggests a log-normal distribution as a proper choice. Specifically, the price at time $t = 0$ of Call option pricing formula reads:

$$C_0 = \int_0^{+\infty} C^{BS}(0, S_0, x)\mathcal{LN}pdf_{\alpha(T-\tau),\nu^2(T-\tau)}(x)\,\mathrm{d}x, \tag{5.9}$$

which can be computed numerically, once parameters $\alpha(T-\tau)$, $\nu(T-S)$, defined in “Appendix B” are obtained. Similarly, the price at $t = 0$ of a Binary Options with terminal value A when in the money, is given by:

$$C_0^{Bin} = A\int_0^{+\infty}\mathcal{N}\left(d_2(T-t, S_0, x)\right)\mathcal{LN}pdf_{\alpha(T-\tau),\nu^2(T-\tau)}(x)\,\mathrm{d}x.$$

6 Numerical applications

6.1 Sensitivity analysis of the pricing formula

In this subsection, we compute European Plain Vanilla and Binary option prices assuming that model parameters are known and considering several strike prices and expiration dates. Since the main features in the suggested model are the introduction of an attention process and a delay in the dependence of the price upon such stochastic factor, we believe it is important to further understand their contribution to the option price formation. To this end, we compute option prices by letting both the initial attention and the delay values change, while other parameters are set to $\mu_P = 0.03$, $\sigma_P = 0.35$, $\sigma_S = 0.04$, $\gamma = 0$ and $r = 0.01$; we assume, without loss of generality, that the Bitcoin price at time $t = 0$ is $S_0 = 450$.

In Table 3, Call option prices are reported for $T = 3$ months, τ= 5 days. Rows correspond to different values of P_0 while columns to different values for the strike price. As expected, Call option prices are increasing with respect to initial attention for the Bitcoin and decreasing with respect to strike price.

In Table 4, Call option prices are summed up, for $P_0 = 100$, by letting the expiration date T and the delay τ vary. Again as expected, for Plain Vanilla Calls the price increases with time to maturity. Increasing the delay reduces option prices; of course, the spread is inversely related to the time to maturity of the option.

Table 3 Call option prices against different strikes K and for different values of P_0: $S_0 = 450, r = 0.01, \mu_P = 0.03, \sigma_P = 0.35, \sigma_S = 0.04, T = 3$ months, $\tau = 1$ week (5 days)

K	400	425	450	475	500
$P_0 = 10$	51.24	28.35	11.46	3.09	0.54
$P_0 = 100$	64.12	48.05	34.94	24.69	16.97
$P_0 = 1000$	128.68	117.75	107.77	98.66	90.35

Table 4 Call option prices against different Strikes K and for different values of T and τ: $S_0 = 450, r = 0.01, \mu_P = 0.03, \sigma_P = 0.35, \sigma_S = 0.04$ and $P_0 = 100$

K	400	425	450	475	500
$T = 1$ month, $\tau = 1$ week	52.85	33.09	18.27	8.81	3.71
$T = 1$ month, $\tau = 2$ weeks	51.58	30.62	15.18	6.13	2.00
$T = 3$ months, $\tau = 1$ week	64.12	48.05	34.94	24.69	16.97
$T = 3$ months, $\tau = 2$ weeks	62.95	46.65	33.42	23.18	15.60

Table 5 Digital Cash or Nothing prices against different Strikes K and for different values of P_0 on Bitcoins

K	400	425	450	475	500
$P_0 = 10$	97.17	82.77	50.31	18.87	4.24
$P_0 = 100$	70.07	58.38	46.58	35.66	26.27
$P_0 = 1000$	45.70	41.77	38.14	34.79	31.72

Market parameters are $S_0 = 450$, $r = 0.01$, $\mu_P = 0.03$, $\sigma_P = 0.35$, $\sigma_S = 0.04$, $T = 3$ months, $\tau = 5$ days. The prize of the option is set to $A = 100$

In Tables 5 and 6, analogous results are reported for Binary Options with outcome $A = 100$; Table 5 sums up Binary Cash or Nothing prices for $S_0 = 450$, $r = 0.01$, $\mu_P = 0.03$, $\sigma_P = 0.35$, $\sigma_S = 0.04$, $T = 3$ months, $\tau = 1$ week (5 working days) against several strikes (in columns). Rows correspond to different values P_0 for the initial attention on Bitcoins. As expected, prices are decreasing with respect to strike prices. Here, *in the money* (ITM) options values are decreasing with respect to P_0 while *out of the money* (OTM) ones are increasing. The difference in ITM and OTM prices is large for low values of P_0, while it is very small for a high level of the initial attention factor in Bitcoins. This may be justified by the fact that, when the attention factor in the Bitcoin is strong, all bets are worth, even the OTM ones, since the underlying value is expected to blow up. Binary Call prices decrease with respect to time to maturity for ITM options and increase for OTM options which become more likely to be exercised. The influence of the delay value is tiny, as for vanilla options, being larger for short time to maturities.

Table 6 Digital Cash or Nothing prices against different Strikes K and for different values of T and τ

K	400	425	450	475	500
$T = 1$ month, $\tau = 1$ week	86.93	69.97	48.27	28.11	13.83
$T = 1$ month, $\tau = 2$ weeks	91.50	74.23	48.69	24.84	9.80
$T = 3$ months, $\tau = 1$ week	70.07	58.38	46.58	35.66	26.27
$T = 3$ months, $\tau = 2$ weeks	71.21	59.10	46.77	35.36	25.62

Market parameters are $S_0 = 450$, $r = 0.01$, $\mu_P = 0.03$, $\sigma_P = 0.35$, $\sigma_S = 0.04$ and $P_0 = 100$. The prize of the option is set to $A = 100$

6.2 Pricing performance on market option prices

In this subsection, we compute option prices in the case where model parameters are not simply assigned by the authors but are estimated on market data for Bitcoin prices. Again we assume that the attention price of risk is $\gamma = 0$ for two main reasons. The first one, as already remarked, because we do not expect the dynamics of the attention factor to change much from a risk-averse to a risk-neutral setting; the second motivation is that this parameter does not appear in the dynamics of the system under the physical measure $\mathbf{P}$, so it can not be estimated on historical data for Bitcoin price and market attention. Of course, its value might be obtained by calibration, i.e., by minimizing a suitable measure for the difference between model and market option prices on a given date, such as the Mean Squared Error (MSE) or its square root (RMSE). However, this approach considers option prices as given data; this is opposite to our purpose of computing market option prices by using data on the underlying. Nevertheless, the closed formula in 5.6 makes it possible to calibrate of parameters within our framework if we considered options prices as given. An example is given in following subsection[5].

Once model prices for options are computed these values are compared to corresponding market price. Data for market prices of options on Bitcoin may be retrieved from online platforms where it is possible to trade on Plain Vanilla and Binary options on the Bitcoin. A relevant platform where bid-ask quotes are publicly available is www.deribit.com, where the underlying is the Bitcoin average price available from `blockchain.info`; we consider the mid-value of the Bid-Ask range in this platform as a benchmark market price for assessing our pricing formula performance. We discard options for which there was no transactions. Every day two different expiration dates are available corresponding to a one month and two months maturity at issue. We are aware of possible synchronicity problems but as a first evaluation of the suggested pricing formula we intentionally neglect this friction. We compute the price of each of the traded options in our sample (from July 20 to July 31, 2017) according to formula (5.9) by plugging proper values for the strike price and the expiration date and for model parameters, estimated according to the profile likelihood method described in Sect. 3. In order to assess pricing performance, we compute the Root Mean Squared Error (RMSE) of model prices with respect to market prices across all the considered

[5] We thank an anonymous referee for this suggestion.

Table 7 Option Price with t=28 July and T=25 August

K	Bid	Ask	Model (Vol.)	Model (Google)	BS
2200	0.2200	0.2455	0.2125	0.2196	0.2110
2300	0.1956	0.2210	0.1820	0.1909	0.1799
2400	0.1600	0.1983	0.1538	0.1645	0.1510
2500	0.1280	0.1774	0.1281	0.1404	0.1248
2600	0.1050	0.1582	0.1052	0.1187	0.1015
2700	0.0850	0.1407	0.0867	0.0995	0.0812
2800	0.0703	0.1247	0.0696	0.0827	0.0640
2900	0.0540	0.1104	0.0482	0.0698	0.0497
RMSE mean bid/ask			0.0266	0.0140	0.0294

The strike prices K are reported in US Dollar while all other data are reported in BTC. The underlying Bitcoin price is 2760.93 \$

sample of options and of suitably chosen subsamples[6]. The same is done when prices are computed with the benchmark Black & Scholes model, where market attention is not accounted for. Of course, the latter is estimated on the same time series of Bitcoin prices; note that only the volatility estimation matters since we set $r = 0$ for both models. It is worth noticing that parameters are given by the estimation procedure on historical data hence the RMSE value is the output of our analysis, not the input function to be minimized, as it is done when a model is calibrated to market option prices in order to derive the underlying parameter. An example in this direction will be given in next subsection.

