

Jump risk premia in the presence of clustered jumps

Francis Liu*
 Natalie Packham†
 Artur Sepp‡

October 27, 2025

Abstract

This paper presents an option pricing model that incorporates clustered jumps using a bivariate Hawkes process. The process captures both self- and cross-excitation of positive and negative jumps, enabling the model to generate return dynamics with asymmetric, time-varying skewness and to produce positive or negative implied volatility skews. This feature is especially relevant for assets such as cryptocurrencies, so-called “meme” stocks, G-7 currencies, and certain commodities, where implied volatility skews may change sign depending on prevailing sentiment. We introduce two additional parameters, namely the positive and negative jump premia, to model the market risk preferences for positive and negative jumps, inferred from options data. This enables the model to flexibly match observed skew dynamics. Using Bitcoin (BTC) options, we empirically demonstrate how inferred jump risk premia exhibit predictive power for both the cost of carry in BTC futures and the performance of delta-hedged option strategies.

JEL classification: D81, G13

Keywords: Volatility risk premium, clustered jumps, Hawkes process, cryptocurrencies

1 Introduction

Jumps are acknowledged as an important risk factor in asset pricing, derivatives pricing, and risk management. Although there is a large body of literature investigating jumps in financial markets for traditional assets, the emergence of option markets for cryptocurrencies and retail-driven trading activity in options on “meme” stocks provide a new environment for this field of study. In this paper, we focus on the dynamics and options market for cryptocurrencies and, specifically, for Bitcoin (BTC).

Empirically, \mathbb{P} -dynamics of BTC exhibit both large positive and negative jumps, while option prices under \mathbb{Q} -dynamics exhibit changes in implied volatility skew from negative (when demand for protective put options increases) to positive (when demand for upside leverage rises).

Figure 1 illustrates the dynamics of Bitcoin (BTC) daily returns, at-the-money (ATM) implied volatilities and call and put skews from April 2019 to May 2024. First, we observe extreme returns

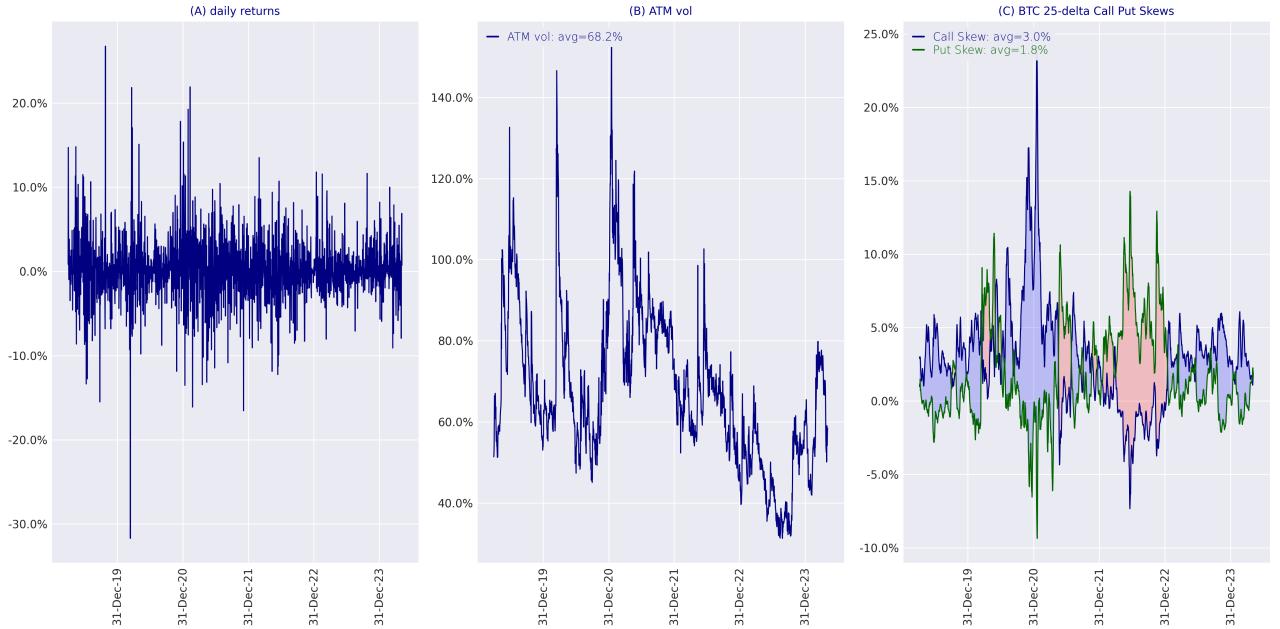
*Department of Business and Economics, Berlin School of Economics and Law, Badensche Str. 52, 10825 Berlin, Germany. Blockchain Research Center, Humboldt-Universität zu Berlin, Germany. International Research Training Group 1792, Humboldt-Universität zu Berlin, Germany. E-mail: francisliutfp@gmail.com.

†Department of Business and Economics, Berlin School of Economics and Law, Badensche Str. 52, 10825 Berlin, Germany. International Research Training Group 1792, Humboldt-Universität zu Berlin, Germany. E-mail: packham@hwr-berlin.de.

‡LGT Bank (Schweiz) AG E-mail: artursepp@gmail.com.

ranging from -30% to 20% , with significant clustering. The implied at-the-money volatility for options with maturity of one month and Black-Scholes delta of 0.5 ranges from 40% to 140% . The call skew is computed as the difference between implied volatility of a call option with 0.25 Black-Scholes delta and implied volatility of a call with 0.5 delta. The put skew is computed as the difference between implied volatility of a put option with -0.25 Black-Scholes delta and the implied volatility of a put with -0.5 delta. Positive call and put skews indicate the higher demand for calls and puts, respectively. We observe that on average both call and put skews are positive. In addition, call skews exceed put skews in most periods, indicating a higher demand for calls than for puts. This behaviour contradicts traditional patterns in options on stocks and stock indices where the call skew is negative, while the put skew is consistently positive.

Figure 1: Subplot (A) shows Bitcoin daily returns from April 2019 to May 2024 with extreme volatility and pronounced clustering during market stress periods. Subplot shows ATM Volatility daily with average of 68.3% . Subplot (C) shows BTC 25-Delta Call/Put Skews with corresponding averages of 3.0% and 1.7% and with blue/pink regions corresponding to periods with relatively higher call/put skews respectively.



To account for empirical features of crypto options markets, we develop a derivatives pricing model that incorporates two Hawkes processes, one for positive and one for negative jumps. These processes capture the clustering of jumps, driven by self- and cross-excitation, effectively capturing the structure of jump arrival times. The model can be calibrated to the joint dynamics of BTC prices and of BTC implied volatility surfaces. In particular, we explore the resulting positive and negative jump risk premia, defined as the difference in expected jump magnitudes between the statistical measure and the risk-neutral measures.

Importantly, our model accommodates changing market preferences for skewness risk, when investors shift their demands between put options to hedge the downside risk and call options to capture positive upside. The preference for call options with upside potential is consistent with evidence from the equity market literature indicating some investors' preferences for assets offering lottery-like pay-offs, see, e.g. Bali et al. (2011), Thaler and Ziemba (1988), Kumar (2009). In cryptocurrency option markets, such preferences for capturing the upside manifest in inflated prices of out-of-the-money call options during bullish periods, which produces positive risk premia of positive jumps.

Furthermore, the positive implied volatility skew is also present in the market data of so-called “meme” stocks. Options activity in these stocks is characterised by strong demand and high volumes for call options and by a positive spot-volatility correlation. Gupta and Pascale (2025) introduce the “Equity Euphoria Indicator”, which measures the percentage of stocks with a positive implied skew ranging from 0% to 100%. The most recent values, 10% as of January 2025 and 11.9% as of September 2025, are close to the levels observed during the Dot.com bubble in late 1990 and the post-Covid rally in 2021. We notice that implied skews on “meme” stocks change sign to negative during severe bear market periods, so that structurally we cannot model the dynamics of these stocks using a stochastic volatility model with a constant correlation parameter. Adding jumps arriving at a constant intensity rate, as for the Bates model (Bates (1996)), would not improve the structural consistency of the model dynamics, unless we specify stochastic intensity rates for positive and negative jumps. Thus, our model represents a simple construction to model the dynamics of stocks with varying signs of implied skew. It is straightforward to incorporate stochastic volatility in our framework; however, we skip this feature to reduce the number of model parameters.

Our work contributes to the existing literature in several areas. First, the model of the price process allows one to identify and understand the factors that drive variations in jump risk premia, in particular, how jump intensities adapt in response to the occurrence of jumps. Extending the work of Hainaut (2016, 2017) on univariate Hawkes process, we assume that the dynamics of these jump intensities are governed by a bivariate Hawkes process. Jump intensities increase with jump arrivals and decrease with time. As such, the model accounts for changes in asset prices and option prices stemming not only from movements in the underlying price, but also from variations in jump intensities.


Second, alongside specifying the pricing model, we derive the measure change between the statistical and risk-neutral measures. This allows us to derive the jump risk premia associated with positive and negative jumps as the difference between expected jump magnitudes under the risk-neutral and statistical measures. A jump risk premium therefore quantifies the difference between market expectations (from options prices) and market realisations (from the price history) and can be interpreted as a measure of market sentiment. As option prices (under the risk-neutral measure) reflect expectations about future jumps and their intensities, in our setting, investors pay a risk premium for call options due to positive jumps if they are optimistic (positive skewness of option prices). Likewise, investors pay a risk premium for put options due to negative jumps if they are pessimistic (negative skewness of option prices).

We further develop an estimation and calibration procedure that takes into account both historical BTC price dynamics and options data. In the time period considered (May 2019 to October 2023), both positive and negative risk premia are observed, with frequent periods where the implied jump risk is higher than the realised jump risk. The last observation is consistent with the findings on the variance risk premia, e.g. (Carr and Wu, 2009; Granelli and Veraart, 2016).

Third, the cryptocurrency market is known to have at times a large basis or cost-of-carry in its futures prices, as well as large swings in the basis. For example, the annualised basis of one-month BTC futures on Deribit on 5 April 2021 was 35%, and – after a sell-off on 19 May 2021 – dropped to –54%. Both positive and negative carry are regularly observed in the market. These can only partially be explained by the funding costs associated with perpetual futures. Therefore, we investigate the relationship between the inferred jump risk premia and the carry of one-month BTC futures to determine whether risk preferences, beyond standard funding costs from perpetual contracts, significantly influence futures pricing dynamics.

Finally, the presence of jumps imposes challenges in risk management and hedging, both in terms of the magnitude of losses and in the unpredictability of jumps. Hedging strategies in crypto markets have been investigated by (e.g. Liu et al., 2023; Matic et al., 2023; Lucic and Sepp, 2024). Here, we investigate whether the risk premia, which capture market participants' forward-looking risk preferences, can be systematically earned through delta-hedged options trading strategies.

The Hawkes process, or self-exciting process, was introduced by (Hawkes, 1971; Hawkes and Oakes, 1974), see also (Errais et al., 2010; Hawkes, 2018; Bacry et al., 2015; Embrechts et al., 2011; Hainaut and Moraux, 2018; Hainaut, 2016) for reviews and applications of Hawkes processes in finance. There is a vast literature on jump risk and jump risk premia, (e.g. Bates, 1996, 2000; Pan, 2002; Bakshi and Kapadia, 2003; Eraker, 2004; Broadie et al., 2007; Santa-Clara and Yan, 2010; Christoffersen et al., 2012; Nguyen et al., 2019). Skewness and their relation to risk premia are treated in (e.g. Bakshi et al., 2003) with a focus on the options market, and from a behavioural perspective in (Bali et al., 2011, e.g.). The literature on risk premia in cryptocurrency markets is growing (e.g. Alexander and Imeraj, 2021, 2023; Almeida et al., 2024; Dobrynskaya, 2024).

The paper is organised as follows. In Section 2, we develop the model and derive a number of results, such as moment-generating functions and the equivalent measure change. Section 3 introduces the estimation method that accompanies the model. In Section 4, we specify the jump risk premia. Section 5 presents the empirical results, that is, exploring the historical jump risk premia, the relationship between the jump risk premia and the futures basis, as well as the ability of the jump risk premia to forecast one-day ahead option portfolio returns that can be used to improve hedging strategies. All calculations in this work can be reproduced with codes available in Quantlet .

2 Model specification

2.1 Price dynamics under \mathbb{P} -measure

Our model extends the univariate Hawks process dynamics considered by Hainaut and Moraux (2018) (also see Hainaut (2016), Hainaut (2017), and chapter 4 of Hainaut (2022) for variants of the univariate model adapted in different situations). To account for the empirical evidence of clustered positive and negative jumps, we incorporate two jump components with self-excitement and cross-excitement.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with information filtration (\mathcal{F}_t) . The adapted price process S_t is modeled as the sum of a continuous part driven by a Brownian motion and two jump components for positive and negative jumps, respectively. The intensities of the jump components are latent processes λ_t^+ and λ_t^- . The arrival times of positive and negative jumps are modeled by two counting processes, abbreviated by $N_t^{(1)}$ and $N_t^{(2)}$, respectively.

We introduce the following price dynamics under the statistical \mathbb{P} measure:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{1}$$

$$+ \left((e^{J^+} - 1) dN_t^{(1)} - \lambda_t^+ \mathbb{E}(e^{J^+} - 1) dt \right) + \left((e^{J^-} - 1) dN_t^{(2)} - \lambda_t^- \mathbb{E}(e^{J^-} - 1) dt \right),$$

$$\begin{pmatrix} d\lambda_t^+ \\ d\lambda_t^- \end{pmatrix} = \begin{pmatrix} \kappa^+ (\theta^+ - \lambda_t^+) \\ \kappa^- (\theta^- - \lambda_t^-) \end{pmatrix} dt + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} J^+ dN_t^{(1)} \\ J^- dN_t^{(2)} \end{pmatrix}, \tag{2}$$

where μ is a constant drift parameter, σ is a constant volatility parameter, W_t is a standard Brownian

motion, $N_t^{(1)}$ and $N_t^{(2)}$ are Poisson processes for the arrival of positive and negative jumps, respectively. Jump sizes J^+ and J^- of positive and negative jumps with probability density functions (PDFs) ϖ^+ and ϖ^- , respectively, given by shifted exponential distributions ϖ^+ and ϖ^- . The jump parts in dynamics (1) are adjusted by deterministic jump compensators, so the expected drift is determined by μ . The coefficients β_{ij} , $i, j = 1, 2$, determine the level of excitement induced by jumps on the intensities. Their signs are set by $\beta_{11}, \beta_{21} \geq 0$ and $\beta_{12}, \beta_{22} \leq 0$, so that all jump realisations lead to jumps in intensity rates. We impose the following constraints on the jump parameters:

$$\kappa^+ \geq \beta_{11}\mathbb{E}J^+ + \beta_{12}\mathbb{E}J^-, \quad \kappa^- \geq \beta_{21}\mathbb{E}J^+ + \beta_{22}\mathbb{E}J^-. \quad (3)$$

These restrictions guarantee that the expected asymptotic values of the intensities are finite, i.e., $\lim_{t \rightarrow \infty} \mathbb{E}\lambda_t^+ < \infty$ and $\lim_{t \rightarrow \infty} \mathbb{E}\lambda_t^- < \infty$. For further details, we refer to Appendix A.1.

The solution to dynamics in (1) is given by:

$$S_t = S_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right) \cdot \exp \left(\left(\sum_{j=1}^{N_t^{(1)}} J_j^- - \mathbb{E}(\mathbf{e}^{J^+} - 1) \int_0^t \lambda_s^+ ds \right) + \left(\sum_{j=1}^{N_t^{(2)}} J_j^- - \mathbb{E}(\mathbf{e}^{J^-} - 1) \int_0^t \lambda_s^- ds \right) \right), \quad (4)$$

$$\lambda_t^+ = \theta^+ + \mathbf{e}^{-\kappa^+ t} (\lambda_0^+ - \theta^+) + \beta_{11} \int_0^t \mathbf{e}^{-\kappa^+(t-s)} J_s^+ dN_s^{(1)} + \beta_{12} \int_0^t \mathbf{e}^{-\kappa^+(t-s)} J_s^- dN_s^{(2)}, \quad (5)$$

$$\lambda_t^- = \theta^- + \mathbf{e}^{-\kappa^- t} (\lambda_0^- - \theta^-) + \beta_{21} \int_0^t \mathbf{e}^{-\kappa^-(t-s)} J_s^+ dN_s^{(1)} + \beta_{22} \int_0^t \mathbf{e}^{-\kappa^-(t-s)} J_s^- dN_s^{(2)}. \quad (6)$$

We also specify the dynamics of the log return $X_t := \ln(S_t/S_0)$ which is given by:

$$dX_t = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t + \left(J^+ dN_t^{(1)} - \lambda_t^+ \mathbb{E}(\mathbf{e}^{J^+} - 1) dt \right) + \left(J^- dN_t^{(2)} - \lambda_t^- \mathbb{E}(\mathbf{e}^{J^-} - 1) dt \right). \quad (7)$$

In summary, the model incorporates three state variables $(X_t, \lambda_t^+, \lambda_t^-)$. The model parameters include a drift μ and the volatility σ of the continuous component; eight parameters for the intensities of positive and negative jumps $(\kappa^+, \theta^+, \beta_{11}, \beta_{12})$ and $(\kappa^-, \theta^-, \beta_{21}, \beta_{22})$, respectively; four parameters for PDFs for jumps in returns: $\varpi^+(j^+)$ and $\varpi^-(j^-)$, as defined next.

We assume that the sizes of positive and negative jumps are driven by shifted exponential distributions with the following PDFs:

$$\begin{aligned} \varpi^+(j) &= \frac{1}{\eta^+} \exp \left(-\frac{1}{\eta^+} (j - \nu^+) \right) \mathbb{1}_{\{j > \nu^+\}}, \\ \varpi^-(j) &= \frac{1}{\eta^-} \exp \left(\frac{1}{\eta^-} (j - \nu^-) \right) \mathbb{1}_{\{j < \nu^-\}}, \end{aligned} \quad (8)$$

with $0 < \eta^+$ and $0 < \eta^-$ controlling the sizes of negative and positive jumps, respectively, and with $\nu^+ > 0$ and $\nu^- < 0$ being thresholds. The Laplace transforms of the jump size distributions (extended

to \mathbb{R}_- for negative jumps) are given by

$$\mathcal{L}^{(+)}(\omega) = \int_{\nu^+}^{\infty} e^{-\omega j} \varpi^+(j) dj = \frac{e^{-\nu^+ \omega}}{1 + \eta^+ \omega}, \quad \text{if } \operatorname{Re}(1/\eta^+ + \omega) > 0 \quad (9)$$

$$\mathcal{L}^{(-)}(\omega) = \int_{-\infty}^{\nu^-} e^{-\omega j} \varpi^-(j) dj = \frac{e^{-\nu^- \omega}}{1 - \eta^- \omega}, \quad \text{if } \operatorname{Re}(\omega) < \operatorname{Re}(1/\eta^-). \quad (10)$$

where $\omega \in \mathbb{C}$ and $\operatorname{Re}(\omega)$ denotes the real part.

