Check for updates

J. Appl. Probab. **63,** 707–728 (2026)
*doi:*10.1017/jpr.2025.10048

FINITE MIXTURE MODELS FOR OPTION PRICING: AN APPLICATION TO BITCOIN OPTIONS

TAK KUEN SIU ,* *Macquarie Business School, Macquarie University*

Abstract

This paper considers option valuation under finite mixture models in a discrete-time economy. Specifically, the Esscher transform is employed to select a pricing kernel. Novel finite mixture models with negative-shifted Gamma and negative-shifted inverse Gaussian distributions are developed. A hybrid finite mixture model that allows different parametric forms for component distributions is introduced to incorporate model uncertainty. An empirical characteristic function estimation method is employed to estimate the finite mixture models. Closed-form pricing formulas for a European call option are obtained for some finite mixture models. Empirical examples using data on the Bitcoin-USD prices are provided to illustrate an application of the proposed models to value Bitcoin options.

Keywords: Option valuation; finite mixture distributions; Esscher transform; empirical characteristic function estimation; Bitcoin options

2020 Mathematics Subject Classification: Primary 91G05
Secondary 91G20; 91G70

1. Introduction

Finite mixture models, such as finite mixture normals, play an important role in modelling empirical behaviour of financial returns. Empirical evidence supporting a mixture of normal distributions for S&P 500 returns data was provided in [60]. The significance of finite mixture models for modelling financial returns was discussed in [73, 80]. The merits of finite mixture models include their analytical tractability and capability of capturing key empirical features of financial returns, such as the skewness and heavy-tailedness of distributions for returns. Finite mixture models also incorporate the impact of uncertainty about market news or events on distributions for financial returns as noted, for example, in [25, 73, 74]. Finite mixture models have important applications in machine learning and big data analytics [8, 20]. For an in-depth statistical analysis of finite mixture models, see, for example, [33, 50, 51, 75]. Finite mixture models are intimately linked with the regime-switching models proposed, for example, in [34, 76, 77]. Specifically, if the finite-state (hidden) Markov chain modulating a Markov regime-switching model is replaced by a mixing random variable, the Markov regime-switching model becomes a finite mixture model. [78] pointed out that the stationary distribution of a

Received 20 August 2025; accepted 20 August 2025.
* Postal address: Department of Actuarial Studies and Business Analytics, Macquarie Business School, Macquarie University, Sydney, NSW 2109, Australia. Email: Ken.Siu@mq.edu.au

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

simple self-exciting threshold autoregressive (SETAR) model is a mixture of normal distributions. See also [23] for exploring links between finite mixture models and regime-switching models.

Option valuation under finite mixture models has been considered in the literature. [57] discussed the pricing of a European call option under a finite mixture of normal distributions and obtained a closed-form pricing formula. With heterogeneous expectations, [32] obtained the Ritchey pricing formula and showed that a finite mixture of lognormal distributions is desirable for recovering a risk-neutral distribution from option prices data. See [7, 47, 58, 59, 81] for some other contributions to option pricing and related problems such as risk-neutral density estimation under finite mixture models. Option valuation under discrete-time regime-switching models has been considered in the literature. [15] proposed a general class of option pricing models based on regime-switching processes, which include the Markov regime-switching models and some GARCH-type models. [42, 45] studied option valuation under discrete-time hidden Markov regime-switching models. [19, 71] considered option pricing under a discrete-time Markov-switching GARCH-type model and a discrete-time SETAR-ARCH model, respectively. [14] considered the pricing of exotic options under a discrete-time, high-order, Markov regime-switching model. [18] introduced a discrete-time, Markov regime-switching, affine term-structure model for pricing interest-rate securities. [68] discussed option pricing under discrete-time doubly Markov regime-switching models with non-normal innovations. [29] proposed a novel approach to deal with path dependence in option valuation under a discrete-time hidden Markov regime-switching model. The approach was improved by [30] based on the extended Girsanov principle allowing for non-normal logarithmic returns. [24] studied an option valuation problem under a discrete-time Markov-switching stochastic volatility model. The aforementioned works focus on models with normal innovations, apart from [30, 65].

Discrete-time finite mixture models offer a simpler solution to an option valuation problem than discrete-time Markov regime-switching models. Fusing techniques in actuarial science, financial mathematics, applied probability, statistics, econometrics, and data science, this paper discusses canonical option valuation under discrete-time finite mixture models. Specifically, the Esscher transform is employed to select a pricing kernel. The Esscher transform is an important tool in actuarial science and its use for option pricing was proposed in the seminal work [26]. For some subsequent developments, see, for example, [3, 11, 12, 16, 27, 28, 31, 62, 70]. The pricing kernel selected by the Esscher transform is justified by maximizing the expected utility based on a state-dependent power utility function. Besides finite mixture normals, we introduce novel finite mixture models including a finite mixture model of negative-shifted Gamma (NegSGa) distributions, a finite mixture model of negative-shifted inverse Gaussian (NegSIG) distributions, as well as a hybrid finite mixture model of normal, NegSGa, and NegSIG distributions. The hybrid finite mixture model incorporates model uncertainty by allowing different parametric forms for its component distributions. The maximum likelihood estimation method is intractable when NegSGa and NegSIG distributions are involved. Given that closed-form expressions for the characteristic functions of normal, NegSGa, and NegSIG distributions are available, an empirical characteristic function (ECF) estimation method is adopted to estimate the finite mixture models. Closed-form pricing formulas for a European call option are obtained for these finite mixture models, up to solving non-linear equations for cases involving NegSIG distributions. Although works on option valuation under mixture normal GARCH-type models exist (see, for example, [4]) and have gained some empirical successes, the finite mixture models considered here lead to closed-form pricing formulas and

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

provide a flexible way to incorporate non-normality. Using data on the daily adjusted close prices of Bitcoin-USD, estimation results for eight models (five finite mixture models and three non-mixture models) are provided. Numerical results are provided and analysed for European call Bitcoin option prices and their implied volatilities.

The rest of the paper is structured as follows. In Section 2, a discrete-time multiperiod finite mixture model is considered. The theoretical results for option valuation under the finite mixture model are presented based on the conditional Esscher transform in Section 3. Analytical solutions to option valuation are obtained for the finite mixture models in Section 4. The estimation procedures and results are provided in Section 5. Numerical results for option prices and implied volatilities are presented and analysed in Section 6. The final section provides concluding remarks and points out some limitations of the proposed models and study. The supplemental material contains Appendix A, an economic justification for the pricing kernel; Appendix B, proofs of lemmas and theorems; Appendix C, exploratory data analysis; and Appendix D, tables and figures.

2. A discrete-time multiperiod model

A discrete-time economy is considered in which uncertainties are described by a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ is a real-world probability measure. Let $\mathcal{T}$ be a finite discrete-time parameter set of the economy, where $\mathcal{T} = \{0, 1, 2, \ldots, T\}$, for some finite horizon $T < \infty$. The price process of a risky asset is described by a discrete-time stochastic process $\{S_t\}_{t\in\mathcal{T}}$ with the state space $\mathbb{R}_+$, (i.e. the set of non-negative real numbers). For each $t = 1, 2, \ldots, T$, let $R_t := \ln(S_t/S_{t-1})$, (i.e. the logarithmic return of the risky asset in the tth period from time $t-1$ to time t).

For each $t = 1, 2, \ldots, T$, R_t takes on values in $\mathbb{R}$, (i.e. the set of real numbers). Write $\mathbb{F}$ for the $\mathbb{P}$-augmentation of the natural filtration $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ generated by $\{R_t\}_{t\in\mathcal{T}\setminus\{0\}}$ such that $\mathcal{F}_t := \sigma\{R_1, R_2, \ldots, R_t\} \vee \mathcal{N}$ and $\mathcal{F}_0 := \sigma\{\varnothing, \Omega\} \vee \mathcal{N}$, where $\mathcal{N}$ is a collection of $\mathbb{P}$-null sets in $\mathcal{F}$, $\mathcal{A} \vee \mathcal{B}$ is the minimal σ-field containing the σ-fields $\mathcal{A}$ and $\mathcal{B}$, and $\varnothing$ is the empty set. Here it is supposed that for each $t = 1, 2, \ldots, T$, R_t has a finite mixture distribution with the cumulative distribution function (CDF) $F(x)$, which is given as follows:

$$F(x) = \sum_{i=1}^{N} p_i F_i(x), \tag{2.1}$$

where $p_i > 0$ for $i = 1, 2, \ldots, N$, and $\sum_{i=1}^{N} p_i = 1$; $F_i(x)$ is the ith-component CDF of the mixture CDF $F(x)$; and p_i is the weight allocated to the ith-component CDF $F_i(x)$. Note that $R_1, R_2, \ldots, R_T$ are identically distributed and have the common finite mixture distribution $F(x)$ defined by (2.1).

Let Θ denote a discrete random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ taking a value on $\{1, 2, \ldots, N\}$. Suppose that the probability mass function of Θ under $\mathbb{P}$ is given by

$$\mathbb{P}(\Theta = i) = p_i, \quad i = 1, 2, \ldots, N. \tag{2.2}$$

Assume that, for each $t \in \mathcal{T} \setminus \{0\}$ and each $i = 1, 2, \ldots, N$, the conditional distribution of R_t given $\Theta = i$ under $\mathbb{P}$ is given by

$$\mathbb{P}(R_t \le x \mid \Theta = i) = F_i(x) \quad \text{for each } x \in \mathbb{R}, \tag{2.3}$$

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

where $F_i(x)$ is the ith-component CDF of the finite mixture CDF $F(x)$ in (2.1). Using the law of total probability, (2.2), and (2.3), for each $t \in \mathcal{T} \setminus \{0\}$, the unconditional CDF of R_t under $\mathbb{P}$ is given by

$$F(x) = \mathbb{P}(R_t \le x) = \sum_{i=1}^{N} \mathbb{P}(R_t \le x \mid \Theta = i)\mathbb{P}(\Theta = i) = \sum_{i=1}^{N} F_i(x)p_i. \tag{2.4}$$

This is the same as the CDF $F(x)$ of the mixture distribution in (2.1). Consequently, the random variable Θ is interpreted as a mixing random variable for the mixture distribution in (2.1).