In Tables 7 and 8 we report market as well as model option prices for July 28, 2017. Black & Scholes prices are also reported and the RMSE is displayed in the last row. On the sample days available maturities were July 27 (options from July 20 to July 27), August 25 (options from July 28 to July 31) and September 29 (all options). Since the bid-ask range on the trading platform is quite large the choice of the best model is subjective and depends on which criteria are prioritized. Tables 7 and 8 evidence that Black and Scholes price is often out of bid-ask range while the proposed model performs better.

In Tables 9 and 10 we report the RMSE for the complete sample of options as well as for specific subsamples obtained according to moneyness and to maturity across all trading days. RMSE values are also computed and reported in the tables for the benchmark Black & Scholes model. Overall our pricing formula does much better than the benchmark in all cases. Highlighted in the tables (in bold) are the few cases where Black & Scholes model does better than the model we suggest, according to the RMSE.

Indeed, the pricing performance of the model depends on the choice for the attention measure to be considered which introduce another source of model risk[7]; however, it

[6] The Root Mean Squared Error is computed as the square root of the sum, across all options in the sample or subsample, of the squared differences between model and market prices.

[7] We thank an anonymous referee for pointing this.

Table 8 Option price with $t = 28$ July and $T = 29$ September

K	Bid	Ask	Model (Vol.)	Model (Google)	BS
2200	0.2108	0.2764	0.2350	0.2981	0.2297
2300	0.1881	0.2547	0.2091	0.2779	0.2029
2400	0.2004	0.2343	0.1851	0.2589	0.1781
2500	0.1700	0.2154	0.1631	0.2411	0.1555
2600	0.1650	0.1950	0.1431	0.2244	0.1349
2700	0.1168	0.1710	0.1249	0.2087	0.1164
2800	0.1044	0.1630	0.1086	0.1941	0.0999
2900	0.0934	0.1490	0.0940	0.1804	0.0853
RMSE mean bid/ask			0.0256	0.0543	0.0329

The strike prices K are reported in US Dollar while all other data are reported in BTC. The underlying Bitcoin price is 2760.93 $

Table 9 RMSE Option Price with Volume of Transactions

Options	Num.	RMSE Mod. (in BTC)	RMSE BS (in BTC)	RMSE Mod. (in %)	RMSE BS (in %)
All	144	0.0257	0.0407	12.55%	14.14%
Very Shorts	16	0.0225	**0.0204**	35.62%	**32.32%**
1 Month	32	0.0200	0.0256	2.00%	2.56%
2 Months	32	0.0231	0.0310	2.31%	3.10%
ITM	54	0.0209	0.0356	3.75%	7.03%
ATM	36	0.0281	0.0437	10.94%	13.22%
OTM	54	0.0283	0.0432	18.05%	19.16%

is worth noticing that when attention is measured by Google Searches the model tends to overprice long-term options while it is the very best for shorter term options as if this attention indicator is driven by enthusiasm giving a sudden impulse to options. This evidence can lead to the proper choice of the attention proxy depending on the expiration date of options to be priced. It would be interesting to consider a weighted mean of these two factors as a proxy of market attention rather than focusing on only one of them; this will be subject of future research.

6.3 Calibration of parameters

The availability of a closed formula for option prices makes it possible to derive parameters' value from market option prices rather than from historical prices of the underlying: this procedure is usually referred to as calibration; alternatively, different subsets of parameters may be derived by using market option data and underlying data, respectively.

Calibration is based on the minimization, with respect to model parameters, of a suitably defined distance between model prices and corresponding market prices for

Table 10 RMSE Option Price with Google Searches

Options	Num.	RMSE Mod. (in BTC)	RMSE BS (in BTC)	RMSE Mod. (in %)	RMSE BS (in %)
All	144	0.0290	0.0416	19.39%	28.40%
Very Shorts	16	0.0130	0.0229	24.71%	45.67%
1 Month	32	0.0110	0.0256	9.63%	24.02%
2 Months	32	0.0426	**0.0310**	27.21%	**19.98%**
ITM	54	0.0268	0.0366	12.13%	16.60%
ATM	36	0.0277	0.0448	16.76%	26.76%
OTM	54	0.0318	0.0441	25.86%	37.39%

a family of financial derivatives for which a market price, resulting from either a trade or a market maker quotation, is observed.

There has been a long debate in the literature whether calibration methodology is a sound estimation method; it is beyond the scope of this paper to go through pros and cons of the procedure. We just mention that calibration is usually performed when the market prices for derivatives are observed in large and trusted derivative exchanges; this is not the case in our setting, where at the time of writing, financial derivatives on Bitcoin are traded on online platforms which are not always trustworthy[8]. Nevertheless, if we want to consider a non-zero price γ for attention risk we need to resort, at least for this parameter, to calibration.

For the sake of completeness, we give a numerical example on how parameter γ can be calibrated on the options that we considered in the previous subsection, under the assumption that their price is given by the mid-value of the bid-ask market range. As for a suitable distance between model and market price we consider the Mean Squared Error.

More precisely, denote with $C_t = C_t(\mu_P, \sigma_P, \gamma, \mu_S, \sigma_S, \tau, K, T)$ the pricing function in t of a Call Option exiting at time T with strike price K where we have stressed the dependence on model parameters. Besides, denote with $C_t^*(T, K)$ the market price of a Call option with same characteristics.

Due to the tiny option market where Bitcoin options are traded, we believe that the PQML estimation procedure should be chosen when it is possible and we apply (6.1). Hence, we estimate $(\mu_P, \sigma_P, \mu_S, \sigma_S, \tau)$ with the PQML method introduced in Sect. 4 and derive γ obtained as:

$$\bar{\gamma} = \arg\min \sum_{j=1}^{J} \left(C_t(\hat{\mu}_P, \hat{\sigma}_P, \gamma, \hat{\mu}_S, \hat{\sigma}_S, \hat{\tau}, K_j, T_j) - C_t^*(K_j, T_j) \right)^2, \tag{6.1}$$

where J is the cardinality of the set of options for which a market value is available, and (K_j, T_j) for $j = 1, 2, \ldots, J$ are the characteristics of these options. By considering options traded on July 28, 2017 and assuming that their market price is the midprice

[8] With the exception of Futures, traded on both the CBOE and CME, which are not suited for calibration.

of the bid-ask displayed, respectively, in Tables 7 and 8 we obtain, according to (6.1), $\bar{\gamma} = -0.90$ and $\bar{\gamma} = 2.05$ for the trading volume and the SVI index, respectively, leading to minimum values for the RMSE of 0.0183 and 0.0121 in the two cases. The difference in sign evidences that in a risk-neutral world the increasing yield of the trading volume would be higher that in the real world why the opposite happens for the internet searches volume.

If we wanted to calibrate all of the parameters their value would be derived as:

$$(\bar{\mu}_P, \bar{\sigma}_P, \bar{\gamma}, \bar{\mu}_S, \bar{\sigma}_S, \bar{\tau}) = \arg\min \sum_{j=1}^{J} \big(C_t(\mu_P, \sigma_P, \gamma, \mu_S, \sigma_S, \tau, K_j, T_j) - C_t^*(K_j, T_j)\big)^2$$

but in this case the number J of options in the sample should be the very large to make estimation consistent, and, as already noticed, the Exchange trading the options should be a reference for the whole market.

7 Concluding remarks

In this paper, we assume that Bitcoin prices are driven by investors' attention as suggested in recent literature; main references in this area are Kristoufek (2013, 2015), Kim et al. (2015), Figà-Talamanca and Patacca (2019). In order to account for such behavior we develop a stochastic model in continuous time describing the dynamics of two factors, one representing the attention index on the Bitcoin system and the other representing the Bitcoin price itself, which is directly affected by the first factor; we also take into account a delay between the attention index and its delivered effect on the Bitcoin price. We investigate statistical properties of the proposed model and we show its arbitrage-free property. Further, under suitable model assumptions we derive a closedform approximation for the joint density of the discretely observed process and we propose a statistical estimation to fit the model to real data. By applying the classical risk-neutral evaluation we are able to derive a quasi-closed formula for European-style derivatives on the Bitcoin with special attention on Plain Vanilla and Binary options for which a market already exists (e.g., https://deribit.com, https://coinut.com). Of course, attention about Bitcoin or, more generally, on cryptocurrencies, is not directly observed but several variables may be considered as indicators. Here, we analyzed the trading volume, see (Kristoufek 2015), and more unconventional attention indicators such as the number of Google searches on the word "bitcoin", as suggested by Da et al. (2011). Firstly, we investigated whether these proxies were consistent with the suggested model and we proved that both the volume of transactions and the number of Google searches give a good fit of the dynamics described in the model. Finally, we fit the model using real data of Bitcoin price with the volume of transactions and the Google searches index, respectively. The ability of our pricing formula on capturing market option prices is also assessed on a sample of options traded on www.deribit.com; the overall performance outperforms substantially the benchmark Black & Scholes pricing formula.