Jump compensators are calculated using Equations (9) and (10) as follows:

$$\begin{aligned} \mathbb{E} \left(e^{J^+} - 1 \right) &= \int_{\nu^+}^{\infty} (e^j - 1) \varpi^+(j) dj = \frac{e^{\nu^+}}{1 - \eta^+} - 1, \\ \mathbb{E} \left(e^{J^-} - 1 \right) &= \int_{-\infty}^{\nu^-} (e^j - 1) \varpi^-(j) dj = \frac{e^{\nu^-}}{1 + \eta^-} - 1. \end{aligned}$$

We specify threshold parameters ν^+ and ν^- for flexibility in defining the minimum size of jumps for empirical inference. The thresholds $\hat{\nu}^+$ and $\hat{\nu}^-$ can be estimated using a peak-over-threshold (POT) approach discussed in Section 3.1.

2.2 Moment-generating function of joint dynamics

The moment generating function (MGF) of the log-return variable $X_t = \ln(S_t/S_0)$ and jump intensities will be used for the equivalent measure change in Section 2.3 and for the option pricing procedure in Section 3.2. The following Proposition extends Proposition 2.2. of Hainaut and Moraux (2018) to the bivariate Hawks process.

Proposition 1. *The MGF of joint variables $(X_T, \lambda_T^+, \lambda_T^-)$, $T \geq t$, generalised to complex arguments, is given by:*

$$\begin{aligned} G^{\mathbb{P}}(\omega, \omega^+, \omega^-; T, x, \lambda^+, \lambda^-) &\equiv \mathbb{E}^{\mathbb{P}} \left(\exp(\omega X_T + \omega^+ \lambda_T^+ + \omega^- \lambda_T^-) \middle| X_t, \lambda_t^+, \lambda_t^- \right) \\ &= \exp \left(A(t, T) + \omega X_t + C(t, T) \lambda_t^+ + D(t, T) \lambda_t^- \right), \end{aligned} \quad (11)$$

for $\omega, \omega^+, \omega^- \in \mathbb{C}$, where the functions A, C , and D solve the following non-linear system of ordinary differential equations (ODEs):

$$\begin{aligned} \frac{\partial}{\partial t} A(t, T) &= -\mu\omega - \frac{\sigma^2}{2} (\omega^2 - \omega) - \kappa^+ \theta^+ C(t, T) - \kappa^- \theta^- D(t, T), \\ \frac{\partial}{\partial t} C(t, T) &= -\mathcal{L}^{(+)}(-\omega - C(t, T)\beta_{11} - D(t, T)\beta_{21}) + \kappa^+ C(t, T) + \omega \left(\frac{e^{\nu^+}}{1 - \eta^+} - 1 \right) + 1, \\ \frac{\partial}{\partial t} D(t, T) &= -\mathcal{L}^{(-)}(-\omega - C(t, T)\beta_{12} - D(t, T)\beta_{22}) + \kappa^- D(t, T) + \omega \left(\frac{e^{\nu^-}}{1 + \eta^-} - 1 \right) + 1, \end{aligned} \quad (12)$$

where $A(T, T) = 0$, $C(T, T) = \omega^+$, $D(T, T) = \omega^-$ with operators \mathcal{L}^+ and \mathcal{L}^- defined in Equations 9 and 10, respectively.

See Appendix A for the proof.

2.3 Equivalent risk-neutral measure \mathbb{Q}

We now construct the model dynamics under an equivalent martingale measure \mathbb{Q} such that the discounted stock price process is a martingale. This is an extension of (Hainaut and Moraux, 2018) to

a bivariate self-exiting Hawks process. Introduce the Radon-Nikodym derivative process defined by:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \frac{M_t}{M_0}, \quad t \geq 0, \quad (13)$$

where $\{M_t\}_{t \geq 0}$ is a local martingale under \mathbb{P} , specified below. By construction, we have the following:

$$\mathbb{E}^{\mathbb{P}} \left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_T} \right) \right] = 1, \quad T \geq 0. \quad (14)$$

Define exponential processes of the form

$$\begin{aligned} M_t(\xi^+, \xi^-, \varphi) \stackrel{\text{def}}{=} & \exp \left\{ \kappa_1^+(\xi^+) \lambda_t^+ + \xi^+ \sum_{j=1}^{N_t^{(1)}} J_j^+ + \kappa_1^-(\xi^-) \lambda_t^- + \xi^- \sum_{j=1}^{N_t^{(2)}} J_j^- \right\} \\ & \cdot \exp \{ q_2(\xi^+, \xi^-) t \} \\ & \cdot \exp \left\{ -\frac{1}{2} \int_0^t \varphi^2(s) ds - \int_0^t \varphi(s) dW_s \right\}, \end{aligned} \quad (15)$$

where $\xi^+, \xi^- \in \mathbb{R}$ are two parameters that specify the risk-premia of positive and negative jumps, respectively, and that are to be estimated from options data. $q_1^+(\xi^+), q_1^-(\xi^-), q_2(\xi^+, \xi^-)$ are real-valued functions, and $\varphi(t)$ is an adapted process adapted to $\{\mathcal{F}_t\}_{t \geq 0}$. In Proposition 5 below, we will choose

$$\varphi(t) = \varphi + \varphi^+ \lambda_t^+ + \varphi^- \lambda_t^-, \quad (16)$$

where $\varphi, \varphi^+, \varphi^- \in \mathbb{R}$ are constants. We note that the process M_t is strictly positive, so that, provided it is a martingale, the probability measures \mathbb{P} and \mathbb{Q} are absolutely continuous. The following proposition specifies martingale conditions for $\{M_t\}_{t \geq 0}$.

Proposition 2. *If, for any ξ^+ and $\xi^- \in \mathbf{R}$, there exists a solution specified as functions $q_1^+(\xi^+)$, $q_1^-(\xi^-)$ and $q_2(\xi^+, \xi^-)$ of the following system:*

$$\begin{aligned} q_1^+(\xi^+) \kappa^+ \theta^+ + q_1^-(\xi^-) \kappa^- \theta^- + q_2(\xi^+, \xi^-) &= 0, \\ \mathcal{L}^{(+)}(-\chi^+) - 1 - q_1^+(\xi^+) \kappa^+ &= 0, \\ \mathcal{L}^{(-)}(-\chi^-) - 1 - q_1^-(\xi^-) \kappa^- &= 0, \end{aligned} \quad (17)$$

then M_t is a local martingale. Here, we introduce new variables χ^+ and χ^- as follows:

$$\begin{aligned} \chi^+ &= q_1^+(\xi^+) \beta_{11} + q_1^-(\xi^-) \beta_{21} + \xi^+, \\ \chi^- &= q_1^+(\xi^+) \beta_{12} + q_1^-(\xi^-) \beta_{22} + \xi^-. \end{aligned} \quad (18)$$

See Appendix A for the proof.

We note that the system of equations in Eq.17 is non-linear because of its variables entering into non-linear functions $\mathcal{L}^{(+)}$ and $\mathcal{L}^{(-)}$. First, we make the following assumption that the risk-premium functions are linear.

Assumption 2.1. *We assume that the real-valued functions are specified by means of a linear system*

of equations as follows:

$$\begin{aligned} q_1^+(\xi^+) &= c^+ + \xi^+ \\ q_1^-(\xi^-) &= c^- + \xi^- \\ q_2(\xi^+, \xi^-) &= c + \xi^+ + \xi^-, \end{aligned} \tag{19}$$

where $c^+, c^-, c \in \mathbb{R}$ are internal parameters which are determined so that the system in Eq (17) is satisfied for any choice of ξ^+ and ξ^- .

Second, we linearise the system of equations in (17), by making the following assumption.

Assumption 2.2. The variables χ^+ and χ^- in Eq. (18) are external variables that specify the risk premium, while variables ξ^+ and ξ^- become internal parameters inferred using Eq.(18).

Corollary 1. The specification of the risk-premium is defined by the two external parameters χ^+ and χ^- . The internal parameters of linear risk-premium functions in Eq. (19), internal risk premia ξ^+ and ξ^- parameters in Eq. (18) along the martingale conditions in Eq. (17) are determined by five internal parameters $(\xi^+, \xi^-, c^+, c^-, c)$ which are obtained by solving the following linear system:

$$\begin{pmatrix} 1 + \beta_{11} & \beta_{21} & \beta_{11} & \beta_{21} & 0 \\ \beta_{12} & 1 + \beta_{22} & \beta_{12} & \beta_{22} & 0 \\ \kappa^+ \theta^+ & \kappa^- \theta^- & \kappa^+ \theta^+ & \kappa^- \theta^- & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \xi^+ \\ \xi^- \\ c^+ \\ c^- \\ c \end{pmatrix} = \begin{pmatrix} \chi^+ \\ \chi^- \\ 0 \\ (\mathcal{L}^{(+)}(-\chi^+) - 1) / \kappa^+ \\ (\mathcal{L}^{(-)}(-\chi^-) - 1) / \kappa^- \end{pmatrix}. \tag{20}$$

Proposition 3. The system of linear equations in Eq. (20) in $(\xi^+, \xi^-, c^+, c^-, c)$ has a unique solution.

Proof. The determinant of the coefficient matrix in Eq.(20) is equal to one, so it is invertible and a solution to the linear system exists and is unique. \square

This chain of results establishes that the equivalent martingale measure \mathbb{Q} is fully characterised by the two external risk-premium parameters χ^+ and χ^- , which can be interpreted as market-implied jump risk premia for positive and negative jumps, respectively. Given any such pair (χ^+, χ^-) , the corresponding internal parameters $(\xi^+, \xi^-, c^+, c^-, c)$ that define the Radon-Nikodym derivative are uniquely determined by solving the linear system in Equation (20). The martingale property of the Radon-Nikodym derivative process M_t is therefore always guaranteed under this construction.

Crucially, the transformation of the model from the physical measure \mathbb{P} to the risk-neutral measure \mathbb{Q} depends only on the values of χ^+ and χ^- . This leads to a clean and tractable separation between the statistical and risk-neutral dynamics, which then enables model calibration and risk-premium estimation from observed option prices. As shown in Proposition 4, the jump intensity dynamics under \mathbb{Q} inherit a similar structure to those under \mathbb{P} , with appropriately adjusted parameters. Thus, the model maintains internal consistency and interpretability under both measures, while enabling flexible inference of jump risk premia from market data.

Proposition 4 (Jump intensity dynamics under the risk-neutral measure \mathbb{Q}). Let $N_t^{(1),\mathbb{Q}}$ and $N_t^{(2),\mathbb{Q}}$ be the counting processes of positive and negative jumps under the measure \mathbb{Q} with intensities

$$\lambda_t^{+,\mathbb{Q}} = \mathcal{L}^{(+)}(-\chi^+) \lambda_t^+, \quad \lambda_t^{-,\mathbb{Q}} = \mathcal{L}^{(-)}(-\chi^-) \lambda_t^-, \tag{21}$$

where χ^+ and χ^- are external risk-premium parameters in Eq.(18).

The evolution of jump intensity rates $\lambda_t^{+,Q}$ and $\lambda_t^{-,Q}$ under \mathbb{Q} are driven by the following dynamics:

$$\begin{aligned} d\lambda_t^{+,Q} &= \kappa^+ \left(\theta^{+,Q} - \lambda_t^{+,Q} \right) dt + \beta_{11}^{Q, \mathbb{Q}} J^{+,Q} dN_t^{(1),Q} + \beta_{12}^{Q, \mathbb{Q}} J^{-,Q} dN_t^{(2),Q} \\ d\lambda_t^{-,Q} &= \kappa^- \left(\theta^{-,Q} - \lambda_t^{-,Q} \right) dt + \beta_{21}^{Q, \mathbb{Q}} J^{+,Q} dN_t^{(1),Q} + \beta_{22}^{Q, \mathbb{Q}} J^{-,Q} dN_t^{(2),Q}, \end{aligned} \quad (22)$$

where

$$\theta^{+,Q} = \mathcal{L}^{(+)}(-\chi^+) \theta^+, \quad \beta_{11}^{Q, \mathbb{Q}} = \mathcal{L}^{(+)}(-\chi^+) \beta_{11}, \quad \beta_{12}^{Q, \mathbb{Q}} = \mathcal{L}^{(+)}(-\chi^+) \beta_{12}, \quad (23)$$

$$\theta^{-,Q} = \mathcal{L}^{(-)}(-\chi^-) \theta^-, \quad \beta_{21}^{Q, \mathbb{Q}} = \mathcal{L}^{(-)}(-\chi^-) \beta_{21}, \quad \beta_{22}^{Q, \mathbb{Q}} = \mathcal{L}^{(-)}(-\chi^-) \beta_{22}, \quad (24)$$

Here, $J^{+,Q}$ and $J^{-,Q}$ are realizations of positive and negative jumps from distributions with probability density functions (PDFs) $\varpi^{+,Q}(j^+)$ and $\varpi^{-,Q}(j^-)$:

$$\varpi^{+,Q}(j) = \frac{1}{\eta^{+,Q}} \exp \left(-\frac{1}{\eta^{+,Q}} (j - \nu^+) \right) \mathbb{1}_{\{j > \nu^+\}}, \quad \varpi^{-,Q}(j) = \frac{1}{\eta^{-,Q}} \exp \left(-\frac{1}{\eta^{-,Q}} (j - \nu^-) \right) \mathbb{1}_{\{j < \nu^-\}}$$

where

$$\eta^{+,Q} = \frac{\eta^+}{1 - \eta^+ \chi^+}, \quad \eta^{-,Q} = \frac{\eta^-}{1 + \eta^- \chi^-}. \quad (25)$$

See Appendix A for the proof. The proposition 4 indicates that the dynamics of the intensities and the PDFs of the jump sizes are specified by the two external parameters χ^+ and χ^- . Accordingly, for model calibration to implied volatility data, we need to infer a pair of value of χ^+ and χ^- that provides the best fit. The following proposition specifies $\varphi(t)$ such that an equivalent risk-neutral measure is obtained.

Proposition 5. *We introduce the adapted process $\varphi(t)$ as follows:*

$$\begin{aligned} \varphi(t) &= \frac{\mu - r}{\sigma} - \frac{\mathbb{E}(e^{J^+} - 1) - \mathcal{L}^{(+)}(-\chi^+) \mathbb{E}^{\mathbb{Q}}(e^{J^{+,Q}} - 1)}{\sigma} \lambda_t^+ - \frac{\mathbb{E}(e^{J^-} - 1) + \mathcal{L}^{(-)}(-\chi^-) \mathbb{E}^{\mathbb{Q}}(e^{J^{-,Q}} - 1)}{\sigma} \lambda_t^-, \\ &= \frac{\mu - r}{\sigma} - \frac{\lambda_t^+ \mathbb{E}(e^{J^+} - 1) - \lambda_t^{+,Q} \mathbb{E}^{\mathbb{Q}}(e^{J^{+,Q}} - 1)}{\sigma} - \frac{\lambda_t^- \mathbb{E}(e^{J^-} - 1) - \lambda_t^{-,Q} \mathbb{E}^{\mathbb{Q}}(e^{J^{-,Q}} - 1)}{\sigma}, \end{aligned} \quad (26)$$

where r is the risk free rate. Then the log return process $X_t = \ln(S_t/S_0)$ is driven under \mathbb{Q} by:

$$\begin{aligned} dX_t &= \left(r - \frac{\sigma^2}{2} \right) dt + \sigma dW_t^{\mathbb{Q}} \\ &\quad + \left(J^{+,Q} dN_t^{(1),Q} - \lambda_t^{+,Q} \mathbb{E}^{\mathbb{Q}}(e^{J^{+,Q}} - 1) dt \right) + \left(J^{-,Q} dN_t^{(2),Q} - \lambda_t^{-,Q} \mathbb{E}^{\mathbb{Q}}(e^{J^{-,Q}} - 1) dt \right), \end{aligned} \quad (27)$$

where $W^{\mathbb{Q}}$ is a standard Brownian motion under \mathbb{Q} , and the discounted price process is a \mathbb{Q} -martingale.

See Appendix A for the proof. The function $\varphi(t)$ specifies the risk premia associated with the continuous component of the process. We will discuss the jump risk premia in Section 4.

Corollary 2 (MGF in Proposition 1 under \mathbb{Q}). *The MGF of log-return X_T with the dynamics in*

Eq.(27) under \mathbb{Q} is given by:

$$\begin{aligned} G^{\mathbb{Q}}(\omega; T, x, \lambda^{+, \mathbb{Q}}, \lambda^{-, \mathbb{Q}}) &\equiv \mathbb{E}^{\mathbb{Q}} \left(\exp(\omega X_T) | X_t, \lambda_t^{+, \mathbb{Q}}, \lambda_t^{-, \mathbb{Q}} \right) \\ &= \exp \left(A(t, T) + \omega X_t + C(t, T) \lambda_t^{+, \mathbb{Q}} + D(t, T) \lambda_t^{-, \mathbb{Q}} \right) \\ &= \exp(\omega X_t) E(t, T; \omega), \end{aligned} \quad (28)$$

where functions A, C , and D solve Eq.(11) with $\omega^+ = 0$, $\omega^- = 0$ and with the following substitution of model parameters:

$$\begin{aligned} \mu &\rightarrow r, \\ \theta^+ &\rightarrow \theta^{+, \mathbb{Q}} = \mathcal{L}^{(+)}(-\chi^+) \theta^+, \quad \beta_{11} \rightarrow \beta_{11}^{\mathbb{Q}} = \mathcal{L}^{(+)}(-\chi^+) \beta_{11}, \quad \beta_{12} \rightarrow \beta_{12}^{\mathbb{Q}} = \mathcal{L}^{(+)}(-\chi^+) \beta_{12}, \\ \theta^- &\rightarrow \theta^{-, \mathbb{Q}} = \mathcal{L}^{(-)}(-\chi^-) \theta^-, \quad \beta_{21} \rightarrow \beta_{21}^{\mathbb{Q}} = \mathcal{L}^{(-)}(-\chi^-) \beta_{21}, \quad \beta_{22}^{\mathbb{Q}} \rightarrow \beta_{22}^{\mathbb{Q}} = \mathcal{L}^{(-)}(-\chi^-) \beta_{22}, \\ \eta^+ &\rightarrow \eta^{+, \mathbb{Q}} = \frac{\eta^+}{1 - \eta^+ \chi^+}, \quad \eta^- \rightarrow \eta^{-, \mathbb{Q}} = \frac{\eta^-}{1 + \eta^- \chi^-}. \end{aligned} \quad (29)$$

Corollary 3. Using the Lewis-Lipton formula (Lewis (2001), Lipton (2002)), the value of a call option, denoted by U^{call} , and a put option, denoted by U^{put} , with strike price K and maturity time T is computed for log-price dynamics in Eq.(27) under the risk-neutral measure \mathbb{Q} by:

$$\begin{aligned} U^{\text{call}}(\tau, S, K) &= e^{-r\tau} S - U(\tau, X, K), \\ U^{\text{put}}(\tau, S, K) &= e^{-r\tau} K - U(\tau, X, K), \end{aligned} \quad (30)$$

where $\tau = T - t$. Here, $U(\tau, X, K)$ is the price of the capped payoff $\min(S, K)$ which is computed by:

$$U(\tau, X, K) = \frac{e^{-r\tau} K}{\pi} \Re \left[\int_0^\infty e^{-(iy - 1/2)X^*} \frac{1}{y^2 + 1/4} E(\tau; \omega = iy - 1/2) dy \right], \quad (31)$$

where $X^* = \ln(S/K)$ is log-moneyness, $\Re(x)$ is the real part of x , and $E(\tau; \omega)$ is specified in Eq.(28).

For the proof, see the proof of Proposition 5.1. in Sepp and Rakhmonov (2023).