For each $t \in \mathcal{T} \setminus \{0\}$, let $M(\eta)$ be the moment-generating function (MGF) of R_t evaluated at η under $\mathbb{P}$ and $M_i(\eta)$ be the (conditional) MGF of R_t given $\Theta = i$ evaluated at η under $\mathbb{P}$. That is,

$$M(\eta) := \mathbb{E}[\mathrm{e}^{\eta R_t}] = \int_{-\infty}^{\infty} \mathrm{e}^{\eta x}\, \mathrm{d}F(x), \tag{2.5}$$

where $\mathbb{E}[\,\cdot\,]$ denotes the expectation under $\mathbb{P}$, and

$$M_i(\eta) := \mathbb{E}[\mathrm{e}^{\eta R_t} \mid \Theta = i] = \int_{-\infty}^{\infty} \mathrm{e}^{\eta x}\, \mathrm{d}F_i(x), \tag{2.6}$$

where $\mathbb{E}[\,\cdot \mid \Theta = i]$ is the conditional expectation given $\Theta = i$ under $\mathbb{P}$. Then it is known that

$$M(\eta) = \sum_{i=1}^{N} p_i M_i(\eta). \tag{2.7}$$

3. Pricing kernel using the conditional Esscher transform

In this section we specify a risk-neutral probability measure (or a pricing kernel) using a version of the conditional Esscher transform [11, 12, 70]. We also identify the respective martingale condition, and the distributions for the returns and mixing random variable under the risk-neutral probability measure. Lastly, we derive a generic expression for the price of a European call option.

Let $h\colon \{1, 2, \ldots, N\} \to \mathbb{R}$ be a real-valued function on the finite set $\{1, 2, \ldots, N\}$. Assume that, for each $i \in \{1, 2, \ldots, N\}$, $M_i(h(i)) < \infty$ (i.e. the conditional MGF of R_t given $\Theta = i$ evaluated at $h(i) \in \mathbb{R}$ under $\mathbb{P}$ exists).

For each $t = 1, 2, \ldots, T$ and $i \in \{1, 2, \ldots, N\}$, let

$$\lambda_t(h(i)) = \frac{\mathrm{e}^{h(i)R_t}}{M_i(h(i))}. \tag{3.1}$$

Then $\lambda_t(h(i))$ is a positive $\mathcal{F}_t$-measurable random variable.

For each $t = 1, 2, \ldots, T$, let $\mathcal{G}_t$ be an enlarged σ-field defined by

$$\mathcal{G}_t := \mathcal{F}_t \vee \sigma\{\Theta\}, \tag{3.2}$$

where $\sigma\{\Theta\}$ denotes the σ-algebra generated by the mixing random variable Θ. Write $\mathbb{G}$ for the respective enlarged filtration $\{\mathcal{G}_t\}_{t\in\mathcal{T}}$ with $\mathcal{G}_0 := \mathcal{F}_0 = \sigma\{\varnothing, \Omega\} \vee \mathcal{N}$.

Similarly to (2.6), let $M_\Theta(\eta)$ denote the conditional MGF of R_t given Θ evaluated at η under $\mathbb{P}$. That is,

$$M_\Theta(\eta) := \mathbb{E}[\mathrm{e}^{\eta R_t} \mid \Theta] = \int_{-\infty}^{\infty} \mathrm{e}^{\eta x}\, \mathrm{d}F_\Theta(x), \tag{3.3}$$

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

where $\mathbb{E}[\,\cdot \mid \Theta]$ is the conditional expectation given Θ under $\mathbb{P}$. Note that $M_\Theta(\eta)$ may be thought of as the (random) MGF of the (random) distribution function $F_\Theta(x)$ under $\mathbb{P}$. For each $i = 1, 2, \ldots, N$, on the set $\{\omega \in \Omega \mid \Theta(\omega) = i\}$, $F_\Theta(x) = F_i(x)$ and $M_\Theta(\eta) = M_i(\eta)$ under $\mathbb{P}$.

Similarly to (3.1), for each $t = 1, 2, \ldots, T$ let

$$\lambda_t(h(\Theta)) = \frac{\mathrm{e}^{h(\Theta)R_t}}{M_\Theta(h(\Theta))}. \tag{3.4}$$

Then $\lambda_t(h(\Theta))$ is a positive $\mathcal{G}_t$-measurable random variable.

We define the $\mathbb{G}$-adapted process $\{\Lambda_t^h(\Theta)\}_{t\in\mathcal{T}}$ by putting

$$\begin{aligned}\Lambda_t^h(\Theta) &:= \prod_{k=1}^{t} \lambda_k(h(\Theta)) \\ &= \prod_{k=1}^{t} \frac{\mathrm{e}^{h(\Theta)R_k}}{M_\Theta(h(\Theta))} = \exp\left(h(\Theta)\sum_{k=1}^{t} R_k - t\ln\left(M_\Theta(h(\Theta))\right)\right), \quad t \in \mathcal{T}\setminus\{0\},\end{aligned} \tag{3.5}$$

and $\Lambda_0^h(\Theta) = 1$.

We impose the following assumption.

Assumption 3.1. *For each $t = 2, 3, \ldots, T$, R_t is conditionally independent of $\mathcal{F}_{t-1}$ given Θ under $\mathbb{P}$.*

Then we have the following lemma.

Lemma 3.1. *Suppose that Assumption 3.1 holds. Then $\{\Lambda_t^h(\Theta)\}_{t\in\mathcal{T}}$ is a positive $(\mathbb{G}, \mathbb{P})$-martingale.*

A probability measure $\mathbb{P}^h$ equivalent to $\mathbb{P}$ on $\mathcal{G}_T$ can then be defined by putting

$$\left.\frac{\mathrm{d}\mathbb{P}^h}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_T} := \Lambda_T^h(\Theta). \tag{3.6}$$

The probability measure $\mathbb{P}^h$ defined in (3.6) is a variant of the Esscher transform. It is intimately linked with the random Esscher transform in [69], and may be thought of as a version of the conditional Esscher transform in discrete time [11, 12, 70]. It is also related to the regime-switching Esscher transform in [16, 45, 68]. The random variable $h(\Theta)$ is interpreted as a random Esscher parameter. In the context of insurance mathematics, we can select an Esscher parameter so that, for a given premium level, the Esscher premium principle with the Esscher parameter is equal to the given premium level. The followng lemma gives a result that will be used for later developments in this paper.

Lemma 3.2. *Suppose that Assumption 3.1 holds. Then $R_1, R_2, \ldots, R_T$ are conditionally independent given Θ under $\mathbb{P}$. Furthermore, $R_1, R_2, \ldots, R_T$ are also conditionally independent given Θ under $\mathbb{P}^h$ defined by (3.6).*

Let $M_\Theta^h(\eta)$ be the conditional MGF of R_t given Θ evaluated at η under the probability measure $\mathbb{P}^h$ defined by (3.6). That is,

$$M_\Theta^h(\eta) := \mathbb{E}^h[\mathrm{e}^{\eta R_t} \mid \Theta], \tag{3.7}$$

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

where $\mathbb{E}^h[\,\cdot \mid \Theta]$ is the conditional expectation given Θ under $\mathbb{P}^h$. Then the following lemma gives the relation between $M^h_\Theta(\eta)$ and $M_\Theta(\eta)$. See, for example, [19, 26, 27, 45, 60] for some related discussions.

Lemma 3.3. *Let $\bar{\eta}\colon \{1, 2, \ldots, N\} \to \mathbb{R}$ be a measurable real-valued function. Then, for each $t \in \mathcal{T} \setminus \{0\}$,*

$$M^h_\Theta(\bar{\eta}(\Theta)) = \frac{M_\Theta(\bar{\eta}(\Theta) + h(\Theta))}{M_\Theta(h(\Theta))}, \quad \mathbb{P}\text{-almost surely.} \tag{3.8}$$

Let r be the constant risk-free force of interest, i.e. the continuously compounded risk-free interest rate. By the fundamental theorem of asset pricing [36–38], the absence of arbitrage opportunities is essentially equivalent to the existence of a martingale measure, say $\mathbb{P}^{h^\dagger}$, equivalent to $\mathbb{P}$ on $\mathcal{G}_T$, such that the price process of the risky asset discounted by the risk-free force of interest r, say $\{e^{-rt}S_t\}_{t\in\mathcal{T}}$, is a positive $(\mathbb{G}, \mathbb{P}^{h^\dagger})$-martingale. The latter statement holds if and only if there exists a random variable $h^\dagger := h^\dagger(\Theta)$, which is $\sigma\{\Theta\}$-measurable, such that, for each $t = 1, 2, \ldots, T$,

$$\mathbb{E}^{h^\dagger}[e^{-rt}S_t \mid \mathcal{G}_{t-1}] = e^{-r(t-1)}S_{t-1}, \tag{3.9}$$

where $\mathbb{E}^{h^\dagger}[\,\cdot \mid \mathcal{G}_{t-1}]$ is the conditional expectation given $\mathcal{G}_{t-1}$ under $\mathbb{P}^{h^\dagger}$. Note that the martingale condition for the discounted share price process is with respect to the enlarged filtration $\mathbb{G}$, which contains information about the share price process and the mixing random variable Θ. The latter may be thought of as private information that is not known to the public. Consequently, the assumption of the strong form of market information efficiency is imposed. It may be noted that some other studies considered the weak form of market infomation efficiency instead. See, for example, [6, 17, 30].

The random variable $h^\dagger(\Theta)$ is interpreted as the risk-neutral random mixing Esscher parameter. In the following theorem, the fundamental theorem of asset pricing is adapted to the current modelling framework based on the possible values taken by the risk-neutral random mixing Esscher parameter $h^\dagger(\Theta)$. These possible values are given by $h^\dagger(i)$ for $i = 1, 2, \ldots, N$.