A calibration example is also provided in order to estimate the price of attention risk on available option prices. There is still space for improvement in the suggested model to take into account stylized facts of observed price evidenced in Catania and Grassi (2017), Guo and Li (2017), Chu et al. (2015). Indeed, even if empirical evidence suggests that market attention/sentiment drives Bitcoin price behavior, other sources of randomness should be added to properly model its stochastic volatility. However, the specification given in Eqs. (2.1)–(2.2) allows to get several nice results, as the quasi-closed pricing formula for European-style derivatives on Bitcoin; the effects of including jumps in the price dynamics, generalizing the approach in Kou (2002) will be investigated. In a companion paper, Cretarola and Figà-Talamanca (2019), we consider a non-zero correlation between the two factors, focusing on the case of no delay; high correlation is shown to be linked to bubbles in the price dynamics, which have been also evidenced for the Bitcoin, see e.g., (Fry and Cheah 2015; Malhotra and Maloo 2014; Donier and Bouchaud 2015; Corbet et al. 2018; Bistarelli et al. 2019b).

Finally, we believe it is important to stress that another specific feature of Bitcoin is the existence of different prices for trades in different online platforms; hence, even considering the bid-ask mid price, several prices are available for the same asset at the same time. This characteristic makes arbitrage opportunities arise across exchanges, since the law of unique price is not fulfilled. In the present paper we ignore this issue by assigning to Bitcoin the average price computed and provided by the Web site https://blockchain.info; this choice is consistent with most literature on Bitcoin price behavior but we plan to give a multivariate specification of the suggested model. Preliminary results in this direction, to account for a multi-exchange framework and possible cross-exchange arbitrage opportunities, are discussed in Bistarelli et al. (2018, 2019a).

Acknowledgements The authors are thankful to Fabio Bellini for valuable discussions on the topic of the present paper and to an anonymous referee for valuable suggestions that helped to significantly improve the paper. Part of the paper was written during the stay of the second-named author at the Department of Finance and Risk Engineering, Tandon School of Engineering, New York University. The second-named author is especially grateful to Peter Carr and Charles Tapiero for valuable suggestions and stimulating discussions. The authors also gratefully acknowledge Stefano Bistarelli for having introduced themselves to the intriguing world of cryptocurrencies. The first-named author is member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM). The authors notify that this paper has circulated under the title "A sentiment-based model for the BitCoin: theory, estimation and option pricing" and, in its preliminary form, under the title "A confidence-based model for asset and derivative prices in the bitcoin market".

Funding The first and the second-named authors are grateful to Bank of Italy and Fondazione Cassa di Risparmio Perugia for the financial support, through Project Numbers 407660/16 and 2015.0459013, respectively.

Appendix A: Levy approximation

In Levy (1992) the author proves that the distribution of the mean integrated Brownian motion $\frac{1}{s}\int_0^s P_u \mathrm{d}u$ can be approximated with a log-normal distribution, at least for suitable values of the model parameters μ_P, σ_P; the parameters of the approximating log-normal distribution are obtained by applying a moment matching technique. Set

$$IP(s) := \int_0^s P_u \mathrm{d}u, \quad s > 0. \tag{A.1}$$

Of course, the distribution of $IP(s)$ can also be approximated by a log-normal for $s > 0$. By applying the moment matching technique the parameters of the corresponding log-Normal distribution for $IP(s)$ are given by

$$\alpha(s) = \log\left(\frac{\mathbb{E}[IP(s)]^2}{\sqrt{\mathbb{E}\left[IP(s)^2\right]}}\right),$$
$$\nu^2(s) = \log\left(\frac{\mathbb{E}\left[IP(s)^2\right]}{\mathbb{E}[IP(s)]^2}\right).$$

The approximate distribution density function $f_{IP(s)}$ of $IP(s)$ is thus given by

$$f_{IP(s)}(x) = \mathcal{LN}pdf_{\alpha(s),\nu^2(s)}(x), \quad \text{if } s > 0,$$

where $\mathcal{LN}pdf_{m,v}$ denotes the probability distribution function of a log-normal distribution with parameters m and v, defined as

$$\mathcal{LN}pdf_{m,v}(y) = \frac{1}{y\sqrt{2\pi v}} e^{-\frac{(log(y)-m)^2}{2v}}, \quad \forall\, y \in \mathbb{R}^+.$$

In the paper the above approximation is applied twice with completely different purposes. In Sect. 4, once $\tau < \Delta$ is assigned, the Levy approximation is applied to derive the distribution of A_1 and of A_i given A_{i-1} where $A_1 = X^\tau_\tau + IP(\Delta - \tau)$ and, for $i \geq 2$, $A_i = \int_{(i-1)\Delta-\tau}^{i\Delta-\tau} P_u du = \int_{-\tau}^{\Delta-\tau} P_{u+(i-1)\Delta} \mathrm{d}u = P_{(i-1)\Delta} \int_{-\tau}^{\Delta-\tau} P_u \mathrm{d}u = P_{(i-1)\Delta}\left(X^\tau_\tau + IP(\Delta - \tau)\right)$.

In Sect. 5 it is applied to derive an approximate distribution for the integrated attention process starting at $t = 0$, i.e., to $X^\tau_{0,T} = X^\tau_\tau + IP(T - \tau)$. Note that $X^\tau_{0,T} - X^\tau_\tau = IP(T - \tau)$ hence the derivations of its distribution is trivial once that of $IP(T - \tau)$ is known.

Appendix B: Technical proofs

The following result provides basic statistical properties for the integrated attention process X^τ defined in (2.3), as well as for its variation in case they are not fully deterministic.

Lemma B.1 *In the market model outlined in Eqs.* (2.1)–(2.2)*, we have:*

(i) *For* $t > \tau$,

$$\mathbb{E}\left[X_t^\tau\right] = X_\tau^\tau + \frac{\phi(0)}{\mu_P}\left(e^{\mu_P(t-\tau)} - 1\right);$$

$$\begin{aligned}\mathbb{V}\mathrm{ar}[X_t^\tau] = &\frac{2\phi^2(0)}{\left(\mu_P + \sigma_P^2\right)\left(2\mu_P + \sigma_P^2\right)}\left(e^{\left(2\mu_P+\sigma_P^2\right)(t-\tau)} - 1\right)\\ &- \frac{2\phi^2(0)}{\mu_P\left(\mu_P + \sigma_P^2\right)}\left(e^{\mu_P(t-\tau)} - 1\right) - \left(\frac{\phi(0)}{\mu_P}\left(e^{\mu_P(t-\tau)} - 1\right)\right)^2.\end{aligned}$$

(ii) *For* $\tau \le t < T$,

$$\mathbb{E}\left[X_{t,T}^\tau\right] = \frac{\phi(0)e^{\mu_P(t-\tau)}}{\mu_P}\left(e^{\mu_P(T-t)} - 1\right);$$

$$\begin{aligned}\mathbb{V}\mathrm{ar}[X_{t,T}^\tau] = &\frac{2\phi^2(0)e^{\left(2\mu_P+\sigma_P^2\right)(t-\tau)}}{\left(\mu_P + \sigma_P^2\right)\left(2\mu_P + \sigma_P^2\right)}\left(e^{\left(2\mu_P+\sigma_P^2\right)(T-t)} - 1\right)\\ &- \frac{2\phi^2(0)e^{\mu_P(t-\tau)}}{\mu_P\left(\mu_P + \sigma_P^2\right)}\left(e^{\mu_P(T-t)} - 1\right)\\ &- \left(\frac{\phi(0)e^{\mu_P(t-\tau)}}{\mu_P}\left(e^{\mu_P(T-t)} - 1\right)\right)^2.\end{aligned}$$

(iii) *For* $t \le \tau < T$,

$$\mathbb{E}\left[X_{t,T}^\tau\right] = \int_{t-\tau}^0 \phi(u)\,\mathrm{d}u + \frac{\phi(0)}{\mu_P}\left(e^{\mu_P(T-\tau)} - 1\right);$$

$$\begin{aligned}\mathbb{V}\mathrm{ar}[X_{t,T}^\tau] = &\frac{2\phi^2(0)}{\left(\mu_P + \sigma_P^2\right)\left(2\mu_P + \sigma_P^2\right)}\left(e^{\left(2\mu_P+\sigma_P^2\right)(T-\tau)} - 1\right)\\ &- \frac{2\phi^2(0)}{\mu_P\left(\mu_P + \sigma_P^2\right)}\left(e^{\mu_P(T-\tau)} - 1\right) - \left(\frac{\phi(0)}{\mu_P}\left(e^{\mu_P(T-\tau)} - 1\right)\right)^2.\end{aligned}$$

In Hull and White (1987) similar outcomes are claimed for $\tau = 0$ without providing a proof; for the sake of clarity, we give here a self-contained proof.

Proof In order to prove the Lemma let us first compute the mean and the variance of $IP(s)$ given in (A.1) for each $s > 0$.