For numerical implementation, the system of ODEs in Eq.(11) must be solved numerically using standard Runge-Kutta methods. Given that the computation of the MGF part $E(\tau; \omega)$ does not depend on the strike price of the option, the function $E(\tau; \omega)$ can be computed on a grid and apply for the numerical calculation of call and put options at the same maturity time with different strike prices¹.

3 Two stage approach calibration of the model

We adopt the following two-stage calibration approach. First, we estimate the model parameters under statistical measure \mathbb{P} from historical price data. Then, we infer the risk premia and the model parameters under the risk-neutral measure \mathbb{Q} using the options data.

¹Github project StochVolModels provides Python code with prototype implementation of option valuation using Eqs (30) and (31), see https://github.com/ArturSepp/StochVolModels/blob/main/stochvolmodels/pricers/hawkes_jd_pricer.py.

3.1 Time series estimation

To filter out jumps from time series of returns, we implement a POT procedure similarly to the approach in Embrechts et al. (2011) and (Hainaut, 2022, Chapter 4). We consider a sample of equally-spaced log-returns observed in the time window $[0, T]$. The POT procedure labels the log-returns that exceed the threshold as jumps. The underlying assumption is that the continuous part of the price process has normally distributed log-returns which implies that sampled returns, filtered by excluding jumps, have zero skewness and zero excess kurtosis.

Let $(X^{(\nu^+, \nu^-)}) := \{X_t : \nu^- < X_t < \nu^+\}$ the set of log-returns bounded by ν^- and ν^+ . The estimate of the thresholds is then given by:

$$(\hat{\nu}^+, \hat{\nu}^-) = \underset{\nu^+, \nu^-}{\operatorname{argmin}} \left[|\operatorname{skew}(X^{(\nu^+, \nu^-)})| + |\operatorname{kurt}(X^{(\nu^+, \nu^-)})| \right], \quad (32)$$

where $\operatorname{skew}(X) = |X|^{-1} \sum_{x \in X} (x - \hat{\mu}_X)^3 / \hat{\sigma}_X^3$, $\operatorname{kurt}(X) = |X|^{-1} \sum_{x \in X} (x - \hat{\mu}_X)^4 / \hat{\sigma}_X^4 - 3$, $\hat{\mu}_X$ and $\hat{\sigma}_X$ are the sample mean and standard deviation, respectively, and $|X|^{-1}$ is the sample size. Given estimated thresholds, we identify the set of positive jump events \mathcal{J}^+ and the set of negative jump events \mathcal{J}^- respectively as follows:

$$\begin{aligned} \mathcal{J}^+ &= \{J^+(s)\}_{s \leq T} = \{X_s \in X : X_s \geq \hat{\nu}^+ \text{ and } s \leq T\}, \\ \mathcal{J}^- &= \{J^-(s)\}_{s \leq T} = \{X_s \in X : X_s \leq \hat{\nu}^- \text{ and } s \leq T\}. \end{aligned} \quad (33)$$

We construct the counting processes for positive jumps and for the negative jumps as follows:

$$\begin{aligned} \widehat{N^{(1)}} &= \# \{X_s \in X : X_s \geq \hat{\nu}^+ \text{ and } s \leq T\}, \\ \widehat{N^{(2)}} &= \# \{X_s \in X : X_s \leq \hat{\nu}^- \text{ and } s \leq T\}. \end{aligned} \quad (34)$$

We denote the corresponding sequences of arrival times of positive jumps \mathcal{T}^+ , of negative jumps \mathcal{T}^- , and an ordered union of jump arrival times \mathcal{T}^\pm by:

$$\begin{aligned} \mathcal{T}^+ &= \{t \in [0, T] : X_t \geq \hat{\nu}^+\} = \left\{ T_1^+, T_2^+, \dots, T_{\widehat{N^{(1)}}}^+ \right\}, \\ \mathcal{T}^- &= \{t \in [0, T] : X_t \leq \hat{\nu}^-\} = \left\{ T_1^-, T_2^-, \dots, T_{\widehat{N^{(2)}}}^- \right\}, \\ \mathcal{T}^\pm &= \left\{ T_{[1]}^\pm, T_{[2]}^\pm, \dots, T_{[\widehat{N^{(1)}} + \widehat{N^{(2)}}]}^\pm \right\}. \end{aligned} \quad (35)$$

Jumps sizes are independent of other random variables and follow shifted exponential distributions. Therefore, given the shift parameter estimates $\hat{\nu}^+$ and $\hat{\nu}^-$ from the POT procedure, the estimators of η^+ and η^- are given by averaging:

$$\begin{aligned} \hat{\eta}^+ &= \widehat{N^{(1)}}^{-1} \sum_{j \in \mathcal{J}^+} (j - \hat{\nu}^+), \\ \hat{\eta}^- &= -\widehat{N^{(2)}}^{-1} \sum_{j \in \mathcal{J}^-} (j + \hat{\nu}^-). \end{aligned} \quad (36)$$

Next, we estimate the parameters of the intensities dynamics via the maximum likelihood estimator (MLE). We deploy a version of the multivariate log-likelihood function documented in Embrechts et al.

(2011) as follows:

$$\begin{aligned} \ln L' = & \sum_{T^+ \in \mathcal{T}^+} \ln \lambda^+(T^+ -) \varpi^+(J^+(T^+)) + \sum_{T^- \in \mathcal{T}^-} \ln \lambda^-(T^- -) \varpi^-(J^-(T^-)) \\ & - \int_0^T \lambda^+(t-) dt - \int_0^T \lambda^-(t-) dt. \end{aligned} \quad (37)$$

We note that the likelihood takes the left-continuous version (indicated by $T-$) of the intensities processes (see (Daley and Vere-Jones, 2003, p. 232)). Since distributions of jumps sizes are already estimated throughout the POT procedure using Eq.(36), we only need a partial likelihood for the intensity processes defined by:

$$\ln L = \sum_{T^+ \in \mathcal{T}^+} \ln \lambda^+(T^+ -) + \sum_{T^- \in \mathcal{T}^-} \ln \lambda^-(T^- -) - \int_0^T \lambda^+(t-) dt - \int_0^T \lambda^-(t-) dt. \quad (38)$$

The relationship between the intensities and the model parameters is specified for $s \in [T_{[k-1]}^\pm, T_{[k]}^\pm)$ and $k \in \{1, 2, \dots, \widehat{N}^{(1)} + \widehat{N}^{(2)}\}$ as follows:

$$\lambda^+(s) = \theta^+ + e^{-\kappa^+(s - T_{[k-1]}^\pm)} \left(\lambda^+(T_{[k-1]}^\pm) - \theta^+ \right), \quad (39)$$

$$\lambda^-(s) = \theta^- + e^{-\kappa^-(s - T_{[k-1]}^\pm)} \left(\lambda^-(T_{[k-1]}^\pm) - \theta^- \right). \quad (40)$$

In the event of positive jump $T^+ \in \mathcal{T}^+$, the intensities jump by:

$$\begin{aligned} \lambda^+(T^+) &= \lambda^+(T^+ -) + \beta_{11} J^+(T^+), \\ \lambda^-(T^+) &= \lambda^-(T^+ -) + \beta_{21} J^+(T^+). \end{aligned} \quad (41)$$

In the event of a negative jump $T^- \in \mathcal{T}^-$, the intensities jump by:

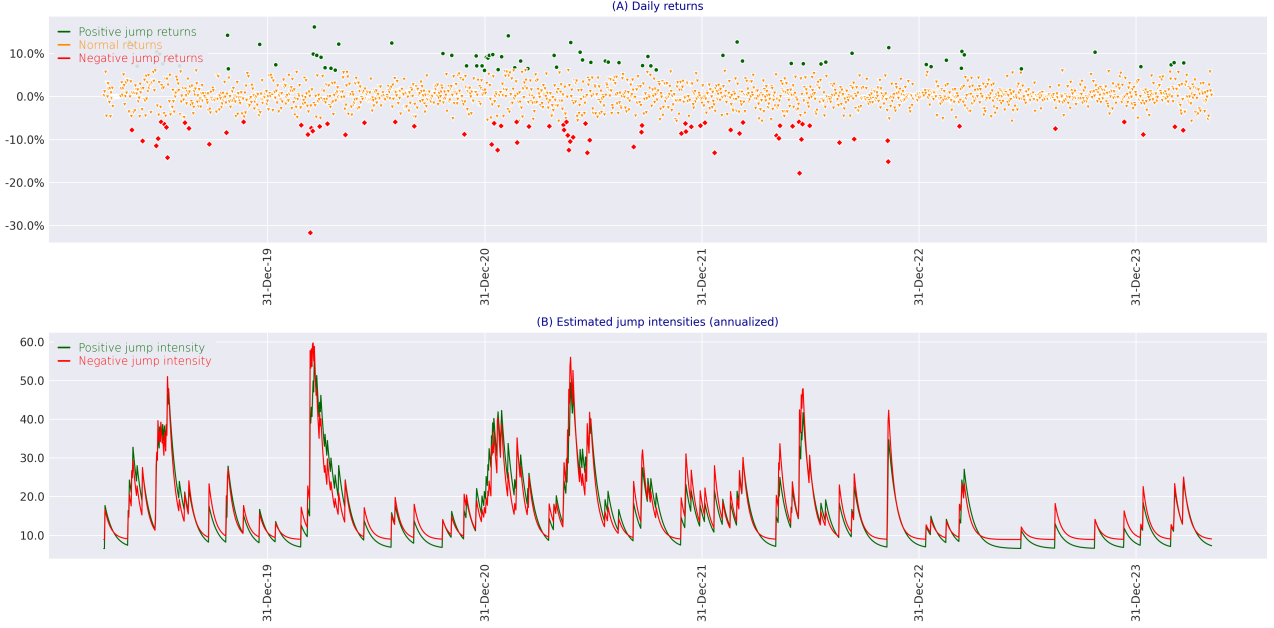
$$\begin{aligned} \lambda^+(T^-) &= \lambda^+(T^- -) + \beta_{12} J^-(T^-), \\ \lambda^-(T^-) &= \lambda^-(T^- -) + \beta_{22} J^-(T^-). \end{aligned} \quad (42)$$

The integrals of the intensities in the partial likelihood Eq.(38) can be computed as follows:

$$\begin{aligned} \int_0^T \lambda^+(t-) dt &= \sum_{k=1}^{\widehat{N}^{(1)} + \widehat{N}^{(2)}} \int_{T_{[k-1]}^\pm}^{T_{[k]}^\pm} \lambda^+(t-) dt \\ &= \sum_{k=1}^{\widehat{N}^{(1)} + \widehat{N}^{(2)}} \theta^+ (T_{[k]}^\pm - T_{[k-1]}^\pm) + \left(\lambda^+(T_{[k]}^\pm) - \theta^+ \right) \frac{1 - e^{-\kappa^+(T_{[k]}^\pm - T_{[k-1]}^\pm)}}{\kappa^+}, \\ \int_0^T \lambda^-(t-) dt &= \sum_{k=1}^{\widehat{N}^{(1)} + \widehat{N}^{(2)}} \theta^- (T_{[k]}^\pm - T_{[k-1]}^\pm) + \left(\lambda^-(T_{[k]}^\pm) - \theta^- \right) \frac{1 - e^{-\kappa^-(T_{[k]}^\pm - T_{[k-1]}^\pm)}}{\kappa^-}. \end{aligned} \quad (43)$$

Therefore, given a set of model parameters, the intensities and their integrals can be computed quickly in a recursive way starting from time 0. We refer to Section 5.2 of Laub et al. (2021) and references therein for the method of directly computing the likelihood. We apply a numerical optimiser to obtain estimates of the jump parameters, $\hat{\kappa}^\pm$, $\hat{\theta}^\pm$, $\{\hat{\beta}_{ij}\}_{i,j=1,2}$, that maximise the likelihood. Figure 2 illustrates the estimation procedure applied to BTC daily returns from 2019-05-30 to 2023-10-03.

Figure 2: Subplot (A) shows BTC daily returns which are classified using the POT procedure as positive jumps (green), negative jumps (red), or non-jumps (yellow). Subplot (B) displays the model intensities estimated via MLE, calibrated using the interarrival times and magnitudes of the jumps identified by POT.



3.2 Calibration to market implied volatilities

Given a set of European option prices and corresponding set of implied volatilities, we deduce volatility σ , and risk-premia χ^+ and χ^- by minimising the mean absolute percentage error (MAPE) as follows:

$$(\hat{\sigma}, \hat{\chi}^+, \hat{\chi}^-) = \underset{\sigma, \chi^+, \chi^-}{\operatorname{argmin}} \sum_{i=1}^I w_i \frac{|\operatorname{IV}_{\text{market}}(\tau_i, K_i, O_i) - \operatorname{IV}_{\text{model}}(\tau_i, K_i, O_i; \sigma, \chi^+, \chi^-)|}{\operatorname{IV}_{\text{market}}(\tau_i, K_i, O_i)},$$

where I is the sample size, w_i is the weight applied to the i^{th} option, $\operatorname{IV}_{\text{market}}(\tau_i, K_i, O_i)$ is the Black-Scholes (BS) implied volatility of the i^{th} option with time-to-maturity τ_i , strike K_i , and option type $O_i \in \{\text{Call}, \text{Put}\}$. Here, $\operatorname{IV}_{\text{model}}(\tau_i, K_i, O_i; \sigma, \chi^+, \chi^-)$ is the corresponding BS implied volatility produced by our model given σ , χ^+ , and χ^- . We take the market BS vega of each option as weight w , thus assigning greater importance to options whose prices are more sensitive to mispricing of implied volatilities.

The computation of model implied volatilities given σ , χ^+ and χ^- consists of the following steps.

1. We compute $\lambda_t^{\pm, \mathbb{Q}}$, $\kappa^{\pm, \mathbb{Q}}$, $\theta^{\pm, \mathbb{Q}}$, $\beta_{ij}^{\mathbb{Q}}$ for $i, j \in [1, 2]$, and $\eta^{\pm, \mathbb{Q}}$ using transformation in Eq (29).
2. We apply the MGF under \mathbb{Q} in Eq.28 along with risk-neutral drift r and volatility σ to value call and put options using Eq.30.
3. We infer model implied volatilities by inverting the Black-Scholes formula.

We use a numerical optimiser to estimate $\hat{\sigma}$, $\hat{\chi}^+$, $\hat{\chi}^-$ that minimise the MAPE.

4 Positive and negative jump risk premia

Following the literature on variance risk premia (e.g. Carr and Wu, 2009; Bollerslev et al., 2009; Almeida et al., 2024), we define jump risk premia as the difference between expected jump sizes under the statistical measure and expected jump sizes under the risk-neutral measure.

Corollary 4 (Using Proposition 5). *The positive jump risk premium denoted by γ_t^+ and negative jump risk premium denoted γ_t^- are specified by:*

$$\gamma_t^+ = \lambda_t^{+, \mathbb{Q}} \mathbb{E}^{\mathbb{Q}}(e^{J^{+, \mathbb{Q}}} - 1) - \lambda_t^+ \mathbb{E}(e^{J^+} - 1), \quad (44)$$

$$\gamma_t^- = \lambda_t^{-, \mathbb{Q}} \mathbb{E}^{\mathbb{Q}}(e^{J^{-, \mathbb{Q}}} - 1) - \lambda_t^- \mathbb{E}(e^{J^-} - 1). \quad (45)$$

According to the model specification for the price process in Equation (1), the terms $\lambda_t^+ \mathbb{E}(e^{J^+} - 1)$ and $\lambda_t^- \mathbb{E}(e^{J^-} - 1)$ are the compensators for positive and negative jumps under statistical measure \mathbb{P} , respectively. By Proposition 4, the parameters χ^+ and χ^- in Eq.(18) specify compensators $\lambda_t^{+, \mathbb{Q}} \mathbb{E}^{\mathbb{Q}}(e^{J^{+, \mathbb{Q}}} - 1)$ and $\lambda_t^{-, \mathbb{Q}} \mathbb{E}^{\mathbb{Q}}(e^{J^{-, \mathbb{Q}}} - 1)$ under the risk-neutral measure \mathbb{Q} . Hereby, the parameters χ^+ and χ^- are inferred from options data.

In this construction, the jump risk premia reflect the difference in forward-looking risk preferences inferred from option prices compared to risk preferences inferred from historical price dynamics. Positive γ_t^+ and negative γ_t^- indicate that the market expects higher future jump risk than was observed in the past, and vice versa for negative γ_t^+ and positive γ_t^- . The jump risk premia are expressed in the same units as the drift rate μ , i.e., as an annualised yield rate.

The separate treatment of positive and negative jumps is motivated by the fact that the market reacts differently to up and down moves in the price process. In particular, the left tail of the risk-neutral distribution is known to be more vulnerable to market downturns than the right tail to market upswings. This is also seen in the impulse response functions in Section 5.3 below, where the responses of positive and negative jump risk premia differ in magnitude and speed of decay in impact.

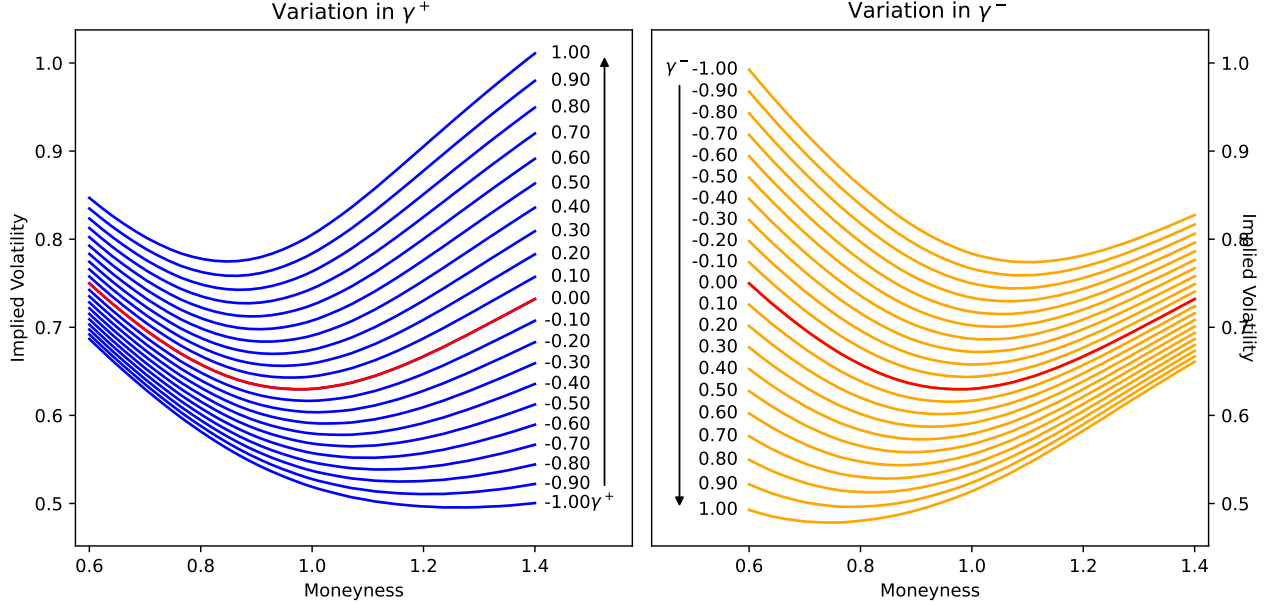
Figure 3 depicts the impact of jump risk premia on the implied volatility profiles of model-generated options with one month to maturity. Generally, the magnitude of the jump risk premia is directly correlated with the overall level of model implied volatilities. A higher positive jump risk premium and a lower negative jump risk premium contribute to a higher level of model implied volatilities. However, as the panel on the top left reveals, the influence of γ^+ is more pronounced on options with a moneyness greater than 1, predominantly in-the-money (ITM) call options. The panel on the right shows the effect of the negative jump risk premium, where one has to bear in mind that a negative γ^- indicates that more negative jump activity is expected under \mathbb{Q} than under \mathbb{P} . In turn, the more negative the γ^- is, the higher the level of the model implied volatilities. Similarly, this change in γ^- shows a stronger impact on options with a moneyness less than 1, typically where the ITM put options are concentrated.