Theorem 3.1. *Suppose an equivalent martingale measure is specified by the probability measure $\mathbb{P}^{h^\dagger}$ defined in (3.6) by setting $h = h^\dagger$. Then the absence of arbitrage opportunities is essentially equivalent to the existence of N real numbers $h^\dagger(i)$, $i = 1, 2, \ldots, N$, such that*

$$e^r = \sum_{i=1}^{N} \frac{M_i(1 + h^\dagger(i))}{M_i(h^\dagger(i))} \mathbf{1}_{\{\Theta=i\}}, \quad \mathbb{P}\text{-almost surely,} \tag{3.10}$$

where $\mathbf{1}_{\{\Theta=i\}}$ is the indicator function of the event $\{\Theta = i\}$.

By slightly modifying the arguments in [27], it may be seen that the N real numbers $h^\dagger_t(i)$, $i = 1, 2, \ldots, N$, that satisfy (3.10) are uniquely determined if they exist. The existence of N real numbers $h^\dagger(i)$, $i = 1, 2, \ldots, N$, that satisfy the martingale condition in (3.10) for some parametric cases will be discussed in Section 4. The pricing kernel selected by the conditional Esscher transform is justified by maximizing the expected state-dependent power utility in the supplemental material (Appendix A).

The following theorem shows that the probability mass function of the mixing random variable Θ remains unchanged when changing the probability measures from $\mathbb{P}$ to $\mathbb{P}^{h^\dagger}$.

Theorem 3.2. *Under $\mathbb{P}^{h^\dagger}$, the mixing random variable Θ has the probability mass function*

$$\mathbb{P}^{h^\dagger}(\Theta = i) = p_i, \quad i = 1, 2, \ldots, N. \tag{3.11}$$

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

Theorem 3.3 provides the conditional distribution of R_t given $\Theta = i$ under $\mathbb{P}^{h^\dagger}$ for each $i = 1, 2, \ldots, N$, and the unconditional distribution of R_t under $\mathbb{P}^{h^\dagger}$.

Theorem 3.3. *For each $t \in \mathcal{T} \setminus \{0\}$ and each $i = 1, 2, \ldots, N$, let $F_i^{h^\dagger}(x)$ denote the conditional probability distribution of R_t given $\Theta = i$ under $\mathbb{P}^{h^\dagger}$, and let $F^{h^\dagger}(x)$ denote the unconditional probability distribution of R_t under $\mathbb{P}^{h^\dagger}$. Then*

$$F_i^{h^\dagger}(x) := \mathbb{P}^{h^\dagger}(R_t \le x \mid \Theta = i) = \frac{\int_{-\infty}^{x} \mathrm{e}^{h^\dagger(i)y} F_i(\mathrm{d}y)}{M_i(h^\dagger(i))}, \tag{3.12}$$

where the integral is interpreted as the Riemann–Stieltjes integral. Furthermore,

$$F^{h^\dagger}(x) := \mathbb{P}^{h^\dagger}(R_t \le x) = \sum_{i=1}^{N} F_i^{h^\dagger}(x) p_i. \tag{3.13}$$

Besides using the conditional Esscher transform to specify a pricing kernel, we can adopt other prominent approaches such as the extended Girsanov principle [17] and the second-order Esscher transform [54] to specify a pricing kernel. However, using the conditional Esscher transform here has several advantages. Firstly, in each of the (parametric) finite mixture distributions to be considered in Section 4, the distribution for returns under the risk-neutral probability measure specified by the conditional Esscher transform is in the same parametric family as the distribution for returns under the real-world probability measure. That is, the (parametric) distribution of returns is closed under changing probability measures via the conditional Esscher transform. Indeed, in continuous-time regime-switching modelling frameworks, it has been shown in some literature that the risk-neutral probability measure selected by the Esscher transform minimizes the relative entropy between an equivalent martingale measure and the real-world probability measure [16, 63]. Note that the relative entropy between an equivalent martingale measure and the real-world probability measure is a measure for the 'distance' between the two probability measures. Consequently, the risk-neutral probability measure selected by the Esscher transform is 'closest' to the real-world probability measure. Secondly, using the conditional Esscher transform, an analytical expression for the price of a European call option is obtained for each of the finite mixture distributions to be considered in Section 4. Thirdly, the pricing kernel specified by the conditional Esscher transform can be justified by an economic equilibrium based on maximization of the expected power utility.

Note that the martingale condition in (3.10) of Theorem 3.1 is equivalent to the following N equations satisfied by $h^\dagger(i)$, $i = 1, 2, \ldots, N$:

$$\mathrm{e}^r = \frac{M_i(1 + h^\dagger(i))}{M_i(h^\dagger(i))}, \quad i = 1, 2, \ldots, N. \tag{3.14}$$

Furthermore, from Lemma 3.2, $R_1, R_2, \ldots, R_T \mid \Theta$ are (conditionally) independent under $\mathbb{P}^{h^\dagger}$. From (3.11) in Theorem 3.2,

$$\mathbb{P}^{h^\dagger}(\Theta = i) = p_i. \tag{3.15}$$

Using (3.8) in Lemma 3.3, for each $i = 1, 2, \ldots, N$ the conditional MGF of R_t given $\Theta = i$ under $\mathbb{P}^{h^\dagger}$ evaluated at η is given by

$$M_i^{h^\dagger}(\eta) = \frac{M_i(\eta + h^\dagger(i))}{M_i(h^\dagger(i))}. \tag{3.16}$$

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

Then the MGF of R_t under $\mathbb{P}^{h^\dagger}$ evaluated at η is given by

$$M^{h^\dagger}(\eta) = \sum_{i=1}^{N} M_i^{h^\dagger}(\eta) p_i = \sum_{i=1}^{N} \frac{M_i(\eta + h^\dagger(i))}{M_i(h^\dagger(i))} p_i. \tag{3.17}$$

A European call option with strike price K and exercise date T is considered. As in [26], we calculate the value of the call option at time $t = 0$. Let C_0 denote the value of the call option at time $t = 0$. Then

$$C_0 = \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+]. \tag{3.18}$$

Using the law of iterated expectations and (3.15),

$$\begin{aligned} C_0 &= \sum_{i=1}^{N} \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+ \mid \Theta = i] \mathbb{P}^{h^\dagger}(\Theta = i) \\ &= \sum_{i=1}^{N} \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+ \mid \Theta = i] p_i. \end{aligned} \tag{3.19}$$

For each $i = 1, 2, \ldots, N$, let

$$C_0(i) := \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+ \mid \Theta = i]. \tag{3.20}$$

Then, from (3.19) and (3.20),

$$C_0 = \sum_{i=1}^{N} C_0(i) p_i. \tag{3.21}$$

4. Analytical solutions

In this section, four parametric cases, namely a finite mixture model of normal distributions, a finite mixture model of negative-shifted Gamma distributions, a finite mixture model of negative-shifted inverse Gaussian distributions, and a hybrid finite mixture model of normal, negative-shifted Gamma and negative-shifted inverse Gaussian distributions are considered. The existence of an equivalent martingale measure, or a pricing kernel, selected by a risk-neutral Esscher measure is discussed for each of the four parametric cases. The derivations presented adopt results from [26, Sections 3 and 4]. Some of the conventions and notation of [26] will also be used in the following.

4.1. Finite mixture model of normal distributions

Assume that, for each $i = 1, 2, \ldots, N$, the ith-component CDF $F_i(x)$ is the CDF of a normal distribution with mean μ_i and variance σ_i^2 (i.e. $N(\mu_i, \sigma_i^2)$). Then

$$M_i(\eta) = \exp\left(\mu_i \eta + \frac{1}{2}\eta^2 \sigma_i^2\right). \tag{4.1}$$

From (3.14) and (4.1),

$$h^\dagger(i) = \frac{r - \mu_i - \frac{1}{2}\sigma_i^2}{\sigma_i^2}, \quad i = 1, 2, \ldots, N. \tag{4.2}$$

Then (4.2) ensures the existence of an equivalent martingale measure $\mathbb{P}^{h^\dagger}$.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

From (3.16), (4.1), and (4.2),

$$M_i^{h^\dagger}(\eta) = \exp\left(\left(r - \frac{1}{2}\sigma_i^2\right)\eta + \frac{1}{2}\eta^2\sigma_i^2\right). \tag{4.3}$$

This implies that under $\mathbb{P}^{h^\dagger}$, the conditional distribution of R_t given $\Theta = i$ is a normal distribution with mean $r - \frac{1}{2}\sigma_i^2$ and variance σ_i^2, (i.e. $R_t \mid \{\Theta = i\} \sim N(r - \frac{1}{2}\sigma_i^2, \sigma_i^2)$), for each $i = 1, 2, \ldots, N$ and each $t = 1, 2, \ldots, T$. Define

$$R_{1,T} := \sum_{t=1}^{T} R_t. \tag{4.4}$$

Note that by Lemma 3.2, $R_1, R_2, \ldots, R_T \mid \Theta$ are (conditionally) independent under $\mathbb{P}^{h^\dagger}$. Furthermore, for each $t = 1, 2, \ldots, T$ and each $i = 1, 2, \ldots, N$, $R_t \mid \{\Theta = i\} \sim N\left(r - \frac{1}{2}\sigma_i^2, \sigma_i^2\right)$ under $\mathbb{P}^{h^\dagger}$. Consequently, $R_{1,T} \mid \{\Theta = i\} \sim N\left(\left(r - \frac{1}{2}\sigma_i^2\right)T, \sigma_i^2 T\right)$ under $\mathbb{P}^{h^\dagger}$. It may be seen that, for each $i = 1, 2, \ldots, N$,

$$C_0(i) = \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+ \mid \Theta = i] = \mathrm{BSM}(S_0, K, T, r, \sigma_i). \tag{4.5}$$

Note that $\mathrm{BSM}(S_0, K, T, r, \sigma_i)$ is the Black–Scholes–Merton call price with volatility parameter σ_i:

$$\mathrm{BSM}(S_0, K, T, r, \sigma_i) = S_0\Phi(d_1(i)) - K\mathrm{e}^{-rT}\Phi(d_2(i)), \tag{4.6}$$

where

$$d_1(i) = \frac{\ln(S_0/K) + \left(r + \frac{1}{2}\sigma_i^2\right)T}{\sigma_i\sqrt{T}}, \qquad d_2(i) = d_1(i) - \sigma_i\sqrt{T}.$$