Fix $s > 0$. Since $P_u > 0$ for each $u \in (0, s]$, by applying the Fubini theorem we get

$$\mathbb{E}[IP(s)] = \mathbb{E}\left[\int_0^s P_u \mathrm{d}u\right] = \int_0^s \mathbb{E}[P_u]\,\mathrm{d}u,$$

where, for each $u \geq 0$, we have

$$\mathbb{E}[P_u] = \phi(0)e^{\left(\mu_P - \frac{\sigma_P^2}{2}\right)u}\mathbb{E}\left[e^{\sigma_P Z_u}\right] = \phi(0)e^{\mu_P u},$$

since P is a geometric Brownian motion with $P_0 = \phi(0)$. Hence

$$\mathbb{E}[IP(s)] = \phi(0)\int_0^s e^{\mu_P u}\mathrm{d}u = \frac{\phi(0)}{\mu_P}\left(e^{\mu_P(s)} - 1\right).$$

As for the variance of $IP(s)$, we have

$$\mathbb{V}\mathrm{ar}[IP(s)] = \mathbb{V}\mathrm{ar}\left[\int_0^s P_u \mathrm{d}u\right] = \mathbb{E}\left[\left(\int_0^s P_u \mathrm{d}u\right)^2\right] - \mathbb{E}\left[\int_0^s P_u \mathrm{d}u\right]^2,$$

with

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^s P_u \mathrm{d}u\right)^2\right] &= 2\mathbb{E}\left[\int_0^s P_v dv \int_0^v P_u \mathrm{d}u\right] = 2\mathbb{E}\left[\int_0^s \int_0^v P_u P_v \mathrm{d}v\mathrm{d}u\right] \\ &= 2\int_0^s \int_0^v \mathbb{E}[P_u P_v]\,\mathrm{d}v\mathrm{d}u, \end{aligned} \tag{B.1}$$

where the last equality again holds thanks to Fubini's theorem. Moreover, by the independence property of the increments of Brownian motion, for $0 < u < v \leq s$, we get

$$\begin{aligned}\mathbb{E}[P_u P_v] &= \mathbb{E}\left[P_u^2 e^{\left(\mu_P - \frac{\sigma_P^2}{2}\right)(v-u) + \sigma_P(Z_v - Z_u)}\right] \\ &= e^{\left(\mu_P - \frac{\sigma_P^2}{2}\right)(v-u)}\mathbb{E}\left[P_u^2 \mathbb{E}\left[e^{\sigma_P(Z_v - Z_u)}|\mathcal{F}_u^P\right]\right] \\ &= e^{\left(\mu_P - \frac{\sigma_P^2}{2}\right)(v-u)}\mathbb{E}\left[P_u^2\right]\mathbb{E}\left[e^{\frac{\sigma_P^2(v-u))}{2}}\right] = e^{\mu_P(v-u)}\mathbb{E}\left[P_u^2\right].\end{aligned}$$

Further,

$$\mathbb{E}\left[P_u^2\right] = \phi^2(0)e^{2\left(\mu_P - \frac{\sigma_P^2}{2}\right)u}\mathbb{E}\left[e^{2\sigma_P Z_u}\right] = \phi^2(0)e^{(2\mu_P + \sigma_P^2)u}.$$

Hence

$$\mathbb{E}[P_u P_v] = \phi^2(0)e^{\mu_P(v-u)}e^{(2\mu_P + \sigma_P^2)u}, \tag{B.2}$$

and by plugging (B.2) into (B.1), we have

$$\begin{aligned}\mathbb{E}\left[IP(s)^2\right] &= 2\phi^2(0)\int_0^s e^{\mu_P v}\int_0^v e^{(\mu_P+\sigma_P^2)u}\mathrm{d}u\mathrm{d}v\\ &= \frac{2\phi^2(0)}{\left(\mu_P+\sigma_P^2\right)\left(2\mu_P+\sigma_P^2\right)}\left(e^{(2\mu_P+\sigma_P^2)s}-1\right)\\ &\quad -\frac{2\phi^2(0)}{\mu_P\left(\mu_P+\sigma_P^2\right)}\left(e^{\mu_P s}-1\right).\end{aligned}$$

Finally, gathering the results we get

$$\begin{aligned}\mathbb{V}\mathrm{ar}[IP(s)] &= \frac{2\phi^2(0)}{\left(\mu_P+\sigma_P^2\right)\left(2\mu_P+\sigma_P^2\right)}\left[e^{(2\mu_P+\sigma_P^2)s}-1\right]\\ &\quad -\frac{2\phi^2(0)}{\mu_P\left(\mu_P+\sigma_P^2\right)}\left(e^{\mu_P s}-1\right)-\left(\frac{\phi(0)}{\mu_P}\left(e^{\mu_P s}-1\right)\right)^2.\end{aligned}$$

Note that X_t^τ, with $t\in[0,\tau]$, and $X_{t,T}^\tau$, with $t<T\le\tau$, are fully deterministic and the computation is trivial.

To prove points (i)–(iii), it suffices to observe that

$$\begin{aligned}X_t^\tau &= X_\tau^\tau + IP(t-\tau),\quad t>\tau,\\ X_{t,T}^\tau &= \int_{t-\tau}^0 \phi(u)\mathrm{d}u + IP(T-\tau),\quad t\le\tau<T,\end{aligned}$$

and the computation easily follows once those of $IP(s)$ are known for $s>0$.

To prove point (ii), it is worth noticing that given $0\le v<s$

$$\begin{aligned}IP(s)-IP(v) &= \int_v^s P_u\mathrm{d}u = \int_v^s P_v e^{\left(\mu_P-\frac{\sigma_P^2}{2}\right)(u-v)+\sigma_P(Z_u-Z_v)}\mathrm{d}u\\ &\overset{\text{(law)}}{=} P_v\int_v^s e^{\left(\mu_P-\frac{\sigma_P^2}{2}\right)(u-v)+\sigma_P Z_{u-v}}\mathrm{d}u\\ &= P_v\int_0^{s-v} e^{\left(\mu_P-\frac{\sigma_P^2}{2}\right)r+\sigma_P Z_r}\mathrm{d}r = P_v IP(s-v),\end{aligned}\tag{B.3}$$

where $r=u-v$. To obtain the desired result, it suffices to note that, for $\tau\le t<T$,

$$X_{t,T}^\tau = IP(T-\tau)-IP(t-\tau)$$

and apply (B.3). The computation of the mean and variance of the above difference is straightforward given the independence of Brownian increments. □

The system given by Eqs. (2.1)–(2.2) is well-defined in $\mathbb{R}_+$, as stated in the following theorem, which also provides its explicit solution.

Theorem B.2 *In the market model outlined in (2.1)–(2.2), the followings hold:*

(i) *the bivariate stochastic delayed differential equation*

$$\begin{cases} \mathrm{d}S_t = \mu_S P_{t-\tau} S_t \mathrm{d}t + \sigma_S \sqrt{P_{t-\tau}} S_t \mathrm{d}W_t, & S_0 = s_0 \in \mathbb{R}_+, \\ \mathrm{d}P_t = \mu_P P_t \mathrm{d}t + \sigma_P P_t \mathrm{d}Z_t, & P_t = \phi(t),\ t \in [-L, 0], \end{cases} \tag{B.4}$$

has a continuous, $\mathbb{F}$-adapted, unique solution $(S, P) = \{(S_t, P_t),\ t \geq 0\}$ given by

$$S_t = s_0 e^{\left(\mu_S - \frac{\sigma_S^2}{2}\right) \int_0^t P_{u-\tau} \mathrm{d}u + \sigma_S \int_0^t \sqrt{P_{u-\tau}} \mathrm{d}W_u}, \quad t \geq 0, \tag{B.5}$$

$$P_t = \phi(0) e^{\left(\mu_P - \frac{\sigma_P^2}{2}\right) t + \sigma_P Z_t}, \quad t \geq 0. \tag{B.6}$$

More precisely, S can be computed step by step as follows: for $k = 0, 1, 2, \ldots$ and $t \in [k\tau, (k+1)\tau]$,

$$S_t = S_{k\tau} e^{\left(\mu_S - \frac{\sigma_S^2}{2}\right) \int_{k\tau}^t P_{u-\tau} \mathrm{d}u + \sigma_S \int_{k\tau}^t \sqrt{P_{u-\tau}} \mathrm{d}W_u}. \tag{B.7}$$

In particular, $P_t \geq 0$ $\mathbf{P}$-a.s. for all $t \geq 0$. If in addition, $\phi(0) > 0$, then $P_t > 0$ $\mathbf{P}$-a.s. for all $t \geq 0$.

(ii) *Further, for every $t \geq 0$, the conditional distribution of S_t, given the integrated attention X_t^τ, is log-Normal with mean $\log(s_0) + \left(\mu_S - \frac{\sigma_S^2}{2}\right) X_t^\tau$ and variance $\sigma_S^2 X_t^\tau$.*

(iii) *Finally, for every $t \in [0, \tau]$, the random variable $\log(S_t)$ has mean $\log(s_0) + \left(\mu_S - \frac{\sigma_S^2}{2}\right) X_t^\tau$ and variance $\sigma_S^2 X_t^\tau$; for every $t > \tau$, $\log(S_t)$ has mean and variance, respectively, given by*

$$\mathbb{E}\left[\log(S_t)\right] = \log(s_0) + \left(\mu_S - \frac{\sigma_S^2}{2}\right) \mathbb{E}\left[X_t^\tau\right];$$

$$\mathbb{V}\mathrm{ar}\left[\log(S_t)\right] = \left(\mu_S - \frac{\sigma_S^2}{2}\right)^2 \mathbb{V}\mathrm{ar}[X_t^\tau] + \sigma_S^2 \mathbb{E}\left[X_t^\tau\right],$$

where $\mathbb{E}\left[X_t^\tau\right]$ and $\mathbb{V}\mathrm{ar}[X_t^\tau]$ are both provided by Lemma B.1, point (i).