Figure 4 illustrates the impact of γ^+ and γ^- on the risk-neutral probability density of one-month log returns. The figure reveals that γ^+ and γ^- jointly govern the asymmetry and tail thickness of the risk-neutral density, encapsulating the market's pricing of upward and downward jump risks.

Figure 3: The panels show model implied volatilities of one-month-maturity options under different positive and negative jump risk premium levels while keeping the opposite jump risk premium zero. The levels of positive jump risk premium are indicated to the right of the corresponding model implied volatilities. The model implied volatilities in red are under zero positive and negative jump risk premia.



IV_smiles_under_diff_risk_premia_level



5 Empirical Results

5.1 Data and empirical study design

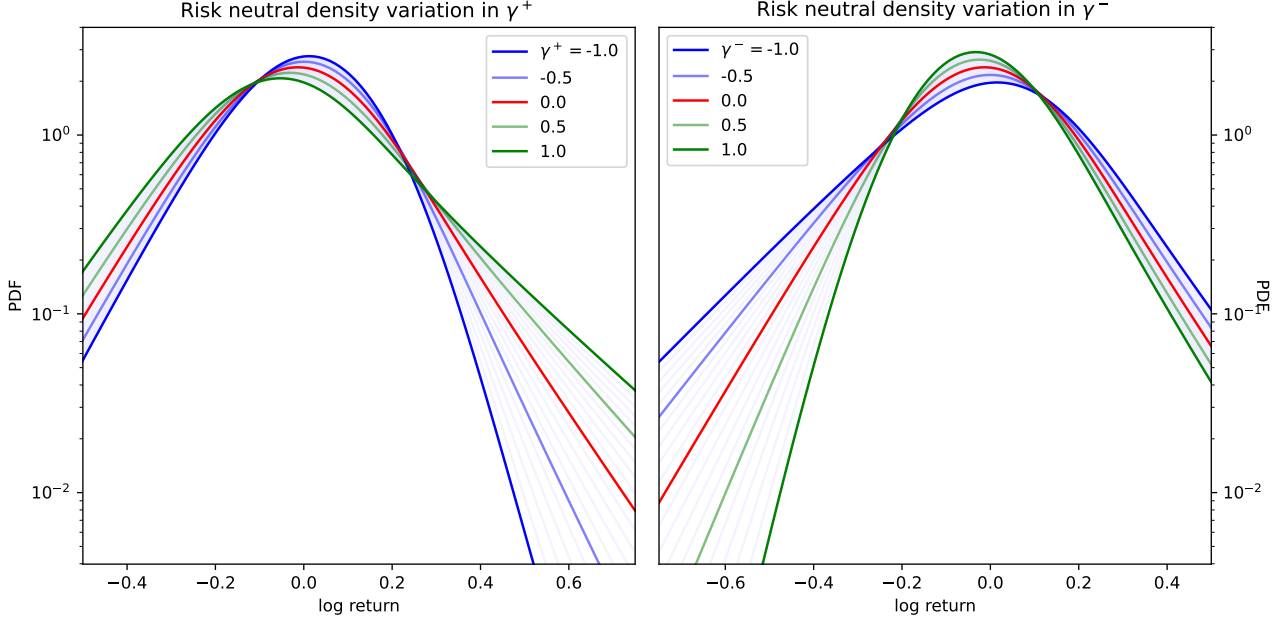
We study the options market for cryptocurrencies, with its dynamics heavily influenced by sentiment-driven jumps and severe price adjustments. In particular, we consider options on Bitcoin (BTC) which are the most liquid in cryptocurrency derivatives. We gather data on both BTC price and BTC options from Deribit, a crypto derivatives exchange established in June 2016. The exchange offers options and futures trading for both BTC and ETH and provides extensive data on BTC prices and option prices through its API. Deribit's comparatively low fees for market makers contribute to the open interest and transaction volume for options traded on the platform.

Our sampled BTC price data spans from 2015-12-31 to 2023-10-04 with the sample of BTC opening prices at 9 am UTC. As for the BTC options data, our collection starts from 2019-05-30 to 2023-10-04. The options data consists of 9 am UTC snapshots of the bid and ask IVs of call and put options. We chose to capture these snapshots at 9 am UTC to minimise the impact of market fluctuations that often occur due to the introduction of new options and futures listings at 8 am UTC.

We split the data into two datasets based on the availability of one-month option data. The first dataset is data observed from 2015-12-31 to 2019-05-29, where 2019-05-30 is the earliest date when 1-month options became available on Deribit. The second dataset is data observed from 2019-05-30 onward.

We design the empirical study as follows. Initially, we use BTC price data from the first dataset for the estimation of model parameters that are associated with jumps, as described in Section 3. In this stage, we apply POT estimation to filter out arrival times and magnitudes of positive and negative

Figure 4: The panels display the risk-neutral probability densities of one-month log return under varying levels of the positive and negative jump risk premia, γ^+ and γ^- . The left panel illustrates the effect of changes in risk-premia for calls γ^+ while holding $\gamma^- = 0$, and the right panel shows the effect of changes in risk-premia for puts γ^- while holding $\gamma^+ = 0$. Increasing γ^+ fattens the right tail of the distribution, leading to greater right-skewness in the distribution. In contrast, variation in γ^- produces the opposite pattern: a fatter or thinner left tail depending on the sign of the premium. Together, γ^+ and γ^- govern the asymmetry and tail behavior of the risk-neutral density, reflecting how markets price the risk of upward and downward jumps.



jumps in the BTC price in the testing data. In this empirical study, the training data are the daily BTC price observed from 2015-12-31 to 2019-05-29. From the time series of BTC price, we compute the jump intensities for the training data, and then generalise the computation of intensities to the testing data based on the estimated model parameters. The model parameters are held fixed to infer jumps intensities to enforce no look-ahead bias in our empirical results.

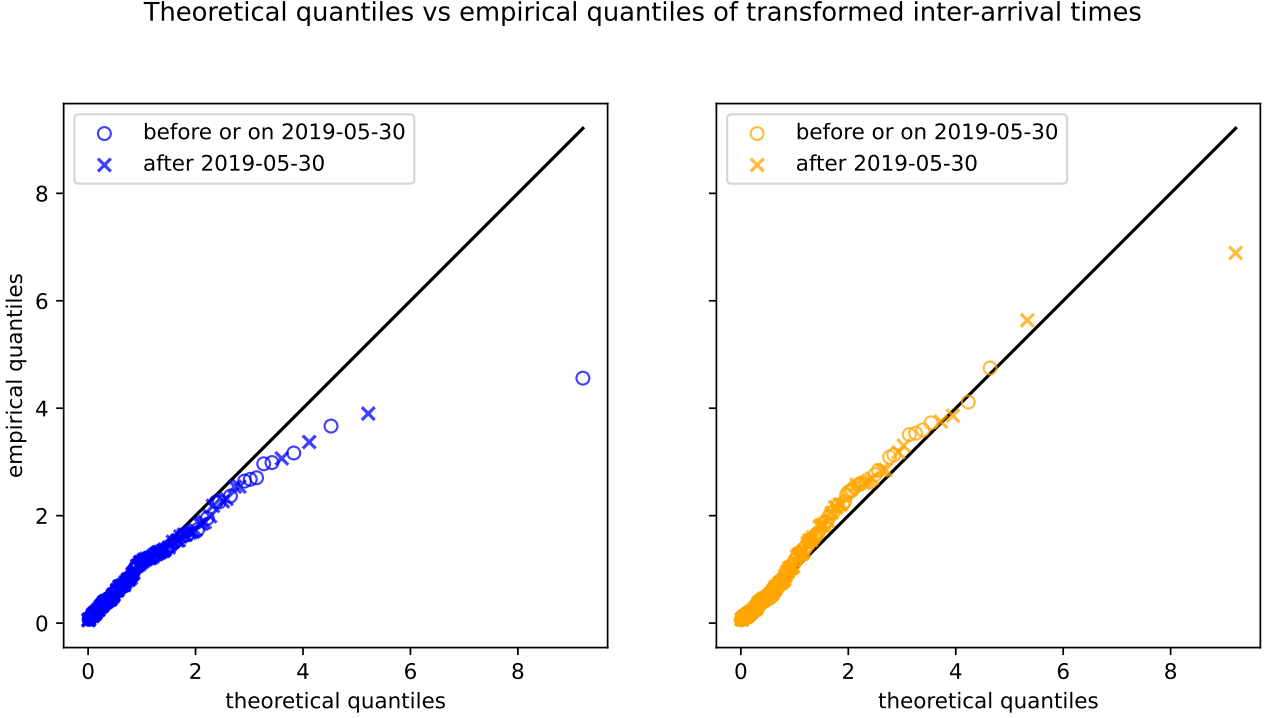
Once a set of model parameters is estimated and the time series of jump intensity rates are inferred, we calibrate model volatility σ , and risk-premium parameters χ^+ and χ^- using options data in the second dataset. Options data include mid quotes of implied volatilities of call and put options, specifically options with maturities that are closest to, but less than one month, and have a moneyness range between 0.8 and 1.2. These options are characterised by high open interest and high transaction volume.

The final step involves the analysis of the jump risk premia computed using Eqs 44 and 45. Section 5.3 presents the evolution of jump risk premia together with the main events in the BTC market using equations. Section 5.4 reports the relationship between the jump risk premia and the cost-of-carry of BTC futures. Section 5.5 shows application of the jump risk premia to profit and loss of delta-hedged option strategies.

5.2 Goodness-of-fit

We visually assess the goodness of fit of our model to the BTC price dynamics through Q-Q plots of the inter-arrival times of unit rate Poisson processes versus the transformed inter-arrival times of

Figure 5: The left panel shows the Q–Q plot of compensator-transformed interarrival times for positive jumps; the right panel shows the same for negative jumps. Crosses and circles represent jumps occurring before (in-sample) and after (out-of-sample) 2019-05-30, respectively. Under the model, the transformed interarrival times should follow a standard exponential distribution. The model is calibrated on data before 2019-05-30, and the estimated parameters are used to infer intensities for periods after 2019-05-30. The close agreement between theoretical and empirical quantiles indicates strong in- and out-of-sample performance. $\mathbb{Q}_{\text{goodness_of_fit_P}}$




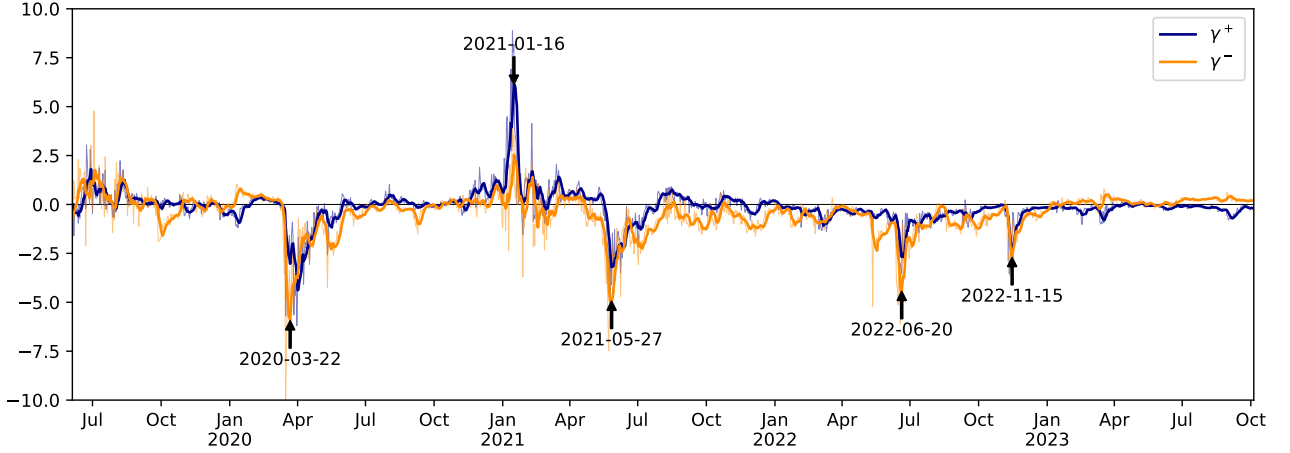
observed BTC price jumps, $\int_{T_{[k-1]}^+}^{T_{[k]}^+} \lambda_s^+ ds$ for $k = 1, 2, \dots, N^{(1)}$ and $\int_{T_{[k-1]}^-}^{T_{[k]}^-} \lambda_s^- ds$ for $k = 1, 2, \dots, N^{(2)}$, where $N^{(1)}$ and $N^{(2)}$ are the numbers of positive and negative jumps observed until time T . Appendix B provides further details of this approach.

Figure 5 presents the Q–Q plot of the model estimates. Model parameters are estimated from 2015-12-31 to 2019-05-29, as mentioned above. These estimates are applied to transform the jumps arrival times of the whole dataset from 2015-12-31 to 2023-10-04. In the Q–Q plots, data points from the first dataset are marked by circles, data points from the testing data are marked by crosses. The model estimates fit reasonably well for most of the negative jumps inter-arrival times.

5.3 Evolution of jump risk premia

The dynamics of jump risk premia, computed using Equations (44) and (45), from 2019-05-30 to 2023-10-04 are illustrated in Figure 6. Overall, the jump risk premia vary strongly over time, while showing a tendency to mean-revert to a mean close to zero. In addition, the negative jump risk premium fluctuates in the same direction as the positive jump risk premium, but with a stronger magnitude. This co-movement suggests the time-varying skewness observed in option prices. When jump premia are negative and put options have a higher demand than call options, the implied volatility skew becomes left-skewed compared to the skew inferred by the statistical measure, and vice versa.

Figure 6: The time series plots illustrate the positive and negative jump risk premia, depicted by light blue and light orange lines, respectively. The darker lines in the plots represent the seven-day moving averages of the corresponding jump risk premia. These moving averages reveal a more distinct pattern in the evolution of the jump risk premia. The jump risk premia are estimated from 2019-05-30 to 2023-10-04.  jumps-premia-evolution



Although Bollerslev and Todorov (2011)’s data and methodology differ significantly from ours, their findings are in line with our results. Using high-frequency data and a model-free approach to decompose volatility into continuous and jump components, they found that investors demand a significantly higher premium for downside jump risk than for upside jump risk. This asymmetry supports our observation that negative jump risk premia tend to have a greater magnitude than positive premia. When pricing longer-dated options, the extremely high jump risk premia are not expected to last very long due to the mean reversion of jump intensities.


A significant peak in premia was observed on 2021-01-16, building on the increase that started in early October 2020, a period marked by heightened institutional activity and a series of all-time high break-outs for BTC price. The troughs and the market situations during the time include the following (i) 2020-03-22: COVID-19 outbreak starting in March 2020; (ii) 2021-05-27: environmental concerns over Bitcoin mining; (iii) 2022-06-20: 40% decline in BTC caused by hawkish interest rate policy of the U.S. Federal Reserve amid high inflation rates; (iv) 2022-11-15: the FTX crisis and subsequent crash of cryptocurrencies.

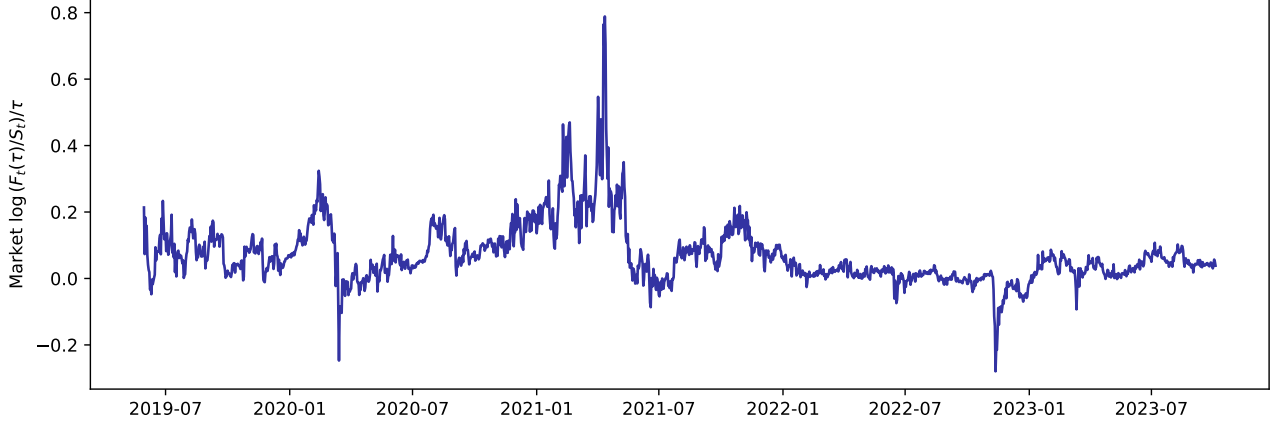
5.4 Jump risk premia as factors in explaining BTC futures cost-of-carry

BTC futures are a popular choice for investors seeking leveraged exposures to BTC dynamics. Here, we investigate whether jump risk premia can serve as explanatory factors for BTC futures pricing, specifically through their relationship with the cost-of-carry. The variable of interest is the BTC basis or cost-of-carry specified as follows:

$$c_t = \ln(F_t(\tau)/S_t)/\tau, \quad (46)$$

where $F_t(\tau)$ is the price at time t of a futures contract with time-to-maturity τ , and S_t is the spot price of BTC. Figure 7 shows the basis c_t , computed using the prevailing futures contract with a time-to-maturity just below one month, based on data from Deribit. The cost-of-carry is annualized and exhibits high volatility, ranging from -20% to 80% . High values of the cost-of-carry arise when

Figure 7: The time series plot shows the BTC futures basis c_t observed in the market. The data is calculated using the price of the future with maturity τ less than, but closest to one month.  `cost_of_carry_factors`



investors seek leverage. A trader can monetise a high value of the cost-of-carry by selling short the futures and going long BTC through a spot market.


In addition to the jump risk premia, the number of positive and negative jumps in the previous week, abbreviated respectively as $N_t^{(1)} - N_{t-7}^{(1)}$ and $N_t^{(2)} - N_{t-7}^{(2)}$, and the average funding rate of the week from Deribit BTC perpetual futures, abbreviated by $r_{t,t-7}^{\text{funding}}$, are included as independent variables in order to support the analysis. The regression of interest is given by

$$\Delta c_t = \beta_1 \Delta \gamma_t^+ + \beta_2 \Delta \gamma_t^- + \beta_3 \Delta (N_t^{(1)} - N_{t-7}^{(1)}) + \beta_4 \Delta (N_t^{(2)} - N_{t-7}^{(2)}) + \beta_5 \Delta r_{t,t-7}^{\text{funding}} + \Delta \varepsilon_t. \quad (47)$$

All variables are differenced to remove potential fixed effects and to mitigate the impact of any fixed effects. To mitigate the noise arising from daily fluctuations and the introduction of new contracts, the data selected for the regression analysis consist of market snapshots captured every Friday at 9 am UTC. This timing is chosen because traders typically adjust their positions on Fridays to manage risk for the upcoming weekend. Additionally, newly issued options and futures contracts typically begin trading at 8 am, allowing the 9 am UTC snapshot to incorporate the initial market responses of these new instruments. The variables are observed every Friday at 9 am UTC from 2019-06-14 to 2023-02-17. Table 1 reports the estimated coefficients, Newey-West robustified t -statistics, and adjusted R^2 coefficients of regressions of c_t on combinations of explanatory variables.