Then, from (3.21) and (4.5), an analytical expression for the call price is

$$C_0 = \sum_{i=1}^{N} \mathrm{BSM}(S_0, K, T, r, \sigma_i)p_i. \tag{4.7}$$

4.2. Finite mixture model of negative-shifted gamma distributions

From the empirical studies in Section 5, the skewness of the logarithmic returns of Bitcoin in USD is negative. Consequently, a shifted Gamma distribution may not be suitable to model the logarithmic returns of Bitcoin in USD since the skewness of a shifted Gamma distribution is positive [26, p. 117]. Consequently, instead of considering a finite mixture model of shifted Gamma distributions, we suppose that the logarithmic returns follow a finite mixture model of negative-shifted Gamma distributions. Specifically, we assume that under the real-world probability measure $\mathbb{P}$, for each $i = 1, 2, \ldots, N$ and each $t = 1, 2, \ldots, T$,

$$R_t \mid \{\Theta = i\} \overset{\mathrm{d}}{=} k_i - Y_i, \tag{4.8}$$

where Y_i follows a Gamma distribution $\mathrm{Ga}(\alpha_i, \beta_i)$ with shape parameter α_i and rate parameter β_i; $\overset{\mathrm{d}}{=}$ means equality in distribution. In this case, for each $i = 1, 2, \ldots, N$ and each $t = 1, 2, \ldots, T$, the conditional distribution of R_t given $\{\Theta = i\}$ under $\mathbb{P}$ is a negative-shifted

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

Gamma distribution with shape parameter α_i, rate parameter β_i, and shifted parameter k_i (i.e. NegSGa(α_i, β_i, k_i)), where k_i, α_i, and β_i are all positive.

Let $G(x \mid \alpha, \beta)$ denote the CDF of a Gamma distribution with shape parameter α and rate parameter β evaluated at $x > 0$. Then, for each $i = 1, 2, \ldots, N$, the ith-component CDF $F_i(x)$ is given by

$$F_i(x) = 1 - G(k_i - x \mid \alpha_i, \beta_i), \quad x < k_i. \tag{4.9}$$

Similarly to [26, (4.1.3)], for each $i = 1, 2, \ldots, N$ the MGF of $F_i(x)$ evaluated at η is given by

$$M_i(\eta) = \left(\frac{\beta_i}{\beta_i + \eta}\right)^{\alpha_i} \mathrm{e}^{k_i \eta}, \quad \eta > -\beta_i. \tag{4.10}$$

From (3.14) and (4.10),

$$\mathrm{e}^r = \left(\frac{\beta_i + h^\dagger(i)}{\beta_i + h^\dagger(i) + 1}\right)^{\alpha_i} \mathrm{e}^{k_i}, \quad i = 1, 2, \ldots, N. \tag{4.11}$$

Take, for each $i = 1, 2, \ldots, N$,

$$\beta_i^\dagger = \beta_i + h^\dagger(i). \tag{4.12}$$

Then, from (4.11) and (4.12),

$$\beta_i^\dagger = \frac{1}{\mathrm{e}^{(k_i - r)/\alpha_i} - 1}, \quad k_i > r. \tag{4.13}$$

Indeed, (4.12) and (4.13) ensure the existence of an equivalent martingale measure $\mathbb{P}^{h^\dagger}$ provided that $k_i > r$. The condition that $k_i > r$ is required to ensure that $\beta_i^\dagger > 0$.

From (3.16), (4.10), and (4.12), for each $i = 1, 2, \ldots, N$ the conditional MGF of R_t given $\Theta = i$ under $\mathbb{P}^{h^\dagger}$ evaluated at η is given by

$$M_i^{h^\dagger}(\eta) = \left(\frac{\beta_i^\dagger}{\beta_i^\dagger + \eta}\right)^{\alpha_i} \mathrm{e}^{k_i \eta}, \quad \eta > -\beta_i^\dagger. \tag{4.14}$$

Consequently, under $\mathbb{P}^{h^\dagger}$, the conditional distribution of R_t given $\Theta = i$ is a negative-shifted Gamma distribution with shape parameter α_i, rate parameter $\beta_i^\dagger$, and shifted parameter k_i. Since $R_1, R_2, \ldots, R_T \mid \Theta$ are (conditionally) independent under $\mathbb{P}^{h^\dagger}$, $R_{1,T} \mid \{\Theta = i\} \sim$ NegSGa($\alpha_i T$, $\beta_i^\dagger$, $k_i T$) under $\mathbb{P}^{h^\dagger}$ for each $i = 1, 2, \ldots, N$.

Define

$$\kappa := \ln\left(\frac{K}{S_0}\right). \tag{4.15}$$

Then, similarly to [26, (4.1.7)],

$$\begin{aligned} C_0(i) &= \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+ \mid \Theta = i] \\ &= S_0 G(k_i T - \kappa \mid \alpha_i T, \beta_i^\dagger + 1) - K\mathrm{e}^{-rT} G(k_i T - \kappa \mid \alpha_i T, \beta_i^\dagger) \\ &= \mathrm{NSGa}(S_0, K, T, r, \alpha_i, \beta_i^\dagger, k_i), \quad \text{say.} \end{aligned} \tag{4.16}$$

Note that for (4.16) to hold, it is required that $k_i T > \kappa$ for each $i = 1, 2, \ldots, N$.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

From (3.21) and (4.16), an analytical expression for the call price is

$$C_0 = \sum_{i=1}^{N} \mathrm{NSGa}(S_0, K, T, r, \alpha_i, \beta_i^{\dagger}, k_i) p_i. \tag{4.17}$$

From (4.13), $\beta_i^{\dagger}$ does not depend on β_i for each $i = 1, 2, \ldots, N$. Consequently, the call price C_0 in (4.17) does not depend on β_i.

4.3. Finite mixture model of negative-shifted inverse Gaussian distributions

Let $J(x \mid a, b)$ denote the CDF of an inverse Gaussian (IG) distribution with parameters a and b evaluated at $x > 0$. Then, from [26, (4.2.1)],

$$J(x \mid a, b) = \Phi\left(-\frac{a}{\sqrt{2x}} + \sqrt{2bx}\right) + \mathrm{e}^{2a\sqrt{b}}\Phi\left(-\frac{a}{\sqrt{2x}} - \sqrt{2bx}\right), \tag{4.18}$$

where $\Phi(\,\cdot\,)$ is the CDF of a standard normal distribution. That is, if μ^{sig} and λ are the mean and the shape parameter of the IG distribution, respectively, then $\mu^{sig} = a/2\sqrt{b}$ and $\lambda = a^2/2$.

We suppose that, under $\mathbb{P}$, for each $i = 1, 2, \ldots, N$ and each $t = 1, 2, \ldots, T$,

$$R_t \mid \{\Theta = i\} \overset{\mathrm{d}}{=} k_i - Y_i, \tag{4.19}$$

where Y_i follows an IG distribution with parameters a_i and b_i. Note that a_i, b_i, and k_i are all positive. In this case, the conditional distribution of R_t given $\{\Theta = i\}$ under $\mathbb{P}$ is a negative-shifted inverse Gaussian distribution with IG parameters a_i and b_i, as well as the shifted parameter k_i. That is, under $\mathbb{P}$, $R_t \mid \{\Theta = i\} \sim \mathrm{NegSIG}(a_i, b_i, k_i)$.

Then, for each $i = 1, 2, \ldots, N$, the ith-component CDF $F_i(x)$ is given by

$$F_i(x) = 1 - J(k_i - x \mid a_i, b_i), \quad x < k_i. \tag{4.20}$$

Similarly to [26, (4.2.3)], for each $i = 1, 2, \ldots, N$ the MGF of $F_i(x)$ evaluated at η is given by

$$M_i(\eta) = \exp\left(a_i(\sqrt{b_i} - \sqrt{b_i + \eta}) + k_i\eta\right), \quad \eta > -b_i. \tag{4.21}$$

From (3.14) and (4.21),

$$r = a_i\left(\sqrt{b_i + h^{\dagger}(i)} - \sqrt{b_i + h^{\dagger}(i) + 1}\right) + k_i. \tag{4.22}$$

Take, for each $i = 1, 2, \ldots, N$,

$$b_i^{\dagger} = b_i + h^{\dagger}(i). \tag{4.23}$$

Then, from (4.22) and (4.23),

$$\sqrt{b_i^{\dagger}} - \sqrt{b_i^{\dagger} + 1} = \frac{r - k_i}{a_i}. \tag{4.24}$$

Consequently, from (4.23), the existence of an equivalent martingale measure $\mathbb{P}^{h^{\dagger}}$ becomes the existence of a solution to the non-linear equation in (4.24) provided that $b_i > -h^{\dagger}(i)$. The latter condition is required to ensure that $b_i^{\dagger} > 0$.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

From (3.16), (4.21), and (4.23), for each $i = 1, 2, \ldots, N$ the conditional MGF of R_t given $\Theta = i$ under $\mathbb{P}^{h^\dagger}$ evaluated at η is given by

$$M_i^{h^\dagger}(\eta) = \exp\left(a_i\left(\sqrt{b_i^\dagger} - \sqrt{b_i^\dagger + \eta}\right) + k_i\eta\right), \quad \eta > -b_i^\dagger. \tag{4.25}$$

Consequently, under $\mathbb{P}^{h^\dagger}$, the conditional distribution of R_t given $\Theta = i$ is a negative-shifted IG distribution with parameters a_i and $b_i^\dagger$, as well as the shifted parameter k_i. Since $R_1, R_2, \ldots, R_T \mid \Theta$ are (conditionally) independent under $\mathbb{P}^{h^\dagger}$, $R_{1,T} \mid \{\Theta = i\} \sim \text{NegSIG}(a_iT, b_i^\dagger, k_iT)$ under $\mathbb{P}^{h^\dagger}$ for each $i = 1, 2, \ldots, N$. Then, similarly to [26, (4.2.7)],

$$\begin{aligned} C_0(i) &= \mathbb{E}^{h^\dagger}[\mathrm{e}^{-rT}(S_T - K)^+ \mid \Theta = i] \\ &= S_0 J(k_iT - \kappa \mid a_iT, b_i^\dagger + 1) - K\mathrm{e}^{-rT} J(k_iT - \kappa \mid a_iT, b_i^\dagger), \\ &= \text{NSIG}(S_0, K, T, r, a_i, b_i^\dagger, k_i). \end{aligned} \tag{4.26}$$

Note that for (4.26) to hold, it is required that $k_iT > \kappa$ for each $i = 1, 2, \ldots, N$.