Proof **Point (i)**. Clearly, S and P, given in (B.5) and (B.6), respectively, are $\mathbb{F}$-adapted processes with continuous trajectories. Similarly to (Mao and Sabanis 2013, Theorem 2.1), we provide existence and uniqueness of a strong solution to the pair of stochastic differential equations in system (B.4) by using forward induction steps of length τ, without the need of checking any assumptions on the coefficients, e.g., the local Lipschitz condition and the linear growth condition.

First, note that the second equation in the system (B.4) does not depend on S, and its solution is well known for all $t \geq 0$. Clearly, Eq. (B.6) says that $P_t \geq 0$ **P**-a.s. for all $t \geq 0$ and that $\phi(0) > 0$ implies that the solution P remains strictly greater than 0 over $[0, +\infty)$, i.e., $P_t > 0$, **P**-a.s. for all $t \geq 0$.
Next, by the first Eq. (B.4) and applying Itô's formula to $\log(S_t)$, we get

$$\mathrm{d}\log(S_t) = \left(\mu_S - \frac{\sigma_S^2}{2}\right) P_{t-\tau}\mathrm{d}t + \sigma_S\sqrt{P_{t-\tau}}\mathrm{d}W_t, \tag{B.8}$$

or equivalently, in integral form

$$\log\left(\frac{S_t}{s_0}\right) = \left(\mu_S - \frac{\sigma_S^2}{2}\right)\int_0^t P_{u-\tau}\mathrm{d}u + \sigma_S\int_0^t \sqrt{P_{u-\tau}}\mathrm{d}W_u, \quad t \geq 0. \tag{B.9}$$

For $t \in [0, \tau]$, (B.9) can be written as

$$\log\left(\frac{S_t}{s_0}\right) = \left(\mu_S - \frac{\sigma_S^2}{2}\right)\int_0^t \phi(u-\tau)\,\mathrm{d}u + \sigma_S\int_0^t \sqrt{\phi(u-\tau)}\mathrm{d}W_u, \tag{B.10}$$

that is, (B.7) holds for $k = 0$.

Given that S_t is now known for $t \in [0, \tau]$, we may restrict the first Eq. (B.4) on $t \in [\tau, 2\tau]$, so that it corresponds to consider (B.8) for $t \in [\tau, 2\tau]$. Equivalently, in integral form,

$$\log\left(\frac{S_t}{S_\tau}\right) = \left(\mu_S - \frac{\sigma_S^2}{2}\right)\int_\tau^t P_{u-\tau}\mathrm{d}u + \sigma_S\int_\tau^t \sqrt{P_{u-\tau}}\mathrm{d}W_u.$$

This shows that (B.7) holds for $k = 1$. Similar computations for $k = 2, 3, \ldots$, give the final result.
Point (ii). Set $Y_t := \int_0^t \sqrt{\phi(t-\tau)}\mathrm{d}W_u$, for $t \in [0, \tau]$ and $Y_t := Y_{k\tau} + \int_{k\tau}^t \sqrt{P_{u-\tau}}\mathrm{d}W_u$, for $t \in [k\tau, (k+1)\tau]$, with $k = 1, 2, \ldots$. Then, by applying the outcomes in Point (i) and the decomposition

$$\log\left(\frac{S_t}{s_0}\right) = \log\left(\frac{S_t}{S_{k\tau}}\right) + \sum_{j=0}^{k-1}\log\left(\frac{S_{(j+1)\tau}}{S_{j\tau}}\right),$$

for $t \in [k\tau, (k+1)\tau]$, with $k = 1, 2, \ldots$, we can write

$$\log(S_t) = \log(s_0) + \left(\mu_S - \frac{\sigma_S^2}{2}\right) X_t^\tau + \sigma_S Y_t, \quad t \geq 0. \tag{B.11}$$

To complete the proof, it suffices to show that, for each $t \geq 0$ the random variable Y_t, conditional on X_t^τ, is Normally distributed with mean 0 and variance X_t^τ. This

is straightforward from (B.10) if $t \in [0, \tau]$. Otherwise, we first observe that since $Z_{u-\tau}$ is independent of W_u for every $\tau < u \leq t$, the distribution of Y_t, conditional on $\{Z_{u-\tau} : \tau < u \leq t - \tau\} = \{P_{u-\tau} : \tau < u \leq t - \tau\} = \mathcal{F}^P_{t-\tau}$, is Normal with mean 0 and variance $\sigma_S^2 X_t^\tau$.
Now, for each $t > \tau$, the moment-generating function of Y_t, conditioned on the history of the process P up to time $t - \tau$, is given by

$$\begin{aligned}\mathbb{E}\left[e^{aY_t}\middle|\mathcal{F}^P_{t-\tau}\right] &= e^{\int_0^t \frac{a^2}{2} P_{u-\tau}\mathrm{d}u} = e^{\frac{a^2}{2}\int_0^t P_{u-\tau}\mathrm{d}u} \\ &= e^{\frac{a^2}{2}\left(\sqrt{X_t^\tau}\right)^2}, \quad a \in \mathbb{R},\end{aligned}$$

that only depends on X_t^τ up to time t, that is,

$$\mathbb{E}\left[e^{aY_t}\middle|\mathcal{F}^P_{t-\tau}\right] = \mathbb{E}\left[e^{aY_t}\middle|X_t^\tau\right], \quad t > \tau.$$

Point (iii). The proof is trivial for $t \in [0, \tau]$. If $t > \tau$, (B.11) and Lemma B.1, point (i), together with the null-expectation property of the Itô integral, give

$$\mathbb{E}\left[\log(S_t)\right] = \log(s_0) + \left(\mu_S - \frac{\sigma_S^2}{2}\right)\left(X_\tau^\tau + \frac{\phi(0)}{\mu_P}\left(e^{\mu_P(t-\tau)} - 1\right)\right).$$

Now, we compute the variance of $\log(S_t)$. Since for each $t > \tau$ the random variable Y_t has mean 0 conditional on $\mathcal{F}^P_{t-\tau}$, we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\log(S_t)\right] &= \mathbb{E}\left[\log^2(S_t)\right] - \left(\mathbb{E}\left[\log(S_t)\right]\right)^2 \\ &= \left(\mu_S - \frac{\sigma_S^2}{2}\right)^2 \mathbb{E}\left[(X_t^\tau)^2\right] + 2\left(\mu_S - \frac{\sigma_S^2}{2}\right)\mathbb{E}\left[X_t^\tau \mathbb{E}\left[Y_t\middle|\mathcal{F}^P_{t-\tau}\right]\right] \\ &\quad + \sigma_S^2 \mathbb{E}\left[X_t^\tau\right] - \left(\mu_S - \frac{\sigma_S^2}{2}\right)^2 \mathbb{E}\left[X_t^\tau\right]^2 \\ &= \left(\mu_S - \frac{\sigma_S^2}{2}\right)^2 \mathbb{V}\mathrm{ar}[X_t^\tau] + \sigma_S^2 \mathbb{E}\left[X_t^\tau\right].\end{aligned}$$

Thus, the proof is complete. □

Let $T > 0$ be a fixed and finite time horizon. The following result ensures that the model given in (2.1)–(2.2) is arbitrage-free.

Lemma B.3 *Let $\phi(t) > 0$, for each $t \in [-L, 0]$, in (2.2). Then, every equivalent martingale measure $\mathbf{Q}$ for S defined on $(\Omega, \mathcal{F}_T)$ has the following density*

$$\left.\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{F}_T} =: L_T^{\mathbf{Q}}, \quad \mathbf{P} - a.s.,$$

where $L_T^{\mathbf{Q}}$ *is the terminal value of the* $(\mathbb{F}, \mathbf{P})$*-martingale* $L^{\mathbf{Q}} = \{L_t^{\mathbf{Q}},\ t \in [0, T]\}$ *given by*

$$L_t^{\mathbf{Q}} := \mathcal{E}\left(-\int_0^{\cdot} \frac{\mu_S P_{s-\tau} - r(s)}{\sigma_S \sqrt{P_{s-\tau}}} \mathrm{d}W_s - \int_0^{\cdot} \gamma_s \mathrm{d}Z_s\right)_t, \quad t \in [0, T], \tag{B.12}$$

for a suitable $\mathbb{F}$*-progressively measurable process* $\gamma = \{\gamma_t,\ t \in [0, T]\}$.