The results of the regression analysis reveal a significant explanatory power of jump risk premia in determining the volatile cost-of-carry for BTC futures. In Regression I, the coefficients for jump risk premia are both positive and statistically significant. This model yields an adjusted R^2 of 14.7%, together with the significant coefficients, supporting the hypothesis that futures prices are directly proportional to the jump risk premia.

The jump risk premia explanatory power is robust against the presence of the number of jumps and the funding rate variables. This conclusion is drawn by comparing the adjusted R^2 values and coefficients from Regression I with those from Regressions IV and V. Combining the results of Regression I with that of Regression IV side by side, which includes the number of jumps from the previous week, the coefficient of the positive jump risk premium remains significant. Inclusion of these variables improves the adjusted R^2 to 19.6%. However, the coefficient of the negative jump risk premium decreases in magnitude and becomes statistically insignificant in this context. A similar pattern is noted

Table 1: The table reports estimated coefficients, Newey-West robustified t -statistics, and adjusted R^2 coefficients of regressions of the futures basis $c_t = \ln(F_t(\tau)/S_t)/\tau$. The futures contract with time-to-maturity τ smaller than, but closest to one month is chosen for each observation time to represent the futures market. The regressors are defined as follows: γ_t^+ and γ_t^- are the jump risk premium. $N_t^{(1)} - N_{t-7}^{(1)}$ and $N_t^{(2)} - N_{t-7}^{(2)}$ are number of positive and negative jumps in the week prior to t , respectively. $r_{t,t-7}^{\text{funding}}$ is the average funding rate of the week from Deribit BTC perpetual futures. The Deribit BTC funding rate is chosen to represent the convenience yield in the crypto market. All variables enter as first differences. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. The variables are observed every Friday at 9 am UTC from 2019-06-14 to 2023-02-17.  cost_of_carry_factors

	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>
γ_t^+	0.0157*** (2.6139)			0.0160*** (3.5339)	0.0071 (1.2400)	0.0075* (1.7027)
γ_t^-	0.0132** (2.2257)			0.0089 (1.4316)	0.0131** (2.5760)	0.0097* (1.7995)
$N_t^{(1)} - N_{t-7}^{(1)}$		0.0102** (2.1840)		0.0112*** (3.0709)		0.0073* (1.8356)
$N_t^{(2)} - N_{t-7}^{(2)}$		-0.0179*** (-2.8429)		-0.0126** (-2.4792)		-0.0114** (-2.2707)
$r_{t,t-7}^{\text{funding}}$			0.0024*** (4.6951)		0.0020*** (3.8976)	0.0018*** (3.7371)
Adj. R^2	0.1467	0.0866	0.2231	0.1956	0.2794	0.3107

in Regression V. When the funding rate is included in the Regression V, the coefficient of negative jump risk premium slightly diminishes while remaining significant, but the coefficient of positive jump risk premium halves and becomes statistically insignificant. The inclusion of the average funding rate improves the adjusted R^2 to 28%. These improvements value of R^2 confirm that the number of jumps and the average funding rate provide additional information to explain the level of BTC futures cost-of-carry in addition to the jump risk premia. However, the shrinkage of the magnitudes and loss in statistical significance observed from the coefficients are probably due to the correlation between the variables. Empirically, the correlation between the positive jump risk premium and the funding rate is around 36.7%; the correlation between the negative jump risk premium and the number of negative jumps is around -24%.

The interpretation of the correlations is as follows. The positive correlation between the positive jump risk premium and the funding rate may be due to the increasing speculative demand for both call options and perpetual futures. In contrast, the negative correlation between the negative jump risk premium and the number of negative jumps can result in higher put option prices during bearish periods with high volatility. This interpretation is consistent with the findings of Schmeling et al. (2023). They attribute high cost-of-carry of BTC futures to the interaction between the trend-chasing behaviour of market traders and the scarcity of arbitrage capital that conducts cash-and-carry trade.

Regression VI offers a comprehensive overview of how jump risk premia affect BTC futures prices, taking into account all relevant variables. This regression reaches an adjusted R^2 value of 31%, with each coefficient demonstrating statistical significance. According to the regression result, a one-unit increase in the positive jump risk premium is associated with a 0.75 basis point increase in the cost of carry of BTC futures c_t . Similarly, a one-unit increase in the negative jump risk premium corresponds to an increase of 0.97 basis points in c_t . The positive signs of both the positive and negative jump risk premia coefficients in Regression VI indicate that a positive change corresponds to (forward-looking)

option and futures market being more optimistic than the price history suggests. The different sizes of the positive and negative jump risk premia coefficients can be attributed to risk aversion, as traders and investors are more sensitive to losses than to gains.

5.5 Jump risk premia realisation

In the spirit of Carr and Madan (1998) who quantifies the return variance risk premium on an asset using the market prices of options, we answer the question of how the risk premia can be realised. To this end, we investigate whether the jump risk premia defined in Equations (44) and (45) can explain the one-week-ahead returns of delta-hedged short-option strategies, following a similar empirical approach to Bakshi and Kapadia (2003). The option strategies are constructed using the same data and approach as Lucic and Sepp (2024), where short positions in calls, puts, straddles and strangles are delta-hedged using Black-Scholes (BS) deltas and rebalanced hourly until expiration.

All strategies are written on options expiring on the subsequent Friday at 8am UTC. As both positive and negative jump risk premia are observed, we define simple trading signals as follows: A positive risk premium for positive jumps, which means that the market prices call options higher than implied by the historical measure, is considered an opportunity to sell call options and call spreads. The trading decision is similar for put options and put spreads when observing a negative risk premium on negative jumps. For portfolios involving straddle and strangles, which are combinations of call and put options, the trading strategy is to sell them if at least one of the risk premia is “active” in the sense described previously. The trading signals for all strategies are determined on each Friday at 9am UTC, and if active, the position is created (and delta-hedged with hourly rebalancing) until expiry the following Friday.²

Using Bakshi and Kapadia (2003)’s empirical methodology, we investigate the relationship between jump risk premia as defined above and one-week-ahead returns of the option strategy portfolios. This is achieved by conducting a regression against the jump risk premia of the one-week-ahead log returns $r_t = \ln(P_t/P_{t-1})$ of each strategy, with P_t the option strategy value at time t . The regression of interest is

$$r_{t+1} = \beta_0 + \beta_1\gamma_t^+ + \beta_2\gamma_t^- + \beta_3\Delta\gamma_t^+ + \beta_4\Delta\gamma_t^- + \beta_5(N_t^{(1)} - N_{t-1}^{(1)}) + \beta_6(N_t^{(2)} - N_{t-1}^{(2)}), \quad (48)$$

where r_{t+1} is the one-week-ahead options strategy return at time t (expressed in basis points), Δ is the first difference operator and $(N_t^{(1)} - N_{t-1}^{(1)})$ and $(N_t^{(2)} - N_{t-1}^{(2)})$ count the number of jumps that occurred within the previous week. The option strategies of interest are delta-hedged short-option strategies of different moneyness, as follows:

- (i) ATM: calls and puts with at-the-money strikes;
- (ii) 10D: options with delta of +0.1 for calls and -0.1 for puts;
- (iii) 25D: options with delta of +0.25 for calls and -0.25 for puts;
- (iv) 25D spread: one unit of call with delta of +0.5 (-0.5 for put) and short two units of call with +0.25 delta (-0.25 delta for put);
- (v) ATM straddle: one unit of call and put each with ATM strikes;

²Long positions in calls and puts are not entered if the jump risk premia are “reversed” due to the loss in time value of the options (negative theta).

Table 2: Regression results of one-week ahead portfolio returns (in bps) regressed on jump risk premia and jump realisations, cf. Eq.(48), heteroskedasticity-robust t -statistics in parentheses. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Sample period: 2019-05-31 to 2023-09-29 (277 observations).

	straddle	10d strangle	25d strangle	call spread	put spread
const	-1.985 (-0.113)	-0.898 (-0.106)	3.827 (0.31)	7.194 (0.854)	-2.822 (-0.388)
γ_t^+	22.652 (0.968)	23.889** (2.15)	26.59 (1.535)	24.41** (2.254)	0.309 (0.027)
γ_t^-	-53.874** (-2.13)	-33.475** (-2.541)	-42.253** (-2.36)	-15.31 (-1.25)	-7.908 (-0.662)
$\Delta\gamma_t^+$	1.646 (0.093)	-27.876* (-1.787)	-23.075 (-1.458)	-28.187** (-2.059)	-12.458* (-1.784)
$\Delta\gamma_t^-$	5.872 (0.277)	35.705* (1.911)	19.97 (1.133)	18.348 (1.066)	3.853 (0.253)
$N_t^{(1)} - N_{t-1}^{(1)}$	8.728 (0.382)	3.249 (0.325)	-0.518 (-0.031)	-3.749 (-0.47)	3.211 (0.284)
$N_t^{(2)} - N_{t-1}^{(2)}$	41.155* (1.733)	7.134 (0.632)	19.118 (1.055)	1.532 (0.299)	-1.201 (-0.122)
Adj. R ²	0.049	0.045	0.03	0.044	-0.002

	atm call	10d call	25d call	atm put	10d put	25d put
const	3.038 (0.46)	0.75 (0.148)	4.182 (0.769)	2.222 (0.342)	0.784 (0.199)	1.459 (0.257)
γ_t^+	19.367** (2.234)	20.446*** (3.049)	22.365*** (3.024)	2.931 (0.324)	2.524 (0.461)	3.992 (0.491)
γ_t^-	-11.39 (-1.277)	-15.932* (-1.744)	-14.445* (-1.78)	-22.474** (-2.074)	-11.587** (-2.23)	-16.64* (-1.793)
$\Delta\gamma_t^+$	-7.976 (-1.238)	-22.573* (-1.79)	-19.624** (-2.267)	-1.039 (-0.13)	-7.417** (-2.035)	-8.926* (-1.735)
$\Delta\gamma_t^-$	10.594 (1.513)	22.222 (1.515)	15.607 (1.53)	-0.67 (-0.071)	7.178** (2.244)	3.187 (0.512)
$N_t^{(1)} - N_{t-1}^{(1)}$	-11.773 (-1.605)	-1.281 (-0.532)	-7.385* (-1.652)	11.508 (1.272)	7.74** (2.047)	6.15 (0.958)
$N_t^{(2)} - N_{t-1}^{(2)}$	14.153* (1.714)	3.682 (1.385)	7.496 (1.53)	5.904 (0.733)	-1.132 (-0.221)	2.745 (0.406)
Adj. R ²	0.041	0.097	0.071	0.048	0.029	0.039

(vi) 25D strangle: one unit of call with delta +0.25 and one unit of put with delta of -0.25;

(vii) 10D strangle: one unit of call with delta +0.1 and one unit of put with delta of -0.1.

Table 2 shows the results of the regression on one-week-ahead returns (in basis points), with robust (HAC) t -statistics in parentheses. The number of positive and negative jumps counted during the week are included as control variables. Recalling that a negative jump risk premium on downward jumps indicates high option prices relative to the statistical history (cf. Figure 3, the mixed strategies (straddle, strangles) earn jump risk premia from downward jumps. The 10-delta strangle also earns

risk premia from upward jumps as well as changes in jump risk premia. Here, a negative first difference in the upward jump risk premium indicates that the jump risk premium decreased over the course of the previous week, i.e., pricing of jump risk of out-of-the-money call options aligns better with historical jump risk. The negative coefficient persists when replacing $\Delta\gamma_t^+$ with the residuals of $\Delta\gamma_t^+$ regressed on γ_t^+ , which rules out multicollinearity effects (likewise for $\Delta\gamma_t^-$ and γ_t^-). Hence, $\Delta\gamma_t^+$ and $\Delta\gamma_t^-$, when statistically significant, carry independent predictive performance driving the P&L. A plausible interpretation of the negative coefficient associated with $\Delta\gamma_t^+$ is that this points to a calming market (expecting less right-tail events), which in turn makes the delta-hedge more robust and allows to earn the time-decay θ inherent in the option positions. The argument is similar for $\Delta\gamma_t^-$ and the positive coefficient.

Overall, on average the short call option strategies earn the positive jump risk premium of upward jumps and the short put option strategies earn the negative jump risk premium of downward jumps. One has to bear in mind, though, that many other factors influence the weekly returns as the adjusted R^2 measures are fairly low for all strategies.

The relative pricing interpretation also aligns with the analysis done by Bakshi and Kapadia (2003). They demonstrate in Eq.(24) of their study that, theoretically, a more pronounced left tail in the underlying price distribution under the risk-neutral measure, as opposed to the “real” measure, enhances the performance of delta-hedged portfolios³.

6 Conclusion

The cryptocurrency market provides a challenging venue for exploring jump risk premia due to its high volatility and sentiment-driven price dynamics. The dynamics of implied volatilities of cryptocurrency options exhibit varying skewness preferences. During bullish periods, the demand for call options increases and the implied skew becomes positive along with high carry costs for perpetual futures. During bearish periods, the demand for puts rises and the implied skew turns negative along with low or negative carry costs for perpetual futures. A similar regime-conditional behaviour of implied skew is present in options on the so-called “meme” stocks, which are characterised by high volumes in options driven by retail investors.

For modeling of regime-conditional implied volatility skew under the risk-neutral measure and skewness of returns under the statistical measure, we have introduced a pricing model using Hawkes processes that incorporates the clustering of positive and negative jumps. We have further included two risk-premium parameters to model specification for modeling of regime-conditional risk-preferences. Here, one parameter is for the external specification of risk-premium for positive jumps, and one parameter is for risk-premium of negative jumps. Both parameters are inferred from the option prices observed on a given date.

For empirical investigation using our model, we have used options on Bitcoin (BTC). We have developed a sequential estimation procedure by first estimating model parameters under the statistical \mathbb{P} measure and then by inferring two risk-premium parameters from options data under the risk-neutral \mathbb{Q} measure. We have defined risk premia for positive and negative jumps as the difference between expected jumps under the risk-neutral and statistical probability measures, respectively. The implications of the jump risk premia estimated using our model are threefold.

³Bakshi and Kapadia (2003) approach the construction of delta-hedged portfolios from a direction opposite to ours. While they engage in strategies that involve buying options and selling the underlying assets, we adopt the reverse strategy, selling options and buying the underlying.

First, the estimated jump-risk premia demonstrate preferences for skewness risk observed in implied volatilities. Specifically for BTC markets, the risk premium of positive jumps exceeded that of negative jumps on several occasions during a few strong bullish periods in BTC. In contrast, the risk premium of negative jumps dominated during bearish periods in BTC. A strong premium for positive jumps indicates that market participants significantly overpay for expected upside during bull periods by buying call options. Vice versa, a significant premium of negative jumps is associated with increased demand for put options during bear periods.

Second, jump risk premia have some explanatory power (with R^2 of 15%) to predict the dynamics of the cost-of-carry inferred from BTC futures. A regression model with the recent number of jumps and BTC perpetual funding rates as control variables yields an R^2 of 31%, with all coefficients being significant. The findings point to interdependencies between the BTC spot, options, and futures markets.

Third, we show that jump risk premia have explanatory power for one-week-ahead profit-and-loss (P&L) of delta-hedged option strategies and portfolios. Simple decision strategies to sell short call options if the risk premium for upward jumps is positive, to sell put option if the risk premium for downward jumps is negative, or to sell straddles if either of the jump risk premia is active, can – on average – earn the respective risk premium. In all cases, the options are hedged with hourly rebalancing.

Finally, we note that the regime-conditional behaviour of the implied option skew is also characteristic of the so-called “meme” stocks, which exhibit positive and negative implied volatility skews dependent on market conditions. In addition, the G-7 currencies tend to exhibit a variation in the sign of implied skew dependent on macro conditions. As a result, our framework can also be applied for analysis of the implied volatility dynamics of “meme” stocks and the G-7 currencies, which we leave for future research. In addition, we assume that the risk-premium parameters are constant for the specification of model dynamics. We leave an extension for time-varying dynamics of model risk-premium parameters as another potential research topic.