From (3.21) and (4.26), an analytical expression for the call price is

$$C_0 = \sum_{i=1}^{N} \text{NSIG}(S_0, K, T, r, a_i, b_i^\dagger, k_i)p_i. \tag{4.27}$$

This analytical pricing formula is up to solving the non-linear equation for $b_i^\dagger$ in (4.24). Furthermore, from (4.24), $b_i^\dagger$ does not depend on b_i for each $i = 1, 2, \ldots, N$. Consequently, the call price C_0 in (4.27) does not depend on b_i.

4.4. Hybrid finite mixture model

In this subsection, a hybrid finite mixture model of normal, negative-shifted Gamma and negative inverse Gaussian distributions is considered. The hybrid model allows for the flexibility that the component distributions may come from different parametric classes of distributions. It provides a feasible way to describe both model uncertainty and parameter uncertainty. The former is interpreted as uncertainty about the parametric form of distributions. Specifically, to describe model uncertainty we consider more than one parametric form for the component distributions and use the mixing mechanism in a finite mixture model to take a 'weighted average' of the component distributions across different parametric distribution classes. The idea may perhaps be related to the use of Bayesian averaging to capture model uncertainty in the context of smooth ambiguity pioneered in [43]. To incorporate parameter uncertainty, the mixing mechanism in a finite mixture model is used to take a 'weighted average' of the component distributions within the same parametric class of distributions. Informally speaking, model uncertainty is incorporated through considering the variation or heterogeneity between different parametric classes of distributions. Parameter uncertainty is captured by considering the variation or heterogeneity within the same parametric class of distributions.

Suppose, for each $i = 1, 2, \ldots, N_1$, the ith-component CDF $F_i(x)$ is the CDF of a normal distribution with mean μ_i and variance σ_i^2. Also, for each $i = N_1 + 1, N_1 + 2, \ldots, N_1 + N_2$, the ith-component CDF $F_i(x)$ is the CDF of a negative-shifted Gamma distribution with shape parameter α_i, rate parameter β_i, and shifted parameter k_i. Assume, for each

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

$i = N_1 + N_2 + 1, N_1 + N_2 + 2, \ldots, N$, the ith-component CDF $F_i(x)$ is the CDF of a negative-shifted IG distribution with IG parameters a_i and b_i, and the shifted parameter k_i. Then the hybrid finite mixture model can be written as

$$F(x) = \sum_{i=1}^{N_1} p_i N(x \mid \mu_i, \sigma_i^2) + \sum_{i=N_1+1}^{N_1+N_2} p_i (1 - G(k_i - x \mid \alpha_i, \beta_i)) + \sum_{i=N_1+N_2+1}^{N} p_i (1 - J(k_i - x \mid a_i, b_i)) \tag{4.28}$$

for $x < \min\{k_{N_1+1}, k_{N_1+2}, \ldots, k_N\}$.

Using the results in Sections 4.1–4.3, an analytical expression for the call price is

$$C_0 = \sum_{i=1}^{N_1} \mathrm{BSM}(S_0, K, T, r, \sigma_i) p_i + \sum_{i=N_1+1}^{N_1+N_2} \mathrm{NSGa}(S_0, K, T, r, \alpha_i, \beta_i^{\dagger}, k_i) p_i + \sum_{i=N_1+N_2+1}^{N} \mathrm{NSIG}(S_0, K, T, r, a_i, b_i^{\dagger}, k_i) p_i. \tag{4.29}$$

Again, the analytical pricing formula in (4.29) is up to solving the non-linear equation for $b_i^{\dagger}$ in (4.24). The existence of an equivalent martingale measure $\mathbb{P}^{h^{\dagger}}$ follows directly from the respective discussions in Sections 4.1–4.3.

5. Estimation procedures and results

In this section we adopt the ECF estimation method to estimate the unknown parameters in eight models that will be used for pricing Bitcoin options in Section 6. A summary of the eight models is provided in Table D.1 in the supplementary material. Models I–V are finite mixture models, while Models VI–VIII are non-mixture models. The latter are used for comparison. Specifically, Model VI is the Black–Scholes–Merton model [9, 52]. Models VII and VIII are based on negative-shifted Gamma and negative-shifted IG distributions, respectively. First, the estimation procedures of the ECF estimation method are briefly reviewed. Then the estimation results based on real Bitcoin price data are provided.

5.1. Estimation procedures

The ECF estimation method [21, 22, 44, 48, 49, 61, 79, 83] has been applied to estimate unknown parameters of finite mixture models of normal distributions [10, 79, 83]. Its key idea is to estimate the unknown parameters by minimizing the 'distance' between the empirical characteristic function and its theoretical counterpart. It was noted in [83] that the ECF estimation method can be considered when the maximum likelihood estimation method is difficult and a closed-form solution of the (theoretical) characteristic function is available. This is why the ECF estimation method is adopted here. Specifically, for the cases of a finite mixture model of negative-shifted Gamma distributions and a finite mixture model of negative-shifted IG distributions, it is possible that the observed logarithmic returns may not be bounded above by the shifted parameters. This renders the maximum likelihood estimation method intractable. Furthermore, we note that closed-form expressions for the (theoretical) characteristic functions of a negative-shifted Gamma distribution and a negative-shifted IG distribution

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

are available and that the (theoretical) characteristic function of a finite mixture model is a mixture of the (theoretical) characteristic functions of individual (theoretical) characteristic functions. Consequently, closed-form expressions for the (theoretical) characteristic functions of the finite mixture model of negative-shifted Gamma distributions and the finite mixture model of negative-shifted IG distributions are obtained. The ECF estimation method is intimately linked with the generalized method of moment (GMM) estimation [35]. See [83] for a detailed discussion on the link between the GMM method and the ECF method. [26] proposed the use of the method of moments to estimate unknown parameters in a shifted Gamma distribution and a shifted IG distribution, and derived the respective closed-form estimators. In the following, the ECF estimation method is briefly discussed in the context of a general finite mixture model.

Let $\{R_1, R_2, \ldots, R_n\}$ denote a set of n observations about the logarithmic returns of Bitcoin. Then the ECF of $\{R_1, R_2, \ldots, R_n\}$ evaluated at $u \in \mathbb{R}$ is defined as

$$C_n(u) := \frac{1}{n} \sum_{j=1}^{n} \exp(\mathrm{i}uR_j), \tag{5.1}$$

where $\mathrm{i} := \sqrt{-1}$.

Let $C(u; \varphi)$ denote the (theoretical) characteristic function of the general finite mixture distribution $F(x)$ in (2.1) evaluated at $u \in \mathbb{R}$ under $\mathbb{P}$, where φ is a vector of unknown parameters in the finite mixture distribution $F(x) := F(x; \varphi)$ and $\varphi \in \Phi \subset \mathbb{R}^k$ for some Euclidean space $\mathbb{R}^k$. For each $j = 1, 2, \ldots, N$, let $C_j(u;\varphi)$ denote the (theoretical) characteristic function of the jth-component CDF $F_j(x)$ of the general finite mixture distribution $F(x)$ evaluated at $u \in \mathbb{R}$ under $\mathbb{P}$. That is, $C_j(u; \varphi)$ is the conditional (theoretical) characteristic function of R_t given $\{\Theta = j\}$ evaluated at $u \in \mathbb{R}$ under $\mathbb{P}$. Note that for each $j = 1, 2, \ldots, N$, the vector of unknown parameters in $F_j(x)$ is obtained by putting some components of φ equal to zero. To simplify the notation, we still adopt φ to denote the vector of unknown parameters in $F_j(x)$. By putting $\eta = \mathrm{i}u$ in (2.7), we have

$$C(u, \varphi) = \sum_{j=1}^{N} p_j C_j(u, \varphi), \tag{5.2}$$

$$C(u, \varphi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}uR_t}] = \int_{\mathcal{S}} \mathrm{e}^{\mathrm{i}ux} \, \mathrm{d}F(x), \tag{5.3}$$

$$C_j(u, \varphi) = \mathbb{E}[\mathrm{e}^{\mathrm{i}uR_t} \mid \Theta = j] = \int_{\mathcal{S}_j} \mathrm{e}^{\mathrm{i}ux} \, \mathrm{d}F_j(x). \tag{5.4}$$

Note that $\mathcal{S}$ is the support of the probability density function (PDF) of the distribution $F(x)$; $\mathcal{S}_j$ is the support of the PDF of the distribution $F_j(x)$ for each $j = 1, 2, \ldots, N$. In the case of a finite mixture of negative-shifted Gamma distributions, a closed-form expression for $C_j(u, \varphi)$ is obtained by putting $\eta = \mathrm{i}u$ in (4.10). In the case of a finite mixture of negative-shifted IG distributions, a closed-form expression for $C_j(u, \varphi)$ is obtained by putting $\eta = \mathrm{i}u$ in (4.21).

Then the ECF estimation method is to estimate φ by solving the following minimization problem:

$$\min_{\varphi \in \Phi} \int_{-\infty}^{\infty} |C_n(u, \varphi) - C(u, \varphi)|^2 g(u) \, \mathrm{d}u, \tag{5.5}$$

where $g(u)$ is some continuous weighting function. Here, as in [83], we choose $g(u) = \exp(-u^2)$.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

In practice, it is often the case that there is no closed-form expression for the integral in (5.5). Consequently, the integral is evaluated via some numerical integration procedures. Here we shall use the function `integrate` in R to compute the integral numerically, which adopts adaptive quadrature of functions and involves a set of discrete points for the variable u of the integration. To solve the minimization problem in (5.5), the differential evolution algorithm is implemented using the `DEoptim` R package [2]. Differential evolution is an evolutionary global optimization algorithm that adopts biology-inspired operations on a population to minimize an objective function over successive generations [2, 53, 56, 72]. More detail about the use of differential evolution to estimate the finite mixture models considered in the current paper is provided in Section 5.2. It will be seen that human or subjective judgement may need to be exercised even when this advanced optimization algorithm is adopted. Finally, it may be noted that the ECF estimator enjoys some desirable asymptotic properties such as strong consistency and asymptotic normality. See [83] for a more detailed discussion.