Proof Firstly, we prove that formula (B.12) defines a probability measure $\mathbf{Q}$ equivalent to $\mathbf{P}$ on $(\Omega, \mathcal{F}_T)$. This means we need to show that $L^{\mathbf{Q}}$ is an $(\mathbb{F}, \mathbf{P})$-martingale, that is, $\mathbb{E}\left[L_T^{\mathbf{Q}}\right] = 1$. Since the $\mathbb{F}$-progressively measurable process γ can be suitably chosen, to prove this relation we can assume $\gamma \equiv 0$, without loss of generality. Set

$$\alpha_t := -\frac{\mu_S P_{t-\tau} - r(t)}{\sigma_S \sqrt{P_{t-\tau}}}, \quad t \in [0, T]. \tag{B.13}$$

We observe that since $\phi(t) > 0$, for each $t \in [-L, 0]$, in (2.2), by Theorem B.2, point (i), we have that $P_{t-\tau} > 0$, $\mathbf{P}$-a.s. for all $t \in [0, T]$, so that the process $\alpha = \{\alpha_t,\ t \in [0, T]\}$ given in (B.13) is well-defined, as well as the random variable $L_T^{\mathbf{Q}}$. Clearly, α is an $\mathbb{F}$-progressively measurable process. Moreover, since the trajectories of the process P are continuous, then P is almost surely bounded on $[0, T]$ and this implies that $\int_0^T |\alpha_u|^2 \mathrm{d}u < \infty$ $\mathbf{P}$-a.s.; on the other hand, the condition $\phi(t) > 0$, for every $t \in [-L, 0]$, implies that almost every path of $\left\{\frac{1}{\sigma_S \sqrt{P_{t-\tau}}},\ t \in [0, T]\right\}$ is bounded on the compact interval $[0, T]$. Set $\mathcal{F}_t^P := \mathcal{F}_0^P = \{\Omega, \emptyset\}$, for $t \leq 0$. Then, α_u, for every $u \in [0, T]$, is $\mathcal{F}_{T-\tau}^P$-measurable. Since $Z_{u-\tau}$ is independent of W_u, for every $u \in [\tau, T]$, the stochastic integral $\int_0^T \alpha_u \mathrm{d}W_u$ conditioned on $\mathcal{F}_{T-\tau}^P$ has a normal distribution with mean zero and variance $\int_0^T |\alpha_u|^2 \mathrm{d}u$. Consequently, the formula for the moment-generating function of a normal distribution implies

$$\mathbb{E}\left[e^{\int_0^T \alpha_u \mathrm{d}W_u} \middle| \mathcal{F}_{T-\tau}^P\right] = e^{\frac{1}{2}\int_0^T |\alpha_u|^2 \mathrm{d}u},$$

or equivalently

$$\mathbb{E}\left[e^{\int_0^T \alpha_u \mathrm{d}W_u - \frac{1}{2}\int_0^T |\alpha_u|^2 \mathrm{d}u} \middle| \mathcal{F}_{T-\tau}^P\right] = 1. \tag{B.14}$$

Taking the expectation of both sides of (B.14) immediately yields $\mathbb{E}\left[L_T^{\mathbf{Q}}\right] = 1$. Now, set $\widetilde{S}_t := \dfrac{S_t}{B_t}$, for each $t \in [0, T]$. It remains to verify that the discounted Bitcoin price process $\widetilde{S} = \{\widetilde{S}_t,\ t \in [0, T]\}$ is an $(\mathbb{F}, \mathbf{Q})$-martingale. By Girsanov's theorem, under the change of measure from $\mathbf{P}$ to $\mathbf{Q}$, we have two independent $(\mathbb{F}, \mathbf{Q})$-Brownian motions $W^{\mathbf{Q}} = \{W_t^{\mathbf{Q}},\ t \in [0, T]\}$ and $Z^{\mathbf{Q}} = \{Z_t^{\mathbf{Q}},\ t \in [0, T]\}$ defined, respectively, by

$$W_t^{\mathbf{Q}} := W_t - \int_0^t \alpha_u \mathrm{d}u, \quad Z_t^{\mathbf{Q}} := Z_t + \int_0^t \gamma_u \mathrm{d}u, \quad t \in [0, T].$$

Under the martingale measure $\mathbf{Q}$, the discounted Bitcoin price process $\widetilde{S}$ satisfies the following dynamics

$$\mathrm{d}\widetilde{S}_t = \widetilde{S}_t \sigma_S \sqrt{P_{t-\tau}} \mathrm{d}W_t^{\mathbf{Q}}, \quad \widetilde{S}_0 = s_0 \in \mathbb{R}_+,$$

which implies that $\widetilde{S}$ is an $(\mathbb{F}, \mathbf{Q})$-local martingale. Finally, proceeding as above it is easy to check that $\widetilde{S}$ is a true $(\mathbb{F}, \mathbf{Q})$-martingale. □

Here $\mathcal{E}(Y)$ denotes the Doléans-Dade exponential of an $(\mathbb{F}, \mathbf{P})$-semimartingale Y. Then, Lemma B.3 ensures that the space of equivalent martingale measures for S is described by (B.12). Note that the attention factor dynamics under $\mathbf{Q}$ in the Bitcoin market is given by

$$\mathrm{d}P_t = (\mu_P - \sigma_P \gamma_t) P_t \mathrm{d}t + \sigma_P P_t \mathrm{d}Z_t^{\mathbf{Q}}, \quad P_t = \phi(t), \ t \in [-L, 0].$$

The process γ can be interpreted as the price for attention risk, i.e., the perception associated to the future direction of investor attention on the Bitcoin market. The probability measure which corresponds to the choice $\gamma \equiv 0$ in (B.12) is the so-called *minimal martingale measure*. Let us focus on the special case of a constant process γ; in this setting the equivalent martingale measure, denoted by $\mathbf{Q}^\gamma$, is parameterized by the constant γ which governs the change of drift of the $(\mathbb{F}, \mathbf{P})$-Brownian motion Z. As a special case of the above dynamics, the two independent $(\mathbb{F}, \mathbf{Q}^\gamma)$-Brownian motions $\widehat{W} = \{\widehat{W}_t, \ t \in [0, T]\}$ and $\widehat{Z} = \{\widehat{Z}_t, \ t \in [0, T]\}$ are defined, respectively, by

$$\begin{aligned} \widehat{W}_t &:= W_t + \int_0^t \frac{\mu_S P_{s-\tau} - r(s)}{\sigma_S \sqrt{P_{s-\tau}}} \mathrm{d}s, \quad t \in [0, T], \\ \widehat{Z}_t &:= Z_t + \gamma\, t, \quad t \in [0, T]. \end{aligned}$$

Denote by $\widetilde{S}_t = \{\widetilde{S}_t, \ t \in [0, T]\}$ the discounted Bitcoin price process defined as $\widetilde{S}_t := \dfrac{S_t}{B_t}$, for each $t \in [0, T]$. Then, on the probability space $(\Omega, \mathcal{F}, \mathbf{Q}^\gamma)$, the pair $(\widetilde{S}, P)$ satisfies the following system of stochastic delayed differential equations:

$$\begin{cases} \mathrm{d}\widetilde{S}_t = \sigma_S \sqrt{P_{t-\tau}} \widetilde{S}_t \mathrm{d}\widehat{W}_t, \quad \widetilde{S}_0 = s_0 \in \mathbb{R}_+, \\ \mathrm{d}P_t = (\mu_P - \gamma\, \sigma_P) P_t \mathrm{d}t + \sigma_P P_t \mathrm{d}\widehat{Z}_t, \quad P_t = \phi(t), \ t \in [-L, 0], \end{cases} \tag{B.15}$$

By Theorem B.2, point (i), the explicit expression of the solution to (B.15), which provides the discounted Bitcoin price $\widetilde{S}_t$, at any time $t \in [0, T]$, is given by

$$\widetilde{S}_t = s_0 e^{\sigma_S \int_0^t \sqrt{P_{u-\tau}} \mathrm{d}\widehat{W}_u - \frac{\sigma_S^2}{2} \int_0^t P_{u-\tau} \mathrm{d}u}, \quad t \in [0, T], \tag{B.16}$$

with the representation of the attention factor P still provided by (B.6) where the parameter μ_P is replaced by $\tilde{\mu}_P = \mu_P - \gamma\, \sigma_P$. All the outcomes on the integrated attention variable $X_{t,T}$ hold under the transformed probability measure $\mathbf{Q}^{\gamma}$ by replacing parameter μ_P with $\tilde{\mu}_P$.

In what follows we prove Lemma 3.1.

Proof of Lemma 3.1 By applying (Levy 1992) we have (see "Appendix A") that the distribution of $A_1 - X_{\tau}^{\tau}$ can be approximated by a log-normal with parameters

$$\alpha_1 = \log\left(\frac{\mathbb{E}[IP(\Delta-\tau)]^2}{\sqrt{\mathbb{E}[IP(\Delta-\tau)^2]}}\right), \qquad \nu_1^2 = \log\left(\frac{\mathbb{E}[IP(\Delta-\tau)^2]}{\mathbb{E}[IP(\Delta-\tau)]^2}\right).$$

By applying the outcomes of Lemma B.1, we have

$$\mathbb{E}^{\mathbf{Q}^{\gamma}}[A_1] = \mathbb{E}[P_0]\frac{e^{\tilde{\mu}_P(\Delta-\tau)}-1}{\tilde{\mu}_P} = P_0\frac{e^{\tilde{\mu}_P(\Delta-\tau)}-1}{\tilde{\mu}_P}$$
$$\mathbb{E}^{\mathbf{Q}^{\gamma}}[A_1^2] = P_0^2\left(\frac{2}{\tilde{\mu}_P+\sigma_P^2}\left[\frac{e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)(\Delta-\tau)}-1}{2\tilde{\mu}_P+\sigma_P^2} - \frac{e^{\tilde{\mu}_P(\Delta-\tau)}-1}{\tilde{\mu}_P}\right]\right).$$