References

- ALEXANDER, C. AND A. IMERAJ (2021): “The Bitcoin VIX and its variance risk premium,” *The Journal of Alternative Investments*, 23, 84–109. 4
- (2023): “Delta hedging bitcoin options with a smile,” *Quantitative Finance*, 23, 799–817. 4
- ALMEIDA, C., M. GRITH, R. MIFTACHOV, AND Z. WANG (2024): “Risk Premia in the Bitcoin Market,” arXiv preprint arXiv:2410.15195. 4, 14
- BACRY, E., I. MASTROMATTEO, AND J.-F. MUZY (2015): “Hawkes Processes in Finance,” *Market Microstructure and Liquidity*, 01. 4
- BAKSHI, G. AND N. KAPADIA (2003): “Delta-hedged gains and the negative market volatility risk premium,” *The Review of Financial Studies*, 16, 527–566. 4, 21, 23
- BAKSHI, G., N. KAPADIA, AND D. MADAN (2003): “Stock return characteristics, skew laws, and the differential pricing of individual equity options,” *The Review of Financial Studies*, 16, 101–143. 4
- BALI, T. G., N. CAKICI, AND R. F. WHITELAW (2011): “Maxing out: Stocks as lotteries and the cross-section of expected returns,” *Journal of financial economics*, 99, 427–446. 2, 4

- BATES, D. S. (1996): “Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options,” *The Review of Financial Studies*, 9, 69–107. 3, 4
- (2000): “Post-’87 crash fears in the S&P 500 futures option market,” *Journal of Econometrics*, 94, 181–238. 4
- BOLLERSLEV, T., G. TAUCHEN, AND H. ZHOU (2009): “Expected stock returns and variance risk premia,” *The Review of Financial Studies*, 22, 4463–4492. 14
- BOLLERSLEV, T. AND V. TODOROV (2011): “Tails, fears, and risk premia,” *The Journal of Finance*, 66, 2165–2211. 18
- BROADIE, M., M. CHERNOV, AND M. JOHANNES (2007): “Model specification and risk premia: Evidence from futures options,” *The Journal of Finance*, 62, 1453–1490. 4
- BROWN, T. C. AND M. G. NAIR (1988): “A simple proof of the multivariate random time change theorem for point processes,” *Journal of Applied Probability*, 25, 210–214. 38
- CARR, P. AND D. MADAN (1998): “Towards a theory of volatility trading,” *Volatility: New estimation techniques for pricing derivatives*, 29, 417–427. 21
- CARR, P. AND L. WU (2009): “Variance risk premiums,” *The Review of Financial Studies*, 22, 1311–1341. 3, 14
- CHRISTOFFERSEN, P., K. JACOBS, AND C. ORNTHANALAI (2012): “Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options,” *Journal of Financial Economics*, 106, 447–472. 4
- DALEY, D. J. AND D. VERE-JONES (2003): *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*, Springer. 12, 38
- DOBRYNSKAYA, V. (2024): “Is downside risk priced in cryptocurrency market?” *International Review of Financial Analysis*, 91, 102947. 4
- EMBRECHTS, P., T. LINIGER, AND L. LIN (2011): “Multivariate Hawkes processes: an application to financial data,” *Journal of Applied Probability*, 48, 367–378. 4, 11
- ERAKER, B. (2004): “Do stock prices and volatility jump? Reconciling evidence from spot and option prices,” *The Journal of finance*, 59, 1367–1403. 4
- ERRAIS, E., K. GIESECKE, AND L. R. GOLDBERG (2010): “Affine point processes and portfolio credit risk,” *SIAM Journal on Financial Mathematics*, 1, 642–665. 4, 28, 29
- GERBER, H. AND E. SHIU (1995): “Option pricing by Esscher transforms,” *Insurance Mathematics and Economics*, 3, 287. 32, 33
- GRANELLI, A. AND A. E. D. VERAART (2016): “Modeling the Variance Risk Premium of Equity Indices: The Role of Dependence and Contagion,” *SIAM Journal on Financial Mathematics*, 7, 382–417. 3
- GUPTA, A. AND S. PASCALE (2025): “Volatility Outlook 2025,” *Barclays Equity Research White Paper*. 3

- HAINAUT, D. (2016): “A bivariate Hawkes process for interest rate modeling,” *Economic Modelling*, 57, 180–196. 3, 4
- (2017): “Clustered Lévy processes and their financial applications,” *Journal of Computational and Applied Mathematics*, 319, 117–140. 3, 4
- (2022): *Continuous Time Processes for Finance*, Springer. 4, 11
- HAINAUT, D. AND F. MORAUX (2018): “Hedging of options in the presence of jump clustering,” *Journal of Computational Finance*, 22, 1–35. 4, 6
- HAWKES, A. G. (1971): “Point spectra of some mutually exciting point processes,” *Journal of the Royal Statistical Society: Series B (Methodological)*, 33, 438–443. 4
- (2018): “Hawkes processes and their applications to finance: a review,” *Quantitative Finance*, 18, 193–198. 4
- HAWKES, A. G. AND D. OAKES (1974): “A cluster process representation of a self-exciting process,” *Journal of applied probability*, 11, 493–503. 4
- JEANBLANC, M., M. YOR, AND M. CHESNEY (2009): *Mathematical Methods for Financial Markets*, Springer Science & Business Media. 29
- KLEBANER, F. C. (2005): *Introduction to Stochastic Calculus with Applications*, Imperial College Press, 2nd ed. 29, 30
- KUMAR, A. (2009): “Who gambles in the stock market?” *The journal of finance*, 64, 1889–1933. 2
- LAUB, P. J., Y. LEE, AND T. TAIMRE (2021): *The Elements of Hawkes Processes*, Springer. 12, 38
- LEWIS, A. L. (2001): “A simple option formula for general jump-diffusion and other exponential Lévy processes,” . 10
- LIPTON, A. (2002): “The vol smile problem,” *Risk*, 15, 61–66. 10
- LIU, F., N. PACKHAM, M.-J. LU, AND W. K. HÄRDLE (2023): “Hedging cryptos with Bitcoin futures,” *Quantitative Finance*, 23, 819–841. 4
- LUCIC, V. AND A. SEPP (2024): “Valuation and hedging of cryptocurrency inverse options,” *Quantitative Finance*, 24, 851–869. 4, 21
- MATIC, J. L., N. PACKHAM, AND W. K. HÄRDLE (2023): “Hedging cryptocurrency options,” *Review of Derivatives Research*, 26, 91–133. 4
- NGUYEN, D. B. B., M. PROKOPCZUK, AND C. W. SIMEN (2019): “The risk premium of gold,” *Journal of International Money and Finance*, 94, 140–159. 4
- PAN, J. (2002): “The jump-risk premia implicit in options: Evidence from an integrated time-series study,” *Journal of Financial Economics*, 63, 3–50. 4
- SANTA-CLARA, P. AND S. YAN (2010): “Crashes, volatility, and the equity premium: Lessons from S&P 500 options,” *The Review of Economics and Statistics*, 92, 435–451. 4

- SCHMELING, M., A. SCHRIMPF, AND K. TODOROV (2023): “Crypto carry,” Tech. rep., Bank for International Settlements. 20
- SEPP, A. AND P. RAKHMONOV (2023): “Log-normal Stochastic Volatility Model with Quadratic Drift,” *International Journal of Theoretical and Applied Finance*, 26, 1–63. 10
- THALER, R. H. AND W. T. ZIEMBA (1988): “Anomalies: Parimutuel betting markets: Racetracks and lotteries,” *Journal of Economic perspectives*, 2, 161–174. 2
- WATANABE, S. (1964): “On discontinuous additive functionals and Lévy measures of a Markov process,” in *Japanese journal of mathematics: transactions and abstracts*, The Mathematical Society of Japan, vol. 34, 53–70. 38

A Proofs

A.1 Conditions for finite expected intensities

The expected values of intensities satisfy the following system of ODEs. For $s < t$,

$$\begin{aligned}\frac{d\mathbb{E}(\lambda_t^+|\mathcal{F}_s)}{dt} &= \kappa^+(\theta^+ - \mathbb{E}(\lambda_t^+|\mathcal{F}_s)) + \beta_{11}\mathbb{E}(\lambda_t^+|\mathcal{F}_s)\mathbb{E}J^+ + \beta_{12}\mathbb{E}(\lambda_t^-|\mathcal{F}_s)\mathbb{E}J^-, \quad \mathbb{E}(\lambda_s^+|\mathcal{F}_s) = \lambda_s^+ \\ \frac{d\mathbb{E}(\lambda_t^-|\mathcal{F}_s)}{dt} &= \kappa^-(\theta^- - \mathbb{E}(\lambda_t^-|\mathcal{F}_s)) + \beta_{21}\mathbb{E}(\lambda_t^+|\mathcal{F}_s)\mathbb{E}J^+ + \beta_{22}\mathbb{E}(\lambda_t^-|\mathcal{F}_s)\mathbb{E}J^-, \quad \mathbb{E}(\lambda_s^-|\mathcal{F}_s) = \lambda_s^-.\end{aligned}$$

See Errais et al. (2010) for proofs. Write the above system of ODEs in matrix form: For $s < t$,

$$\begin{pmatrix} \frac{d\mathbb{E}(\lambda_t^+|\mathcal{F}_s)}{dt} \\ \frac{d\mathbb{E}(\lambda_t^-|\mathcal{F}_s)}{dt} \end{pmatrix} = \Phi \begin{pmatrix} \mathbb{E}(\lambda_t^+|\mathcal{F}_s) \\ \mathbb{E}(\lambda_t^-|\mathcal{F}_s) \end{pmatrix} + C, \quad \begin{pmatrix} \mathbb{E}(\lambda_s^+|\mathcal{F}_s) \\ \mathbb{E}(\lambda_s^-|\mathcal{F}_s) \end{pmatrix} = \begin{pmatrix} \lambda_s^+ \\ \lambda_s^- \end{pmatrix} \quad (49)$$

where

$$\Phi \stackrel{\text{def}}{=} \begin{pmatrix} -\kappa^+ + \beta_{11}\mathbb{E}J^+ & \beta_{12}\mathbb{E}J^- \\ \beta_{21}\mathbb{E}J^- & -\kappa^- + \beta_{21}\mathbb{E}J^+ \end{pmatrix} \text{ and } C \stackrel{\text{def}}{=} \begin{pmatrix} \kappa^+\theta^+ \\ \kappa^-\theta^- \end{pmatrix}.$$

If Φ is invertible, the solution to Eq.(49) is

$$\begin{pmatrix} \mathbb{E}(\lambda_s^+|\mathcal{F}_s) \\ \mathbb{E}(\lambda_s^-|\mathcal{F}_s) \end{pmatrix} = \expm(\Phi t) \begin{pmatrix} \lambda_s^+ \\ \lambda_s^- \end{pmatrix} + \Phi^{-1} (\expm(\Phi t) - I) C,$$

where $\expm()$ is the matrix exponential, Φ^{-1} is the matrix inverse of Φ , and I is a 2×2 identity matrix. The matrix exponential can be computed via eigenvectors and eigenvalues. Let $S\Lambda S^{-1}$ be the eigendecomposition of Φ , where S is the full matrix of eigenvectors and $\Lambda = \text{diag}(\lambda_1, \lambda_2)$ is the diagonal matrix of eigenvalues of Φ , λ_1 and λ_2 ; By the definition of matrix exponential as a sum of powers,

$$\expm(\Phi t) = \sum_{n=0}^{\infty} \frac{(\Phi t)^n}{n!} = \sum_{n=0}^{\infty} \frac{(S\Lambda S^{-1}t)^n}{n!} = S \expm(\Lambda t) S^{-1},$$

where $\expm(\Lambda t)$ is simply $\text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t})$. It is clear that if both λ_1 and λ_2 are negative and real (such that $\lim_{t \rightarrow \infty} \expm(\Phi t) = 0$), the asymptotic expected values of intensities is

$$\lim_{t \rightarrow \infty} \begin{pmatrix} \mathbb{E}(\lambda_t^+|\mathcal{F}_s) \\ \mathbb{E}(\lambda_t^-|\mathcal{F}_s) \end{pmatrix} = -\Phi^{-1}C.$$

By the Gershgorin circle theorem, the eigenvalues of Φ lie within

$$\{z : |z - (-\kappa^+ + \beta_{11}\mathbb{E}J^+)| \leq |\beta_{12}\mathbb{E}J^-|\} \cup \{z : |z - (-\kappa^- + \beta_{22}\mathbb{E}J^-)| \leq |\beta_{21}\mathbb{E}J^+|\}.$$

Since $\beta_{12}\mathbb{E}J^-$ and $\beta_{21}\mathbb{E}J^+$ are always positive, the eigenvalues of Φ lie within

$$\{z : |z - (-\kappa^+ + \beta_{11}\mathbb{E}J^+)| \leq \beta_{12}\mathbb{E}J^-\} \cup \{z : |z - (-\kappa^- + \beta_{22}\mathbb{E}J^-)| \leq \beta_{21}\mathbb{E}J^+\}.$$

Therefore, the sufficient conditions for the eigenvalues of Φ being negative are

$$\kappa^+ \geq \beta_{11}\mathbb{E}J^+ + \beta_{12}\mathbb{E}J^- \text{ and } \kappa^- \geq \beta_{21}\mathbb{E}J^+ + \beta_{22}\mathbb{E}J^-. \quad (50)$$

This means that if the mean reverting rates κ^+ and κ^- are larger than their corresponding sum of excitements from jumps which are average in size, the asymptotic expected values of intensities are finite.

A.2 Proof of Proposition 1

Let $G(t, X_t, \lambda_t^+, \lambda_t^-) = \mathbb{E}(\exp(\omega X_T + \omega^+ \lambda_T^+ + \omega^- \lambda_T^-) | X_t, \lambda_t^+, \lambda_t^-)$. The process $(G(t, X_t, \lambda_t^+, \lambda_t^-))_{t \geq 0}$ is a (Doob-Lévy) martingale (see e.g. Theorem 2.31 of Klebaner (2005)). By the predictable Itô formula (e.g. Sections 10.2 and 8.4.3 of (Jeanblanc et al., 2009)) and the martingale property of G ,

$$\begin{aligned} 0 = & G_t \\ & + \left(\mu - \frac{\sigma^2}{2} - \lambda_t^+ \mathbb{E}(e^{J^+} - 1) - \lambda_t^- \mathbb{E}(e^{J^-} - 1) \right) G_x + \frac{\sigma^2}{2} G_{xx} \\ & + \kappa^+(\theta^+ - \lambda_t^+) G_{\lambda^+} + \kappa^-(\theta^- - \lambda_t^-) G_{\lambda^-} \\ & + \lambda_t^+ \left[\int_{\mathbb{R}} G(t, X_t + j^+, \lambda_t^+ + \beta_{11}j^+, \lambda_t^- + \beta_{21}j^+) \varpi^+(j^+) dj^+ - G(t, X_t, \lambda_t^+, \lambda_t^-) \right] \\ & + \lambda_t^- \left[\int_{\mathbb{R}} G(t, X_t + j^-, \lambda_t^+ + \beta_{12}j^-, \lambda_t^- + \beta_{22}j^-) \varpi^-(j^-) dj^- - G(t, X_t, \lambda_t^+, \lambda_t^-) \right], \end{aligned} \quad (51)$$

where $G_t, G_x, G_{xx}, G_{\lambda^+}, G_{\lambda^-}$ denote the partial derivatives of G . Assume that G has an exponential affine form (see e.g. Errais et al. (2010)):

$$G(t, x, \lambda^+, \lambda^-) = \exp(A(t, T) + B(t, T)x + C(t, T)\lambda^+ + D(t, T)\lambda^-),$$

where the functions A, B, C, D are time dependent functions with terminal conditions: $A(T, T) = 0$, $B(T, T) = \omega$, $C(T, T) = \omega^+$, and $D(T, T) = \omega^-$. Inserting the partial derivatives into Equation (51),

$$\begin{aligned} 0 = & (A_t + B_t X_t + C_t \lambda_t^+ + D_t \lambda_t^-) G \\ & + \left(\mu - \frac{\sigma^2}{2} - \lambda_t^+ \mathbb{E}(e^{J^+} - 1) - \lambda_t^- \mathbb{E}(e^{J^-} - 1) \right) BG \\ & + \frac{\sigma^2}{2} B^2 G \\ & + \kappa^+(\theta^+ - \lambda_t^+) C + \kappa^-(\theta^- - \lambda_t^-) DG \\ & + \lambda_t^+ \left[\int_{\mathbb{R}} G(t, X_t + j^+, \lambda_t^+ + \beta_{11}j^+, \lambda_t^- + \beta_{21}j^+) \varpi^+(j^+) dj^+ - G(t, X_t, \lambda_t^+, \lambda_t^-) \right] \\ & + \lambda_t^- \left[\int_{\mathbb{R}} G(t, X_t + j^-, \lambda_t^+ + \beta_{12}j^-, \lambda_t^- + \beta_{22}j^-) \varpi^-(j^-) dj^- - G(t, X_t, \lambda_t^+, \lambda_t^-) \right]. \end{aligned}$$

Grouping terms by the state variables and dividing both sides by G (recall $G > 0$) gives,

$$\begin{aligned} 0 = & A_t + \mu B + \frac{\sigma^2}{2} (B^2 - B) + \kappa^+ \theta^+ C + \kappa^- \theta^- D \\ & + B_t X_t \\ & + \lambda_t^+ \left[\mathcal{L}^{(+)}(-B - C\beta_{11} - D\beta_{21}) - 1 + C_t - \mathbb{E}(e^{J^+} - 1)B - \kappa^+ C \right] \\ & + \lambda_t^- \left[\mathcal{L}^{(-)}(-B - C\beta_{12} - D\beta_{22}) - 1 + D_t - \mathbb{E}(e^{J^-} - 1)B - \kappa^- D \right], \end{aligned}$$

with $\mathcal{L}^{(+)}$ and $\mathcal{L}^{(-)}$ the MGFs of positive and negative jump sizes, cf. Equations (9) and (10). The above equation holds for all values of the state variables, x, λ^+, λ^- , so the following system of PDEs must hold

$$\begin{aligned} 0 &= A_t + \mu B + \frac{\sigma^2}{2} (B^2 - B) + \kappa^+ \theta^+ C + \kappa^- \theta^- D \\ 0 &= B_t \\ 0 &= \mathcal{L}^{(+)}(-B - C\beta_{11} - D\beta_{21}) - 1 + \partial_t C - \mathbb{E}(e^{J^+} - 1)B - \kappa^+ C \\ 0 &= \mathcal{L}^{(-)}(-B - C\beta_{12} - D\beta_{22}) - 1 + \partial_t D - \mathbb{E}(e^{J^-} - 1)B - \kappa^- D. \end{aligned}$$

The solution of the above system of PDEs is, for $t \geq 0$

$$\begin{aligned} A_t(t, T) &= -\mu\omega - \frac{\sigma^2}{2} (\omega^2 - \omega) - \kappa^+ \theta^+ C(t, T) - \kappa^- \theta^- D(t, T) \\ B_t(t, T) &= 0, \\ C_t(t, T) &= -\mathcal{L}^{(+)}(-\omega - C(t, T)\beta_{11} - D(t, T)\beta_{21}) + 1 + \mathbb{E}(e^{J^+} - 1)\omega + \kappa^+ C(t, T), \\ D_t(t, T) &= -\mathcal{L}^{(-)}(-\omega - C(t, T)\beta_{12} - D(t, T)\beta_{22}) + 1 + \mathbb{E}(e^{J^-} - 1)\omega + \kappa^- D(t, T), \end{aligned}$$

where the terminal conditions are $A(T, T) = 0$, $B(T, T) = \omega$, $C(T, T) = \omega^+$, and $D(T, T) = \omega^-$.

A.3 Proof of Proposition 2

Let $m_t = \ln(M_t)$, so

$$\begin{aligned} m_t &= q_1^+(\xi^+) \lambda_t^+ + q_1^-(\xi^-) \lambda_t^- + \xi^+ \sum_{j=1}^{N_t^{(1)}} J_j^+ + \xi^- \sum_{j=1}^{N_t^{(2)}} J_j^- + q_2(\xi^+, \xi^-) t \\ &\quad - \frac{1}{2} \int_0^t \varphi(s)^2 ds - \int_0^t \varphi(s) dW_s \end{aligned}$$

with dynamics

$$\begin{aligned} dm_t &= q_1^+(\xi^+) \kappa^+ (\theta^+ - \lambda_t^+) dt + (q_1^+(\xi^+) \beta_{11} + \xi^+) J^+ dN_t^{(1)} + q_1^+(\xi^+) \beta_{12} J^- dN_t^{(2)} \\ &\quad + q_1^-(\xi^-) \kappa^- (\theta^- - \lambda_t^-) dt + q_1^-(\xi^-) \beta_{21} J^+ dN_t^{(1)} + (q_1^-(\xi^-) \beta_{22} + \xi^-) J^- dN_t^{(2)} \\ &\quad + q_2(\xi^+, \xi^-) dt - \frac{1}{2} \varphi(t)^2 dt - \varphi(t) dW_t \\ &= q_1^+(\xi^+) \kappa^+ (\theta^+ - \lambda_t^+) dt + \chi^+ J^+ dN_t^{(1)} + q_1^-(\xi^-) \kappa^- (\theta^- - \lambda_t^-) dt + \chi^- J^- dN_t^{(2)} \\ &\quad + q_2(\xi^+, \xi^-) dt - \frac{1}{2} \varphi(t)^2 dt - \varphi(t) dW_t. \end{aligned}$$

Denote the random measures of J^\pm by Ξ^\pm , so that $J^+ = \int_0^\infty \Xi^+(dz)$ and $J^- = \int_{-\infty}^0 \Xi^-(dz)$. Applying Ito's Lemma for semimartingales (e.g. Section 8.10 of Klebaner, 2005) to M_t gives

$$\begin{aligned} dM_t &= M_{t-} dm_t + \frac{1}{2} M_t d[m_t, m_t] \\ &\quad + M_{t-} \int_0^\infty \left(e^{\chi^+ z} - 1 - \chi^+ z \right) \Xi^+(dz) dN_t^{(1)} + M_{t-} \int_{-\infty}^0 \left(e^{\chi^- z} - 1 - \chi^- z \right) \Xi^-(dz) dN_t^{(2)}. \end{aligned}$$

Expanding the dm_t -term and compensating the jumps gives

$$\begin{aligned}
dM_t &= M_t(q_1^+(\xi^+)\kappa^+\theta^+ + q_1^-(\xi^-)\kappa^-\theta^- + q_2(\xi^+, \xi^-))dt - M_t\varphi(t)dW_t \\
&\quad - M_t\lambda_t^+ \left(q_1^+(\xi^+)\kappa^+ - \int_0^\infty (e^{\chi^+z} - 1) \varpi^+(dz) \right) dt \\
&\quad - M_t\lambda_t^- \left(q_1^-(\xi^-)\kappa^- - \int_{-\infty}^0 (e^{\chi^-z} - 1) \varpi^-(dz) \right) dt \\
&\quad + M_{t-} \int_0^\infty (e^{\chi^+z} - 1) (\Xi^+(dz) dN_t^{(1)} - \lambda_t^+ \varpi^+(dz) dt) \\
&\quad + M_{t-} \int_{-\infty}^0 (e^{\chi^-z} - 1) (\Xi^-(dz) dN_t^{(2)} - \lambda_t^- \varpi^-(dz) dt).
\end{aligned}$$

The drift terms vanish if the conditions hold.