5.2. Estimation results

A dataset on daily Bitcoin-USD (BTC-USD) adjusted close prices is used in empirical studies on the eight models in Table D.1. Each of Models I–V has two mixing components. The time period for the BTC-USD prices data is from 1 January 2020 to 29 May 2025 (1976 observations). The dataset was extracted from Yahoo Finance using R. It covers some periods of the Covid pandemic. In [64, 66], Bitcoin price data covering some periods of the Covid pandemic were considered. However, [64] focused on risk evaluation based on Value at Risk and Expected Shortfall. [66] studied the impacts of a long memory in the conditional volatility and conditional non-normality on pricing Bitcoin options. With the Bitcoin logarithmic returns data (1975 observations), where the empirical studies were performed using R. The estimation results are discussed in this subsection.

The estimation results for the five two-component finite mixture models (Models I–V) are provided in Tables D.2–D.6 in the supplementary material, respectively. Specifically, for each of the five two-component finite mixture models, the estimates of the unknown parameters, the optimal value(s) of the objective function corresponding to the parameter estimates, and the lower and upper boundaries of the parameters are presented. For the finite mixture model of normal distributions (Model I), one set of parameter estimates is provided based on one set of lower and upper boundaries of the parameters. For the other four finite mixture models with non-normal mixing components (Models II–V), two sets of parameter estimates are provided based on two sets of lower and upper boundaries of the parameters. In each of Tables D.2–D.6, 'fn' represents the optimal value of the objective function in differential evolution evaluated at the parameter estimates. All the parameter estimates in Tables D.2–D.6 are in daily units. It will be seen from the results that the parameter estimates obtained from differential evolution may be different when different sets of lower and upper boundaries of the parameters are adopted.

In Table D.2, the lower and upper boundaries of μ_1 and μ_2 are given by -100 and 100 times the sample mean of the logarithmic returns, respectively. The upper boundaries of σ_1 and σ_2 are 100 times the sample standard deviation of the logarithmic returns. From the estimation results, around 34.87% weight is allocated to the first component. The first component corresponds to a 'bad' economic regime with a negative mean return and high volatility. The second component corresponds to a 'good' economic regime with a positive mean return and low volatility.

There are two sets (Sets 1 and 2) of estimation results in each of Tables D.3–D.6. In each of Tables D.3 and D.4, the upper boundaries of parameters in Set 1 are 5 times their respective method of moments estimates from the sample logrithmic returns data. The upper boundaries

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

of parameters in Set 2 are 10 times their respective method of moments estimates from the sample data. The estimates of a_i and b_i, $i = 1, 2$, in Table D.4 are obtained from the respective estimates of μ_i^{sig} and λ_i. In each of Tables D.5 and D.6, for each of Sets 1 and 2, the lower and upper boundaries of μ are -100 and 100 times the sample mean of the logarithmic returns, respectively; the upper boundaries of σ in Sets 1 and 2 are 100 times the sample standard deviation of the logarithmic returns. Again, the upper boundaries of negative SGa and SIG parameters in Set 1 are five times their respective method of moment estimates from the sample logrithmic returns data. The upper boundaries of negative SGa and SIG parameters in Set 2 are 10 times their respective method of moments estimates from the sample data. The estimates of a and b in Table D.6 are obtained from the respective estimates of μ^{sig} and λ. There are two reasons why the upper boundaries of negative SGa and SIG parameters are set as either 5 or 10 times their respective method of moment estimates. Firstly, some of the parameters may become large if their upper boundaries are large. Specifically, the parameters β_i (or β) in the negative SGa distributions may become large if their upper boundaries are large. Secondly, the optimal value 'fn' may not improve significantly even if the upper boundaries are large. However, the computational time increases if the upper boundaries are large. In each of Tables D.3–D.6, the estimation results from Sets 1 and 2 could be quite different. This illustrates that the parameter estimates obtained from differential evolution depend on the chosen boundaries of the parameters. We may need to exercise subjective judgements even though the advanced optimization algorithm is adopted.

In Table D.3, the optimal value 'fn' (i.e. the objective function in (5.5)) increases from $1.699\,928 \times 10^{-10}$ to $2.519\,128 \times 10^{-9}$ when moving from Set 1 to Set 2. This indicates that the optimal value 'fn' may not improve when parameter bounds are relaxed. This happens because the solver was not able to find the global optimum and led to suboptimal convergence. Although the estimates of β_1 in Sets 1 and 2 are large, they are still interior optimums (i.e. strictly less than the respective upper boundaries). Although the parameter estimates in Set 1 provide a better fit to the data than those in Set 2 according to the optimal value 'fn', the parameter estimates in Set 2 will be used to calculate the option prices in Section 6 to illustrate the mixture effect. In Table D.4, the optimal value 'fn' increases from $1.512\,852 \times 10^{-9}$ to $3.604\,965 \times 10^{-9}$ when moving from Set 1 and Set 2. This also indicates that the optimal value 'fn' may not improve when the upper boundaries become larger. Since the parameter estimate of b in Set 2 looks more reasonable than that in Set 1, the former will be used to calculate the option prices in Section 6. In Table D.5, the parameter estimates in Set 1 provide a better fit to the data than those in Set 2. Furthermore, the parameter estimate of σ in Set 2 is very small (0.002 944 167), which may result in unreasonably small option prices. Consequently, the parameter estimates in Set 1 will be used to calculate the option prices in Section 6. In Table D.6, the estimates of p in Sets 1 and 2 are 0.012 423 47 and 0.578 201 806, respectively. To illustrate the mixture effect, the parameter estimates in Set 2 may be better than those in Set 1. However, the estimate of σ in Set 2 is very small (0.002 358 598), which may result in unreasonably small option prices. Furthermore, the parameter estimates in Set 1 provide a better fit to the data than those in Set 2. Consequently, the parameter estimates in Set 1 will be used to calculate the option prices in Section 6.

Tables D.7–D.9 present the estimation results for the three non-mixture models (Models VI–VIII), respectively.

In Table D.7, both the maximum likelihood estimates (MLE) and the ECF estimates are provided. For the ECF estimates, the lower and upper boundaries of μ are -100 and 100 times the sample mean of the logarithmic returns, respectively. The upper boundary of σ is

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

100 times the sample standard deviation of the logarithmic returns. The optimal value 'fn' (i.e. the objective function in (5.5)) evaluated at the ECF estimates is smaller than the optimal value 'fn' evaluated at the MLE, though the optimal values evaluated at both the MLE and the ECF estimates are reasonably small. This indicates that the ECF estimation method may be better than the MLE method when estimating the mean and standard deviation of a normal distribution. However, it is possible that the solver was not able to find the global optimum and led to suboptimal convergence. The ECF estimates will be used to calculate the option prices in Section 6. In each of Tables D.8 and D.9, the upper boundaries for the parameters in Sets 1 and 2 are 5 and 10 times their respective method of moment estimates, respectively. The estimates of a and b in Table D.9 are obtained from those of μ^{sig} and λ. In Table D.8, the parameter estimates in Set 2 provide a better fit to the data and will be used to calculate the option prices in Section 6. In Table D.9, the parameter estimates in Set 1 provide a better fit to the data and will be used to calculate the option prices in Section 6. Using 'fn' as a criterion, the fitting performances of Models I–VIII look similar, though Model VIII is the best among them.

6. An application to Bitcoin option pricing

Recently, there has been considerable interest in developing quantitative models for pricing Bitcoin options and options written on cryptocurrencies. For example, [40, 42, 66, 67, 81] considered the pricing of Bitcoin options under stochastic volatility and GARCH-type models. [65] constructed Bayesian lower and upper estimates for the prices of ether options under GARCH models by leveraging Bayesian non-linear expectation proposed in [64]. [13, 39] discussed the pricing of Bitcoin options under jump-diffusion models. [45, 55] adopted the machine learning and neural network approaches to price Bitcoin options, respectively. Using the Black–Scholes–Merton model, [1] discussed the pricing of crypto quanto and inverse options. From Figure C.1 (panel B) in the supplementary material, the daily logarithmic returns exhibit volatility clustering. Consequently, models with conditional heteroscedasticity and stochastic volatility may provide a better fit to the returns data than finite mixture models. There are two reasons why we consider finite mixture models for pricing Bitcoin options. First, closed-form expressions for the European call Bitcoin options are obtained under finite mixture models (Section 4). Secondly, pricing Bitcoin options under models with conditional heteroscedasticity and stochastic volatility was considered in the aforementioned literature. Lastly, by considering the finite mixture models, we can focus on the impact of non-normality of Bitcoin returns on option pricing.

In this section, the proposed finite mixture models are applied to price Bitcoin options. Specifically, the estimated models (Models I–V) based on real Bitcoin returns data in Section 5 are used to compute the prices of European-style call Bitcoin options. Three estimated non-mixture models (Models VI–VIII) are used for comparison. See Table D.1 for the detail. For each of the eight models, a matrix of option prices with different strike prices and maturities is computed. The possible strike prices are $0.8s_0$, $0.85s_0$, $0.9s_0$, $0.95s_0$, s_0, $1.05s_0$, $1.1s_0$, $1.15s_0$, and $1.2s_0$, where s_0 is the current Bitcoin price, (i.e. the Bitcoin price at the end of the dataset). The maturities are 21, 42, 63, 126, and 252 trading days. They correspond to one, two, three, six, and twelve months, respectively. Since analytical expressions for the option prices are obtained for Models I–VIII, the option prices are quickly obtained. All the computations for the option prices were calculated with R. The risk-free interest rate is assumed to be 4.33% per annum (i.e. 0.0433/252 per trading day), as taken from the US Federal Funds Effective Rate (FEDFUNDS) in May 2025. Note that the interest rate on Bitcoin is zero [5, 67].