Hence

$$\alpha_1 = \log\left(\frac{P_0\left(\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P}\right)^2}{\sqrt{\frac{2}{\tilde{\mu}_P+\sigma_P^2}\left[\frac{e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)\Delta}-1}{2\tilde{\mu}_P+\sigma_P^2}-\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P}\right]}}\right)$$
$$\nu_1^2 = \log\left(\frac{\frac{2}{\tilde{\mu}_P+\sigma_P^2}\left[\frac{e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)\Delta}-1}{2\tilde{\mu}_P+\sigma_P^2}-\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P}\right]}{\left(\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P}\right)^2}\right).$$

By applying simple computation, we get the outcomes for part (i). Moreover,

$$\mathbb{E}^{\mathbf{Q}^{\gamma}}[A_i] = \mathbb{E}^{\mathbf{Q}^{\gamma}}[P_{i-1}]\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P} = \mathbb{E}^{\mathbf{Q}^{\gamma}}[P_{i-2}]\,e^{\tilde{\mu}_P\Delta}\,\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P} = \mathbb{E}[A_{i-1}]\,e^{\tilde{\mu}_P\Delta}$$
$$\mathbb{E}^{\mathbf{Q}^{\gamma}}[A_i^2] = \mathbb{E}^{\mathbf{Q}^{\gamma}}[P_{i-1}^2]\left(\frac{2}{\tilde{\mu}_P+\sigma_P^2}\left[\frac{e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)\Delta}-1}{2\tilde{\mu}_P+\sigma_P^2}-\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P}\right]\right)$$
$$= \mathbb{E}^{\mathbf{Q}^{\gamma}}[P_{i-2}^2]e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)\Delta}\left(\frac{2}{\tilde{\mu}_P+\sigma_P^2}\left[\frac{e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)\Delta}-1}{2\tilde{\mu}_P+\sigma_P^2}-\frac{e^{\tilde{\mu}_P\Delta}-1}{\tilde{\mu}_P}\right]\right)$$
$$= \mathbb{E}^{\mathbf{Q}^{\gamma}}[A_{i-1}^2]\,e^{\left(2\tilde{\mu}_P+\sigma_P^2\right)\Delta}.$$

Then

$$\begin{aligned}\alpha_i &= \log\left(\frac{\mathbb{E}^{\mathbf{Q}^\gamma}[A_i]^2}{\sqrt{\mathbb{E}^{\mathbf{Q}^\gamma}[A_i^2]}}\right) = \log\left(\frac{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}]^2\left(e^{\tilde{\mu}_P\Delta}\right)^2}{\sqrt{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}^2]\,e^{(2\tilde{\mu}_P+\sigma_P^2)\Delta}}}\right)\\ &= \log\left(\frac{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}]^2}{\sqrt{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}^2]}}\right) + \log\left(\frac{\left(e^{\tilde{\mu}_P\Delta}\right)^2}{\sqrt{e^{(2\tilde{\mu}_P+\sigma_P^2)\Delta}}}\right) = \alpha_{i-1} + \left(\tilde{\mu}_P - \frac{\sigma_P^2}{2}\right)\Delta,\\ \nu_i^2 &= \log\left(\frac{\mathbb{E}^{\mathbf{Q}^\gamma}[A_i^2]}{\sqrt{\mathbb{E}^{\mathbf{Q}^\gamma}[A_i]^2}}\right) = \log\left(\frac{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}^2]\,e^{(2\tilde{\mu}_P+\sigma_P^2)\Delta}}{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}]^2\,e^{(\tilde{\mu}_P\Delta)^2}}\right)\\ &= \log\left(\frac{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}^2]}{\sqrt{\mathbb{E}^{\mathbf{Q}^\gamma}[A_{i-1}]^2}}\right) + \log\left(\frac{e^{(2\tilde{\mu}_P+\sigma_P^2)\Delta}}{e^{(\tilde{\mu}_P\Delta)^2}}\right) = \nu_{i-1}^2 + \sigma_P^2\Delta.\end{aligned}$$

Conditioning to A_{i-1}, we get

$$\begin{aligned}\alpha_i &= \log(A_{i-1}) + \left(\tilde{\mu}_P - \frac{\sigma_P^2}{2}\right)\Delta,\\ \nu_i^2 &= \sigma_P^2\Delta,\end{aligned}$$

which gives part (ii). □

Appendix C: Finite sample behavior of QML estimates

In order to check the goodness of the log-likelihood approximation introduced in Theorem 3.2, we apply the proposed estimation method to simulated data and assume, for the sake of simplicity, $\tau = 0$. We simulate m samples of length n for the processes in (2.1) and (2.2) assuming a constant finer observation step δ; we extract corresponding samples for $\mathbf{R}$, $\mathbf{A}$ at a lower frequency, with observation step $\Delta = r\delta$. In the numerical exercise we choose $n = 730$, $m = 1000$, $\delta = \frac{1}{365}$ (daily observations), $\Delta = 7\delta$ (weekly observations); parameters values are set as $\mu_P = 2$, $\sigma_P = 0.5$ $\mu_S = 0.05$, $\sigma_S = 0.3$. We end with 1000 samples of 104 observations for $(\mathbf{R}, \mathbf{A})$; for each sample we estimate the parameters by means of the quasi-maximum likelihood as suggested in previous subsection. The results are summed up in Table 11.

We also performed a t test in order to check for estimation bias. The fitted values of μ_P, μ_S, σ_S are close in mean to their theoretical value and with a reasonable standard deviation; the p-values of the t test confirm that estimated are not biased. Different conclusions are in order as for parameter σ_P which estimations is by no doubt biased. In Fig. 2 we plot the histograms of the estimated as well as the fitted normal distribution and the expected mean of the asymptotic distribution. Pictures confirm the biasedness of the estimator for σ_P but all other estimates perform well and outcomes may become better by increasing the sample length. The simulation exercise have been repeated

Table 11 Parameter fit with simulated data of QML method

Variable	Theor. value	Fitted value	SE	t value	$P(>\|t\|)$	RMSE
μ_P	2.0000	1.9759	0.3675	−0.0655	0.9478	11.6475
σ_P	0.5000	0.4089	0.0302	−3.0170	0.0026	3.0358
μ_S	0.0500	0.0497	0.0112	−0.0297	0.9763	0.3544
σ_S	0.3000	0.2978	0.0210	−0.1032	0.9178	0.6687

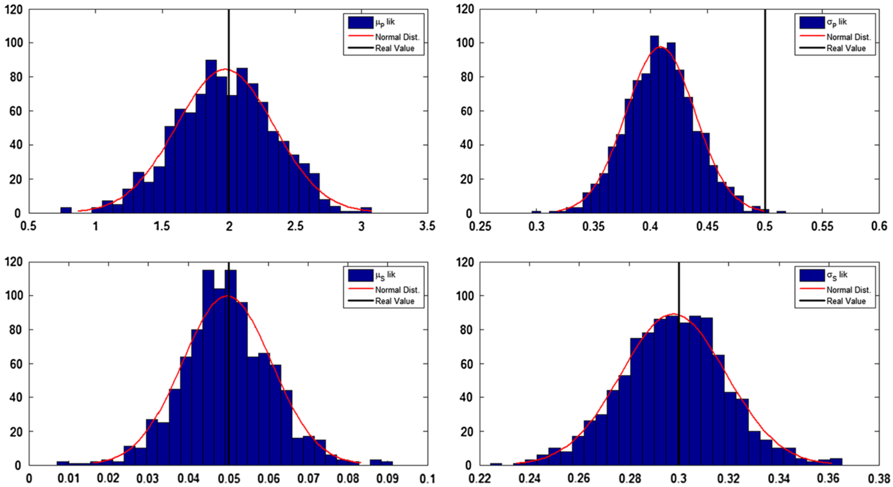

Fig. 2 Histogram of parameter fit with simulated data of QML method

Table 12 Parameter fit with simulated data of Moments method

Variable	Theor. value	Fitted value	SE	t value	$P(>\|t\|)$	RMSE
μ_P	2.0000	2.0104	0.3638	0.0286	0.9772	11.5034
σ_P	0.5000	0.4995	0.0135	−0.0385	0.9693	0.4282

by letting the parameters values, the number and the sample length vary obtaining analogous qualitative results.

In order to disentangle the contribution of the Levy approximation (Levy 1992) to the estimation bias, we suggest to apply the method of moments to estimate μ_P and σ_P considering the whole sample for P generated at the finer observation step δ to compute the sample mean and sample variance of the attention realizations. If we then we plug the estimated values in the likelihood (3.2) in order to estimate $\{\mu_S, \sigma_S\}$ these two estimates remain unchanged; in fact the likelihood may be maximized separately with respect to μ_P, σ_P and μ_S, σ_S since each of the two addend in the likelihood expression depends on just one of this pairs. The results of this alternative estimation method are reported in Table 12.