A.4 Proof of Proposition 4

The initial guess of the relationship between λ^+ and $\lambda^{+, \mathbb{Q}}$, and λ^- and $\lambda^{-, \mathbb{Q}}$ under the measure \mathbb{P} is

$$\begin{aligned}
\lambda_t^{+, \mathbb{Q}} &= \mathcal{L}^{(+)}(-\chi^+)\lambda_t^+, \\
\lambda_t^{-, \mathbb{Q}} &= \mathcal{L}^{(-)}(-\chi^-)\lambda_t^-, \quad t \geq 0.
\end{aligned}$$

We verify the above by comparing the MGFs under the two measures. Start with the MGF of jump intensities:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}} \left[\exp(\omega^+ \lambda_T^{+, \mathbb{Q}} + \omega^- \lambda_T^{-, \mathbb{Q}}) | \mathcal{F}_t \right] &= e^{-m} \mathbb{E} \left(\exp(m_T + \omega^+ \lambda_T^{+, \mathbb{Q}} + \omega^- \lambda_T^{-, \mathbb{Q}}) | \lambda_t^+ = \lambda^+, \lambda_t^- = \lambda^-, m_t = m \right) \\
&\stackrel{\text{def}}{=} e^{-m} H(t\omega^+, \omega^-; \lambda^+, \lambda^-, m, T).
\end{aligned}$$

Focus on the form of H that is comparable to the MGF under measure \mathbb{P} stated in Proposition 1:

$$H(t, \lambda^+, \lambda^-, m, T) = \exp \left(A(t, T) + C(t, T) \mathcal{L}^{(+)}(-\chi^+)\lambda^+ + D(t, T) \mathcal{L}^{(-)}(-\chi^-)\lambda^- + E(t, T)m \right),$$

where the functions A, C and D are defined in Proposition 1 and have the same terminal conditions; E is a time-dependent function with terminal condition $E(T, T) = 1$. Applying Ito's lemma and by martingale property of conditional expectation,

$$\begin{aligned}
0 &= H_t \\
&\quad + \left(\kappa_1^+(\xi^+)\kappa^+(\theta^+ - \lambda^+) + \kappa_1^-(\xi^-)\kappa^-(\theta^- - \lambda^-) + q_2(\xi^+, \xi^-) - \frac{1}{2}\varphi^2(t) \right) H_m + \frac{1}{2}\varphi^2(t)H_{mm} \\
&\quad + \kappa^+(\theta^+ - \lambda^+)H_{\lambda^+} + \kappa^-(\theta^- - \lambda^-)H_{\lambda^-} \\
&\quad + \lambda^+ \left[\int_{\mathbb{R}} H(t, \lambda^+ + \beta_{11}j^+, \lambda^- + \beta_{21}j^+, m + \chi^+j^+, T) \varpi^+(j^+)dj^+ - H(t, \lambda^+, \lambda^-, m, T) \right] \\
&\quad + \lambda^- \left[\int_{\mathbb{R}} H(t, \lambda^+ + \beta_{12}j^-, \lambda^- + \beta_{22}j^-, m + \chi^-j^-, T) \varpi^-(j^-)dj^- - H(t, \lambda^+, \lambda^-, m, T) \right].
\end{aligned}$$

By the first martingale condition of process $\{M_t\}_{t \geq 0}$, Eq.(17), we simplify the above equation:

$$\begin{aligned}
0 = & H_t \\
& - \left(\kappa_1^+(\xi^+) \kappa^+ \lambda^+ + \kappa_1^-(\xi^-) \kappa^- \lambda^- + \frac{1}{2} \varphi^2(t) \right) H_m + \frac{1}{2} \varphi^2(t) H_{mm} \\
& + \kappa^+(\theta^+ - \lambda^+) H_{\lambda^+} + \kappa^-(\theta^- - \lambda^-) H_{\lambda^-} \\
& + \lambda^+ \left[\int_{\mathbb{R}} H(t, \lambda^+ + \beta_{11} j^+, \lambda^- + \beta_{21} j^+, m + \chi^+ j^+, T) \varpi^+(j^+) dj^+ - H(t, \lambda^+, \lambda^-, m, T) \right] \\
& + \lambda^- \left[\int_{\mathbb{R}} H(t, \lambda^+ + \beta_{12} j^-, \lambda^- + \beta_{22} j^-, m + \chi^- j^-, T) \varpi^-(j^-) dj^- - H(t, \lambda^+, \lambda^-, m, T) \right].
\end{aligned}$$

Inserting the partial derivatives of H into the above equation and dividing both sides by H ,

$$\begin{aligned}
0 = & A_t + C_t \mathcal{L}^{(+)}(-\chi^+) \lambda^+ + D_t \mathcal{L}^{(-)}(-\chi^-) \lambda^- + E_t m \\
& - \left(\kappa_1^+(\xi^+) \kappa^+ \lambda^+ + \kappa_1^-(\xi^-) \kappa^- \lambda^- + \frac{1}{2} \varphi^2(t) \right) E + \frac{1}{2} \varphi^2(t) E^2 \\
& + \kappa^+(\theta^+ - \lambda^+) \mathcal{L}^{(+)}(-\chi^+) C + \kappa^-(\theta^- - \lambda^-) \mathcal{L}^{(-)}(-\chi^-) D \\
& + \lambda^+ \left[\mathcal{L}^{(+)}(-C \mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D \mathcal{L}^{(-)}(-\chi^-) \beta_{21} - E \chi^+) - 1 \right] \\
& + \lambda^- \left[\mathcal{L}^{(-)}(-C \mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D \mathcal{L}^{(-)}(-\chi^-) \beta_{22} - E \chi^-) - 1 \right]
\end{aligned}$$

Grouping terms by the state variables, x , λ^+ , and λ^- ,

$$\begin{aligned}
0 = & A_t + \frac{1}{2} \varphi^2(t) (E^2 - E) + \kappa^+ \theta^+ \mathcal{L}^{(+)}(-\chi^+) C + \kappa^- \theta^- \mathcal{L}^{(-)}(-\chi^-) D \\
& + E_t m \\
& + \lambda^+ \left[\mathcal{L}^{(+)}(-C \mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D \mathcal{L}^{(-)}(-\chi^-) \beta_{21} - E \chi^+) - 1 + C_t \mathcal{L}^{(+)}(-\chi^+) - \kappa_1^+(\xi^+) \kappa^+ - \kappa^+ C \mathcal{L}^{(+)}(-\chi^+) \right] \\
& + \lambda^- \left[\mathcal{L}^{(-)}(-C \mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D \mathcal{L}^{(-)}(-\chi^-) \beta_{22} - E \chi^-) - 1 + D_t \mathcal{L}^{(-)}(-\chi^-) - \kappa_1^-(\xi^-) \kappa^- - \kappa^- D \mathcal{L}^{(-)}(-\chi^-) \right]
\end{aligned}$$

The above equation holds for all values of x , λ^+ and λ^- , we infer

$$\begin{aligned}
\partial_t A(t, T) = & -\kappa^+ \theta^+ \mathcal{L}^{(+)}(-\chi^+) C(t, T) - \kappa^- \theta^- \mathcal{L}^{(-)}(-\chi^-) D(t, T) \\
\partial_t C(t, T) = & -\frac{\mathcal{L}^{(+)}(-C(t, T) \mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D(t, T) \mathcal{L}^{(-)}(-\chi^-) \beta_{21} - \chi^+)}{\mathcal{L}^{(+)}(-\chi^+)} + \frac{1 + \kappa_1^+(\xi^+) \kappa^+}{\mathcal{L}^{(+)}(-\chi^+)} + \kappa^+ C(t, T), \\
\partial_t D(t, T) = & -\frac{\mathcal{L}^{(-)}(-C(t, T) \mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D(t, T) \mathcal{L}^{(-)}(-\chi^-) \beta_{22} - \chi^-)}{\mathcal{L}^{(-)}(-\chi^-)} + \frac{1 + \kappa_1^-(\xi^-) \kappa^-}{\mathcal{L}^{(-)}(-\chi^-)} + \kappa^- D(t, T), \\
\partial_t E(t, T) = & 0.
\end{aligned}$$

The terminal conditions are $A(T, T) = 0$, $C(T, T) = \omega^+$, $D(T, T) = \omega^-$, and $E(T, T) = 1$.

The first term of $\partial_t C$, $\frac{\mathcal{L}^{(+)}(-C(t, T) \mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D(t, T) \mathcal{L}^{(-)}(-\chi^-) \beta_{21} - \chi^+)}{\mathcal{L}^{(+)}(-\chi^+)}$, is in fact an Esscher transformation applied onto the positive jumps size density ϖ^+ with $-\chi^+$ as the Esscher parameter (A general form of Esscher transform can be found in Eq. (2.6) in Gerber and Shiu (1995)). Similarly, the first term of $\partial_t D$ is an Esscher transform to the negative jumps size density with $-\chi^-$ as the

Esscher parameter. We denote the transformed MGFs of the jumps sizes as

$$\begin{aligned}\mathcal{L}^{(+),\mathbb{Q}}(x) &\stackrel{\text{def}}{=} \frac{\mathcal{L}^{(+)}(x - \chi^+)}{\mathcal{L}^{(+)}(-\chi^+)}, \text{ and} \\ \mathcal{L}^{(-),\mathbb{Q}}(x) &\stackrel{\text{def}}{=} \frac{\mathcal{L}^{(-)}(x - \chi^-)}{\mathcal{L}^{(-)}(-\chi^-)}.\end{aligned}$$

On the other hand, the second terms of $\partial_t C$ and $\partial_t D$ can be simplified by the martingale conditions in Eq.(17).

To summarise,

$$\begin{aligned}\partial_t C(t, T) &= -\mathcal{L}^{(+),\mathbb{Q}} \left(-C(t, T)\mathcal{L}^{(+)}(-\chi^+)\beta_{11} - D(t, T)\mathcal{L}^{(-)}(-\chi^-)\beta_{21} \right) + 1 + \kappa^+ C(t, T) \\ \partial_t D(t, T) &= -\mathcal{L}^{(-),\mathbb{Q}} \left(-C(t, T)\mathcal{L}^{(+)}(-\chi^+)\beta_{12} - D(t, T)\mathcal{L}^{(-)}(-\chi^-)\beta_{22} \right) + 1 + \kappa^- D(t, T).\end{aligned}$$

Next, we work out the Esscher transformed densities of the jump sizes (see Eq.(2.5) of Gerber and Shiu (1995)),

$$\begin{aligned}\varpi^{+,\mathbb{Q}}(j) &\stackrel{\text{def}}{=} \frac{e^{-\chi^+ j} \varpi^+(j)}{\int_{\mathbb{R}} e^{-\chi^+ y} \varpi^+(y) dy} = \frac{1}{\eta^{+,\mathbb{Q}}} \exp \left(-\frac{1}{\eta^{+,\mathbb{Q}}} (j - \nu^+) \right) \mathbb{1}_{\{j > \nu^+\}} \\ \varpi^{-,\mathbb{Q}}(j) &\stackrel{\text{def}}{=} \frac{e^{-\chi^- j} \varpi^-(j)}{\int_{\mathbb{R}} e^{-\chi^- y} \varpi^-(y) dy} = \frac{1}{\eta^{-,\mathbb{Q}}} \exp \left(\frac{1}{\eta^{-,\mathbb{Q}}} (j - \nu^-) \right) \mathbb{1}_{\{j < \nu^-\}},\end{aligned}$$

where

$$\eta^{+,\mathbb{Q}} = \frac{\eta^+}{1 - \eta^+ \chi^+}, \text{ and } \eta^{-,\mathbb{Q}} = \frac{\eta^-}{1 + \eta^- \chi^-}.$$

To summarise the parameter changes, the MGF of the jump intensities under the measure \mathbb{Q} is

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}} \left(e^{\omega^+ \lambda_T^+ + \omega^- \lambda_T^-} | \mathcal{F}_t \right) &= \exp \left(A(t, T) + C(t, T)\mathcal{L}^{(+)}(-\chi^+)\lambda_t^+ + D(t, T)\mathcal{L}^{(-)}(-\chi^-)\lambda_t^- \right) \\ &= \exp \left(A(t, T) + C(t, T)\lambda_t^{+,\mathbb{Q}} + D(t, T)\lambda_t^{-,\mathbb{Q}} \right),\end{aligned}$$

where the function A, C and D solve the following system of PDEs

$$\begin{aligned}\partial_t A(t, T) &= -\kappa^+ \theta^+ \mathcal{L}^{(+)}(-\chi^+)C(t, T) - \kappa^- \theta^- \mathcal{L}^{(-)}(-\chi^-)D(t, T), & A(T, T) &= 0 \\ \partial_t C(t, T) &= -\mathcal{L}^{(+),\mathbb{Q}}(-C(t, T)\mathcal{L}^{(+)}(-\chi^+)\beta_{11} - D(t, T)\mathcal{L}^{(-),\mathbb{Q}}(-\chi^-)\beta_{21}) + 1 + \kappa^+ C(t, T), & C(T, T) &= \omega^+ \\ \partial_t D(t, T) &= -\mathcal{L}^{(-),\mathbb{Q}}(-C(t, T)\mathcal{L}^{(+)}(-\chi^+)\beta_{12} - D(t, T)\mathcal{L}^{(-),\mathbb{Q}}(-\chi^-)\beta_{22}) + 1 + \kappa^- D(t, T), & D(T, T) &= \omega^-.\end{aligned}$$

By comparing the above MGF under \mathbb{Q} to the MGF under \mathbb{P} in Proposition 1 (when $\omega = 0$), we conclude that the measure change

- alters the positive and negative jump intensities by multipliers $\mathcal{L}^{(+)}(-\chi^+)$ and $\mathcal{L}^{(-)}(-\chi^-)$, respectively.
- shifts the rates of the positive and negative jump sizes from η^+ to $\eta^+/(1 - \eta^+ \chi^+)$ and from η^- to $\eta^-/(1 + \eta^- \chi^-)$, respectively.
- preserves the mean reversion rates of jump intensities κ^+ and κ^- and the shift parameters of jump sizes ν^+ and ν^- .

A.5 Proof of Proposition 5

The proof relies on comparing the the MGFs of X under \mathbb{Q} and \mathbb{P} . First, we derive the MGF of X under \mathbb{Q} . Denote $m_t = \ln(M_t)$. Given $X_t = x, \lambda_t^+ = \lambda^+, \lambda_t^- = \lambda^-$ and $m_t = m$, the MGF is

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}(e^{\omega X_T} | \mathcal{F}_t) &= \mathbb{E}(e^{m_T - m + \omega X_T} | \mathcal{F}_t) \\ &= e^{-m} \mathbb{E}(e^{m_T + \omega X_T} | X_t = x, \lambda_t^+ = \lambda^+, \lambda_t^- = \lambda^-, M_t = m, T),\end{aligned}$$

where $m_t = \ln(M_t)$. Let $H(t\omega; x, \lambda^+, \lambda^-, m, T) = \mathbb{E}(e^{M_T + \omega X_T} | X_t = x, \lambda_t^+ = \lambda^+, \lambda_t^- = \lambda^-, M_t = m)$. By the tower law, $\{H(t)\}_{t \geq 0}$ is a martingale. By Ito's lemma and martingale property of conditional expectation,

$$\begin{aligned}0 &= \partial_t H \\ &+ \left(\mu - \frac{\sigma^2}{2} - \lambda^+ \mathbb{E}(e^{J^+} - 1) - \lambda^- \mathbb{E}(e^{J^-} - 1) \right) \partial_x H + \frac{\sigma^2}{2} \partial_x^2 H \\ &+ \left(\kappa_1^+(\xi^+) \kappa^+(\theta^+ - \lambda^+) + \kappa_1^-(\xi^-) \kappa^-(\theta^- - \lambda^-) + q_2(\xi^+, \xi^-) - \frac{1}{2} \varphi^2(t) \right) \partial_m H + \frac{1}{2} \varphi^2(t) \partial_m^2 H \\ &- \sigma \varphi(t) \partial_{xm}^2 \\ &+ \kappa^+(\theta^+ - \lambda^+) \partial_+ H + \kappa^-(\theta^- - \lambda^-) \partial_- H \\ &+ \lambda^+ \left[\int_{\mathbb{R}} H(t, \omega^+, \omega^-; \lambda^+ + \beta_{11} j^+, \lambda^- + \beta_{21} j^+, m + \chi^+ j^+, T) \varpi^+(j^+) dj^+ - H(t, \omega^+, \omega^-; \lambda^+, \lambda^-, m, T) \right] \\ &+ \lambda^- \left[\int_{\mathbb{R}} H(t, \omega^+, \omega^-; \lambda^+ + \beta_{12} j^-, \lambda^- + \beta_{22} j^-, m + \chi^- j^-, T) \varpi^-(j^-) dj^- - H(t, \omega^+, \omega^-; \lambda^+, \lambda^-, m, T) \right].\end{aligned}$$

By the first martingale condition of process $\{M_t\}_{t \geq 0}$, Eq.(17), we simplify the above equation:

$$\begin{aligned}0 &= \partial_t H \\ &+ \left(\mu - \frac{\sigma^2}{2} - \lambda^+ \mathbb{E}(e^{J^+} - 1) - \lambda^- \mathbb{E}(e^{J^-} - 1) \right) \partial_x H + \frac{\sigma^2}{2} \partial_x^2 H \\ &- \left(\kappa_1^+(\xi^+) \kappa^+ \lambda^+ + \kappa_1^-(\xi^-) \kappa^- \lambda^- + \frac{1}{2} \varphi^2(t) \right) \partial_m H + \frac{1}{2} \varphi^2(t) \partial_m^2 H - \sigma \varphi(t) \partial_{xm}^2 \\ &+ \kappa^+(\theta^+ - \lambda^+) \partial_+ H + \kappa^-(\theta^- - \lambda^-) \partial_- H \\ &+ \lambda^+ \left[\int_{\mathbb{R}} H(t, \omega^+, \omega^-; \lambda^+ + \beta_{11} j^+, \lambda^- + \beta_{21} j^+, m + \chi^+ j^+, T) \varpi^+(j^+) dj^+ - H(t, \omega^+, \omega^-; \lambda^+, \lambda^-, m, T) \right] \\ &+ \lambda^- \left[\int_{\mathbb{R}} H(t, \omega^+, \omega^-; \lambda^+ + \beta_{12} j^-, \lambda^- + \beta_{22} j^-, m + \chi^- j^-, T) \varpi^-(j^-) dj^- - H(t, \omega^+, \omega^-; \lambda^+, \lambda^-, m, T) \right].\end{aligned}\tag{52}$$