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

Tables D.10–D.17 in the supplementary material show the prices of European-style call Bitcoin options with different levels of moneyness and maturities from Models I to VIII, respectively.

From Tables D.10–D.17, for each of the eight models (Models I–VIII), the option prices increase as the maturity T increases, and they decrease as the moneyness K/s_0 increases. These are necessary no-arbitrage conditions [52]. From Tables D.10 and D.11, the option prices from Model I are higher than those from Model II. One possible explanation for this is that the negative-shifted Gamma distribution has a finite upper limit of its support, which limits the impact of upside movements of the underlying Bitcoin price on the call price. This results in lower call prices. By comparing the option prices in Table D.11 with those in Table D.12, the option prices from Model II are lower than those from Model III. One potential explanation for this is that the negative-shifted IG distribution has a heavier tail than the negative-shifted Gamma distribution. Since the negative-shifted IG distribution is heavy-tailed, this may explain why some option prices from Model III are higher than the respective prices from Model I, as can be seen by comparing Tables D.10 and D.12. From Table D.13, the option prices are unreasonably small when the maturity is short ($T = 21$ days) as well as when the option is deep out-of-money ($K/s_0 = 1.2$). This indicates that Model IV may not price short-lived options and out-of-money options accurately. For this reason, it will not be used to compute the Black–Scholes–Merton implied volatilities in the later part of this section. By comparing Tables D.10 and D.13, option prices from Model IV are lower than those from Model I. This may be partly explained by a relatively small volatility in component 1 of Model IV and a finite upper limit of the support of the negative-shifted Gamma distribution. By comparing Table D.14 with Tables D.10–D.13, Model V gives the most conversative estimates for the option prices among all five finite mixture models. This may be attributed to the combined effect of the heavy-tailedness of a negative-shifted IG distribution and the unbounded support of a normal distribution. By comparing the option prices in Tables D.15–D.17 with those in Tables D.10–D.14, the option prices from non-mixture models are more conversative than the respective prices from the finite mixture models. This reflects that the finite mixture models provide the flexibility of incorporating two possible economic regimes, one of which gives less conversative option prices.

Figure D.1 in the supplementary material plots the Black–Scholes–Merton implied volatilities against moneyness from Models I–III and Model V for three different maturities. Figure D.2 plots the BSM implied volatilities against maturity from Models I–III and Model V for in-the-money (ITM), at-the-money (ATM), and out-of-money (OTM) options. For each moneyness and each maturity, the BSM implied volatility from a model is computed by minimizing the absolute difference between the option price from the model and the theoretical option price from the BSM model evaluated at the implied volatility. Again, the differential evolution algorithm is used for the minimization, with the lower and upper boundaries being 0 and 1, respectively.

From Figure D.1, all four finite mixture models can produce implied volatility smirks or skews. That is, the BSM implied volatilities for ITM options are higher than those for ATM and OTM options. The BSM implied volatilities from Model V are higher than those from the other three finite mixture models. This is in line with what was observed by comparing Table D.14 with Tables D.10–D.13. From Figure D.2, the four finite mixture models produce consistent term structures of the BSM implied volatilities for ITM, ATM, and OTM options. Specifically, when the options are ITM, ATM, and OTM options, all four finite mixture models produce downward-sloping, nearly flat, and upward-sloping term structures of the BSM implied volatilities, respectively.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

7. Conclusion

Finite mixture models were adopted for the canonical valuation of Bitcoin options in a discrete-time economy by integrating knowledge in actuarial science, financial mathematics, applied probability, statistics, econometrics, and data science. Specifically, the time-honored tool in actuarial science, namely the Esscher transform, was used to specify a pricing kernel. Novel finite mixture models were introduced to capture non-normality of returns and model uncertainty. Specifically, a finite mixture model of negative SGa distributions, a finite mixture model of negative SIG distributions, as well as a hybrid finite mixture model of normal, negative SGa, and negative SIG distributions were introduced. The finite mixture models were estimated using the empirical characteristic function estimation method. Analytical pricing formulas for a European call option were obtained for some finite mixture models. The pricing formulas provide quick and convenient ways to compute matrices of option prices and to calibrate the models to real option price data. Using real data on the adjusted close prices of Bitcoin-USD, the estimation results of eight models (five finite mixture models and three non-mixture models) were provided. Numerical results illustrating applications of the eight models to pricing European call Bitcoin options were presented. The results reveal that the hybrid mixture model with a normal distribution and a negative-shifted IG distribution gives the most conservative pricing result. It is also found that four finite mixture models produce implied volatility smirks/skews, and that the finite mixture models produce different patterns for term structures of the implied volatility for ITM, ATM, and OTM options. These results could be helpful to stakeholders in Bitcoin option markets such as traders, regulators, and individual investors.

There are some limitations of the proposed finite mixture models and studies that may provide some potential ideas for further research. The models cannot incorporate conditional heteroscedasticity and stochastic volatility. Finite mixture models with either negative-shifted Gamma distributions or negative-shifted IG distributions impose finite upper limits of the supports of the distributions. This limits the capability of the models for modelling large up movements in the Bitcoin returns. The estimation results based on differential evolution depend on the choice of the boundaries of the unknown parameters, which involves subjective judgement. Real Bitcoin option price data were not used in this study.

Acknowledgements

The author sincerely thanks the two reviewers for their positive and meticulous comments.

Funding information

The author also wishes to acknowledge the support from the Discovery Grant from the Australian Research Council (ARC) (Project Number: DP190102674).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

Data

The data related to the empirical results found in Section 5 can be found at Yahoo Finance; they were extracted using R.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

Supplementary material

The supplementary material for this article can be found at http://doi.org/10.1017/jpr.2025.10048.

References

[1] ALEXANDER, C., CHEN, D. AND IMERAJ, A. (2023). Crypto quanto and inverse options. *Math. Finance* **33**, 1005–1043.

[2] ARDIA, D., MULLEN, K., PETERSON, B., ULRICH, J. AND BOUDT, K. (2025). Global optimization by differential evolution. R package `DEoptim`. January 20, 2025. Version 2.2-8.

[3] BADESCU, A., ELLIOTT, R. J. AND SIU, T. K. (2009). Esscher transforms and consumption-based models. *Insurance Math. Econom.* **45**, 337–347.

[4] BADESCU, A., KULPERGER, R. AND LAZAR, E. (2008). Option valuation with normal mixture GARCH models. *Studies Nonlinear Dynam. Econometrics* **12**, 1580–1580.

[5] BARIVIERA, A. F., BASGALLA, M. J., HASPERUÉ, W. AND NAIOUF, M. (2017). Some stylized facts of the Bitcoin market. *Physica A* **484**, 82–90.

[6] BÉGIN, J. F. AND GODIN, F. (2023). Option pricing under stochastic volatility models with latent volatility. *Quant. Finance* **23**, 1079–1097.

[7] BHAT, H. S. and KUMAR, N. (2012). Option pricing under a normal mixture distribution derived from the Markov tree model. *Europ. J. Operat. Res.* **223**, 762–774.

[8] BISHOP, C. (2006). *Pattern Recognition And Machine Learning*. Springer, New York.

[9] BLACK, F. AND SCHOLES, M. (1973). The pricing of options and corporate liabilities. *J. Political Econom.* **81**, 637–659.

[10] BRYANT, J. L. AND PAULSON, A. S. (1983). Estimation of mixing properties via distance between characteristic functions. *Commun. Statist. Theory Meth.* **12**, 1009–1029.

[11] BÜHLMANN, H., DELBAEN, F., EMBRECHTS, P. AND SHIRYAEV, A. N. (1996). No arbitrage, change of measure and conditional Esscher transforms. *CWI Quart.* **9**, 291–317.

[12] BÜHLMANN, H., DELBAEN, F., EMBRECHTS, P. AND SHIRYAEV, A. N. (1998). On Esscher transforms in discrete finance models. *ASTIN Bull.* **28**, 171–186.

[13] CAO, M. AND CELIK, B. (2021). Valuation of Bitcoin options. *J. Futures Markets* **41**, 1007–1026.

[14] CHING, W K., SIU, T. K. AND LI, L. M. (2007). Pricing exotic options under a high-order Markovian regime switching model. *J. Appl. Math. Decision Sci.* **2007**, 18014. DOI: 10.1155/2007/18014.

[15] DUAN, J. C., POPOVA, I. AND RITCHKEN, P. (2001). Option pricing under regime switching. *Quant. Finance* **2**, 116–132.

[16] ELLIOTT, R. J., CHAN, L. L. AND SIU, T. K. (2005). Option pricing and Esscher transform under regime switching. *Ann. Finance* **1**, 423–432.

[17] ELLIOTT, R. J. AND MADAN, D. B. (1998). A discrete time equivalent martingale measure. *Math. Finance* **8**, 127–152.

[18] ELLIOTT, R. J., SIU, T. K. AND BADESCU, A. (2011). Bond valuation under a discrete-time regime-switching term-structure model and its continuous-time extension. *Managerial Finance* **37**, 1025–1047.

[19] ELLIOTT, R. J., SIU, T. K. AND CHAN, L. L. (2006). Option pricing for GARCH models with Markov switching. *Internat. J. Theoret. Appl. Finance* **9**, 825–841.

[20] FAN, J. (2014). Features of big data and sparsest solution in high confidence set. In *Past, Present, and Future of Statistical Science*, eds X. LIN, C. GENEST, D. L. BANKS, G. MOLENBERGHS, D. W. SCOTT AND J. L. WANG. Chapman & Hall, New York, pp. 507–523.

[21] FEUERVERGER, A. AND MCDUNNOUGH, P. (1981). On the efficiency of empirical characteristic function procedures. *J. R. Statist. Soc. B* **43**, 20–27.