To visualize the bias of $\{\sigma_P\}$ we plot in Fig. 3 the histogram of parameter $\{\mu_P, \sigma_P\}$ fit with simulated data using the two methods. As we can see using the two step

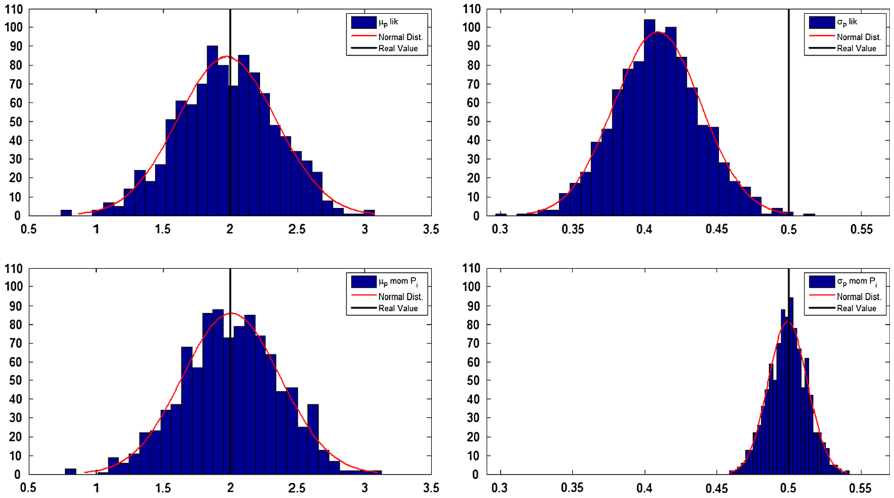

Fig. 3 Histogram of parameter $\{\mu_P, \sigma_P\}$ fit with simulated data of QML method and Moments method at finer step δ

procedure we obtain better estimates of $\{\sigma_P\}$ both in terms of expected value and standard deviation. It is evident from Table 12 and Fig. 3 that the estimation of σ_P is not biased in this case hence the estimation bias may essentially be attributed to the aggregation of the attention over time intervals and to the corresponding approximating distribution. Hence, whether the attention factor is observed at a finer step than the price, the above separate estimation is more reliable.

References

Barber, B.M., Odean, T.: All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. **21**(2), 785–818 (2007)

Bistarelli, S., Cretarola, A., Figà-Talamanca, G., Mercanti, I., Patacca, M.: Is arbitrage possible in the bitcoin market? In: Coppola, M., Carlini, E., D'Agostino, D., Altmann, J., Bañares, J.Á.: editors, Economics of Grids, Clouds, Systems, and Services—15th International Conference, GECON 2018, Pisa, Italy, September 18–20, 2018. Springer International Publishing. https://doi.org/10.1007/978-3-030-13342-9_21 (2018)

Bistarelli, S., Cretarola, A., Figà-Talamanca, G., Patacca, M.: Model-based arbitrage in multi-exchange models for Bitcoin price dynamics Digit. Finance (2019a). https://doi.org/10.1007/s42521-019-00001-2

Bistarelli, S., Figà-Talamanca, G., Lucarini, F., Mercanti, I.: Studying forward looking bubbles in Bitcoin/USD exchange rates. In: Proceedings of the 23rd International Database Applications & Engineering Symposium. ACM (2019b)

Black, F., Scholes, M.: The pricing of options and corporate liabilities. J. Polit. Econ. pp. 637–654 (1973)

Bukovina, J., Martiček, M.: Sentiment and bitcoin volatility. Technical report, Mendel University in Brno, Faculty of Business and Economics (2016)

Catania, L., Grassi, S.: Modelling crypto-currencies financial time-series. CEIS Working Paper, (2017)

Chu, J., Nadarajah, S., Chan, S.: Statistical analysis of the exchange rate of bitcoin. PLoS ONE **10**(7), e0133678 (2015)

Corbet, S., Lucey, B., Yarovaya, L.: Datestamping the Bitcoin and Ethereum bubbles. Finance Res. Lett. **26**, 81–88 (2018)

Cretarola, A., Figà-Talamanca, G., Patacca, M.: A continuous time model for Bitcoin price dynamics. In: Corazza, M., Durbán, M., Grané, A., Perna, C., Sibillo, M. editors, Mathematical and Statistical Methods for Actuarial Sciences and Finance - MAF 2018, pp. 273–277. Springer International Publishing, https://doi.org/10.1007/978-3-319-89824-7_49 (2018)
Cretarola, A., Figà-Talamanca, G.: Detecting bubbles in Bitcoin price dynamics via market exuberance. Ann. Oper. Res. (2019). https://doi.org/10.1007/s10479-019-03321-z
Da, Z., Engelberg, J., Gao, P.: In search of attention. J. Finance **66**(5), 1461–1499 (2011)
Davison, A.C.: Statistical Models, vol. 11. Cambridge University Press, Cambridge (2003)
Donier, J., Bouchaud, J.-P.: Why do markets crash? Bitcoin data offers unprecedented insights. PLoS ONE **10**(10), e0139356 (2015)
Figà-Talamanca, G., Patacca, M.: Does market attention affect Bitcoin returns and volatility? Decisions Econ. Finan. (2019). https://doi.org/10.1007/s10203-019-00258-7
Föllmer, H., Schweizer, M.: Hedging of contingent claims under incomplete information. In: Davis, M.H.A., Elliot, R.J. (eds.) Applied Stochastic Analysis. volume 5, pp. 389–414. Gordon and Breach, New York (1991)
Föllmer, H., Schweizer, M.: Minimal martingale measure. In: Encyclopedia of Quantitative Finance, Wiley Online Library (2010)
Fry, J., Cheah, E.-T.: Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of bitcoin. Econ. Lett. **130**, 32–36 (2015)
Gervais, S., Kaniel, R., Mingelgrin, D.H.: The high-volume return premium. J. Finance **56**(3), 877–919 (2001)
Gourieroux, C., Monfort, A., Trognon, A.: Pseudo maximum likelihood methods: theory. Econometrica **52**(3), 681–700 (1984)
Guo, L., Li, XJ.: Risk analysis of cryptocurrency as an alternative asset class. In: Applied Quantitative Finance, pp. 309–329. Springer (2017)
Hou, K., Xiong, W., Peng, L.: A tale of two anomalies: The implications of investor attention for price and earnings momentum. SSRN Electr. J. (2009)
Hull, J., White, A.: The pricing of options on assets with stochastic volatilities. J. Finance **42**(2), 281–300 (1987)
Kim, Y.B., Lee, S.H., Kang, S.J., Choi, M.J., Lee, J., Kim, C.H.: Virtual world currency value fluctuation prediction system based on user sentiment analysis. PLoS ONE **10**(8), e0132944 (2015)
Kou, S.G.: A jump-diffusion model for option pricing. Manage. Sci. **48**(8), 1086–1101 (2002)
Kristoufek, L.: BitCoin meets Google trends and Wikipedia: quantifying the relationship between phenomena of the internet era. Sci. Rep. **3**, 3415 (2013)
Kristoufek, L.: What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis. PLoS ONE **10**(4), e0123923 (2015)
Levy, E.: Pricing European average rate currency options. J. Int. Money Finance **11**(5), 474–491 (1992)
Malhotra, A., Maloo, M.: Bitcoin-is it a bubble? Evidence from unit root tests. SSRN Electr. J. (2014)
Mao, X., Sabanis, S.: Delay geometric Brownian motion in financial option valuation. Stoch. Int. J. Probab. Stoch. Process. **85**(2), 295–320 (2013)
Massey Jr., F.J.: The Kolmogorov–Smirnov test for goodness of fit. J. Am Stat. Assoc. **46**(253), 68–78 (1951)
Milevsky, M.A., Posner, S.E.: Asian options, the sum of lognormals, and the reciprocal gamma distribution. J. Financ. Quant. Anal. **33**(03), 409–422 (1998)
Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. Working Paper (2008)
Pascucci, A.: PDE and Martingale Methods in Option Pricing. Springer Science & Business Media, New York (2011)
Pawitan, Y.: In All Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford University Press, Oxford (2001)
Protter, P.E.: Stochastic Integration and Differential Equations, volume 21 of Stochastic Modelling and Applied Probability. Springer, Berlin, 3rd corrected printing, 2nd edn (2005)
The Wall Street Journal. CBOE Teams Up with Winklevoss Twins for Bitcoin Data. https://www.wsj.com/articles/cboe-teams-up-with-winklevoss-twins-for-bitcoin-data-1501675200 (2017a)
The Wall Street Journal. Bitcoin Options Exchange Wins Approval from CFTC. https://www.wsj.com/articles/bitcoin-options-exchange-wins-approval-from-cftc-1500935886 (2017b)
Tsay, R.S.: Analysis of Financial Time Series, vol. 543. Wiley, New York (2005)
White, H.: Maximum likelihood estimation of misspecified models. Econometrica, pp. 1–25, (1982)

Yermack, D.: Is bitcoin a real currency? An economic appraisal. In: Handbook of Digital Currency, chapter second, pp. 31–43. Elsevier, Amsterdam (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Alessandra Cretarola[1] · Gianna Figà-Talamanca[2] · Marco Patacca[3,4]

Gianna Figà-Talamanca
gianna.figatalamanca@unipg.it

Marco Patacca
marco.patacca@devinci.fr

1 Department of Mathematics and Computer Science, University of Perugia, Via Luigi Vanvitelli, 1, 06123 Perugia, Italy

2 Department of Economics, University of Perugia, Via Alessandro Pascoli, 06123 Perugia, Italy

3 Léonard de Vinci Pôle Universitaire, Research Center, 92916 Paris La Défense Cedex, France

4 Department of Economics, University of Perugia, Via Alessandro Pascoli, 06123 Perugia, Italy