Next, we write the form of H such that it is comparable to the MGF under measure \mathbb{P} stated in Proposition 1:

$$H(t; \omega, x, \lambda^+, \lambda^-, m, T) = \exp \left(A(t, T) + B(t, T)x + C(t, T)\mathcal{L}^{(+)}(-\chi^+) \lambda^+ + D(t, T)\mathcal{L}^{(-)}(-\chi^-) + E(t, T)m \right),$$

where functions A , B , C , and D are defined in Proposition 1 with terminal conditions $A(T, T) = 0$, $B(T, T) = \omega$, and $C(T, T) = D(T, T) = 0$; E is a time-dependent function with terminal condition

$E(T, T) = 1$. Inserting the partial derivatives of H into Eq.(52)

$$\begin{aligned}
0 = & \partial_t A + \partial_t Bx + \partial_t C\mathcal{L}^{(+)}(-\chi^+) \lambda^+ + \partial_t D\mathcal{L}^{(-)}(-\chi^-) \lambda^- + \partial_t Em \\
& + \left(\mu - \frac{\sigma^2}{2} - \lambda^+ \mathbb{E} \left(e^{J^+} - 1 \right) - \lambda^- \mathbb{E} \left(e^{J^-} - 1 \right) \right) B + \frac{\sigma^2}{2} B^2 \\
& - \left(\kappa_1^+ (\xi^+) \kappa^+ \lambda^+ + \kappa_1^- (\xi^-) \kappa^- \lambda^- + \frac{1}{2} \varphi^2(t) \right) E + \frac{1}{2} \varphi^2(t) E^2 \\
& + \kappa^+ (\theta^+ - \lambda^+) \mathcal{L}^{(+)}(-\chi^+) C + \kappa^- (\theta^- - \lambda^-) \mathcal{L}^{(-)}(-\chi^-) D \\
& - BE\sigma\varphi(t) \\
& + \lambda^+ \left[\mathcal{L}^{(+)} \left(-B - C\mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{21} - E\chi^+ \right) - 1 \right] \\
& + \lambda^- \left[\mathcal{L}^{(-)} \left(-B - C\mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{22} - E\chi^- \right) - 1 \right]
\end{aligned}$$

Group terms by state variables

$$\begin{aligned}
0 = & \partial_t A + \mu B - \frac{\sigma^2}{2} (B^2 - B) + \frac{1}{2} \varphi^2(t) (E^2 - E) + \kappa^+ \theta^+ \mathcal{L}^{(+)}(-\chi^+) C + \kappa^- \theta^- \mathcal{L}^{(-)}(-\chi^-) D - BE\sigma\varphi(t) \\
& + \partial_t Bx + \partial_t Em \\
& + \lambda^+ \left[\mathcal{L}^{(+)} \left(-B - C\mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{21} - E\chi^+ \right) - 1 \right. \\
& \quad \left. + \partial_t C\mathcal{L}^{(+)}(-\chi^+) - \mathbb{E} \left(e^{J^+} - 1 \right) B - \kappa_1^+ (\xi^+) \kappa^+ E - \kappa^+ \mathcal{L}^{(+)}(-\chi^+) C \right] \\
& + \lambda^- \left[\mathcal{L}^{(-)} \left(-B - C\mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{22} - E\chi^- \right) - 1 \right. \\
& \quad \left. + \partial_t D\mathcal{L}^{(-)}(-\chi^-) - \mathbb{E} \left(e^{J^-} - 1 \right) B - \kappa_1^- (\xi^-) \kappa^- E - \kappa^- \mathcal{L}^{(-)}(-\chi^-) D \right].
\end{aligned}$$

Since the above equation holds for all value of state variables, x , λ^+ , λ^- , we infer for $t \leq T$, $\partial_t B(t, T) = \partial_t E(t, T) = 0$, and thus $B(t, T) = \omega$ and $E(t, T) = 1$. The above equation becomes

$$\begin{aligned}
0 = & \partial_t A + \mu B - \frac{\sigma^2}{2} (\omega^2 - \omega) + \kappa^+ \theta^+ \mathcal{L}^{(+)}(-\chi^+) C + \kappa^- \theta^- \mathcal{L}^{(-)}(-\chi^-) D - \omega\sigma\varphi(t) \\
& + \lambda^+ \left[\mathcal{L}^{(+)} \left(-\omega - C\mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{21} - \chi^+ \right) - 1 \right. \\
& \quad \left. + \partial_t C\mathcal{L}^{(+)}(-\chi^+) - \mathbb{E} \left(e^{J^+} - 1 \right) \omega - \kappa_1^+ (\xi^+) \kappa^+ - \kappa^+ \mathcal{L}^{(+)}(-\chi^+) C \right] \\
& + \lambda^- \left[\mathcal{L}^{(-)} \left(-\omega - C\mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{22} - \chi^- \right) - 1 \right. \\
& \quad \left. + \partial_t D\mathcal{L}^{(-)}(-\chi^-) - \mathbb{E} \left(e^{J^-} - 1 \right) \omega - \kappa_1^- (\xi^-) \kappa^- - \kappa^- \mathcal{L}^{(-)}(-\chi^-) D \right].
\end{aligned}$$

Recall we choose the \mathcal{F}_t -adapted process as $\varphi(t) = \varphi + \varphi^+ \lambda_t^+ + \varphi^- \lambda_t^-$, so expanding the $\varphi(t)$ term in the above equation gives

$$\begin{aligned}
0 = & \partial_t A + \mu\omega + \frac{\sigma^2}{2} (\omega^2 - \omega) + \kappa^+ \theta^+ \mathcal{L}^{(+)}(-\chi^+) C + \kappa^- \theta^- \mathcal{L}^{(-)}(-\chi^-) D - \omega\sigma\varphi \\
& + \lambda^+ \left[\mathcal{L}^{(+)} \left(-\omega - C\mathcal{L}^{(+)}(-\chi^+) \beta_{11} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{21} - \chi^+ \right) - 1 \right. \\
& \quad \left. + \partial_t C\mathcal{L}^{(+)}(-\chi^+) - \mathbb{E} \left(e^{J^+} - 1 \right) \omega - \kappa_1^+ (\xi^+) \kappa^+ - \kappa^+ \mathcal{L}^{(+)}(-\chi^+) C - \omega\sigma\varphi^+ \right] \\
& + \lambda^- \left[\mathcal{L}^{(-)} \left(-\omega - C\mathcal{L}^{(+)}(-\chi^+) \beta_{12} - D\mathcal{L}^{(-)}(-\chi^-) \beta_{22} - \chi^- \right) - 1 \right. \\
& \quad \left. + \partial_t D\mathcal{L}^{(-)}(-\chi^-) - \mathbb{E} \left(e^{J^-} - 1 \right) \omega - \kappa_1^- (\xi^-) \kappa^- - \kappa^- \mathcal{L}^{(-)}(-\chi^-) D - \omega\sigma\varphi^- \right]
\end{aligned}$$

The above equation holds for all values of state variables. Together with the parameter changes stated in Proposition 4, we have

$$\begin{aligned}\partial_t A &= -(\mu - \sigma\varphi)\omega - \frac{\sigma^2}{2}(\omega^2 - \omega) - \kappa^+ \theta^{+, \mathbb{Q}} C - \kappa^- \theta^{-, \mathbb{Q}} D \\ \partial_t C &= -\mathcal{L}^{+, \mathbb{Q}}(-\omega - C\beta_{11}^{\mathbb{Q}} - D\beta_{21}^{\mathbb{Q}} - \chi^+) + 1 + \frac{\mathbb{E}(e^{J^+} - 1) + \sigma\varphi^+}{\mathcal{L}^{(+)}(-\chi^+)}\omega + \kappa^+ C \\ \partial_t D &= -\mathcal{L}^{-, \mathbb{Q}}(-\omega - C\beta_{21}^{\mathbb{Q}} - D\beta_{22}^{\mathbb{Q}} - \chi^-) + 1 + \frac{\mathbb{E}(e^{J^-} - 1) + \sigma\varphi^-}{\mathcal{L}^{(-)}(-\chi^-)}\omega + \kappa^- D\end{aligned}$$

By choosing

$$\begin{aligned}\varphi &= \sigma^{-1}(\mu - r), \\ \varphi^+ &= \sigma^{-1} \left[\mathcal{L}^{(+)}(-\chi^+) \mathbb{E}^{\mathbb{Q}}(e^{J^+} - 1) - \mathbb{E}(e^{J^+} - 1) \right], \\ \varphi^- &= \sigma^{-1} \left[\mathcal{L}^{(-)}(-\chi^-) \mathbb{E}^{\mathbb{Q}}(e^{J^-} - 1) - \mathbb{E}(e^{J^-} - 1) \right],\end{aligned}$$

the MGF of $\{X_t\}_{t \geq 0}$ under measure \mathbb{Q} can be written as

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}(e^{\omega X_T} | \mathcal{F}_t) &= \exp \left(A(t, T) + \omega x + C(t, T) \mathcal{L}^{(+)}(-\chi^+) \lambda_t^+ + D(t, T) \mathcal{L}^{(-)}(-\chi^-) \lambda_t^- \right) \\ &= \exp \left(A(t, T) + \omega x + C(t, T) \lambda_t^{+, \mathbb{Q}} + D(t, T) \lambda_t^{-, \mathbb{Q}} \right),\end{aligned}$$

where the functions A , C , and D solve the following system of PDEs:

$$\begin{aligned}\partial_t A(t, T) &= -r\omega - \frac{\sigma^2}{2}(\omega^2 - \omega) - \kappa^+ \theta^{+, \mathbb{Q}} C(t, T) - \kappa^- \theta^{-, \mathbb{Q}} D(t, T) \\ \partial_t C(t, T) &= -\mathcal{L}^{+, \mathbb{Q}}(-\omega - C(t, T)\beta_{11}^{\mathbb{Q}} - D(t, T)\beta_{21}^{\mathbb{Q}} - \chi^+) + 1 + \mathbb{E}^{\mathbb{Q}}(e^{J^+} - 1)\omega + \kappa^+ C(t, T) \\ \partial_t D(t, T) &= -\mathcal{L}^{-, \mathbb{Q}}(-\omega - C(t, T)\beta_{21}^{\mathbb{Q}} - D(t, T)\beta_{22}^{\mathbb{Q}} - \chi^-) + 1 + \mathbb{E}^{\mathbb{Q}}(e^{J^-} - 1)\omega + \kappa^- D(t, T).\end{aligned}$$

Finally, by comparing the above MGF under \mathbb{Q} to that under \mathbb{P} stated in Proposition 1, we conclude that the dynamics of the log returns $\{X_t\}_{t \geq 0}$ under measure \mathbb{Q} can be retrieved by the substitution of model parameters stated in Eq.29.

A.6 Derivation of the likelihood

Suppose we observe $k - 1$ jump events and collect the following:

- positive and negative jump event times $\mathcal{T}^+ = \{T_{[1]}^+, T_{[2]}^+, \dots, T_m^+\}$ and $\mathcal{T}^- = \{T_{[1]}^-, T_{[2]}^-, \dots, T_{[n]}^-\}$ respectively,
- the corresponding jump sizes of positive and negative jump events $\{J^+(T_{[1]}^+), J^+(T_{[2]}^+), \dots, J^+(T_{[m]}^+)\}$ and $\{J^-(T_{[1]}^-), J^-(T_{[2]}^-), \dots, J^-(T_{[n]}^-)\}$ respectively, and,
- a set of ordered set of union jump times $\mathcal{T}^\pm = \mathcal{T}^+ \cup \mathcal{T}^- = \{T_{[1]}^\pm, T_{[2]}^\pm, \dots, T_{[k-1]}^\pm\}$,

the hazard rate of observing the k^{th} jump between time $T_{[k-1]}^{\pm}$ and $s > T_{[k-1]}^{\pm}$ is

$$\begin{aligned} F^{\pm}(s|T_{[k-1]}^{\pm}) &= 1 - \exp\left(-\int_{T_{[k-1]}^{\pm}}^s (\lambda_t^+ + \lambda_t^-)dt\right) \\ &= 1 - \exp\left(-I^+(T_{[k-1]}^{\pm}, s) - I^-(T_{[k-1]}^{\pm}, s)\right), \end{aligned}$$

where

$$\begin{aligned} I^+(T_{[k-1]}^{\pm}, s) &= \int_{T_{[k-1]}^{\pm}}^s \lambda_t^+ dt = \theta^+(s - T_{[k-1]}^{\pm}) + (\lambda^+(T_{[k-1]}^{\pm}) - \theta^+) \frac{1 - e^{-\kappa^+(s - T_{[k-1]}^{\pm})}}{\kappa^+} \\ I^-(T_{[k-1]}^{\pm}, s) &= \int_{T_{[k-1]}^{\pm}}^s \lambda_t^- dt = \theta^-(s - T_{[k-1]}^{\pm}) + (\lambda^-(T_{[k-1]}^{\pm}) - \theta^-) \frac{1 - e^{-\kappa^-(s - T_{[k-1]}^{\pm})}}{\kappa^-}. \end{aligned}$$

The pds of the inter-arrival time given the $(k-1)^{\text{th}}$ jump is

$$f^{\pm}(s|T_{[k-1]}^{\pm}) = (\lambda_s^+ + \lambda_s^-) \exp(-I^+ - I^-).$$

Suppose the next jump happens at time $T_{[k]}^{\pm}$, the probability of this jump being a positive jump is

$$\mathbb{P}\left(dN^{(1)}(s) = 1 | s = T_{[k]}^{\pm}\right) = \frac{\lambda^+(T_{[k]}^{\pm}-)}{\lambda^+(T_{[k]}^{\pm}-) + \lambda^-(T_{[k]}^{\pm}-)},$$

where $dN^{(1)}(s) = N^{(1)}(s) - N^{(1)}(s-)$; And similarly for negative jump

$$\mathbb{P}\left(dN^{(2)}(s) = 1 | s = T_{[k]}^{\pm}\right) = \frac{\lambda^-(T_{[k]}^{\pm}-)}{\lambda^+(T_{[k]}^{\pm}-) + \lambda^-(T_{[k]}^{\pm}-)}.$$

Accordingly, the pdf of a positive jump occurring at time s is

$$f^+(s|T_{[k-1]}^{\pm}) = f^{\pm}(s|T_{[k-1]}^{\pm}) \mathbb{P}\left(dN^{(1)}(s) = 1 | s = T_{[k]}^{\pm}\right) = \lambda_s^+ \exp(-I^+ - I^-),$$

and similarly for that of a negative jump

$$f^-(s|T_{[k-1]}^{\pm}) = \lambda_s^- \exp(-I^+ - I^-).$$

Therefore, the likelihood function of $\{\lambda_t^+\}_{t \leq T_{[k-1]}^{\pm}}$, $\{\lambda_t^-\}_{t \leq T_{[k-1]}^{\pm}}$, and the jumps sizes densities ϖ^+ and ϖ^- to the observed inter-arrival times is

$$\begin{aligned} &\mathbb{P}\left(\mathcal{T}^+, \mathcal{T}^- | \{\lambda_t^+\}_{t \leq T_{[k-1]}^{\pm}}, \{\lambda_t^-\}_{t \leq T_{[k-1]}^{\pm}}, \varpi^+, \varpi^-\right) \\ &= \prod_{j=1}^m \lambda^+(T_{[m]}^+) \varpi^+(J^-(T_{[m]}^+)) \\ &\times \prod_{j=1}^n \lambda^-(T_{[n]}^-) \varpi^-(J^-(T_{[n]}^-)) \\ &\times \prod_{j=1}^{k-1} \exp\left(-I^+(T_{[j]}^{\pm}, T_{[j-1]}^{\pm}) - I^-(T_{[j]}^{\pm}, T_{[j-1]}^{\pm})\right). \end{aligned}$$

B Goodness of fit assessment by Q-Q plot

Our assessment relies on the work by Watanabe (1964) who states that a Poisson process can be characterized by the form of its compensator, $\Lambda(T) = \int_0^T \lambda(t)dt$. This characterisation is later on known as the random time change theorem, see Section 7.4 of (Daley and Vere-Jones, 2003) and reference therein. We rewrite the multivariate random time change theorem stated in Section 9.3 of Laub et al. (2021) in a bidimensional version to suit our purpose here.

Proposition 6. *Bidimensional random time change. (Theorem 9.3 of Laub et al. (2021))*

Suppose we observe two counting processes, $N^{(1)}(t)$ and $N^{(2)}(t)$, which are characterised by intensities λ_t^+ and λ_t^- , respectively. The event arrival times of the two counting processes are labelled as $\{T_1^{(1)}, T_2^{(1)}, \dots\}$ and $\{T_1^{(2)}, T_2^{(2)}, \dots\}$. In addition, suppose $\lambda_t = \lambda_t^+ + \lambda_t^-$ is positive over $[0, T]$ and $\int_0^T \lambda(t)dt < \infty$ almost surely, then the transformed event arrival times

$$\left\{ \int_0^{T_1^{(1)}} \lambda_t^+ dt, \int_0^{T_2^{(1)}} \lambda_t^+ dt, \dots \right\} \text{ and } \left\{ \int_0^{T_1^{(2)}} \lambda_t^- dt, \int_0^{T_2^{(2)}} \lambda_t^- dt, \dots \right\}$$

are events arrival times of two Poisson processes with unit rate.

In order words, the transformed events inter-arrival times,

$$\left\{ \int_0^{T_1^{(1)}} \lambda_t^+ dt, \int_{T_1^{(1)}}^{T_2^{(1)}} \lambda_t^+ dt, \int_{T_2^{(1)}}^{T_3^{(1)}} \lambda_t^+ dt, \dots \right\} \text{ and } \left\{ \int_0^{T_1^{(2)}} \lambda_t^- dt, \int_{T_1^{(2)}}^{T_2^{(2)}} \lambda_t^- dt, \int_{T_2^{(2)}}^{T_3^{(2)}} \lambda_t^- dt, \dots \right\}$$

are two random variables that follows exponential distribution with unit rate. Note that the event of $N^{(2)}$ can arrive between two events of $N^{(1)}$ and vice versa, e.g. it is possible that $T_k^{(1)} < T_j^{(2)} < T_{k+1}^{(1)}$ for some integers k and j . The random time change theorem enables us to perform residual analysis to assess the goodness-of-fit of time-inhomogeneous Poisson processes.

Theorem 1. *Consider an unbounded, increasing sequence of time points $\{t_1, t_2, \dots\}$ in $(0, \infty)$, and a monotone, continuous function $\Lambda(\cdot)$ such that $\lim_{t \rightarrow \infty} \Lambda(t) = \infty$ almost surely. The transformed sequence $\{\Lambda(t_1), \Lambda(t_2), \dots\}$ is a realisation of a unit rate Poisson process if and only if the original sequence $\{t_1, t_2, \dots\}$ is a realisation from the point process characterised by $\Lambda(\cdot)$.*

Proof. See Brown and Nair (1988). □

We refer readers to Section 9 of Laub et al. (2021) for further details regarding the application of random time change theorem on Hawkes processes.