[22] FEUERVERGER, A. AND MUREIKA, R. A. (1977). The empirical characteristic function and its applications. *Ann. Statist.* **5**, 88–97.

[23] FRÜHWIRTH-SCHNATTER, S. (2006). *Finite Mixture and Markov Switching Models*, Springer, New York.

[24] FU, M. C., LI, B., WU, R. AND ZHANG, T. (2022). Option pricing under a discrete-time Markov switching stochastic volatility with co-jump model. *Frontiers Math. Finance* **1**, 137–160. DOI: 10.3934/fmf.2021005.

[25] GEMMILL, G., and SAFLEKOS, A. (2000). How useful are implied distributions? Evidence from stock-index options. *J. Derivatives* **7**, 83–91. DOI: 10.3905/jod.2000.319123.

[26] GERBER, H. U. AND SHIU, E. S. W. (1994). Option pricing by Esscher transforms (with discussions). *Trans. Soc. Actuaries* **46**, 99–191.

[27] GERBER, H. U. AND SHIU, E. S. W. (1995). An actuarial bridge to option pricing. In *Securitization of Insurance Risk: The 1995 Bowles Symposium*. The Society of Actuaries, United States, pp. 45–62.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

[28] GERBER, H. U. AND SHIU, E. S. W. (1996). Actuarial bridges to dynamic hedging and option pricing. *Insurance Math. Econom.* **18**, 183–218.
[29] GODIN, F., LAI, V. S. AND TROTTIER, D. A. (2019). Option pricing under regime-switching models: Novel approaches removing path-dependence. *Insurance Math. Econom.* **87**, 130–142.
[30] GODIN, F., and TROTTIER, D. A. (2021). Option pricing in regime-switching frameworks with the extended Girsanov principle. *Insurance Math. Econom.* **99**, 116–129.
[31] GOOVAERTS, M. J. AND LAEVEN, R. J. A. (2008). Actuarial risk measures for financial derivative pricing. *Insurance Math. Econom.* **42**, 540–547.
[32] GUO, C. (2005). Option pricing with heterogeneous expectations. *Financial Rev.* **33**, 81–92.
[33] HALL, P. AND TITTERINGTON, D. M. (1984). Efficient nonparametric estimation of mixture proportions. *J. R. Statist. Soc. B* **46**, 465–473.
[34] HAMILTON, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**, 357–384.
[35] HANSEN, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50**, 1029–1054.
[36] HARRISON, J. M. AND KREPS, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *J. Econom. Theory* **20**, 381–408.
[37] HARRISON, J. M. AND PLISKA, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stoch. Process. Appl.* **11**, 215–280.
[38] HARRISON, J. M. AND PLISKA, S. R. (1983). A stochastic calculus model of continuous trading: Complete markets. *Stoch. Process. Appl.* **15**, 313–316.
[39] HILLIARD, J. E. AND NGO, J. T. D. (2022). Bitcoin: Jumps, convenience yields, and option prices. *Quant. Finance* **22**, 2079–2091.
[40] HOU, A. J., WANG, W., CHEN, C. Y. H. AND HÄRDLE, W. K. (2020). Pricing cryptocurrency options. *J. Financial Econometrics* **18**, 250–279.
[41] ISHIJIMA, H. AND KIHARA, T. (2005). *Option pricing with hidden Markov models*. Institute of Finance, Waseda University, Tokyo.
[42] JALAN, A., MATKOVSKYY, R. AND AZIZ, S. (2021). The Bitcoin options market: A first look at pricing and risk. *Applied Economics* **53**, 2026–2041.
[43] KLIBANOFF, P., MARINACCI, M. AND MUKERJI, S. (2005). A smooth model of decision making under ambiguity. *Econometrica* **73**, 1849–1892.
[44] KNIGHT, J. L. AND YU, J. (2002). Empirical characteristic function in time series estimation. *Econometric Theory* **18**, 691–721.
[45] LI, L., ARAB, A., LIU, J., LIU, J. AND HAN, Z. (2019). Bitcoin options pricing using LSTM-based prediction model and blockchain statistics. In *Proc. 2019 IEEE Int. Conf. Blockchain*, pp. 67–74.
[46] LIEW, C. C. AND SIU, T. K. (2010). A hidden Markov regime-switching model for option valuation. *Insurance Math. Econom.* **47**, 374–384.
[47] LIU, X., SHACKLETON, M. B., TAYLOR, S. J. AND XU, X. (2007). Closed-form transformations from risk-neutral to real-world distributions. *J. Banking Finance* **31**, 1501–1520.
[48] MADAN, D. B. AND SENETA, E. (1987). Simulation of estimates using the empirical characteristic function. *Internat. Statist. Rev.* **55**, 153–161.
[49] MADAN, D. B. AND SENETA, E. (1990). The variance gamma (V.G.) model from share market returns. *J. Business* **63**, 511–524.
[50] MCLACHLAN, G. J., LEE, S. X. AND RATHNAYAKE, S. I. (2019). Finite mixture models. *Ann. Rev. Statist. Appl.* **6**, 355–378.
[51] MCLACHLAN, G. J. AND PEEL, D. (2000). *Finite Mixture Models*. John Wiley, Hoboken, NJ.
[52] MERTON, R. C. (1973). Theory of rational option pricing. *Bell J. Econom. Manag. Sci.* **4**, 141–183.
[53] MITCHELL, M. (1998). *An Introduction to Genetic Algorithms*. MIT Press.
[54] MONFORT, A. AND PEGORARO, F. (2012). Asset pricing with second-order Esscher transforms. *J. Banking* **36**, 1678–1687.
[55] PAGNOTTONI, P. (2019). Neural network models for Bitcoin option pricing. *Frontiers Artificial Intell.* **2**, 5. DOI: 10.3389/frai.2019.00005.
[56] PRICE, K. V., STORN, R. M. AND LAMPINEN J. A. (2006). *Differential Evolution: A Practical Approach To Global Optimization*. Springer, Berlin.
[57] RITCHEY, R. J. (1990). Call option valuation for discrete normal mixtures. *J. Financial Res.* **13**, 285–296.
[58] ROMBOUTS, J. V. K. and STENTOFT, L. (2014). Bayesian option pricing using mixed normal heteroskedasticity models. *Comput. Statist. Data Anal.* **76**, 588–605.
[59] ROMBOUTS, J. V. K. and STENTOFT, L. (2015). Option pricing with asymmetric heteroskedastic normal mixture models. *Int. J. Forecasting* **31**, 635–650.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press

[60] RYDÉN, T., TERÄSVIRTA, T. AND ÅSBRINK, S. (1998). Stylized facts of daily return series and the hidden Markov model. *J. Appl. Econometrics* **13**, 217–244.
[61] SINGLETON, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *J. Econometrics* **102**, 111–141.
[62] SIU, T. K. (2005). Fair valuation of participating policies with surrender options and regime switching. *Insurance Math. Econom.* **37**, 533–552.
[63] SIU, T. K. (2011). Regime-switching risk: To price or not to price? *Internat. J. Stoch. Anal.* **2011**, 843246.
[64] SIU, T. K. (2023). Bayesian nonlinear expectation for time series modelling and its application to Bitcoin. *Empirical Economics* **64**, 505–537.
[65] SIU, T. K. (2024). Bayesian lower and upper estimates for Ether option prices with conditional heteroscedasticity and model uncertainty. *J. Risk Financial Manag.* **17**, 436.
[66] SIU, T. K. (2025). Market consistent valuation for Bitcoin options with long memory in conditional volatility and conditional non-normality. *J. Futures Markets* **45**, 917–945.
[67] SIU, T. K. AND ELLIOTT, R. J. (2021). Bitcoin option pricing with a SETAR-GARCH model. *Europ. J. Finance* **27**, 564–595.
[68] SIU, T. K., FUNG, E. S. AND NG, M. K. (2011). Option valuation with a discrete-time double Markovian regime-switching model. *Appl. Math. Finance* **18**, 473–490.
[69] SIU, T. K., TONG, H. AND YANG, H. (2001). Bayesian risk measures for derivatives via random Esscher transform. *North Amer. Actuarial J.* **5**, 78–91.
[70] SIU, T. K., TONG, H. AND YANG, H. (2004). On pricing derivatives under GARCH models: A dynamic Gerber–Shiu approach. *North Amer. Actuarial J.* **8**, 17–31.
[71] SIU, T. K., TONG, H. AND YANG, H. (2006). Option pricing under threshold autoregressive models by threshold Esscher transform. *J. Indust. Manag. Optim.* **2**, 177–197.
[72] STORN, R. AND PRICE, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.* **11**, 341–359.
[73] TAYLOR, S. J. (2005). *Asset Price Dynamics, Volatility, And Prediction*. Princeton University Press.
[74] TAYLOR, S. J. (2008). *Modelling Financial Time Series*, 2nd edn. World Scientific, Singapore.
[75] TITTERINGTON, D. M., SMITH, A. F. M. AND MAKOV, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley, Chichester.
[76] TONG, H. (1983). *Threshold Models In Non-Linear Time Series Analysis*. Springer, New York.
[77] TONG, H. (1990). *Non-Linear Time Series: A Dynamical System Approach*. Oxford University Press.
[78] TONG, H. (2007). Exploring volatility from a dynamical system perspective. In: *Proc. 56th Session Int. Statist. Institute*. International Statistical Institute, The Hague.
[79] TRAN, K. C. (1998). Estimating mixtures of normal distributions via empirical characteristic function. *Econometric Rev.* **17**, 167–183.
[80] TSAY, R. S. (2012). *An Introduction To Analysis Of Financial Data With R*. John Wiley, Hoboken, NJ.
[81] VENTER, P. J., and MARÉ, E. (2021). Univariate and multivariate GARCH models applied to Bitcoin futures option pricing. *J. Risk Financial Manag.* **14**, 261.
[82] WILKENS, S. (2005). Option pricing based on mixtures of distributions: Evidence from the Eurex index and interest rate futures options market. *Derivatives Use, Trading Regulation* **11**, 213–231.
[83] YU, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Rev.* **23**, 93–123.

https://doi.org/10.1017/jpr.2025.10048 Published online by Cambridge University Press